DONACION CEI. Abril · 99

# Álgebra y geometría


**EUGENIO HERNÁNDEZ**
*Facultad de Ciencias*
*Universidad Autónoma de Madrid*





**ADDISON-WESLEY/UNIVERSIDAD AUTÓNOMA DE MADRID**

Argentina • Brasil • Chile • Colombia
Ecuador • España • Estados Unidos • México
Perú • Puerto Rico • Venezuela

Copublicación de Addison-Wesley Iberoamericana, S. A.
y la Universidad Autónoma de Madrid



Impreso en Estados Unidos de América. *Printed in U.S.A.*

ISBN 0-201-62586-5

4 5 6 7 8 9 10-MA-00 99 98 97 96

# INDICE

# PROLOGO



La matemática no es un deporte para espectadores; el lector debe acercarse a este texto con un lapicero en su mano y un papel a su lado, para verificar con sus propios razonamientos y su espíritu crítico las afirmaciones que contiene.

De la misma manera que lograr un nivel adecuado en el juego del tenis requiere tiempo y práctica, y conseguir tocar cualquier pieza de música clásica requiere esfuerzo, recompensado por la belleza que su música proporciona, la matemática es una ciencia cuyo aprendizaje requiere esfuerzo y práctica y cuya recompensa se alcanza por la elegancia con la que permite resolver problemas propios y de otras Ciencias.

Esperamos que el lector se esfuerce en comprender los conceptos y resultados que se exponen en este libro, porque ellos son la base para poder apreciar posteriormente varias de las aportaciones que la Ciencia ha dado a la humanidad a través de los tiempos y de manera especial en este siglo XX.

Este libro ha surgido de las clases de Álgebra y Geometría impartidas durante varios años a los alumnos de primer curso de las licenciaturas de Ciencias Físicas y Ciencias Matemáticas en la Universidad Autónoma de Madrid. Ha crecido con la colaboración de varios colegas del Departamento de Matemáticas de la citada Universidad; unos aportando soluciones para la mejor exposición de algunas lecciones; otros mejorando ideas ya plasmadas en papel; otros, finalmente, corrigiendo varias versiones del manuscrito. A todos ellos agradezco su desinteresada aportación en la elaboración de este libro.

En él se ha pretendido seguir un esquema que permita al lector adivinar los resultados e intuir su demostración: para ello se dan varios ejemplos antes de enunciar un resultado y aportar las razones convincentes que lo demuestran. Estas razones son puramente geométricas cuando ello ha sido posible, como en la demostración de las propiedades de las secciones cónicas (capítulo 11) o en la clasificación de los movimientos en el plano (capítulo 10).

Ejemplos de aplicaciones se dan en varias ocasiones después de haber concluido la demostración de un importante resultado. Con todo ello se intenta lograr una participación activa del lector en el descubrimiento de las ideas principales de cada capítulo, a la vez que la oportunidad para que vaya comprobando su nivel de conocimientos.

Este nivel de conocimientos puede comprobarse también intentando solucionar los numerosos problemas que se proponen al final de casi todas las secciones y de aquellos que, a modo de repaso, se incluyen después de los capítulos 7 y 13. El completo aprendizaje de las teorías matemáticas se consigue después de haber resuelto numerosos ejercicios. El lector debe intentar resolverlos todos, con la seguridad de que estos intentos, aunque sean fallidos, le proporcionarán grandes beneficios.

De muchos de los problemas se incluyen los resultados al final del libro. La elaboración de estos resultados ha contado con la participación de varios Ayudantes del Departamento de Matemáticas, a quienes también agradezco su contribución.

Las Rozas de Madrid
Agosto de 1987

## CAPITULO 1

# RESOLUCION DE SISTEMAS DE ECUACIONES LINEALES. OPERACIONES CON MATRICES



Este capítulo está dedicado a la resolución de sistemas de ecuaciones lineales; el problema geométrico más sencillo en el cual surge la necesidad de resolver sistemas de ecuaciones lineales es el de conocer la intersección de dos rectas en el plano. Así, por ejemplo, los números que satisfacen el sistema

$$\left.\begin{array}{r} x+y=2 \\ 3x-y=2 \end{array}\right\}$$

determinan el punto de intersección de las rectas $x+y=2$ y $3x-y=2$, representadas en la figura 1.1.

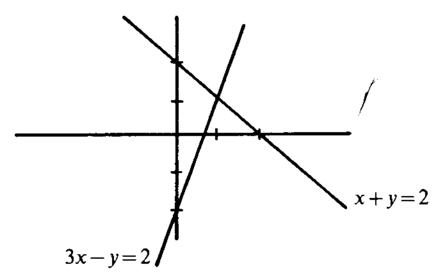



Figura 1.1

Es posible que el lector esté familiarizado con la resolución de sistemas de ecuaciones lineales como el anterior utilizando uno cualquiera de los siguientes métodos:

1) *Método de eliminación*, que consiste en realizar «operaciones» con las ecuaciones dadas hasta eliminar una de las incógnitas.
2) *Método de sustitución*, que consiste en despejar una incógnita de una de las ecuaciones y sustituirla en la otra.
3) *Método de Cramer o de los determinantes*, que consiste en encontrar las soluciones del sistema anterior como un cociente de dos determinantes.

De todos estos métodos el que resulta menos engorroso cuando se trata de resolver sistemas de un gran número de incógnitas es el método de *eliminación*, que recibe el nombre de *método de eliminación de Gauss* (Carl Friedrich Gauss fue uno de los más prestigiosos matemáticos de comienzos del siglo XIX).

Comenzaremos exponiendo este método seguidamente.

## 1.1. RESOLUCION DE SISTEMAS DE ECUACIONES LINEALES: METODO DE ELIMINACION DE GAUSS

EJEMPLO A. Tratemos de resolver el sistema

$$\left.\begin{aligned} x+y&=2 \\ 3x-y&=2 \end{aligned}\right\}$$

de dos ecuaciones con dos incógnitas; una solución de este sistema es un par de números $(a, b)$ que satisface las dos ecuaciones simultáneamente. El primer paso es multiplicar por 3 la primera ecuación y restarla de la segunda para obtener el sistema

$$\left.\begin{aligned} x+y&=\phantom{-}2 \\ -4y&=-4 \end{aligned}\right\}$$

A continuación dividimos entre $-4$ (o bien multiplicamos por $-\frac{1}{4}$) la segunda ecuación para obtener el sistema

$$\left.\begin{aligned} x+y&=2 \\ y&=1 \end{aligned}\right\}$$

Restando la segunda ecuación de la primera se obtiene

$$\left.\begin{aligned} x&=1 \\ y&=1 \end{aligned}\right\}$$



que nos permite obtener fácilmente el par de números (1, 1), que es solución del sistema dado.

*Nota.* Observar que $x=1$, $y=1$ satisfacen ambas ecuaciones; el lector debe comprobar siempre que el resultado obtenido es correcto sustituyendo los valores encontrados en el sistema dado.

EJEMPLO B. Tratemos de resolver el sistema

$$\begin{aligned} x_1+3x_2+x_3&=-3 \\ 3x_1+9x_2+4x_3&=-7 \\ 2x_1-x_2+x_3&=6 \end{aligned}$$

de tres ecuaciones lineales con tres incógnitas.

Comenzamos eliminando $x_1$ de las ecuaciones segunda y tercera; esto se consigue multiplicando por 3 la primera ecuación y restándola de la segunda y multiplicando por 2 la primera ecuación y restándola de la tercera. Realizando estas operaciones, el sistema anterior se transforma en

$$\left.\begin{aligned} x_1+3x_2+x_3&=-3 \\ x_3&=2 \\ -7x_2-x_3&=12 \end{aligned}\right\}$$

Intercambiando las ecuaciones segunda y tercera se obtiene:

$$\left.\begin{aligned} x_1+3x_2+x_3&=-3 \\ -7x_2-x_3&=12 \\ x_3&=2 \end{aligned}\right\}$$

A continuación eliminamos $x_3$ de la primera y la segunda de las ecuaciones restando la tercera de la primera y sumando la tercera y la segunda. Obtenemos

$$\left.\begin{aligned} x_1+3x_2&=-5 \\ -7x_2&=14 \\ x_3&=2 \end{aligned}\right\}$$

Multiplicando por $-\frac{1}{7}$ la segunda ecuación se obtiene:

$$\left.\begin{aligned} x_1+3x_2&=-5 \\ x_2&=-2 \\ x_3&=2 \end{aligned}\right\}$$

Finalmente, eliminamos $x_2$ de la primera ecuación multiplicando por 3 la segunda y restándola de la primera; obtenemos:

$$\left.\begin{aligned} x_1&=1\\ x_2&=-2\\ x_3&=2 \end{aligned}\right\}$$

que nos da la solución $(1, -2, 2)$ del sistema.

*Comprobación.* $1-6+2=-3$; $3-18+8=-7$; $2+2+2=6$.

* * *

El lector se preguntará la razón por la que el método utilizado produce la solución del sistema. La respuesta es que las «operaciones» realizadas con las ecuaciones transforman un sistema en otro *equivalente*, es decir que tiene las mismas soluciones. Recapitulemos las operaciones realizadas en los ejemplos anteriores que, de ahora en adelante, serán llamadas *operaciones elementales*:

*i)* Multiplicar una ecuación por un número real no nulo.
*ii)* Intercambiar dos ecuaciones.
*iii)* Sumar o restar un múltiplo de una ecuación a otra.

No resulta complicado comprobar que las operaciones elementales transforman un sistema en otro equivalente. Por ejemplo, si $(a, b, c)$ es solución de $x_1+3x_2+x_3=-3$, también es solución de $3x_1+9x_2+3x_3=-9$, ya que

$$3a+9b+3c=3(a+3b+c)=3(-3)=-9$$

De manera similar, si $(a, b, c)$ es también solución de $3x_1+9x_2+4x_3=-7$, cinco veces la ecuación $x_1+3x_2+x_3=-3$ más la ecuación $3x_1+9x_2+4x_3=-7$ nos da

$$8x_1+24x_2+9x_3=-22$$

que tiene $(a, b, c)$ como solución, ya que

$$8a+24b+9c=(5a+15b+5c)+(3a+9b+4c)=5(-3)-7=-22.$$

Resumiendo, podemos decir que el método de *eliminación de Gauss* consiste en reducir un sistema dado a otro equivalente, lo más sencillo posible, mediante operaciones elementales.

Puede observarse que la repetición de las incógnitas y de los signos + en los ejemplos anteriores es innecesaria. Si eliminamos estos símbolos en el ejemplo B, el sistema queda reducido a la siguiente ordenación rectangular, que llamaremos *matriz*:

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & -3\\ 3 & 9 & 4 & -7\\ 2 & -1 & 1 & 6 \end{array}\right)$$



La matriz

$$\begin{pmatrix} 1 & 3 & 1\\ 3 & 9 & 4\\ 2 & -1 & 1 \end{pmatrix}$$

recibe el nombre de *matriz de los coeficientes del sistema* y la matriz anterior recibe el nombre de *matriz ampliada del sistema*. En ésta, la línea vertical separa la matriz de los coeficientes de los términos independientes.

Utilizando la matriz ampliada del sistema e indicando con *i)*, *ii)* o *iii)* las operaciones elementales que se realizan sobre las filas de la matriz, de acuerdo con las operaciones elementales que se realizan sobre las ecuaciones, el ejemplo B puede resumirse de la siguiente manera:

$$\left(\begin{array}{ccc|c} 1 & 3 & 1 & -3\\ 3 & 9 & 4 & -7\\ 2 & -1 & 1 & 6 \end{array}\right) \xrightarrow[\substack{F_2-3F_1\\ F_3-2F_1}]{iii)} \left(\begin{array}{ccc|c} 1 & 3 & 1 & -3\\ 0 & 0 & 1 & 2\\ 0 & -7 & -1 & 12 \end{array}\right) \xrightarrow{ii)} \left(\begin{array}{ccc|c} 1 & 3 & 1 & -3\\ 0 & -7 & -1 & 12\\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow[\substack{F_1-F_3\\ F_2+F_3}]{iii)}$$

$$\left(\begin{array}{ccc|c} 1 & 3 & 0 & -5\\ 0 & -7 & 0 & 14\\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow[-\frac{1}{7}F_2]{i)} \left(\begin{array}{ccc|c} 1 & 3 & 0 & -5\\ 0 & 1 & 0 & -2\\ 0 & 0 & 1 & 2 \end{array}\right) \xrightarrow[F_1-3F_2]{iii)} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1\\ 0 & 1 & 0 & -2\\ 0 & 0 & 1 & 2 \end{array}\right)$$

La última matriz es la matriz del sistema

$$\left.\begin{aligned} x_1&=1\\ x_2&=-2\\ x_3&=2 \end{aligned}\right\}$$

que nos da las soluciones.

EJEMPLO C. Para resolver el sistema

$$\left.\begin{aligned} x_1+3x_2-x_3+x_4&=1\\ -2x_1+x_2+2x_3&=7\\ x_2-x_4&=0 \end{aligned}\right\}$$

de tres ecuaciones con cuatro incógnitas escribimos la matriz del sistema y la reducimos a la más sencilla posible mediante operaciones elementales. El lector no tendrá dificultad en seguir los pasos realizados.

$$\left(\begin{array}{cccc|c} 1 & 3 & -1 & 1 & 1\\ -2 & 1 & 2 & 0 & 7\\ 0 & 1 & 0 & -1 & 0 \end{array}\right) \xrightarrow[F_2+2F_1]{iii)} \left(\begin{array}{cccc|c} 1 & 3 & -1 & 1 & 1\\ 0 & 7 & 0 & 2 & 9\\ 0 & 1 & 0 & -1 & 0 \end{array}\right) \xrightarrow{ii)}$$

$$\left(\begin{array}{cccc|c} 1 & 3 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 7 & 0 & 2 & 9 \end{array}\right) \xrightarrow[F_3-7F_2]{iii)} \left(\begin{array}{cccc|c} 1 & 3 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 9 & 9 \end{array}\right) \xrightarrow[\frac{1}{9}F_3]{i)} \left(\begin{array}{cccc|c} 1 & 3 & -1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

$$\xrightarrow[\substack{F_1-F_3\\ F_2+F_3}]{iii)} \left(\begin{array}{cccc|c} 1 & 3 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right) \xrightarrow[F_1-3F_2]{iii)} \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & -3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

Esta última matriz corresponde al sistema

$$\left.\begin{array}{r} x_1-x_3=-3 \\ x_2=1 \\ x_4=1 \end{array}\right\}$$

Claramente $x_2=1$ y $x_4=1$, pero de la primera ecuación no pueden encontrarse $x_1$ y $x_3$; simplemente, dado un valor, $c$, a $x_3$ se obtiene un valor $x_1=-3+c$ para $x_1$ y las soluciones del sistema son:

$$\begin{aligned} x_1&=-3+c \\ x_2&=1 \\ x_3&=c \\ x_4&=1 \end{aligned}$$

o bien $(-3+c,\ 1,\ c,\ 1)$, para todo número real $c$. Este sistema tiene *infinitas* soluciones que se obtienen dando valores a $c$. ¡Realizar la comprobación!

* * *

Las dos matrices finales de los ejemplos B y C pueden escribirse de la forma

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right) \quad \text{y} \quad \left(\begin{array}{cccc|c} 1 & 0 & -1 & 0 & -3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right)$$

Estas dos matrices tienen en común que por ellas puede trazarse una «escalera descendente» tal que:

1) Cada peldaño tiene altura uno.
2) Debajo de la escalera todos los elementos de la matriz son cero.
3) En cada esquina de un peldaño aparece el número 1.
4) Toda columna que contiene un 1 en una esquina de un peldaño tiene todos los demás elementos nulos.



Toda matriz que posee estas propiedades será denominada una *matriz escalonada. El método de eliminación de Gauss consiste entonces en reducir la matriz de un sistema dado a una matriz escalonada mediante operaciones elementales.*

EJEMPLO D. Para resolver el sistema

$$\left.\begin{array}{r} x_1-x_2+x_3=1 \\ 2x_1+x_2+x_3=0 \\ 2x_1-2x_2+2x_3=3 \end{array}\right\}$$

escribimos su matriz ampliada y la transformamos mediante operaciones elementales hasta reducirla a una matriz escalonada:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 2 & 1 & 1 & 0 \\ 2 & -2 & 2 & 3 \end{array}\right) \xrightarrow[\substack{F_2-2F_1\\ F_3-2F_1}]{iii)} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 3 & -1 & -2 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\frac{1}{3}F_2]{i)} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 1 \\ 0 & 1 & -1/3 & -2/3 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[F_1+F_2]{iii)} \left(\begin{array}{ccc|c} 1 & 0 & 2/3 & 1/3 \\ 0 & 1 & -1/3 & -2/3 \\ 0 & 0 & 0 & 1 \end{array}\right) \xrightarrow[\substack{F_1-\frac{1}{3}F_3\\ F_2+\frac{2}{3}F_3}]{iii)} \left(\begin{array}{ccc|c} 1 & 0 & 2/3 & 0 \\ 0 & 1 & -1/3 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$$

Esta última matriz corresponde al sistema

$$\left.\begin{array}{r} x_1+(2/3)\ x_3=0 \\ x_2-(1/3)\ x_3=0 \\ 0=1 \end{array}\right\}$$

que, obviamente, *no tiene solución*, ya que $0\neq 1$.

* * *

*Observación.* El proceso que se sigue al aplicar a un sistema particular el método de eliminación de Gauss no es único, como el lector habrá podido comprobar si ha intentado realizar por su cuenta alguno de los ejemplos C o D. Sin embargo, la matriz escalonada de un sistema dado es única, ya que produce las soluciones del sistema. La demostración completa de este resultado no es sencilla.

* * *

A continuación damos algunas definiciones relativas a los sistemas hasta ahora estudiados. Una expresión de la forma

$$\left.\begin{array}{r} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \vdots \qquad\qquad \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{array}\right\} \qquad \text{(I)}$$

donde los $a_{ij}$ son números reales, se denomina un *sistema de m ecuaciones lineales con n incógnitas*. Una solución de (I) son $n$ números $(s_1, s_2, ..., s_n)$ tal que al sustituir $x_j$ por $s_j$ se obtiene una igualdad en todas las ecuaciones del sistema.

Si el sistema posee al menos una solución se dice que es *compatible* y si no posee ninguna solución se dice que es *incompatible*; si un sistema es compatible y tiene una única solución se dice que es *determinado* y si tiene más de una solución se dice que es *indeterminado*.

Los sistemas de los ejemplos A, B y C son compatibles, mientras que el del ejemplo D es incompatible. En los casos de compatibilidad, el del ejemplo C es indeterminado y los sistemas de los ejemplos A y B son determinados.

Los distintos casos que pueden presentarse en un sistema quedan caracterizados por la matriz escalonada que se obtiene en cada sistema. Recordemos que

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

se denomina la *matriz de los coeficientes del sistema*, y

$$\bar{A}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{pmatrix}$$

es la *matriz ampliada del sistema*. Las matrices escalonadas de la matriz de los coeficientes y de la matriz ampliada de los sistemas de los ejemplos B, C y D son las siguientes:

| | Matriz escalonada de la matriz de los coeficientes | Matriz escalonada de la matriz ampliada |
|---|---|---|
| Ejemplo B | $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ | $\left(\begin{array}{ccc\|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right)$ |
| Ejemplo C | $\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ | $\begin{pmatrix} 1 & 0 & -1 & 0 & -3 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$ |
| Ejemplo D | $\begin{pmatrix} 1 & 0 & 2/3 \\ 0 & 1 & -1/3 \\ 0 & 0 & 0 \end{pmatrix}$ | $\left(\begin{array}{ccc\|c} 1 & 0 & 2/3 & 0 \\ 0 & 1 & -1/3 & 0 \\ 0 & 0 & 0 & 1 \end{array}\right)$ |

Si convenimos en escribir:

$p$ = número de peldaños de una matriz escalonada de $A$,
$\bar{p}$ = número de peldaños de una matriz escalonada de $\bar{A}$,
$n$ = número de incógnitas del sistema,



las diferencias, en cuanto a soluciones, de los ejemplos anteriores quedan plasmadas en el siguiente cuadro:

| | | $p$ | $\bar{p}$ | $n$ |
|---|---|---|---|---|
| Ejemplo B | Sistema compatible determinado | 3 | 3 | 3 |
| Ejemplo C | Sistema compatible indeterminado | 3 | 3 | 4 |
| Ejemplo D | Sistema incompatible | 2 | 3 | 3 |

La observación detenida de este cuadro sugiere el siguiente resultado:

TEOREMA 1 (*Rouche-Frobenius*)

1) Un sistema es *compatible determinado* si y sólo si $p=\bar{p}=n$.
2) Un sistema es *compatible indeterminado* si y sólo si $p=\bar{p}<n$.
3) Un sistema es *incompatible* si y sólo si $p<\bar{p}$.

*Nota.* Las definiciones de $p, \bar{p}$ y $n$ pueden encontrarse entre los dos cuadros anteriores.

*Demostración.* Observemos en primer lugar que basta demostrar las implicaciones:

*a*) si $p=\bar{p}=n$, el sistema es compatible determinado,
*b*) si $p=\bar{p}<n$, el sistema es compatible indeterminado,
*c*) si $p<\bar{p}$, el sistema es incompatible,

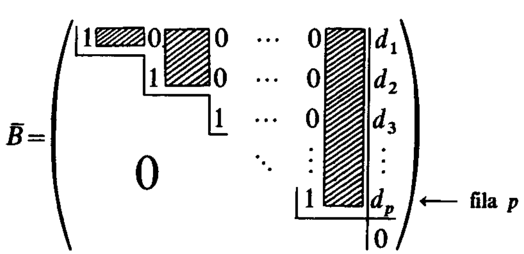
ya que el resto de las implicaciones contenidas en el teorema se deducen de estas tres. Se anima al lector a que compruebe esta última afirmación.

*a*) Si $p=\bar{p}=n$, la matriz escalonada $\bar{B}$ de este sistema es de la forma

$$\begin{array}{cl} \left(\begin{array}{ccccc|c} 1 & 0 & 0 & \cdots & 0 & d_1 \\ & 1 & 0 & & 0 & d_2 \\ & & 1 & \cdots & \vdots & \vdots \\ & 0 & & \ddots & 0 & \vdots \\ & & & & 1 & d_n \end{array}\right) & \begin{array}{l} \text{columna } n \\ \\ \\ \\ \leftarrow \text{fila } n \end{array} \end{array}$$

Por tanto, $x_1=d_1 \; x_2=d_2, ..., x_n=d_n$ es su única solución y el sistema es compatible y determinado.

*b)* Si $p=\bar{p}<n$, la matriz escalonada $\bar{B}$ de este sistema es de la forma




donde los rectángulos sombreados representan matrices. Si las columnas con 1 en una esquina de un peldaño corresponden a las incógnitas $x_1=x_{n_1}, x_{n_2}, \dots, x_{n_p}$ se tiene que

$$x_{n_1}=d_1, x_{n_2}=d_2, \dots, x_{n_p}=d_p, \quad x_j=0, \quad \forall j\neq n_1, n_2, \dots, n_p$$

es una solución del sistema. Por tanto, el sistema es compatible.

Falta demostrar que el sistema es indeterminado. Si todas las matrices sombreadas son nulas, el sistema es equivalente a

$$\left.\begin{array}{l} x_{n_1}=d_1 \\ x_{n_2}=d_2 \\ \vdots \\ x_{n_p}=d_p \end{array}\right\}$$

Puesto que $p<n$, alguna incógnita, digamos $x_k$, no aparece en el sistema anterior y, por tanto, $x_k$ puede tomar cualquier valor, con lo cual se obtienen infinitas soluciones del sistema.

Supongamos, por tanto, que no todas las matrices sombreadas son nulas. Sea

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_l \end{pmatrix}$$

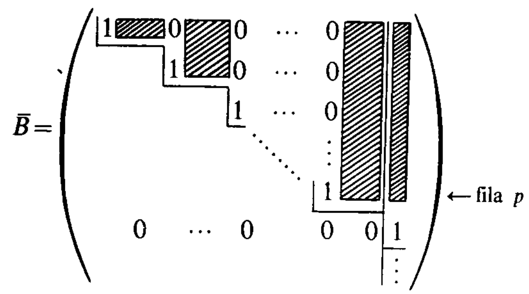
una columna no nula de una de estas matrices correspondiente a la incógnita $x_k$. Dando al resto de las incógnitas distintas de $x_{n_j}$, $j=1, 2, \dots, p$, y $x_k$ el valor cero, tenemos

$$x_{n_1}+e_1x_k=d_1;\ x_{n_2}+e_2x_k=d_2;\ \dots;\ x_{n_e}+e_lx_k=d_l;\ x_{n_{e+1}}=d_{e+1};\cdots;\ x_{n_p}=d_p$$

lo cual nos da infinitas soluciones del sistema dando valores a $x_k$. El sistema, por tanto, es indeterminado.



*c)* Si $p<\bar{p}$, la matriz escalonada $\bar{B}$ de este sistema es de la forma




La $(p+1)$-ésima fila de $\bar{B}$ nos da la ecuación $0=1$ y, por tanto, el sistema es incompatible. ■

* * *

Finalmente, el sistema (I) se dice que es *homogéneo* si todos los $b_j=0$ y *no homogéneo* si al menos uno de los $b_j$ es distinto de cero.

Al realizar operaciones elementales con la matriz ampliada

$$\bar{A}=\left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 \end{array}\right)$$

de un sistema homogéneo, la última columna está siempre formada por ceros; por tanto, nunca puede tener más peldaños que la matriz escalonada de $A$. Así pues, para un sistema *homogéneo* $p=\bar{p}$ y por el teorema de Rouche-Frobenius el sistema es siempre *compatible.*

Este resultado puede demostrarse más fácilmente observando que $x_1=0$, $x_2=0, \dots, x_n=0$ es siempre solución de un sistema homogéneo; esta solución se denomina *trivial.* Si un sistema homogéneo posee soluciones distintas de la trivial es indeterminado, y para esto es necesario y suficiente, según el teorema de Rouche-Frobenius, que $p=\bar{p}<n$.

EJEMPLO E. Queremos demostrar que el sistema homogéneo

$$\left.\begin{array}{r} x_1+x_2-x_3-x_4=0 \\ 2x_1-x_2+x_3-2x_4=0 \end{array}\right\}$$

es indeterminado. Realizando operaciones elementales con su matriz ampliada se tiene:

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 2 & -1 & 1 & -2 & 0 \end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 0 & -3 & 3 & 0 & 0 \end{array}\right) \xrightarrow[ii)]{i)}$$

$$\left(\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \end{array}\right)$$

Como $2=p=\bar{p}<n=4$, se obtiene el resultado deseado.

Sus soluciones se obtienen escribiendo el sistema correspondiente a la matriz escalonada:

$$\left.\begin{aligned} x_1 \qquad -x_4&=0\\ x_2-x_3 \qquad &=0\end{aligned}\right\}$$

Haciendo $x_3=c_1$, $x_4=c_2$ se tiene:

$$\begin{aligned} x_1&=c_2\\ x_2&=c_1\\ x_3&=c_1\\ x_4&=c_2\end{aligned}$$

donde $c_1$ y $c_2$ son números reales cualesquiera.

EJERCICIOS 1.1

**1.** Encontrar todas las soluciones (si existen) de cada uno de los siguientes sistemas de ecuaciones lineales mediante el método de eliminación de Gauss:

*a)* $\left.\begin{aligned} x_1+x_2&=1\\ x_1-x_2&=-1\end{aligned}\right\}$ *b)* $\left.\begin{aligned} x_1+x_2&=1\\ 2x_1+x_2&=0\end{aligned}\right\}$

**2.** Utilizar operaciones elementales para reducir las matrices dadas a su matriz escalonada.

*a)* $\begin{pmatrix} 1 & 3 & 7 & 4\\ 6 & -5 & -1 & 2\\ 7 & -2 & 6 & 1\end{pmatrix}$ *b)* $\begin{pmatrix} 0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$

*c)* $\begin{pmatrix} 3 & -1 & 3\\ 1 & 0 & -1\\ 3 & -1 & 1\\ -1 & 1 & -1\end{pmatrix}$ *d)* $\begin{pmatrix} 2 & 3 & 1 & -4 & 2\\ 1 & 2 & 1 & -2 & 3\\ 1 & 1 & 0 & 2 & 1\end{pmatrix}$

**3.** Resolver los siguientes sistemas mediante el método de eliminación de Gauss:

*a)* $\left.\begin{aligned} 2x_1+x_2+x_3&=3\\ x_1-x_2+x_3&=0\\ 3x_1-x_2+2x_3&=2\end{aligned}\right\}$ *b)* $\left.\begin{aligned} x_1+2x_2+4x_3&=1\\ x_1+x_2+3x_3&=2\\ 2x_1+5x_2+9x_3&=1\end{aligned}\right\}$



*c)* $\left.\begin{aligned} 2x_1-x_2+3x_3&=9\\ 3x_1-5x_2+x_3&=-4\\ 4x_1-7x_2+x_3&=5\end{aligned}\right\}$ *d)* $\left.\begin{aligned} 2x_1-2x_2+x_4&=-3\\ 2x_1+3x_2+x_3-3x_4&=-6\\ 3x_1+4x_2-x_3+2x_4&=0\\ x_1+3x_2+x_3-x_4&=2\end{aligned}\right\}$

*e)* $\left.\begin{aligned} 3x_1+x_2-x_3+2x_4-7&=0\\ 2x_1-2x_2+5x_3-7x_4-1&=0\\ -4x_1-4x_2+7x_3-11x_4+13&=0\end{aligned}\right\}$ *f)* $\left.\begin{aligned} x_1+x_2+x_3&=3\\ 2x_1-x_2+3x_3&=4\\ 3x_1+x_2-x_3&=3\\ 2x_1-2x_3&=0\end{aligned}\right\}$

*g)* $\left.\begin{aligned} x_1+2x_2+3x_3&=2\\ x_1-x_2+x_3&=0\\ x_1+3x_2-x_3&=-2\\ 3x_1+4x_2+3x_3&=0\end{aligned}\right\}$ *h)* $\left.\begin{aligned} x_1-3x_2+2x_3&=0\\ -x_1-2x_2+2x_3&=0\\ -2x_1+x_2&=0\end{aligned}\right\}$

*i)* $\left.\begin{aligned} x_1+x_2+x_3+x_4&=0\\ x_2-x_4&=5\\ x_1+x_3+2x_4&=1\\ x_1+2x_2&=0\end{aligned}\right\}$ *j)* $\left.\begin{aligned} x_1+x_2+x_3-6&=0\\ x_1+2x_2+2x_3-9&=0\\ x_1+2x_2+3x_3-10&=0\end{aligned}\right\}$

**4.** Utilizar el teorema de Rouche-Frobenius para realizar un estudio del número de soluciones de los siguientes sistemas:

*a)* $\left.\begin{aligned} x_1+2x_2-x_3&=1\\ x_2+x_3&=2\\ x_1+3x_2&=3\end{aligned}\right\}$ *b)* $\left.\begin{aligned} 3x_1+2x_2+x_3+x_4-x_5&=3\\ x_1+2x_3+x_5&=3\end{aligned}\right\}$

*c)* $\left.\begin{aligned} 2x_1+x_2&=5\\ x_1-x_2&=1\\ x_1+2x_2&=0\end{aligned}\right\}$

**5.** Demostrar que el sistema

$$\left.\begin{aligned} a_{11}x_1+a_{12}x_2&=0\\ a_{21}x_1+a_{22}x_2&=0\end{aligned}\right\}$$

es compatible indeterminado si y sólo si $a_{11}a_{22}-a_{12}a_{21}=0$.

## 1.2. RANGO DE UNA MATRIZ. ESTRUCTURA DE LAS SOLUCIONES DE UN SISTEMA

En la sección anterior se observó que el número de peldaños de la matriz escalonada de un sistema es importante para determinar la compatibilidad o incompatibilidad de un sistema. En esta sección se relacionará este número con el importante concepto de *rango de una matriz*. Además, se estudiará la estructura de las soluciones de un sistema de ecuaciones lineales.

En los ejemplos de la sección anterior se ha observado que las soluciones de un sistema pueden escribirse de la forma

$$(a_1, a_2, \ldots, a_n)$$

donde $a_1, a_2, \ldots, a_n$ son números reales. Un elemento de esta forma recibe el nombre de *vector* y se denota por

$$\vec{a}=(a_1, a_2, \ldots, a_n)$$

Los números reales $a_j, j=1, 2, \ldots, n$, reciben el nombre de *componentes* del vector $\vec{a}$; $a_1$ se denomina *primera componente*, $a_2$ *segunda componente*, y así sucesivamente. El vector $\vec{0}$ es el vector cuyas componentes son todas nulas. Dos vectores $\vec{a}=(a_1, a_2, \ldots, a_n)$ y $\vec{b}=(b_1, b_2, \ldots, b_n)$ son iguales, y escribiremos $\vec{a}=\vec{b}$, cuando $a_1=b_1, a_2=b_2, \ldots, a_n=b_n$.

Los vectores no sólo aparecen como soluciones de un sistema de ecuaciones lineales, sino que también aparecen en las filas o en las columnas de una matriz, en cuyo caso reciben el nombre de vectores fila o vectores columna de la matriz. Por ejemplo, la matriz

$$\begin{pmatrix} 1 & 2 & 0 & 1 \\ 2 & 1 & 3 & 0 \\ -1 & 2 & -1 & 1 \end{pmatrix}$$

tiene como vectores fila

$$\vec{a}=(1, 2, 0, 1), \quad \vec{b}=(2, 1, 3, 0), \quad \vec{c}=(-1, 2, -1, 1)$$

y como vectores columna:

$$\vec{d}=(1, 2, -1), \quad \vec{e}=(2, 1, 2), \quad \vec{f}=(0, 3, -1), \quad \vec{g}=(1, 0, 1)$$

*Nota.* A veces los vectores columna de una matriz se escriben en notación vertical de la forma

$$\vec{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$



Las definiciones que daremos con vectores no dependen de la notación que se utilice para representarlos. En este texto aparecerán las dos notaciones indistintamente.

Al resolver un sistema por el método de eliminación de Gauss se han realizado ciertas operaciones con los vectores fila de su matriz ampliada, que recibían el nombre de operaciones elementales. Estas operaciones son la *suma de vectores*:

$$\vec{a}+\vec{b}=(a_1, a_2, \ldots, a_n)+(b_1, b_2, \ldots, b_n)=(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$$

y la *multiplicación de un vector par un número real c*:

$$c\vec{a}=c(a_1, a_2, \ldots, a_n)=(ca_1, ca_2, \ldots, ca_n)$$

Las operaciones con vectores poseen las siguientes propiedades, que se dejan como ejercicio para el lector:



($S_1$) $\vec{a}+\vec{b}=\vec{b}+\vec{a}$ (conmutativa)
($S_2$) $(\vec{a}+\vec{b})+\vec{c}=\vec{a}+(\vec{b}+\vec{c})$ (asociativa)
($S_3$) $\vec{0}+\vec{a}=\vec{a}+\vec{0}=\vec{a}$
($S_4$) $\vec{a}+(-\vec{a})=(-\vec{a})+\vec{a}=\vec{0}$, donde $(-\vec{a})=(-1)\vec{a}$
($M_1$) $c(\vec{a}+\vec{b})=c\vec{a}+c\vec{b}$
($M_2$) $(c+d)\vec{a}=c\vec{a}+d\vec{a}$
($M_3$) $c(d\vec{a})=(cd)\vec{a}$
($M_4$) $1\vec{a}=\vec{a}$

*Nota.* El vector $(-\vec{a})=(-1)\vec{a}$ se denomina *opuesto* de $\vec{a}$.

Dado un conjunto de vectores $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_s\}$, una expresión de la forma

$$d_1\vec{a}_1+d_2\vec{a}_2+\cdots+d_s\vec{a}_s$$

donde los $d_j, j=1, \ldots, s$, son números reales, se dice que es una *combinación lineal* de los vectores dados.

Un vector $\vec{a}$ se dice que es *combinación lineal* de los vectores $\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_s\}$ si existen números reales $d_1, d_2, \ldots, d_s$ tal que

$$\vec{a}=d_1\vec{a}_1+d_2\vec{a}_2+\cdots+d_s\vec{a}_s$$

El vector $\vec{0}$ es combinación lineal de cualquier conjunto de vectores, ya que basta tomar todos los $d_j=0$. La expresión anterior puede escribirse de la forma

$$d_1\vec{a}_1+d_2\vec{a}_2+\cdots+d_s\vec{a}_s+(-1)\vec{a}=\vec{0}$$

donde no todos los coeficientes de los vectores son nulos (el coeficiente de $\vec{a}$ es $-1$).

Esto sugiere la siguiente definición.

DEFINICIÓN 1

*i)* Un conjunto de vectores $\vec{a}_1, \vec{a}_2, ..., \vec{a}_k$ es *linealmente dependiente* si existen números $d_1, d_2, ..., d_k$, no todos nulos, tal que

$$d_1\vec{a}_1 + d_2\vec{a}_2 + \cdots + d_k\vec{a}_k = \vec{0}$$

*ii)* Un conjunto de vectores $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$ se dice *linealmente independiente* si no son linealmente dependientes, es decir, cualquier expresión del tipo

$$d_1\vec{a}_1 + d_2\vec{a}_2 + \cdots + d_k\vec{a}_k = \vec{0}$$

implica necesariamente que $d_1 = d_2 = \cdots = d_k = 0$.

EJEMPLO A. Queremos estudiar si los vectores $\vec{a}_1 = (2, 1)$ y $\vec{a}_2 = (1, 1)$ son linealmente dependientes o independientes; para ello tratemos de encontrar dos números $d_1$ y $d_2$ tal que

$$d_1(2, 1) + d_2(1, 1) = (0, 0)$$

Puesto que dos vectores son iguales si y sólo si sus componentes correspondientes coinciden, la igualdad anterior puede escribirse de la forma

$$\left.\begin{aligned} 2d_1 + d_2 &= 0 \\ d_1 + d_2 &= 0 \end{aligned}\right\}$$

que es un sistema de dos ecuaciones con dos incógnitas. Realizando operaciones elementales con la matriz de este sistema se tiene:

$$\left(\begin{array}{cc|c} 2 & 1 & 0 \\ 1 & 1 & 0 \end{array}\right) \xrightarrow{ii)} \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 2 & 1 & 0 \end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & -1 & 0 \end{array}\right)$$

$$\xrightarrow{i)} \left(\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \end{array}\right) \longrightarrow \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right)$$

Como $p = \bar{p} = 2 =$ número de incógnitas y el sistema es homogéneo, sólo tiene la solución trivial $d_1 = d_2 = 0$. Esto implica que los vectores $\vec{a}_1, \vec{a}_2$ son *linealmente independientes.*

* * *

EJEMPLO B. Para estudiar si los vectores $\vec{a}_1 = (1, 1, 3)$, $\vec{a}_2 = (0, 1, 2)$ y $\vec{a}_3 = (1, 2, 5)$ son linealmente dependientes o independientes formamos la expresión

$$d_1(1, 1, 3) + d_2(0, 1, 2) + d_3(1, 2, 5) = (0, 0, 0)$$



que puede escribirse de la forma:

$$\left.\begin{aligned} d_1 + d_3 &= 0 \\ d_1 + d_2 + 2d_3 &= 0 \\ 3d_1 + 2d_2 + 5d_3 &= 0 \end{aligned}\right\}$$

Puesto que

$$\left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 3 & 2 & 5 & 0 \end{array}\right) \xrightarrow{iii)} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \end{array}\right) \xrightarrow{iii)} \left(\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

se tiene que $p = \bar{p} = 2 < 3 =$ número de incógnitas y, por tanto, el sistema posee soluciones no triviales: los vectores dados son *linealmente dependientes.*

Si queremos encontrar una combinación lineal de los vectores anteriores basta resolver el sistema anterior; dicho sistema es equivalente a

$$\left.\begin{aligned} d_1 + d_3 &= 0 \\ d_2 + d_3 &= 0 \end{aligned}\right\}$$

de acuerdo con las operaciones elementales anteriormente realizadas. Haciendo $d_3 = c$, se tiene $d_1 = -c$, $d_2 = -c$. Como caso particular de $c$ podemos tomar $c = -1$, con lo cual $d_1 = 1$, $d_2 = 1$, y se tiene:

$$1(1, 1, 3) + 1(0, 1, 2) - 1(1, 2, 5) = (0, 0, 0)$$

como fácilmente puede comprobarse.

* * *

En los ejemplos anteriores se habrá observado que únicamente es necesario escribir la matriz cuyas columnas son los vectores dados y realizar en ella operaciones elementales para estudiar la dependencia o independencia lineal de un conjunto de vectores. Si el número de peldaños coincide con el número de columnas (o equivalentemente la matriz escalonada es la identidad) los vectores son linealmente independientes y en caso contrario son linealmente dependientes.

DEFINICIÓN 2 (*Rango de un conjunto de vectores y rango de una matriz*)

*i)* Se denomina *rango* de un conjunto de vectores al mayor número de ellos que son linealmente independientes.

*ii)* Se denomina *rango de una matriz A*, y se denota por $r(A)$, al rango de sus vectores columna.

El rango de los vectores del ejemplo A es 2, ya que son linealmente independientes.

El rango de los vectores del ejemplo B es menor que 3, ya que son linealmente dependientes.

En estos momentos es conveniente resaltar la forma de calcular el rango de una matriz. Dada la matriz

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

se consideran los vectores

$$\vec{a}_1=\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \quad \vec{a}_2=\begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \ldots, \vec{a}_n=\begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}$$

de entre los cuales hay que determinar el mayor número de ellos que sean linealmente independientes.

En primer lugar es necesario estudiar si los $n$ vectores son linealmente independientes, es decir, si el sistema homogéneo

$$x_1\vec{a}_1+x_2\vec{a}_2+\cdots+x_n\vec{a}_n=\vec{0} \tag{1}$$

posee únicamente la solución nula. En caso de que posea únicamente la solución nula, el rango de la matriz es $n$. Observar que para determinar si (1) posee soluciones no nulas únicamente es necesario reducir la matriz $A$ a su forma escalonada.

Si los $n$ vectores son linealmente dependientes, es necesario estudiar si alguno de los posibles subconjuntos de $n-1$ vectores de entre los $n$ anteriores es linealmente independiente.

Si estos subconjuntos de $n-1$ vectores son todos linealmente dependientes es necesario estudiar todos los subconjuntos de $n-2$ vectores. El proceso termina cuando encontremos por primera vez unos cuantos vectores de entre los anteriores que sean linealmente independientes.

En teoría puede parecer complicado y largo calcular el rango de una matriz. En la práctica, sin embargo, resulta sencillo, como se muestra en el ejemplo siguiente.

EJEMPLO C. Para encontrar el rango de la matriz

$$A=\begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 0 & -1 & -2 \\ 3 & 4 & 1 & -2 \end{pmatrix}$$



comenzamos resolviendo el sistema

$$x_1\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}+x_2\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}+x_3\begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}+x_4\begin{pmatrix} 1 \\ -1 \\ -2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

de tres ecuaciones con cuatro incógnitas. Puesto que

$$\begin{pmatrix} 1 & 3 & 2 & 1 \\ 1 & 0 & -1 & -2 \\ 3 & 4 & 1 & -2 \end{pmatrix} \xrightarrow{iii)} \begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & -3 & -3 & -3 \\ 0 & -5 & -5 & -5 \end{pmatrix} \xrightarrow{iii)} \begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & -3 & -3 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{i)} \begin{pmatrix} 1 & 3 & 2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

el sistema posee soluciones no triviales. Por tanto, los cuatro vectores columna de la matriz son linealmente dependientes.

Si tomamos tres de ellos, por ejemplo, los tres primeros, la matriz escalonada del sistema que ellos forman es

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

ya que se obtiene realizando las mismas operaciones elementales que antes sobre las tres primeras columnas de la matriz. Estos tres vectores son, por tanto, linealmente dependientes.

El lector puede comprobar que cualesquiera tres vectores de entre los anteriores son linealmente dependientes

Finalmente, si tomamos los dos primeros vectores, el sistema

$$x_1\begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix}+x_2\begin{pmatrix} 3 \\ 0 \\ 4 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

tiene como matriz escalonada

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

y, por tanto, son linealmente independientes.

Concluimos, entonces, que $r(A)=2$.

* * *

El lector habrá podido observar en el ejemplo anterior que el rango de una matriz coincide con el número de peldaños de su matriz escalonada. Este resultado se demuestra a continuación.

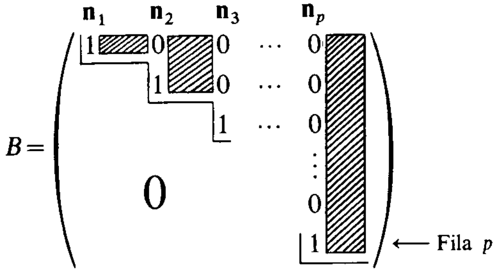
TEOREMA 1

> El rango de una matriz coincide con el número de peldaños de su matriz escalonada.

*Demostración.* Sea $A$ una matriz de la forma

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

y denotemos por $\vec{a}_1, \vec{a}_2, ..., \vec{a}_n$ sus vectores columna. Sea $p$ el número de peldaños de una matriz escalonada de $A$. Esta matriz escalonada es de la forma:




El sistema homogéneo

$$x_1\vec{a}_{n_1}+x_2\vec{a}_{n_2}+\cdots+x_p\vec{a}_{n_p}=\vec{0}$$

correspondiente a las columnas no sombreadas de la matriz tiene únicamente la solución trivial, ya que su matriz escalonada es

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ & 1 & 0 & \cdots & 0 \\ & & 1 & & \vdots \\ & & & \ddots & 0 \\ & 0 & & & 1 \end{pmatrix} \begin{array}{l} \text{Columna } p \\ \\ \\ \\ \leftarrow \text{Fila } p \end{array}$$

Hemos demostrado, por tanto, que la matriz $A$ posee al menos $p$ vectores columna linealmente independientes.



El teorema quedará demostrado si probamos que no existen más de $p$ vectores columna de $A$ que sean linealmente independientes.

Tomemos $q$ vectores columna $\vec{a}_{m_1}, \vec{a}_{m_2}, ..., \vec{a}_{m_q}$ de $A$, con $q>p$. El sistema

$$x_1\vec{a}_{m_1}+x_2\vec{a}_{m_2}+\cdots+x_q\vec{a}_{m_q}=\vec{0} \tag{2}$$

posee una matriz escalonada con $p$ peldaños como máximo, ya que la matriz escalonada de $A$ tiene $p$ peldaños. Puesto que el sistema anterior tiene $q$ incógnitas $y$, $q$ es mayor que $p$, (2) es indeterminado. Por tanto, los $q$ vectores dados son linealmente dependientes. ■

EJEMPLO D. Para calcular el rango de los vectores $\vec{a}_1=(1, 1, 3)$, $\vec{a}_2=(2, 2, 6)$ y $\vec{a}_3=(2, -1, 5)$ escribimos la matriz

$$\begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ 3 & 6 & 5 \end{pmatrix}$$

cuyas columnas son los vectores dados, y realizamos sobre ella operaciones elementales:

$$\begin{pmatrix} 1 & 2 & 2 \\ 1 & 2 & -1 \\ 3 & 6 & 5 \end{pmatrix} \xrightarrow{iii)} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & -3 \\ 0 & 0 & -1 \end{pmatrix} \xrightarrow{i)} \begin{pmatrix} 1 & 2 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

$$\xrightarrow{iii)} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Puesto que el número de peldaños de la matriz escalonada es 2, del teorema anterior deducimos que el rango de los tres vectores dados es 2.

* * *

El teorema 1 nos permite escribir el teorema de *Rouche-Frobenius* utilizando el rango de una matriz, tal y como se hace en la mayor parte de la literatura matemática.

TEOREMA DE ROUCHE-FROBENIUS (véase la sección 1.1)

> Dado un sistema de $m$ ecuaciones lineales con $n$ incógnitas, con matriz de sus coeficientes $A$ y matriz ampliada $\bar{A}$, se tienen los siguientes resultados:
>
> *i)* El sistema es *compatible determinado* si y sólo si
>
> $$r(A)=r(\bar{A})=n$$

*ii)* El sistema es *compatible indeterminado* si y sólo si

$$r(A)=r(\bar{A})<n$$

*iii)* Un sistema es *incompatible* si y sólo si

$$r(A)<r(\bar{A})$$

* * *

A continuación se realiza el estudio de la estructura de las soluciones de un sistema. Comenzamos con un sistema homogéneo de $m$ ecuaciones y $n$ incógnitas:

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{array}\right\} \qquad \text{(I)}$$

Sabemos que todo sistema homogéneo posee la solución trivial $\vec{0}=(0, 0, ..., 0)$ y, por tanto, es compatible. Demostraremos a continuación que si es indeterminado ha de tener infinitas soluciones.

PROPOSICIÓN 2

*i)* Si $\vec{u}=(u_1, u_2, ..., u_n)$ es una solución del sistema homogéneo (I), $c\vec{u}$ es también solución del mismo sistema para todo número real $c$.

*ii)* Si $\vec{u}=(u_1, u_2, ..., u_n)$ y $\vec{v}=(v_1, v_2, ..., v_n)$ son soluciones del sistema homogéneo (I), $\vec{u}+\vec{v}$ también lo es.

*Demostración.* Si $\vec{u}$ es solución de (I) se tienen las igualdades:

$$\begin{array}{l} a_{11}u_1+a_{12}u_2+\cdots+a_{1n}u_n=0 \\ a_{21}u_1+a_{22}u_2+\cdots+a_{2n}u_n=0 \\ \vdots \\ a_{m1}u_1+a_{n2}u_2+\cdots+a_{mn}u_n=0 \end{array}$$

Multiplicando por el número real $c$ cada una de estas igualdades se obtiene el resultado deseado.

Si, además, $\vec{v}$ es solución de (I) se tienen las igualdades:

$$\begin{array}{l} a_{11}v_1+a_{12}v_2+\cdots+a_{1n}v_n=0 \\ a_{21}v_1+a_{22}v_2+\cdots+a_{2n}v_n=0 \\ \vdots \\ a_{m1}v_1+a_{n2}v_2+\cdots+a_{mn}v_n=0 \end{array}$$



Sumando las correspondientes igualdades se obtiene:

$$\begin{array}{l} a_{11}(u_1+v_1)+a_{12}(u_2+v_2)+\cdots+a_{1n}(u_n+v_n)=0 \\ a_{21}(u_1+v_1)+a_{22}(u_2+v_2)+\cdots+a_{2n}(u_n+v_n)=0 \\ \\ a_{m1}(u_1+v_1)+a_{m2}(u_2+v_2)+\cdots+a_{mn}(u_n+v_n)=0 \end{array}$$

lo cual prueba que $\vec{u}+\vec{v}$ es también solución del mismo sistema. ■

PROPOSICIÓN 3

Si el sistema homogéneo (I) es indeterminado existen $k$ vectores, $\vec{u}_1, \vec{u}_2, ..., \vec{u}_k$, linealmente independientes de manera que todas las soluciones de (I) son de la forma

$$c_1\vec{u}_1+c_2\vec{u}_2+\cdots+c_k\vec{u}_k$$

con los $c_j$ números reales. Además, $k=n-r(A)$, donde $A$ denota la matriz de los coeficientes del sistema (I).

Es conveniente ilustrar con un ejemplo el resultado de la proposición 3.

EJEMPLO E. Tratemos de encontrar las soluciones del sistema homogéneo

$$\left.\begin{array}{r} 2x_1-3x_2+x_3-x_4+2x_5=0 \\ 3x_1-x_3+x_4=0 \\ x_1+3x_2-2x_3+2x_4-2x_5=0 \end{array}\right\}$$

Realizando operaciones elementales con las filas de la matriz de sus coeficientes se obtiene:

$$\begin{pmatrix} 2 & -3 & 1 & -1 & 2 \\ 3 & 0 & -1 & 1 & 0 \\ 1 & 3 & -2 & 2 & -2 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & -2 \\ 3 & 0 & -1 & 1 & 0 \\ 2 & -3 & 1 & -1 & 2 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & -2 \\ 0 & -9 & 5 & -5 & 6 \\ 0 & -9 & 5 & -5 & 6 \end{pmatrix}$$

$$\longrightarrow \begin{pmatrix} 1 & 3 & -2 & 2 & -2 \\ 0 & 1 & -\frac{5}{9} & \frac{5}{9} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -\frac{1}{3} & \frac{1}{3} & -3 \\ 0 & 1 & -\frac{5}{9} & \frac{5}{9} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Haciendo $x_3=c_1$, $x_4=c_2$ y $x_5=c_3$ se tiene que

$$x_1=\frac{1}{3}c_1-\frac{1}{3}c_2+3c_3$$

$$x_2=\frac{5}{9}c_1-\frac{5}{9}c_2-\frac{1}{3}c_3$$

Las soluciones de este sistema pueden escribirse de la forma:

$$(x_1, x_2, x_3, x_4, x_5)=\left(\frac{1}{3}c_1-\frac{1}{3}c_2+3c_3,\ \frac{5}{9}c_1-\frac{5}{9}c_2-\frac{1}{3}c_3,\ c_1, c_2, c_3\right)=$$

$$=c_1\left(\frac{1}{3}, \frac{5}{9}, 1, 0, 0\right)+c_2\left(-\frac{1}{3}, -\frac{5}{9}, 0, 1, 0\right)+$$

$$+c_3\left(3, -\frac{1}{3}, 0, 0, 1\right).$$

Los vectores

$$\vec{u}_1=\left(\frac{1}{3}, \frac{5}{9}, 1, 0, 0\right),\quad \vec{u}_2=\left(-\frac{1}{3}, -\frac{5}{9}, 0, 1, 0\right)\quad \text{y}\quad \vec{u}_3=\left(3, -\frac{1}{3}, 0, 0, 1\right)\ \text{son lineal-}$$

mente independientes, ya que si tenemos

$$d_1\vec{u}_1+d_2\vec{u}_2+d_3\vec{u}_3=\vec{0},$$

igualando las tres últimas componentes de la izquierda a cero se deduce que $d_1=d_2=d_3=0$.

Observar, finalmente, que el número de vectores linealmente independientes es $3=n-r(A)$.

* * *

*Demostración de la proposición* 3. Supongamos, para simplificar, que la matriz escalonada de este sistema es de la forma

$$\begin{array}{cc} & \phantom{B=}\qquad\qquad\qquad\qquad \text{Columna } n=p+k \\ B= & \left(\begin{array}{ccccc|ccc} 1 & 0 & 0 & \cdots & 0 & u_{11} & \cdots & u_{1k} \\ & 1 & 0 & \cdots & 0 & u_{21} & \cdots & u_{2k} \\ & & 1 & & \vdots & \vdots & & \vdots \\ & & & & 0 & & & \\ & & & & 1 & u_{p1} & \cdots & u_{pk} \end{array}\right) \leftarrow \text{Fila } p \end{array}$$



donde $p=r(A)$. Tomar

$$\vec{u}_1=(-u_{11}, -u_{21}, \ldots, -u_{p1}, 1, 0, \ldots, 0)$$
$$\vec{u}_2=(-u_{11}, -u_{22}, \ldots, -u_{p2}, 0, 1, \ldots, 0)$$
$$\vdots$$
$$\vec{u}_k=(-u_{1k}, -u_{2k}, \ldots, -u_{pk}, 0, 0, \cdots, 1)$$

Por un razonamiento análogo al del ejemplo E se concluye que toda solución de (I) es de la forma

$$d_1\vec{u}_1+d_2\vec{u}_2+\cdots+d_k\vec{u}_k.$$

Además, los $k$-vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ son linealmente independientes, ya que si tenemos una combinación lineal de ellos de la forma

$$d_1\vec{u}_1+d_2\vec{u}_2+\cdots+d_k\vec{u}_k=\vec{0}$$

las $n-p$ últimas componentes de cada uno de ellos producen las igualdades $d_1=d_2=\cdots=d_k=0$. Por tanto, los $k$-vectores dados son linealmente independientes. ■

* * *

La estructura de las soluciones de un sistema no homogéneo se deduce de la estructura de las soluciones de un sistema homogéneo.

Sea

$$\left.\begin{array}{l} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \vdots \qquad\qquad\qquad\qquad \vdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_n \end{array}\right\} \qquad \text{(II)}$$

un sistema de $m$ ecuaciones con $n$ incógnitas. Se denomina *sistema homogéneo asociado* a (II) el sistema que se obtiene sustituyendo los $b_j$ de la derecha del sistema por ceros.

PROPOSICIÓN 4

Si $\vec{v}$ es una solución de (II), todas sus soluciones son de la forma $\vec{v}+\vec{u}$, donde $\vec{u}$ es solución de su sistema homogéneo asociado.

*Demostración.* Si $\vec{w}$ es otra solución de (II) escribimos

$$\vec{w}=\vec{v}+(\vec{w}-\vec{v})$$

y observamos que $\vec{u}=\vec{w}-\vec{v}$ es una solución del sistema homogéneo asociado, ya que $\vec{w}$ y $\vec{v}$ son ambas soluciones del sistema (II). ■

Las proposiciones 3 y 4 nos permiten enunciar el siguiente resultado.

TEOREMA 5

Sea $\vec{v}$ una solución de (II). Existen $k$ vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ linealmente independientes, tal que todas las soluciones de (II) son de la forma

$$\vec{v} + c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k$$

donde los $c_k$ son números reales y $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ son soluciones del sistema homogéneo asociado a (II).

Además, $k = n - r(A)$, donde $A$ es la matriz de los coeficientes del sistema.

*Nota.* La expresión $\vec{v} + c_1\vec{u}_1 + c_2\vec{u}_2 + \cdots + c_k\vec{u}_k$ se denomina *solución general del sistema* y $\vec{v}$ se denomina una *solución particular* del sistema.

## EJERCICIOS 1.2

**1.** Demostrar las propiedades $(S_1)$, $(S_2)$, $(S_3)$ y $(S_4)$, $(M_1)$, $(M_2)$, $(M_3)$ y $(M_4)$ de la suma de vectores y de la multiplicación de vectores por un número real.

**2.** Determinar si los siguientes conjuntos de vectores son linealmente dependientes o independientes y en caso de que sean linealmente dependientes, encontrar una combinación lineal entre ellos:

*a)* $\{(1, 2), (2, 4)\}$
*b)* $\{(3, 5, 1), (2, 1, 3)\}$
*c)* $\{(1, 2, 3), (1, 3, 2), (0, -1, 1)\}$
*d)* $\{(1, 0, 1, 0), (2, 1, 3, 1), (0, 1, 1, 1), (2, 2, 4, 2)\}$

**3.** Calcular el rango de los siguientes conjuntos de vectores:

*a)* $\begin{pmatrix}1\\2\end{pmatrix}, \begin{pmatrix}3\\2\end{pmatrix}$ *b)* $\begin{pmatrix}4\\5\end{pmatrix}, \begin{pmatrix}1\\5/4\end{pmatrix}$ *c)* $\begin{pmatrix}1\\3\\7\end{pmatrix}, \begin{pmatrix}6\\-5\\-1\end{pmatrix}, \begin{pmatrix}7\\-2\\6\end{pmatrix}$

*d)* $\begin{pmatrix}1\\0\\4\end{pmatrix}, \begin{pmatrix}7\\1\\3\end{pmatrix}, \begin{pmatrix}-6\\0\\1\end{pmatrix}$ *e)* $\begin{pmatrix}5\\9\\1\\6\end{pmatrix}, \begin{pmatrix}-4\\3\\2\\-2\end{pmatrix}, \begin{pmatrix}1\\2\\1\\2\end{pmatrix}$



**4.** Calcular el rango de las siguientes matrices:

*a)* $\begin{pmatrix}1 & 3 & 7 & 4\\ 6 & -5 & -1 & 2\\ 7 & -2 & 6 & 1\end{pmatrix}$ *b)* $\begin{pmatrix}4 & -3 & 2 & 5 & 7\\ 6 & -9 & 4 & 2 & 3\\ -1 & 0 & 5 & 7 & 4\\ 0 & 1 & 0 & 3 & 2\end{pmatrix}$

*c)* $\begin{pmatrix}1 & 2 & 1 & 4 & 3 & 6\\ 0 & 1 & 4 & 7 & 3 & 2\\ 1 & 1 & 3 & 2 & -1 & 0\\ 3 & 2 & -1 & -2 & 0 & 1\\ 0 & 1 & 0 & 2 & 0 & -1\end{pmatrix}$ *d)* $\begin{pmatrix}2 & 3 & 1 & -4 & 2\\ 1 & 2 & 1 & -2 & 3\\ 2 & 3 & 0 & 2 & 1\\ 1 & 1 & 0 & 2 & 1\end{pmatrix}$

**5.** Estudiar su compatibilidad y encontrar la solución general de los sistemas

$$\left.\begin{aligned} x_1 + x_2 - x_3 &= 1\\ x_1 - 2x_2 + x_3 &= 0\\ x_2 - x_3 &= 0\end{aligned}\right\} \qquad \left.\begin{aligned} x_1 \qquad + x_3 - x_4 &= 5\\ x_2 + x_3 + x_4 &= 2\end{aligned}\right\}$$

**6.** Demostrar que todo conjunto con $n+1$ vectores de $n$ componentes cada uno es linealmente dependiente.

## 1.3. APLICACIONES LINEALES DE $\mathbb{R}^n$ EN $\mathbb{R}^m$ Y OPERACIONES CON MATRICES

En esta sección deduciremos las operaciones con matrices a partir de las operaciones que pueden realizarse con aplicaciones lineales. Tales operaciones con matrices serán necesarias para un estudio posterior de la resolución de sistemas de ecuaciones lineales. Comenzaremos con el concepto de aplicación entre conjuntos.

Dados dos conjuntos $S$ y $T$, toda ley que asocia a cada uno de los elementos de $S$ un elemento de $T$ como máximo se denomina una *aplicación de $S$ en $T$.* Si a esta ley la representamos con la letra $f$ se acostumbra a escribir $f: S \mapsto T$, lo cual se lee «*$f$ es una aplicación de $S$ en $T$*». El elemento de $T$ asociado con el elemento $s$ de $S$ se escribe mediante $f(s)$ y recibe el nombre de *imagen del elemento s.*

EJEMPLO A. 1) Si $\mathbb{R}$ denota el conjunto de los números reales, $f: \mathbb{R} \mapsto \mathbb{R}$ dada por $f(x) = 3x$, es decir, a cada número real le asocia su triple, es una aplicación.

2) Si $\mathbb{N}$ denota el conjunto de los números naturales, 0, 1, 2, 3, ..., la ley $f: \mathbb{N} \mapsto \mathbb{N}$ que asocia a todo número natural $n$ su cuadrado, es decir, $f(n) = n^2$, es una aplicación.

3) Si $S = \{1, 2, 3, 4\}$ y $T = \{1, 2\}$ y definimos $f(1) = 1$, $f(2) = 1$, $f(3) = 2$, $f(4) = 2$ (obsérvese la figura 1.2), se tiene una aplicación de $S$ en $T$.

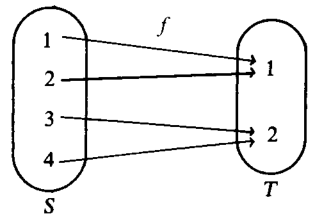



Figura 1.2

*Nota.* Cuando una aplicación está definida en un conjunto de números (naturales, enteros, racionales, reales, etc.) y sus imágenes están también en un conjunto de números, se suele utilizar la palabra *función* en lugar de «aplicación». En algunos casos se suele utilizar la palabra *transformación* en lugar de «aplicación».

Dada una aplicación $f$: $S \mapsto T$ se denomina *imagen de f*, y se denota por $im(f)$ al conjunto de todas las imágenes de los elementos de $S$, es decir:

$$\mathrm{im}(f) = \{f(s) / s \in S\}$$

En el ejemplo A.1), $\mathrm{im}(f) = \mathbb{R}$; en el A.2), $\mathrm{im}(f) = \{0, 1, 4, 9, 16, 25, \ldots\}$, y en el A.3), $\mathrm{im}(f) = \{1, 2\}$.

Cuando el conjunto $\mathrm{im}(f)$ coincide con el conjunto final $T$ se dice que $f$ es una aplicación *suprayectiva*; A.1) y A.3) son ejemplos de aplicaciones suprayectivas, mientras que A.2) no lo es. Otra forma de comprobar que $f$ es *suprayectiva* es estudiando si todo elemento de $T$ es imagen de algún elemento de $S$, es decir,

«para todo $t \in T$, existe $s \in S$, tal que $f(s) = t$»

Una aplicación $f$: $S \mapsto T$ se denomina *inyectiva* si dos elementos distintos cualesquiera de $S$ tienen distintas imágenes, es decir, $s \neq s'$, $s, s' \in S$, $\Rightarrow f(s) \neq f(s')$. Otra forma de comprobar que $f$ es inyectiva es utilizando la negación de la implicación anterior, a saber:

$$«f(s) = f(s') \Rightarrow s = s'»$$

La aplicación del ejemplo A.1) es inyectiva, ya que $f(x) = f(y) \Leftrightarrow 3x = 3y \Leftrightarrow x = y$; la aplicación del ejemplo A.2) es también inyectiva, ya que $f(n) = f(m) \Leftrightarrow n^2 = m^2 \Leftrightarrow n = m$ (puesto que $n, m \in \mathbb{N}$); sin embargo, la aplicación del ejemplo A.3) *no* es inyectiva, ya que $f(1) = f(2)$ y $1 \neq 2$.

Una aplicación que es a la vez suprayectiva e inyectiva recibe el nombre de *biyectiva*. La aplicación del ejemplo A.1) es biyectiva, y no lo son ninguna de las aplicaciones de los ejemplos A.2) y A.3).

* * *



La operación básica con las funciones es la *composición*. Dadas dos aplicaciones $f$: $S \mapsto T$ y $g$: $T \mapsto U$, se denomina *composición de f y g*, y se denota por $g \circ f$ al resultado de aplicar $g$ a la imagen mediante $f$ de cualquier elemento de $S$, es decir,

$$g \circ f(s) = g(f(s))$$

EJEMPLO B. Si $f$: $\mathbb{R} \mapsto \mathbb{R}$ está dada por $f(x) = x^2 - 1$ y $g$: $\mathbb{R} \mapsto \mathbb{R}$ está dada por $g(x) = x + 5$, se tiene que

$$g \circ f(x) = g(f(x)) = f(x) + 5 = x^2 - 1 + 5 = x^2 + 4$$

Para que la composición de $f$ y $g$ pueda definirse es necesario que el conjunto final de $f$, es decir, $T$, coincida con el conjunto inicial de $g$; así pues, $f \circ g$ no está definida a menos que $U = S$. En el caso en que $g \circ f$ puedan definirse cabe preguntarse si $g \circ f$ coincide con $f \circ g$, es decir, si la composición de aplicaciones es *conmutativa*. La respuesta es, en general, negativa, ya que en el ejemplo B tenemos

$$f \circ g(x) = f(g(x)) = [g(x)]^2 - 1 = (x + 5)^2 - 1 = x^2 + 10x + 24$$

que no coincide con $g \circ f$.

Sin embargo, la composición de aplicaciones satisface la propiedad *asociativa*, es decir:

$$(h \circ g) \circ f = h \circ (g \circ f)$$

si $f$, $g$ y $h$ son tres aplicaciones para las cuales tienen sentido las composiciones anteriores. Este resultado se deduce inmediatamente de la definición de composición, ya que

$$\begin{aligned}(h \circ g) \circ f(s) &= (h \circ g)(f(s)) = h(g(f(s))) = \\ &= h(g \circ f(s)) = h \circ (g \circ f)(s).\end{aligned}$$

El comportamiento de la composición de aplicaciones con respecto a los tipos de aplicaciones citados anteriormente queda expuesto en las siguientes propiedades, que se dejan como ejercicio para el lector. Si $f$: $S \mapsto T$ y $g$: $T \mapsto U$ se tiene:

(C.1) $f$ inyectiva y $g$ inyectiva $\Rightarrow$ $g \circ f$ inyectiva.
(C.2) $f$ y $g$ suprayectivas $\Rightarrow$ $g \circ f$ suprayectiva.
(C.3) $f$ y $g$ biyectivas $\Rightarrow$ $g \circ f$ biyectiva.

* * *

Una forma de definir una aplicación es utilizando matrices. Sea, por ejemplo, la matriz

$$A = \begin{pmatrix} 1 & -1 \\ 2 & 3 \end{pmatrix}$$

y sea $\mathbb{R}^2$ el conjunto de todos los vectores $\vec{x}=(x_1, x_2)$ con dos componentes; podemos definir $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$ mediante

$$f(\vec{x})=f(x_1, x_2)=(x_1-x_2, 2x_1+3x_2)$$

Así, por ejemplo,

$$f(1, 5)=(1-5, 2\cdot 1+3\cdot 5)=(-4, 17)$$

En general, definimos $\mathbb{R}^n$ como el conjunto de los vectores $\vec{x}=(x_1, x_2, ..., x_n)$ de $n$ componentes. Dada una matriz

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

de $m$ filas y $n$ columnas, se denomina *aplicación lineal asociada con* $A$ a la aplicación $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ dada por

$$f(\vec{x})=f(x_1, x_2, ..., x_n)=$$
$$=(a_{11}+a_{12}x_2+\cdots+a_{1n}x_n, a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n, ..., a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n)$$

EJEMPLO C. La matriz

$$A=\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

tiene como aplicación lineal asociada a $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$ dada por

$$f(x_1, x_2)=(0x_1+1x_2, 1x_1+0x_2)=(x_2, x_1)$$

La aplicación $f$ intercambia las componentes de todo vector $\vec{x}=(x_1, x_2)$ de $\mathbb{R}^2$. Geométricamente, $f$ refleja cada vector $\vec{x}=(x_1, x_2)$ de $\mathbb{R}^2$ en la recta $x_1=x_2$ (Fig. 1.3). (Esto es fácil de probar: ¡inténtalo!)

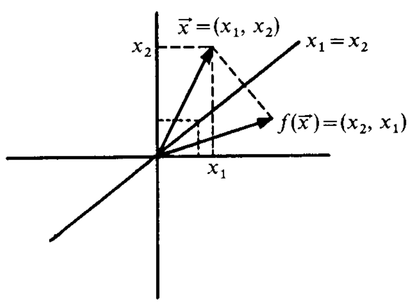



Figura 1.3



EJEMPLO D. La matriz

$$A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

tiene como aplicación lineal asociada a $f$: $\mathbb{R}^3 \mapsto \mathbb{R}^3$ dada por

$$f(\vec{x})=f(x_1, x_2, x_3)=$$
$$=(1\cdot x_1+0\cdot x_2+0\cdot x_3, 0\cdot x_1+1\cdot x_2+0\cdot x_3, 0\cdot x_1+0\cdot x_2+0\cdot x_3)=(x_1, x_2, 0)$$

Geométricamente, $f$ proyecta todo vector de $\mathbb{R}^3$ en un vector del plano $x_1x_2$ cuyos dos primeras componentes son las dos primeras componentes de $\vec{x}$ y su tercera componente es nula (Fig. 1.4).

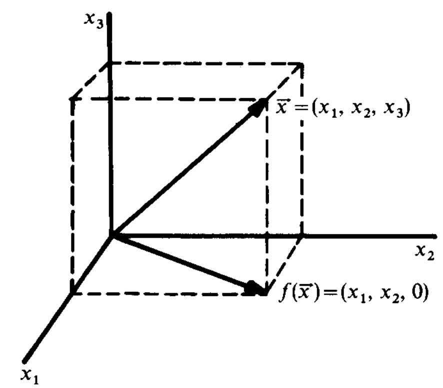



Figura 1.4

Las aplicaciones lineales tienen un buen comportamiento con respecto a la suma de vectores y a la multiplicación de éstos por números reales:

TEOREMA 1

Sea $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ una aplicación lineal; para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$ y para todo número real $r$ se tiene:

*i)* $f(\vec{x}+\vec{y})=f(\vec{x})+f(\vec{y})$
*ii)* $f(r\vec{x})=rf(\vec{x})$.

*Demostración.* Haremos la demostración para una aplicación lineal $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$, puesto que las ideas principales de la demostración en el caso general están incluidas en este caso particular. Sea, por tanto:

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

la matriz de $f$ y $\overrightarrow{x}=(x_1, x_2)$, $\overrightarrow{y}=(y_1, y_2)$. Se tiene que

$$\begin{aligned} f(\overrightarrow{x}+\overrightarrow{y}) &= f(x_1+y_1,\ x_2+y_2)=(a(x_1+y_1)+b(x_2+y_2),\ c(x_1+y_1)+ \\ &\quad +d(x_2+y_2))=(ax_1+ay_1+bx_2+by_2,\ cx_1+cy_1+dx_2+dy_2)= \\ &= (ax_1+bx_2,\ cx_1+dx_2)+(ay_1+by_2,\ cy_1+dy_2)= \\ &= f(x_1, x_2)+f(y_1, y_2)=f(\overrightarrow{x})+f(\overrightarrow{y}) \end{aligned}$$

Esto demuestra *i*). Para demostrar *ii*) sea $\overrightarrow{x}=(x_1, x_2)$ y $r$ un número real. Se tiene que

$$\begin{aligned} f(r\overrightarrow{x}) &= f(rx_1, rx_2)=(a(rx_1)+b(rx_2),\ c(rx_1)+d(rx_2))= \\ &= r(ax_1+bx_2,\ cx_1+dx_2)=rf(\overrightarrow{x}) \end{aligned}$$

Esto termina la demostración de *ii*) y, por tanto, la demostración del teorema. ■

*Nota.* Combinando *i*) y *ii*) del teorema 1 se tiene que

$$f(r\overrightarrow{x}+s\overrightarrow{y})=rf(\overrightarrow{x})+sf(\overrightarrow{y})$$

para todo $\overrightarrow{x}, \overrightarrow{y}\in\mathbb{R}^n$ y para cualesquiera números reales $r$ y $s$.

El teorema 1 se muestra gráficamente en las figuras 1.5 y 1.6.

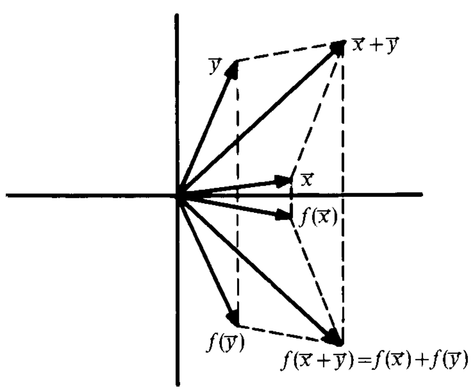



Figura 1.5

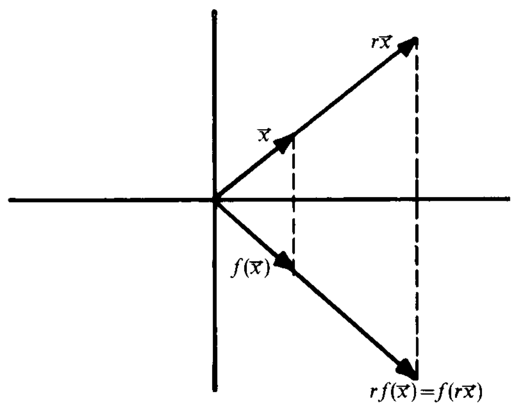



Figura 1.6

Este teorema permite demostrar que la imagen de una recta en $\mathbb{R}^n$ mediante una aplicación lineal es otra recta o un punto; en efecto, si $\overrightarrow{x}+r\overrightarrow{v}$ es la ecuación de la recta que pasa por el extremo de $\overrightarrow{x}$ en la dirección de $\overrightarrow{v}$ se tiene que

$$f(\overrightarrow{x}+r\overrightarrow{v})=f(\overrightarrow{x})+rf(\overrightarrow{v})$$

debido al teorema 1 (Fig. 1.7). Si $f(\overrightarrow{v})=\overrightarrow{0}$ se tiene el punto $f(\overrightarrow{x})$, y si $f(\overrightarrow{v})\neq\overrightarrow{0}$ se obtiene una recta que pasa por el extremo de $f(\overrightarrow{x})$ en la dirección de $f(\overrightarrow{v})$.



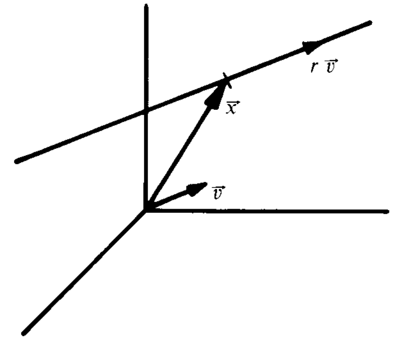



Figura 1.7

EJEMPLO E. Queremos hallar la imagen del cuadrado de vértices (0, 0), (0, 1), (1, 1) y (1, 0) mediante la aplicación lineal $f$ dada por la matriz

$$A=\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

La recta (1) (véase figuras 1.8 y 1.9) tiene por ecuación $\overrightarrow{0}+r\overrightarrow{e}_1$ y, por tanto, se transforma en (1′), que tiene por ecuación

$$f(\overrightarrow{0}+r\overrightarrow{e}_1)=f(\overrightarrow{0})+rf(\overrightarrow{e}_1)=\overrightarrow{0}+r(2\cdot 1+0,\ 1\cdot 1+2\cdot 0)=\overrightarrow{b}+r(2, 1)$$

Análogamente:

$$\begin{cases} (2)\colon \overrightarrow{0}+r\overrightarrow{e}_2 \quad \text{se transforma en} \quad (2')\colon \overrightarrow{0}+r(1, 2) \\ (3)\colon (1, 0)+r\overrightarrow{e}_2 \quad \text{se transforma en} \quad (3')\colon (2, 1)+r(1, 2) \\ (4)\colon (0, 1)+r\overrightarrow{e}_1 \quad \text{se transforma en} \quad (4')\colon (1, 2)+r(2, 1) \end{cases}$$

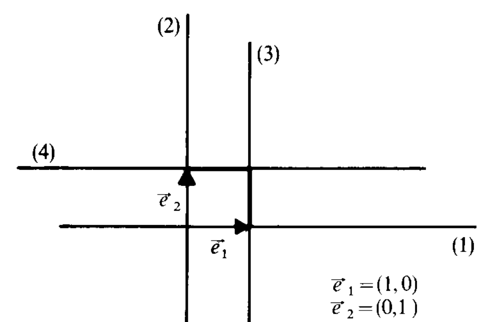



Figura 1.9

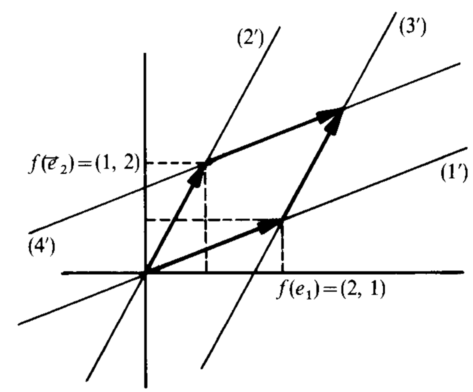



Figura 1.8

Con estos resultados se tiene que la imagen del cuadrado mediante $f$ es el paralelogramo limitado por las rectas (1'), (2'), (3') y (4'), que tiene como vértices (0, 0), (2, 1), (3, 3) y (1, 2).

* * *

Las propiedades *i)* y *ii)* dadas en el teorema 1 caracterizan a las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$. Se tiene el siguiente resultado:

TEOREMA 2

Sea $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ una aplicación que satisface:

*i)* $f(\vec{x}+\vec{y})=f(\vec{x})+f(\vec{y})$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$, y
*ii)* $f(r\vec{x})=rf(\vec{x})$ para todo $\vec{x} \in \mathbb{R}^n$ y todo número real $r$.

Entonces, $f$ es una aplicación *lineal* con matriz $A$ cuyas columnas vienen dadas por los vectores $f(\vec{e}_1), f(\vec{e}_2), \ldots, f(\vec{e}_n)$, donde $e_j$ es el vector de $\mathbb{R}^n$ con todas sus componentes nulas excepto la que ocupa el lugar $j$, que es 1.

*Demostración.* Al igual que en la demostración del teorema 1 vamos a suponer que $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$; supongamos que $f(\vec{e}_1)=f(1, 0)=(a, b)$ y $f(\vec{e}_2)=f(0, 1)=(c, d)$. Utilizando *i)* y *ii)* se tiene que

$$\begin{aligned} f(\vec{x}) &= f(x_1, x_2)=f(x_1(1, 0)+x_2(0, 1))=f(x_1\vec{e}_1+x_2\vec{e}_2)= \\ &= x_1 f(\vec{e}_1)+x_2 f(e_2)=x_1(a, b)+x_2(c, d)= \\ &= (ax_1+cx_2, bx_1+dx_2) \end{aligned}$$

Por tanto, $f$ es una aplicación lineal que tiene como matriz

$$A=\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

Observar que la primera columna de $A$ es $f(\vec{e}_1)$ y la segunda es $f(\vec{e}_2)$. ∎

EJEMPLO F. Tratemos de encontrar la matriz de un giro de 90°, $g$, en $\mathbb{R}^2$ (el giro se considera, salvo indicación contraria, que se realiza en sentido positivo, es decir, contrario al de las agujas del reloj) (Fig. 1.10). La aplicación $g$ satisface *i)* y *ii)* del

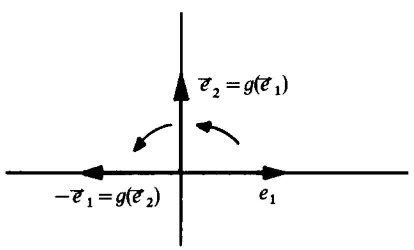



Figura 1.10



teorema 2 (¡tratar de demostrarlo geométricamente!). Por el teorema 2, $g$ es una aplicación lineal y su matriz tiene como columnas

$$f(\vec{e}_1)=\vec{e}_2=(0, 1)$$
$$f(\vec{e}_2)=-e_1=(-1, 0)$$

Por tanto,

$$A=\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

* * *

Las aplicaciones lineales pueden sumarse y multiplicarse por números reales: dadas $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ y $g$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ dos aplicaciones lineales y $c$ un número real, definimos la *suma* de $f$ y $g$ como una aplicación $f+g$: $\mathbb{R}^n \to \mathbb{R}^m$ tal que

$$(f+g)(\vec{x})=f(\vec{x})+g(\vec{x}), \quad \vec{x} \in \mathbb{R}^n$$

y la *multiplicación de f por c* como una aplicación $cf$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ tal que

$$(cf)(\vec{x})=c(f(\vec{x})), \quad \vec{x} \in \mathbb{R}^n$$

EJEMPLO G. Dadas dos aplicaciones lineales $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^3$ y $g$: $\mathbb{R}^2 \mapsto \mathbb{R}^3$ mediante

$$f(x_1, x_2)=(x_1, x_1+x_2, 2x_2)$$

y

$$g(x_1, x_2)=(0, x_1, x_2)$$

se tiene que

$$\begin{aligned} (f-2g)(\vec{x}) &= (f+(-2)g)(x_1, x_2)=f(x_1, x_2)+(-2)g(x_1, x_2)= \\ &= (x_1, x_1+x_2, 2x_2)+(0, -2x_1, -2x_2)=(x_1, x_2-x_1, 0) \end{aligned}$$

* * *

La suma de dos aplicaciones lineales $f$ y $g$ es una aplicación lineal, ya que

$$\begin{aligned} (f+g)(\vec{x}+\vec{y}) &= f(\vec{x}+\vec{y})+g(\vec{x}+\vec{y})=f(\vec{x})+f(\vec{y})+g(\vec{x})+g(\vec{y}) \\ &= f(\vec{x})+g(\vec{x})+f(\vec{y})+g(\vec{y})=(f+g)(\vec{x})+(f+g)(\vec{y}) \end{aligned}$$

y

$$(f+g)(r\vec{x})=f(r\vec{x})+g(r\vec{x})=rf(\vec{x})+rg(\vec{x})=r(f+g)(\vec{x});$$

por el teorema 2 esto basta para probar que $f+g$ es lineal.

De manera similar puede comprobarse que $cf$ es una aplicación lineal si $f$ lo es.

Si $A$ es la matriz de la aplicación lineal $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ y $B$ es la matriz de la aplicación lineal $g$: $\mathbb{R}^n \mapsto \mathbb{R}^m$, $f+g$ tendrá una matriz cuya $j$-ésima columna está dada por $(f+g)(\vec{e}_j)$, debido al teorema 2. Puesto que $(f+g)(e_j)=f(e_j)+g(e_j)$ la $j$-ésima columna de la matriz de $f+g$ es la suma de las $j$-ésimas columnas de las matrices de $f$ y $g$. A la matriz de $f+g$ se le denomina *matriz suma de A y B*, y, por tanto, si

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix}$$

se tiene que

$$A+B=\begin{pmatrix} a_{11}+b_{11}, & a_{12}+b_{12}, & \ldots, & a_{1n}+b_{1n} \\ a_{21}+b_{21}, & a_{22}+b_{22}, & \ldots, & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1}, & a_{m2}+b_{m2}, & \ldots, & a_{mn}+b_{mn} \end{pmatrix}$$

Observar que para poder sumar matrices el número de filas de $A$ y $B$ debe coincidir, así como el número de columnas de ambas matrices. Una matriz con $m$ filas y $n$ columnas se dice que es una matriz de *orden* $m\times n$, la cual corresponde a una aplicación de $\mathbb{R}^n$ en $\mathbb{R}^m$. Por tanto, la suma de matrices sólo es posible si ambas son del mismo orden.

Sea $A$ una matriz de orden $m\times n$ asociada con la aplicación lineal $f$: $\mathbb{R}^n \mapsto \mathbb{R}^m$ y $c$ *un número real. Puesto que* $cf$ es una aplicación lineal, tiene asociada una matriz, que se simboliza mediante $cA$ *—multiplicación de A por el número real c—* cuya $j$-ésima columna es $(cf)(e_j)=c(f(e_j))$, es decir, $c$ veces la $j$-ésima columna de $A$:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \Rightarrow cA=\begin{pmatrix} ca_{11} & ca_{22} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{pmatrix}$$

En el ejemplo G las matrices de $f$ y $g$ son

$$A=\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$



Por tanto:

$$A-2B=\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}+(-2)\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 2 \end{pmatrix}+\begin{pmatrix} 0 & 0 \\ -2 & 0 \\ 0 & -2 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ -1 & 1 \\ 0 & 0 \end{pmatrix}$$

que corresponde a la aplicación lineal $f-2g$ (¡comprobarlo!)

* * *

Estas operaciones que acabamos de definir con aplicaciones lineales y matrices tienen propiedades similares a las propiedades $(S_1)$-$(S_4)$ y $(M_1)$-$(M_4)$ de la suma de vectores y de la multiplicación de vectores por un número real (ver sección 1.2). Estas propiedades se enuncian en el siguiente cuadro y se dejan como ejercicio para el lector.

| Suma de aplicaciones lineales | Suma de matrices |
|---|---|
| $(S_1)$ $f+g=g+f$<br>$(S_2)$ $(f+g)+h=f+(g+h)$<br>$(S_3)$ $0+f=f+0=f$, donde 0 es la aplicación nula<br>$(S_4)$ $f+(-f)=(-f)+f=0$, donde $-f=(-1)f$ | $(S_1)$ $A+B=B+A$ (conmutativa)<br>$(S_2)$ $(A+B)+C=A+(B+C)$ (asociativa)<br>$(S_3)$ $0+A=A+0=A$, donde 0 es la matriz nula<br>$(S_4)$ $A+(-A)=(-A)+A=0$ |

| Multiplicación de aplicaciones lineales por un número real | Multiplicación de matrices por un número real |
|---|---|
| $(M_1)$ $c(f+g)=cf+cg$<br>$(M_2)$ $(c+d)f=cf+df$<br>$(M_3)$ $c(df)=(cd)f$<br>$(M_4)$ $1f=f$ | $(M_1)$ $c(A+B)=cA+cB$<br>$(M_2)$ $(c+d)A=cA+dA$<br>$(M_3)$ $c(dA)=(cd)A$<br>$(M_4)$ $1A=A$ |

* * *

El último resultado de esta sección es obtener una operación con matrices que corresponda a la composición de aplicaciones lineales. Comenzamos con el caso particular de aplicaciones de $\mathbb{R}^2$ en $\mathbb{R}^2$ o matrices (cuadradas) de orden $2\times 2$.

Sean

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

las matrices de $f$, $g$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$. Tenemos que

$$g\circ f(\overline{x})=g\circ f(x_1, x_2)=g(f(x_1, x_2))=g(a_{11}x_1+a_{12}x_2,\ a_{21}x_1+a_{22}x_2)=$$
$$=(b_{11}(a_{11}x_1+a_{12}x_2)+b_{12}(a_{21}x_1+a_{22}x_2),\quad b_{21}(a_{11}x_1+a_{12}x_2)+b_{22}(a_{21}x_1+a_{22}x_2))=$$
$$=((b_{11}a_{11}+a_{21}b_{12})x_1+(b_{11}a_{12}+b_{12}\cdot a_{22})x_2, (b_{21}a_{11}+b_{22}a_{21})x_1+(b_{21}\cdot a_{12}+b_{22}a_{22})x_2)$$

Por tanto, $g \circ f$ es una aplicación lineal que tiene como matriz

$$\begin{pmatrix} b_{11}a_{11}+a_{21}b_{12}, & b_{11}a_{12}+b_{12}a_{22} \\ b_{21}a_{11}+b_{22}a_{21}, & b_{21}a_{12}+b_{22}a_{22} \end{pmatrix}$$

que recibe el nombre de *producto de B por A*.

EJEMPLO H. Dadas $A=\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ y $B=\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}$, tenemos que

$$BA=\begin{pmatrix} 3 & 0 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 3\cdot 1+0\cdot 0 & 3\cdot 2+0\cdot 1 \\ 1\cdot 1+2\cdot 0 & 1\cdot 2+2\cdot 1 \end{pmatrix}=\begin{pmatrix} 3 & 6 \\ 1 & 4 \end{pmatrix}$$

y

$$A^2=A\cdot A=\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}=\begin{pmatrix} 1\cdot 1+2\cdot 0 & 1\cdot 2+2\cdot 1 \\ 0\cdot 1+1\cdot 0 & 0\cdot 2+1\cdot 1 \end{pmatrix}=\begin{pmatrix} 1 & 4 \\ 0 & 1 \end{pmatrix}$$

* * *

Recordemos que para que $g \circ f$ tenga sentido es necesario que el conjunto final de $f$ coincida con el original de $g$. Por tanto,

$$f\colon \mathbb{R}^n \mapsto \mathbb{R}^m \quad , \quad g\colon \mathbb{R}^m \mapsto \mathbb{R}^p$$

Como la matriz $A$ de $f$ es de orden $m \times n$ y la matriz $B$ de $g$ es de orden $p \times m$, para poder calcular $BA$ es necesario que el número de columnas de $B$ coincida con el número de filas de $A$. Gráficamente:

$$\underset{p\times \boxed{m}}{B} \; , \; \underset{\boxed{m}\times n}{A} \; \to \underset{p\times n}{BA}$$

Dada una matriz $A$ el elemento que ocupa el lugar $(i, j)$, es decir, la intersección de la $i$-ésima fila con la $j$-ésima columna, se denota por $a_{ij}$. Esto nos permite escribir una matriz $A$ abreviadamente como $(a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,m}}$ si es de orden $m \times n$.

DEFINICIÓN 1 (*Multiplicación de matrices*)

Dadas las matrices

$$B=(b_{ij})_{\substack{i=1,\dots,p\\ j=1,\dots,m}} \quad \text{y} \quad A=(a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$$

de órdenes $p\times m$ y $m \times n$, respectivamente, definimos $BA$ como la matriz de orden $p \times n$ cuyo elemento que ocupa el lugar $(i, j)$ está dado por

$$\sum_{k=1}^{n} b_{ik}a_{kj}=b_{i1}a_{1j}+b_{i2}a_{2j}+\cdots+b_{in}a_{nj}$$



/

EJEMPLO I. Dadas las matrices

$$B=\begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 0 \end{pmatrix} \quad \text{y} \quad A=\begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix}$$

podemos calcular $BA$, ya que $2\times\boxed{3}\,\boxed{3}\times 3$, pero no $AB$, ya que $3\times\boxed{3}\,\boxed{2}\times 3$. El elemento que aparece en el lugar (2, 1) de $BA$ es

$$b_{21}a_{11}+b_{22}a_{21}+b_{23}a_{31}=3\cdot 0+1\cdot 1+0\cdot(-2)=1$$

Además,

$$BA=\begin{pmatrix} 1 & 2 & 4 \\ 3 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 & 1 \\ 1 & 3 & 0 \\ -2 & 1 & 1 \end{pmatrix}=$$

$$=\begin{pmatrix} 1\cdot 0+2\cdot 1+4(-2) & 1\cdot 2+2\cdot 3+4\cdot 1 & 1\cdot 1+2\cdot 0+4\cdot 1 \\ 3\cdot 0+1\cdot 1+0(-2) & 3\cdot 2+1\cdot 3+0\cdot 1 & 3\cdot 1+1\cdot 0+0\cdot 1 \end{pmatrix}=\begin{pmatrix} -6 & 12 & 5 \\ 1 & 9 & 3 \end{pmatrix}$$

* * *

Como el lector puede suponer, el producto de matrices corresponde a la composición de las aplicaciones lineales que ellas determinan, en el orden adecuado.

PROPOSICIÓN 3

Si $f\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ tiene $A=(a_{ij})_{\substack{i=1,\dots,m\\ j=1,\dots,n}}$ como matriz y $g\colon \mathbb{R}^n \mapsto \mathbb{R}^p$ tiene $B=(b_{ij})_{\substack{i=1,\dots,p\\ j=1,\dots,m}}$ como matriz, $g\circ f\colon \mathbb{R}^n \mapsto \mathbb{R}^p$ es una aplicación lineal que tiene como matriz $BA$.

*Demostración.* La $j$-ésima columna de la matriz de $g \circ f$ es

$$\begin{aligned} g\circ f(e_j)&=g(f(e_j))=g(a_{1j}, a_{2j}, \dots, a_{mj})= \\ &=a_{1j}g(\vec{e}_1)+a_{2j}g(\vec{e}_2)+\cdots+a_{mj}g(\vec{e}_m) \end{aligned}$$

en donde se ha utilizado la linealidad de $g$. Por tanto,

$$g\circ f(e_j)=a_{1j}(b_{11}, b_{21}, \dots, b_{n1})+a_{2j}(b_{12}, b_{22}, \dots, b_{n2})+\cdots+a_{mj}(b_{1m}, b_{2m}, \dots, b_{nm})=$$
$$=(a_{1j}b_{11}+a_{2j}b_{12}+\cdots+a_{mj}b_{1m}, a_{1j}b_{21}+a_{2j}b_{22}+\cdots+a_{mj}b_{2m}, \dots, a_{1j}b_{n1}+a_{2j}b_{n2}+\cdots+a_{nj}b_{nj})=$$

$$\left(\sum_{k=1}^{n} b_{1k}a_{kj}, \sum_{k=1}^{n} b_{2k}a_{kj}, \dots, \sum_{k=1}^{n} b_{nk}a_{kj}\right)$$

De aquí se deduce que el elemento que ocupa el lugar $(i, j)$ en $BA$ es $\sum_{k=1}^{n} b_{ik}a_{kj}$, que era lo que queríamos demostrar. La demostración de que $g \circ f$ es una aplicación lineal se deja para el lector. ■

Para finalizar esta sección damos algunas propiedades de la composición de aplicaciones lineales y de la multiplicación de matrices. Ya sabemos que la propiedad *asociativa* se cumple para la composición de aplicaciones y, por tanto, se cumple también para la multiplicación de matrices.

La propiedad conmutativa puede que no tenga sentido, como en el ejemplo I, en el que no puede calcularse $AB$, pero sí $BA$. Incluso si $AB$ y $BA$ pueden calcularse la propiedad conmutativa no es cierta:

$$\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 1 \\ 4 & 2 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}=\begin{pmatrix} 1 & 2 \\ 2 & 2 \end{pmatrix}$$

La asociativa y otras propiedades de estas operaciones se resumen a continuación.

| | |
|---|---|
| $(C_1)$ $(h \circ g) \circ f = h \circ (g \circ f)$ | $(C_1)$ $(CB)A = C(BA)$ |
| $(C_2)$ $h \circ (f+g) = h \circ f + h \circ g$ | $(C_2)$ $C(A+B) = CA + CB$ |
| $(C_3)$ $(h+g) \circ f = h \circ f + g \circ f$ | $(C_3)$ $(C+B)A = CA + BA$ |
| $(C_4)$ $(cg) \circ f = g \circ (cf) = c(g \circ f)$ | $(C_4)$ $(cB)A = B(cA) = c(BA)$ |

La multiplicación de matrices nos permite escribir un sistema de ecuaciones lineales de una forma muy sencilla. Dado el sistema

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \qquad \text{(I)}$$

de $m$ ecuaciones lineales con $n$ incógnitas, si escribimos

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

se tiene que

$$A\vec{x} = \vec{b}$$

es una forma abreviada de escribir (I).



EJEMPLO J. El sistema

$$\left.\begin{array}{r} 2x_1 + x_2 - x_3 = 1 \\ x_1 + 2x_3 = -2 \end{array}\right\}$$

se escribe con notación matricial de la forma

$$\begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

## EJERCICIOS 1.3

**1.** Estudiar si las siguientes aplicaciones son inyectivas, suprayectivas o biyectivas:

*a)* $f\colon \mathbb{R} \mapsto \mathbb{R}$, $f(x) = x^2$

*b)* $g\colon \mathbb{N}^+ \mapsto \mathbb{R}$, $g(n) = n^3 - n$

*c)* $h\colon \mathbb{Z} \mapsto 2\mathbb{Z}$, $h(z) = 2z$, donde $\mathbb{Z}$ denota el conjunto de los números enteros 0, $\pm 1$, $\pm 2$, $\pm 3$, ... y $2\mathbb{Z} = \{0, \pm 2, \pm 4, \pm 6, ...\}$

**2.** Encontrar todas las biyecciones del conjunto $\{1, 2, 3\}$ en sí mismo.

**3.** Dadas $f\colon S \mapsto T$ y $g\colon T \mapsto U$, demostrar que:

*a)* $f$, $g$ inyectivas $\Rightarrow$ $g \circ f$ inyectiva

*b)* $f$, $g$ suprayectivas $\Rightarrow$ $g \circ f$ suprayectiva

*c)* $f$, $g$ biyectivas $\Rightarrow$ $g \circ f$ biyectiva

**4.** Dadas $f(x) = x^2 + 7$, $g(x) = 3x - 5$ y $h(x) = \operatorname{sen} x$, funciones de $\mathbb{R}$ en $\mathbb{R}$, calcular:

*a)* $h \circ g \circ f$ *b)* $g \circ h \circ g$ *c)* $g \circ h \circ f$

**5.** Escribir las matrices de las siguientes aplicaciones:

*a)* $f(x_1, x_2, x_3) = (x_1 + x_2, 2x_1 - x_3)$

*b)* $f(x_1, x_2) = (x_1 - x_2, x_1 + 2x_2, 2x_1 - 3x_2)$

*c)* $f(x_1, x_2) = (x_1, x_2, x_2, x_1)$

*d)* $f(x_1, x_2, x_3) = 3x_1 - x_2 + x_3$

**6.** Dada $f\colon \mathbb{R}^2 \mapsto \mathbb{R}^2$ mediante $f(x_1, x_2) = (x_1 + x_2, 2x_1 - x_2)$, hallar la imagen mediante $f$ de las siguientes regiones:

*a)* $\{(x_1, x_2)/1 \leq x_1 \leq 2, 0 \leq x_2 \leq 1\}$

*b)* $\{(x_1, x_2)/-1 \leq x_1 \leq 1, -1 \leq x_2 \leq 1\}$

**7.** Demostrar que $f\colon \mathbb{R}^2 \mapsto \mathbb{R}^2$ dada por $f(x_1, x_2) = (x_1, x_2^2)$ *no* es lineal. Dibujar la imagen de la recta $\vec{x} = r(1, 1)$ mediante $f$.

**8.** Hallar la matriz de un giro de ángulo $\phi$ en sentido positivo en $\mathbb{R}^2$.

**9.** Hallar la matriz (en $\mathbb{R}^3$) de la simetría con respecto al plano $x=y$.

**10.** Dadas las matrices

$$A=\begin{pmatrix}2 & 1 & 8\\ 3 & 2 & 7\end{pmatrix},\quad B=\begin{pmatrix}1 & 5 & 0\\ 2 & 3 & 1\\ 4 & 0 & 6\end{pmatrix},\quad C=\begin{pmatrix}4 & 1 & 3\\ 0 & 2 & 4\\ 7 & 6 & 5\end{pmatrix}$$

$$D=\begin{pmatrix}1\\ 1\\ 0\end{pmatrix},\quad E=(3,\ 2),\quad F=\begin{pmatrix}5 & 4\\ 1 & 2\end{pmatrix},\quad G=\begin{pmatrix}3 & 1\\ 1 & 2\end{pmatrix}$$

calcular:

*a)* $A((B+C)D)$ *b)* $(B+C)^2$ *c)* $((F^2-G^2)A)D$ *d)* $(B^2+C^2)3D$
*e)* $AB^2$ *f)* $2E(F^2-G^2)$ *g)* $G^5$

**11.** Dadas las aplicaciones lineales $f(x_1, x_2)=(x_1+x_2, x_1-x_2)$, $g(x_1, x_2)=(3x_1-x_2, 2x_1)$ y $h(x_1, x_2)=(x_1, -x_2, x_1-3x_2)$, calcular:

*a)* $h\circ(g+f)$ *b)* $h\circ g\circ g$ *c)* $h\circ f\circ g$

**12.** Demostrar las propiedades ($S_1$), ($S_2$), ($S_3$) y ($S_4$) y ($M_1$), ($M_2$), ($M_3$) y ($M_4$) de la suma de aplicaciones y matrices y de la multiplicación de éstas por números reales.

**13.** Demostrar las propiedades ($C_2$), ($C_3$) y ($C_4$) de la composición de aplicaciones lineales y de la multiplicación de matrices.

**14.** Escribir en forma matricial los siguientes sistemas de ecuaciones lineales:

*a)* $\left.\begin{aligned}x+y+z&=7\\ x-3y+2z&=3\end{aligned}\right\}$ *b)* $\left.\begin{aligned}x+2z&=4\\ 3y+2z+x&=1\\ z-7y&=3\end{aligned}\right\}$ *c)* $\left.\begin{aligned}ax+3z-2y&=a^2\\ 2z-ay+x&=3\\ y-z+ax&=a\end{aligned}\right\}$

**15.** Desarrollar los siguientes sistemas escritos en forma matricial:

*a)* $\begin{pmatrix}3 & 1 & 4\\ 3 & 2 & 6\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}3\\ 0\end{pmatrix}$ *b)* $\begin{pmatrix}3 & 0\\ 4 & 2\\ 7 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}1\\ 0\\ 5\end{pmatrix}$

## 1.4. INVERSA DE UNA APLICACION E INVERSA DE UNA MATRIZ

Dada una aplicación $f\colon S\mapsto T$ decimos que $g\colon T\mapsto S$ es *una inversa de f* si:

*a)* $g\circ f(s)=s$, para todo $s\in S$, y
*b)* $f\circ g(t)=t$, para todo $t\in T$.



Si definimos $I_S(s)=s$ para todo $s\in S$ e $I_T(t)=t$ para todo $t\in T$ (las cuales reciben el nombre de *aplicaciones identidad*) las condiciones anteriores *a)* y *b)* se escriben de la forma

$$g\circ f=I_S \quad \text{y} \quad f\circ g=I_T$$

EJEMPLO A. 1. Si $f(x)=2x+3$ es una aplicación de $\mathbb{R}$ en $\mathbb{R}$, $g(x)=\frac{1}{2}x-\frac{3}{2}$ es una inversa de $f$, ya que

$$g\circ f(x)=g(f(x))=\frac{1}{2}f(x)-\frac{3}{2}=\frac{1}{2}(2x+3)-\frac{3}{2}=x$$

y

$$f\circ g(x)=f(g(x))=2g(x)+3=2\left(\frac{1}{2}x-\frac{3}{2}\right)+3=x$$

2. Si $f\colon \mathbb{R}^2\mapsto\mathbb{R}^2$ es la simetría con respecto a una recta, $f$ es su propia inversa, ya que

$$f\circ f=f^2=I_{\mathbb{R}^2}$$

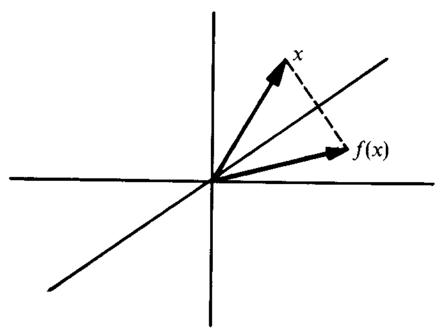



Figura 1.11

La inversa de una aplicación no siempre existe. Si $f\colon \{1, 2, 3\}\mapsto\{1, 2\}$ es la aplicación dada por $f(1)=f(3)=1$, $f(2)=2$, $f$ no posee inversa. En efecto, si $g\colon \{1, 2\}\mapsto\{1, 2, 3\}$ fuera una inversa de $f$ tendríamos que

$$1=g\circ f(1)=g\circ f(3)=3$$

lo cual es una contradicción.

Si existe la inversa de una aplicación $f$ se dice que $f$ es *invertible* y la inversa de $f$ se denota por $f^{-1}$.

La inversa de $f$, cuando existe, es única. Para probar este resultado supongamos que $g$: $T \mapsto S$ y $h$: $T \mapsto S$ son dos inversas de $f$: $S \mapsto T$. Tenemos que para todo $t \in T$,

$$h(t) = (g \circ f)(h(t)) = g(f(h(t))) = g(f \circ h)(t) = g(t)$$

y, por tanto, $h = g$.

A continuación damos una condición necesaria y suficiente para que exista la inversa de una aplicación.

**TEOREMA 1**

Sea $f$: $S \mapsto T$ una aplicación. $f$ es invertible si y sólo si $f$ es biyectiva.

*Demostración.* Supongamos, primero, que $f$ es invertible y sea $g$: $T \mapsto S$ su inversa. Si $f(s) = f(s')$ *tenemos que* $g \circ f(s) = g \circ f(s')$ y, por tanto, $s = s'$; esto prueba que $f$ es inyectiva. Para probar que $f$ es suprayectiva, sea $t \in T$ y tomar $s = g(t)$; entonces $f(s) = f(g(t)) = (f \circ g)(t) = t$. esto prueba que $f$ es biyectiva.

Supongamos ahora que $f$ es biyectiva. Dado $t \in T$ definimos $g(t) = s$ de manera que $f(s) = t$. La aplicación $g$ es la inversa de $f$, ya que

$$g \circ f(s) = g(f(s)) = g(t) = s$$

y

$$f \circ g(t) = f(g(t)) = f(s) = t$$

■

EJEMPLO B. 1. Sea $f$: $\mathbb{R}^+ \mapsto \mathbb{R}$ dada por $f(x) = \log x$ (log denota siempre logaritmo neperiano), donde $\mathbb{R}^+$ es el conjunto de los números reales positivos. Esta aplicación es inyectiva ya que $\log x = \log y \Rightarrow x = y$, y es suprayectiva ya que si $y \in \mathbb{R}$, tomamos $x = e^y$ y tenemos que $\log x = \log e^y = y$. Por tanto, $f$ es invertible.

Claramente, su inversa es la aplicación $g$: $\mathbb{R} \mapsto \mathbb{R}^+$ dada por $g(x) = e^x$.

2. Sea $f$: $\mathbb{R} \mapsto \mathbb{R}$ dada por $f(x) = 2x - 1$. Esta aplicación es inyectiva ya que $f(x) = f(y) \Leftrightarrow 2x - 1 = 2y - 1$, de donde se deduce que $x = y$; es, además, suprayectiva ya que dado $y \in \mathbb{R}$ podemos encontrar $x \in \mathbb{R}$ tal que $2x - 1 = y$; basta tomar $x = \frac{1}{2}(y+1)$. Por el teorema 1, $f$ es invertible.

Su inversa se ha encontrado en la demostración de la suprayectividad:

$$f^{-1}(x) = \frac{1}{2}(x+1)$$

* * *

Pasamos ahora a calcular la inversa de aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$; comenzamos demostrando que si una aplicación lineal $f$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es invertible, su inversa $f^{-1}$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es también una aplicación lineal.



**TEOREMA 2**

Si $f$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es una aplicación lineal invertible, $f^{-1}$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es también una aplicación lineal invertible y $(f^{-1})^{-1} = f$.

*Demostración.* Lo único que es necesario demostrar es que $f^{-1}$ es lineal. Si $\overline{x}, \overline{y} \in \mathbb{R}^n$ se tiene

$$\begin{aligned} f(f^{-1}(\overline{x}+\overline{y})) &= f \circ f^{-1}(\overline{x}+\overline{y}) = \overline{x}+\overline{y} = f \circ f^{-1}(\overline{x}) + f \circ f^{-1}(\overline{y}) = \\ &= f(f^{-1}(\overline{x})) + f(f^{-1}(\overline{y})) = f(f^{-1}(\overline{x}) + f^{-1}(y)) \end{aligned}$$

Puesto que $f$ es inyectiva por el teorema 1,

$$f^{-1}(\overline{x}+\overline{y}) = f^{-1}(\overline{x}) + f^{-1}(\overline{y}).$$

Finalmente, si $\overline{x} \in \mathbb{R}^n$, $r \in \mathbb{R}$,

$$f(f^{-1}(r\overline{x})) = f \circ f^{-1}(r\overline{x}) = r\overline{x} = rf \circ f^{-1}(\overline{x}) = rf(f^{-1}(\overline{x})) = f(rf^{-1}(\overline{x}))$$

y de nuevo puesto que $f$ es inyectiva debido al teorema 1 se tiene que

$$f^{-1}(r\overline{x}) = rf^{-1}(\overline{x})$$

Estas dos propiedades son suficientes para asegurar que $f^{-1}$ es lineal debido al teorema 2 de la sección 1.3.

* * *

Si $f$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es una aplicación lineal, su matriz asociada $A$ es de orden $n \times n$ y, por tanto, el número de filas coincide con el número de columnas. Estas matrices se denominan *cuadradas* y se dicen de *orden* $n$ en lugar de orden $n \times n$.

Si $f$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es invertible, su inversa $f^{-1}$: $\mathbb{R}^n \mapsto \mathbb{R}^n$ es también una aplicación lineal y su matriz asociada $B$ es también cuadrada y del mismo orden que $A$. La matriz $B$ recibe el nombre de *inversa de* $A$ y se denota mediante $B = A^{-1}$.

Puesto que la matriz de $f \circ f^{-1}$ es $AB$ y la matriz de $f^{-1} \circ f$ es $BA$ (ver los resultados obtenidos en la sección 1.3) se tiene que $B$ es la inversa de $A$ si y sólo si

$$AB = I_n \quad \text{y} \quad BA = I_n$$

donde $I_n$ es la matriz identidad de orden $n$, es decir, la matriz con *unos* en la diagonal principal y *ceros* en el resto:

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

EJEMPLO C. Intentamos calcular la inversa de la matriz

$$A=\begin{pmatrix}2 & 1\\ 4 & 3\end{pmatrix}$$

Sea $B=\begin{pmatrix}x_1 & x_2\\ x_3 & x_4\end{pmatrix}$ su inversa. Puesto que $AB=I_2$, se tiene

$$\begin{pmatrix}2 & 1\\ 4 & 3\end{pmatrix}\begin{pmatrix}x_1 & x_2\\ x_3 & x_4\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \Leftrightarrow \begin{pmatrix}2x_1+x_3 & 2x_2+x_4\\ 4x_1+3x_3 & 4x_2+3x_4\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

Igualando los elementos de las matrices se tienen los siguientes sistemas de dos ecuaciones lineales con dos incógnitas:

$$\left.\begin{aligned}2x_1+x_3&=1\\ 4x_1+3x_3&=0\end{aligned}\right\}\ (\mathrm{I}) \quad ; \quad \left.\begin{aligned}2x_2+x_4&=0\\ 4x_2+3x_4&=1\end{aligned}\right\}\ (\mathrm{II})$$

Para resolver (I) escribimos:

$$\left(\begin{array}{cc|c}2 & 1 & 1\\ 4 & 3 & 0\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|c}2 & 1 & 1\\ 0 & 1 & -2\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|c}2 & 0 & 3\\ 0 & 1 & -2\end{array}\right) \xrightarrow{i)} \left(\begin{array}{cc|c}1 & 0 & 3/2\\ 0 & 1 & -2\end{array}\right)$$

de donde deducimos $x_1=3/2$, $x_3=-2$.

Para resolver (II) escribimos:

$$\left(\begin{array}{cc|c}2 & 1 & 1\\ 4 & 3 & 0\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|c}2 & 1 & 0\\ 0 & 1 & 1\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|c}2 & 0 & -1\\ 0 & 1 & 1\end{array}\right) \xrightarrow{i)} \left(\begin{array}{cc|c}1 & 0 & -1/2\\ 0 & 1 & 1\end{array}\right)$$

de donde deducimos $x_2=-1/2$, $x_4=1$.

Por tanto:

$$B=\begin{pmatrix}3/2 & -1/2\\ -2 & 1\end{pmatrix}$$

Podemos comprobar que $BA=I_2$:

$$\begin{pmatrix}3/2 & -1/2\\ -2 & 1\end{pmatrix}\begin{pmatrix}2 & 1\\ 4 & 3\end{pmatrix}=\begin{pmatrix}3-2 & 3/2-3/2\\ -4+4 & -2+3\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$

* * *

En el ejemplo anterior se observa que las operaciones elementales que se han realizado para resolver los sistemas (I) y (II) son las mismas. Esto sugiere que el cálculo



de la inversa de $A$ podría haberse realizado a la vez con una matriz que incluyera las columnas $\begin{pmatrix}1\\ 0\end{pmatrix}$ y $\begin{pmatrix}0\\ 1\end{pmatrix}$; es decir:

$$\left(\begin{array}{cc|cc}2 & 1 & 1 & 0\\ 4 & 3 & 0 & 1\end{array}\right)$$

que es una matriz de orden $2\times 4$ de la forma $(A \mid I_2)$. Realizando operaciones elementales en esta matriz se tiene:

$$\left(\begin{array}{cc|cc}2 & 1 & 1 & 0\\ 4 & 3 & 0 & 1\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|cc}2 & 1 & 1 & 0\\ 0 & 1 & -2 & 1\end{array}\right) \xrightarrow{iii)} \left(\begin{array}{cc|cc}2 & 0 & 3 & -1\\ 0 & 1 & -2 & 1\end{array}\right) \xrightarrow{i)} \left(\begin{array}{cc|cc}1 & 0 & 3/2 & -1/2\\ 0 & 1 & -2 & 1\end{array}\right)$$

es decir, $(I_2 \mid A^{-1})$.

> Por tanto, para calcular la inversa de una matriz cuadrada $A$ de orden $n$ se reduce la matriz $(A \mid I_n)$ a la matriz $(I_n \mid B)$ mediante operaciones elementales. Entonces $A^{-1}=B$.

EJEMPLO D. Para calcular la inversa de la matriz

$$A=\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 3\\ 2 & -1 & -8\end{pmatrix}$$

escribimos

$$\left(\begin{array}{ccc|ccc}1 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 3 & 0 & 1 & 0\\ 2 & -1 & -8 & 0 & 0 & 1\end{array}\right) \mapsto \left(\begin{array}{ccc|ccc}1 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 3 & 0 & 1 & 0\\ 0 & -5 & -8 & -2 & 0 & 1\end{array}\right) \mapsto \left(\begin{array}{ccc|ccc}1 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 3 & 0 & 1 & 0\\ 0 & 0 & 7 & -2 & 5 & 1\end{array}\right)$$

$$\mapsto \left(\begin{array}{ccc|ccc}1 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 3 & 0 & 1 & 0\\ 0 & 0 & 1 & -2/7 & 5/7 & 1/7\end{array}\right) \mapsto \left(\begin{array}{ccc|ccc}1 & 2 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 6/7 & -8/7 & -3/7\\ 0 & 0 & 1 & -2/7 & 5/7 & 1/7\end{array}\right)$$

$$\mapsto \left(\begin{array}{ccc|ccc}1 & 0 & 0 & -5/7 & 16/7 & 6/7\\ 0 & 1 & 0 & 6/7 & -8/7 & -3/7\\ 0 & 0 & 1 & -2/7 & 5/7 & 1/7\end{array}\right)$$

Por tanto, $A$ es invertible y

$$A^{-1}=\begin{pmatrix} -5/7 & 16/7 & 6/7 \\ 6/7 & -8/7 & -3/7 \\ -2/7 & 5/7 & 1/7 \end{pmatrix}$$

(Comprobarlo calculando $A^{-1}\cdot A$).

EJEMPLO E. Tratamos de encontrar la inversa de la aplicación $f$: $\mathbb{R}^2 \mapsto \mathbb{R}^2$ dada por $f(x_1, x_2)=(3x_1+2x_2, x_1+2x_2)$. Puesto que la matriz de $f$ es

$$A=\begin{pmatrix} 3 & 2 \\ 1 & 2 \end{pmatrix}$$

tenemos

$$(A \mid I_2)=\left(\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right) \mapsto \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & -4 & 1 & -3 \end{array}\right)$$

$$\mapsto \left(\begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 1 & -1/4 & 3/4 \end{array}\right) \mapsto \left(\begin{array}{cc|cc} 1 & 0 & 1/2 & -1/2 \\ 0 & 1 & -1/4 & 3/4 \end{array}\right)$$

Por tanto, $f^{-1}$ tiene como matriz

$$A^{-1}=\begin{pmatrix} 1/2 & -1/2 \\ -1/4 & 3/4 \end{pmatrix}$$

y en consecuencia:

$$f^{-1}(x_1, x_2)=\left(\frac{1}{2}x_1-\frac{1}{2}x_2, -\frac{1}{4}x_1+\frac{3}{4}x_2\right)$$

(¡comprobarlo!).

* * *

El lector se preguntará si existe alguna forma de determinar si una matriz cuadrada $A$ posee inversa sin necesidad de calcularla. La respuesta es afirmativa y la produce el teorema de Rouché-Frobenius.

Supongamos que $A=(a_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ y $X=(x_{ij})_{\substack{i=1,\dots,n\\ j=1,\dots,n}}$ es la inversa de $A$; puesto que $A\cdot X=I_n$ se tiene que

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix}=\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{lugar } j=1, 2, \dots, n$$

lo cual constituyen $n$ sistemas de $n$ ecuaciones cada uno con $n$ incógnitas.



Si $A$ posee una inversa, cada uno de los $n$ sistemas anteriores posee solución única. Por el teorema de Rouché-Frobenius se ha de tener que $r(A)=n$.

Recíprocamente, si $r(A)=n$, cada uno de los sistemas anteriores posee solución única, ya que

$$r(A)=r\left(\begin{array}{c|c} & 0 \\ & \vdots \\ A & 1 \\ & \vdots \\ & 0 \end{array}\right)=n$$

Hemos obtenido el siguiente resultado:

TEOREMA 3

Una matriz cuadrada $A$ de orden $n$ es invertible si y sólo si $r(A)=n$.

Como aplicación, observar que la matriz

$$A=\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 3 & -1 \end{pmatrix}$$

no posee inversa, ya que

$$\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 1 & 3 & -1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & -1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$$

y, por tanto, $r(A)=2\neq 3$.

* * *

PROPOSICIÓN 4

Si $f$: $S\mapsto T$ y $g$: $T\mapsto U$ son invertibles, $g\circ f$: $S\mapsto U$ es invertible y $(g\circ f)^{-1}=f^{-1}\circ g^{-1}$.

*Nota.* De la proposición 4 se deduce que si $f$ y $g$ son aplicaciones lineales con matrices $A$ y $B$, respectivamente, $(AB)^{-1}=B^{-1}A^{-1}$, siempre que existan las inversas.

*Demostración.* Basta probar que $(f^{-1}\circ g^{-1})\circ(g\circ f)=I_S$ y $(g\circ f)\circ(f^{-1}\circ g^{-1})=I_U$. Tenemos que

$$(f^{-1}\circ g^{-1})\circ(g\circ f)(s)=f^{-1}(g^{-1}(g(f(s))))=f^{-1}(g^{-1}\circ g(f(s))=f^{-1}(f(s))=s$$

y análogamente:

$$(g\circ f)\circ(f^{-1}\circ g^{-1})(u)=u$$

■

EJEMPLO F. Tratemos de calcular $(AB)^{-1}$, donde

$$A=\begin{pmatrix}2 & 1\\ 4 & 3\end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix}3 & 2\\ 1 & 2\end{pmatrix}$$

Puesto que

$$A^{-1}=\begin{pmatrix}3/2 & -1/2\\ -2 & 1\end{pmatrix} \text{ (ejemplo C) y } B^{-1}=\begin{pmatrix}1/2 & -1/2\\ -1/4 & 3/4\end{pmatrix} \text{ (ejemplo E)}$$

tenemos que

$$(AB)^{-1}=B^{-1}A^{-1}=\begin{pmatrix}1/2 & -1/2\\ -1/4 & 3/4\end{pmatrix}\begin{pmatrix}3/2 & -1/2\\ -2 & 1\end{pmatrix}=$$

$$=\begin{pmatrix}3/4+1 & -1/4-1/2\\ -3/8-6/4 & 1/8+3/4\end{pmatrix}=\begin{pmatrix}7/4 & -3/4\\ -15/8 & 7/8\end{pmatrix}$$

* * *

Para finalizar esta sección aplicamos los conocimientos adquiridos para resolver algunos sistemas de ecuaciones lineales.

Supongamos que $A$ es una matriz cuadrada de orden $n$ e invertible y tratemos de resolver el sistema $A\vec{x}=\vec{b}$. Si $A^{-1}$ denota la inversa de $A$ tenemos que

$$A^{-1}(A\vec{x})=A^{-1}\vec{b}$$

y, por tanto, $\vec{x}=A^{-1}\vec{b}$.

EJEMPLO G. Tratemos de resolver el sistema

$$\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 3\\ 2 & -1 & -8\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}14\\ 7\\ 0\end{pmatrix}$$



dado en forma matricial. La matriz de este sistema es invertible como se ha visto en el ejemplo D, en donde también se ha calculado su inversa. Por tanto:

$$\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 3\\ 2 & -1 & -8\end{pmatrix}^{-1}\begin{pmatrix}14\\ 7\\ 0\end{pmatrix}=\begin{pmatrix}-5/7 & 16/7 & 6/7\\ 6/7 & -8/7 & -3/7\\ -2/7 & 5/7 & 1/7\end{pmatrix}\begin{pmatrix}14\\ 7\\ 0\end{pmatrix}=\begin{pmatrix}6\\ 4\\ 1\end{pmatrix}$$

(¡comprobar el resultado!).

## EJERCICIOS 1.4

**1.** Demostrar que las aplicaciones $f\colon \mathbb{R}^2 \mapsto \mathbb{R}^2$ dada por $f(x_1, x_2)=(x_1+x_2, x_1+2x_2)$ y $g\colon \mathbb{R}^2 \mapsto \mathbb{R}^2$ dada por $g(x_1, x_2)=(2x_1-x_2, -x_1+x_2)$ son inversas una de la otra.

**2.** Dadas $f(x)=3x-2$ y $g(x)=\cos x$, aplicaciones de $\mathbb{R}$ en $\mathbb{R}$, calcular $g\circ f^{-1}$ y $f^{-1}\circ f^{-1}\circ g$.

**3.** Se denomina rango de una aplicación al rango de su matriz asociada. Calcular el rango de las siguientes aplicaciones lineales:

*a)* $f(x_1, x_2)=(0, x_1, x_1)$
*b)* $f(x_1, x_2, x_3)=(x_1+x_2+x_3, x_1+2x_2+3x_3, x_1-x_2-3x_3)$
*c)* $f(x_1, x_2, x_3)=(x_1+2x_2+x_3, -x_1+2x_2+x_3)$
*d)* $f(x_1, x_2)=(x_1+x_2, 0, x_1+x_2)$
*e)* $f(x_1, x_2, x_3, x_4)=(x_1, x_1+x_2, x_1+x_2+x_3, x_1+x_2+x_3+x_4)$

**4.** Encontrar, si es posible, la inversa de las siguientes matrices:

*a)* $\begin{pmatrix}1 & 2\\ -1 & 1\end{pmatrix}$ *b)* $\begin{pmatrix}1 & 2\\ 3/2 & 3\end{pmatrix}$ *c)* $\begin{pmatrix}1 & 0 & 0\\ 2 & 1 & -1\\ 3 & 1 & 4\end{pmatrix}$

*d)* $\begin{pmatrix}2 & 0 & 1\\ 1 & 3 & 2\\ 1 & -3 & 3\end{pmatrix}$ *e)* $\begin{pmatrix}1 & 2 & 0\\ 0 & -1 & 1\\ 3 & 1 & 5\end{pmatrix}$ *f)* $\begin{pmatrix}1 & 2 & 3 & 4\\ 1 & 3 & 2 & 6\\ -2 & -2 & -6 & 0\\ 1 & 4 & 1 & 7\end{pmatrix}$

**5.** Encontrar la inversa de cada una de las siguientes matrices:

*a)* $\begin{pmatrix}1 & 0\\ a & 1\end{pmatrix}$ *b)* $\begin{pmatrix}1 & 1/a\\ a & -1\end{pmatrix}$, $a\neq 0$ *c)* $\begin{pmatrix}1 & 0 & a\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$ *d)* $\begin{pmatrix}0 & 0 & a\\ 0 & a & 0\\ a & 0 & 0\end{pmatrix}$, $a\neq 0$

**6.** Encontrar, si es posible, la inversa de las siguientes aplicaciones lineales:

*a)* $f(x_1, x_2, x_3)=(x_1+x_2+x_3,\ 2x_1+x_3,\ x_1+x_2+2x_3)$

*b)* $f(x_1, x_2)=(x_1+cx_2,\ x_1-cx_2),\ c\in\mathbb{R}$

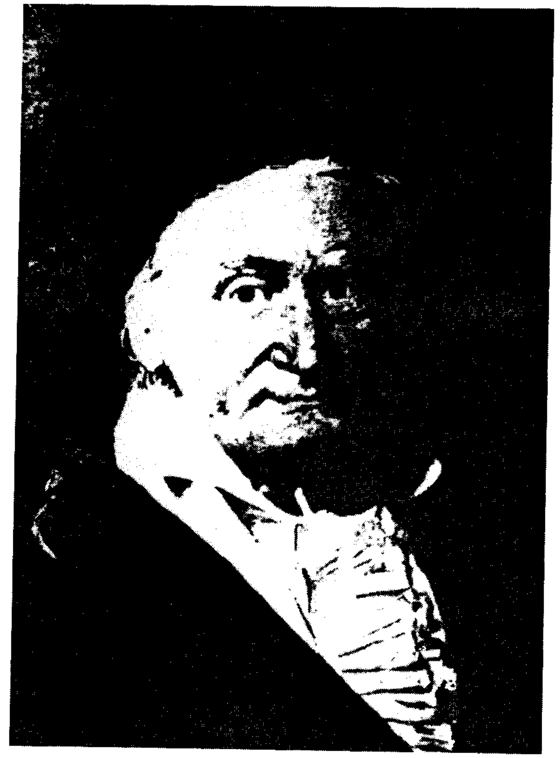
**7.** Encontrar $A^{-1}$ y resolver el sistema $A\vec{x}=\vec{b}$ para:

*a)* $A=\begin{pmatrix}1 & a\\ 0 & 1\end{pmatrix},\ \vec{b}=\begin{pmatrix}1\\ 2\end{pmatrix}$ *b)* $A=\begin{pmatrix}-3 & 1 & 1\\ 2 & 6 & 5\\ -5 & 7 & 6\end{pmatrix},\ \vec{b}=\begin{pmatrix}16\\ -16\\ 32\end{pmatrix}$

*c)* $A=\begin{pmatrix}2 & 1 & 5\\ 0 & 3 & 1\\ 0 & 0 & -1\end{pmatrix},\ \vec{b}=\begin{pmatrix}\frac{1}{6}\\ -\frac{2}{6}\\ 0\end{pmatrix}$



## BIOGRAFIA



*Carl Friedrich Gauss* nació el 30 de abril de 1777 en Brunswick, un pueblecito que actualmente pertenece a Alemania Oriental. Impresionados por su habilidad para las matemáticas y los idiomas, su madre y sus profesores le recomendaron al Duke de Brunswick, quien le proporcionó la ayuda económica necesaria para estudiar en la Universidad de Göttingen.

Cuando solamente tenía 19 años Gauss realizó uno de los descubrimientos más espectaculares de las matemáticas del siglo XVIII: fue el primer matemático que construyó un polígono regular de 17 lados utilizando solamente regla y compás. Euclides sabía como construir polígonos regulares de 3, 4, 5 y 15 lados, así como aquellos que se obtienen duplicando éstos. Animado por su descubrimiento, Gauss logró encontrar una solución algebraica al problema de construir con regla y compás un polígono regular de $n$ lados y desarrolló un criterio basado en la teoría de números con el cual puede decidirse si un polígono regular de un cierto número de lados puede construirse geométricamente.

En 1799 se le concedió el título de Doctor por la Universidad de Helmsted. En su tesis, Gauss dio una demostración del Teorema Fundamental del Algebra, en el que se muestra que toda ecuación algebraica con coeficientes complejos tiene soluciones complejas. Los números complejos, cuya formulación actual se debe, entre otros, a Gauss, se estudiarán en el capítulo 4.

A los 24 años publicó una de las obras más completas de la historia de las matemáticas; se tituló *Disquisitiones Aritmeticae* y en ella formuló conceptos y métodos de enorme importancia en el desarrollo posterior de la Teoría de Números.

El Duke de Brunswick financió tan generosamente sus investigaciones que en 1803 pudo renunciar a una oferta de profesor en la Academia de Ciencias de San Petersburgo. En 1807 fue nombrado profesor de astronomía y director del observatorio de la Universidad de Göttingen, en donde permaneció durante toda su vida.

En 1801, Gauss tuvo la oportunidad de aplicar su habilidad matemática y sus ideas en el campo de la astronomía. El primer día del año fue descubierto un asteroide, que posteriormente sería llamado Ceres. A pesar de que pudo ser observado durante 40 días, ningún astrónomo fue capaz de calcular su órbita. Con solamente tres observaciones, Gauss desarrolló una técnica para calcular su órbita de manera que Ceres pudo ser localizado fácilmente a finales de 1801 y comienzos de 1802. El método utilizado en sus cálculos fue publicado en 1809 en su libro *Theoria Motus Corporum Coelestium*; este método continúa utilizándose actualmente, con ligeras modificaciones, en los cálculos que se realizan con ordenadores.

Carl Friedrich Gauss contribuyó de manera decisiva al desarrollo de varias ramas de la matemática durante el siglo XIX.

Murió el 23 de febrero de 1855 y poco después de su muerte se pusieron en circulación varias monedas en su honor.

## CAPITULO 2

# DETERMINANTES Y SUS APLICACIONES





En el capítulo anterior se ha llegado a la conclusión de que un sistema de ecuaciones lineales de la forma

$$A\vec{x} = \vec{b}$$

es muy fácil de resolver si se conoce la inversa de la matriz $A$. También allí se dio un método para calcular la inversa de una matriz, basado en la realización de operaciones elementales con sus filas.

En este capítulo introduciremos el *concepto de determinante de una matriz*, que nos permitirá obtener una fórmula para calcular la matriz inversa de una dada. Las propiedades de los determinantes, básicas en este capítulo y en muchos de los restantes, serán estudiadas minuciosamente.

En particular, estableceremos la relación entre el rango de una matriz y la anulación de ciertos determinantes, lo cual nos permitirá encontrar una fórmula para resolver sistemas de ecuaciones compatibles.

## 2.1. DETERMINANTES DE MATRICES DE ORDEN 2 Y DE ORDEN 3

Dado el sistema

$$\left.\begin{aligned} ax_1+bx_2=e_1 \\ cx_1+dx_2=e_2 \end{aligned}\right\} \qquad \text{(I)}$$

de dos ecuaciones con dos incógnitas, intentamos buscar condiciones sobre los coeficientes del sistema para que posea solución única. Por el teorema de Rouché-Frobenius esto se cumple solamente cuando

$$r\begin{pmatrix} a & b \\ c & d \end{pmatrix}=2$$

Realizando operaciones elementales con las filas de la matriz de los coeficientes, se tiene:

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{cF_1;\, aF_2} \begin{pmatrix} ac & bc \\ ca & da \end{pmatrix} \xrightarrow{F_2-F_1} \begin{pmatrix} ac & bc \\ 0 & ad-bc \end{pmatrix}$$

Si $ad-bc$ es nulo, la matriz $A$ de los coeficientes del sistema tiene rango inferior a 2 y, por tanto, el sistema (I) no tiene solución única.

Bajo la condición de que el número $ad-bc$ sea *no* nulo, el sistema (I) puede resolverse. Multiplicando la primera ecuación por $d$ y la segunda por $b$ y restándolas se obtiene

$$(ad-bc)x_1=e_1d-e_2b$$

Por tanto:

$$x_1=\frac{e_1d-e_2b}{ad-bc}.$$

Multiplicando la primera ecuación por $c$ y la segunda por $a$ y restando la primera de la segunda se obtiene

$$(ad-bc)x_2=ae_2-ce_1$$

Por tanto,

$$x_2=\frac{ae_2-ce_1}{ad-bc}.$$



Dada la matriz

$$A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

el número $ad-bc$, a ella asociado, recibe el nombre de *determinante de A* y se denota mediante $|A|$. Por tanto, hemos demostrado el siguiente resultado:

PROPOSICIÓN 1

El sistema (I) posee solución única si y sólo si el determinante de la matriz de sus coeficientes es no nulo. Además, sus soluciones se calculan mediante las fórmulas

$$x_1=\frac{\begin{vmatrix} e_1 & b \\ e_2 & d \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}} \quad \text{y} \quad x_2=\frac{\begin{vmatrix} a & e_1 \\ c & e_2 \end{vmatrix}}{\begin{vmatrix} a & b \\ c & d \end{vmatrix}}$$

Las fórmulas anteriores reciben el nombre de *regla de Cramer* para la resolución de sistemas compatibles determinados de dos ecuaciones con dos incógnitas.

Observar que las soluciones se obtienen como fracciones que tienen como denominador el determinante de $A$; el numerador de la fracción de $x_1$ es el determinante de la matriz que se obtiene sustituyendo la primera columna de la matriz $A$ por el vector columna que aparece a la derecha del sistema (I); el numerador de la fracción que determina $x_2$ es el determinante de la matriz que se obtiene sustituyendo la segunda columna de la matriz $A$ por el vector columna que aparece a la derecha del sistema (I).

EJEMPLO A. El sistema

$$\left.\begin{aligned} 2x_1-x_2=1 \\ x_1+3x_2=0 \end{aligned}\right\}$$

tiene solución única ya que

$$\begin{vmatrix} 2 & -1 \\ 1 & 3 \end{vmatrix}=2\cdot 3-(-1)\cdot 1=6+1=7$$

Su solución es

$$x_1=\frac{\begin{vmatrix} 1 & -1 \\ 0 & 3 \end{vmatrix}}{7}=\frac{3}{7} \quad ; \quad x_2=\frac{\begin{vmatrix} 2 & 1 \\ 1 & 0 \end{vmatrix}}{7}=\frac{-1}{7}$$

* * *

La principal propiedad de los determinantes es su linealidad; esta propiedad, entre otras, se demuestra en el siguiente resultado.

TEOREMA 2

*i)* $\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix}$ y $\begin{vmatrix} a & b \\ c+c' & d+d' \end{vmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a & b \\ c' & d' \end{vmatrix}$

*ii)* Si $B$ es la matriz que se obtiene a partir de una matriz $A$ multiplicando una cualquiera de sus filas por un número real $r$ se tiene que $|B|=r|A|$.

*iii)* Si $B$ es la matriz que se obtiene a partir de $A$ intercambiando dos de sus filas se tiene que $|B|=-|A|$.

*iv)* Si $B$ es la matriz que se obtiene a partir de $A$ sumando un múltiplo de una fila de $A$ a otra fila de $A$ se tiene que $|B|=|A|$.

*v)* El determinante de la matriz identidad es 1.

*Nota.* Las propiedades *i)* y *ii)* son las que determinan la linealidad por filas del determinante. Las propiedades *ii), iii)* y *iv)* nos muestran el comportamiento del determinante frente a las operaciones elementales con las filas de una matriz.

*Demostración.* Unicamente demostraremos la primera parte de *i)* dejando el resto de las demostraciones para el lector debido a su sencillez. Utilizando la definición de determinante se tiene que:

$$\begin{vmatrix} a+a' & b+b' \\ c & d \end{vmatrix} = (a+a')d - c(b+b') = ad + a'd - cb - cb' =$$
$$= (ad-cb) + (a'd-cb') = \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \begin{vmatrix} a' & b' \\ c & d \end{vmatrix} \qquad \blacksquare$$

que era lo que queríamos demostrar.

*Observación.* El determinante de una matriz de orden 2 con dos filas iguales es nulo:

$$\begin{vmatrix} a & b \\ a & b \end{vmatrix} = ab - ab = 0$$

* * *

A continuación pasamos a estudiar una posible definición del determinante de una matriz de orden 3. Para ello supongamos que el sistema

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}\right\} \qquad \text{(II)}$$



de tres ecuaciones con tres incógnitas tiene solución única. Para encontrar esta solución multiplicamos la primera ecuación por

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}$$

la segunda por

$$-\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

y la tercera por

$$\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

y las sumamos. Después de realizadas las simplificaciones adecuadas se obtiene:

$$\left(a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{21}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\right)x_1 =$$
$$= b_1\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - b_2\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + b_3\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}$$

Por analogía con el caso del sistema de dos ecuaciones con dos incógnitas denominamos *determinante de la matriz*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

a la expresión que multiplica a $x_1$ en la fórmula anterior.

Una forma sencilla de recordar esta definición es haciendo uso de los conceptos de matriz *adjunta* de un elemento y de *menor*. Dada una matriz $(a_{ij})_{i,j=1,2,\dots,n}$ se denomina *matriz adjunta* del elemento que ocupa el lugar $(i, j)$ a la matriz que se obtiene eliminando la fila $i$ y la columna $j$ de la matriz dada, y se denota por $A_{ij}$. El determinante de $A_{ij}$ se denomina *menor* del elemento que ocupa el lugar $(i, j)$, es decir, $a_{ij}$.

Por ejemplo, en la matriz

$$A = \begin{pmatrix} 1 & 0 & -2 \\ 2 & 1 & 3 \\ 5 & -1 & 7 \end{pmatrix}$$

$$A_{12} = \begin{pmatrix} \boxed{1} & \boxed{0} & \boxed{-2} \\ 2 & \boxed{1} & 3 \\ 5 & \boxed{-1} & 7 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix} \quad \text{y} \quad |A_{12}| = 14 - 15 = -1$$

mientras que

$$A_{22}=\begin{pmatrix} 1 & \boxed{0} & -2 \\ \boxed{2} & \boxed{1} & \boxed{3} \\ 5 & \boxed{-1} & 7 \end{pmatrix}=\begin{pmatrix} 1 & -2 \\ 5 & 7 \end{pmatrix} \quad \text{y} \quad |A_{22}|=7+10=17$$

Con estas notaciones podemos dar la siguiente definición.

DEFINICIÓN *(Determinante de una matriz de orden 3)*

Dada la matriz

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

definimos su *determinante* mediante la fórmula

$$|A|=a_{11}|A_{11}|-a_{21}|A_{21}|+a_{31}|A_{31}|$$

EJEMPLO B

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ -1 & 1 & 2 \end{vmatrix}=1\begin{vmatrix} \boxed{1} & \boxed{2} & \boxed{0} \\ \boxed{2} & 1 & 3 \\ \boxed{-1} & 1 & 2 \end{vmatrix}-2\begin{vmatrix} \boxed{1} & 2 & 0 \\ \boxed{2} & \boxed{1} & \boxed{3} \\ \boxed{-1} & 1 & 2 \end{vmatrix}+(-1)\begin{vmatrix} \boxed{1} & 2 & 0 \\ \boxed{2} & 1 & 3 \\ \boxed{-1} & \boxed{1} & \boxed{2} \end{vmatrix}=$$

$$=1\begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix}-2\begin{vmatrix} 2 & 0 \\ 1 & 2 \end{vmatrix}+(-1)\begin{vmatrix} 2 & 0 \\ 1 & 3 \end{vmatrix}=1\cdot(-1)-2\cdot(4)+(-1)\cdot 6=-15$$

* * *

Si desarrollamos la expresión que nos da el determinante de una matriz de orden 3 se obtienen los siguientes seis productos:

$$\begin{aligned} |A|&=a_{11}(a_{22}a_{33}-a_{23}a_{32})-a_{21}(a_{12}a_{33}-a_{13}a_{32})+ \\ &+a_{31}(a_{12}a_{23}-a_{13}a_{22})=a_{11}a_{22}a_{33}+a_{21}a_{13}a_{32}+ \\ &+a_{31}a_{12}a_{23}-a_{31}a_{13}a_{22}-a_{21}a_{12}a_{33}-a_{11}a_{23}a_{32} \end{aligned}$$

Los términos positivos de esta expresión son los productos de los términos de la diagonal principal de $A$ y los productos de los términos que ocupan los vértices de los triángulos de la figura adjunta:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \text{(términos positivos)}$$



Un resultado análogo se obtiene para los términos negativos:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \text{(términos negativos)}$$

Esta regla nemotécnica puede emplearse para calcular determinantes, pero sólo funciona con matrices de orden 3.

EJEMPLO C. Los términos positivos del determnante de la matriz

$$A=\begin{pmatrix} 1 & -2 & 3 \\ 5 & 0 & -2 \\ -1 & 1 & 1 \end{pmatrix}$$

son

$$\begin{vmatrix} 1 & -2 & 3 \\ 5 & 0 & -2 \\ -1 & 1 & 1 \end{vmatrix}: 1\cdot 0\cdot 1+(-2)(-2)(-1)+5\cdot 1\cdot 3=11$$

y los términos negativos son

$$\begin{vmatrix} 1 & -2 & 3 \\ 5 & 0 & -2 \\ -1 & 1 & 1 \end{vmatrix}: +(-1)\cdot 0\cdot 3+1\cdot 1\cdot(-2)+5(-2)\cdot 1=-12$$

Por tanto, $|A|=11-(-12)=23$ (¡comprobar el resultado utilizando la definición!).

* * *

A continuación probamos que el determinante de una matriz de orden 3 tiene propiedades análogas a las enunciadas en el teorema 2 para determinantes de matrices de orden 2. Concretamente:

TEOREMA 3

Las propiedades análogas a *i*), *ii*), *iii*), *iv*) y *v*) del teorema 2 son ciertas para matrices de orden 3.

*Demostración.* *i*) Si la primera fila es suma de dos se tiene:

$$\begin{vmatrix} a_{11}+a'_{11} & a_{12}+a'_{12} & a_{13}+a'_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}=(a_{11}+a'_{11})\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix}-a_{21}\begin{vmatrix} a_{12}+a'_{12} & a_{13}+a'_{13} \\ a_{32} & a_{33} \end{vmatrix}+$$

$$+a_{31}\begin{vmatrix} a_{12}+a'_{12} & a_{13}+a'_{13}\\ a_{22} & a_{23}\end{vmatrix} = a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} - a_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix} + a_{31}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix} +$$

$$a'_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} - a_{21}\begin{vmatrix}a'_{12}&a'_{13}\\a_{32}&a_{33}\end{vmatrix} + a_{31}\begin{vmatrix}a'_{12}&a'_{13}\\a_{22}&a_{23}\end{vmatrix} =$$

$$=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix} + \begin{vmatrix}a'_{11}&a'_{12}&a'_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}$$

en donde en la segunda igualdad se ha utilizado la propiedad *i*) del teorema 2. Igualmente puede demostrarse cuando la suma aparece en las filas segunda o tercera.

*ii*) Si multiplicamos una fila de una matriz (por ejemplo, la segunda) por un número real $r$ se obtiene:

$$\begin{vmatrix}a_{11}&a_{12}&a_{13}\\ra_{21}&ra_{22}&ra_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix} = a_{11}\begin{vmatrix}ra_{22}&ra_{23}\\a_{32}&a_{33}\end{vmatrix} - ra_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix} + a_{31}\begin{vmatrix}a_{12}&a_{13}\\ra_{22}&ra_{23}\end{vmatrix} =$$

$$= ra_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} - ra_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix} + ra_{31}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix} = r\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}$$

en donde se ha utilizado la propiedad *ii*) del teorema 2.

*iii*) Intercambiando, por ejemplo, la primera y la segunda fila de la matriz $A$ obtenemos:

$$\begin{vmatrix}a_{21}&a_{22}&a_{23}\\a_{11}&a_{12}&a_{13}\\a_{31}&a_{32}&a_{33}\end{vmatrix} = a_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix} - a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} + a_{31}\begin{vmatrix}a_{22}&a_{23}\\a_{12}&a_{13}\end{vmatrix} =$$

$$= -a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} + a_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{32}&a_{33}\end{vmatrix} - a_{31}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix} =$$

$$= -\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{vmatrix}$$

*iv*) Tomemos

$$B=\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}+ra_{11}&a_{32}+ra_{12}&a_{33}+ra_{13}\end{vmatrix}$$



que se ha obtenido de $A$ sumando a la tercera fila $r$ veces la primera. Debido a las propiedades *i*) y *ii*) ya demostradas se tiene que

$$|B| = |A| + r\begin{vmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{11}&a_{12}&a_{13}\end{vmatrix} = |A| + r\left[a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{12}&a_{13}\end{vmatrix} - a_{21}\begin{vmatrix}a_{12}&a_{13}\\a_{12}&a_{13}\end{vmatrix} +\right.$$

$$\left. + a_{11}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix}\right] = |A| + r\left[-a_{11}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{23}\end{vmatrix} + a_{11}\begin{vmatrix}a_{12}&a_{13}\\a_{22}&a_{-3}\end{vmatrix}\right] = |A|$$

que era lo que se quería demostrar.

La propiedad *v*) es inmediata a partir de la definición. ■

Terminamos esta sección con un ejemplo que muestra cómo pueden utilizarse las propiedades *iii*), *iv*) y *v*) del teorema 3 para calcular determinantes.

EJEMPLO D

$$\begin{vmatrix}1&2&0\\2&1&3\\-1&1&2\end{vmatrix} \overset{iv)}{=\!=} \begin{vmatrix}1&2&0\\0&-3&3\\0&3&2\end{vmatrix} \overset{iv)}{=\!=} \begin{vmatrix}1&2&0\\0&-3&3\\0&0&5\end{vmatrix} \overset{iii)}{=\!=} (-3)(5)\begin{vmatrix}1&2&0\\0&1&-1\\0&0&1\end{vmatrix}$$

$$\overset{iv)}{=\!=} -15\begin{vmatrix}1&2&0\\0&1&0\\0&0&1\end{vmatrix} \overset{iv)}{=\!=} -15\begin{vmatrix}1&0&0\\0&1&0\\0&0&1\end{vmatrix} \overset{v)}{=\!=} -15\cdot 1 = -15$$

que es el mismo resultado que se obtuvo en el ejemplo B.

## EJERCICIOS 2.1

**1.** Calcular los siguientes determinantes de matrices de orden 2:

*a*) $\begin{vmatrix}2&1\\1&2\end{vmatrix}$ *b*) $\begin{vmatrix}3&1\\0&4\end{vmatrix}$ *c*) $\begin{vmatrix}a&1\\a&b\end{vmatrix}$

*d*) $\begin{vmatrix}\cos\alpha&-\operatorname{sen}\alpha\\\operatorname{sen}\alpha&\cos\alpha\end{vmatrix}$ *e*) $\begin{vmatrix}\cos\alpha&\operatorname{sen}\alpha\\\operatorname{sen}\alpha&-\cos\alpha\end{vmatrix}$

**2.** Hallar la inversa de las matrices *c*) y *d*) anteriores.

**3.** Utilizar la regla de Cramer para resolver los sistemas $A\vec{x}=\vec{b}$, donde

$$A=\begin{pmatrix}2&4\\2&3\end{pmatrix}$$

y

*a*) $\vec{b}=\begin{pmatrix}2\\1\end{pmatrix}$ *b*) $\vec{b}=\begin{pmatrix}\pi\\\pi\end{pmatrix}$ *c*) $\vec{b}=\begin{pmatrix}0\\0\end{pmatrix}$

**4.** Utilizar la definición de determinante para calcular:

*a)* $\begin{vmatrix} 1 & 4 & 3 \\ 2 & 1 & 0 \\ 0 & 3 & 1 \end{vmatrix}$ *b)* $\begin{vmatrix} 3 & 2 & -1 \\ 7 & 0 & 7 \\ 1 & 4 & -7 \end{vmatrix}$ *c)* $\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ x^2 & y^2 & z^2 \end{vmatrix}$ *d)* $\begin{vmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{vmatrix}$

**5.** Calcular *a)* y *b)* anteriores utilizando la regla nemotécnica para calcular determinantes de orden 3.

**6.** Calcular el determinante de la matriz *c)* del ejercicio 4 realizando operaciones elementales en sus filas y utilizando las propiedades de los determinantes.

**7.** Calcular

$$\begin{vmatrix} \cos\alpha & -\operatorname{sen}\alpha & 1 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha & \cos\alpha \end{vmatrix}$$

**8.** Demostrar que si $a$, $b$ y $c$ son números reales, las raíces de la ecuación

$$\begin{vmatrix} a-x & b \\ b & c-x \end{vmatrix} = 0$$

son reales.

## 2.2. DEFINICION GENERAL DE DETERMINANTE. PROPIEDADES

En esta sección daremos la definición de determinante de una matriz cuadrada de orden $n$ de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Esta definición es la generalización de la definición de determinante de matrices de orden 3 dada en la sección 2.1. Recordemos que $|A_{ij}|$ denota el determinante de la matriz de orden $n-1$ que se obtiene suprimiendo la fila $i$ y la columna $j$ de la matriz $A$; $|A_{ij}|$ se denomina *menor* del elemento que ocupa el lugar $(i, j)$ en $A$.

DEFINICIÓN (*Determinante de una matriz de orden n*)

Dada una matriz $A$ como arriba, su determinante se define de la forma

$$|A| = a_{11}|A_{11}| - a_{21}|A_{21}| + \cdots + (-1)^{n+1}a_{n1}|A_{n1}|$$



*Observación.* En la definición que acabamos de dar el determinante de una matriz de orden $n$ se reduce a calcular $n$ determinantes de matrices de orden $n-1$; cada determinante de orden $n-1$ se reduce a calcular $n-1$ determinantes de orden $n-2$. Continuando este proceso se llega a obtener determinantes de orden 2, que son fáciles de calcular.

EJEMPLO A

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 4 & 2 & 0 & 3 \\ 2 & 2 & 1 & 1 \end{vmatrix} = 1\begin{vmatrix} 2 & 2 & 3 \\ 2 & 0 & 3 \\ 2 & 1 & 1 \end{vmatrix} - 0\begin{vmatrix} 2 & 3 & 4 \\ 2 & 0 & 3 \\ 2 & 1 & 1 \end{vmatrix} + 4\begin{vmatrix} 2 & 3 & 4 \\ 2 & 2 & 3 \\ 2 & 1 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 & 4 \\ 2 & 2 & 3 \\ 2 & 0 & 3 \end{vmatrix}$$

$$\begin{vmatrix} 2 & 2 & 3 \\ 2 & 0 & 3 \\ 2 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} 0 & 3 \\ 1 & 1 \end{vmatrix} - 2\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} + 2\begin{vmatrix} 2 & 3 \\ 0 & 3 \end{vmatrix} = -6 + 2 + 12 = 8$$

$$\begin{vmatrix} 2 & 3 & 4 \\ 2 & 2 & 3 \\ 2 & 1 & 1 \end{vmatrix} = 2\begin{vmatrix} 2 & 3 \\ 1 & 1 \end{vmatrix} - 2\begin{vmatrix} 3 & 4 \\ 1 & 1 \end{vmatrix} + 2\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix} = -2 + 2 + 2 = 2$$

$$\begin{vmatrix} 2 & 3 & 4 \\ 2 & 2 & 3 \\ 2 & 0 & 3 \end{vmatrix} = 2\begin{vmatrix} 2 & 3 \\ 0 & 3 \end{vmatrix} - 2\begin{vmatrix} 3 & 4 \\ 0 & 3 \end{vmatrix} + 2\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix} = 12 - 18 + 2 = -4$$

Por tanto

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 4 & 2 & 0 & 3 \\ 2 & 2 & 1 & 1 \end{vmatrix} = 1\cdot 8 - 0 + 4\cdot 2 - 2(-4) = 8 + 8 + 8 = 24.$$

* * *

El cálculo de determinantes utilizando la definición dada resulta larguísimo cuando se trata de matrices de orden elevado. El estudio de las propiedades del determinante aliviará este problema.

Para demostrar estas propiedades utilizaremos un razonamiento denominado *método de inducción.*

Este método puede aplicarse cuando se trata de demostrar una propiedad $P(n)$ que depende de un número natural $n$. Para ello se procede como sigue:

1.º Se demuestra que $P(n)$ es cierta para un cierto número natural $n_0$, que generalmente es pequeño.

2.º Se acepta que la propiedad es cierta para todo número $k > n_0$ e inferior a $n$ (*hipótesis de inducción*) y utilizando esto se demuestra que también es cierta para $n$.

Estos dos resultados permiten concluir que la propiedad es cierta para todo $n \geqslant n_0$. En efecto, por 1.° es cierta para $n_0$; utilizando 2.° es también cierta para $n_0+1$; utilizando de nuevo 2.°, $P(n)$ es cierta para $n_0+2$, y este proceso puede continuarse hasta alcanzar cualquier número natural $n$ por muy elevado que sea.

Mostraremos con un ejemplo sencillo cómo se aplica el método de inducción. Se trata de demostrar que *la suma de los $n$ primeros números impares es $n^2$*.

Sea $P(n)=1+3+5+\cdots+(2n-1)=n^2$ la propiedad que queremos demostrar.

1.° $P(1)=1=1^2$ es cierta.

Para convencerse de que el resultado es cierto, el lector puede comprobar que $P(2)=1+3=4=2^2$ y $P(3)=1+3+5=9=3^2$.

2.° Supongamos que $P(k)$ es cierta para todo $k<n$, es decir se tiene

$$1+3+5+\cdots+(2k-1)=k^2, \quad k=1, 2, 3, \ldots, n-1$$

Ahora,

$$1+3+5+\cdots+(2n-1)=1+3+5+\cdots+(2n-3)+(2n-1)=$$

$$=(1+3+5+\cdots+[2(n-1)-1])+(2n-1) \overset{(*)}{=} (n-1)^2+(2n-1)=$$

$$=n^2-2n+1+2n-1=n^2$$

que era lo que queríamos demostrar. En la igualdad (*) es donde se ha utilizado la hipótesis de inducción, ya que se ha aplicado que $P(n-1)$ es cierta.

Algunos resultados que pueden probarse mediante el método de inducción se proponen como ejercicio al final de esta sección.

* * *

PROPIEDAD 1(P.1)

> Si una matriz tiene una fila de ceros, su determinante es nulo.

*Demostración.* Esta propiedad es fácilmente comprobable para matrices de orden 2:

$$\begin{vmatrix} a & b \\ 0 & 0 \end{vmatrix} = a\cdot 0 - b\cdot 0 = 0 \quad \text{y} \quad \begin{vmatrix} 0 & 0 \\ c & d \end{vmatrix} = 0\cdot d - 0\cdot c = 0$$

Aceptemos que es cierta para toda matriz de orden $n-1$ y tratemos de demostrarlo para una matriz de orden $n$ con una fila de ceros. Supongamos que los ceros ocupan la fila $i$ de la matriz, es decir:

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i-1,1} & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$



De la definición de determinante deducimos

$$|A|=a_{11}|A_{11}|-\cdots+(-1)^i a_{i-1,1}|A_{i-1,1}|+(-1)^{i+1}0|A_{i,1}|+\cdots+(-1)^{n+1}a_{n,1}|A_{n1}|$$

Todos los adjuntos $A_{k1}$ con $k\neq i$ tienen una fila de ceros y, por tanto, $|A_{k1}|=0$ por la hipótesis de inducción. Para $k=i$, $|A_{i1}|$ está multiplicado por 0 y, por tanto, todos los sumandos de la igualdad anterior son nulos. ■

PROPIEDAD 2 (P.2)

$$|C|=\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} +$$

$$+\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = |A|+|B|$$

para todo $i=1, 2, \ldots, n$.

*Demostración.* Esta propiedad se ha demostrado para matrices de orden 2 y de orden 3 en la sección anterior. Aceptemos que es cierta para toda matriz de orden menor que $n$. Tenemos

$$|C|=a_{11}|C_{11}|-\cdots+(-1)^{i+1}[a_{i1}+b_{i1}]|C_{i1}|+\cdots+(-1)^{n+1}a_{n1}|C_{n1}|$$

Cada adjunto $c_{k1}$ con $k\neq i$ es una matriz de orden $n-1$ con una fila que es una suma; por tanto:

$$|C_{k1}|=|A_{k1}|+|B_{k1}|$$

donde $A_{ki}$ y $B_{ki}$ son adjuntos de $A$ y $B$, respectivamente. Sustituyendo este resultado en la igualdad anterior se tiene:

$$|C|=a_{11}[|A_{1i}|+|B_{1i}|]-\cdots+(-1)^{i+1}[a_{i1}+b_{i1}]|C_{i1}|+\cdots+(-1)^{n+1}a_{n1}[|A_{n1}|+|B_{n1}|]$$

Puesto que $|C_{i1}|=|A_{i1}|=|B_{i1}|$ se tiene que

$$|C|=a_{11}|A_{11}|+\cdots+(-1)^{i+1}a_{i1}|A_{i1}|+\cdots+(-1)^{n+1}a_{n1}|A_{n1}|+$$
$$+a_{11}|B_{11}|+\cdots+(-1)^{i+1}b_{i1}|B_{i1}|+\cdots+(-1)^{n+1}a_{n1}|B_{n1}|=|A|+|B|$$ ■

PROPIEDAD 3 (P.3)

> Si $B$ es la matriz que se obtiene a partir de una matriz $A$ multiplicando por un número real $r$ una de sus filas se tiene que $|B|=r|A|$.

*Demostración.* En la sección anterior se demostró que el resultado es cierto para matrices de orden 2 y de orden 3. Aceptemos la hipótesis de inducción y supongamos, para simplificar la demostración que

$$B=\begin{pmatrix} ra_{11} & ra_{12} & \cdots & ra_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Tenemos que

$$|B|=ra_{11}|B_{11}|-a_{21}|B_{21}|+\cdots+(-1)^{n+1}a_{n1}|B_{n1}|$$

El adjunto $B_{11}$ coincide con $A_{11}$ y, por tanto, $|B_{11}|=|A_{11}|$; los adjuntos $B_{k1}$ con $k\neq 1$ tienen una fila que está multiplicada por $r$ y, por tanto, $|B_{k1}|=r|A_{k1}|$ debido a la hipótesis de inducción. Sustituyendo estos resultados en la igualdad anterior obtenemos

$$|B|=ra_{11}|A_{11}|-ra_{21}|A_{21}|+\cdots+(-1)^{n+1}ra_{n-1}|A_{n-1}|=r|A| \qquad \blacksquare$$

PROPIEDAD 4 (P.4)

> Si $B$ es la matriz que se obtiene de $A$ intercambiando dos de sus filas, se tiene que $|B|=-|A|$.

*Demostración.* Comenzaremos demostrando el resultado cuando las dos filas que se intercambian son consecutivas. Al igual que en las demostraciones anteriores utilizaremos el método de demostración por inducción. El resultado es cierto para matrices de orden 2 y de orden 3 puesto que se ha demostrado en la sección anterior. Sea, por tanto:

$$B=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i+1,1} & a_{i+1,2} & \cdots & a_{i+1,n} \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$



la matriz que se ha obtenido de $A$ intercambiando las filas $i$ e $i+1$. Tenemos que

$$|B|=a_{11}|B_{11}|-\cdots+(-1)^{i+1}a_{i+1,1}|B_{i1}|+(-1)^{i+2}a_{i1}|B_{i+1,1}|+\cdots+(-1)^{n+1}a_{n1}|B_{n1}|$$

Se tiene que

$$B_{i1}=\begin{pmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}=A_{i+1,1} \quad \text{y} \quad B_{i+1,1}=\begin{pmatrix} a_{12} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i+1,1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \end{pmatrix}=A_{i1}$$

Por otro lado, $B_{k1}$ con $k\neq i,\, i+1$, es una matriz de orden $n-1$ con dos filas consecutivas intercambiadas con respecto a $A_{k1}$; por la hipótesis de inducción tenemos que $|B_{k1}| = -|A_{k1}|$.

Sustituyendo estos resultados en la igualdad anterior se obtiene:

$$\begin{aligned} |B| &= -a_{11}|A_{11}|+\cdots+(-1)^{i+1}a_{i+1,1}|A_{i+1,1}|+(-1)^{i+2}a_{i1}|A_{i1}|+\cdots-(-1)^{n+1}a_{n1}|A_{n1}|= \\ &= -[a_{11}|A_{11}|-\cdots+(-1)^{i+1}a_{i1}|A_{i1}|+(-1)^{i+2}a_{i+1,1}|A_{i+1,1}|+\cdots+(-1)^{n+1}a_{n1}|A_{n1}|] \\ &= -|A| \end{aligned}$$

Hemos demostrado el resultado cuando se intercambian dos filas consecutivas. Supongamos ahora que se intercambian las filas $i$ y $j=i+k$ de una matriz. Sea

$$B=\begin{pmatrix} a_{11} & \cdots\cdots & a_{1n} \\ \vdots & & \vdots \\ a_{j1} & \cdots\cdots & a_{jn} \\ \vdots & & \vdots \\ a_{i1} & \cdots\cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots\cdots & a_{nn} \end{pmatrix}$$

la matriz que se obtiene de $A$ intercambiando las filas $i$ y $j$. La matriz $B$ puede obtenerse a partir de $A$ realizando varios intercambios de filas consecutivas. En primer lugar, la fila $i$ de $A$ puede colocarse en la fila $j=i+k$ realizando $k$ intercambios de filas consecutivas; se obtiene así la matriz

$$\bar{A}=\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i+1,1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{j1} & \cdots & a_{jn} \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{array}{l} \\ \\ \longleftarrow \textit{Fila } i \\ \\ \\ \longleftarrow \text{Fila } j=i+k \\ \\ \\ \end{array}$$

Para pasar de $\bar{A}$ a $B$ mediante intercambios de filas consecutivas es necesario intercambiar la fila $(a_{j1}, ..., a_{jn})$ que ocupa el lugar $j-1=i+(k-1)$, $k-1$ veces para colocarla en la fila $i$. En definitiva, de $A$ se pasa a $B$ mediante $k+(k-1)=2k-1$ intercambios de filas consecutivas. Cada uno de estos intercambios cambia de signo el determinante y puesto que se hacen un número impar de intercambios se tiene que

$$|B|=(-1)^{2k-1}|A|=-|A| \qquad \blacksquare$$

**PROPIEDAD 5 (P.5)**

> Si una matriz tiene dos filas iguales, su determinante es nulo.

*Demostración.* Sea

$$A=\begin{array}{cl} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \longleftarrow \text{Fila } i \\ \\ \longleftarrow \text{Fila } j \\ \\ \\ \end{array} \end{array}$$

una matriz con dos filas iguales, a saber, las filas $i$ y $j$; intercambiando estas filas se obtiene la misma matriz; sin embargo, por la propiedad 4, el determinante cambia de signo. Por tanto, $|A|=-|A|$; de aquí se deduce que $|A|=0$, que era lo que queríamos demostrar. $\blacksquare$

EJEMPLO B. La matriz

$$A=\begin{pmatrix} 2 & 1 & 3 & 5 & 1 \\ 0 & 1 & -1 & 0 & 2 \\ 3 & 1 & -3 & 1 & 1 \\ 0 & 1 & -1 & 0 & 2 \\ -1 & 5 & 2 & 7 & 0 \end{pmatrix}$$

tiene determinante nulo ya que su segunda y su cuarta fila coinciden.

**PROPIEDAD 6 (P.6)**

> Si $B$ es la matriz que se obtiene a partir de $A$ sumando un múltiplo de una fila de $A$ a otra fila de $A$ se tiene que $|B|=|A|$.



*Demostración.* Multipliquemos la fila $i$ de $A$ por $r$ y sumémosla a la fila $k$ para obtener $B$; por las propiedades 2 y 3:

$$|B|=|A|+r\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{in} \\ \vdots & & \vdots \end{vmatrix} \rangle \text{iguales}$$

La matriz cuyo determinante aparece el último en la igualdad anterior tiene dos filas iguales y, por tanto, es nulo por la propiedad 5. Este razonamiento demuestra el resultado. $\blacksquare$

Las propiedades 3 y 6 de los determinantes junto con

$$(*) \qquad |I_n|=\begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix}=1$$

permiten calcular determinantes de una matriz cuadrada de cualquier orden. El siguiente ejemplo muestra este hecho.

En este ejemplo y en otros de cálculo de determinantes en los que se considere necesario, se utilizarán los siguientes símbolos aclaratorios de las igualdades que se escriben:

1) $rF_j$ significa que la $j$-ésima fila de la matriz ha sido multiplicada por el número real $r$.
2) $F_j+rF_k$ significa que a la $j$-ésima fila de la matriz se le ha sumado $r$ veces la fila $k$.
3) La propiedad que utilicemos en cada caso se indicará bajo el signo de igualdad.

EJEMPLO B

$$\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 4 & 2 & 0 & 3 \\ 2 & 2 & 1 & 1 \end{vmatrix} \underset{\text{P.6}}{\overset{\substack{F_3-4F_1 \\ F_4-2F_1}}{=\!=}} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 0 & -6 & -12 & -13 \\ 0 & -2 & -5 & -7 \end{vmatrix} \underset{\text{P.6}}{\overset{\substack{F_3+3F_2 \\ F_4+F_2}}{=\!=}}$$

$$\underset{\text{P.6}}{\overset{F_4-\frac{1}{2}F_3}{=\!=}} \begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & -6 & -4 \\ 0 & 0 & 0 & -2 \end{vmatrix} \underset{\text{P.3}}{=\!=} 2\cdot(-6)(-2)\begin{vmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 1 & 3/2 \\ 0 & 0 & 1 & 2/3 \\ 0 & 0 & 0 & 1 \end{vmatrix}$$

$$\xlongequal[\text{Aplicar varias veces P.6}]{} 2\cdot(-6)(-2)\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} = 24 \ (*)$$

Observar que el resultado es el mismo que el obtenido en el ejemplo A.

* * *

*Nota.* En el ejemplo anterior solamente se han necesitado P.3, P.6 y (*) para calcular el determinante. Este es un hecho general, es decir, las propiedades P.3, P.6 y (*) determinan de manera única un número al cual se le denomina «determinante» y que satisface todas las propiedades y fórmulas que daremos aquí. El lector interesado puede encontrar una exposición de este tipo en el libro de A. G. Kurosch, *Curso de álgebra superior*, Editorial Mir, 1977.

En el ejemplo B puede observarse que cuando tenemos que calcular el determinante de una matriz como

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & 2 & 3 \\ 0 & 0 & -6 & -4 \\ 0 & 0 & 0 & -2 \end{pmatrix}$$

el resultado se obtiene multiplicando los elementos de su diagonal principal, es decir, la diagonal de izquierda a derecha y de arriba a abajo.

Este es un resultado general. Una matriz se dice *triangular superior* si todos sus elementos por debajo de su diagonal principal son nulos.

PROPIEDAD 7 (P.7)

El determinante de una matriz triangular superior es el producto de los elementos de su diagonal.

*Demostración.* Sea

$$A=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & & & \vdots \\ 0 & 0 & & \cdots & a_{nn} \end{pmatrix}$$



una matriz triangular superior. De la definición de determinante aplicada repetidas veces se tiene:

$$|A|=a_{11}\begin{vmatrix} a_{22} & a_{23} & \cdots & a_{2n} \\ 0 & a_{33} & \cdots & a_{3n} \\ \vdots & & & \vdots \\ 0 & & & a_{nn} \end{vmatrix}=a_{11}a_{22}\begin{vmatrix} a_{33} & \cdots & a_{3n} \\ \vdots & & \vdots \\ 0 & \cdots & a_{nn} \end{vmatrix}=\cdots=a_{11}a_{22}\cdots a_{nn} \qquad \blacksquare$$

De la propiedad 7 se deduce que el determinante de una matriz diagonal, es decir, una matriz cuyos elementos son todos nulos excepto los de su diagonal principal, es el producto de los elementos de la diagonal:

$$\begin{vmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ 0 & a_{22} & 0 & \cdots & 0 \\ 0 & 0 & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & a_{nn} \end{vmatrix}=a_{11}a_{22}a_{33}\cdots a_{nn}$$

Es conveniente que el ejercicio siguiente se realice en este momento y sin conocer más propiedades de los determinantes, puesto que más adelante resultará muy sencillo.

*Ejercicio* 1. Una matriz *triangular inferior* es una matriz que tiene todos sus términos nulos por encima de la diagonal principal. Demostrar que el determinante de una matriz triangular inferior es el producto de los elementos de la diagonal principal.

EJEMPLO C. Para calcular el determinante de la matriz

$$A=\begin{pmatrix} 2 & 3 & 5 & 2 & 0 \\ 0 & 1 & 2 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 4 & 1 & -3 & 0 & -4 \\ 6 & -1 & 1 & 2 & 0 \end{pmatrix}$$

la reducimos a una matriz triangular superior mediante operaciones elementales en sus filas y aplicamos las propiedades ya conocidas de los determinantes. Tenemos:

$$|A| \xlongequal[\text{P.6}]{\substack{F_5-2F_1 \\ F_6-2F_3}} \begin{vmatrix} 2 & 3 & 5 & 2 & 0 \\ 0 & 1 & 2 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 0 & -5 & -13 & -4 & -4 \\ 0 & -1 & -1 & 2 & -2 \end{vmatrix} \xlongequal[\text{P.6}]{F_1-F_3} \begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 3 & 0 & 1 & 0 & 1 \\ 0 & -5 & -13 & -4 & -4 \\ 0 & -1 & -1 & 2 & -2 \end{vmatrix}$$

$$\underset{P.6}{\overset{F_3+3F_1}{=\!=\!=\!=}}\begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 9 & 13 & 6 & -2 \\ 0 & -5 & -13 & -4 & -4 \\ 0 & -1 & -1 & 2 & -2 \end{vmatrix} \underset{P.6}{\overset{\substack{F_3-9F_2\\ F_4+5F_2\\ F_5+F_2}}{=\!=\!=\!=}} \begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & -5 & 6 & -11 \\ 0 & 0 & -3 & -4 & 1 \\ 0 & 0 & 1 & 2 & -1 \end{vmatrix}$$

$$\underset{P.6}{\overset{\substack{F_3+5F_5\\ F_4+3F_5}}{=\!=\!=\!=}}\begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 16 & -16 \\ 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 1 & 2 & -1 \end{vmatrix} \underset{P.4}{\overset{F_3\leftrightarrow F_5}{=\!=\!=\!=}} - \begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 16 & -16 \end{vmatrix}$$

$$\underset{P.6}{\overset{F_5-8F_4}{=\!=\!=\!=}} - \begin{vmatrix} -1 & 3 & 4 & 2 & -1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & -1 \\ 0 & 0 & 0 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{vmatrix} \underset{P.7}{=\!=\!=\!=} -(-1)\cdot 1\cdot 1\cdot 2\cdot 0=0$$

* * *

En la definición dada de determinante se ha utilizado la primera columna de la matriz y sus adjuntos. Nuestro próximo objetivo es demostrar que también puede calcularse el determinante mediante una fórmula que incluya cualquier otra columna o incluso cualquier fila de la matriz y sus adjuntos. Estos resultados serán de gran utilidad para simplificar el cálculo de determinantes.

Comenzaremos demostrando que puede utilizarse la primera fila de la matriz para calcular determinantes.

TEOREMA 1

Si

$$A=\begin{pmatrix} 0 & \cdots & a_{1j} & \cdots & 0 \\ a_{21} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & & \vdots & & \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{pmatrix}$$

es una matriz con todos los elementos de su primera fila nulos excepto el que ocupa el lugar $j$, se tiene que

$$|A| \equiv (-1)^{j+1}a_{1j}|A_{1j}|$$



*Demostración.* Procedemos por inducción en el orden de la matriz. Si $A$ es una matriz de orden 2 puede ser de la forma

$$A=\begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix} \quad \text{o} \quad B=\begin{pmatrix} 0 & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

En el primer caso se tiene $|A|=a_{11}a_{22}=a_{11}|A_{11}|$ y en el segundo $|B|=-a_{12}a_{21}=-a_{12}|A_{12}|$, lo cual prueba que la fórmula es cierta en este caso.

Para que el lector pueda seguir mejor el caso general hacemos un caso particular de una matriz de orden 3. Sea

$$A=\begin{pmatrix} 0 & a_{12} & 0 \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Tenemos que

$$|A|=-a_{21}\begin{vmatrix} a_{12} & 0 \\ a_{32} & a_{33} \end{vmatrix}+a_{31}\begin{vmatrix} a_{12} & 0 \\ a_{22} & a_{23} \end{vmatrix}=-a_{21}a_{12}a_{33}+a_{31}a_{12}a_{23}$$

$$=-a_{12}[a_{21}a_{33}-a_{31}a_{23}]=-a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$$

que era lo que queríamos demostrar.

Aceptemos ahora que el resultado es cierto para toda matriz de orden inferior a $n$ y sea $A$ una matriz de orden $n$. Si el elemento no nulo ocupa el primer lugar de la primera fila se tiene

$$A=\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

De la definición de determinante deducimos que

$$|A|=a_{11}|A_{11}|-a_{21}|A_{21}|+\cdots+(-1)^{n+1}a_{n1}|A_{n1}|=a_{11}|A_{11}|$$

ya que cada adjunto $A_{j1}$, con $j>1$, tiene una fila de ceros y, por tanto, su determinante es nulo (propiedad 1).

Supongamos ahora que el elemento no nulo de la primera fila ocupa el lugar $j$, con $j>1$. Tenemos

$$A=\begin{pmatrix} 0 & 0 & \cdots & 0 & a_{1j} & 0 & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2j} & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}$$

De la definición de determinante deducimos que

$$|A| = -a_{21}|A_{21}| + a_{31}|A_{31}| - \cdots + (-1)^{n+1}a_{n1}|A_{n1}|$$

El adjunto $A_{i1}$, $i > 1$, de la fórmula anterior es de la forma

$$B^{(i)} = A_{i1} = \begin{pmatrix} 0 & \cdots & a_{1j} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{i-1,2} & \cdots & a_{i-1,j} & \cdots & a_{i+1,n} \\ a_{i+1,2} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n2} & \cdots & a_{n,j} & \cdots & a_{n,n} \end{pmatrix}$$

y de orden $n-1$; por la hipótesis de inducción su determinante es

$$|A_{i1}| = (-1)^j a_{1j}|B^{(i)}_{1,j-1}|$$

donde $B^{(i)}_{1,j-1}$ denota la matriz que se obtiene de $B^{(i)} = A_{i1}$ suprimiendo la primera fila y la $(j-1)$-ésima columna. Por tanto:

$$\begin{aligned} |A| &= -a_{21}(-1)^j a_{1j}|B^{(2)}_{1,j-1}| + a_{31}(-1)^j a_{1j}|B^{(3)}_{1,j-1}| + \cdots\cdots + (-1)^{n+1}a_{n1}(-1)^j a_{1j}|B^{(n)}_{1,j-1}| = \\ &= (-1)^{j+1}a_{1j}[a_{21}|B^{(2)}_{1,j-1}| - a_{31}|B^{(3)}_{1,j-1}| + \cdots + (-1)^n a_{n1}|B^{(n)}_{1,j-1}|] \end{aligned}$$

Solamente falta probar que la expresión entre paréntesis coincide con $|A_{1j}|$. Puesto que

$$B^{(j)} = A_{1j} = \begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2n} \\ a_{31} & a_{32} & \cdots & a_{3,j-1} & a_{3,j+1} & \cdots & a_{3n} \\ \vdots & \vdots & & & & & \\ a_{n1} & a_{n2} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}_{(n-1)}$$

se tiene que

$$|A_{1j}| = a_{21}|B^{(j)}_{11}| - a_{31}|B^{(j)}_{21}| + \cdots + (-1)^n a_{n1}|B^{(j)}_{n-1,1}|$$

En $B^{(j)}_{11}$ se han suprimido la primera fila y la primera columna de $A_{1j}$ y, por tanto, se han suprimido las dos primeras filas de $A$ y las columnas primera y $j$-ésima de $A$; de la misma manera, en $B^{(2)}_{1,j-1}$ se han suprimido la primera fila y la $j-1$ columna de $A_{21}$ y, por tanto, se han suprimido las dos primeras filas y la primera fila y la $j$-ésima columna de $A$. Por tanto:



$$B^{(j)}_{11} = B^{(2)}_{1,j-1}$$

De manera similar se demuestra que $B^{(j)}_{21} = B^{(3)}_{1,j-1}$, etc., lo cual demuestra el resultado. ■

EJEMPLO D. El determinante de la matriz

$$A = \begin{pmatrix} 0 & 0 & -3 & 0 \\ 1 & 2 & 1 & -4 \\ 3 & 1 & 0 & -1 \\ -1 & -5 & 2 & 0 \end{pmatrix}$$

es

$$|A| = (-1)^{3+1}(-3)\begin{vmatrix} 1 & 2 & -4 \\ 3 & 1 & -1 \\ -1 & -5 & 0 \end{vmatrix} \underset{\text{P.6}}{=\!=} -3\begin{vmatrix} 1 & 2 & 4 \\ 0 & -5 & 11 \\ 0 & -3 & -4 \end{vmatrix} = -3\cdot 1\begin{vmatrix} -5 & 11 \\ -3 & -4 \end{vmatrix} =$$
$$= (-3)[20+33] = -159$$

TEOREMA 2 (*Desarrollo por la primera fila*)

Si

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

se tiene que

$$|A| = a_{11}|A_{11}| - a_{12}|A_{12}| + \cdots + (-1)^{n+1}a_{1n}|A_{1n}|$$

*Demostración.* Escribimos

$$A = \begin{pmatrix} a_{11}+0+\cdots+0 & 0+a_{12}+\cdots+0 & \cdots & 0+0+\cdots+a_{1n} \\ a_{21} & a_{22} & & \vdots \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

Si $A_j$, $j=1, 2, ..., n$, es la matriz que tiene todas sus filas iguales a $A$, excepto la primera, en la que todos sus elementos son nulos excepto el que ocupa el lugar $j$ que es $a_{1j}$, de la propiedad 2 de los determinantes aplicada reiteradas veces se tiene

$$|A|=|A_1|+|A_2|+\cdots+|A_n|$$

Por el teorema 1, $|A_j|=(-1)^{j+1}a_{1j}|A_{1j}|$, lo cual prueba el resultado deseado. ■

EJEMPLO E. Para calcular el determinante de la matriz

$$A=\begin{pmatrix} 2 & 1 & -1 \\ 3 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix}$$

utilizamos el teorema 2 para obtener

$$|A|=2\begin{vmatrix} 0 & 1 \\ 2 & 1 \end{vmatrix}-1\begin{vmatrix} 3 & 1 \\ -1 & 1 \end{vmatrix}+(-1)\begin{vmatrix} 3 & 0 \\ -1 & 2 \end{vmatrix}=2\cdot(-2)-1(3+1)-1\cdot 6=$$
$$=-4-4-6=-14$$

* * *

Sea $B$ la matriz que se obtiene de $A=(a_{ij})_{\substack{i=1,...,n \\ j=1,...,n}}$ intercambiando las filas primera y segunda.

Por la propiedad 5 tenemos que $|A|=-|B|$ y puesto que

$$B=\begin{pmatrix} a_{21} & a_{22} & \cdots & a_{2n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

utilizando el teorema 1 para desarrollar el determinante de $B$ por su primera fila se tiene que

$$|A|=-|B|=-[a_{21}|B_{11}|-a_{22}|B_{12}|+\cdots+(-1)^{n+1}a_{2n}|B_{1n}|]$$

Puesto que $B_{1j}=A_{2j}$ para todo $j=1, 2, ..., n$ tenemos que

$$|A|=-a_{21}|A_{21}|+a_{22}|A_{22}|+\cdots-(-1)^{n+1}a_{2n}|A_{2n}|$$

lo cual nos da una fórmula para desarrollar el determinante de $A$ por la segunda fila. La principal diferencia con el teorema 1 es que el primer término es negativo.

Si ahora intercambiamos la segunda y la tercera fila, el determinante cambia de signo por la propiedad 5 y, por tanto, se tendría una fórmula para desarrollar el determinante por la tercera fila, comenzando con un sumando positivo.

Este razonamiento demuestra el siguiente resultado.



TEOREMA 3 (*Desarrollo por la i-ésima fila*)

Si $A=(a_{ij})_{i,j=1,2,...,n}$, se tiene que

$$|A|=(-1)^{i+1}a_{i1}|A_{i1}|+(-1)^{i+2}a_{i2}|A_{i2}|+\cdots+(-1)^{i+n}a_{in}|A_{in}|$$

No es de extrañar que exista un resultado análogo para calcular el determinante utilizando un desarrollo por cualquiera de las columnas. Ello requiere la definición de matriz *traspuesta* de una dada.

Si $A$ es una matriz de orden $m \times n$, se denomina matriz *traspuesta* de $A$, y se denota por $A^t$, a la matriz de orden $n \times m$ que se obtiene poniendo como $i$-ésima fila a la $i$-ésima columna de la matriz $A$.

Por ejemplo, la matriz traspuesta de

$$A=\begin{pmatrix} 1 & 2 & -2 \\ 3 & 4 & 1 \end{pmatrix}$$

es

$$A^t=\begin{pmatrix} 1 & 3 \\ 2 & 4 \\ -2 & 1 \end{pmatrix}$$

El resultado principal que relaciona la operación de trasposición con el producto de matrices es el siguiente:

$$(AB)^t=B^tA^t$$

siempre que las multiplicaciones anteriores tengan sentido (ver ejercicio 6 al final de esta sección).

TEOREMA 4

Si $A$ es una matriz cuadrada, $|A|=|A^t|$.

*Demostración.* El resultado es cierto para matrices de orden 2 puesto que

$$|A|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{12}a_{21}$$

y

$$|A^t|=\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}=a_{11}a_{22}-a_{21}a_{12}$$

Supongamos que el resultado es cierto para matrices de orden inferior a $n$ y sea $A = (a_{ij})_{i,j=1,2,\ldots,n}$ una matriz de orden $n$. Puesto que

$$B = A^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix}$$

por el teorema 2 se tiene que

$$|A^t| = a_{11}|B_{11}| - a_{21}|B_{12}| + \cdots + (-1)^{n+1}a_{n1}|B_{1n}|$$

Puesto que $B_{1j} = A^t_{j1}$, $j = 1, 2, \ldots, n$, utilizando la hipótesis de inducción se tiene que

$$|A^t| = a_{11}|A_{11}| - a_{21}|A_{21}| + \cdots + (-1)^{n+1}a_{n1}|A_{n1}| = |A| \qquad \blacksquare$$

Los teoremas 3 y 4 producen el siguiente resultado:

TEOREMA 5 (*Desarrollo por la j-ésima columna*)

Si $A = (a_{ij})_{i,j=1,2,\ldots,n}$, se tiene que

$$|A| = (-1)^{j+1}a_{1j}|A_{1j}| + (-1)^{j+2}a_{2j}|A_{2j}| + \cdots + (-1)^{j+n}a_{nj}|A_{nj}|$$

EJEMPLO F. Para calcular el determinante de la matriz

$$A = \begin{pmatrix} 1 & 3 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 3 & 0 & 1 & 4 \\ 4 & 0 & 2 & 1 \end{pmatrix}$$

desarrollamos por la segunda columna para obtener

$$|A| = \begin{vmatrix} 1 & 3 & 5 & 2 \\ 0 & 0 & 3 & 0 \\ 3 & 0 & 1 & 4 \\ 4 & 0 & 2 & 1 \end{vmatrix} = -3\begin{vmatrix} 0 & 3 & 0 \\ 3 & 1 & 4 \\ 4 & 2 & 1 \end{vmatrix} = -3(-3)\begin{vmatrix} 3 & 4 \\ 4 & 1 \end{vmatrix} =$$

$$= -3(-3)\begin{vmatrix} 3 & 4 \\ 4 & 1 \end{vmatrix} = 9(-13) = -117$$

en donde el determinante de orden 3 se ha desarrollado por la primera fila.

* * *



El teorema 4 produce otro resultado interesante. En las propiedades 1 a 6 de los determinantes puede sustituirse la palabra «fila» por «columna» y los resultados siguen siendo válidos. Esto permite realizar operaciones elementales con columnas para calcular determinantes.

La verificación de estos resultados se deja para el lector.

En el ejemplo anterior se observa que el cálculo de determinantes se simplifica si desarrollamos por una fila o columna que presente el mayor número de ceros. Si una matriz no posee ningún elemento nulo es posible realizar operaciones elementales con sus filas o columnas, de manera que el determinante no cambie, pero que la nueva matriz posea varios ceros en alguna de sus filas o columnas.

El método combinado de realizar operaciones elementales y utilizar las fórmulas anteriores es el más utilizado para calcular determinantes, debido a que simplifica el número de operaciones a realizar.

EJEMPLO G. Sea

$$A = \begin{pmatrix} -1 & 4 & 3 & -2 \\ -3 & 1 & 0 & 5 \\ 2 & 1 & -1 & 3 \\ 2 & 1 & 1 & 1 \end{pmatrix}$$

Multiplicando por 2 la segunda columna y restándosela a la primera se tiene

$$|A| \underset{\text{P.6}}{\overset{C_1 - 2C_2}{=\!=\!=}} \begin{vmatrix} -9 & 4 & 3 & -2 \\ -5 & 1 & 0 & 5 \\ 0 & 1 & -1 & 3 \\ 0 & 1 & 1 & 1 \end{vmatrix}$$

Desarrollando por la primera columna se tiene

$$|A| = -9\begin{vmatrix} 1 & 0 & 5 \\ 1 & -1 & 3 \\ 1 & 1 & 1 \end{vmatrix} + 5\begin{vmatrix} 4 & 3 & -2 \\ 1 & -1 & 3 \\ 1 & 1 & 1 \end{vmatrix}$$

Para calcular el primero de estos determinantes restamos la primera fila de la segunda y de la tercera para obtener

$$\begin{vmatrix} 1 & 0 & 5 \\ 1 & -1 & 3 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 5 \\ 0 & -1 & -2 \\ 0 & 1 & -4 \end{vmatrix} = \begin{vmatrix} -1 & -2 \\ 1 & -4 \end{vmatrix} = 4 + 2 = 6$$

Para calcular el segundo restamos la primera columna de la segunda y la tercera para obtener

$$\begin{vmatrix} 4 & 3 & -2 \\ 1 & -1 & 3 \\ 1 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 4 & -1 & -6 \\ 1 & -2 & 2 \\ 1 & 0 & 0 \end{vmatrix} = \begin{vmatrix} -1 & -6 \\ -2 & 2 \end{vmatrix} = -2-12=-14$$

Por tanto:

$$|A| = -9\cdot 6+5\cdot(-14) = -54-70 = -124$$

* * *

Es conveniente escribir una matriz con los signos que se asocian a cada elemento en el desarrollo de los teoremas 3 y 5 con motivo de ayudar al lector a recordarlos. Esta matriz es

$$\begin{pmatrix} + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ + & - & + & - & + & \cdots \\ - & + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

EJEMPLO H. Tratemos de calcular el determinante de la matriz

$$A = \begin{pmatrix} x & a & a \\ a & x & a \\ a & a & x \end{pmatrix}$$

Las operaciones elementales que se realizan se indican adecuadamente encima del signo de igualdad.

$$|A| = \begin{vmatrix} x & a & a \\ a & x & a \\ a & a & x \end{vmatrix} \overset{F_3-F_2}{=\!=\!=} \begin{vmatrix} x & a & a \\ a & x & a \\ 0 & a-x & x-a \end{vmatrix} = x\begin{vmatrix} x & a \\ a-x & x-a \end{vmatrix} - a\begin{vmatrix} a & a \\ a-x & x-a \end{vmatrix} =$$

$$= x[x(x-a)-a(a-x)]-a[a(x-a)-a(a-x)] =$$

$$= (x-a)[x^2+xa-a^2-a^2] = (x-a)(x-a)(x+2a)$$



EJERCICIOS 2.2

**1.** Calcular los siguientes determinantes mediante su desarrollo por la primera columna:

*a)* $\begin{vmatrix} 4 & 2 & 1 \\ 7 & 6 & 5 \\ 1 & 2 & 3 \end{vmatrix}$ *b)* $\begin{vmatrix} 3 & 0 & 1 \\ 4 & 3 & 2 \\ 1 & 2 & 1 \end{vmatrix}$ *c)* $\begin{vmatrix} 1 & 0 & 0 & 7 \\ 0 & 1 & 2 & 3 \\ 2 & 1 & 1 & 1 \\ 0 & 3 & 2 & 1 \end{vmatrix}$

**2.** Calcular los siguientes determinantes reduciéndolos a una matriz triangular superior mediante operaciones elementales:

*a)* $\begin{vmatrix} 2 & 4 & -2 \\ 3 & 5 & 2 \\ 2 & 5 & 1 \end{vmatrix}$ *b)* $\begin{vmatrix} 3 & 6 & -3 & 9 \\ 2 & 1 & 0 & 4 \\ 1 & 3 & 2 & 5 \\ 0 & 1 & 3 & 2 \end{vmatrix}$ *c)* $\begin{vmatrix} 1 & 0 & 3 & 5 \\ 2 & 1 & 2 & 7 \\ 2 & 2 & 0 & 4 \\ 3 & 2 & 1 & 2 \end{vmatrix}$ *d)* $\begin{vmatrix} 2 & 0 & -2 & 4 \\ 0 & 3 & 9 & -6 \\ 1 & 2 & -1 & 0 \\ 0 & 3 & 4 & 2 \end{vmatrix}$

**3.** Calcular los siguientes determinantes usando sus propiedades y efectuando un número reducido de computaciones.

*a)* $\begin{vmatrix} 1 & 0 & 2 \\ 3 & 4 & 6 \\ 5 & 7 & 10 \end{vmatrix}$ *b)* $\begin{vmatrix} 1 & 2 & 2 & 1 \\ 0 & 4 & 2 & 1 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 1 & 1 \end{vmatrix}$ *c)* $\begin{vmatrix} 2 & 1 & 3 & 2 \\ 1 & 0 & 0 & 1 \\ 4 & -1 & 2 & 1 \\ 3 & -1 & 0 & 1 \end{vmatrix}$ *d)* $\begin{vmatrix} 2 & 0 & 0 & 3 & 5 \\ 1 & -1 & 2 & 1 & 2 \\ 0 & 1 & 3 & -1 & 1 \\ 4 & 1 & 1 & -1 & 0 \\ -2 & 3 & 1 & -3 & 0 \end{vmatrix}$

*e)* $\begin{vmatrix} 3 & 0 & 0 & 3 \\ 0 & 3 & 0 & 3 \\ 0 & 0 & 3 & 3 \\ 3 & 3 & 0 & 0 \end{vmatrix}$ *f)* $\begin{vmatrix} 1-\lambda & 2 & 1 \\ 0 & 3-\lambda & 1 \\ 3 & 1 & 1-\lambda \end{vmatrix}$ *g)* $\begin{vmatrix} 1 & 1 & 0 & 1 \\ 2 & x & 0 & 1 \\ 0 & 1 & y & 1 \\ 0 & 1 & 1 & 1 \end{vmatrix}$

**4.** Demostrar, sin necesidad de calcularlos, que los siguientes determinantes son nulos:

*a)* $\begin{vmatrix} 1 & 2 & 3 \\ -1 & -2 & -3 \\ 3 & 1 & -4 \end{vmatrix}$ *b)* $\begin{vmatrix} x & y & 2x+3y \\ 4 & 3 & 17 \\ z & t & 2z+3t \end{vmatrix}$ *c)* $\begin{vmatrix} \operatorname{sen}^2 a & 1 & \cos^2 a \\ \operatorname{sen}^2 b & 1 & \cos^2 b \\ \operatorname{sen}^2 c & 1 & \cos^2 c \end{vmatrix}$

*d)* $\begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \end{vmatrix}$ *e)* $\begin{vmatrix} 1^2 & 2^2 & 3^2 & 4^2 \\ 2^2 & 3^2 & 4^2 & 5^2 \\ 3^2 & 4^2 & 5^2 & 6^2 \\ 4^2 & 5^2 & 6^2 & 7^2 \end{vmatrix}$

**5.** Sabiendo que los números 58.786, 30.628, 12.831, 80.743 y 16.016 son todos divisibles por 13, demostrar que el determinante de la siguiente matriz es también divisible por 13:

$$A=\begin{pmatrix} 5 & 8 & 7 & 8 & 6 \\ 3 & 0 & 6 & 2 & 8 \\ 1 & 2 & 8 & 3 & 1 \\ 8 & 0 & 7 & 4 & 3 \\ 1 & 6 & 0 & 1 & 6 \end{pmatrix}$$

**6.** Demostrar que $(AB)^t = B^t A^t$ siempre que las multiplicaciones anteriores tengan sentido.

**7.** *a)* Demostrar que si $A$ es una matriz cuadrada de orden $n$, y $r$ es un número real $|rA| = r^n|A|$. *b)* Utilizar *a)* para decir cómo cambia el determinante de una matriz de orden $n$ si los signos de todos sus elementos se cambian.

**8.** Una matriz cuadrada $A$ se dice *antisimétrica* si $A = -A^t$.

*a)* Dar un ejemplo de una matriz antisimétrica de orden 4.

*b)* Demostrar que si $A$ es antisimétrica y de orden impar su determinante es nulo. (*Sugerencia*: utilizar 7.*b*)).

*c)* ¿Es cierto el resultado de la parte *b)* si $A$ es de orden par?

**9.** Demostrar que $\begin{vmatrix} 1 & x & y \\ 1 & a_1 & b_1 \\ 1 & a_2 & b_2 \end{vmatrix} = 0$ es la ecuación de una recta en el plano que pasa por los puntos $A_1 = (a_1, b_1)$ y $A_2 = (a_2, b_2)$.

Los siguientes ejercicios se refieren al método de demostración por inducción.

**10.** Demostrar la fórmula $1+2+3+\cdots+n = \dfrac{n+1}{2}\,n$.

**11.** Demostrar la fórmula $1^2+2^2+3^2+\cdots+n^2 = \dfrac{n(n+1)(2n+1)}{6}$.

**12.** Un polígono se denomina *convexo* si las líneas que unen dos cualesquiera de sus vértices están en el interior del polígono. Demostrar, mediante el método de inducción, que la suma de los ángulos interiores de un polígono convexo de $n$ lados es $180(n-2)$.

**13.** Deducir una fórmula para calcular

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^n$$

y demostrarla mediante el método de inducción.

**14.** Si $A$ es una matriz triangular inferior de orden 3 cuyos elementos son números reales, todos los elementos de su diagonal principal son iguales y $A^2 = I$, demostrar que $A$ es $I$ ó $-I$.



## 2.3. DETERMINANTE DE UN PRODUCTO DE MATRICES. CALCULO DE DETERMINANTES DE ORDEN $n$

El primer resultado que queremos probar en esta sección es que el determinante de un producto de matrices cuadradas del mismo orden es el producto de los determinantes de cada una de ellas.

Para probar este resultado mostraremos primero que las operaciones elementales con las filas de una matriz pueden expresarse mediante multiplicación por matrices adecuadas.

Estudiaremos cada operación elemental por separado.

*a)* Si multiplicamos la fila $i$ de la matriz $A = (a_{ij})_{i,j=1,\dots,n}$ por un número real $c$ no nulo el resultado es el mismo que si multiplicamos la matriz $A$ por la izquierda por la matriz

$$E=\begin{array}{cl} \begin{pmatrix} 1 & & & & & \\ & \ddots & & & 0 & \\ & & 1 & & & \\ & & & c & & \\ & 0 & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \\ \leftarrow \text{Fila } i \\ \\ \\ \\ \end{array} \\ \text{Columna } i & \end{array}$$

En efecto:

$$EA=\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ ca_{i1} & ca_{i2} & \cdots & ca_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{array}{l} \\ \\ \leftarrow \text{Fila } i \\ \\ \\ \end{array}$$

*b)* El resultado de intercambiar las filas $i$ y $k$ de la matriz $A = (a_{ij})_{i,j=1,\dots,n}$ es el mismo que si multiplicamos la matriz $A$ por la izquierda por la matriz

$$E=\begin{array}{cl} \begin{pmatrix} 1 & & & & & 0 \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \\ & & \vdots & & \vdots & \\ & & 1 & \cdots & 0 & \\ & 0 & & & & \ddots \\ & & & & & & 1 \end{pmatrix} & \begin{array}{l} \\ \\ \leftarrow \text{Fila } i \\ \\ \leftarrow \text{Fila } k \\ \\ \\ \end{array} \\ \text{Columna } i \quad \text{Columna } k & \end{array}$$

En efecto:

$$EA=\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & 0 & \\ & & 0 & \cdots & 1 & & \\ & & \vdots & & \vdots & & \\ & & 1 & \cdots & 0 & & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}=\begin{array}{c} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \end{array}\begin{array}{l} \\ \\ \leftarrow \text{Fila } i \\ \\ \leftarrow \text{Fila } k \\ \\ \\ \end{array}$$

*c*) El resultado de multiplicar por $c$ la fila $i$ de la matriz $A=(a_{ij})_{i,j=1,2,\ldots,n}$ y sumárselo a la fila $k$ es el mismo que si multiplicamos la matriz $A$ por la izquierda por la matriz

$$E=\begin{array}{c} \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & 0 & \\ & & \vdots & \ddots & & \\ & & c & \cdots & 1 & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \\ \text{Columna } i \quad \text{Columna } k \end{array}\begin{array}{l} \\ \\ \leftarrow \text{Fila } i \\ \\ \leftarrow \text{Fila } k \\ \\ \\ \end{array}$$

En efecto:

$$EA=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & 0 & \\ & & \vdots & \ddots & & \\ & & c & \cdots & 1 & \\ & 0 & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}=$$

$$=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ ca_{i1}+a_{k1} & ca_{i2}+a_{k2} & \cdots & ca_{in}+a_{kn} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$



Las matrices $E$ que aparecen en cada uno de los casos anteriores reciben el nombre de *matrices elementales*. Observar que toda matriz elemental se obtiene a partir de la matriz identidad realizando la operación elemental correspondiente a su caso. Así, por ejemplo, la matriz elemental del caso *b*) se obtiene intercambiando las filas $i$ y $k$ de la matriz identidad.

Son ejemplos de matrices elementales las siguientes:

$$E=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad ; \quad E=\begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix} \quad ; \quad E=\begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Puesto que toda matriz puede reducirse a una matriz escalonada mediante una sucesión de operaciones elementales, se tiene el siguiente resultado:

RESULTADO 1

Toda matriz cuadrada $A$ puede reducirse mediante una sucesión de matrices elementales $E_1, E_2, \ldots, E_k$ a una matriz escalonada $B$ tal que

$$E_k \cdots E_2 E_1 A = B$$

EJEMPLO A. La matriz



$$A=\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$$

se transforma en la identidad mediante la siguiente sucesión de operaciones elementales:

$$\begin{pmatrix} 1 & 2 & 1 \\ 3 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow{F_2-3F_1} \begin{pmatrix} 1 & 2 & 1 \\ 0 & -5 & -1 \\ 0 & 1 & 1 \end{pmatrix} \xrightarrow[F_3]{F_2\leftrightarrow F_3} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & -5 & -1 \end{pmatrix} \xrightarrow{F_3+5F_2} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 4 \end{pmatrix} \xrightarrow{\frac{1}{4}F_3}$$

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{F_2-F_3} \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{F_1-F_3} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{F_1-2F_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Por tanto, podemos escribir la igualdad

$$\begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1/4\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 5 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}A=I$$

que el lector puede comprobar.

EJEMPLO B. La matriz

$$A=\begin{pmatrix}2 & 1\\ 4 & 2\end{pmatrix}$$

se transforma en una matriz escalonada mediante la siguiente sucesión de operaciones elementales:

$$\begin{pmatrix}2 & 1\\ 4 & 2\end{pmatrix}\xrightarrow{F_2-2F_1}\begin{pmatrix}2 & 1\\ 0 & 0\end{pmatrix}\xrightarrow{\frac{1}{2}F_1}\begin{pmatrix}1 & 1/2\\ 0 & 0\end{pmatrix}$$

Por tanto:

$$\begin{pmatrix}1/2 & 0\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix}A=\begin{pmatrix}1 & 1/2\\ 0 & 0\end{pmatrix}$$

* * *

PROPOSICIÓN 2

Toda matriz elemental posee una inversa, que es también una matriz elemental.

*Demostración.* Toda matriz elemental $E$ se obtiene a partir de la identidad mediante una operación elemental. Mediante la operación inversa de la realizada, que es también una operación elemental, $E$ se transforma en la identidad. ■

EJEMPLO C. La inversa de $E=\begin{pmatrix}c & 0\\ 0 & 1\end{pmatrix}$ es $\begin{pmatrix}1/c & 0\\ 0 & 1\end{pmatrix}$. La inversa de $E=\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}$ es ella misma. La inversa de $\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 5 & 1\end{pmatrix}$ es $\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -5 & 1\end{pmatrix}$. La comprobación se deja para el lector.

* * *



PROPOSICIÓN 3

Si $A$ es una matriz invertible, $A$ puede factorizarse como un producto de matrices elementales.

*Demostración.* Puesto que $A$ es invertible puede reducirse a la identidad mediante una sucesión de operaciones elementales; por el resultado 1 existen matrices elementales $E_1, E_2, \ldots, E_k$ tal que

$$E_k \cdots E_2 E_1 A = I$$

Por tanto, $A=E_1^{-1}E_2^{-1}\cdots E_k^{-1}$ y cada una de las matrices $E_j^{-1}$ es una matriz elemental debido a la proposición 2. ■

EJEMPLO D. Si $A$ es la matriz del ejemplo $A$ se tiene:

$$A=\begin{pmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 5 & 1\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1/4\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}^{-1}$$

$$\begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}^{-1}=\begin{pmatrix}1 & 0 & 0\\ 3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -5 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 4\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$$

* * *

TEOREMA 4

Si $A$ y $B$ son matrices cuadradas del mismo orden,

$$|AB|=|A||B|$$

*Demostración.* La demostración la realizaremos en tres etapas.

*Etapa* 1. Si $A$ es una matriz elemental $E$ el resultado es cierto. Para probarlo estudiamos cada uno de los posibles casos de matrices elementales.

En el caso *a*), $|EB|=c|B|$ por la propiedad 3 de la sección 2.2; además, $|E|=c$ y, por tanto, $|EB|=c|B|=|E||B|$.

En el caso *b*), $|EB|=-|B|$ por la propiedad 4 de la sección 2.2; además, $|E|=-1$ por la misma propiedad y, por tanto, $|EB|=|E||B|$.

En el caso *c*), $|EB|=|B|$ por la propiedad 6 de los determinantes; además, $|E|=1$ por la misma propiedad y, por tanto, $|EB|=|E||B|$ también en este caso.

*Etapa* 2. Si $A$ es invertible, por la proposición 3 podemos escribir $A=E_1E_2\cdots E_k$, donde cada $E_j$ es una matriz elemental. Aplicando repetidas veces el resultado obtenido en la etapa 1 se tiene:

$$|AB|=|E_1E_2\cdots E_kB|=|E_1||E_2\cdots E_kB|=\cdots=|E_1||E_2|\cdots|E_k||B|=$$
$$=|E_1E_2\cdots E_k||B|=|A||B|$$

*Etapa* 3. Si $A$ no es invertible, por el teorema 3 de la sección 1.4, el rango de $A$ es inferior a $n$ y, por tanto, $|A|=0$. En este caso también $|AB|=0$ ya que $A$ puede escribirse de la forma $A = E_1'E_2' \dots E_k'A'$, donde las $E_j'$ son matrices elementales y $A'$ tiene una fila de ceros. Por tanto, $|AB| = |E_1'E_2' \dots E_k'A'B| = |E_1'||E_2'| \dots |E_k'||A'B| = 0$, ya que $A'B$ tiene también una fila de ceros. ∎

EJEMPLO E. El teorema sobre el determinante de un producto es útil para calcular algunos determinantes. El determinante de la matriz

$$A=\begin{pmatrix} b^2+c^2 & ab & ca \\ ab & c^2+a^2 & bc \\ ca & bc & a^2+b^2 \end{pmatrix}$$

puede calcularse fácilmente escribiéndola como un producto de dos matrices de orden 3. En efecto:

$$A=\begin{pmatrix} b & c & 0 \\ a & 0 & c \\ 0 & a & b \end{pmatrix}\begin{pmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{pmatrix}$$

y puesto que

$$\begin{vmatrix} b & c & 0 \\ a & 0 & c \\ 0 & a & b \end{vmatrix} = -2abc \quad \text{y} \quad \begin{vmatrix} b & a & 0 \\ c & 0 & a \\ 0 & c & b \end{vmatrix} = -2abc$$

se tiene que

$$|A| = 4a^2b^2c^2$$

* * *

La segunda parte de esta sección está dedicada a dar algunas sugerencias relativas al cálculo de determinantes de matrices de orden $n$ y de matrices de orden elevado que pueden reducirse al cálculo de determinantes de matrices de orden inferior.

Comenzamos con el determinante de la matriz de orden $n$:

$$A_n=\begin{pmatrix} x & a & a & \cdots & a \\ a & x & a & \cdots & a \\ a & a & x & \cdots & a \\ \vdots & \vdots & \vdots & & \vdots \\ a & a & a & \cdots & x \end{pmatrix}$$

El desarrollo de $|A_n|$ por una cualquiera de sus filas o columnas no conduce a ningún resultado positivo. Lo más adecuado es realizar operaciones elementales con sus filas o



columnas y utilizar las propiedades de los determinantes estudiadas en la sección anterior, hasta convertirla en una matriz triangular.

Restando la primera fila de cada una de las restantes se tiene:

$$|A_n|=\begin{vmatrix} x & a & a & \cdots & a \\ a-x & x-a & 0 & \cdots & 0 \\ a-x & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ a-x & 0 & 0 & \cdots & x-a \end{vmatrix}$$

Sumando a la primera columna todas las restantes se tiene:

$$|A_n|=\begin{vmatrix} x+(n-1)a & a & a & \cdots & a \\ 0 & x-a & 0 & \cdots & 0 \\ 0 & 0 & x-a & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & x-a \end{vmatrix} = (x-a)^{n-1}[x+(n-1)a]$$

*Nota.* Ver el ejemplo H de la sección anterior, donde se ha calculado este determinante para $n=3$.

* * *

La demostración por inducción también puede aplicarse para calcular determinantes de matrices de orden $n$. Expondremos este método calculando el determinante de la matriz

$$A_n=\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{pmatrix}$$

que se conoce en la literatura matemática con el nombre de *determinante de Vandermonde.*

Multiplicando cada fila por $x_1$ y restándosela a la siguiente se tiene:

$$V_n(x_1, \dots, x_n)=|A_n|=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 0 & x_2-x_1 & x_3-x_1 & \cdots & x_n-x_1 \\ 0 & x_2(x_2-x_1) & x_3(x_3-x_1) & \cdots & x_n(x_n-x_1) \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & x_2^{n-2}(x_2-x_1) & x_3^{n-2}(x_3-x_1) & \cdots & x_n^{n-2}(x_n-x_1) \end{vmatrix} =$$

$$=(x_2-x_1)\cdot(x_3-x_1)\cdot\dots\cdot(x_n-x_1)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_2 & x_3 & \cdots & x_n \\ \vdots & \vdots & & \vdots \\ x_2^{n-2} & x_3^{n-2} & \cdots & x_n^{n-2} \end{vmatrix}$$

$$=\left[\prod_{k=2}^{n}(x_k-x_1)\right]V_{n-1}(x_2, x_3, \dots, x_n)$$

donde el símbolo $\prod_{k=2}^{n}$ significa el producto de todos los factores que aparecen variando $k$ de 2 a $n$.

De la misma manera se deducirá la fórmula

$$V_{n-1}(x_2, x_3, \dots, x_n)=\left[\prod_{k=3}^{n}(x_k-x_2)\right]V_{n-2}(x_3, \dots, x_n)$$

Se deducirá de aquí que $V_n(x_1, x_2, \dots, x_n)$ es el producto de todos los factores de la forma $(x_k-x_j)$ con $k>j$, es decir:

$$V_n(x_1, \dots, x_n)=\prod_{1\leq j<k\leq n}(x_k-x_j)$$

* * *

Finalmente mostramos un método para calcular determinantes de orden elevado siempre que se puedan relacionar con determinantes de orden inferior del mismo tipo. Por ejemplo, para calcular

$$D_6=\begin{vmatrix} 3 & 4 & 0 & 0 & 0 & 0 \\ 2 & 3 & 4 & 0 & 0 & 0 \\ 0 & 2 & 3 & 4 & 0 & 0 \\ 0 & 0 & 2 & 3 & 4 & 0 \\ 0 & 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 & 2 & 3 \end{vmatrix}$$

observamos que

$$D_6=3D_5-2\begin{vmatrix} 4 & 0 & 0 & 0 & 0 \\ 2 & 3 & 4 & 0 & 0 \\ 0 & 2 & 3 & 4 & 0 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 & 3 \end{vmatrix}=3D_5-8D_4$$



La fórmula $D_6=3D_5-8D_4$ se denomina *fórmula de recurrencia* y nos permite calcular $D_6$ conociendo $D_5$ y $D_4$. La misma fórmula de recurrencia aplicada a $D_5$ y $D_4$ nos da:

$$D_5=3D_4-8D_3, \quad D_4=3D_3-8D_2$$

Por tanto:

$$D_6=3[3D_4-8D_3]-8[3D_3-8D_2]=9D_4-2\cdot 3\cdot 8D_3+8^2D_2$$

Puesto que $D_3=3D_2-8D_1$ se tiene finalmente que

$$\begin{aligned} D_6&=9[3D_3-8D_2]-2\cdot 3\cdot 8D_3+8^2D_2=3^3[3D_2-8D_1]-9\cdot 8D_2- \\ &-2\cdot 3\cdot 8[3D_2-8D_1]+8^2D_2=[3^4-3^2\cdot 8-2\cdot 3^2\cdot 8+8^2]D_2+ \\ &+[-3^3\cdot 8+2\cdot 3\cdot 8^2]D_1=-71D_2+168D_1 \end{aligned}$$

Se tiene que

$$D_2=\begin{vmatrix} 3 & 4 \\ 2 & 3 \end{vmatrix}=9-8=1 \quad \text{y} \quad D_1=3$$

por tanto:

$$D_6=-71\cdot 1+168\times 3=-71+504=433$$

* * *

## EJERCICIOS 2.3

**1.** Encontrar una sucesión de matrices elementales $E_1, E_2, \dots, E_k$ tal que $E_k\cdots E_2E_1A$ sea una matriz escalonada, donde:

*a)* $A=\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 1 \end{pmatrix}$ *b)* $A=\begin{pmatrix} 1 & 2 & -1 \\ 3 & -1 & 2 \\ -1 & 5 & 4 \end{pmatrix}$ *c)* $A=\begin{pmatrix} 1 & 2 \\ -1 & 1 \\ 0 & 3 \end{pmatrix}$

*d)* $A=\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ *e)* $A=\begin{pmatrix} 2 & 1 & 0 \\ 3 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix}$

**2.** Expresar las matrices del ejercicio 1 que sean invertibles como un producto de matrices elementales.

**3.** Utilizar el teorema sobre el determinante de un producto para demostrar la fórmula

$$|A^{-1}|=\frac{1}{|A|}$$

cuando $A$ es una matriz invertible.

**4.** Sean $A$, $B$ y $C$ matrices cuadradas de orden $n$ y formemos la matriz

$$D=\begin{pmatrix} A & 0 \\ C & B \end{pmatrix}$$

de orden $2n$. *a)* Demostrar que

$$D=\begin{pmatrix} A & 0 \\ C & I_n \end{pmatrix}\begin{pmatrix} I_n & 0 \\ 0 & B \end{pmatrix}$$

*b)* Utilizar este resultado para demostrar que $|D|=|A||B|$.

**5.** Calcular el determinante de la matriz

$$A=\begin{pmatrix} a & b & c & d \\ -b & a & d & -c \\ -c & -d & a & b \\ -d & c & -b & a \end{pmatrix}$$

hallando $AA^t$. [*Sol.*: $(a^2+b^2+c^2+d^2)^2$.]

**6.** Escribir la matriz

$$A=\begin{pmatrix} 1+x_1y_1 & 1+x_1y_2 & 1+x_1y_3 \\ 1+x_2y_1 & 1+x_2y_2 & 1+x_2y_3 \\ 1+x_3y_1 & 1+x_3y_2 & 1+x_3y_3 \end{pmatrix}$$

como un producto de matrices y calcular su determinante. [*Sol.*: 0.]

**7.** Calcular los siguientes determinantes reduciéndolos al determinante de una matriz triangular (todas las matrices son de orden $n$).

*a)* $$\begin{vmatrix} 2 & 2 & 2 & \cdots & 2 \\ 2 & 2-x & 2 & \cdots & 2 \\ 2 & 2 & 2-x & \cdots & 2 \\ \vdots & \vdots & \vdots & & \vdots \\ 2 & 2 & 2 & \cdots & 2-x \end{vmatrix}$$

*b)* $$\begin{vmatrix} 1 & 2 & 3 & \cdots & n \\ 1 & x+1 & 3 & \cdots & n \\ 1 & 2 & x+1 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 2 & 3 & \cdots & x+1 \end{vmatrix}$$

*c)* $$\begin{vmatrix} n & 1 & 1 & \cdots & 1 \\ 1 & n & 1 & \cdots & 1 \\ 1 & 1 & n & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & n \end{vmatrix}$$

[*Sol.*: *a)* $2(-x)^{n-1}$. *b)* $(x-1)(x-2)\cdots(x-n+1)$. *c)* $(n-1)^{n-1}[2n-1]$.]



**8.** Demostrar que

$$\begin{vmatrix} 1 & n & n & \cdots & n \\ n & 2 & n & \cdots & n \\ n & n & 3 & \cdots & n \\ \vdots & \vdots & \vdots & & \vdots \\ n & n & n & \cdots & n \end{vmatrix}=(-1)^{n-1}n!$$

**9.** Encontrar una fórmula de recurrencia para calcular

$$D_n=\begin{vmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \end{vmatrix}$$

en función de $D_{n-1}$ y $D_{n-2}$. Utilizarla para calcular $D_8$ y $D_9$. [*Sol.*: $D_8=0$, $D_9=-1$.]

**10.** Una sucesión de Fibonacci (Fibonacci fue un matemático italiano del siglo XIII) es una sucesión de números que comienza con 1 y 2 y en la que cada número es la suma de los dos inmediatamente anteriores. Es decir, 1, 2, 3, 5, 8, ... Demostrar que el $n$-ésimo término de la sucesión de Fibonacci coincide con

$$|F_n|=\begin{vmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{vmatrix}$$

donde $F_n$ es una matriz de orden $n$.

**11.** Calcular

$$\begin{vmatrix} x-1 & 1 & 1 & 1 \\ 1 & x-1 & 1 & 1 \\ 1 & 1 & x-1 & 1 \\ 1 & 1 & 1 & x-1 \end{vmatrix}$$

**12.** Calcular el determinante de la matriz de orden $n$, $A=(a_{ij})$, donde $a_{ij}=|i-j|$. (Sugerencia: hacerlo para $n=2$, 3, 4, e intentar deducir la fórmula general para demostrarla por inducción.)

**13.** Calcular

$$\begin{vmatrix} 1 & x & x^2 & x^3 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 9 & 16 \end{vmatrix}$$

## 2.4. INVERSA DE UNA MATRIZ. REGLA DE CRAMER

En esta sección comenzaremos dando una fórmula para calcular la inversa de una matriz utilizando la teoría de los determinantes. A continuación usaremos esta fórmula para resolver sistemas de ecuaciones lineales cuya matriz de sus coeficientes es invertible.

Comenzaremos dando una condición necesaria y suficiente para la invertibilidad de una matriz haciendo uso del concepto de determinante.

**TEOREMA 1**

Sea $A$ una matriz cuadrada de orden $n$. Las siguientes condiciones son equivalentes:

1) $A$ es invertible.
2) rango $(A)=n$.
3) $|A|\neq 0$.

*Demostración.* Ya sabemos que 1) y 2) son equivalentes por el teorema 3 de la sección 1.4 (capítulo 1). Basta demostrar, por tanto, que 1) y 3) son equivalentes.

Si $A$ es invertible, $A$ puede escribirse como un producto de matrices elementales de la forma $A=E_1E_2\cdots E_k$ debido a los resultados de la sección anterior. Puesto que el determinante de un producto es el producto de los determinantes se tiene que

$$|A|=|E_1||E_2|\cdots|E_k|$$

La parte derecha de esta igualdad es no nula, ya que toda matriz elemental tiene determinante no nulo; por tanto, $|A|\neq 0$. Esto prueba que 1) implica 3).

En la etapa 3 del teorema 4 de la sección anterior se ha demostrado que si $A$ no es invertible, $|A|=0$; esto prueba que 3) implica 1), ya que si $|A|$ es no nulo y $A$ no fuera invertible se tendría una contradicción. ■

Suponiendo que $|A|\neq 0$, nos proponemos encontrar una fórmula para calcular $A^{-1}$. Si $A=(a_{ij})_{i,j=1,\dots,n}$, del desarrollo de su determinante por la fila $i$-ésima tenemos que

$$|A|=(-1)^{i+1}a_{i1}|A_{i1}|+(-1)^{i+2}a_{i2}|A_{i2}|+\cdots+(-1)^{i+n}a_{in}|A_{in}|$$

El número $c_{ij}=(-1)^{i+j}|A_{ij}|$, $i=1, \dots, n$, $j=1, \dots, n$, se denomina *cofactor* del elemento



que ocupa el lugar $(i, j)$ de la matriz $A$. Con esta notación la fórmula anterior se escribe de la forma

$$|A|=a_{i1}c_{i1}+a_{i2}c_{i2}+\cdots+a_{in}c_{in}$$

Sea

$$\operatorname{cof}(A)=C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{pmatrix}$$

la matriz de los cofactores de los elementos de $A$, que se denomina *matriz de cofactores de A*.

Con esta notación se tiene que la matriz $AC^t$ tiene como valor $|A|$ en todos los lugares de su diagonal principal. Es decir,

$$AC^t=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} c_{11} & \cdots & c_{i1} & \cdots & c_{n1} \\ \vdots & & \vdots & & \vdots \\ c_{12} & \cdots & c_{i2} & \cdots & c_{n2} \\ \vdots & & \vdots & & \vdots \\ c_{1n} & \cdots & c_{in} & \cdots & c_{nn} \end{pmatrix}=\begin{array}{c} \begin{pmatrix} |A| & & ? \\ & |A| & \\ ? & & |A| \end{pmatrix} \begin{array}{l} \\ \leftarrow \text{Fila } i \\ \\ \end{array} \\ \uparrow \\ \text{Columna } i \end{array}$$

Si logramos probar que los términos de $AC^t$ que no están en la diagonal principal son todos nulos, se tendría que

$$AC^t=|A|I_n$$

y, por tanto,

$$A^{-1}=\frac{1}{|A|}C^t$$

lo cual nos da una fórmula para calcular la inversa de la matriz $A$.

La demostración completa de la fórmula anterior para calcular $A^{-1}$ se da en el siguiente teorema:

**TEOREMA 2**

Si $A$ es una matriz invertible se tiene que

$$A^{-1}=\frac{1}{|A|}C^t$$

donde $C$ es la matriz de cofactores de $A$.

*Demostración.* De la discusión precedente se deduce que la demostración se completa si probamos que los elementos de $AC^t$ que no están en su diagonal son nulos. Sea $i \neq j$ e intentemos calcular

$$a_{i1}c_{j1}+a_{i2}c_{j2}+\cdots+a_{in}c_{jn}.$$

Esta expresión coincide con el desarrollo por la fila $j$ del determinante de la matriz

$$B=\begin{array}{cl}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \longleftarrow \text{Fila } i \\ \\ \longleftarrow \text{Fila } j \\ \\ \\ \end{array}\end{array}$$

que se obtiene de $A$ sustituyendo la fila $j$ por la $i$. Puesto que $B$ posee dos filas iguales, $|B| = 0$ y, por tanto:

$$a_{i1}c_{j1}+a_{i2}c_{j2}+\cdots+a_{in}c_{jn}=0$$

si $i \neq j$, que era lo que queríamos demostrar. ■

EJEMPLO A. La matriz

$$A=\begin{pmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 0 & 1 & 3 \end{pmatrix}$$

es invertible puesto que

$$|A|=\begin{vmatrix} 1 & 2 & 0 \\ -1 & 1 & 2 \\ 0 & 1 & 3 \end{vmatrix}=\begin{vmatrix} 1 & 2 & 0 \\ 0 & 3 & 2 \\ 0 & 1 & 3 \end{vmatrix}=1\begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix}=7\neq 0$$

Para calcular su inversa comenzamos calculando la matriz de sus cofactores:

$$c_{11}=\begin{vmatrix} 1 & 2 \\ 1 & 3 \end{vmatrix}=1 \quad ; \quad c_{12}=-\begin{vmatrix} -1 & 2 \\ 0 & 3 \end{vmatrix}=3 \quad ; \quad c_{13}=\begin{vmatrix} -1 & 1 \\ 0 & 1 \end{vmatrix}=-1$$

$$c_{21}=-\begin{vmatrix} 2 & 0 \\ 1 & 3 \end{vmatrix}=-6 \quad ; \quad c_{22}=\begin{vmatrix} 1 & 0 \\ 0 & 3 \end{vmatrix}=3 \quad ; \quad c_{23}=-\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix}=-1$$

$$c_{31}=\begin{vmatrix} 2 & 0 \\ 1 & 2 \end{vmatrix}=4 \quad ; \quad c_{32}=-\begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix}=-2 \quad ; \quad c_{33}=\begin{vmatrix} 1 & 2 \\ -1 & 1 \end{vmatrix}=3$$



Por tanto:

$$C=\mathrm{cof}(A)=\begin{pmatrix} 1 & 3 & -1 \\ -6 & 3 & -1 \\ 4 & -2 & 3 \end{pmatrix}$$

Por el teorema 2:

$$A^{-1}=\frac{1}{|A|}C^t=\frac{1}{7}\begin{pmatrix} 1 & -6 & 4 \\ 3 & 3 & -2 \\ -1 & -1 & 3 \end{pmatrix}=\begin{pmatrix} 1/7 & -6/7 & 4/7 \\ 3/7 & 3/7 & -2/7 \\ -1/7 & -1/7 & 3/7 \end{pmatrix}$$

(¡Comprobarlo!)

* * *

*Notas.* 1) Una matriz con determinante nulo se dice que es una matriz *singular*, de manera que una matriz invertible también recibe el nombre de matriz *no singular*.

2) Es conveniente que el lector se ejercite en el cálculo de inversas de matrices de orden 3 hasta el punto de que pueda calcularlas sin necesidad de escribir la matriz de los cofactores.

* * *

Utilizando el teorema 2 pueden resolverse sistemas de ecuaciones lineales de la forma:

$$\left.\begin{array}{c} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad \vdots \\ a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n \end{array}\right\}$$

escrito $A\vec{x}=\vec{b}$ en notación matricial, siempre que $A$ sea no singular o invertible. En efecto, por el teorema 2:

$$\vec{x}=A^{-1}\vec{b}=\frac{1}{|A|}C^t\vec{b}=\frac{1}{|A|}\begin{pmatrix} c_{11} & c_{21} & \cdots & c_{n1} \\ \vdots & \vdots & & \vdots \\ c_{1j} & c_{2j} & \cdots & c_{nj} \\ \vdots & \vdots & & \vdots \\ c_{1n} & c_{2n} & \cdots & c_{nn} \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{pmatrix}=$$

$$=\frac{1}{|A|}\begin{array}{cl}\begin{pmatrix} b_1c_{11}+ & b_2c_{21}+\cdots+ & b_nc_{n1} \\ \vdots & \vdots & \vdots \\ b_1c_{1j}+ & b_2c_{2j}+\cdots+ & b_nc_{nj} \\ \vdots & \vdots & \vdots \\ b_1c_{1n}+ & b_2c_{2n}+\cdots+ & b_nc_{nn} \end{pmatrix} & \begin{array}{l} \\ \\ \longleftarrow \text{Componente } j \\ \\ \\ \end{array}\end{array}$$

Por tanto, para todo $j=1, 2, \ldots, n$ se tiene que

$$x_j=\frac{1}{|A|}[b_1c_{1j}+b_2c_{2j}+\cdots+b_nc_{nj}]$$

La expresión entre paréntesis es el desarrollo del determinante de la matriz

$$A_j=\begin{pmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & & & \vdots & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{pmatrix}$$
$$\begin{array}{c}\uparrow\\ \text{Columna } j\end{array}$$

por la columna $j$.

Hemos probado el siguiente resultado:

**TEOREMA 3** *(Regla de Cramer)*

Si $A$ es una matriz invertible de orden $n$, el sistema de ecuaciones lineales $A\vec{x}=\vec{b}$ es compatible determinado y sus soluciones se calculan mediante la fórmula

$$x_j=\frac{|A_j|}{|A|}, \quad j=1, 2, \ldots, n$$

donde $A_j$ es la matriz $A$ en la que se ha sustituido la columna $j$ por el vector $\vec{b}$.

EJEMPLO B. El sistema

$$\left.\begin{array}{r} 2x_1+3x_2=8 \\ x_1+x_3=1 \\ 2x_2-x_4=1 \\ x_1+x_4=4 \end{array}\right\}$$

tiene solución única, ya que

$$\begin{vmatrix} 2 & 3 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 1 \end{vmatrix}=-\begin{vmatrix} 2 & 3 & 0 \\ 0 & 2 & -1 \\ 1 & 0 & 1 \end{vmatrix}=-4+3=-1\neq 0$$



Por el teorema 3:

$$x_1=\frac{1}{-1}\begin{vmatrix} 8 & 3 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & -1 \\ 4 & 0 & 0 & 1 \end{vmatrix}=\begin{vmatrix} 8 & 3 & 0 \\ 1 & 2 & -1 \\ 4 & 0 & 1 \end{vmatrix}=4\begin{vmatrix} 3 & 0 \\ 2 & -1 \end{vmatrix}+\begin{vmatrix} 8 & 3 \\ 1 & 2 \end{vmatrix}=-12+13=1$$

$$x_2=\frac{1}{-1}\begin{vmatrix} 2 & 8 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \\ 1 & 4 & 0 & 1 \end{vmatrix}=\begin{vmatrix} 2 & 8 & 0 \\ 0 & 1 & -1 \\ 1 & 4 & 1 \end{vmatrix}=2\begin{vmatrix} 1 & -1 \\ 4 & 1 \end{vmatrix}+\begin{vmatrix} 8 & 0 \\ 1 & -1 \end{vmatrix}=10-8=2$$

$$x_3=\frac{1}{-1}\begin{vmatrix} 2 & 3 & 8 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 1 & -1 \\ 1 & 0 & 4 & 1 \end{vmatrix}=-\begin{vmatrix} 2 & 3 & 8 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 2 & 5 & 0 \\ 1 & 0 & 4 & 1 \end{vmatrix}=-\begin{vmatrix} 2 & 3 & 8 \\ 1 & 0 & 1 \\ 1 & 2 & 5 \end{vmatrix}=-\begin{vmatrix} 0 & -1 & -2 \\ 0 & -2 & -4 \\ 1 & 2 & 5 \end{vmatrix}=0$$

$$x_4=\frac{1}{-1}\begin{vmatrix} 2 & 3 & 0 & 8 \\ 1 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 \\ 1 & 0 & 0 & 4 \end{vmatrix}=\begin{vmatrix} 2 & 3 & 8 \\ 0 & 2 & 1 \\ 1 & 0 & 4 \end{vmatrix}=2\cdot\begin{vmatrix} 2 & 1 \\ 0 & 4 \end{vmatrix}+\begin{vmatrix} 3 & 8 \\ 2 & 1 \end{vmatrix}=16+(3-16)=3$$

EJEMPLO C. Para determinar si el sistema homogéneo

$$\left.\begin{array}{r} x_1+3x_2+x_3+x_4=0 \\ 2x_1+2x_2+2x_3=0 \\ 2x_1+x_3=0 \\ -x_1-x_2-x_4=0 \end{array}\right\}$$

tiene soluciones no triviales basta calcular el determinante de la matriz $A$ de sus coeficientes. Si $|A|\neq 0$ solamente posee la solución trivial y si $|A|=0$ posee soluciones no triviales, ya que en este caso $r(A)<n$ debido al teorema 1.

Puesto que

$$\begin{vmatrix} 1 & 3 & 1 & 1 \\ 2 & 2 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ -1 & -1 & 0 & -1 \end{vmatrix} \overset{F_2-F_1-F_3}{\underset{\text{P.6}}{=\!=\!=}} \begin{vmatrix} 1 & 3 & 1 & +1 \\ -1 & -1 & 0 & -1 \\ 2 & 0 & 1 & 0 \\ -1 & -1 & 0 & -1 \end{vmatrix} \overset{F_2=F_4}{\underset{\text{P.2}}{=\!=\!=}} 0$$

el sistema posee infinitas soluciones.

EJERCICIOS 2.4

**1.** Hallar las inversas de las siguientes matrices, calculando primero la matriz de cofactores (comprobar los resultados):

*a)* $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 1 & -1 \\ 3 & 1 & 4 \end{pmatrix}$ *b)* $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ *c)* $\begin{pmatrix} 2 & 0 & 1 \\ 1 & 3 & 2 \\ 1 & -3 & 3 \end{pmatrix}$ *d)* $\begin{pmatrix} 1 & 3 & 0 & 3 \\ 2 & 1 & -1 & 1 \\ 0 & 1 & 4 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$

**2.** Demostrar que si $A=\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ con $|A|\neq 0$, $A^{-1}=\dfrac{1}{ad-cb}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ utilizando el teorema sobre la inversa de una matriz.

**3.** Demostrar que $(A^{-1})^t=(A^t)^{-1}$.

**4.** Probar que $|\operatorname{cof}(A)|=|A|^{n-1}$.

**5.** Hallar la inversa de las siguientes matrices:

*a)* $\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & -1 & & & & 0 \\ & 1 & -1 & & & \\ & & 1 & -1 & & \\ & & & 1 & -1 & \\ & & & & 1 & -1 \\ 0 & & & & & 1 \end{pmatrix}$ *c)* $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & a & -b \\ 0 & 0 & b & a \end{pmatrix}$, $\begin{matrix} a\neq 0 \\ b\neq 0 \end{matrix}$

**6.** Hallar la inversa de las siguientes matrices de orden $n$:

*a)* $\begin{pmatrix} 1 & 2 & 2^2 & \cdots & 2^{n-1} \\ 0 & 1 & 2 & \cdots & 2^{n-2} \\ 0 & 0 & 1 & \cdots & 2^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$ *b)* $\begin{pmatrix} 1+x & 1 & 1 & \cdots & 1 \\ 1 & 1+x & 1 & \cdots & 1 \\ 1 & 1 & 1+x & \cdots & 1 \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & 1 & 1 & \cdots & 1+x \end{pmatrix}$

**7.** Encontrar los valores de $a$ y $b$ para los que las matrices siguientes son invertibles:

*a)* $\begin{pmatrix} a+1 & 1 & 1 \\ 1 & a-1 & 1 \\ 0 & 1 & a+2 \end{pmatrix}$ *b)* $\begin{pmatrix} 2 & 0 & 0 \\ 0 & a+1 & -1 \\ 0 & 1 & a-3 \end{pmatrix}$ *c)* $\begin{pmatrix} 1 & a & b \\ 1 & a^2 & b^2 \\ 1 & a^3 & b^3 \end{pmatrix}$



**8.** Dada la aplicación lineal $T\colon \mathbb{R}^3 \mapsto \mathbb{R}^3$ mediante $T(x, y, z)=(x+2y+z,\ 3x+y+2z,\ 3x+2y+2z)$, demostrar que $T$ es invertible y encontrar su inversa.

**9.** Sea $A$ una matriz invertible de orden $n$ y triangular superior. Demostrar que $A^{-1}$ es también triangular superior.

**10.** Resolver los siguientes sistemas de ecuaciones utilizando la regla de Cramer.

*a)* $\left.\begin{aligned} x_1+2x_2-x_3&=7 \\ 2x_1+x_2+x_3&=6 \\ x_1-x_2+3x_3&=-1 \end{aligned}\right\}$ *b)* $\left.\begin{aligned} 2x_1+3x_2-5x_3-2&=0 \\ 3x_1-x_2+2x_3+1&=0 \\ 5x_1+4x_2-6x_3-3&=0 \end{aligned}\right\}$

*c)* $\left.\begin{aligned} 2x_1+x_2+4x_3+8x_4&=-1 \\ x_1+3x_2-6x_3+2x_4&=3 \\ 3x_1-2x_2+2x_3-2x_4&=8 \\ 2x_1-x_2+2x_3&=4 \end{aligned}\right\}$

**11.** Hallar los valores de $m$ para los que el siguiente sistema posee soluciones no triviales:

$$\left.\begin{aligned} 2x-y+z&=0 \\ x+my-z&=0 \\ x+y+z&=0 \end{aligned}\right\}$$

**12.** Hallar la inversa de cada una de las siguientes matrices:

*a)* $\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\operatorname{sen}\alpha \\ 0 & \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$ *c)* $\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{pmatrix}$

Un polinomio de grado $n$ con coeficientes reales es una expresión de la forma $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$, con $a_j\in\mathbb{R}$, $j=0, 1, \ldots, n$ y $a_n\neq 0$.

**13.** Encontrar un polinomio $f$ de grado 2 tal que $f(1)=2$, $f(-1)=4$ y $f(3)=16$.

**14.** Demostrar que $n+1$ valores distintos de la variable $x$ determinan de manera única un polinomio de grado $n$. [*Sugerencia:* reducir el problema a demostrar que el determinante de una cierta matriz es no nulo y observar que esta matriz es la que origina el determinante de Vandermonde.]

**15.** Demostrar que si un polinomio de grado $n$ tiene $n+1$ soluciones reales distintas ha de ser el polinomio nulo. [*Nota:* $x_0\in\mathbb{R}$ es solución de $f$ si $f(x_0)=0$.]

**16.** Decir para qué valores de $a$ es invertible la matriz $A=(a_{ij})$, donde $a_{ij}=i-j$ si $i>j$ y $a_{ij}=a$ en el resto de los casos.

## 2.5. RANGO DE UNA MATRIZ. RESOLUCION DE SISTEMAS COMPATIBLES E INDETERMINADOS

En la sección 1.2 del capítulo 1 definimos el rango de una matriz como el mayor número de sus vectores columna que son linealmente independientes. En el teorema 1 de la misma sección se demostró que el rango de una matriz coincide con el número de peldaños de su matriz escalonada.

En esta sección utilizamos los determinantes para obtener el rango de una matriz. Demostraremos también el sorprendente resultado de que el rango de los vectores filas de una matriz coincide con el rango de sus vectores columna.

Dada una matriz $A$ de orden $m \times n$ se denomina *menor de orden $k$ de $A$* al determinante de cualquier matriz de orden $k$ formada con los elementos de la intersección de $k$ cualesquiera de sus filas y $k$ cualesquiera de sus columnas.

EJEMPLO A. Dada

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix}$$

un menor de orden 3 es

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix} \mapsto \begin{vmatrix} 1 & 2 & 4 \\ -1 & 2 & 1 \\ 1 & 2 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 2 & 4 \\ 0 & 4 & 5 \\ 0 & 0 & -2 \end{vmatrix} = -8$$

y menores de orden 2 son, por ejemplo:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix} \mapsto \begin{vmatrix} -1 & 1 \\ 1 & 0 \end{vmatrix} = -1 \quad \text{y} \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 2 \end{pmatrix} \mapsto \begin{vmatrix} 2 & 4 \\ 2 & 2 \end{vmatrix} = -4$$

* * *

Los menores de orden 1 de una matriz son sus elementos. Si $A$ es una matriz cuadrada de orden $n$ solamente hay un menor de $A$ de orden $n$ que coincide con $|A|$. En este mismo caso los menores de los elementos de $A$ son menores de orden $n-1$.

Sea $A$ una matriz no nula; *siempre podemos encontrar un único número $R$ que satisface las siguientes condiciones:*

1) *$A$ posee al menos un menor no nulo de orden $R$.*
2) *Todo menor de $A$ de orden mayor que $R$ es nulo.*

Cualquier menor de $A$ de orden $R$ que no es nulo se denomina *menor básico;* las columnas de la matriz de las que proviene este menor básico se denominan *columnas básicas y a las filas se les denomina filas básicas.*



EJEMPLO B. En la matriz

$$A = \begin{pmatrix} 2 & 2 & 4 & 6 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 3 & 4 \end{pmatrix}$$

los menores de orden tres son:

$$\begin{vmatrix} 2 & 2 & 4 \\ 0 & 0 & 1 \\ 1 & 1 & 3 \end{vmatrix}, \quad \begin{vmatrix} 2 & 2 & 6 \\ 0 & 0 & 1 \\ 1 & 1 & 4 \end{vmatrix}, \quad \begin{vmatrix} 2 & 4 & 6 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{vmatrix} \quad \text{y} \quad \begin{vmatrix} 2 & 4 & 6 \\ 0 & 1 & 1 \\ 1 & 3 & 4 \end{vmatrix}$$

que son todos nulos. Sin embargo, el menor de orden 2:

$$\begin{vmatrix} 2 & 4 \\ 0 & 1 \end{vmatrix} = 2$$

es no nulo y, por tanto, es un menor básico de $A$; sus columnas básicas son la primera y la tercera y sus filas básicas son la primera y la segunda.

Es fácil encontrar otros menores básicos de orden 2 de esta matriz.

* *

El rango de la matriz $A$ del ejemplo $B$ es 2, ya que

$$\begin{pmatrix} 2 & 2 & 4 & 6 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 3 & 4 \end{pmatrix} \xrightarrow{i)} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 3 & 4 \end{pmatrix} \xrightarrow{iii)} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\xrightarrow{iii)} \begin{pmatrix} 1 & 1 & 2 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{iii)} \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

que coincide con el número $R$ de la matriz $A$.

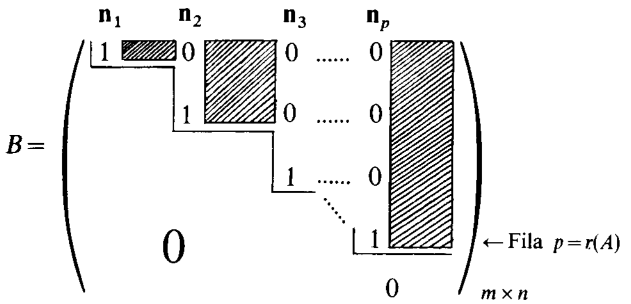
Esto nos hace pensar en la certeza del siguiente resultado:

TEOREMA 1

Dada una matriz no nula, el número $R$ anteriormente definido coincide con su rango.

*Demostración.* El número $R$ anteriormente definido no se altera mediante transformaciones elementales en la matriz; la demostración de esta afirmación se deja como ejercicio para el lector (ver el ejercicio 7 al final de esta sección).

Así pues, el número $R$ para una matriz $A$ no nula coincide con el número $R$ para su matriz escalonada $B$. La matriz escalonada $B$ es de la forma




donde las matrices sombreadas pueden ser no nulas. El número $R$ para la matriz $B$ no puede ser superior a $p$, ya que cualquier menor de orden $p+1$ incluiría una fila de ceros.

Por otro lado, ha de ser igual a $p$, puesto que podemos tomar las $p$ primeras filas y las columnas $n_1, n_2, n_3, \ldots, n_p$ para obtener

$$\begin{vmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots \ 1 \end{vmatrix} = 1 \neq 0$$

Por tanto, $R$ coincide con el número de peldaños de $B$; por el teorema 1 de la sección 1.2 este número coincide con el rango de la matriz. ■

EJEMPLO C. El rango de la matriz $A=\begin{pmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 3 & 1 & 5 \end{pmatrix}$ es 2, ya que

$$\begin{vmatrix} 1 & 2 & 0 \\ 2 & 1 & 3 \\ 3 & 1 & 5 \end{vmatrix} \underset{C_2-2C_1}{=\!=\!=\!=} \begin{vmatrix} 1 & 0 & 0 \\ 2 & -3 & 3 \\ 3 & -5 & 5 \end{vmatrix} = -\begin{vmatrix} 1 & 0 & 0 \\ 2 & 3 & 3 \\ 3 & 5 & 5 \end{vmatrix} = 0 \quad \text{y} \quad \begin{vmatrix} 1 & 2 \\ 2 & 1 \end{vmatrix} = -3 \neq 0.$$

* * *

TEOREMA 2 *(Teorema del menor básico)*

Sea $A$ una matriz no nula. 1) Cualquier columna de $A$ es combinación lineal de sus columnas básicas. 2) El mismo resultado es cierto para las filas de $A$.



*Demostración.* Debido a las propiedades de los determinantes se puede suponer, sin pérdida de generalidad, que el menor básico ocupa las $R$ primeras filas y las $R$ primeras columnas de la matriz $A$, donde $r(A)=R$. Es decir:

$$A=\left(\begin{array}{ccc|cc} a_{11} & \cdots & a_{1R} & \cdots & a_{1n} \\ \vdots & C & \vdots & & \vdots \\ a_{R1} & \cdots & a_{RR} & \cdots & a_{Rn} \\ \hline \vdots & & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mR} & \cdots & a_{mn} \end{array}\right)$$

donde $|C|\neq 0$. Sea $\vec{v}_j$, $j=1, 2, \ldots, n$, el $j$-ésimo vector columna de la matriz $A$. Para demostrar 1) basta probar que cualquier vector $\vec{v}_j$ es combinación lineal de los vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_R$.

Si $1\leq j\leq R$, $\vec{v}_j$ es uno de los vectores $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_R$ y, por tanto, es combinación lineal de ellos.

Si $j>R$ consideramos el sistema homogéneo

$$\begin{pmatrix} a_{11} & \cdots & a_{1R} & a_{1j} \\ \vdots & & \vdots & \vdots \\ a_{R1} & \cdots & a_{RR} & a_{Rj} \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mR} & a_{mj} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_R \\ x_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix} \tag{I}$$

de $m$ ecuaciones con $R+1$ incógnitas. Puesto que el rango de la matriz de los coeficientes de este sistema es $R$, que es menor que el número de incógnitas, el sistema tiene infinitas soluciones. De todas estas soluciones siempre podemos elegir una de la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_R \\ x_j \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_R \\ d_j \end{pmatrix}$$

con $d_j\neq 0$, ya que si todas las soluciones tuvieran $d_j=0$ el sistema (I) sería equivalente al sistema

$$\begin{pmatrix} a_{11} & \cdots & a_{1R} \\ \vdots & & \vdots \\ a_{R1} & \cdots & a_{RR} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mR} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_R \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

que únicamente posee la solución nula. Sustituyendo esta solución en (I) tenemos la igualdad

$$\begin{pmatrix} a_{11} & \cdots & a_{1R} & a_{1j} \\ \vdots & & \vdots & \vdots \\ a_{R1} & \cdots & a_{RR} & a_{Rj} \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mR} & a_{mj} \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_R \\ d_j \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

Esta igualdad puede escribirse de la forma

$$d_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{R1} \\ \vdots \\ a_{m1} \end{pmatrix} + \cdots + d_R \begin{pmatrix} a_{1R} \\ \vdots \\ a_{RR} \\ \vdots \\ a_{mR} \end{pmatrix} + d_j \begin{pmatrix} a_{1j} \\ \vdots \\ a_{Rj} \\ \vdots \\ a_{mj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

que coincide con $d_1\vec{v}_1 + \cdots + d_R\vec{v}_R + d_j\vec{v}_j = \vec{0}$. Puesto que $d_j \neq 0$ se deduce que

$$\vec{v}_j = -\frac{d_1}{d_j}\vec{v}_1 - \cdots - \frac{d_R}{d_j}\vec{v}_R$$

lo cual prueba que $\vec{v}_j$ es combinación lineal de $\vec{v}_1, ..., \vec{v}_R$.

El resultado por filas se obtiene de considerar la matriz traspuesta $A^t$ de $A$, que ha de tener el mismo rango que $A$ debido al teorema 4 de la sección 2.2. ■

EJEMPLO D. La matriz

$$A = \begin{pmatrix} 2 & 1 \\ 3 & 1 \\ 1 & 1 \\ 4 & 0 \end{pmatrix}$$

tiene rango 2 puesto que $\begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = 2 - 3 = -1$. Por tanto, las filas tercera y cuarta son combinación lineal de sus filas primera y segunda. Estas combinaciones lineales se encuentran resolviendo los sistemas

$$(1, 1) = c_1(2, 1) + c_2(3, 1) \quad \text{y} \quad (4, 0) = d_1(2, 1) + d_2(3, 1)$$



o equivalentemente:

$$\left.\begin{aligned} 2c_1 + 3c_2 &= 1 \\ c_1 + c_2 &= 1 \end{aligned}\right\} \quad \text{y} \quad \left.\begin{aligned} 2d_1 + 3d_2 &= 4 \\ d_1 + d_2 &= 0 \end{aligned}\right\}$$

lo cual es un buen ejercicio para el lector.

* * *

COROLARIO 3

El número máximo de columnas linealmente independientes de una matriz es igual al número máximo de sus filas linealmente independientes.

*Demostración.* $r(A) = r(A^t)$.

COROLARIO 4

El determinante de una matriz (cuadrada) es nulo si y sólo si una de sus filas (columna) es combinación lineal de las restantes filas (columnas) de la matriz.

*Demostración.* Si $A$ es una matriz cuadrada con $|A| = 0$, $r(A) < n$ (teorema 1 de la sección 2.4); por el teorema 2 se obtiene el resultado deseado.

Por otro lado, si $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ son los vectores columna de $A$ y $\vec{v}_n = \alpha_1\vec{v}_1 + \alpha_2\vec{v}_2 + \cdots + \alpha_{n-1}\vec{v}_{n-1}$ realizando las operaciones elementales:

$$C_n - \alpha_1 C_1 - \alpha_2 C_2 - \cdots - \alpha_{n-1} C_{n-1}$$

sobre sus columnas se tiene que

$$|A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1,n-1} & 0 \\ a_{21} & a_{22} & \cdots & a_{2,n-1} & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{vmatrix} = 0$$ ■

EJEMPLO E. El determinante de la matriz $\begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1 \end{pmatrix}$ es nulo, ya que su tercera columna es dos veces la primera menos la segunda. ¿Qué fila es combinación lineal de las restantes?

* * *

Los resultados que hemos obtenido sobre el rango de una matriz se aplican a continuación para resolver sistemas compatibles e indeterminados.

EJEMPLO F. Deseamos resolver el sistema

$$\left.\begin{aligned} x_1+2x_2&=1\\ 2x_1+x_2+3x_3&=2\\ 3x_1+x_2+5x_3&=3\end{aligned}\right\} \qquad \text{(I)}$$

Sea $A=\begin{pmatrix}1&2&0\\2&1&3\\3&1&5\end{pmatrix}$ la matriz de sus coeficientes y $\bar{A}=\begin{pmatrix}1&2&0&1\\2&1&3&2\\3&1&5&3\end{pmatrix}$ la matriz ampliada. Puesto que $\begin{vmatrix}1&2&1\\2&1&3\\3&1&5\end{vmatrix}=0$ y $\begin{vmatrix}1&2\\2&1\end{vmatrix}=-3$, $r(A)=2$, y puesto que la cuarta columna de $\bar{A}$ coincide con la primera, $r(\bar{A})=2$. Por el teorema de Rouche-Frobenius, el sistema es compatible e indeterminado.

Puesto que $\bar{A}$ tiene como básicas sus dos primeras filas, por el teorema 2 la tercera debe de ser combinación lineal de las dos restantes; por tanto, la tercera ecuación del sistema no introduce nuevas soluciones en (I).

Podemos suprimirla del sistema y tenemos el sistema equivalente:

$$\left.\begin{aligned} x_1+2x_2&=1\\ 2x_1+x_2+3x_3&=2\end{aligned}\right\}$$

Poniendo a la derecha de las igualdades las incógnitas correspondientes a las columnas no básicas se tiene

$$\left.\begin{aligned} x_1+2x_2&=1\\ 2x_1+x_2&=2-3x_3\end{aligned}\right\}$$

Resolviendo este sistema por la regla de Cramer en función de $x_3$ se tiene:

$$x_1=\frac{\begin{vmatrix}1&2\\2-3x_3&1\end{vmatrix}}{\begin{vmatrix}1&2\\2&1\end{vmatrix}}=\frac{1-4+6x_3}{-3}=1-2x_3$$

$$x_2=\frac{\begin{vmatrix}1&1\\2&2-3x_3\end{vmatrix}}{-3}=\frac{2-3x_3-2}{-3}=x_3$$

Haciendo $x_3=c$ se tiene

$$(x_1,\ x_2,\ x_3)=(1-2c,\ c,\ c)=(1,\ 0,\ 0)+c(-2,\ 1,\ 1)$$



EJEMPLO G. El sistema

$$\left.\begin{aligned} 2x_1-3x_2+5x_3+7x_4&=1\\ 4x_1-6x_2+2x_3+3x_4&=2\\ 2x_1-3x_2-11x_3-15x_4&=1\end{aligned}\right\}$$

es compatible e indeterminado, ya que $r(A)=2=r(\bar{A})<4=$número de incógnitas (¡comprobarlo!). Puesto que $\begin{vmatrix}5&7\\2&3\end{vmatrix}=1\neq 0$ podemos suprimir la tercera fila y pasar al segundo miembro de las igualdades las incógnitas $x_1$ y $x_2$ para obtener

$$\left.\begin{aligned} 5x_3+7x_4&=1-2x_1+3x_2\\ 2x_3+3x_4&=2-4x_1+6x_2\end{aligned}\right\}$$

Resolviendo por la regla de Cramer se tiene

$$x_3=\frac{\begin{vmatrix}1-2x_1+3x_2&7\\2-4x_1+6x_2&3\end{vmatrix}}{1}=3-6x_1+9x_2-14+28x_1-42x_2=$$
$$=22x_1-33x_2-11$$

$$x_4=\frac{\begin{vmatrix}5&1-2x_1+3x_2\\2&2-4x_1+6x_2\end{vmatrix}}{1}=10-20x_1+30x_2-2+4x_1-6x_2=$$
$$=-16x_1+24x_2+8$$

Haciendo $x_1=c_1$, $x_2=c_2$ se tiene:

$$(x_1,\ x_2,\ x_3,\ x_4)=(c_1,\ c_2,\ 22c_1-33c_2-11,\ -16c_1+24c_2+8)=$$
$$=(0,\ 0,\ -11,\ 8)+c_1(1,\ 0,\ 22,\ -16)+c_2(0,\ 1,\ -33,\ 24)$$

* * *

Resumimos a continuación los pasos a seguir para resolver un sistema compatible e indeterminado:

1) Detectar un menor básico de $A$.
2) Suprimir las ecuaciones que corresponden a filas no básicas de $A$.
3) Poner a la derecha de las ecuaciones aquellas incógnitas que corresponden a columnas no básicas de $A$.
4) Resolver por la regla de Cramer.

* * *

En la sección 1.2 (capítulo 1) se demostró que todo sistema homogéneo e indeterminado tiene $k=n-r(A)$ soluciones linealmente independientes de manera que cualquier otra solución del sistema es combinación lineal de las anteriores (véase la proposición 3 de la citada sección).

Para terminar esta sección daremos una nueva demostración de este resultado.

Sea $A\vec{x}=\vec{0}$ un sistema homogéneo y sea $R=r(A)$. Podemos suponer, sin pérdida de generalidad, que un menor básico de $A$ está en su parte superior izquierda. Resolviendo $A\vec{x}=\vec{0}$ con las reglas dadas para resolver sistemas compatibles e indeterminadas se tiene el sistema equivalente:

$$\left.\begin{array}{l} a_{11}x_1+\cdots+a_{1R}x_R=-a_{1,R+1}x_{R+1}-\cdots-a_{1n}x_n \\ \vdots \qquad\quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\ a_{m1}x_1+\cdots+a_{mR}x_R=-a_{m,R+1}x_{R+1}-\cdots-a_{mn}x_n \end{array}\right\} \tag{I}$$

Sea $\vec{v}_j=(v_{j1}, v_{j2}, ..., v_{jR})$, $j=1, ..., n-R$, la solución que se obtiene de (I) haciendo $x_{R+j}=1$ y $x_{R+k}=0$ si $k\neq j$, $k=1, 2, ..., n-R$, en la parte derecha del sistema.

Los vectores

$$\vec{e}_j=(v_{j1}, v_{j2}, ..., v_{jR}, 0, ..., 0, \overset{R+j}{1}, 0, 0),$$

$j=1, ..., n-R$, son linealmente independientes ya que el rango de la matriz

$$\begin{pmatrix} v_{11} & v_{12} & ..., & v_{1R} & 1 & 0 & ... & 0 \\ v_{21} & v_{22} & ..., & v_{2R} & 0 & 1 & ... & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ v_{n-R,1} & v_{n-R,2} & ..., & v_{n-R,R} & 0 & 0 & ... & 1 \end{pmatrix} \longleftarrow \text{Menor básico de orden } n-R$$

es $n-R$.

Falta demostrar que cualquier otra solución de $A\vec{x}=\vec{0}$ es combinación lineal de $\vec{e}_1, \vec{e}_2, ..., \vec{e}_{n-R}$.

Sea

$$\vec{e}=(u_1, u_2, ..., u_R, u_{R+1}, ..., u_n)$$

otra solución de $A\vec{x}=\vec{0}$. Consideremos el vector

$$\vec{e}_0=\vec{e}-u_{R+1}\vec{e}_1-\cdots-u_n\vec{e}_{n-R}$$

El vector $\vec{e}_0$ es de la forma

$$\vec{e}_0=(z_1, z_2, ..., z_R, 0, 0, ..., 0)$$

como fácilmente puede comprobarse. Además, $\vec{e}_0$ es solución de $A\vec{x}=\vec{0}$, ya que es combinación lineal de soluciones de este mismo sistema (proposición 2 de la sección 1.2).



Por tanto, $(z_1, z_2, ..., z_R)$ es solución del sistema

$$\left.\begin{array}{lll} a_{11}x_1 & +\cdots+ & a_{1R}x_R=0 \\ \vdots & & \vdots \\ a_{R1}x_1 & +\cdots+ & a_{RR}x_R=0 \end{array}\right\}$$

que se ha obtenido de (I) poniendo $x_{R+1}=0, ..., x_n=0$, como corresponde a las $n-R$ últimas componentes de $\vec{e}_0$. Puesto que el determinante de la matriz de este sistema es no nulo (¿por qué?), el sistema solamente posee la solución trivial. Por tanto, $\vec{e}_0=0$ y tenemos

$$\vec{e}=u_{R+1}\vec{e}_1+\cdots+u_n\vec{e}_{n-R}$$

que era lo que deseábamos demostrar. ∎

EJERCICIOS 2.5

**1.** Encontrar, mediante menores, el rango de las siguientes matrices:

*a)* $\begin{pmatrix} 2 & -1 & 3 & -2 & 4 \\ 4 & -2 & 5 & 1 & 7 \\ 2 & -1 & 1 & 8 & 2 \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 1 & 1 & -2 \end{pmatrix}$ *c)* $\begin{pmatrix} 1 & 2 \\ 3 & 1 \\ 4 & 0 \end{pmatrix}$

*d)* $\begin{pmatrix} 4 & 3 & -5 & 2 & 3 \\ 8 & 6 & -7 & 4 & 2 \\ 4 & 3 & -8 & 2 & 7 \\ 4 & 3 & 1 & 2 & -5 \\ 8 & 6 & -1 & 4 & -6 \end{pmatrix}$

**2.** Encontrar el rango de la matriz

$$\begin{pmatrix} 1 & a & -1 & 2 \\ 2 & -1 & a & 5 \\ 1 & 10 & -6 & 1 \end{pmatrix}$$

para los distintos valores del número real $a$.

**3.** Demostrar que $n+1$ vectores cualesquiera de $\mathbb{R}^n$ son siempre linealmente dependientes.

**4.** Probar que los siguientes sistemas son compatibles e indeterminados y resolverlos.

*a)* $\left.\begin{array}{r} 3x_1+4x_2+x_3+2x_4=3 \\ 6x_1+8x_2+2x_3+5x_4=7 \\ 9x_1+12x_2+3x_3+10x_4=13 \end{array}\right\}$ *b)* $\left.\begin{array}{r} 4x_1+x_2-2x_3+x_4=3 \\ x_1-2x_2-x_3+2x_4=2 \\ 2x_1+5x_2-x_4=-1 \\ 3x_1+3x_2-x_3-3x_4=1 \end{array}\right\}$

*c*) $$\left.\begin{aligned} x_1+x_2+x_3&=1\\ -x_1+3x_2&=0\\ 2x_1+6x_2+3x_3&=3 \end{aligned}\right\}$$

**5.** Estudiar mediante menores y aplicando el teorema de Rouche-Frobenius la compatibilidad de los siguientes sistemas:

*a*) $$\left.\begin{aligned} x+2y+2z&=3\\ y+2z&=2\\ 3x+2y&=1\\ 5x+7y+6z&=2 \end{aligned}\right\}$$ *b*) $$\left.\begin{aligned} 2x+z&=0\\ 4x-6y&=2\\ -2x+12y+z&=7 \end{aligned}\right\}$$

*c*) $$\left.\begin{aligned} 2x+y&=2\\ 2x+3z&=4\\ x+2y+5t&=0 \end{aligned}\right\}$$ *d*) $$\left.\begin{aligned} x+5y+2z&=3\\ 2y+5z+3t&=2\\ -2x+4y+2t&=1 \end{aligned}\right\}$$

**6.** Dado el sistema de ecuaciones lineales

$$\left.\begin{aligned} 3x_1+x_2-8x_3+2x_4+x_5&=0\\ 2x_1-2x_2-3x_3-7x_4+2x_5&=0\\ x_1+11x_2-12x_3+34x_4-5x_5&=0\\ x_1-5x_2+2x_2-16x_4+3x_5&=0 \end{aligned}\right\}$$

encontrar el mayor número de soluciones que sean linealmente independientes. Indicar uno de estos conjuntos linealmente independiente de soluciones.

**7.** *a*) Sea $E$ una matriz elemental de la forma

$$E=\begin{pmatrix} 1 & & & & \\ & 1 & & \text{\Large 0} & \\ & & \ddots & & \\ & & & c & \\ & \text{\Large 0} & & & \ddots \\ & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow \text{fila } i; \\ \\ \\ \end{array}$$

demostrar que el número $R$ para $EA$ coincide con el número $R$ para $A$.

*b*) Sea $E$ una matriz elemental que se ha obtenido intercambiando dos filas de la matriz identidad; demostrar que el número $R$ para $EA$ coincide con el número $R$ para $A$.

*c*) Sea $E$ la matriz elemental de la forma



$$E=\begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & \vdots & \ddots & & \\ & & c & \cdots & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \begin{array}{l} \\ \\ \longleftarrow \text{fila } i \\ \\ \longleftarrow \text{fila } j\ ; \\ \\ \\ \end{array}$$

demostrar que el número $R$ para $EA$ coincide con el número $R$ para $A$.

*Notas.* En este problema, $A$ es una matriz cuadrada. Las demostraciones de estos resultados deben realizarse sin utilizar el teorema 1.

Este problema completa la demostración del teorema 1 de esta sección.

## 2.6. DETERMINANTES Y PERMUTACIONES

Presentaremos a continuación una nueva fórmula para calcular el determinante de una matriz cuadrada. Esta nueva fórmula requiere el concepto de *permutación.*

Recibe el nombre de *permutación* de los números naturales 1, 2, ..., $n$ toda aplicación biyectiva del conjunto $\{1, 2, ..., n\}$ en sí mismo. Las permutaciones se representarán con las letras griegas $\alpha$, $\beta$, $\gamma$, ...

EJEMPLO A. Las permutaciones de los números naturales 1 y 2 son $\alpha$: $\{1, 2\}\mapsto\{1, 2\}$ dada por $\alpha(1)=1$ y $\alpha(2)=2$, y $\beta$: $\{1, 2\}\mapsto\{1, 2\}$ dada por $\beta(1)=2$ y $\beta(2)=1$.

A continuación escribimos las permutaciones de los números naturales 1, 2 y 3:

$$\begin{array}{llll} \alpha_1: & \alpha_1(1)=1, & \alpha_1(2)=2, & \alpha_1(3)=3\\ \alpha_2: & \alpha_2(1)=2, & \alpha_2(2)=1, & \alpha_2(3)=3\\ \alpha_3: & \alpha_3(1)=3, & \alpha_3(2)=1, & \alpha_3(3)=2\\ \alpha_4: & \alpha_4(1)=3, & \alpha_4(2)=2, & \alpha_4(3)=1\\ \alpha_5: & \alpha_5(1)=2, & \alpha_5(2)=3, & \alpha_5(3)=1\\ \alpha_6: & \alpha_6(1)=1, & \alpha_6(2)=3, & \alpha_6(3)=2 \end{array}$$

Al lector se le ha pedido todas estas aplicaciones en el ejercicio 2 de la sección 1.3.

* * *

En el ejemplo anterior puede observarse que para escribir una permutación no es necesario escribir los elementos del conjunto inicial; basta con escribir ordenadamente la imagen de los elementos en su orden natural. Así, por ejemplo:

$$\alpha_3=(3\quad 1\quad 2)$$

y

$$\alpha_5=(2\quad 3\quad 1)$$

donde $\alpha_3$ y $\alpha_5$ son como en el ejemplo $A$.

Si llamamos $S_n$ al conjunto de todas las permutaciones de los $n$ primeros números naturales y $\alpha \in S_n$ podemos escribir

$$\alpha = (\alpha(1) \quad \alpha(2) \quad \cdots \quad \alpha(n))$$

Dada una permutación $\alpha$ de $S_n$, diremos que en $\alpha$ se ha *realizado una inversión* si se han intercambiado entre sí dos elementos de su imagen. El resultado de realizar una inversión en una permutación es, de nuevo, una permutación.

EJEMPLO B. $\beta = (14325)$ es una permutación que se obtiene de $\alpha = (54321)$ realizando una inversión; a saber, se han invertido el 1 y el 5.

* * *

Sea $k_\alpha$ el número de inversiones que es necesario realizar en la permutación $\alpha$ para transformarla en la permutación identidad; el número $k_\alpha$ no es fijo para una permutación $\alpha \in S_n$, sino que varía con el orden en que se realicen las inversiones. Sin embargo, si por un procedimiento se ha necesitado un número par (impar) de inversiones cualquier otro procedimiento debe necesitar un número par (impar) de inversiones.

Por tanto, el número

$$\operatorname{sig}(\alpha) = (-1)^{k_\alpha}$$

que recibe el nombre de *signatura de* $\alpha$, no depende de la forma en que se realicen las inversiones. Este resultado puede encontrarse demostrado en cualquier libro de teoría de grupos.

Si $k_\alpha$ es par, la signatura de $\alpha$ es $+1$ y entonces diremos que $\alpha$ es una permutación *par;* si $k_\alpha$ es impar, la signatura de $\alpha$ es $-1$ y entonces diremos que $\alpha$ es una permutación *impar*.

EJEMPLO C. Dada la permutación $\alpha = (53142)$ podemos invertir el 1 y el 5 para obtener $\alpha_1 = (13542)$; a continuación invertimos el 2 y el 3 para obtener $\alpha_2 = (12543)$; finalmente invertimos el 3 y el 5 para obtener $\alpha_3 = (12345)$, que coincide con la permutación identidad. Por tanto,

$$\operatorname{sig}(\alpha) = (-1)^3 = -1$$

El proceso que hemos seguido para hallar la signatura de $\alpha$ puede esquematizarse de la siguiente forma:

$$\alpha = (53142) \xrightarrow{5 \leftrightarrows 1} (13542) \xrightarrow{3 \leftrightarrows 2} (12543) \xrightarrow{5 \leftrightarrows 3} (12345)$$

* * *

Pasamos a continuación a encontrar una fórmula para el determinante, que utiliza el concepto de permutación. Si comenzamos con matrices de orden 3 sabemos que

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32}$$
$$- a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32}$$



Observar que esta «suma algebraica» posee 6 elementos y que cada uno de sus términos tiene 3 elementos, de manera que cada uno de estos tres elementos proviene de distinta fila y distinta columna que los restantes. Además, el subíndice de las filas está siempre ordenado, mientras que el subíndice de las columnas forma una permutación de los elementos 1, 2 y 3. Por tanto, los 6 términos anteriores pueden escribirse de la forma

$$a_{1\alpha(1)}a_{2\alpha(2)}a_{3\alpha(3)}, \quad \alpha \in S_3$$

afectados de un signo «positivo» o «negativo». Para determinar el signo de cada uno de estos sumandos observamos que los términos positivos corresponden a las permutaciones pares $\alpha_1$, $\alpha_2$ y $\alpha_3$ del ejemplo A y los términos negativos corresponden a las permutaciones impares $\alpha_4$, $\alpha_5$ y $\alpha_6$ del ejemplo A. Por tanto:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{31} & a_{33} \end{vmatrix} = \sum_{\alpha \in S_3} (\operatorname{sig}(\alpha)) a_{1\alpha(1)}a_{2\alpha(2)}a_{3\alpha(3)}$$

Este resultado sugiere el siguiente:

TEOREMA 1

Sea $A = (a_{ij})_{i,j=1,\dots,n}$ una matriz de orden $n$; si definimos

$$\Delta(A) = \sum_{\alpha \in S_n} (\operatorname{sig}(\alpha)) a_{1\alpha(1)}a_{2\alpha(2)} \cdots a_{n\alpha(n)}$$

se tiene que $|A| = \Delta(A)$.

Antes de demostrar el teorema 1 necesitamos los siguientes resultados:

LEMA 2

Si $A$ es una matriz $(a_{ij})_{i,j=1,\dots,n}$ en la que la $j$-ésima fila es una suma de la forma $a_{ij} = b_{ij} + b'_{ij}$, $j = 1, 2, \dots, n$ y escribimos

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \quad ; \quad B' = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b'_{i1} & b'_{i2} & \cdots & b'_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

se tiene que

$$\Delta(A) = \Delta(B) + \Delta(B')$$

*Demostración.* Por la definición de $A$ se tiene que

$$\begin{aligned}
\Delta(A) &= \sum_{\alpha\in S_n} (\mathrm{sig}\,(\alpha)) a_{1\alpha(1)} \cdots a_{i\alpha(i)} \cdots a_{n\alpha(n)} = \\
&= \sum_{\alpha\in S_n} (\mathrm{sig}\,(\alpha)) a_{1\alpha(1)} \cdots [b_{i\alpha(i)} + b'_{i\alpha(i)}] \cdots a_{n\alpha(n)} = \\
&= \sum_{\alpha\in S_n} (\mathrm{sig}\,(\alpha)) a_{1\alpha(1)} \cdots b_{i\alpha(i)} \cdots a_{n\alpha(n)} + \sum_{\alpha\in S_n} (\mathrm{sig}\,(\alpha)) a_{1\alpha(1)} \cdots b'_{i\alpha(i)} \cdots a_{n\alpha(n)} = \\
&= \Delta(B) + \Delta(B')
\end{aligned}$$

■

LEMA 3

Si $A=(a_{ij})_{i,j=1,\dots,n}$ y $A_j$ denota la matriz

$$A_j = \begin{pmatrix} a_{11} & \cdots & a_{ij} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{nj} & \cdots & 0 \end{pmatrix}$$

que se obtiene de $A$ sustituyendo todos los elementos de la última fila por cero excepto el que ocupa la columna $j$, se tiene que

$$\Delta(A_j) = (-1)^{n+j} a_{nj} \Delta(A_{nj})$$

donde $A_{nj}$ denota la matriz adjunta del elemento que ocupa el lugar $(n, j)$.

*Demostración.* Comenzamos observando que si $B$ es la matriz que se obtiene de $A$ intercambiando dos columnas se tiene que $\Delta(B) = -\Delta(A)$; la demostración de este resultado se pide en el ejercicio 3 al final de esta sección.

Intercambiamos la columna $j$ de $A_j$ con todas las columnas posteriores a ella hasta llevarla a la última columna de la matriz; se obtiene así la matriz

$$B_j = \begin{pmatrix} b_{11} & \cdots & b_{1,j-1} & b_{1j} & \cdots & b_{1,n-1} & b_{1n} \\ b_{21} & \cdots & b_{2,j-1} & b_{2j} & \cdots & b_{2,n-1} & b_{2n} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & b_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} & a_{ij} \\ a_{21} & \cdots & a_{2,j-1} & a_{2,j+1} & \cdots & a_{2,n} & a_{2j} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 & a_{nj} \end{pmatrix}$$



Puesto que la columna $j$ se ha intercambiado $n-j$ veces para llevarla a la última columna, se tiene que

$$\Delta(B_j) = (-1)^{n-j} \Delta(A_j)$$

Observando que $(-1)^{-n} = (-1)^n$ se tiene que

$$\Delta(A_j) = (-1)^{-n+j} \Delta(B_j) = (-1)^{n+j} \Delta(B_j) \qquad (*)$$

Para calcular $\Delta(B_j)$ utilizamos la definición de $\Delta$ y se obtiene:

$$\Delta(B_j) = \sum_{\substack{\alpha\in S_n \\ \alpha(n)=n}} (\mathrm{sig}\,\alpha) b_{1\alpha(1)} \cdots b_{n-1,\alpha(n-1)} \cdot b_{n\alpha(n)}$$

ya que si $\alpha(n) \neq n$, $b_{n\alpha(n)} = 0$. Toda permutación $\alpha \in S_n$ que satisface $\alpha(n) = n$ puede interpretarse como una permutación $\beta \in S_{n-1}$ sin más que definir $\beta(k) = \alpha(k)$, $1 \leqq k \leqq n-1$; recíprocamente, si $\beta \in S_{n-1}$ podemos definir $\alpha \in S_n$ tal que $\alpha(k)$, $1 \leqq k \leqq n-1$ y $\alpha(n) = n$; además, $\mathrm{sig}\,(\alpha) = \mathrm{sig}\,(\beta)$. Por tanto:

$$\Delta(B_j) = b_{nn} \left[ \sum_{\beta\in S_{n-1}} (\mathrm{sig}\,\beta) b_{1\alpha(1)} \cdots b_{n-1,\beta(n-1)} \right] = b_{nn} \Delta(B_{nn})$$

Puesto que $b_{nn} = a_{nj}$ y $B_{nn} = A_{nj}$, se tiene que

$$\Delta(B_j) = a_{nj} \Delta(A_{nj})$$

Sustituyendo este resultado en la fórmula (*) se obtiene de resultado deseado. ■

*Demostración del teorema 1.* Realizaremos la demostración mediante el método de inducción. Si $A$ es de orden 3 ya hemos comprobado anteriormente que el resultado es cierto; el lector puede comprobar que el resultado es también cierto si $A$ es de orden 2.

Supongamos que el teorema es cierto para toda matriz de orden inferior a $n$ y sea $A = (a_{ij})_{i,j=1,\dots,n}$ una matriz de orden $n$. La matriz $A$ puede escribirse de la forma

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} + 0 + \cdots + 0 & 0 + a_{n2} + \cdots + 0 & \cdots & 0 + 0 + \cdots + a_{nn} \end{pmatrix}$$

Aplicando repetidamente el lema 2 se tiene que

$$\Delta(A) = \Delta(A_1) + \Delta(A_2) + \cdots + \Delta(A_j) + \cdots + \Delta(A_n)$$

donde $A_j$, $j = 1, 2, \dots, n$, es como en el enunciado del lema. Por el lema 3 se tiene que

$$\Delta(A) = (-1)^{n+1} a_{n1} \Delta(A_{n1}) + \cdots + (-1)^{n+j} a_{nj} \Delta(A_{nj}) + \cdots + (-1)^{2n} a_{nn} \Delta(A_{nn})$$

Puesto que las matrices $A_{nj}$, $j=1, 2, ..., n$, son de orden $n-1$, por la hipótesis de inducción se tiene que

$$\Delta(A)=(-1)^{n+1}a_{n1}|A_{n1}|+\cdots+(-1)^{n+j}a_{nj}|A_{nj}|+\cdots+(-1)^{2n}a_{nn}|A_{nn}|=|A|$$

donde la última igualdad se debe al teorema 3 de la sección 2.2 (desarrollo del determinante por la última fila). Esto termina la demostración. ■

La fórmula del determinante resulta engorrosa cuando se trata de matrices de orden superior a tres si queremos utilizarla en el cálculo práctico. Por otro lado, es bastante eficiente para realizar algunas de las demostraciones de las propiedades de los determinantes, que se expusieron en la sección 2 de este mismo capítulo.

Por ejemplo, la propiedad 1, que expresa que si una matriz tiene una fila (columna) de ceros su determinante es nulo, es fácil de demostrar. Basta observar que en todos los términos de $\Delta(A)$ aparece un elemento nulo y, por tanto, $|A|=\Delta(A)=0$. Las demostraciones de otras propiedades se piden en los ejercicios al final de esta sección.

## EJERCICIOS 2.6

**1.** Dadas las permutaciones

$$\alpha=(23154) \quad , \quad \beta=(13245) \quad , \quad \gamma=(53412)$$
$$\lambda=(31856472) \quad , \quad \delta=(4321)$$

calcular la signatura de cada una de ellas y la signatura de $\beta\circ\alpha$, $\beta^{-1}$ y $\gamma\circ\beta^{-1}$.

**2.** Indicar el signo de los términos $a_{12}a_{23}a_{31}a_{44}$ y $a_{14}a_{21}a_{32}a_{43}$ en el desarrollo del determinante de una matriz de orden 4.

**3.** *a*) Demostrar que si $\beta$ es la permutación que se obtiene de $\alpha$ intercambiando $\alpha(i)$ y $\alpha(k)$, $i\neq k$, se tiene que $\mathrm{sig}(\beta)=-\mathrm{sig}(\alpha)$.
*b*) Demostrar que si $B$ es la matriz que se obtiene de $A$ intercambiando dos columnas se tiene que $\Delta(B)=-\Delta(A)$.

**4.** Demostrar que si $B$ es la matriz que se obtiene de $A$ multiplicando una de sus filas por $r$ se tiene que $\Delta(B)=r\Delta(A)$.

**5.** Utilizar el problema 3 para demostrar que si $A$ es una matriz con dos filas idénticas se tiene $\Delta(A)=0$.

CAPITULO 3

# LA GEOMETRIA DEL PLANO Y DEL ESPACIO

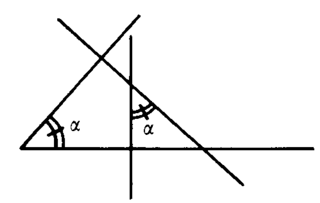


La geometría es la rama de la matemática que estudia la forma y el tamaño de las figuras, así como las trasformaciones que sobre ellas se ejercen. Dependiendo de que las figuras estén en un plano o en el espacio se obtienen las geometrías descritas en el título de este capítulo.

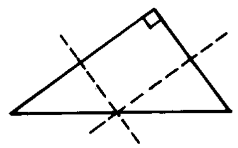
La antigua civilización griega poseía muchos conocimientos acerca de la geometría. Estos conocimientos fueron recogidos por uno de sus mejores exponentes, Euclides, que los recopiló en un libro denominado *Los elementos.* La traducción de este libro, que llegó a Europa a través de la civilización árabe, ha sido la base de todo el estudio de la geometría hasta finales del siglo XIX. *Los elementos* están basados en un sistema de verdades evidentes, denominadas *axiomas,* a partir de las cuales se deducen las propiedades de las figuras mediante un razonamiento lógico.

Ejemplos de propiedades de las figuras son los siguientes:

1) Los ángulos formados por rectas perpendiculares entre sí son iguales.



2) Las mediatrices de los catetos de un triángulo rectángulo se cortan en el punto medio de la hipotenusa.



En el siglo XVII el filósofo y matemático francés R. Descartes introdujo la noción de coordenada de un punto; todo punto tiene una representación con respecto a unas rectas dadas que se cortan. Los trabajos de los matemáticos durante los dos siglos siguientes mostraron que las propiedades geométricas de las figuras pueden demostrarse más fácilmente utilizando el sistema de representación mediante coordenadas cartesianas. Esta forma de estudiar la geometría se denomina *geometría analítica*.

La geometría analítica sustituyó a la geometría de *Los elementos* de Euclides a finales del siglo XIX y actualmente es la forma más extendida de estudiar la geometría.

La geometría analítica del plano y del espacio ocupará gran parte de nuestra exposición en este capítulo; dedicaremos también alguna sección a estudiar propiedades de las figuras desde el punto de vista euclídeo.

## 3.1. RECTAS EN UN PLANO

Dadas dos rectas en un plano que se cortan en un punto $O$, cualquier otro punto del plano queda determinado por dos números que reciben el nombre de *componentes* con respecto a las rectas dadas. Estas componentes se obtienen de la siguiente manera. Fijada una unidad de medida $u$ en cada una de las rectas $l_1$, $l_2$, todo punto $P$ del plano tiene como componentes $x$ e $y$, donde $x$ es la distancia, medida con la unidad $u$, del punto $P$ a la recta $l_2$ siguiendo una paralela a $l_1$, e $y$ es la distancia del punto $P$ a la recta $l_1$ siguiendo una paralela a $l_2$, midiendo la distancia con respecto a la unidad de medida $u$ (figura 1).

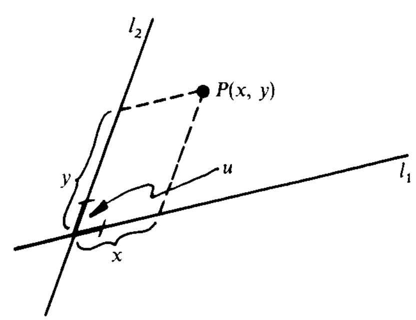



Fig. 1



Si las rectas $l_1$ y $l_2$ son perpendiculares, diremos que tenemos un sistema de coordenadas *rectangulares* o *cartesianas*. La primera componente se denomina *abscisa* del punto y la segunda *ordenada*. De acuerdo con esta nomenclatura la recta $l_1$ se denomina *eje de abscisas* y $l_2$ se denomina *eje de ordenadas*, escribiéndose también por $OX$ y $OY$, respectivamente. A la izquierda de $O$ en la recta $l_1$ y por debajo de $O$ en la recta $l_2$ se toman medidas negativas. El punto de intersección $O$ se denomina *origen de coordenadas* (figura 2)

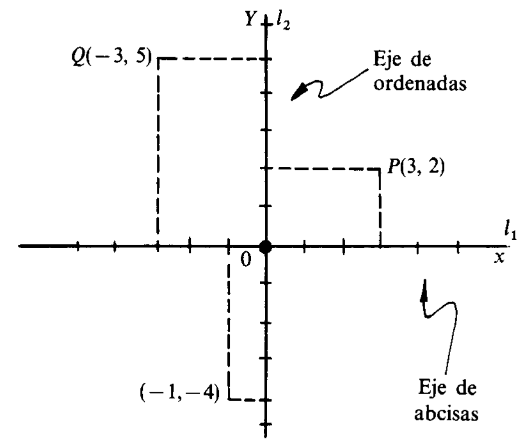



Figura 2

Un plano dotado de un sistema de coordenadas se designa por $\mathbb{R}^2$.

Otra forma de tratar a los puntos de $\mathbb{R}^2$ es considerarlos como el extremo de un vector cuyo origen es el origen de coordenadas (véase la figura 3).

El álgebra de los vectores se ha estudiado en la sección 1.2 del capítulo 1. La suma de los vectores $\overline{v}_1$ y $\overline{v}_2$ es la diagonal del paralelogramo que tiene como lados los vectores $\overline{v}_1$ y $\overline{v}_2$ (véase figura 4.*a*)). El producto de un vector $\overline{v}$ por un número real $c$ es

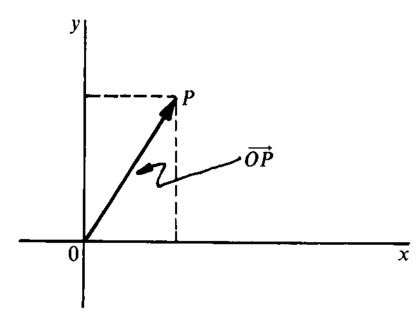



Fig. 3

otro vector cuyo extremo es el punto que tiene como componentes $c$ veces las componentes de $\vec{v}$; el vector $c\vec{v}$ tiene el mismo sentido que $\vec{v}$ *si* $c$ es positivo y distinto sentido si $c$ es negativo (ver figura 4.*b*)).

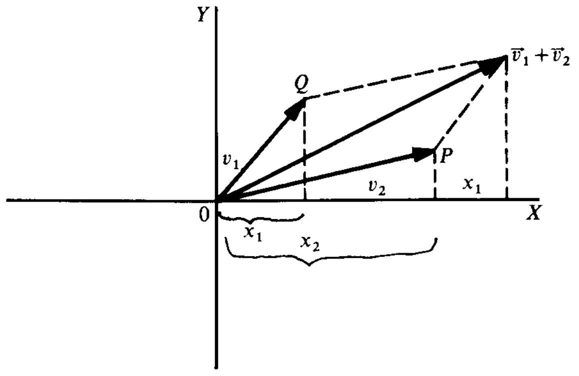



Fig. 4.*a*)

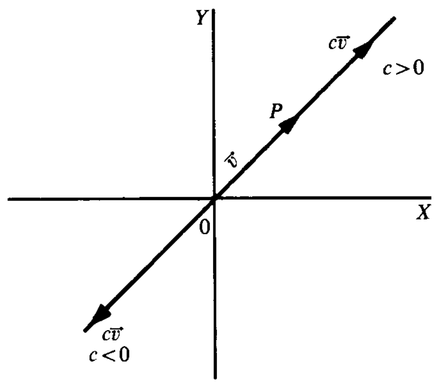



Fig. 4.*b*)

Dados dos puntos $P=(x_1, y_1)$ y $Q=(x_2, y_2)$, se denomina vector $\overrightarrow{PQ}$, al vector con origen en $P$ y extremo en $Q$; sus componentes son

$$\overrightarrow{PQ}=(x_2-x_1, y_2-y_1)$$

y, por tanto, es equivalente al vector que tiene como origen $O$ y extremo un punto $R$ de coordenadas $(x_2-x_1, y_2-y_1)$.

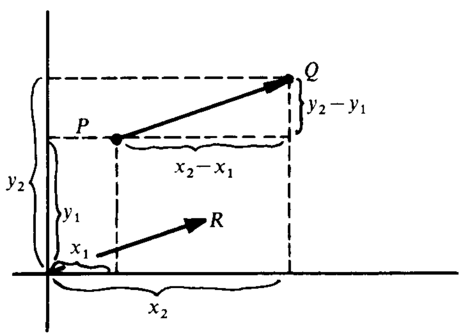


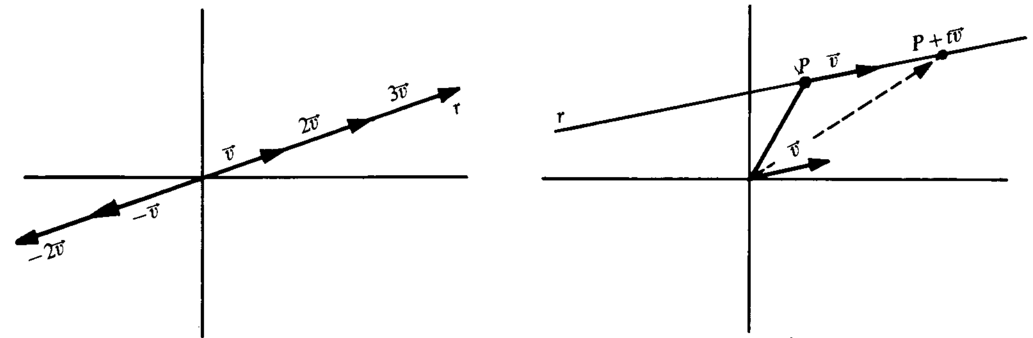

Fig. 5

Dado un vector $\vec{v}$, el conjunto de puntos $t\vec{v}$, donde $t$ es un número real, determina la recta que pasa por el origen de coordenadas y tiene a $\vec{v}$ como vector director (ver figura 6.*a*)).

La recta $r$ que pasa por el punto $P$ y tiene a $\vec{v}$ como vector director se representa mediante el conjunto de puntos

$$P+t\vec{v}$$






*a*) Fig. 6 *b*)

donde $t$ es un número real. Si $P=(p_1, p_2)$ y $\vec{v}=(v_1, v_2)$ cualquier punto $(x, y)$ de la recta $r$ se escribe de la forma

$$(x, y)=(p_1, p_2)+t(v_1, v_2).$$

Igualando componentes se tiene:

$$\left.\begin{aligned} x&=p_1+tv_1\\ y&=p_2+tv_2 \end{aligned}\right\}$$

que se denominan *ecuaciones paramétricas* de la recta $r$.

EJEMPLO A. Para encontrar las ecuaciones paramétricas de la recta que pasa por los puntos $P=(2, 1)$ y $Q=(-1, 3)$ observamos que uno de sus vectores directores es $\overrightarrow{PQ}=(-3, 2)$ y puesto que pasa por el punto $P=(2, 1)$ se tiene

$$(x, y)=(2, 1)+t(-3, 2)$$

o bien

$$\left.\begin{aligned} x&=2-3t\\ y&=1+2t \end{aligned}\right\}$$

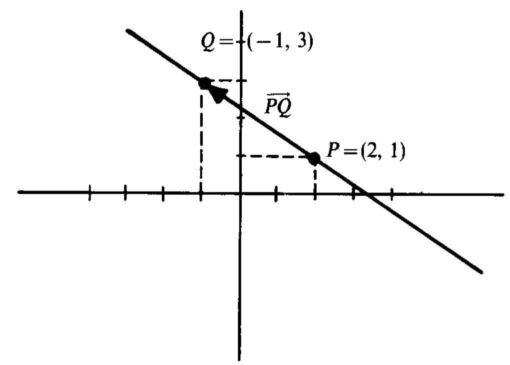



Fig. 7

*Nota.* La representación paramétrica de una recta no es única; en el ejemplo anterior podríamos tomar $\frac{1}{2}\vec{PQ}=\vec{v}$ como vector director y $Q$ como uno de los puntos de la recta; se obtendrían en este caso las ecuaciones

$$(x, y)=(-1, 3)+t\left(-\frac{3}{2}, 1\right)$$

que representan la misma recta.

* * *

En la recta de ecuaciones paramétricas

$$\left.\begin{aligned} x&=p_1+tv_1\\ y&=p_2+tv_2\end{aligned}\right\}$$

podemos eliminar la variable real $t$ despejándola en cada una de las ecuaciones e igualando los resultados; se obtiene

$$t=\frac{x-p_1}{v_1} \quad , \quad t=\frac{y-p_2}{v_2}$$

y, por tanto:

$$\frac{x-p_1}{v_1}=\frac{y-p_2}{v_2}$$

siempre que $v_1$ y $v_2$ sean no nulos. La igualdad anterior puede escribirse de la forma

$$v_2x-v_1y+v_1p_2-v_2p_1=0$$

o bien

$$ax+by+c=0,$$

que se denomina *ecuación general de la recta* en coordenadas cartesianas.

Si $v_1=0$, las ecuaciones paramétricas de una recta se escriben de la forma

$$\left.\begin{aligned} x&=p_1\\ y&=p_2+tv_1\end{aligned}\right\}.$$

Puesto que de la primera de las ecuaciones ya se ha eliminado el parámetro $t$, la ecuación general de esta recta es

$$x=p_1$$



Si $v_2=0$, la ecuación general de la recta es $x=p_2$. Estos casos particulares corresponden a rectas paralelas a los ejes de coordenadas (ver figura 8).

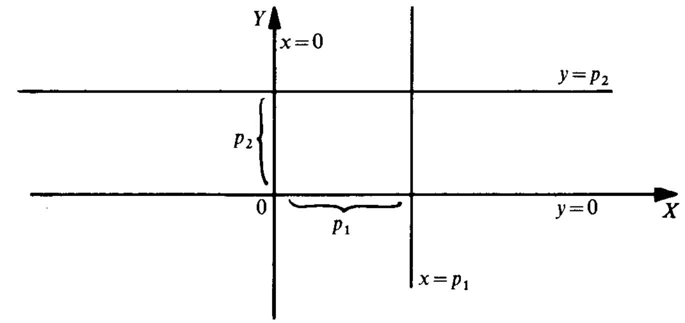



Figura 8

EJEMPLO B. Para encontrar las ecuaciones cartesianas de la recta que pasa por el punto $P=(2, 1)$ y tiene como vector director $(-1, 5)$ escribimos sus ecuaciones paramétricas

$$\left.\begin{aligned} x&=2-t\\ y&=1+5t\end{aligned}\right\} \qquad \text{(I)}$$

y eliminamos $t$:

$$t=\frac{x-2}{-1}=\frac{y-1}{5}.$$

Por tanto:

$$5x+y-11=0$$

Otra forma de eliminar $t$ en (I) es considerar que tenemos un sistema de dos ecuaciones con una incógnita $t$, mientras que $x$ e $y$ son constantes, es decir, el sistema

$$\left.\begin{aligned} -t&=x-2\\ 5t&=y-1\end{aligned}\right\}.$$

Puesto que fijados $x$ e $y$ este sistema tiene una sola solución en $t$, del teorema de Rouche-Frobenius se deduce que

$$r\begin{pmatrix}-1\\ 5\end{pmatrix}=r\begin{pmatrix}-1 & x-2\\ 5 & y-1\end{pmatrix}$$

Puesto que $r\begin{pmatrix}-1\\5\end{pmatrix}=1$ hemos de tener que $r\begin{pmatrix}-1 & x-2\\5 & y-1\end{pmatrix}=1$, y esto es equivalente a

$$\begin{vmatrix}-1 & x-2\\5 & y-1\end{vmatrix}=0.$$

Desarrollando este determinante se obtiene

$$-y+1-5x+10=0$$

que es el mismo resultado que obtuvimos anteriormente.

* * *

Dadas dos rectas en un plano pueden darse los siguientes casos:

*a*) Las dos rectas se cortan en un solo punto.
*b*) Las dos rectas son paralelas.
*c*) Las dos rectas coinciden.

Si se conocen las ecuaciones de las rectas en coordenadas cartesianas podemos determinar su posición relativa, es decir, cada uno de los casos anteriores, estudiando las soluciones del sistema formado por las dos ecuaciones. Si las ecuaciones son

$$\left.\begin{matrix}r_1\equiv a_1x+b_1y+c_1=0\\ r_2\equiv a_2x+b_2y+c_2=0\end{matrix}\right\} \qquad \text{(II)}$$

tendremos el caso *a*) si el sistema (II) posee una sola solución; por el teorema de Rouché-Frobenius esto sucede solamente cuando

$$\begin{vmatrix}a_1 & b_1\\a_2 & b_2\end{vmatrix}\neq 0$$

En el caso *b*) el sistema (II) no posee solución, y, por tanto, se ha de tener

$$r\begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix}\neq r\begin{pmatrix}a_1 & b_1 & -c_1\\a_2 & b_2 & -c_2\end{pmatrix}$$

Finalmente, si las rectas coinciden el sistema (II) posee infinitas soluciones y, por tanto, se ha de tener que

$$r\begin{pmatrix}a_1 & b_1\\a_2 & b_2\end{pmatrix}=r\begin{pmatrix}a_1 & b_1 & -c_1\\a_2 & b_2 & -c_2\end{pmatrix}<2$$

En los ejercicios no es conveniente memorizar los resultados obtenidos, sino obtenerlos en cada caso mediante la utilización del teorema de Rouché-Frobenius.



EJEMPLO C. Las rectas de ecuaciones $3x+2y-7=0$ y $6x+4y+1=0$ son paralelas, ya que

$$r\begin{pmatrix}3 & 2\\6 & 4\end{pmatrix}=1 \quad \text{y} \quad r\begin{pmatrix}3 & 2 & 7\\6 & 4 & -1\end{pmatrix}=2$$

* * *

Si las rectas están dadas en coordenadas paramétricas podemos estudiar su posición relativa pasándolas a coordenadas cartesianas. Sin embargo, también puede conocerse su posición a partir de sus coordenadas paramétricas, como se muestra en el siguiente ejemplo.

EJEMPLO D. Dadas las rectas de ecuaciones

$$r_1\colon (x, y)=(1, 0)+t(-1, 2)$$
$$r_2\colon (x, y)=(-2, 1)+s(-1, -3)$$

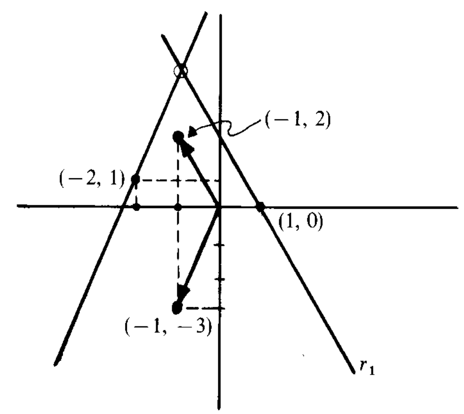



Figura 9

ambas se cortan si y sólo si existe un único valor de $t$ y un único valor de $s$ para los cuales se obtiene el mismo resultado de $x$ e $y$. Por tanto, el sistema

$$(1, 0)+t(-1, 2)=(-2, 1)+s(-1, -3)$$

tiene solución única en $t$ y $s$. Igualando componentes se tiene

$$\left.\begin{matrix}1-t=-2-s\\ 2t=1-3s\end{matrix}\right\} \Leftrightarrow \left.\begin{matrix}-t+s=-3\\ 2t+3s=1\end{matrix}\right\}$$

Este sistema posee solución única, ya que

$$\begin{vmatrix} -1 & 1 \\ 2 & 3 \end{vmatrix} = -3-2=-5\neq 0$$

Por tanto, las rectas dadas se cortan en un único punto. Para encontrar las coordenadas del punto de corte encontramos los valores de $t$ y $s$ que satisfacen el sistema anterior y sustituimos en las ecuaciones de las rectas:

$$t=\frac{\begin{vmatrix} -3 & 1 \\ 1 & 3 \end{vmatrix}}{-5}=\frac{-10}{-5}=2 \quad ; \quad s=\frac{\begin{vmatrix} -1 & -3 \\ 2 & 1 \end{vmatrix}}{-5}=\frac{5}{-5}=-1$$

$$(1, 0)+2(-1, 2)=(-1, 4) \quad ; \quad (-2, 1)+(-1)(-1, -3)=(-1, 4)$$

Se trata, por tanto, del punto $(-1, 4)$. Observar que solamente es necesario calcular el valor de $t$ o el de $s$, pero no ambos, para obtener el punto de corte. El hallar los dos valores sirve de adecuada comprobación.

* * *

Supongamos que las rectas de ecuaciones

$$P+t\vec{u}=(p_1, p_2)+t(u_1, u_2)$$
$$Q+s\vec{v}=(q_1, q_2)+s(v_1, v_2)$$

son paralelas o coincidentes. En el primer caso el sistema

$$\left.\begin{aligned} p_1+tu_1&=q_1+sv_1 \\ p_2+tu_2&=q_2+sv_2 \end{aligned}\right\} \qquad \text{(III)}$$

no posee soluciones y en el segundo posee infinitas soluciones. En cualquier caso, aplicando el teorema de Rouché-Frobenius el rango de la matriz de los coeficientes de sus incógnitas ha de ser 1 y, por tanto, se tiene

$$\begin{vmatrix} u_1 & -v_1 \\ u_2 & -v_2 \end{vmatrix}=0 \Leftrightarrow \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}=0$$

Puesto que los vectores $\vec{u}$ y $\vec{v}$ son no nulos, por el teorema del menor básico (teorema 2 de la sección 2.5) se obtiene que $\vec{u}$ es una combinación lineal de $\vec{v}$, es decir:

$$\vec{u}=c\vec{v}$$

Por tanto, los vectores $\vec{u}$ y $\vec{v}$ son proporcionales.

Recíprocamente, si $\vec{u}$ y $\vec{v}$ son dos vectores proporcionales, las rectas que los tienen como vectores directores son paralelas o coincidentes.



EJEMPLO E. Las rectas $(1, 2)+t(-1, 4)$ y $(-1, 0)+s(2, -8)$ son paralelas o coincidentes, ya que $(2, -8)=-2(-1, 4)$. Para determinar si son paralelas o coincidentes escribimos el sistema que se obtiene de igualar las dos ecuaciones:

$$\left.\begin{aligned} 1-t&=-1+2s \\ 2+4t&=-8s \end{aligned}\right\} \quad ; \quad \left.\begin{aligned} -t-2s&=-2 \\ 4t+8s&=-2 \end{aligned}\right\}.$$

Puesto que

$$\begin{vmatrix} -1 & -2 \\ 4 & 8 \end{vmatrix}=0$$

tenemos que

$$r\begin{pmatrix} -1 & -2 \\ 4 & 8 \end{pmatrix}=1,$$

y puesto que

$$\begin{vmatrix} -1 & -2 \\ 4 & -2 \end{vmatrix}=2+8=10\neq 0$$

se tiene que

$$r\begin{pmatrix} -1 & -2 & -2 \\ 4 & 8 & -2 \end{pmatrix}=2.$$

Las rectas, por tanto, son paralelas.

* * *

Para terminar esta sección observamos que la ecuacion

$$\begin{vmatrix} 1 & x & y \\ 1 & a_1 & a_2 \\ 1 & b_1 & b_2 \end{vmatrix}=0$$

determina una recta que pasa por los puntos $A=(a_1, a_2)$ y $B=(b_1, b_2)$. Para probar este resultado observar que desarrollando el determinante por la primera fila se tiene:

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} - x\begin{vmatrix} 1 & a_2 \\ 1 & b_2 \end{vmatrix} + y\begin{vmatrix} 1 & a_1 \\ 1 & b_1 \end{vmatrix}=0$$

que es la forma general de la ecuación de una recta en coordenadas cartesianas; además, la recta pasa por los puntos $A=(a_1, a_2)$ y $B=(b_1, b_2)$, ya que al sustituir $x$ e $y$

por los correspondientes valores de $A$ y $B$ se obtiene un determinante con dos filas iguales.

EJERCICIOS 3.1

**1.** Dados $P=(1, 0)$, $Q=(-1, 2)$, $\vec{u}=(1, -1)$ y $\vec{v}=(3, -2)$, hallar las ecuaciones paramétricas y cartesianas de las siguientes rectas, y dibujarlas:

*a)* Recta que pasa por $P$ y tiene como vector director $2\vec{u}$.
*b)* Recta que pasa por $P$ y tiene como vector director $\vec{u}+\vec{v}$.
*c)* Recta que pasa por $Q$ y tiene como vector director $\vec{u}-2\vec{v}$.

**2.** Demostrar que la ecuación de una recta que pasa por los puntos $P=(p_1, p_2)$ y $Q=(q_1, q_2)$ puede escribirse de la forma

$$y-p_2=\frac{q_2-p_2}{q_1-p_1}(x-p_1), \quad \text{si} \quad q_1-p_1\neq 0$$

[*Nota*: El número $(q_2-p_2)/(q_1-p_1)$ se denomina *pendiente* de la recta.]

**3.** Determinar la posición relativa de los siguientes pares de rectas y encontrar el punto de intersección si se cortan:

*a)* $(x, y)=(2, 1)+t(1, 1)$ y $(x, y)=(1, 0)+s(-5, -5)$
*b)* $x+y=1$ y $2x-y=2$
*c)* $2x-y=4$ y $(x, y)=(2, 0)+t(-2, -4)$
*d)* $(x, y)=t(-1, 2)$ y $(x, y)=(1, 0)+s(2, 3)$

**4.** Determinar ecuaciones paramétricas de las siguientes rectas, y dibujarlas:

*a)* $x-2y=5$ *b)* $2x-y=3$ *c)* $x=3$ *d)* $y=0$

[*Nota:* El resultado no es necesariamente único.]

**5.** Dada la recta de ecuación $ax+by+c=0$, demostrar que todas las rectas paralelas a ella tienen una ecuación de la forma $ax+by+c'=0$ con $c'\in\mathbb{R}$. [*Nota:* El conjunto de las rectas paralelas a una dada se denomina *haz de rectas* con la propiedad citada.] Encontrar la ecuación del haz de rectas paralelas a la recta que pasa por $(1, 2)$ y tiene $(-1, 1)$ como vector director.

**6.** Encontrar las ecuaciones paramétricas y cartesianas de las siguientes rectas, comprobando el resultado si se cree posible:

*a)* Paralela a $(x, y)=(3, 3)+t(2, 1)$ que pasa por $(1, 0)$.
*b)* Paralela a $2x-y=5$ que pasa por $(1, -2)$.
*c)* Paralela por el punto $(-2, -3)$ a la recta que pasa por $(1, 0)$ y $(0, 1)$.



**7.** Demostrar que la recta que pasa por los puntos $P$ y $Q$ tiene como ecuaciones paramétricas

$$(1-t)P+tQ$$

o bien

$$sP+tQ \quad \text{con} \quad s+t=1$$

**8.** Dadas las rectas secantes de ecuaciones

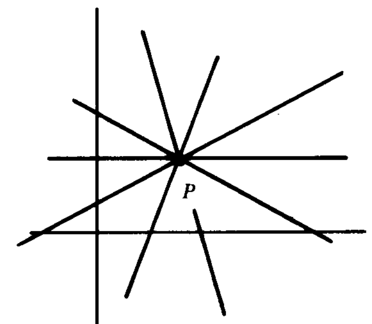
$$a_1x+b_1y+c_1=0$$
$$a_2x+b_2y+c_2=0$$

comprobar que para cada par de valores $t$, $s$ la ecuación

$$t(a_1x+b_1y+c_1)+s(a_2x+b_2y+c_2)=0 \quad (*)$$

es una recta que pasa por el punto $P$ de intersección de las anteriores y que todas ellas se obtienen así.

[*Nota:* (*) se denomina la ecuación del *haz de rectas* que pasan por $P$.]



Haz de rectas por $P$

**9.** Utilizar el problema anterior para encontrar:

*a)* La recta que pasa por $(1, 1)$ y por la intersección de $x+y=2$, $2x+y=5$.
*b)* Ecuaciones del haz de rectas que pasan por el punto $(1, 4)$.
*c)* Ecuaciones del haz de rectas que pasan por el punto $(0, 0)$.

## 3.2. RECTAS Y PLANOS EN EL ESPACIO

De manera análoga a como se hizo en el plano puede introducirse en el espacio un sistema de coordenadas de manera que todo punto del espacio quede determinado por tres números, que se denominan sus *componentes*. Basta para ello considerar tres rectas que se cortan en un solo punto, que será el *origen*, y cualquier otro punto queda determinado por su «distancia» a cada uno de los planos que determinan dos de las rectas dadas tomada paralelamente a la tercera (ver figuras 1 y 2).

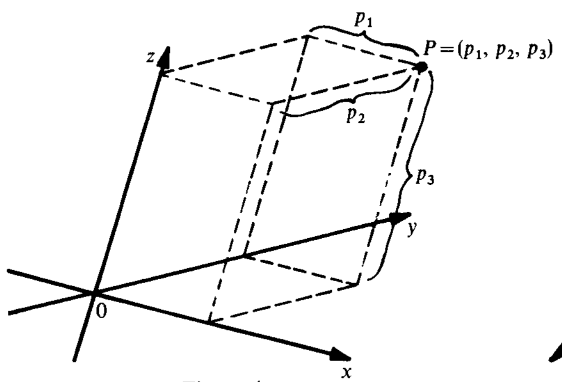


Figura 1

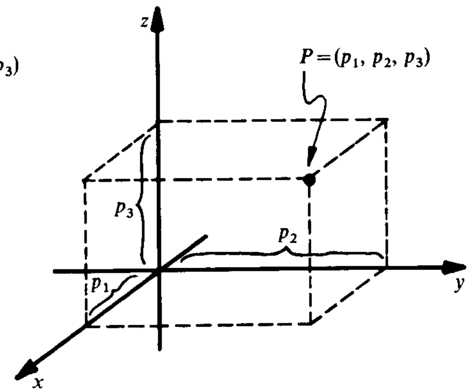


Figura 2

Si las rectas dadas, que se denominan *ejes de coordenadas*, son perpendiculares, tenemos un sistema de coordenadas rectangulares o cartesiano. Cada punto del espacio queda determinado por sus tres componentes, que se denotan por $(x, y, z)$ o $(x_1, x_2, x_3)$.

Denotaremos por $\mathbb{R}^3$ al espacio con un sistema de coordenadas rectangulares. Todo punto $P$ de $\mathbb{R}^3$ puede interpretarse como un vector que tiene como origen el origen de coordenadas y como extremo el punto $P$. La suma de vectores y el producto de un vector por un número real tienen en $\mathbb{R}^3$ la misma interpretación geométrica que en el plano (ver sección 3.1).

Dados un punto $P \in \mathbb{R}^3$ y un vector $\overline{v} \in \mathbb{R}^3$, el conjunto de puntos que se representa de la forma

$$P + t\overline{v}$$

donde $t$ es un número real, determina la *recta que pasa por P en la dirección de* $\overline{v}$. Si $P = (p_1, p_2, p_3)$ y $\overline{v} = (v_1, v_2, v_3)$ podemos escribir

$$\left.\begin{aligned} x &= p_1 + tv_1 \\ y &= p_2 + tv_2 \\ z &= p_3 + tv_3 \end{aligned}\right\} \qquad (1)$$

EJEMPLO A. La recta que pasa por el punto $P = (1, 2, 3)$ y tiene a $\overline{v} = (-1, 1, 2)$ como vector director tiene como ecuaciones paramétricas

$$\left.\begin{aligned} x &= 1 - t \\ y &= 2 + t \\ z &= 3 + 2t \end{aligned}\right\}$$



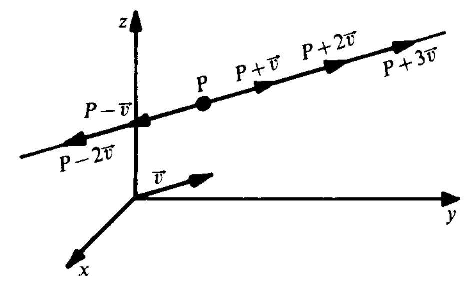


Figura 3

* * *

Dados dos puntos $P$ y $Q$ en $\mathbb{R}^3$, la ecuación de la recta que pasa por $P$ y $Q$ tiene como ecuaciones paramétricas

$$P + t(Q - P)$$

ya que uno de sus vectores de dirección es $\overline{v} = Q - P$. Por tanto, la ecuación paramétrica de una recta que pasa por $P$ y $Q$ puede escribirse de la forma

$$(1-t)P + tQ$$

o bien

$$sP + tQ \quad \text{con} \quad s + t = 1.$$

Dadas dos rectas en el espacio, existen cuatro posiciones en las cuales pueden encontrarse (ver figura 4). O bien se cortan en un solo punto, o se cruzan sin cortarse, o son paralelas en el sentido de que sus vectores directores son proporcionales o son coincidentes.

La posición relativa de dos rectas en el espacio puede determinarse estudiando el sistema formado al igualar las correspondientes coordenadas de las dos rectas dadas y determinando si sus vectores directores son o no proporcionales en los casos segundo y tercero.

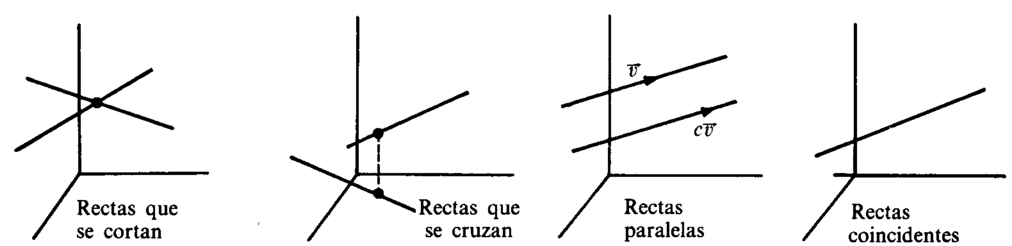


Figura 4

EJEMPLO B. Para determinar la posición relativa de las rectas

$$(x, y, z)=(1, 0, 1)+t(-1, 1, 0)$$

y

$$(x, y, z)=(0, 1, 2)+s(2, 0, 1)$$

igualamos las correspondientes coordenadas para obtener el sistema

$$(1, 0, 1)+t(-1, 1, 0)=(0, 1, 2)+s(2, 0, 1)$$

o equivalentemente:

$$\left.\begin{aligned}-t-2s&=-1\\ t&=1\\ -s&=1\end{aligned}\right\}$$

Puesto que

$$r\begin{pmatrix}-1 & -2\\ 1 & 0\\ 0 & -1\end{pmatrix}=2$$

y

$$r\begin{pmatrix}-1 & -2 & -1\\ 1 & 0 & 1\\ 0 & -1 & 1\end{pmatrix}=3$$

ya que

$$\begin{vmatrix}-1 & -2 & -1\\ 1 & 0 & 1\\ 0 & -1 & 1\end{vmatrix}=\begin{vmatrix}-1 & -2 & -1\\ 0 & -2 & 0\\ 0 & -1 & 1\end{vmatrix}=-1\begin{vmatrix}-2 & 0\\ -1 & 1\end{vmatrix}=2\neq 0$$

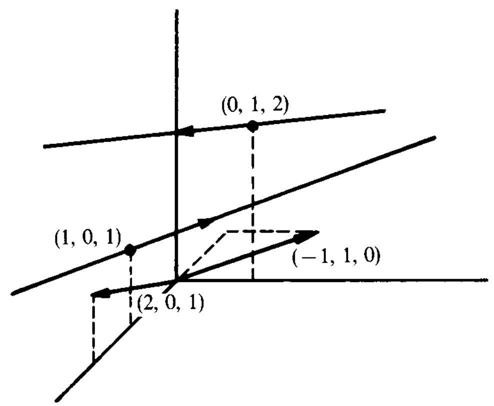



Figura 5



el sistema no posee ninguna solución. Por tanto, las rectas dadas se cruzan o son paralelas. Puesto que sus vectores directores no son proporcionales (¿por qué?), ambas rectas se cruzan.

* * *

Dada una recta en coordenadas paramétricas como en (1), podemos despejar $t$ de cada una de las igualdades para obtener:

$$t=\frac{x-p_1}{v_1}, \quad t=\frac{y-p_2}{v_2}, \quad t=\frac{z-p_3}{v_3}$$

siempre que $v_1$, $v_2$ y $v_3$ sean no nulos. Tenemos, por tanto, las igualdades

$$\frac{x-p_1}{v_1}=\frac{y-p_2}{v_2}=\frac{z-p_3}{v_3} \tag{2}$$

que se transforman en las tres igualdades siguientes:

$$\left.\begin{aligned}v_2x-v_1y&=v_2p_1-v_1p_2\\ v_3y-v_2z&=v_3p_2-v_2p_3\\ v_3x-v_1z&=v_3p_1-v_1p_3\end{aligned}\right\}$$

cada una de las cuales es deducible de las otras dos como se deduce fácilmente de la forma en que se han obtenido. Por tanto, podemos escribir

$$\left.\begin{aligned}v_2x-v_1y&=v_2p_1-v_1p_2\\ v_3y-v_2z&=v_3p_2-v_2p_3\end{aligned}\right\} \tag{3}$$

como las ecuaciones de la recta dada, que reciben el nombre de *ecuaciones cartesianas.*

EJEMPLO C. Para hallar las ecuaciones cartesianas de la recta que pasa por los puntos $P=(1, 0, 1)$ y $Q=(-2, 1, 2)$ escribimos sus ecuaciones paramétricas

$$(x, y, z)=(1, 0, 1)+t(-3, 1, 1).$$

Tenemos, por tanto:

$$\frac{x-1}{-3}=\frac{y-0}{1}=\frac{z-1}{1}$$

o bien

$$\left.\begin{aligned}x+3y&=1\\ y-z&=-1\end{aligned}\right\}$$

* * *

Si alguna o algunas de las componentes del vector $\vec{v}$ son nulas (¡no todas pueden ser nulas!) se obtienen rectas en posiciones especiales con respecto a los ejes y planos coordenados. Por ejemplo, si $v_1 = v_2 = 0$, $v_3 \neq 0$ se obtiene una recta paralela al eje $OZ$ que pasa por el punto $P=(p_1, p_2, p_3)$; esta recta tiene como ecuaciones cartesianas

$$x = p_1, \quad y = p_2$$

Se invita al lector a buscar otros casos particulares.

* * *

Dados dos vectores $\vec{u}$ y $\vec{v}$ en $\mathbb{R}^3$ no proporcionales y un punto $P$ de $\mathbb{R}^3$ el conjunto de puntos que se representa de la forma

$$P + t\vec{u} + s\vec{v}$$

con $t$ y $s$ números reales es un plano que pasa por el punto $P$ y tiene $\vec{u}$ y $\vec{v}$ como *vectores directores.*

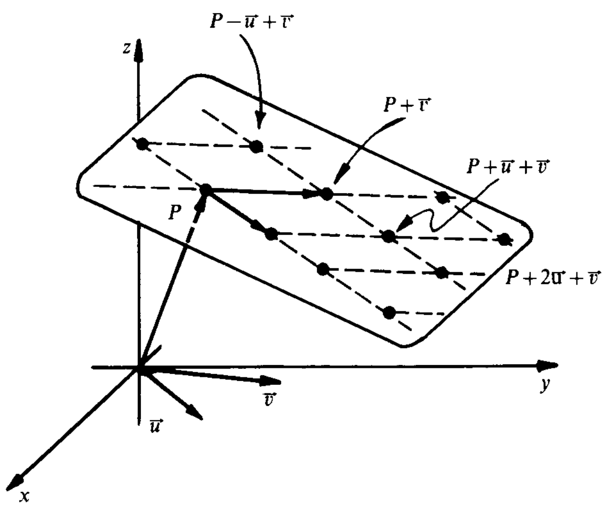



Figura 6

Si $P=(p_1, p_2, p_3)$, $\vec{u}=(u_1, u_2, u_3)$ y $\vec{v}=(v_1, v_2, v_3)$ la ecuación del plano que pasa por $P$ y tiene $\vec{u}$ y $\vec{v}$ como vectores directores puede escribirse con las igualdades

$$\left.\begin{array}{l} x = p_1 + tu_1 + sv_1 \\ y = p_2 + tu_2 + sv_2 \\ z = p_3 + tu_3 + sv_3 \end{array}\right\} \qquad (4)$$

que reciben el nombre de *ecuaciones paramétricas* del plano considerado.



Si queremos encontrar las *ecuaciones cartesianas* de este plano, es decir, ecuaciones que solamente contengan $x$, $y$ y $z$ y no los parámetros $t$ y $s$, basta con eliminar estos parámetros entre las tres ecuaciones de (4). Esto puede hacerse mediante el método de sustitución, pero es más conveniente considerar (4) como un sistema de tres ecuaciones con incógnitas $t$ y $s$ que, fijadas $x$, $y$, $z$, sólo posee una solución; por el teorema de Rouche-Frobenius se ha de tener

$$r\begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix} = r\begin{pmatrix} u_1 & v_1 & x-p_1 \\ u_2 & v_2 & y-p_2 \\ u_3 & v_3 & z-p_3 \end{pmatrix}.$$

Para que esto se cumpla es necesario que

$$\begin{vmatrix} u_1 & v_1 & x-p_1 \\ u_2 & v_2 & y-p_2 \\ u_3 & v_3 & z-p_3 \end{vmatrix} = 0$$

Por la propiedad 3 de los determinantes (ver sección 2.2) para columnas se tiene

$$\begin{vmatrix} u_1 & v_1 & x \\ u_2 & v_2 & y \\ u_3 & v_3 & z \end{vmatrix} - \begin{vmatrix} u_1 & v_1 & p_1 \\ u_2 & v_2 & p_2 \\ u_3 & v_3 & p_3 \end{vmatrix} = 0$$

Desarrollando el primero de los determinantes por la última columna se tiene

$$\begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix} x - \begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix} y + \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} z - \begin{vmatrix} u_1 & v_1 & p_1 \\ u_2 & v_2 & p_2 \\ u_3 & v_3 & p_3 \end{vmatrix} = 0$$

que podemos escribir de la forma

$$ax + by + cz + d = 0$$

Esta ecuación recibe el nombre de *ecuación cartesiana* o en *coordenadas rectangulares* de un plano.

Casos particulares de esta ecuación son, por ejemplo, $x = -d$ o $z = 0$. El primero es un plano que contiene a todos los puntos de la forma $(-d, p_2, p_3)$ con $p_2$, $p_3$ cualesquiera números reales y, por tanto, es un plano paralelo al plano determinado por los ejes $OY$ y $OZ$ que pasa a distancia $-d$ del origen (ver figura 7). El segundo es el plano determinado por los ejes $OX$ y $OY$ (ver figura 7).

Se invita al lector a tratar de visualizar geométricamente otros casos particulares de la ecuación cartesiana de un plano.

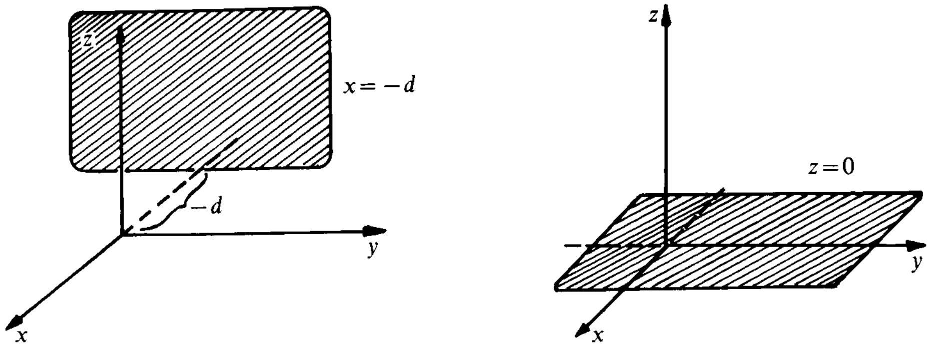



Figura 7

EJEMPLO D. Las ecuaciones paramétricas del plano que pasa por el punto $P=(1, 1, 0)$ y tiene como vectores directores $\vec{u}=(-1, 0, 2)$ y $\vec{v}=(1, 3, 3)$ son

$$\left.\begin{aligned} x&=1-t+s\\ y&=1+3s\\ z&=2t+3s \end{aligned}\right\}.$$

Su ecuación cartesiana es

$$\begin{vmatrix} -1 & 1 & x-1\\ 0 & 3 & y-1\\ 2 & 3 & z \end{vmatrix}=0$$

Tenemos, por tanto:

$$0=\begin{vmatrix} -1 & 1 & x-1\\ 0 & 3 & y-1\\ 2 & 3 & z \end{vmatrix}=\begin{vmatrix} -1 & 1 & x-1\\ 0 & 3 & y-1\\ 0 & 5 & z+2x-2 \end{vmatrix}=-\begin{vmatrix} 3 & y-1\\ 5 & z+2x-2 \end{vmatrix}=$$

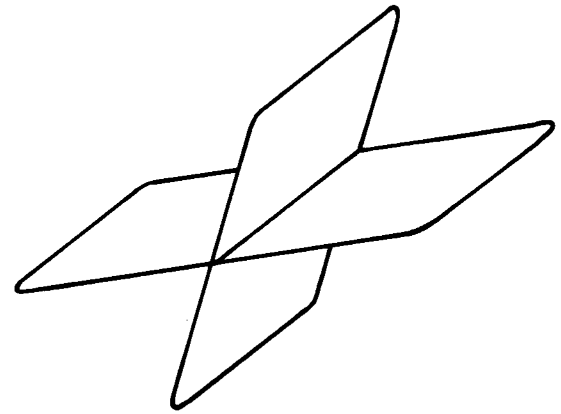
$$=-3z-6x+6+5y-5=-6x+5y-3z+1$$

* * *

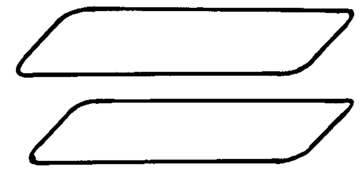
Las posibles posiciones relativas de dos planos en el espacio son las siguientes (véase figura 8):

*a*) Los dos planos se cortan en una recta.
*b*) Los dos planos son paralelos.
*c*) Ambos planos coinciden.





Planos que se cortan en una recta



Planos paralelos

Figura 8

La posición relativa de dos planos puede determinarse a partir de sus ecuaciones, utilizando el teorema de Rouché-Frobenius para determinar el número de puntos en común que ambas poseen. Un estudio de este tipo se realiza en los dos próximos ejemplos.

EJEMPLO E. Para hallar la posición relativa de los planos de ecuaciones $2x+3y-z=1$ y $-x-2y+z=0$ estudiamos el sistema

$$\left.\begin{aligned} 2x+3y-z&=1\\ -x-2y+z&=0 \end{aligned}\right\}$$

Puesto que

$$r\begin{pmatrix} 2 & 3 & -1\\ -1 & -2 & 1 \end{pmatrix}=2=r\begin{pmatrix} 2 & 3 & -1 & 1\\ -1 & -2 & 1 & 0 \end{pmatrix}<3=\text{número de incógnitas}$$

el sistema tiene infinitas soluciones. Puesto que $\begin{pmatrix} 2 & 3\\ -1 & -2 \end{pmatrix}$ es un menor básico de la matriz de los coeficientes, las soluciones dependen de un solo parámetro y, por tanto, los planos se cortan en una recta.

Para determinar la ecuación de esta recta escribimos el sistema en la forma

$$\left.\begin{aligned} 2x+3y&=1+z\\ -x-2y&=-z \end{aligned}\right\}$$

y lo resolvemos utilizando la regla de Cramer:

$$x=\frac{\begin{vmatrix}1+z & 3\\ -z & -2\end{vmatrix}}{\begin{vmatrix}2 & 3\\ -1 & -2\end{vmatrix}}=\frac{-2-2z+3z}{-1}=2-z$$

$$y=\frac{\begin{vmatrix}2 & 1+z\\ -1 & -z\end{vmatrix}}{\begin{vmatrix}2 & 3\\ -1 & -2\end{vmatrix}}=\frac{-2z+1+z}{-1}=z-1.$$

Haciendo $z=t$, se tiene

$$(x,\ y,\ z)=(2-t,\ t-1,\ t)=(2,\ -1,\ 0)+t(-1,\ 1,\ 1)$$

que son las ecuaciones paramétricas de la recta intersección de los dos planos dados.

* * *

EJEMPLO F. Dados los planos de ecuaciones paramétricas

$$(x,\ y,\ z)=(1,\ 0,\ 1)+t(1,\ 1,\ 0)+s(0,\ 1,\ -1)$$
$$(x,\ y,\ z)=(0,\ 0,\ 3)+l(1,\ 2,\ -1)+m(2,\ 1,\ 1)$$

tratamos de determinar su posición relativa. Igualando las correspondientes coordenadas obtenemos el siguiente sistema:

$$\left.\begin{aligned}1+t&=l+2m\\ t+s&=2l+m\\ 1-s&=3-l+m\end{aligned}\right\}\Leftrightarrow\left.\begin{aligned}t-l-2m&=-1\\ t+s-2l-m&=0\\ -s+l-m&=2\end{aligned}\right\}$$

Puesto que

$$\begin{pmatrix}1 & 0 & -1 & -2\\ 1 & 1 & -2 & -1\\ 0 & -1 & 1 & -1\end{pmatrix}\xrightarrow{F_2-F_1}\begin{pmatrix}1 & 0 & -1 & -2\\ 0 & 1 & -1 & 1\\ 0 & -1 & 1 & -1\end{pmatrix}\xrightarrow{F_3-F_2}\begin{pmatrix}1 & 0 & -1 & -2\\ 0 & 1 & -1 & 1\\ 0 & 0 & 0 & 0\end{pmatrix}$$

el rango de la matriz de los coeficientes del sistema anterior es 2. Por otro lado:

$$\begin{vmatrix}1 & 0 & -1\\ 1 & 1 & 0\\ 0 & -1 & 2\end{vmatrix}=\begin{vmatrix}1 & 0 & -1\\ 0 & 1 & 1\\ 0 & -1 & 2\end{vmatrix}=3$$



y, por tanto, el rango de la matriz ampliada es 3. Esto nos lleva a deducir que no existen soluciones del sistema anterior y, por tanto, los *planos dados son paralelos.*

* * *

Observar que en el ejemplo F los vectores directores de la segunda recta son combinación lineal de los vectores directores de la primera recta, ya que

$$(1,\ 2,\ -1)=(1,\ 1,\ 0)+(0,\ 1,\ -1)$$

y

$$(2,\ 1,\ 1)=2(1,\ 1,\ 0)-(0,\ 1,\ -1)$$

Este resultado es cierto en general: *si dos planos son paralelos, los vectores directores de uno de ellos son combinación lineal de los vectores directores del otro y el recíproco también es cierto.* La demostración de este resultado se pide en el ejercicio 7 al final de esta sección.

EJEMPLO G. Tratemos de hallar el plano paralelo al de ecuación $x+y+z=0$ que pasa por el punto $P=(1,\ 1,\ 1)$.

Escribimos el plano dado en ecuaciones paramétricas; para ello basta con determinar tres de sus puntos que no estén alineados; por ejemplo:

$$A=(0,\ 0,\ 0),\quad B=(1,\ -1,\ 0),\quad C=(0,\ 1,\ -1)$$

Estos puntos no están alineados, ya que los vectores

$$\vec{u}=\vec{AB}=(1,\ -1,\ 0)\ \text{ y }\ \vec{v}=\vec{AC}=(0,\ 1,\ -1)$$

no son proporcionales (¿por qué?). Puesto que el plano dado pasa por el punto $A=(0,\ 0,\ 0)$ y tiene $\vec{u}$ y $\vec{v}$ como vectores directores su ecuación paramétrica es

$$(x,\ y,\ z)=(0,\ 0,\ 0)+t(1,\ -1,\ 0)+s(0,\ 1,\ -1).$$

Como vectores directores de un plano paralelo a éste pueden tomarse $\vec{u}$ y $\vec{v}$ y debido a que el plano pedido pasa por el punto $P=(1,\ 1,\ 1)$ sus ecuaciones paramétricas son

$$(x,\ y,\ z)=(1,\ 1,\ 1)+t(1,\ -1,\ 0)+s(0,\ 1,\ -1)$$

Su ecuación en coordenadas cartesianas es

$$0=\begin{vmatrix}1 & 0 & x-1\\ -1 & 1 & y-1\\ 0 & -1 & z-1\end{vmatrix}=\begin{vmatrix}1 & 0 & x-1\\ 0 & 1 & x+y-2\\ 0 & -1 & z-1\end{vmatrix}=z-1+x+y-2=x+y+z-3$$

es decir,

$$x+y+z=3.$$

Otra forma de resolver este problema es utilizando el resultado del ejercicio 8 al final de esta sección; en él se pide demostrar que las ecuaciones de todos los planos paralelos al plano de ecuación $ax+by+cz+d=0$ pueden escribirse de la forma $ax+by+cz+d'=0$, donde $d'$ toma valores en los números reales.

Aceptando este resultado, la ecuación del plano que se busca es de la forma $x+y+z+d'=0$; el valor de $d'$ se calcula imponiendo que el punto $P=(1, 1, 1)$ sea un punto de este nuevo plano; por tanto:

$$1+1+1+d'=0.$$

La ecuación del plano buscado es

$$x+y+z-3=0$$

que es el mismo resultado que se obtuvo anteriormente.

* * *

OBSERVACIÓN 1. Dados dos planos de ecuaciones

$$\begin{aligned} ax+by+cz&=d\\ a'x+b'y+c'z&=d' \end{aligned} \tag{5}$$

ambos determinan una recta si y sólo si

$$r\begin{pmatrix} a & b & c\\ a' & b' & c' \end{pmatrix}=2 \tag{6}$$

Por tanto, dos ecuaciones de la forma que aparecen en (5) son también ecuaciones cartesianas de una recta siempre que se cumpla la condición (6).

OBSERVACIÓN 2. Son numerosos los diferentes ejercicios que pueden hacerse referentes a rectas y planos en el espacio. Al lector se le invita a que realice todos los ejercicios propuestos al final de esta sección y que también realice algunos de su propia cosecha.

Es conveniente también que, siempre que sea posible, se realicen adecuadas representaciones geométricas de las rectas y planos que aparecen en un problema.

## EJERCICIOS 3.2

**1.** Dados $P=(1, 1, -1)$, $Q=(0, 1, 2)$, $\overline{u}=(-1, 2, 0)$ y $\overline{v}=(1, -1, -1)$, hallar las ecuaciones paramétricas y cartesianas de las siguientes rectas en $\mathbb{R}^3$, y dibujarlas:

*a)* Recta que pasa por $P$ con vector director $\overline{u}-\overline{v}$.

*b)* Recta que pasa por $P$ y $Q$.

*c)* Recta que pasa por $Q$ con vector director $3\overline{v}$.



**2.** Dados los siguientes pares de rectas en $\mathbb{R}^3$, determinar su posición relativa y si se cortan, encontrar el punto de intersección:

*a)* $(x, y, z)=(-1, 2, 1)+t(4, 3, 2)$ y $(x, y, z)=(0, 1, 0)+s(1, 3, 2)$

*b)* $(x, y, z)=(0, 1, 5)+t(1, 3, -2)$ y $(x, y, z)=(5, 5, 3)+s(4, 1, 0)$

*c)* $\left.\begin{aligned} 2x+3y-z&=3\\ x-3y&=4 \end{aligned}\right\}$ y $\left.\begin{aligned} x+y&=2\\ y-z&=-1 \end{aligned}\right\}$

**3.** Hallar la ecuación de la recta paralela a la de ecuación

$$\left.\begin{aligned} 3x-y+z&=1\\ x+y-3z&=0 \end{aligned}\right\}$$

que pasa por el punto $(1, 1, 1)$.

**4.** De todas las rectas que pasan por el punto $P=(0, 2, -1)$, hallar la que corta a las rectas de ecuaciones

$$(1, 1, 2)+t(2, -1, 0)$$

y

$$(0, 1, 1)+s(-3, 1, 2)$$

**5.** Dados $P=(1, 2, 3)$, $Q=(-1, -2, -3)$, $R=(0, 1, -1)$, $\overline{u}=(0, 1, -1)$ y $\overline{v}=(5, 1, 2)$, hallar las ecuaciones paramétricas y cartesianas de los siguientes planos:

*a)* Plano que pasa por $P$, $Q$ y $R$.

*b)* Plano que pasa por $P$ y $R$ y es paralelo a la recta que pasa por $Q$ y tiene $\overline{u}-\overline{v}$ como vector director.

*c)* Plano que contiene a $R$ en la dirección de $u+2\overline{v}$ y $2\overline{u}+\overline{v}$.

**6.** Determinar, si es posible, la intersección de los siguientes pares de planos en $\mathbb{R}^3$:

*a)* $x-y+z=1$ y $2x+2y-3z=4$

*b)* $(x, y, z)=t(1, 1, -1)+s(0, 1, -2)$ y $(x, y, z)=(0, 1, 0)+l(0, 1, -1)+m(2, 3, 5)$

*c)* $(x, y, z)=(2, 3, 1)+(0, 1, 2)+s(3, 1, -5)$ y $x-6y+3z+1=0$

**7.** Demostrar que dos planos son paralelos o coinciden si y sólo si los vectores directores de uno de ellos son combinación lineal de los vectores directores del otro.

**8.** Demostrar que las ecuaciones de los planos paralelos al plano de ecuación $ax+by+cz+d=0$ pueden escribirse de la forma $ax+by+cz+d'=0$ con $d'$ un número real.

**9.** Escribir las ecuaciones paramétricas y cartesianas de los siguientes planos y rectas en $\mathbb{R}^3$:

*a)* Recta que pasa por $Q=(0, 2, 1)$ y es paralela al plano

$$\pi:\ 3x-z+2=0$$

y al plano que pasa por $P=(1, 0, 1)$, $Q$ y el origen.

*b*) Plano paralelo a la recta

$$r\colon (2, 1, 0)+t(1, 2, 1)$$

por el punto $Q$ y que contiene al punto $(1, 3, 0)$.

*c*) Todas las rectas que pasan por $P$ y son paralelas a $\pi$ y, entre ellas, la que corta a $r$.

*d*) El plano que pasa por $R=(2, 1, 1)$ y contiene a la recta

$$s\colon \left.\begin{array}{r} 3x+y-z=2 \\ -2y+z=1 \end{array}\right\}$$

*e*) La recta que une $R$ con la intersección de $s$ y el plano $x-y=1$.

*f*) Haz de rectas que une $R$ con los puntos de $s$.

*g*) Haz de planos que pasa por $R$. [*Sugerencia:* generalizar el problema 8 de la sección anterior.]

**10.** Demostrar que $\begin{vmatrix} 1 & x & y & z \\ 1 & a_1 & a_2 & a_3 \\ 1 & b_1 & b_2 & b_3 \\ 1 & c_1 & c_2 & c_3 \end{vmatrix}=0$ es la ecuación de un plano que pasa por los puntos $A=(a_1, a_2, a_3)$, $B=(b_1, b_2, b_3)$ y $C=(c_1, c_2, c_3)$.

## 3.3. DISTANCIAS Y ANGULOS. PRODUCTO ESCALAR

Dados dos puntos $P$ y $Q$ en el plano, de coordenadas

$$P=(p_1, p_2) \quad , \quad Q=(q_1, q_2)$$

con respecto a un sistema de coordenadas rectangulares, del teorema de Pitágoras se deduce que la distancia de $P$ a $Q$ se puede calcular mediante la fórmula

$$d(P, Q)=\sqrt{(q_1-p_1)^2+(q_2-p_2)^2} \quad \text{(ver figura 1)}$$

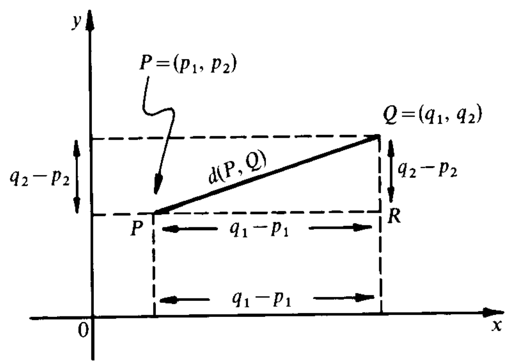



Figura 1



Si los puntos $P$ y $Q$ están en $\mathbb{R}^3$ y tienen como coordenadas

$$P=(p_1, p_2, p_3)$$
$$Q=(q_1, q_2, q_3)$$

aplicando el teorema de Pitágoras a los triángulos rectángulos $PQS$ y $PSR$ se obtiene el siguiente resultado:

$$d(P, Q)=\sqrt{(q_1-p_1)^2+(q_2-p_2)^2+(q_3-p_3)^2} \quad \text{(véase figura 2)}$$

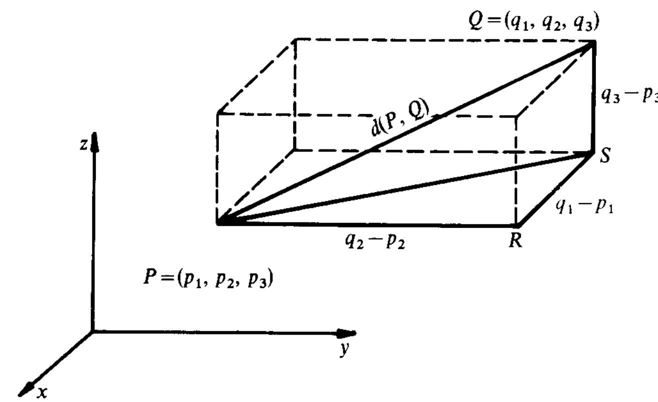



Figura 2

Dado un vector $\vec{v}=\overrightarrow{AB}$ en el plano o en el espacio se denomina *longitud* de $\vec{v}$ y se designa por $\|\vec{v}\|$ a la distancia entre $A$ y $B$. Palabras sinónimas de longitud de un vector son *módulo de* un vector y *norma* de un vector.

EJEMPLO A. La longitud del vector $\vec{v}=\overrightarrow{AB}$, donde $A=(1, -1, 2)$ y $B=(3, 4, -5)$, es

$$\|\vec{v}\|=\sqrt{(3-1)^2+(4+1)^2+(-5-2)^2}=\sqrt{4+25+49}=\sqrt{78}$$

* * *

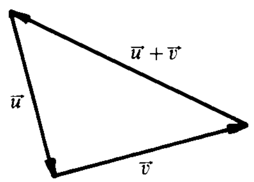
La longitud de un vector posee las siguientes propiedades:

1) $\|c\vec{v}\|=|c|\|\vec{v}\|$, donde $c$ es un número real y $|c|$ denota el valor absoluto de $c$.

2) $\|\vec{u}+\vec{v}\|\leqslant\|\vec{u}\|+\|\vec{v}\|$.

La segunda propiedad se denomina *desigualdad triangular* y geométricamente se expresa diciendo que cualquier lado de un triángulo es menor que la suma de los otros dos. Esta propiedad puede demostrarse utilizando la ley del coseno, a saber:

$$\|\vec{u}+\vec{v}\|^2=\|\vec{u}\|^2+\|\vec{v}\|^2-2\|\vec{u}\|\|\vec{v}\|\cos \sphericalangle(\vec{u}, \vec{v})$$



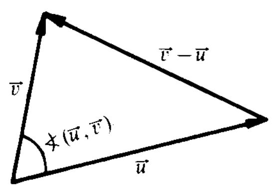

Puesto que $\cos \sphericalangle (u, v) \geqq -1$, se obtiene que

$$\|\vec{u}+\vec{v}\|^2 \leqq \|\vec{u}\|^2+\|\vec{v}\|^2+2\|\vec{u}\|\|\vec{v}\|=(\|\vec{u}\|+\|\vec{v}\|)^2$$

de donde se deduce el resultado tomando raíces cuadradas en ambos lados.

El *ángulo* que forman dos vectores $\vec{u}$ y $\vec{v}$ en $\mathbb{R}^2$ o en $\mathbb{R}^3$ puede obtenerse utilizando el teorema del coseno; ello nos permite calcular el coseno del ángulo que forman, el cual queda unívocamente determinado si suponemos que $0 \leqq \sphericalangle (\vec{u}, \vec{v}) \leqq \pi$. Puede obtenerse una fórmula para calcular $\cos \sphericalangle (\vec{u}, \vec{v})$; si $\vec{u}=(u_1, u_2)$ y $\vec{v}=(v_1, v_2)$, del teorema del coseno se deduce que

$$(v_1-u_1)^2+(v_2-u_2)^2=u_1^2+u_2^2+v_1^2+v_2^2-2\|\vec{u}\|\|\vec{v}\| \cos \sphericalangle (\vec{u}, \vec{v})$$




Simplificando obtenemos

$$-2u_1v_1-2u_2v_2=-2\|\vec{u}\|\|\vec{v}\| \cos \sphericalangle (\vec{u}, \vec{v})$$

de donde se deduce que

$$\cos \sphericalangle (\vec{u}, \vec{v})=\frac{u_1v_1+u_2v_2}{\|\vec{u}\|\|\vec{v}\|} \tag{1}$$

La expresión que aparece en el numerador de la parte derecha de esta fórmula recibe el nombre de *producto escalar* de los vectores $\vec{u}$ y $\vec{v}$ y se denota por $(\vec{u}, \vec{v})$. Si los vectores tienen dos componentes tenemos

$$(\vec{u}, \vec{v})=u_1v_1+u_2v_2$$

Si tienen tres componentes se tendría

$$(\vec{u}, \vec{v})=u_1v_1+u_2v_2+u_3v_3$$

y la fórmula para calcular el ángulo que forman es similar a la (1).



EJEMPLO B. El coseno del ángulo que forman los vectores $\vec{u}=(1, 1, 1)$ y $\vec{v}=(-1, 1, 1)$ en $\mathbb{R}^3$ es

$$\cos \sphericalangle (\vec{u}, \vec{v})=\frac{(\vec{u}, \vec{v})}{\|\vec{u}\|\|\vec{v}\|}=\frac{1\cdot(-1)+1\cdot 1+1\cdot 1}{\sqrt{3}\sqrt{3}}=\frac{1}{3}$$

Observar que el ángulo es agudo, ya que su coseno es positivo.

* * *

*Nota.* De acuerdo con la fórmula (1) *dos vectores son perpendiculares si y sólo si su producto escalar es cero*, ya que es en este caso cuando su coseno es cero y, por tanto, el ángulo que forman es $\pi/2$.

De la definición de producto escalar pueden deducirse las siguientes propiedades:

1) $(\vec{u}, \vec{v})=(\vec{v}, \vec{u})$ para todo par de vectores $\vec{u}$ y $\vec{v}$.
2) $(\vec{u}+\vec{v}, \vec{\omega})=(\vec{u}, \vec{\omega})+(\vec{v}, \vec{\omega})$ para cualesquiera tres vectores $\vec{u}$, $\vec{v}$ y $\vec{\omega}$ (propiedad lineal).
3) $(c\vec{u}, \vec{v})=c(\vec{u}, \vec{v})$ para todo número real $c$ y todo par de vectores $\vec{u}$ y $\vec{v}$.
4) $(\vec{u}, \vec{u})=\|\vec{u}\|^2 \geqq 0$ para todo vector $\vec{u}$.

Con la relación de producto escalar pueden atacarse problemas en el plano y en el espacio sobre perpendicularidad y ángulos que forman algunas de las figuras anteriormente estudiadas. Lo que sigue es una serie de problemas relativos a estos temas; esta serie no es exhaustiva y sería conveniente que el lector se propusiera sus propios problemas sobre estas nociones.

* * *

Comenzamos hallando las *ecuaciones cartesianas de una recta en un plano* sin necesidad de hallar sus ecuaciones paramétricas; supongamos que la recta pasa por el punto $P=(p_1, p_2)$ y tiene a $\vec{u}=(u_1, u_2)$ como vector director. Un vector perpendicular a $\vec{u}$ es

$$\vec{v}=(-u_2, u_1)$$

ya que

$$(\vec{u}, \vec{v})=-u_1u_2+u_2u_1=0$$

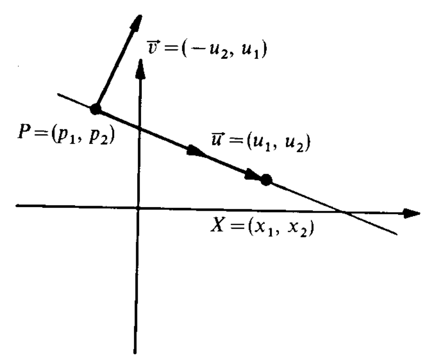



Figura 3

Un punto $X=(x_1, x_2)$ está en la recta pedida si y sólo si $\overrightarrow{PX}$ es perpendicular a $\vec{v}$; por tanto:

$$0=(\overrightarrow{PX}, \vec{v})=-(x_1-p_1)u_2+(x_2-p_2)u_1=-u_2x_1+u_1x_2+c$$

donde $c=p_1u_2-p_2u_1$. Esta es la ecuación cartesiana de la recta que tiene a $\vec{u}=(-u_2, u_1)$ como vector perpendicular y pasa por el punto $P$. Un vector perpendicular a una recta dada se denomina también un *vector característico* de la recta.

Dada la ecuación $a_1x_1+a_2x_2=c$ que determina una recta en un plano, un vector perpendicular a ella es $\vec{a}=(a_1, a_2)$. Para demostrar esto basta escribir la ecuación de la recta en la forma $(\vec{a}, \vec{x})=c$, donde $\vec{x}=(x_1, x_2)$; si $P=(p_1, p_2)$ es un punto de la recta se tiene $(\vec{a}, \vec{p})=c$, con $\vec{p}=(p_1, p_2)$; por tanto, tenemos $(\vec{a}, \vec{x})=(\vec{a}, \vec{p})$, o equivalentemente:

$$(\vec{a}, \vec{x}-\vec{p})=0$$

Esto nos dice que $(\vec{a}, \vec{x})=c$ es la ecuación de la recta que pasa por $P$ y tiene a $\vec{a}$ como vector perpendicular.

EJEMPLO C. Para hallar la ecuación de una recta que pasa por el punto $P=(1, -2)$ y tiene a $\vec{u}=(1, 3)$ como vector director, observamos que $\vec{v}=(-3, 1)$ es un vector perpendicular a $\vec{u}$; por tanto:

$$(\vec{v}, \vec{x})=-3x_1+x_2=c$$

es la forma de la ecuación pedida; para determinar $c$ sustituimos el punto $P=(1, -2)$ y obtenemos

$$-3-2=c$$

Por tanto, $-3x_1+x_2=-5$ es la ecuación de la recta.

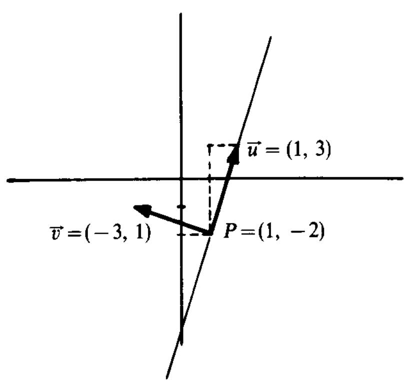



Figura 4

* * *



Utilizando el producto escalar puede encontrarse la *ecuación cartesiana de un plano* en el espacio cuando se conocen un punto del plano y un vector perpendicular a él. Si $P=(p_1, p_2, p_3)$ es un punto del plano y $\vec{u}=(u_1, u_2, u_3)$ es un vector perpendicular a él, cualquier punto $X=(x_1, x_2, x_3)$ del plano satisface

$$(\vec{u}, \overrightarrow{XP})=0$$

como puede fácilmente observarse en la figura adjunta. La igualdad anterior puede escribirse de la forma

$$u_1x_1+u_2x_2+u_3x_3=c$$

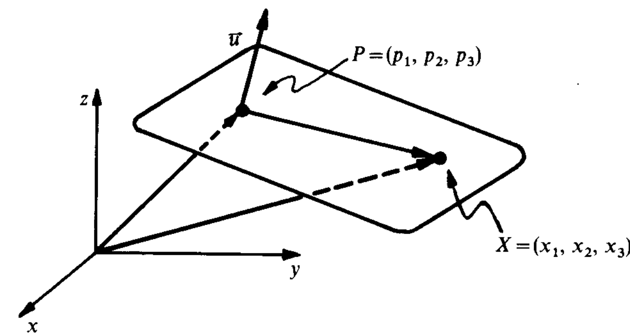



Figura 5

Recíprocamente, una fórmula como la anterior determina un plano que es perpendicular al vector $\vec{u}=(u_1, u_2, u_3)$. La demostración es similar a la realizada para la recta en el plano y se deja como ejercicio.

EJEMPLO D. El plano que pasa por el punto $P=(1, 3, -2)$ y tiene a $\vec{u}=(2, -1, 4)$ como vector perpendicular o característico tiene como ecuación una expresión de la forma

$$2x_1-x_2+4x_3=c$$

El valor de $c$ se determina imponiendo que $P$ pertenezca al plano: $2\cdot 1-3+4(-2)=-9=c$. Por tanto, $2x_1-x_2+4x_3=-9$ es la ecuación del plano.

* * *

Sea $\vec{u}$ un vector cualquiera de $\mathbb{R}^2$ o $\mathbb{R}^3$. El producto escalar puede utilizarse para *descomponer todo vector $\vec{v}$ de $\mathbb{R}^2$ o $\mathbb{R}^3$ en una suma de dos vectores*

$$\vec{v}=\vec{a}+\vec{b}$$

*donde $\vec{a}$ es un múltiplo de $\vec{u}$ y $\vec{b}$ es perpendicular a $\vec{u}$* (ver figura 6).

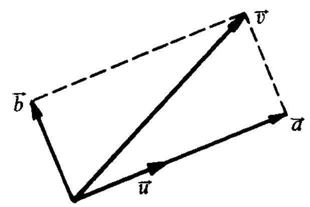



Figura 6

Puesto que $\vec{a}$ es un múltiplo de $\vec{u}$ hemos de tener $\vec{a}=c\vec{u}$, con $c\in\mathbb{R}$, y, por tanto:

$$\vec{v}=c\vec{u}+\vec{b}$$

Multiplicando escalarmente los dos lados de la igualdad anterior por $\vec{u}$ y teniendo en cuenta que $(\vec{b}, \vec{u})=0$, ya que $\vec{b}$ es perpendicular a $\vec{u}$, se tiene que

$$(\vec{v}, \vec{u})=c(\vec{u}, \vec{u})$$

De aquí deducimos que

$$\vec{a}=c\vec{u}=\frac{(\vec{v}, \vec{u})}{(\vec{u}, \vec{u})}\vec{u} \tag{1}$$

y, por tanto:

$$b=\vec{v}-\vec{a}=\vec{v}-\frac{(\vec{v}, \vec{u})}{\|\vec{u}\|^2}\vec{u}.$$

De la fórmula (1) y de la fórmula del coseno del ángulo que forman dos vectores deducimos que

$$\vec{a}=\frac{\|\vec{v}\|}{\|\vec{u}\|}[\cos \sphericalangle(\vec{u}, \vec{v})]\vec{u}$$

y si $\vec{u}$ es un vector *unitario*, es decir, de longitud 1, se tiene

$$\vec{a}=\|\vec{v}\|[\cos \sphericalangle(\vec{u}, \vec{v})]\vec{u}$$

lo cual demuestra que $\vec{a}$ es un vector en la dirección de $\vec{u}$ *que tiene* $\|\vec{v}\|[\cos \sphericalangle(\vec{u}, \vec{v})]$ como módulo (ver figura 7).

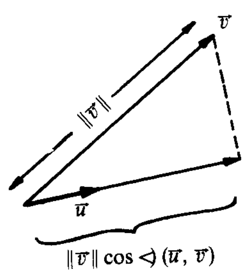



Figura 7



EJEMPLO E. Para descomponer $\vec{v}=(1, 3, -1)$ en un vector paralelo a $\vec{u}=(-1, 0, 1)$ y otro perpendicular a él, escribimos

$$\vec{v}=\vec{a}+\vec{b}=c\vec{u}+\vec{b}$$

donde $\vec{b}$ es perpendicular a $\vec{u}$. Del razonamiento anterior deducimos que

$$c=\frac{(\vec{v}, \vec{u})}{\|\vec{u}\|^2}=\frac{1\cdot(-1)+3\cdot 0+(-1)\cdot 1}{(-1)^2+0^2+1^2}=\frac{-2}{2}=-1.$$

Por tanto:

$$\vec{a}=-1\vec{u}=(1, 0, -1) \quad \text{y} \quad \vec{b}=\vec{v}-\vec{a}=(0, 3, 0)$$

* * *

Dado un punto $P$ se denomina *proyección* de $P$ sobre una recta o un plano a un punto $P'$, de la recta o del plano, tal que $\overrightarrow{PP'}$ es perpendicular a la recta o al plano dados (véase figura 8).

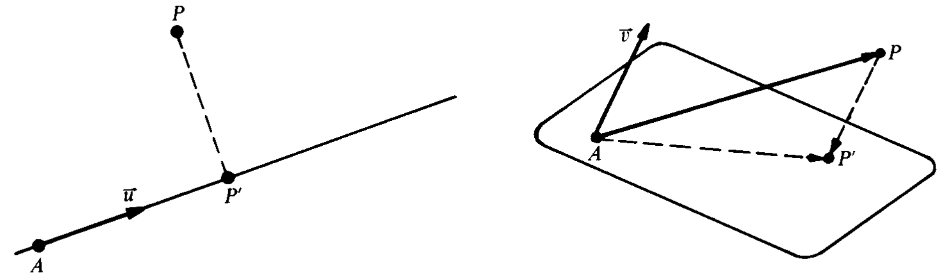



Figura 8

El problema de hallar la proyección de un punto sobre una recta o un plano se resuelve con métodos análogos a los que se utilizaron para descomponer un vector en sus componentes paralela y perpendicular a un vector fijado. En lugar de encontrar fórmulas de manera teórica preferimos realizar algunos ejemplos.

EJEMPLO F. Queremos hallar la proyección del punto $P=(1, 2)$ sobre la recta $r$ de ecuación

$$-2x+3y=-6$$

en el plano. Un vector perpendicular a la recta dada es $\vec{v}=(-2, 3)$ y, por tanto, $\vec{u}=(3, 2)$ es un vector director de la recta. Si tomamos un punto de la recta, por ejemplo, $A=(0, -2)$, $P'$ será de la forma: $P'=A+c\vec{u}$. Además, $PP'$ es perpendicular a $\vec{u}$ y, por tanto,

$$\begin{aligned}0&=(\overrightarrow{PP'}, \vec{u})=(\overrightarrow{PA}+c\vec{u}, \vec{u})=(\overrightarrow{PA}, \vec{u})+c(\vec{u}, \vec{u})=\\&=((-1, -4), (3, 2))+c((3, 2), (3, 2))=-11+c\cdot 13\end{aligned}$$

De aquí deducimos que $c=11/13$ y, por tanto,

$$P'=(0,\ -2)+\frac{11}{13}(3,\ 2)=\left(\frac{33}{13},\ -\frac{4}{13}\right)$$

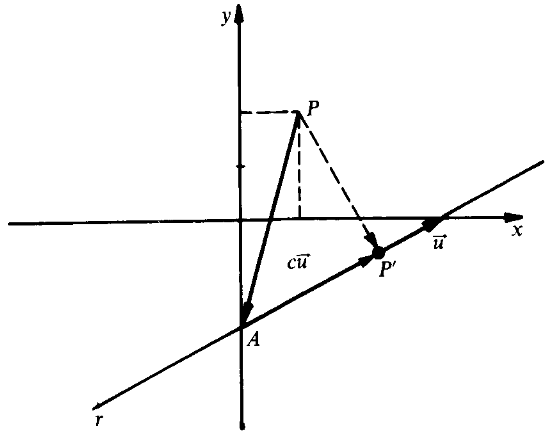



Figura 9

EJEMPLO G. Queremos hallar la proyección del punto $P=(1, 2, 3)$ sobre el plano $\pi$ de ecuación $x+y-2z=3$; sabemos que un vector perpendicular a este plano es $\vec{v}=(1, 1, -2)$. Si $P'$ es la proyección de $P$ sobre el plano $\pi$ hemos de tener $PP'=c\vec{v}$; además (ver figura 8), si $A$ es un punto del plano, por ejemplo, $A=(1, 2, 0)$, se tiene que $\overrightarrow{AP'}$ es perpendicular a $\vec{v}$. Por tanto:

$$0=(\overrightarrow{AP'},\ \vec{v})=(\overrightarrow{AP}+\overrightarrow{PP'},\ \vec{v})=(\overrightarrow{AP},\ \vec{v})+c(\vec{v},\ \vec{v})=$$
$$=((0,\ 0,\ 3),\ (1,\ 1,\ -2))+c((1,\ 1,\ -2),\ (1,\ 1,\ -2))=-6+c6$$

Así pues, $c=1$ y $\overrightarrow{PP'}=\vec{v}=(1, 1, -2)$. Como $P'=P+\overrightarrow{PP'}$, se tiene

$$P'=(1,\ 2,\ 3)+(1,\ 1,\ -2)=(2,\ 3,\ 1)$$

* * *

Se denomina *distancia de un punto P a una recta o a un plano* a la menor de las distancias del punto $P$ a cada uno de los puntos de la recta o del plano. Utilizando el teorema de Pitágoras se demuestra que la distancia de $P$ a una recta o a un plano coincide con la distancia de $P$ a la proyección $P'$ de $P$ sobre la recta o el plano considerados. Para demostrar esto basta observar que si $Q$ es otro punto de la recta o del plano (ver figura 10) se tiene

$$\|\overrightarrow{PQ}\|^2=\|\overrightarrow{PP'}\|^2+\|\overrightarrow{P'Q}\|^2>\|\overrightarrow{PP'}\|^2$$



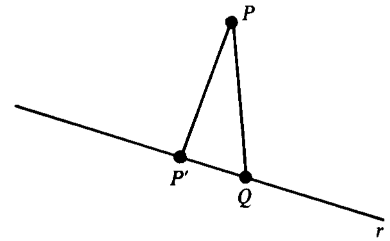



Figura 10

si $Q\neq P'$. Puesto que ya sabemos calcular la proyección $P'$ de $P$ también sabemos calcular la distancia de un punto a una recta o a un plano.

En los casos de la distancia de un punto a una recta en un plano o de un punto a un plano en el espacio se obtiene una fórmula sencilla, que es conveniente encontrar.

PROPOSICIÓN 1 *(Distancia de un punto a un plano)*

La distancia del punto $P=(p_1, p_2, p_3)$ al plano de ecuación $a_1x_1+a_2x_2+a_3x_3+c=0$ está dada por la fórmula

$$\frac{|a_1p_1+a_2p_2+a_3p_3+c|}{\sqrt{a_1^2+a_2^2+a_3^2}}$$

*Demostración.* La ecuación del plano puede escribirse de la forma

$$(\vec{a},\ \vec{x})+c=0$$

donde $\vec{a}=(a_1, a_2, a_3)$ y $\vec{x}=(x_1, x_2, x_3)$. Puesto que $\vec{a}$ es un vector perpendicular al plano, la distancia $d(P, \pi)$ del punto $P$ al plano $\pi$ es de la forma $\|t\vec{a}\|$, donde $t$ debe de ser elegido de manera que

$$P+t\vec{a}=P'\in\pi$$

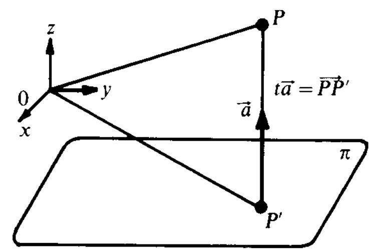



Figura 11

Para obtener $t$ sustituimos en la ecuación del plano:

$$0=(\vec{a}, \overrightarrow{OP'})+c=(\vec{a}, \overrightarrow{OP}+t\vec{a})+c=(\vec{a}, \overrightarrow{OP})+t(\vec{a}, \vec{a})+c$$

De aquí deducimos que

$$t=\frac{-(\vec{a}, \overrightarrow{OP})-c}{\|\vec{a}\|^2}$$

y, por tanto,

$$d(P, \pi)=\|t\vec{a}\|=\left\|\frac{-(\vec{a}, \overrightarrow{OP})-c}{\|\vec{a}\|^2}\vec{a}\right\|=\frac{|(\vec{a}, \overrightarrow{OP})+c|}{\|\vec{a}\|}$$

que coincide con la fórmula anunciada en la proposición. ■

Un razonamiento similar al anterior permite obtener el siguiente resultado, que se deja como ejercicio.

PROPOSICIÓN 2 *(Distancia de un punto a una recta en el plano)*

La distancia del punto $P=(p_1, p_2)$ a la recta de ecuación $a_1x_1+a_2x_2+c=0$ está dada por la fórmula

$$\frac{|a_1p_1+a_2p_2+c|}{\sqrt{a_1^2+a_2^2}}$$

EJEMPLO H. La distancia del punto $P=(1, 3, -2)$ al plano de ecuación

$$\pi: 3x_1-2x_2+7x_3=5$$

es

$$d(P, \pi)=\frac{|3\cdot 1-2\cdot 3+7\cdot(-2)-5|}{\sqrt{3^2+(-2)^2+7^2}}=\frac{|-22|}{\sqrt{62}}=\frac{22}{\sqrt{62}}$$

* * *

Tratamos de hallar ahora las *bisectrices de dos rectas dadas* que se cortan. Tanto en el plano como en el espacio sus ecuaciones serán de la forma

$$A+t\vec{u} \quad \text{y} \quad A+s\vec{v}$$

donde $A$ es el punto de corte. Puesto que $\dfrac{\vec{u}}{\|\vec{u}\|}$ y $\dfrac{\vec{v}}{\|\vec{v}\|}$ son vectores de igual longirud es fácil demostrar que

$$\frac{\vec{u}}{\|\vec{u}\|}+\frac{\vec{v}}{\|\vec{v}\|}$$



y

$$\frac{\vec{u}}{\|\vec{u}\|}-\frac{\vec{v}}{\|\vec{v}\|}$$

son vectores que tienen la dirección de cada una de las bisectrices buscadas. Basta para ello observar que los triángulos $ABD$ y $ACD$ son isósceles e iguales. Las ecuaciones serán

$$A+t\left(\frac{\vec{u}}{\|\vec{u}\|}\pm\frac{\vec{v}}{\|\vec{v}\|}\right)$$

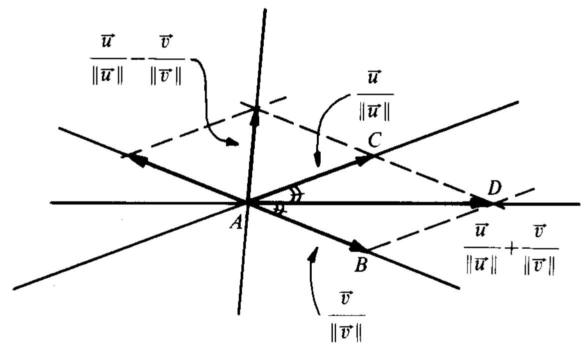



Figura 12

EJEMPLO I. Deseamos hallar las bisectrices de las rectas de ecuaciones

$$r_1: x-y=-1, \quad r_2: 2x+y=4$$

Un vector perpendicular a $r_1$ es $(1, -1)$, luego un vector director de $r_1$ es

$$\vec{u}=(1, 1)$$

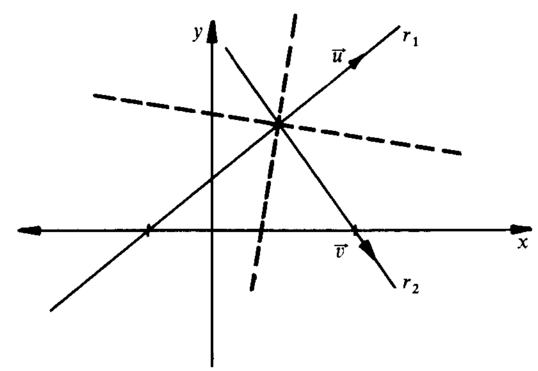



Figura 13

Un vector perpendicular a $r_2$ es (2, 1), luego un vector director de $r_2$ es

$$\vec{v}=(1, -2)$$

Los vectores directores de las bisectrices serán

$$\frac{\vec{u}}{\|\vec{u}\|}\pm\frac{\vec{v}}{\|\vec{v}\|}=\frac{(1, 1)}{\sqrt{2}}\pm\frac{(1, -2)}{\sqrt{5}}$$

Un punto por el que pasan las bisectrices es el punto de intersección de las rectas dadas; el lector puede comprobar que este punto es $A=(1, 2)$. Por tanto:

$$(x, y)=(1, 2)+t\left(\frac{1}{\sqrt{2}}\pm\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{2}}\pm\frac{-2}{\sqrt{5}}\right)$$

son las ecuaciones paramétricas de las bisectrices.

* * *

Para terminar esta sección recordamos que el ángulo que forman dos planos se define como el ángulo que forman sus vectores perpendiculares (ver figura 13). Por tanto, el coseno del ángulo que forman dos planos puede calcularse mediante la fórmula obtenida al comienzo de esta sección.

Si el lector desea realizar algún ejercicio relacionado con este concepto puede intentar hallar el ángulo que forman los planos de ecuaciones

$$3x-2y+z=0 \quad \text{y} \quad 2x+y-3z=0$$

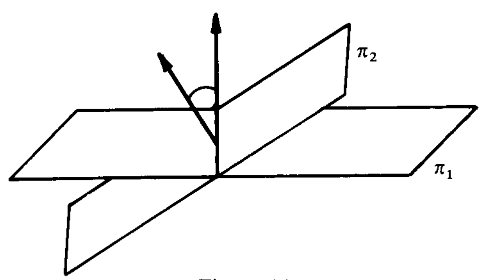



Figura 14

## EJERCICIOS 3.3

**1.** Dados los puntos $P=(1, 0)$, $Q=(-1, 1)$ y $R=(2, 1)$ y las rectas

$$r_1\colon 3x+y=4, \quad r_2\colon x-2y=-1$$

se pide:



*a*) Distancia de $P$ a $Q$.
*b*) El coseno del ángulo que forman los vectores $\overrightarrow{PQ}$ y $\overrightarrow{PR}$.
*c*) El coseno del ángulo que forman las rectas $r_1$ y $r_2$.
*d*) Ecuación cartesiana de la recta que pasa por $P$ y es perpendicular a $r_1$.
*e*) Ecuación cartesiana de la recta que pasa por $R$ y es perpendicular a $r_2$.
*f*) Descomponer el vector $\overrightarrow{PQ}$ en un vector paralelo a la recta $r_1$ y en otro perpendicular a ella.
*g*) Hallar la proyección del punto $Q$ sobre la recta $r_2$.
*h*) Hallar la distancia de $Q$ a $r_2$ y de $P$ a $r_1$.
*i*) Hallar el simétrico del punto $P$ con respecto a la recta $r_1$.
*j*) Encontrar las ecuaciones cartesianas de las bisectrices de $r_1$ y $r_2$.

**2.** Dados los puntos $P=(1, 0, 0)$, $Q=(1, 0, 1)$ y $R=(2, 3, -2)$ y los planos

$$r_1\colon 3x-y+z=3, \quad r_2\colon x+2y-z=2$$

se pide:

*a*) Distancia de $P$ a $Q$ y de $Q$ a $R$.
*b*) El coseno del ángulo que forman los planos $r_1$ y $r_2$.
*c*) Ecuación del plano perpendicular al vector $\overrightarrow{PQ}$ que pasa por $R$.
*d*) La proyección del punto $P$ sobre el plano $r_1$.
*e*) Descomponer el vector $\overrightarrow{QR}$ en un vector perpendicular a $r_2$ y otro paralelo a $r_2$.
*f*) Encontrar las ecuaciones cartesianas de la recta perpendicular al plano $r_1$ que pasa por $R$.
*g*) Hallar la ecuación del plano perpendicular a la recta intersección de los dos planos dados y que pasa por $Q$.
*h*) Distancia de $Q$ a $r_2$ y de $P$ a $r_1$.
*i*) Hallar el punto simétrico de $P$ respecto del plano $r_1$.

**3.** Dados los puntos $P=(0, 0, 1)$ y $Q=(0, 3, 0)$ y las rectas de ecuaciones:

$$r_1\colon (1, 1, 0)+t(2, 0, 1), \quad r_2\colon (0, 0, -2)+s(1, -1, 3)$$

se pide:

*a*) Demostrar que $r_1$ y $r_2$ se cruzan.
*b*) Ecuación del plano perpendicular a $r_1$ que pasa por $P$.
*c*) Ecuaciones paramétricas y cartesianas de la recta paralela a $r_1$ que corta a $r_2$ en el punto (0, 0, $-2$).
*d*) Ecuaciones de las bisectrices de la recta $r_2$ y la encontrada en el apartado anterior.
*e*) La proyección del punto $Q$ sobre $r_1$.
*f*) Distancia de $Q$ a $r_1$ y de $P$ a $r_2$.

g) Descomponer el vector $\overline{PQ}$ en un vector paralelo y otro perpendicular a $r_1$.
h) Ecuación del plano paralelo a $r_1$ que contiene a $r_2$.
i) Ecuación del plano paralelo a $r_2$ que contiene a $r_1$.
j) Distancia de $r_1$ a $r_2$. [*Sugerencia:* hallar la distancia entre los planos de los apartados *h*) e *i*).]

**4.** Utilizar vectores para demostrar que las diagonales de un rectángulo son perpendiculares si y sólo si el rectángulo es un cuadrado.

**5.** Demostrar que las bisectrices de dos rectas que se cortan son perpendiculares.

**6.** Dadas las rectas de ecuaciones

$$r_1\colon \left.\begin{array}{r} x-y+z=1 \\ 2x+y=3 \end{array}\right\}, \quad r_2\colon \left.\begin{array}{r} y+z=2 \\ x-2y+z=0 \end{array}\right\}$$

demostrar que se cortan y hallar las ecuaciones paramétricas de sus bisectrices.

**7.** Hallar la ecuación del plano cuyo punto más cercano al origen es $P=(1, -2, 1)$.

**8.** Hallar la ecuación de un plano que determine con los ejes coordenados segmentos de longitudes 2, 3 y 1, respectivamente.

**9.** Hallar la ecuación del plano paralelo al plano de ecuación $2x-3y-6z-14=0$ y que dista 5 unidades del origen.

**10.** Encontrar la ecuación del plano perpendicular al de ecuación $x-2y-2z+9=0$ que pasa por los puntos $P=(2, -1, 6)$ y $Q=(1, -2, 4)$. [*Sol.:* $2x+4y-3z+18=0$.]

## 3.4. FIGURAS SENCILLAS EN EL PLANO Y EN EL ESPACIO: SUS ECUACIONES

En las secciones anteriores hemos estudiado las rectas y los planos por medio de sus ecuaciones. En esta sección estudiaremos otras figuras en el plano y en el espacio. En todas ellas se observará que si sus ecuaciones paramétricas sólo dependen de un parámetro las figuras son curvas o trozos de curvas, mientras que si depende de dos parámetros se trata de superficies o trozos de superficies.

Comenzamos con el segmento que une los puntos $P$ y $Q$, donde $P$ y $Q$ pueden ser puntos del plano o del espacio. La recta que contiene a $P$ y $Q$ tiene como ecuación paramétrica

$$P+t(\overrightarrow{PQ})=P+t(Q-P)=(1-t)P+tQ$$

donde $t$ es un número real. Si imponemos la restricción

$$0\leqq t\leqq 1$$



Figura 1

se obtienen todos los puntos del segmento $\overline{PQ}$ y solamente éstos. Por tanto, el segmento $\overline{PQ}$ puede representarse de la forma

$$(1-t)P+tQ \quad \text{con} \quad 0\leqq t\leqq 1$$

Dado un punto $X$ en el segmento $\overline{PQ}$ se denomina *razón de* $X$ con *respecto a* $P$ y $Q$ al cociente

$$r=\frac{\|\overrightarrow{PX}\|}{\|\overrightarrow{QX}\|}$$

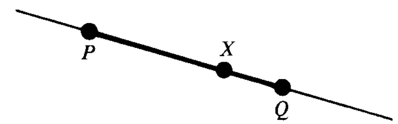


Figura 2

La razón de un punto con respecto a otros dos es, por tanto, un número real positivo.

Dado un número real positivo $r$ siempre puede encontrarse un punto $X$ tal que su razón con respecto a dos puntos dados $P$ y $Q$ es $r$. Para encontrar $X$ observamos que

$$r=\frac{\|\overrightarrow{PX}\|}{\|\overrightarrow{QX}\|}$$

y además:

$$X=(1-t)P+tQ=P+t\overrightarrow{QP}$$

para algún $t\in(0, 1)$. Para encontrar este valor de $t$ sustituimos $X$ en la fórmula que nos da la razón $r$ y obtenemos

$$r=\frac{\|t\overrightarrow{QP}\|}{\|(1-t)\overrightarrow{QP}\|}=\frac{t}{1-t}$$

De aquí podemos despejar $t$ en función de $r$:

$$t=\frac{r}{1+r}$$

El punto medio del segmento $\overline{PQ}$ es un punto cuya razón es 1 con respecto a $P$ y $Q$ y, por tanto, $t=1/2$. De aquí deducimos que el punto medio de $\overline{PQ}$ es

$$X=\left(1-\frac{1}{2}\right)P+\frac{1}{2}Q=\frac{P+Q}{2}$$

EJEMPLO A. Queremos hallar un punto $X$ que tenga razón 2 con respecto a los puntos $P=(1, 2, 0)$ y $Q=(-1, 0, 3)$. El punto $X$ será de la forma $X=(1-t)P+tQ$, donde

$$t=\frac{r}{1+r}=\frac{2}{3}$$

Por tanto:

$$X=\left(1-\frac{2}{3}\right)(1, 2, 0)+\frac{2}{3}(-1, 0, 3)=\left(-\frac{1}{3}, \frac{2}{3}, 2\right).$$

Observar que $X$ es un punto que dista $\frac{2}{3}\|PQ\|$ de $P$ y $\frac{1}{3}\|PQ\|$ de $Q$.

* * *

Dados tres puntos no alineados $A$, $B$ y $C$, en el plano o en el espacio, determinan un triángulo. Los puntos del interior del triángulo pueden determinarse paramétricamente. Los puntos del lado $\overline{BC}$ tienen como ecuación paramétrica

$$(1-t)B+tC, \quad 0 \leqq t \leqq 1$$

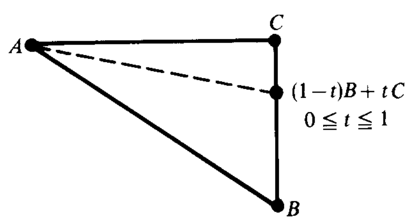



Figura 3

Uniendo $A$ con todos los puntos del lado $\overline{BC}$ se tiene el interior del triángulo, que satisface la ecuación

$$(1-s)A+s[(1-t)B+tC], \quad 0 \leqq s \leqq 1, \ 0 \leqq t \leqq 1$$

Esto puede escribirse de la forma

$$(1-s)A+s(1-t)B+stC, \quad 0 \leqq s \leqq 1, \ 0 \leqq t \leqq 1$$



o bien

$$aA+bB+cC, \quad a+b+c=1, \ 0 \leqq a, b, c \leqq 1.$$

La demostración de esta última afirmación se deja como ejercicio para el lector.

Cuatro puntos $A$, $B$, $C$ y $D$ no coplanarios en el espacio determinan un poliedro de cuatro caras triangulares. Puede demostrarse con un razonamiento análogo al anterior que los puntos de su interior se representan paramétricamente de la forma

$$aA+bB+cC+dD, \quad a+b+c+d=1, \ 0 \leqq a, b, c, d \leqq 1$$

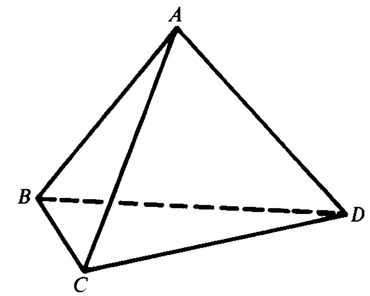



Figura 4

EJEMPLO B. Dados $A$, $B$ y $C$, tres puntos en el plano o en el espacio, queremos hallar las ecuaciones paramétricas de la región sombreada. El segmento que une $B$ con $C$ tiene como ecuaciones paramétricas

$$(1-t)B+tC, \quad 0 \leqq t \leqq 1.$$

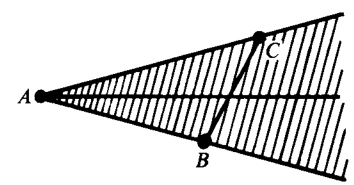



Figura 5

Trazando las semirrectas que unen $A$ con cualquiera de los puntos del segmento $\overline{BC}$ se obtiene la región sombreada. Estas semirrectas tienen por ecuación

$$(1-s)A+s[(1-t)B+tC], \quad 0 \leqq s, \ 0 \leqq t \leqq 1$$

Dado un triángulo cualquiera, se denominan *medianas* a las rectas que unen un vértice con el punto medio del lado opuesto.

Demostraremos a continuación que las tres medianas de un triángulo se cortan en un punto cuya distancia a cualquiera de los vértices del triángulo es 2/3 de la distancia del vértice al punto medio del lado opuesto.

Para demostrar este resultado sean $A$, $B$ y $C$ los vértices del triángulo; el punto medio del lado $\overline{BC}$ es $\dfrac{B+C}{2}$. El punto que se encuentra del vértice $A$ a $\dfrac{2}{3}$ de la distancia de $A$ al punto medio del lado $BC$ es (ver ejemplo A):

$$\left(1-\frac{2}{3}\right)A+\frac{2}{3}\,\frac{B+C}{2}=\frac{1}{3}\,A+\frac{B+C}{3}=\frac{A+B+C}{3}$$

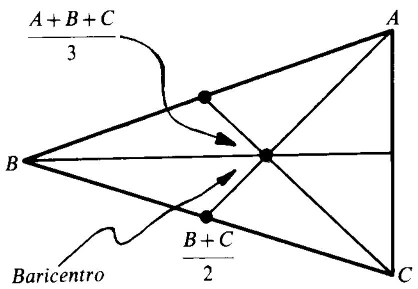



Figura 6

Realizando el razonamiento con cualquier otro vértice se obtendrá el mismo resultado. Esto demuestra nuestra afirmación.

El punto donde se cortan las *medianas* de un triángulo se denomina *baricentro* o *centro de gravedad* del triángulo.

Otros puntos característicos de un triángulo se obtienen como intersección de sus alturas, sus bisectrices o sus mediatrices. Los resultados que a ellos les conciernen se proponen en los ejercicios del final de esta sección.

* * *

Pasamos a continuación a estudiar las figuras más sencillas en el plano y en el espacio que no pueden formarse con segmentos de rectas.

En el plano tenemos la *circunferencia*, que es el lugar geométrico de los puntos que equidistan de un punto fijo, llamado *centro*. La distancia del centro a uno cualquiera de los puntos de la circunferencia recibe el nombre de *radio* de la circunferencia.

Si $C=(c_1, c_2)$ es el centro de la circunferencia y $r$ el radio, un punto $X=(x_1, x_2)$ está en la circunferencia si y sólo si

$$d(X, C)=r.$$

Por tanto:

$$(x_1-c_1)^2+(x_2-c_2)^2=r^2$$

es la ecuación implícita de la circunferencia de centro $c$ y radio $r$.



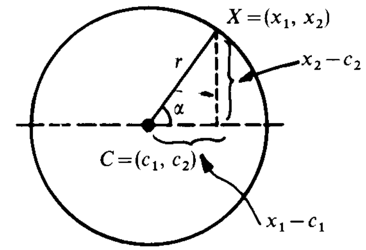



Figura 7

Ecuaciones paramétricas de la circunferencia pueden encontrarse utilizando el ángulo $\alpha$ que forman el radio de la circunferencia con una recta fijada que pase por el centro de la circunferencia. Se tiene que

$$x_1-c_1=r\cos\alpha,\quad x_2-c_2=r\operatorname{sen}\alpha.$$

Por tanto,

$$\begin{cases} x_1=c_1+r\cos\alpha \\ x_2=c_2+r\operatorname{sen}\alpha \end{cases}$$

son las ecuaciones paramétricas buscadas.

La figura que tiene a

$$\begin{cases} x_1=c_1+a\cos\alpha \\ x_2=c_2+b\operatorname{sen}\alpha \end{cases}$$

como ecuaciones paramétricas recibe el nombre de *elipse* (ver figura 8). Para encontrar sus ecuaciones cartesianas observamos que

$$\cos\alpha=\frac{x_1-c_1}{a}\quad \text{y}\quad \operatorname{sen}\alpha=\frac{x_2-c_2}{b}.$$

Por tanto,

$$1=\cos^2\alpha+\operatorname{sen}^2\alpha=\frac{(x_1-c_1)^2}{a^2}+\frac{(x_2-c_2)^2}{b^2}.$$

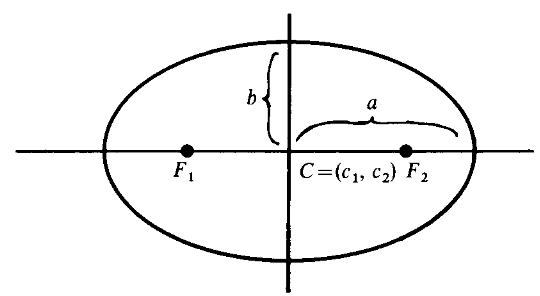



Figura 8

La elipse puede definirse de la siguiente manera: el lugar geométrico de los puntos de un plano cuya suma de distancias a dos puntos fijos, llamados focos, $F_1$ y $F_2$, es constante.

* * *

Otra curva que posee una definición «similar» a la de la elipse es la *lemniscata de Bernouilli:* es el lugar geométrico de los puntos del plano cuyas distancias a dos puntos fijos, llamados focos, $F_1$ y $F_2$, tienen producto igual a $c^2$, con $2c$ la distancia entre ambos focos.

Si $F_1=(-1, 0)$ y $F_2=(1, 0)$, la lemniscata de Bernouilli tiene por ecuación

$$[(x+1)^2+y^2][(x-1)^2+y^2]=1$$

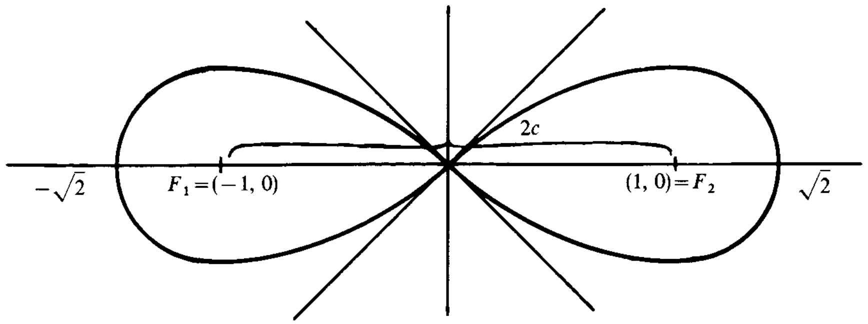



Figura 9

Simplificando esta igualdad se obtiene

$$(x^2+y^2)^2-2(x^2-y^2)=0$$

Escribiendo $x=r\cos\alpha$, $y=r \operatorname{sen}\alpha$ tenemos que

$$r^4=2r^2(\cos^2\alpha-\operatorname{sen}^2\alpha)$$

Por tanto:

$$r^2=2\cos 2\alpha.$$

Esta ecuación, denominada ecuación en coordenadas polares de la lemniscata de Bernouilli, permite obtener la representación gráfica de la figura adjunta.

* * *

La curva que describe un punto de una circunferencia cuando rueda sin deslizar sobre una recta tiene unas ecuaciones paramétricas sencillas. Esta curva recibe el nombre de *cicloide,* y su gráfica se aprecia en la figura adjunta.



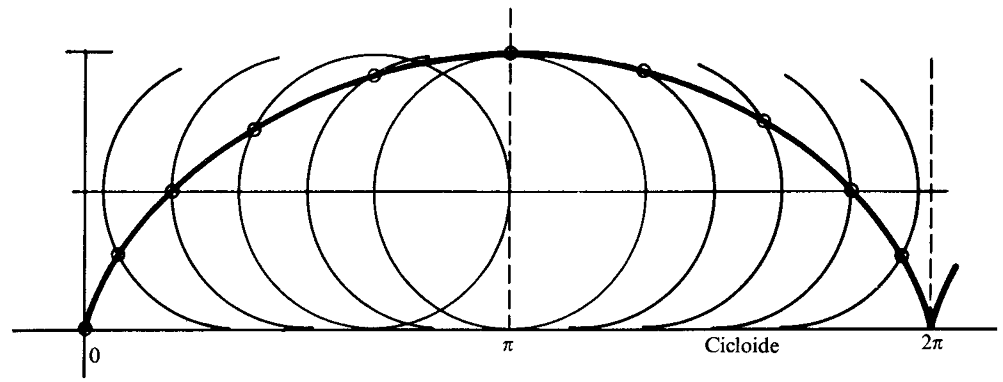



Figura 10

Si la circunferencia tiene radio 1 y está situada inicialmente como se muestra en la figura 11, después de que el centro de la circunferencia haya recorrido una longitud $t$, el punto $P$ se transforma en $P'$, cuyas coordenadas $x$ e $y$ satisfacen

$$x=t-d(B, C)=t-\cos\alpha=t-\cos\left(t-\frac{\pi}{2}\right)=t-\operatorname{sen} t$$

$$y=1+d(C, P')=1+\operatorname{sen}\alpha=1+\operatorname{sen}\left(t-\frac{\pi}{2}\right)=1-\cos t$$

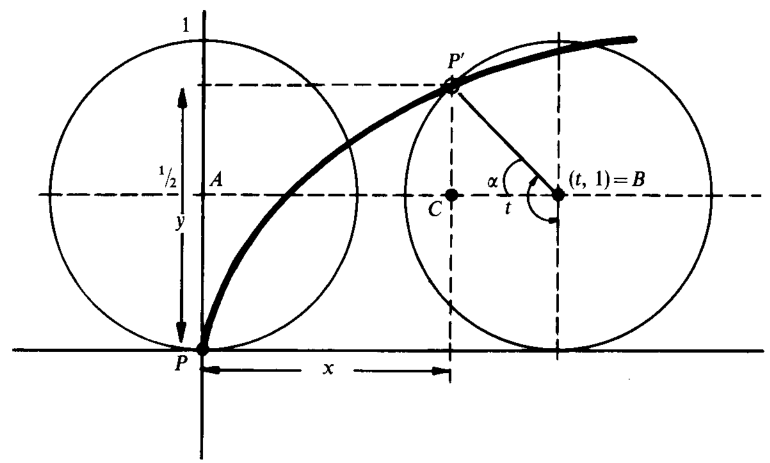



Figura 11

* * *

En el espacio la figura más sencilla es la *esfera,* que posee una definición similar a la de la circunferencia. La *esfera* es el lugar geométrico de los puntos del espacio que equidistan de un punto fijo, llamado *centro.* La distancia del centro a un punto cualquiera de la esfera recibe el nombre de *radio* de la esfera.

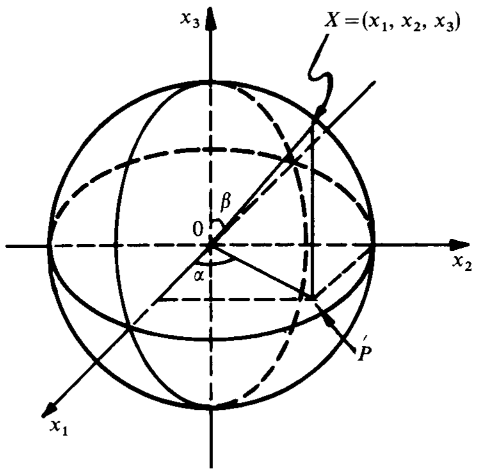



Figura 12

Si $C=(c_1, c_2, c_3)$ es el centro de la esfera y $r$ su radio, cualquier punto $X=(x_1, x_2, x_3)$ de ella satisface la igualdad

$$d(X, C)=r$$

Por tanto:

$$(x_1-c_1)^2+(x_2-c_2)^2+(x_3-c_3)^2=r^2$$

son las ecuaciones paramétricas de la esfera de centro $C$ y radio $r$.

Ecuaciones paramétricas de la esfera pueden encontrarse en función de los ángulos $\alpha$ y $\beta$ de la figura 12. Tenemos

$$x_3-c_3=r \cos \beta$$

mientras que $\|\overrightarrow{CP}\|=r$ sen $\beta$. De aquí deducimos que

$$\begin{cases} x_1-c_1=r \cos \alpha \operatorname{sen} \beta \\ x_2-c_2=r \operatorname{sen} \alpha \operatorname{sen} \beta \\ x_3-c_3=r \cos \beta \end{cases}$$

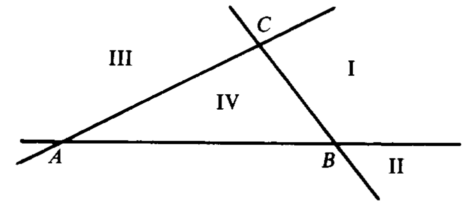
que son las ecuaciones paramétricas buscadas. Observar que estas ecuaciones paramétricas dependen de dos parámetros, $\alpha$ y $\beta$, lo cual está en concordancia con el hecho de que determinan una superficie en el espacio.

## EJERCICIOS 3.4

**1.** Encontrar los puntos que dividen al segmento $AB$ en tres partes iguales, donde $A=(-1, 0, -3)$ y $B=(4, 3, 5)$.



**2.** Describir paramétricamente las regiones I, II, III y IV de la figura adjunta, donde $A$, $B$ y $C$ son tres puntos no alineados.




**3.** Hallar la intersección de las siguientes figuras en el plano:

*a)* La recta $x=y+1$ y la circunferencia $(x-2)^2+y^2=6$.

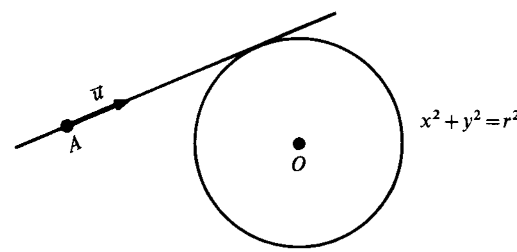
*b)* El semiplano $y-\frac{1}{\sqrt{5}}x \geqq \frac{4}{\sqrt{5}}$ y la circunferencia $(x-2)^2+y^2=6$.

*c)* El sector $s(4, 1)+(1-s-t)(1, 2)$, $t>0$, $s>0$ y el semiplano $x+2y-5<0$.

**4.** Demostrar que un criterio para que la recta $A+t\overline{u}$ toque a la circunferencia $x^2+y^2=r^2$ en un solo punto es que se cumpla la igualdad

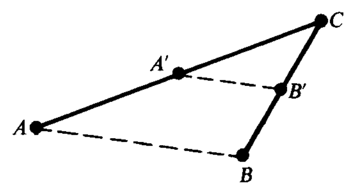
$$r^2\|\overline{u}\|^2=\begin{vmatrix} u_1 & u_2 \\ a_1 & a_2 \end{vmatrix}^2$$

donde $\overline{u}=(u_1, u_2)$ y $A=(a_1, a_2)$.

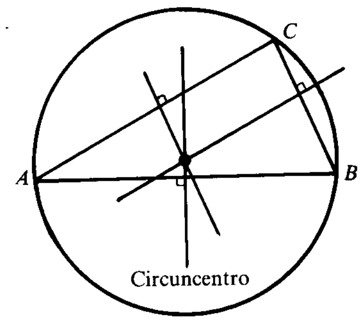



**5.** Demostrar que si $A'$ y $B'$ son los puntos medios de los segmentos $\overline{AC}$ y $\overline{BC}$, se tiene que

$$\|\overrightarrow{A'B'}\|=\frac{1}{2}\|\overrightarrow{AB}\|$$

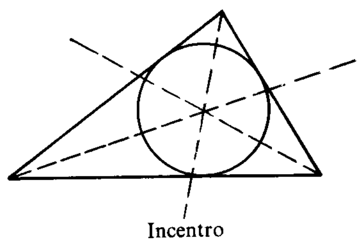


**6.** Se denomina *mediatriz* de un triángulo a la perpendicular por el punto medio a uno cualquiera de sus lados. Demostrar que las tres mediatrices de un triángulo se cortan en un punto que coincide con el centro de la circunferencia circunscrita al triángulo. (Este punto recibe el nombre de *circuncentro* del triángulo dado.)

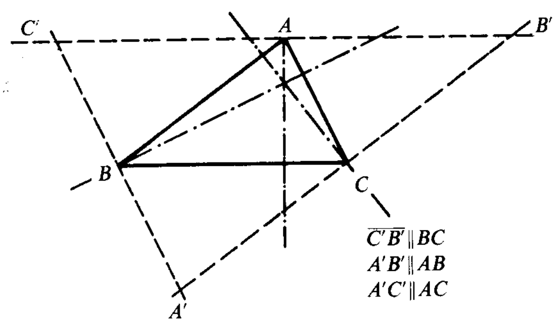



**7.** Demostrar que el circuncentro (ver problema anterior) de un triángulo rectángulo se halla en el punto medio de su hipotenusa.

**8.** Se denomina *bisectriz* de un triángulo a cualquier recta que divide a uno cualquiera de sus ángulos interiores en dos partes iguales. Demostrar que las bisectrices de un triángulo se cortan en un solo punto que coincide con el centro de la circunferencia inscrita en él. (Este punto recibe el nombre de *incentro* del triángulo dado.)




**9.** Cada una de las rectas que pasan por el vértice de un triángulo y son perpendiculares al lado opuesto reciben el nombre de *alturas* del triángulo $ABC$ de la figura. Estas alturas coinciden con las mediatrices del triángulo $A'B'C'$ de la figura adjunta. Deducir de este resultado y del problema 6 que las alturas de un triángulo se cortan en un punto. (Este punto se denomina *ortocentro* del triángulo dado.)






**10.** Dado el triángulo $ABC$ en el plano cuyos vértices tienen como coordenadas $A=(-3, 0)$, $B=(2, 1)$ y $C=(0, 3)$, hallar:

*a*) Las coordenadas del baricentro.
*b*) El circuncentro y la ecuación de la circunferencia circunscrita.
*c*) El incentro y la ecuación de la circunferencia inscrita.
*d*) El ortocentro.

(Ver los problemas 6, 8 y 9 para las definiciones.)

**11.** Hallar las ecuaciones de la altura relativa al vértice $A$ del prisma de base triangular $BCD$, donde

$$A=(1, 1, 3), \quad B=(0, -1, 0), \quad C=(1, 0, 0) \quad \text{y} \quad D=(0, 3, 1)$$

**12.** Encontrar las ecuaciones paramétricas y cartesianas de las siguientes figuras:

*a*) La circunferencia que pasa por los puntos $A=(2, 1)$, $B=(0, 3)$ y $C=(-1, 0)$.
*b*) La esfera que pasa por los puntos $A=(1, 2, 0)$, $B=(1, 0, 2)$, $C=(3, 0, 0)$ y $D=(2, 1, \sqrt{2})$.

## 3.5. AREAS Y VOLUMENES. PRODUCTO VECTORIAL

Comenzaremos encontrando una fórmula para determinar el área de un paralelogramo situado en el plano o en el espacio. Observamos, en primer lugar, que un paralelogramo queda determinado por dos vectores $\vec{a}$ y $\vec{b}$ que son linealmente independientes (ver figura 1).

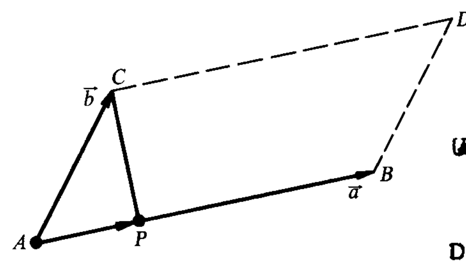



Figura 1



Sabemos que el área de un paralelogramo se obtiene multiplicando la longitud de su base por la altura; por tanto:

$$\text{área} = \|\vec{a}\| \cdot \|\overrightarrow{CP}\|$$

donde $P$ es la proyección del punto $C$ sobre la recta que contiene a $A$ y $B$. Puesto que $\|\overrightarrow{CP}\| = \|\vec{b}\| \operatorname{sen} \sphericalangle(\vec{a}, \vec{b})$, deducimos que

$$(\text{área})^2 = \|\vec{a}\|^2\|\vec{b}\|^2 \operatorname{sen}^2 \sphericalangle(\vec{a}, \vec{b}) = \|\vec{a}\|^2\|\vec{b}\|^2[1-\cos^2 \sphericalangle(\vec{a}, \vec{b})] =$$

$$= \|\vec{a}\|^2\|\vec{b}\|^2\left[1-\frac{(\vec{a}, \vec{b})^2}{\|\vec{a}\|^2\|\vec{b}\|^2}\right] = \|\vec{a}\|^2\|\vec{b}\|^2-(\vec{a}, \vec{b})^2.$$

Hemos obtenido el siguiente resultado:

PROPOSICIÓN 1

Si $\vec{a}$ y $\vec{b}$ son dos vectores linealmente independientes en el plano o en el espacio, el área $A$ del paralelogramo determinado por los vectores $\vec{a}$ y $\vec{b}$ se obtiene mediante la fórmula

$$A = \sqrt{\|\vec{a}\|^2\|\vec{b}\|^2-(\vec{a}, \vec{b})^2} = \sqrt{\begin{vmatrix} (\vec{a}, \vec{a}), & (\vec{a}, \vec{b}) \\ (\vec{a}, \vec{b}), & (\vec{b}, \vec{b}) \end{vmatrix}}$$

Si $\vec{a} = (a_1, a_2)$ y $\vec{b} = (b_1, b_2)$ son dos vectores linealmente independientes en *el plano* tenemos que

$$\|\vec{a}\|^2\|\vec{b}\|^2-(\vec{a}, \vec{b})^2 = (a_1^2+a_2^2)(b_1^2+b_2^2)-(a_1b_1+a_2b_2)^2 =$$
$$= a_1^2b_1^2+a_1^2b_2^2+a_2^2b_1^2+a_2^2b_2^2-a_1^2b_1^2-2a_1b_1a_2b_2-a_2^2b_2^2 =$$
$$= (a_1b_2-a_2b_1)^2 = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2$$

Por tanto, *el área de un paralelogrtamo en el plano coincide con el valor absoluto del determinante formado por las componentes de los vectores que lo determinan.*

En el espacio también existe una fórmula sencilla, pero es necesario utilizar un concepto que se introducirá más adelante.

El área de un triángulo de vértices $A$, $B$ y $C$ coincide con la mitad del área del paralelogramo determinado por los vectores $\overrightarrow{AC}$ y $\overrightarrow{AB}$, como puede observarse en la figura 2. Por tanto, las fórmulas para el área de un triángulo se obtienen dividiendo entre 2 las fórmulas para el área del paralelogramo que él determina.

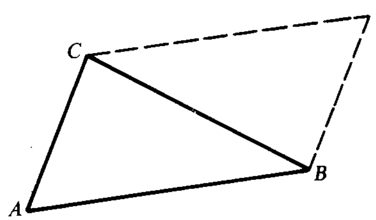


Figura 2



Finalmente, recordamos que el área de cualquier figura limitada por segmentos es fácil de calcular a partir del área de triángulos, ya que toda «figura» puede «triangularse» (ver figura 3).

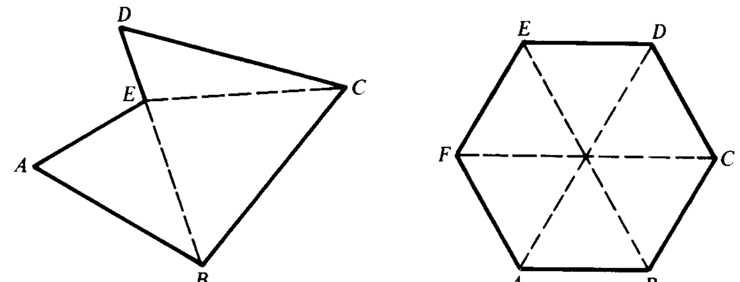


Figura 3

EJEMPLO A. Tratemos de encontrar el área del triángulo de vértices $A=(1, 2)$, $B=(-2, 0)$ y $C=(1, -3)$. Si consideramos los vectores

$$\vec{a} = \overrightarrow{AC} = (0, -5)$$

y

$$\vec{b} = \overrightarrow{AB} = (-3, -2)$$

se tiene que

$$\begin{vmatrix} 0 & -5 \\ -3 & -2 \end{vmatrix} = -15$$

Por tanto, el área pedida es 15/2.

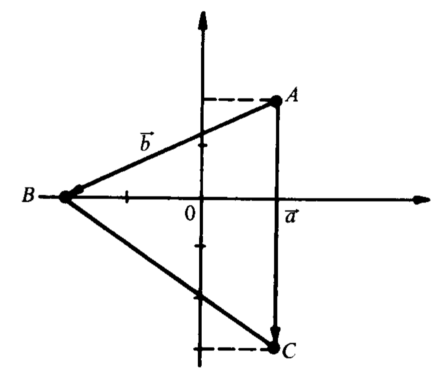


Figura 4

EJEMPLO B. Queremos encontrar el área del triángulo que tiene como vértices los puntos de intersección del plano de ecuación $2x+y+3z=6$ con los ejes coordenados (figura 5). Estos puntos son

$$A=(3, 0, 0), \quad B=(0, 6, 0) \quad \text{y} \quad C=(0, 0, 2)$$

como fácilmente puede comprobarse.

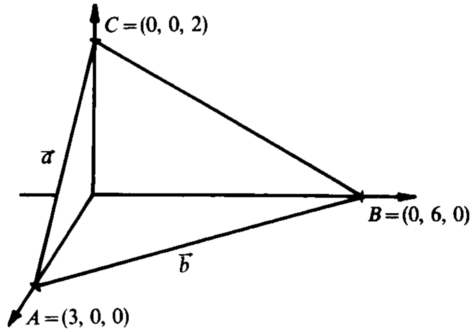



Figura 5

Si tomamos

$$\vec{a}=(-3, 0, 2) \quad \text{y} \quad \vec{b}=(-3, 6, 0)$$

de la proposición 1 deducimos que

$$A=\frac{1}{2}\sqrt{(9+0+4)(9+36+0)-(9+0+0)^2}=\frac{1}{2}\sqrt{13\cdot 45-81}=$$

$$=\frac{1}{2}\sqrt{504}=3\sqrt{14}$$

Observar que en este caso no podemos utilizar el determinante, mientras que sí se utilizó en el ejemplo A.

* * *

Una fórmula en la que interviene un determinante de orden 3 puede utilizarse para calcular el volumen de un paralelepípedo en el espacio. Un paralelepípedo queda determinado por tres vectores linealmente independientes (ver figura 6). Si estos vectores son

$$\vec{a}=(a_1, a_2, a_3), \quad \vec{b}=(b_1, b_2, b_3) \quad \text{y} \quad \vec{c}=(c_1, c_2, c_3)$$

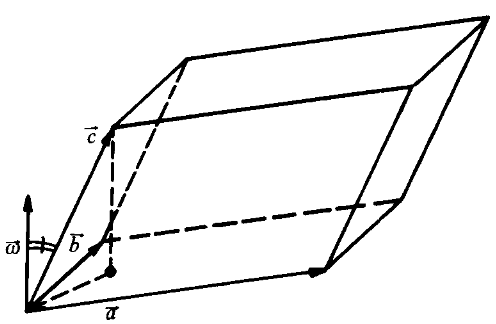



Figura 6



demostraremos que el volumen del paralelepípedo que ellos determinan se obtiene como el valor absoluto del determinante de la matriz.

$$(\vec{a}, \vec{b}, \vec{c})=\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix}$$

Para demostrarlo basta con encontrar un vector $\vec{\omega}$ perpendicular al plano determinado por los vectores $\vec{a}$ y $\vec{b}$, ya que entonces tendremos que

$$V=\text{volumen}=(\text{área de la base})\cdot \|\vec{c}\|\,|\cos \sphericalangle (\vec{c}, \vec{\omega})|$$

ya que la altura del paralelepípedo coincide con $\|\vec{c}\|\,|\cos \sphericalangle (\vec{c}, \vec{\omega})|$.

Para encontrar un vector $\vec{\omega}=(x_1, x_2, x_3)$ que sea perpendicular a $\vec{a}=(a_1, a_2, a_3)$ y $\vec{b}=(b_1, b_2, b_3)$ basta con resolver el sistema

$$\left.\begin{array}{l}(\vec{a}, \vec{\omega})=a_1x_1+a_2x_2+a_3x_3=0\\ (\vec{b}, \vec{\omega})=b_1x_1+b_2x_2+b_3x_3=0\end{array}\right\}$$

Puesto que los vectores $\vec{a}$ y $\vec{b}$ son linealmente independientes:

$$r\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}=2$$

y, por tanto, el sistema tiene infinitas soluciones. Supongamos que

$$\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\neq 0$$

y sea $x_3=t$; tenemos que

$$\omega=\frac{1}{\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}}\left(t\begin{vmatrix} -a_3 & a_2 \\ -b_3 & b_2 \end{vmatrix}, t\begin{vmatrix} a_1 & -a_3 \\ b_1 & -b_3 \end{vmatrix}, t\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\right)$$

Tomando $t=\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\neq 0$, por ejemplo, se tiene que

$$\vec{\omega}=\left(\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\right)$$

que recibe el nombre de *producto vectorial de* $\vec{a}$ y $\vec{b}$

DEFINICIÓN 1 *(Producto vectorial de dos vectores)*

Dados los vectores $\vec{a}=(a_1, a_2, a_3)$ y $\vec{b}=(b_1, b_2, b_3)$ en el espacio definimos su *producto vectorial* como el vector

$$\vec{a}\times\vec{b}=\left(\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, -\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}, \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}\right)$$

Una forma de recordar las componentes del vector producto vectorial de $\vec{a}$ y $\vec{b}$ es observar que corresponden al resultado de eliminar la primera, la segunda y la tercera columna, respectivamente, de la matriz

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix}$$

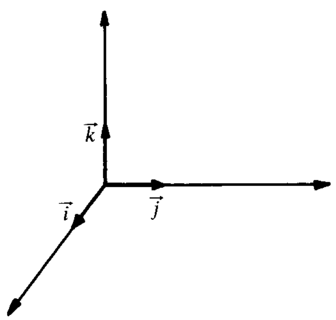
teniendo siempre cuidado de que a la segunda componente es necesario cambiarle el signo.

Otra forma de recordarlo, procedente de la física, es la siguiente: sean $\vec{i}=(1, 0, 0)$, $\vec{j}=(0, 1, 0)$ y $\vec{k}=(0, 0, 1)$ los vectores coordenados unitarios; podemos escribir

$$\vec{a}=a_1\vec{i}+a_2\vec{j}+a_3\vec{k}$$

y

$$\vec{b}=b_1\vec{i}+b_2\vec{j}+b_3\vec{k}$$




El vector $\vec{a}\times\vec{b}$ se obtiene desarrollando «formalmente» el determinante

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

por la primera fila.



EJEMPLO C. Tratemos de hallar la ecuación cartesiana del plano que pasa por el punto $A=(2, 1, 3)$ y tiene $\vec{u}=(1, 0, -1)$ y $\vec{v}=(2, -1, 1)$ como vectores directores. Un vector perpendicular a $\vec{u}$ y $\vec{v}$ es

$$\vec{u}\times\vec{v}=\left(\begin{vmatrix} 0 & -1 \\ -1 & 1 \end{vmatrix}, -\begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}, \begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix}\right)=(-1, -3, -1);$$

la ecuación de todos los planos perpendiculares a $\vec{u}\times\vec{v}$ es

$$-x-3y-z=d$$

El valor de $d$ se calcula con la condición de que el punto $A$ pertenezca al plano; por tanto, tenemos $-2-3-3=-8=d$. La ecuación del plano es

$$x+3y+z=8$$

* * *

El módulo o longitud del producto vectorial de dos vectores tiene una bonita interpretación geométrica. Si los vectores dados son $\vec{a}=(a_1, a_2, a_3)$ y $\vec{b}=(b_1, b_2, b_3)$, y calculamos $\|\vec{a}\times\vec{b}\|^2$ se tiene que

$$\begin{aligned}\|\vec{a}\times\vec{b}\|^2&=\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}^2+\begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix}^2+\begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}^2=(a_2b_3-a_3b_2)^2+(a_1b_3-a_3b_1)^2+\\ &+(a_1b_2-a_2b_1)^2=a_2^2b_3^2+a_3^2b_2^2+a_1^2b_3^2+a_3^2b_1^2+a_1^2b_2^2+a_2^2b_2^2-\\ &-2[a_2b_3a_3b_2+a_1b_3a_3b_1+a_1b_2a_2b_1]=(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2)-\\ &-a_1^2b_1^2-a_2^2b_2^2-a_3^2b_3^2-2[a_2b_3a_3b_2+a_1b_3a_3b_1+a_1b_2a_2b_1]=\\ &=(a_1^2+a_2^2+a_3^2)(b_2^2+b_2^2+b_3^2)-(a_1b_1+a_2b_2+a_3b_3)^2=\\ &=\|\vec{a}\|^2\|\vec{b}\|^2-(\vec{a}, \vec{b})^2\end{aligned}$$

De la proposición 1 deducimos el siguiente resultado:

PROPOSICIÓN 2 *(Area de un paralelogramo en el espacio)*

El área $A$ de un paralelogramo en el espacio determinado por dos vectores $\vec{a}$ y $\vec{b}$ coincide con el módulo del producto vectorial de los vectores, es decir:

$$A=\|\vec{a}\times\vec{b}\|=\|\vec{a}\|\|\vec{b}\|\operatorname{sen}\measuredangle(\vec{a}, \vec{b})$$

El área del triángulo del ejemplo B podemos calcularla ahora utilizando el producto vectorial. En efecto, tenemos que

$$A=\frac{1}{2}\|\vec{a}\times\vec{b}\|=\frac{1}{2}\left[\begin{vmatrix}0 & 2\\ 6 & 0\end{vmatrix}^2+\begin{vmatrix}-3 & 2\\ -3 & 0\end{vmatrix}^2+\begin{vmatrix}-3 & 0\\ -3 & 6\end{vmatrix}^2\right]^{1/2}=$$

$$=\frac{1}{2}[144+36+324]^{1/2}=\frac{1}{2}\sqrt{504}=3\sqrt{14}$$

que coincide con el resultado encontrado en el ejemplo B.

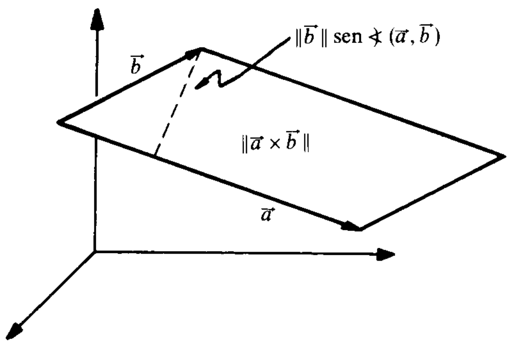



Figura 7

* * *

El problema que nos ha conducido a los resultados anteriores es el de encontrar una fórmula para determinar el volumen de un paralelepípedo en $\mathbb{R}^3$. Este problema puede ser resuelto ahora de una forma elegante. Para el paralelepípedo de la figura 8 se tiene que su volumen es

$$V=(\text{área de la base})\cdot\|\overrightarrow{CP}\||\cos\sphericalangle(\vec{a}\times\vec{b},\,\vec{c})|$$

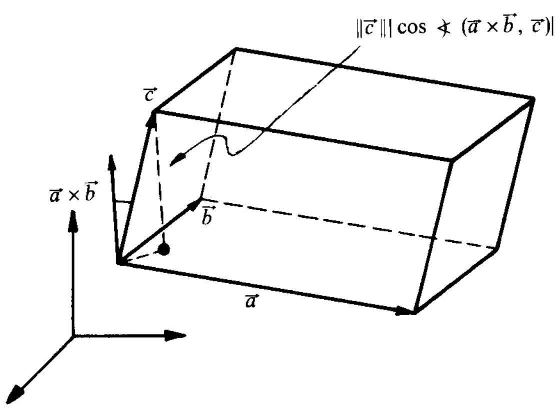



Figura 8



Debido a la proposición 2 se tiene que

$$V=\|\vec{a}\times\vec{b}\|\cdot\|\vec{c}\||\cos\sphericalangle(\vec{a}\times\vec{b},\,\vec{c})|$$

Si recordamos la fórmula para calcular el coseno del ángulo que forman dos vectores (ver sección 3.3), obtenemos que

$$V=|(\vec{a}\times\vec{b},\,\vec{c})|$$

El producto escalar de $\vec{a}\times\vec{b}$ con $\vec{c}$ recibe el nombre de *producto mixto* de los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$. El producto mixto de tres vectores puede calcularse utilizando un determinante, ya que

$$(\vec{a}\times\vec{b},\,\vec{c})=\left(\left(\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix},\,-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix},\,\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}\right),\,(c_1,c_2,c_3)\right)=$$

$$=\begin{vmatrix}a_2 & a_3\\ b_2 & b_3\end{vmatrix}c_1-\begin{vmatrix}a_1 & a_3\\ b_1 & b_3\end{vmatrix}c_2+\begin{vmatrix}a_1 & a_2\\ b_1 & b_2\end{vmatrix}c_3=\begin{vmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{vmatrix}$$

Todos estos resultados quedan resumidos en la siguiente proposición.

**PROPOSICIÓN 3** *(Volumen de un paralelepípedo)*

El volumen del paralelepípedo determinado por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$ en el espacio puede calcularse mediante la fórmula

$$V=|(\vec{a}\times\vec{b},\,\vec{c})|$$

y coincide con el valor absoluto del determinante de la matriz

$$\begin{pmatrix}a_1 & a_2 & a_3\\ b_1 & b_2 & b_3\\ c_1 & c_2 & c_3\end{pmatrix}$$

donde $\vec{a}=(a_1,a_2,a_3)$, $\vec{b}=(b_1,b_2,b_3)$ y $\vec{c}=(c_1,c_2,c_3)$.

EJEMPLO D. El volumen del paralelepípedo determinado por los vectores $\vec{a}=(1,2,-3)$, $\vec{b}=(0,1,2)$ y $\vec{c}=(1,-2,-1)$ es el valor absoluto de

$$\begin{vmatrix}1 & 2 & -3\\ 0 & 1 & 2\\ 1 & -2 & -1\end{vmatrix}=\begin{vmatrix}1 & 2\\ -2 & -1\end{vmatrix}+\begin{vmatrix}2 & -3\\ 1 & 2\end{vmatrix}=-1+4+4+3=10$$

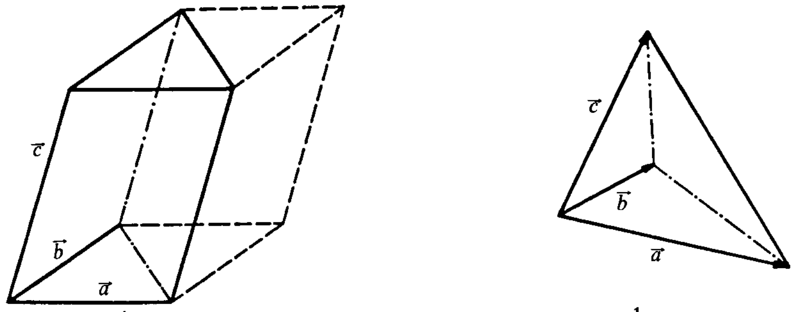
Por tanto, $V=10$.

* * *

Utilizando la proposición 3 puede calcularse el volumen de otras figuras en el espacio. Si se trata, por ejemplo, de un prisma triangular determinado por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$ (véase figura 9) se tiene que su volumen es

$$V=\frac{1}{2}|(\vec{a}\times\vec{b},\ \vec{c})|$$

ya que dos de ellos unidos por una de sus caras laterales forman un paralelepípedo determinado por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$.




$V=\frac{1}{2}|(\vec{a}\times\vec{b},\ \vec{c})|$ $\qquad$ $V=\frac{1}{6}|(\vec{a}\times\vec{b},\ \vec{c})|$

Figura 9

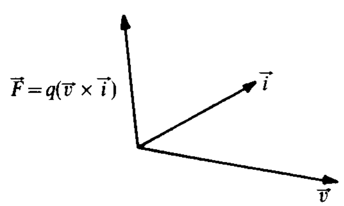
Si consideramos una pirámide de base triangular determinada por los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$, y tenemos en cuenta que el volumen de una pirámide es un tercio del área de la base por la altura, deducimos que

$$V=\frac{1}{6}|(\vec{a}\times\vec{b},\ \vec{c})|.$$

Otra forma de llegar a esta conclusión es observando que seis pirámides de base triangular e iguales forman un prisma de base rectangular, como puede apreciarse en las figuras 10 y 11. (Los dibujos de estas figuras han sido realizados por Manuel Moreno, 1.° de Físicas UAM, 1985.)

El producto vectorial, que nos ha aparecido al intentar calcular áreas y volúmenes de figuras tridimensionales, aparece de manera natural en algunos fenómenos físicos. Se justifica experimentalmente que una carga eléctrica positiva de magnitud $q$, que se mueve con una velocidad $\vec{v}$, dentro de un campo magnético de intensidad $\vec{i}$, se encuentra sometida a una fuerza $\vec{F}$ cuyo módulo y dirección están dados por

$$\vec{F}=q(\vec{v}\times\vec{i}).$$





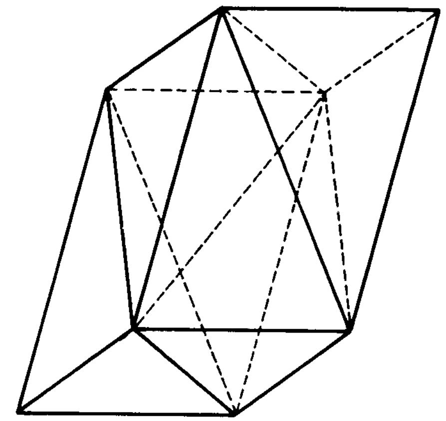



Figura 10

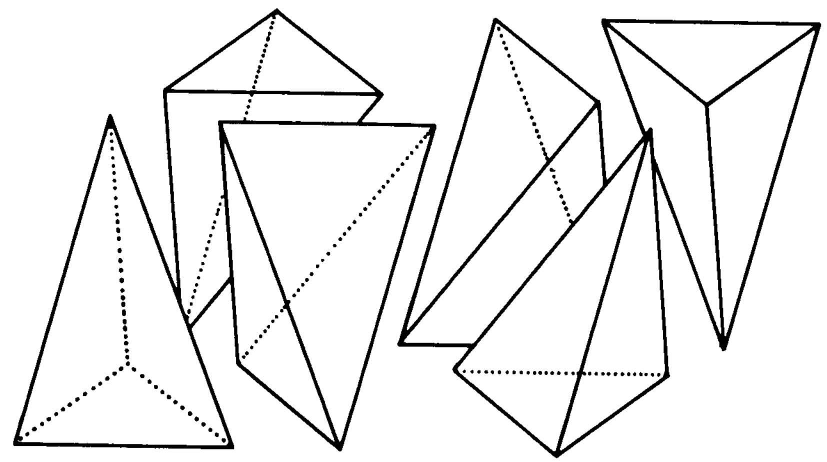
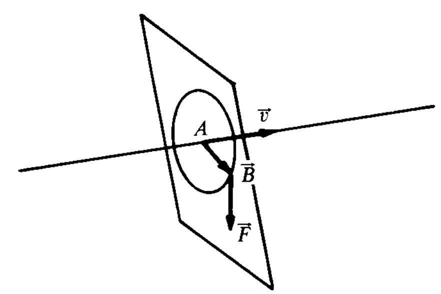


Figura 11

Una corriente que circula por un hilo recto produce un campo magnético cuyo vector $\vec{F}$ tiene el sentido del producto vectorial del vector que señala el sentido de la corriente, $\vec{v}$, en el hilo y el vector $\vec{AB}$ de la figura adjunta.



Estudiaremos a continuación las propiedades del producto vectorial. Sabemos que el producto vectorial tiene dirección perpendicular a cada uno de los vectores y su módulo o longitud coincide con el área del paralelogramo que ambos vectores determinan. Sin embargo, esto no determina completamente el producto vectorial de $\vec{a}$ y $\vec{b}$, ya que la dirección perpendicular a $\vec{a}$ y $\vec{b}$ posee dos sentidos opuestos. Se plantea entonces la pregunta de saber cuál es el sentido del vector $\vec{a}\times\vec{b}$.

Realizamos algunos ejemplos sencillos. Si $e_1=(1, 0, 0)$, $e_2=(0, 1, 0)$ y $e_3=(0, 0, 1)$, tenemos que

$$e_1\times e_2=\left(\begin{vmatrix}0 & 0\\ 1 & 0\end{vmatrix}, -\begin{vmatrix}1 & 0\\ 0 & 0\end{vmatrix}, \begin{vmatrix}1 & 0\\ 0 & 1\end{vmatrix}\right)=(0, 0, 1)=e_3$$

y

$$e_2\times e_1=\left(\begin{vmatrix}1 & 0\\ 0 & 0\end{vmatrix}, -\begin{vmatrix}0 & 0\\ 1 & 0\end{vmatrix}, \begin{vmatrix}0 & 1\\ 1 & 0\end{vmatrix}\right)=-e_3$$


En estos ejemplos se observa que el sentido del producto vectorial de $\vec{a}$ y $\vec{b}$ está dado por el sentido de avance de un tornillo que gira yendo de $\vec{a}$ hacia $\vec{b}$.

Otra forma de recordar el sentido del producto vectorial es la *regla de la mano derecha:* el sentido de $\vec{a}\times\vec{b}$ es el sentido en el que apunta el dedo pulgar cuando los dedos de la mano derecha están curvados de $\vec{a}$ hacia $\vec{b}$ (véase figura 12).

Las propiedades que conocemos del producto vectorial le determinan totalmente, es decir, el producto vectorial $\vec{a}\times\vec{b}$ de dos vectores $\vec{a}$ y $\vec{b}$ en $\mathbb{R}^3$ que son linealmente independientes es el único vector que satisface las siguientes propiedades:

1) Es perpendicular a $\vec{a}$ y $\vec{b}$.
2) Su longitud es el área del paralelogramo determinado por $\vec{a}$ y $\vec{b}$.
3) Su sentido está dado por la regla del tornillo.

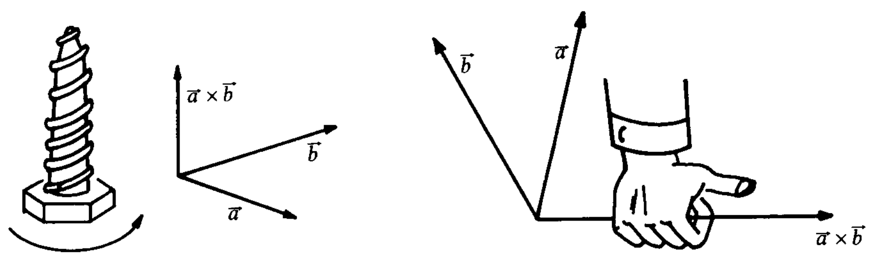



Figura 12



PROPOSICIÓN 4 *(Propiedades del producto vectorial)*

Cualesquiera que sean los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$ en $\mathbb{R}^3$ se tiene:

*a)* $\vec{a}\times\vec{b}=-\vec{b}\times\vec{a}$
*b)* $(\vec{a}\times\vec{b}, \vec{a})=0$ y $(\vec{a}\times\vec{b}, \vec{b})=0$
*c)* Si $\vec{a}$ y $\vec{b}$ son no nulos, $\vec{a}\times\vec{b}=\vec{0}$ si y sólo si $\vec{a}$ y $\vec{b}$ son paralelos.
*d)* $(\vec{a}+\vec{b})\times\vec{c}=\vec{a}\times\vec{c}+\vec{b}\times\vec{c}$
*e)* $(\vec{a}\times\vec{b}, \vec{c})=(\vec{a}, \vec{b}\times\vec{c})=(\vec{b}, \vec{c}\times\vec{a})$
*f)* $\vec{a}\times(\vec{b}\times\vec{c})=(\vec{a}, \vec{c})\vec{b}-(\vec{a}, \vec{b})\vec{c}$

*Nota.* Las propiedades *a), b)* y *c)* se deducen inmediatamente de las propiedades ya conocidas del producto vectorial. El resto de las propiedades pueden demostrarse directamente utilizando la definición de producto vectorial, y se dejan como ejercicio para el lector.

Existen varias formas de calcular la distancia entre dos rectas que se cruzan en el espacio; una de ellas se sugiere en el apartado *j)* del problema 3 de la sección 3.3. Otra de ellas puede obtenerse utilizando la proposición 3 de esta sección. Dadas la rectas de ecuaciones

$$r_1\colon A+t\vec{u} \quad \text{y} \quad r_2\colon B+s\vec{v}$$

formamos un paralelepípedo como se muestra en la figura 13. El volumen del paralelepípedo se obtiene multiplicando el área de la base por la altura $a$. De las proposiciones 2 y 3 deducimos que

$$V=|(\vec{u}\times\vec{v}, \overrightarrow{AB})|=\|\vec{u}\times\vec{v}\|\cdot a$$

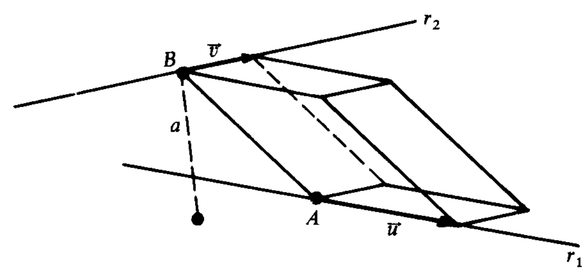



Figura 13

Puesto que la altura $a$ del paralelepípedo coincide con la distancia de $B$ al plano que contiene a $r_1$ y es paralelo a $r_2$ y ésta, a su vez, coincide con la distancia entre las rectas dadas, la distancia buscada es

$$d(r_1, r_2)=a=|(\vec{u}\times\vec{v}, \overrightarrow{AB})|/\|\vec{u}\times\vec{v}\|$$

EJEMPLO E. Para calcular la distancia entre las rectas de ecuaciones

$$r_1\colon (1, 0, 1)+t(2, 0, 1) \quad \text{y} \quad r_2\colon (0, 1, 1)+s(1, 3, -1)$$

calculamos

$$(\vec{u}\times\vec{v}, \overrightarrow{AB})=\begin{vmatrix} 2 & 0 & 1\\ 1 & 3 & -1\\ -1 & 1 & 0\end{vmatrix}=\begin{vmatrix} 0 & 2 & 1\\ 0 & 4 & -1\\ -1 & 1 & 0\end{vmatrix}=-\begin{vmatrix} 2 & 1\\ 4 & -1\end{vmatrix}=6$$

y

$$\|\vec{u}\times\vec{v}\| = \|(2, 0, 1)\times(1, 3, -1)\| = \left\| \left( \begin{vmatrix} 0 & 1\\ 3 & -1\end{vmatrix}, -\begin{vmatrix} 2 & 1\\ 1 & -1\end{vmatrix}, \begin{vmatrix} 2 & 0\\ 1 & 3\end{vmatrix} \right) \right\| =$$

$$= \|(-3, 3, 6\| = (9+9+36)^{1/2} = \sqrt{54}.$$

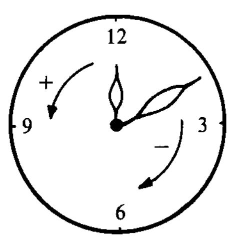
Por tanto,

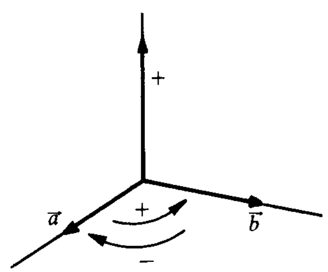
$$d(r_1, r_2)=\frac{6}{\sqrt{54}}=\frac{2}{\sqrt{6}}$$

* * *

Terminamos esta sección haciendo algunos comentarios sobre la orientación del plano y del espacio.

En un plano no hay una orientación privilegiada: podemos decir cuándo dos pares de vectores tienen la misma orientación, pero no decir cuál es la orientación de cada uno de ellos. Sin embargo, decimos que hay un sentido positivo y un sentido negativo de giro porque siempre miramos nuestro plano de un lado fijo: el papel por el lado en que dibujamos o las agujas del reloj desde fuera del mismo. ¡Si usamos un plano transparente y lo miramos desde atrás, los sentidos de giro cambian!



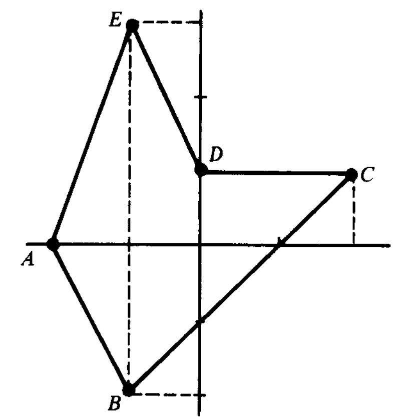

En el espacio no tiene sentido preguntar si el giro en torno a un eje es positivo o negativo y sólo tiene una aparente respuesta si en el eje se ha elegido una dirección positiva y el observador se sitúa en la dirección positiva de este eje. A pesar de no existir una respuesta satisfactoria para definir un sentido de giro positivo o negativo en el espacio, siempre podemos decidir si tres vectores están *positivamente orientados* utilizando el producto vectorial: los vectores $\vec{a}$, $\vec{b}$ y $\vec{c}$ están *positivamente orientados* si el vector $\vec{c}$ tiene el mismo sentido que el vector $\vec{a}\times\vec{b}$.



## EJERCICIOS 3.5

**1.** Hallar el área de la figura de vértices $ABCDE$, donde

$$A=(-2, 0),\quad B=(-1, -2),\quad C=(2, 1),\quad D=(0, 1),\quad E=(-1, 3)$$




**2.** Hallar las ecuaciones cartesianas de los siguientes conjuntos de puntos:

*a*) Plano que pasa por los puntos $A=(1, 2, 1)$, $B=(-1, 3, 0)$ y $C=(2, 1, 3)$.

*b*) Recta perpendicular a las rectas

$$(3, 3, 2)+t(3, -1, 2) \quad \text{y} \quad (1, 1, 9)+s(5, 1, -5)$$

y que pasa por su punto de intersección.

*c*) Plano perpendicular a la recta de ecuación $(1, 2, 1)+t(-1, 2, 0)$ que pasa por el punto $(-1, 2, 3)$.

**3.** Demostrar que la distancia de un punto $A$ a la recta $B+t\vec{u}$ en el plano puede calcularse mediante la fórmula

$$d=\frac{\begin{vmatrix} u_1 & u_2\\ a_1-b_1 & a_2-b_2\end{vmatrix}}{\|\vec{u}\|}$$

donde $A=(a_1, a_2)$, $B=(b_1, b_2)$ y $\vec{u}=(u_1, u_2)$.

[*Sugerencia:* utilizar que el área de un paralelogramo en el plano puede expresarse como un determinante.]

**4.** Dada la pirámide de base $ABCD$ y vértice $E$, donde $A=(2, 0, 0)$, $B=(3, 1, 0)$, $C=(0, 1, 0)$, $D=(-1, 0, 0)$ y $E=(1, 1, 3)$, hallar:

*a)* El área de la cara $ABE$.
*b)* El área de la base.
*c)* El volumen del prisma.
*d)* La distancia entre las rectas $EB$ y $DC$.
*e)* El valor de la altura.

**5.** Hallar el volumen del prisma determinado por los vectores

$$\vec{a}=(1, 2, -1), \quad \vec{b}=(0, 1, 2) \quad \text{y} \quad \vec{c}=(1, 2, -3)$$

**6.** Demostrar las siguientes propiedades del producto vectorial:

*a)* $(\vec{a}\times\vec{b}, \vec{c})=(\vec{a}, \vec{b}\times\vec{c})=(\vec{b}, \vec{c}\times\vec{a})$
*b)* $a\times(\vec{b}\times\vec{c})=(\vec{a}, \vec{c})\vec{b}-(\vec{a}, \vec{b})\vec{c}$

## CAPITULO 4

# LOS NUMEROS COMPLEJOS



Las sucesivas ampliaciones de los sistemas de números se han realizado para acomodar resultados sorprendentes en los sistemas de números conocidos. Estos resultados sorprendentes provienen, en la mayor parte de los casos, de la resolución de ecuaciones algebraicas.

Por ejemplo, la ecuación $x+7=5$, en la que solamente aparecen números naturales, no posee ningún número natural como solución; su solución es el número negativo $-2$. Los números naturales, junto con los números negativos, forman el sistema de los números enteros. Este sistema de números es insuficiente para resolver todas las ecuaciones algebraicas; la ecuación $3x=5$ no posee como solución ningún número entero; su solución es el número fraccionario $5/3$. Los números enteros, junto con los números fraccionarios, forman el conjunto de los números racionales. Estos números resultan insuficientes para resolver ecuaciones cuadráticas; por ejemplo, la ecuación $x^2=2$ no tiene un número racional como solución; sus soluciones son los números irracionales $\sqrt{2}$ y $-\sqrt{2}$. Los números racionales, junto con los irracionales, forman el sistema de los números reales.

En todos estos sistemas de números están definidas las operaciones de suma, multiplicación, resta y división, que generalizan las operaciones con números naturales.

Los números reales no son, sin embargo, suficientes para acoger en su seno las soluciones de toda ecuación cuadrática. La ecuación $x^2 = -1$ carece de toda solución real, ya que el cuadrado de un número real es siempre un número positivo. Nos vemos en la necesidad de ampliar el concepto de número para incluir aquellos que permitan resolver esta ecuación. La idea más sencilla, y a la vez genial, es definir un «nuevo» número, $i$, de manera que satisfaga la relación fundamental $i^2 = -1$. El nuevo sistema de números que se obtiene añadiendo éste y sus combinaciones a los números reales recibe el nombre de *sistema de números complejos*. Las operaciones que con ellos se realicen deben ser una generalización de las correspondientes operaciones con números reales. Este capítulo está dedicado al estudio de los números complejos.

## 4.1. LOS NUMEROS COMPLEJOS Y SUS PROPIEDADES

Un número complejo es toda expresión de la forma

$$z = a + bi$$

donde $a$ y $b$ son números reales e $i$ es un número que satisface $i^2 = -1$. El número real $a$ recibe el nombre de *parte real* del número complejo $z$, y se representa mediante $\mathrm{real}(z)$, y el número real $b$ recibe el nombre de *parte imaginaria* de $z$ y se representa mediante $\mathrm{img}(z)$.

La suma de dos números complejos es otro número complejo, que tiene como parte real la suma de las partes reales de cada uno de ellos, y como parte imaginaria la suma de las partes imaginarias de cada uno de ellos; así pues, si

$$z = a + bi, \quad z' = c + di$$

se tiene que

$$z + z' = (a + c) + (b + d)i.$$

El producto de dos números complejos es otro número complejo que se obtiene multiplicando los números complejos como si fueran polinomios en la variable $i$ y sustituyendo $i^2$ por $-1$ siempre que aparezca. Tenemos, entonces, que

$$z \cdot z' = (a + bi)(c + di) = ac + adi + bci + bdi^2 = (ac - bd) + (ad + bc)i.$$

EJEMPLO A. Queremos calcular $[(3+2i)^3 + (1-i)]^2$; podemos utilizar el binomio de Newton para calcular el cubo de $(3+2i)$ y obtenemos

$$(3+2i)^3 = 3^3 + 3 \cdot 3^2 \cdot 2i + 3 \cdot 3 \cdot (2i)^2 + (2i)^3 = 27 + 54i - 36 - 8i = -9 + 46i.$$



Por tanto:

$$[(3+2i)^3 + (1-i)]^2 = [(-9+46i) + (1-i)]^2 = (-8+45i)^2 = (-8)^2 - 45^2 + \\ + 2 \cdot (-8) \cdot 45i = -1.961 - 720i.$$

EJEMPLO B. Puesto que $i^2 = -1$, tenemos que $i^4 = (-1)^2 = 1$; por tanto, si $n$ es un número natural que al dividirlo entre 4 da de resto $r$, $0 \leqq r < 4$, es decir, $n = 4l + r$, donde $l$ es un número natural, se tiene que

$$i^n = i^{4l} i^r = i^r$$

Así, por ejemplo, $i^{431} = i^3$, ya que $431 = 4 \times 107 + 3$; puesto que $i^3 = i^2 \cdot i = -i$, se tiene que $i^{431} = -i$.

* * *

El cociente entre dos números complejos, de los cuales el divisor es no nulo, es otro número complejo; si deseamos hallar el cociente entre los números complejos $z = a + bi$ y $z' = c + di \neq 0$, escribimos

$$\frac{z}{z'} = \frac{a+bi}{c+di} = x + iy.$$

El número complejo $x + iy$ es el cociente entre $z$ y $z'$ si se cumple que

$$(c + di)(x + iy) = a + bi.$$

Utilizando la definición de multiplicación de números complejos e igualando las partes reales e imaginarias se tiene que

$$\left.\begin{array}{l} cx - dy = a \\ dx + cy = b \end{array}\right\}.$$

Es fácil obtener las soluciones de este sistema:

$$x = \frac{ca + db}{c^2 + d^2}, \quad y = \frac{cb - ad}{c^2 + d^2}$$

(Observar que la solución es única, ya que $c^2 + d^2 \neq 0$ por ser $z' \neq 0$.) Por tanto:

$$\frac{a+bi}{c+di} = \frac{ca+db}{c^2+d^2} + i\frac{cb-ad}{c^2+d^2}.$$

La fórmula para calcular el cociente de dos números complejos no es muy difícil de recordar; sin embargo, es mucho más fácil calcular el cociente de dos números complejos utilizando la noción de conjugado. Se denomina *conjugado* del número complejo $z'=c+di$ al número complejo

$$\bar{z'}=c-di$$

que posee la misma parte real que $z'$, pero su parte imaginaria se ha cambiado de signo.

Al multiplicar un número complejo $z'=c+di$ por su conjugado se tiene

$$z'\cdot\bar{z'}=(c+di)(c-di)=c^2+d^2$$

que coincide con el denominador de las partes real e imaginaria del cociente entre $z$ y $z'$. Esto sugiere que para calcular el cociente entre dos números complejos hasta multiplicar el numerador y el denominador por el conjugado del denominador. En efecto, este cálculo proporciona el resultado adecuado, ya que

$$\frac{a+bi}{c+di}=\frac{a+bi}{c+di}\cdot\frac{c-di}{c-di}=\frac{(ac+bd)+i(bc-ad)}{c^2+d^2}=\frac{ac+bd}{c^2+d^2}+i\frac{bc-ad}{c^2+d^2}$$

EJEMPLO C. Queremos expresar el número complejo

$$\frac{(1-i)(1+2i)}{1+i}$$

en la forma $a+bi$; el numerador da como resultado

$$(1-i)(1+2i)=1+2i-i-2i^2=1+2+i=3+i.$$

Multiplicando numerador y denominador por $1-i$ se tiene

$$\frac{(1-i)(1+2i)}{1+i}=\frac{3+i}{1+i}\cdot\frac{1-i}{1-i}=\frac{3-3i+i-i^2}{2}=\frac{4-2i}{2}=2-i$$

* * *

La suma de números complejos cumple las siguientes propiedades, que son semejantes a las propiedades de la suma de números reales:

(S1) Asociativa:

$$(a+bi)+[(c+di)+(e+fi)]=[(a+bi)+(c+di)]+(e+fi)$$

(S2) Conmutativa:

$$(a+bi)+(c+di)=(c+di)+(a+bi)$$



(S3) Existencia de elemento neutro: el número complejo $0=0+0i$ satisface

$$(a+bi)+0=0+(a+bi)=a+bi$$

(S4) Existencia de elemento opuesto: el opuesto de $a+bi$ es $-a-bi$.

La multiplicación de números complejos tiene propiedades análogas a las de la suma; son las siguientes:

(M1) Asociativa:

$$(a+bi)[(c+di)(e+fi)]=[(a+bi)(c+di)](e+fi)$$

(M2) Conmutativa:

$$(a+bi)(c+di)=(c+di)(a+bi)$$

(M3) Existencia de elemento unidad: el número complejo $1=1+0i$ satisface

$$(a+bi)1=1(a+bi)=a+bi$$

(M4) Existe el inverso de todo número complejo distinto de cero: el inverso de un número complejo $z=a+bi$, no nulo, es otro número complejo $x+yi$ tal que

$$(a+bi)(x+yi)=1.$$

Por tanto,

$$x+yi=\frac{1}{a+bi}=\frac{a-bi}{a^2+b^2}=\frac{a}{a^2+b^2}+i\frac{(-b)}{a^2+b^2}$$

de acuerdo con los resultados anteriores sobre el cociente de dos números complejos.

Finalmente, la suma y la multiplicación de números complejos están relacionadas mediante la siguiente propiedad:

(D1) Distributiva:

$$(a+bi)[(c+di)+(e+fi)]=(a+bi)(c+di)+(a+bi)(e+fi).$$

El lector no tendrá gran dificultad en comprobar las propiedades anteriormente expuestas. Cuando un sistema de números con operaciones en él definidas satisface las propiedades anteriores recibe el nombre de *cuerpo*. Podemos hablar, entonces, del cuerpo de los números complejos o del cuerpo de los números reales.

* * *

La validez de las operaciones con números complejos era cuestionada por varios matemáticos anteriores al siglo XIX. El nombre de «imaginarios» que aún se da a los números complejos cuya parte real es nula es un vestigio de este escepticismo. Sin embargo, a comienzos del siglo XIX una sencilla interpretación geométrica de las operaciones con números complejos hizo desaparecer estas sospechas. Esta interpretación geométrica fue encontrada simultáneamente por Wessel (1745-1818), Argand (1768-1822) y Gauss (1777-1855).

Esta interpretación geométrica consiste en colocar la parte real de un número complejo en un eje y su parte imaginaria en otro eje perpendicular al primero. El primero de estos ejes se denomina *eje real* y el segundo *eje imaginario*. De esta forma todo números complejo queda representado en un plano mediante un vector, como puede apreciarse en la figura 1.

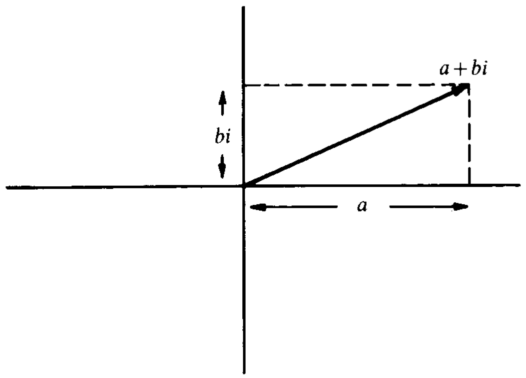



Figura 1

La suma de dos números complejos es un número complejo cuya representación gráfica coincide con la diagonal del paralelogramo que se obtiene con los dos vectores dados (véase figura 2).

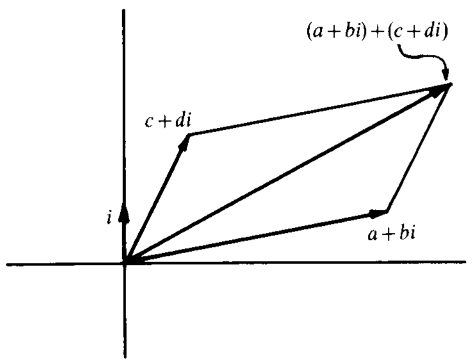



Figura 2



Para poder interpretar geométricamente la multiplicación de números complejos es necesario recurrir a una nueva forma de escribirlos. Esto se hará en la próxima sección.

EJERCICIOS 4.1

**1.** Calcular:

*a)* $[(2-3i)^3-i]^2$ *b)* $\left(\frac{1}{2}-i\frac{\sqrt{3}}{2}\right)^3$ *c)* $[(2+i)(2-i)]^2$

**2.** Expresar los siguientes números complejos en la forma $a+bi$:

*a)* $\dfrac{1-i}{1+i}$ *b)* $\dfrac{(3-i)(2+i)}{3+i}$ *c)* $\dfrac{(2-i)^2}{(3-i)^2}$

**3.** Encontrar las partes reales e imaginarias de

$$z^2, \quad \frac{1}{z^2} \quad \text{y} \quad \frac{z+1}{z-1}$$

donde $z=a+bi$.

**4.** Encontrar dos números complejos tales que su cuadrado sea $8-6i$.

**5.** Demostrar las siguientes igualdades:

*a)* $\overline{z_1+z_2}=\bar{z}_1+\bar{z}_2$ y *b)* $\overline{z_1 z_2}=\bar{z}_1\bar{z}_2$

donde $z_1$ y $z_2$ son dos números complejos cualesquiera.

## 4.2. FORMAS TRIGONOMETRICA Y POLAR DE UN NUMERO COMPLEJO

Se denomina *módulo* de un número complejo $z=a+bi$ a la longitud del vector que le representa, y se escribe de la forma $|z|$. Utilizando el teorema de Pitágoras se tiene que

$$r=|z|=\sqrt{a^2+b^2}$$

Todo número complejo posee un módulo positivo, excepto el número complejo $z=0$, que posee módulo nulo.

El ángulo que forma la dirección positiva del eje real con el vector que representa al número complejo $z$ se denomina *argumento* de $z$ y se representa mediante

$$\alpha=\arg(z).$$

El argumento de un número complejo $z$ no está determinado unívocamente, sino que puede variar en un múltiplo de $360° = 2\pi$ radianes. En cualquier caso, la trigonometría elemental nos permite obtener:

$$\operatorname{tg} \alpha = \frac{b}{a}$$

siempre que $a$ sea no nulo.

En el triángulo de la figura 1 puede observarse que

$$a = r \cos \alpha \quad \text{y} \quad b = r \operatorname{sen} \alpha$$

y, por tanto:

$$z = r[\cos \alpha + i \operatorname{sen} \alpha].$$

que recibe el nombre de *forma trigonométrica* del número complejo $z$.

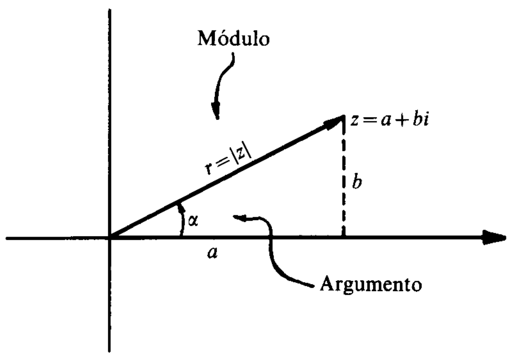



Figura 1

EJEMPLO A. Queremos escribir el número complejo $z = -2 + 2i$ en forma trigonométrica; su módulo es $r = \sqrt{4+4} = 2\sqrt{2}$ y su argumento $\alpha$ satisface $\operatorname{tg} \alpha = -\frac{2}{2} = -1$; por tanto, como el número complejo está en el segundo cuadrante, se tiene $\alpha = 135°$. Así pues,

$$-2 + 2i = 2\sqrt{2}\,[\cos 135° + i \operatorname{sen} 135°].$$

El lector puede también comprobar que

$$-2i = 2[\cos 270° + i \operatorname{sen} 270°]$$

obteniéndose así la forma trigonométrica del número complejo $z = -2i$.



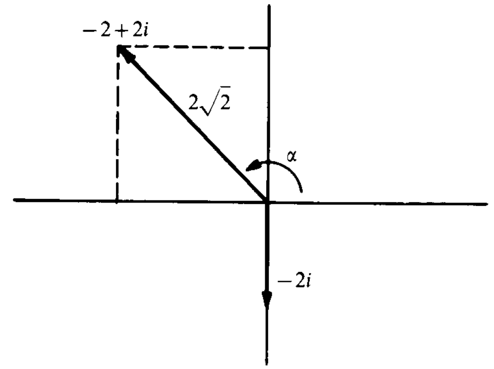



Figura 2

* * *

Dados los números complejos $z_1 = r(\cos \alpha + i \operatorname{sen} \alpha)$ y $z_2 = s(\cos \beta + i \operatorname{sen} \beta)$ en su forma trigonométrica el resultado de multiplicarlos produce

$$z_1 \cdot z_2 = rs[(\cos \alpha \cos \beta - \operatorname{sen} \alpha \operatorname{sen} \beta) + i(\cos \alpha \operatorname{sen} \beta + \operatorname{sen} \alpha \cos \beta)].$$

Utilizando las fórmulas trigonométricas para el coseno y el seno de una suma de ángulos se tiene que

$$z_1 \cdot z_2 = rs[\cos(\alpha + \beta) + i \operatorname{sen}(\alpha + \beta)]$$

De aquí se deduce que el módulo de un producto de números complejos es el producto de los módulos de cada uno de ellos y que su argumento es la suma de los argumentos de cada uno de ellos. Esta afirmación tiene una representación geométrica sencilla que puede observarse en la figura 3.

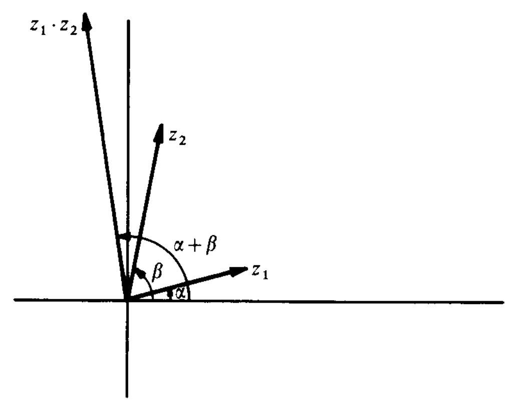



Figura 3

Para interpretar geométricamente el cociente $z_1/z_2$, donde $z_2$ es nulo, calculamos primero el inverso de $z_2$. Puesto que

$$z_2 = s(\cos\beta + i\,\mathrm{sen}\,\beta)$$

se tiene que

$$\frac{1}{z_2} = \frac{1}{s(\cos\beta + i\,\mathrm{sen}\,\beta)} = \frac{s(\cos\beta - i\,\mathrm{sen}\,\beta)}{s^2} = \frac{1}{s}[\cos(-\beta) + i\,\mathrm{sen}(-\beta)].$$

Por tanto,

$$\frac{z_1}{z_2} = z_1 \cdot \frac{1}{z_2} = \frac{r}{s}[\cos(\alpha-\beta) + i\,\mathrm{sen}(\alpha-\beta)].$$

De aquí se deduce que el módulo de un cociente de números complejos es el cociente de los módulos de cada uno de ellos y que su argumento se obtiene restando del argumento del dividendo el argumento del divisor.

Hemos probado que el módulo del producto de dos números complejos es el producto de los módulos de cada uno de ellos. El módulo de la suma es, sin embargo, menor o igual que la suma de los módulos de cada uno de ellos, es decir:

$$|z_1 + z_2| \leqq |z_1| + |z_2|$$

dándose la igualdad únicamente en algunos casos particulares. La desigualdad anterior es una inmediata consecuencia de aplicar la desigualdad triangular al triángulo $OAB$ (véase sección 3 del capítulo 3).

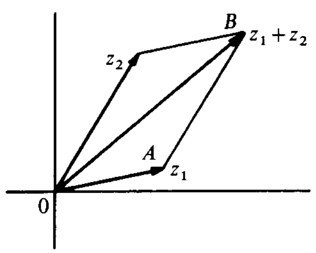



Figura 4

Particularmente interesantes son algunos resultados que se obtienen utilizando el conjugado de un número complejo $z = a + bi$. Se tiene que

$$z + \bar{z} = (a+bi) + (a-bi) = 2a = 2\,\mathrm{real}\,(z)$$
$$z - \bar{z} = (a+bi) - (a-bi) = 2ib = 2i\,\mathrm{img}\,(z)$$

y

$$z \cdot \bar{z} = (a+bi)(a-bi) = a^2 + b^2 = |z|^2$$

* * *



Finalizamos esta sección dando una nueva forma de escribir un número complejo. En cualquier libro de análisis matemático puede encontrarse la fórmula

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots, \quad x \in \mathbb{R}$$

lo cual es una aplicación de la fórmula de Taylor para funciones de una variable. De la misma manera puede demostrarse que

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$$

y

$$\mathrm{sen}\,x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots.$$

Por tanto:

$$r[\cos\alpha + i\,\mathrm{sen}\,\alpha] = r\left[\left(1 - \frac{\alpha^2}{2!} + \frac{\alpha^4}{4!} - \frac{\alpha^6}{6!} + \cdots\right) + i\left(\alpha - \frac{\alpha^3}{3!} + \frac{\alpha^5}{5!} - \frac{\alpha^7}{7!} + \cdots\right)\right] =$$

$$= r\left[1 + i\alpha + \frac{(i\alpha)^2}{2!} + \frac{(i\alpha)^3}{3!} + \frac{(i\alpha)^4}{4!} + \frac{(i\alpha)^5}{5!} + \cdots\right].$$

La similitud de esta última expresión entre corchetes, con la fórmula para $e^x$ anteriormente dada, nos lleva a dar la siguiente definición:

$$re^{i\alpha} = r[\cos\alpha + i\,\mathrm{sen}\,\alpha],$$

que recibe el nombre de fórmula de Euler.

La expresión $re^{i\alpha}$ se denomina *forma polar* de un número complejo; en ella está contenida toda la información necesaria para que el número complejo quede determinado: están dados su módulo y su argumento.

## EJERCICIOS 4.2

**1.** Calcular su módulo, su argumento y expresar los siguientes números complejos en sus formas trigonométrica y polar:

$$1+i, \quad \frac{1}{2} - \frac{\sqrt{3}}{2}i, \quad -\frac{\sqrt{3}}{2} - \frac{1}{2}i, \quad -2-2i$$

**2.** Calcular $e^{\pi i}$, $e^{2\pi i}$, $e^{\frac{\pi}{4}i}$ y $e^{-\frac{\pi}{4}i}$.

4. Hallar el módulo y el argumento de los siguientes números complejos, sin realizar las operaciones:

$$(1+i)(1-i), \quad \frac{1+\sqrt{3}i}{1-\sqrt{3}i}, \quad i^{431}(3-3i)$$

5. Si $z_1$ y $z_2$ son dos números complejos, demostrar que:

*a)* $|z_1|-|z_2| \leqq |z_1+z_2|$

*b)* $||z_1|-|z_2|| \leqq |z_1-z_2|$

6. Probar, sin efectuar cálculos, que si $a-bi$ es no nulo, $a+bi/a-bi$ tiene siempre módulo 1.

7. Demostrar que

$$\cos\alpha = \frac{e^{i\alpha}+e^{-i\alpha}}{2} \quad \text{y} \quad \operatorname{sen}\alpha = \frac{e^{i\alpha}-e^{-i\alpha}}{2i}$$

utilizando la fórmula de Euler.

## 4.3. RAICES DE NUMEROS COMPLEJOS

Tratemos ahora de encontrar la raíz $n$-ésima de un número complejo; este problema tiene una solución sencilla si el número complejo está escrito en forma trigonométrica.

Observemos, primero, que si deseamos calcular la potencia $n$-ésima del número complejo

$$z=r[\cos\alpha+i\operatorname{sen}\alpha]$$

se tiene que

$$z^n=r^n[\cos n\alpha+i\operatorname{sen} n\alpha].$$

Esta fórmula, que recibe el nombre de *fórmula de De Moivre* (matemático inglés, 1667-1754), se deduce del hecho de que al multiplicar dos números complejos, el módulo es el producto de cada uno de sus módulos y su argumento es la suma de cada uno de sus argumentos. Este resultado fue demostrado en la sección anterior.

Demostraremos a continuación que todo número complejo posee $n$ raíces $n$-ésimas que son también números complejos. Dado el número complejo

$$\omega=s[\cos\beta+i\operatorname{sen}\beta]$$



una raíz $n$-ésima de $\omega$ es cualquier número complejo $z$ que satisface $z^n=\omega$. Para encontrar $z$ escribimos

$$z=r[\cos\alpha+i\operatorname{sen}\alpha]$$

e imponemos la condición de que su potencia $n$-ésima sea $\omega$; utilizando la fórmula de De Moivre se deduce que

$$r^n[\cos n\alpha+i\operatorname{sen} n\alpha]=s[\cos\beta+i\operatorname{sen}\beta].$$

Igualando los módulos y los argumentos de estos dos números complejos se tiene que

$$r^n=s$$

y

$$n\alpha=\beta+2k\pi, \quad k=0, \pm1, \pm2, \pm\cdots$$

en donde se ha tenido en cuenta que dos números complejos iguales poseen argumentos que difieren en un múltiplo entero de $360°=2\pi$ radianes.

Puesto que $r$ y $s$ son números reales positivos se tiene que

$$r=\sqrt[n]{s}$$

donde $\sqrt[n]{s}$ denota la única raíz real positiva del número real positivo $s$. Además,

$$\alpha=\frac{\beta}{n}+\frac{2k\pi}{n}, \quad k=0, \pm1, \pm2, \ldots .$$

Por tanto, los números complejos

$$\omega_k=\sqrt[n]{s}\left[\cos\left(\frac{\beta}{n}+\frac{2k\pi}{n}\right)+i\operatorname{sen}\left(\frac{\beta}{n}+\frac{2k\pi}{n}\right)\right], \quad k=0, \pm1, \pm2, \ldots$$

son raíces $n$-ésimas de $\omega$.

En la colección infinita de números complejos $\{\omega_k\}$ hay una cantidad infinita de ellos que se repiten. Observar, por ejemplo, que si $k=qn+r$, $k \geqq n$, $0 \leqq r \leqq n-1$, se tiene que

$$\cos\left(\frac{\beta}{n}+\frac{2k\pi}{n}\right)=\cos\left(\frac{\beta}{n}+2q\pi+\frac{2r\pi}{n}\right)=\cos\left(\frac{\beta}{n}+\frac{2r\pi}{n}\right)$$

y, análogamente,

$$\operatorname{sen}\left(\frac{\beta}{n}+\frac{2k\pi}{n}\right)=\operatorname{sen}\left(\frac{\beta}{n}+2q\pi+\frac{2r\pi}{n}\right)=\operatorname{sen}\left(\frac{\beta}{n}+\frac{2r\pi}{n}\right).$$

Por tanto, $\omega_k=\omega_r$. De manera similar puede demostrarse que si $k$ es un entero negativo, $\omega_k$ coincide con algún $\omega_r$, con $0 \leqq r \leqq n-1$. Además, todos los $\omega_r$, con $r \in \{0, 1, 2, ..., n-1\}$ son distintos, cómo puede comprobarse fácilmente. Así pues, $\omega$ posee $n$-raíces complejas que escritas en forma polar son

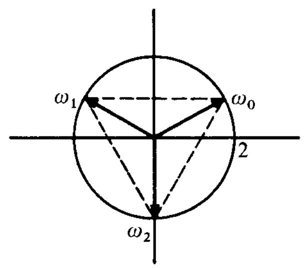
$$\omega_k=\sqrt[n]{s}\,e^{\frac{\beta+2k\pi}{n}}, \quad k=0, 1, 2, ..., n-1$$

EJEMPLO A. Queremos calcular las raíces cúbicas del número complejo $8i$; puesto que $8i=8\left[\cos\frac{\pi}{2}+i\operatorname{sen}\frac{\pi}{2}\right]$ las tres raíces cúbicas de $i$ son:

$$\omega_0=2\left[\cos\frac{\pi}{6}+i\operatorname{sen}\frac{\pi}{6}\right]=2\left[\frac{\sqrt{3}}{2}+i\frac{1}{2}\right]$$

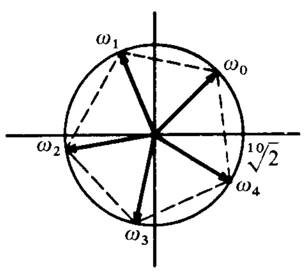
$$\omega_1=2\left[\cos\left(\frac{\pi}{6}+\frac{2\pi}{3}\right)+i\operatorname{sen}\left(\frac{\pi}{6}+\frac{2\pi}{3}\right)\right]=2\left[-\frac{\sqrt{3}}{2}+i\frac{1}{2}\right]$$

$$\omega_2=2\left[\cos\left(\frac{\pi}{6}+\frac{4\pi}{3}\right)+i\operatorname{sen}\left(\frac{\pi}{6}+\frac{4\pi}{3}\right)\right]=-2i$$




EJEMPLO B. Calculamos a continuación las raíces quintas del número complejo $z=-1-i$; puesto que

$$-1-i=\sqrt{2}\left(\cos\frac{5\pi}{4}+i\operatorname{sen}\frac{5\pi}{4}\right)$$






se tiene que las raíces pedidas son:

$$\omega_0=\sqrt[10]{2}\left[\cos\frac{\pi}{4}+i\operatorname{sen}\frac{\pi}{4}\right]=\sqrt[10]{2}\left[\frac{1}{\sqrt{2}}+i\frac{1}{\sqrt{2}}\right]$$

$$\omega_1=\sqrt[10]{2}\left[\cos\left(\frac{\pi}{4}+\frac{2\pi}{5}\right)+i\operatorname{sen}\left(\frac{\pi}{4}+\frac{2\pi}{5}\right)\right]=\sqrt[10]{2}\,[\cos 117^\circ+i\operatorname{sen}117^\circ]$$

$$\omega_2=\sqrt[10]{2}\left[\cos\left(\frac{\pi}{4}+\frac{4\pi}{5}\right)+i\operatorname{sen}\left(\frac{\pi}{4}+\frac{4\pi}{5}\right)\right]=\sqrt[10]{2}\,[\cos 189^\circ+i\operatorname{sen}189^\circ]$$

$$\omega_3=\sqrt[10]{2}\left[\cos\left(\frac{\pi}{4}+\frac{6\pi}{5}\right)+i\operatorname{sen}\left(\frac{\pi}{4}+\frac{6\pi}{5}\right)\right]=\sqrt[10]{2}\,[\cos 261^\circ+i\operatorname{sen}261^\circ]$$

$$\omega_4=\sqrt[10]{2}\left[\cos\left(\frac{\pi}{4}+\frac{8\pi}{5}\right)+i\operatorname{sen}\left(\frac{\pi}{4}+\frac{8\pi}{5}\right)\right]=\sqrt[10]{2}\,[\cos 333^\circ+i\operatorname{sen}333^\circ]$$

* * *

De la fórmula para calcular las raíces $n$-ésimas de un número complejo se deduce que todas ellas tienen el mismo módulo, a saber, $\sqrt[n]{r}$, y sus argumentos difieren en $2\pi/n$ radianes; por tanto, los extremos de estas raíces son los vértices de un polígono regular de $n$ lados inscrito en una circunferencia de radio $\sqrt[n]{r}$. Este comportamiento puede observarse en los ejemplos anteriores.

Un caso particular interesante es calcular las races $n$-ésimas de 1, las cuales se denominan *raíces n-ésimas de la unidad.* Puesto que 1 es un número complejo que tiene módulo 1, de la observación anterior se deduce que los extremos de las raíces $n$-ésimas de la unidad son los vértices de un polígono regular de $n$ lados inscrito en una circunferencia de radio 1.

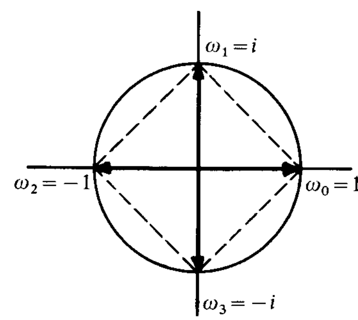
Claramente, una raíz $n$-ésima de la unidad es 1; el resto pueden calcularse con la fórmula anterior. Por ejemplo, las raíces cuartas de la unidad son:

$$\omega_0=1[\cos 0^\circ+i\operatorname{sen}0^\circ]=1$$

$$\omega_1=1\left[\cos\frac{2\pi}{4}+i\operatorname{sen}\frac{2\pi}{4}\right]=i$$

$$\omega_2=1\left[\cos\frac{4\pi}{4}+i\operatorname{sen}\frac{4\pi}{4}\right]=-1$$

$$\omega_3=1\left[\cos\frac{6\pi}{4}+i\operatorname{sen}\frac{6\pi}{4}\right]=-i$$




* * *

EJERCICIOS 4.3

1. Calcular los diferentes valores de

$$\sqrt{1-i},\quad \sqrt[3]{-8},\quad \sqrt[3]{-i},\quad \sqrt[4]{16i},\quad \sqrt{-9}$$

2. Calcular las raíces sextas de la unidad.

3. Calcular las raíces cúbicas de la unidad. Sean éstas $\omega_0=1$, $\omega_1$ y $\omega_2$; demostrar que $\omega_1$ y $\omega_2$ satisfacen la ecuación $x^2+x+1=0$ y ambas son conjugadas.

## 4.4. RESOLUCION DE ECUACIONES ALGEBRAICAS

Una ecuación algebraica es una ecuación de la forma

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=0 \tag{1}$$

donde los $a_j$ son números complejos; en particular, si los $a_j$ son números reales la ecuación anterior se dice que es una ecuación con coeficientes reales.

Se denomina *solución* de la ecuación algebraica (1) a cualquier número complejo $z$ tal que al sustituir $x$ por $z$ en la parte izquierda de la ecuación y realizar las operaciones indicadas se obtiene 0 como resultado.

Inicialmente podría pensarse que toda ecuación algebraica con coeficientes reales posee soluciones reales. Nada más lejos de la verdad, puesto que ya sabemos que la ecuación $x^2+1=0$ no posee soluciones reales. Sin embargo, posee las soluciones complejas $i$ y $-i$. Esto nos lleva a pensar que toda ecuación algebraica posee soluciones complejas. Este resultado se conoce con el nombre de *teorema fundamental del álgebra*, y se enuncia así:

*Toda ecuación algebraica de la forma (1) posee una solución compleja.*

La primera demostración correcta de este teorema es debida a C. F. Gauss; la demostración más sencilla requiere conocimientos relativos a funciones de variable compleja y se incluye en cualquier curso dedicado a este tema.

Si $z_1$ es una solución de (1), existen números complejos $b_0, b_1, \ldots, b_{n-1}$ tal que

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=(b_{n-1}x^{n-1}+\cdots+b_1x+b_0)(x-z_1).$$

Para demostrar este resultado multiplicamos los polinomios de la parte derecha de esta igualdad e igualamos sus coeficientes con los del polinomio de la parte izquierda. Por este procedimiento se obtienen las siguientes igualdades:

$$a_n=b_{n-1};\ a_{n-1}=b_{n-2}-z_1b_{n-1};\ a_{n-2}=b_{n-3}-z_1b_{n-2};\ \ldots;\ a_1=b_0-z_1b_1;\ a_0=-z_1b_0.$$



La primera igualdad nos permite encontrar $b_{n-1}(=a_n)$; conocido $b_{n-1}$ podemos encontrar $b_{n-2}$ de la segunda igualdad, ya que $b_{n-2}=a_{n-1}+z_1b_{n-1}=a_{n-1}+z_1a_n$; conocido $b_{n-2}$ podemos encontrar $b_{n-3}$ de la tercera igualdad, ya que

$$b_{n-3}=a_{n-2}+z_1b_{n-2}=a_{n-2}+z_1(a_{n-1}+z_1a_n)$$

Este proceso continúa hasta que se ha calculado $b_0$ de la penúltima igualdad:

$$b_0=a_1+z_1b_1=a_1+z_1(a_2+z_1b_2)\pm\cdots=a_1+a_2z_1+a_3z_1^2+\cdots+a_nz_1^{n-1}.$$

Esta igualdad coincide con $a_0=-z_1b_0$, ya que $z_1$ es solución de la ecuación (1). Este procedimiento para calcular el polinomio $b_{n-1}x^{n+1}+\cdots+b_1x+b_0$ se conoce con el nombre de *regla de Ruffini* y puede esquematizarse como sigue:

$$\begin{array}{r|cccccc}
 & a_n & a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \\
z_1) & & z_1a_n & z_1b_{n-2} & \cdots & z_1b_1 & z_1b_0 \\
\hline
 & a_n & b_{n-2} & b_{n-3} & \cdots & b_0 & \multicolumn{1}{|c}{0.}
\end{array}$$

Utilizando de nuevo el teorema fundamental del álgebra para la ecuación $b_{n-1}x^{n-1}+\cdots+b_1x+b_0=0$ se obtiene otra solución compleja $z_2$ de (1). Repitiendo este proceso $n$ veces podemos escribir

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0=a_n\cdot(x-z_1)(x-z_2)\cdots(x-z_n).$$

Algunas de las soluciones $z_j$ pueden repetirse, con lo cual podemos escribir

$$a_nx^n+a_{n-1}+\cdots+a_1x+a_0=a_n\cdot(x-z_1)^{c_1}(x-z_2)^{c_2}\cdots(x-z_k)^{c_k}.$$

El número $c_j$ se denomina *multiplicidad de la raíz* $z_j$, $j=1, 2, \ldots, k$. Tenemos, por tanto, el siguiente resultado:

*Toda ecuación algebraica de la forma (1) posee n soluciones complejas, donde cada solución está contada el número de veces que indica su multiplicidad.*

Encontrar las $n$ soluciones de una ecuación algebraica es un problema imposible de resolver en general. Sin embargo, en algunos casos particulares, que aparecen con mucha frecuencia, es posible encontrar las soluciones.

Si la ecuación algebraica (1) es de grado 2 y con coeficientes reales, es decir, de la forma

$$a_2x^2+a_1x+a_0=0,\quad a_0, a_1, a_2\in\mathbb{R}$$

la conocida fórmula

$$\frac{-a_1\pm\sqrt{a_1^2-4a_2a_0}}{2a_2}$$

nos permite obtener sus soluciones.

Si la ecuación algebraica es de grado 3 ó 4 y con coeficientes reales, también existen fórmulas para calcular sus soluciones; estas fórmulas resultan bastante complicadas. El lector interesado puede encontrarlas en el libro de A. G. Kurosch, *Curso de álgebra superior* (Editorial Mir, 1977).

Ecuaciones algebraicas de la forma

$$x^n - a = 0$$

donde $a$ es un número complejo se han resuelto en la sección anterior, ya que simplemente se trata de calcular las raíces $n$-ésimas del número complejo $a$.

Otros casos particulares pueden verse en los ejercicios al final de esta sección.

Para finalizar es conveniente exponer un método para encontrar las soluciones racionales, en particular las enteras, de una ecuación algebraica con coeficientes enteros, si es que ésta posee algunas soluciones del mencionado tipo.

Supongamos que la ecuación algebraica con coeficientes enteros

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0 \tag{1}$$

posee una solución racional de la forma $z = M/N$; siempre podemos suponer que $M$ y $N$ son primos entre sí, ya que en caso contrario podemos eliminar los factores comunes en ambos. Sustituyendo $z$ en la ecuación se tiene:

$$a_n \frac{M^n}{N^n} + a_{n-1} \frac{M^{n-1}}{N^{n-1}} + \cdots + a_1 \frac{M}{N} + a_0 = 0$$

o bien

$$a_n M^n + a_{n-1} M^{n-1} N + \cdots + a_1 M N^{n-1} + a_0 N^n = 0. \tag{2}$$

La igualdad (2) puede escribirse de la forma

$$(a_n M^{n-1} + a_{n-1} M^{n-2} N + \cdots + a_1 N^{n-1}) M = -a_0 N^n$$

Por tanto, $M$ divide a $a_0 N^n$; como $M$ y $N$ son primos entre sí, deducimos que *$M$ ha de ser un divisor de $a_0$*.

La igualdad (2) también puede escribirse de la forma

$$N(a_{n-1} M^{n-1} + \cdots + a_1 M N^{n-2} + a_0 N^{n-1}) = -a_n M^n.$$

De aquí se deduce que $N$ divide a $a_n M^n$; como $M$ y $N$ son primos entre sí, deducimos que *$N$ ha de ser un divisor de $a_n$*.

Resumiendo, *las soluciones racionales de la ecuación algebraica con coeficientes enteros (1) se encuentran entre aquellos números fraccionarios cuyo numerador es un divisor de $a_0$ y cuyo denominador es un divisor de $a_n$*.



EJEMPLO A. Para encontrar las soluciones de la ecuación

$$3x^3 - 14x^2 + 23x - 10 = 0$$

intentamos averiguar si posee soluciones racionales. Los divisores de $a_0 = -10$ son

$$1,\ -1,\ 2,\ -2,\ 5,\ -5,\ 10 \text{ y } -10$$

Los divisores de $a_3 = 3$ son

$$1,\ -1,\ 3,\ -3.$$

Las posibles soluciones racionales de la ecuación son

$$1,\ -1,\ 2,\ -2,\ 5,\ -5,\ 10,\ -10$$

$$\frac{1}{3},\ \frac{-1}{3},\ \frac{2}{3},\ -\frac{2}{3},\ \frac{5}{3},\ -\frac{5}{3},\ \frac{10}{3} \text{ y } -\frac{10}{3}$$

Uno de estos números es solución si al sustituirlo en la parte izquierda de la ecuación y realizar las operaciones indicadas se obtiene cero como resultado. Por ejemplo, 2 no es solución, ya que

$$3 \cdot 2^3 - 14 \cdot 2^2 + 23 \cdot 2 - 10 = 24 - 56 + 46 - 10 = 4 \neq 0.$$

Sin embargo, 2/3 es solución, ya que

$$3 \cdot \frac{2^3}{3^3} - 14 \frac{2^2}{3^2} + 23 \cdot \frac{2}{3} - 10 = \frac{8}{9} - \frac{56}{9} + \frac{46}{3} - 10 = 0.$$

Haciendo uso de la regla de Ruffini o simplemente dividiendo el polinomio $3x^3 - 14x^2 + 23x - 10$ entre $x - 2/3$ se tiene

$$3x^3 - 14x^2 + 23x - 10 = 3(x^2 - 4x + 5)\left(x - \frac{2}{3}\right).$$

Por tanto, las otras dos soluciones de la ecuación son también solución de la ecuación $x^2 - 4x + 5 = 0$; éstas pueden obtenerse mediante la fórmula cuadrática:

$$\frac{4 \pm \sqrt{16-20}}{2} = \frac{4 \pm \sqrt{-4}}{2} = 2 \pm i.$$

Así pues, $z_1 = \frac{2}{3}$, $z_2 = 2 + i$ y $z_3 = 2 - i$ son las tres soluciones de la ecuación dada.

EJEMPLO B. Queremos encontrar las soluciones de la ecuación

$$x^4-x^3-x^2-x-2=0.$$

Si posee soluciones racionales, deben ser números enteros puesto que el denominador ha de ser $\pm 1$ (observar que el coeficiente de $x^4$ es 1). Estas posibles soluciones son los divisores de $-2$, es decir:

$$1, \ -1, \ 2 \text{ y } -2$$

Después de algunos intentos se concluye que $x=-1$ y $x=2$ son soluciones de la ecuación dada; en efecto:

$$(-1)^4-(-1)^3-(-1)^2-(-1)-2=1+1-1+1-2=0.$$

y

$$2^4-2^3-2^2-2-2=16-8-4-2-2=0$$

Una vez dividido el polinomio $x^4-x^3-x^2-x-2$ entre $(x+1)$ y $(x-2)$ se tiene como divisor al polinomio $x^2+1$. Las soluciones de $x^2+1=0$ son $i$ y $-i$. Por tanto, las soluciones de la ecuación dada son

$$z_1=-1, \ z_2=2, \ z_3=i \ \text{ y } \ z_4=-i$$

## EJERCICIOS 4.4

**1.** Resolver las siguientes ecuaciones algebraicas:

*a)* $x^4-10x^3+35x^2-50x+24=0$. [*Sol.:* 1, 2, 3, 4.]

*b)* $x^4+x^3-x-1=0$. [*Sol.:* $-1$, $1$, $-\frac{1}{2}+\frac{\sqrt{3}}{2}i$, $-\frac{1}{2}-\frac{\sqrt{3}}{2}i$.)

*c)* $x^3-x^2-3x+6=0$. [*Sol.:* $-2$, $\frac{3\pm\sqrt{3}i}{2}$.]

*d)* $4x^3-4x^2-5x+3=0$. [*Sol.:* $\frac{3}{2}$, $\frac{1}{2}$, $-1$.]

*e)* $x^4+7x^2-144=0$. [*Sugerencia:* hacer $x^2=t$.] [*Sol.:* 3, $-3$, $4i$, $-4i$.]

**2.** Demostrar la igualdad $(x^4+x^3+x^2+x+1)(x-1)=x^5-1$. Utilizar este resultado para encontrar las soluciones de la ecuación $x^4+x^3+x^2+x+1=0$.

**3.** Demostrar que si $z$ es una solución compleja de una ecuación algebraica con coeficientes reales, $\bar{z}$ es también solución de la misma ecuación.

**4.** Encontrar las raíces de la ecuación $x^2+2x+1-2i=0$ completando cuadrados. ¿Contradice el resultado de este problema lo demostrado en el problema anterior?



## 4.5. EJERCICIOS DE ALGEBRA LINEAL CON NUMEROS COMPLEJOS

Todos los conceptos introducidos en los capítulos 1 y 2, así como los resultados allí obtenidos, pueden ser generalizados a los números complejos. Por recordar algunos citaremos el método de Gauss para la resolución de sistemas de ecuaciones lineales, el concepto de dependencia e independencia lineal de vectores, el determinante de una matriz y la regla de Cramer.

En esta sección proponemos ejercicios relativos a estos conceptos y resultados, en los cuales se utilizan números complejos. Estos ejercicios sirven, además, de repaso de los capítulos 1 y 2.

En los ejercicios que siguen $\mathbb{C}$ denota el conjunto de los números complejos. De acuerdo con esta notación, $\mathbb{C}^n$ es el conjunto de todos los elementos de la forma $(z_1, z_2, \ldots, z_n)$, donde cada $z_i \in \mathbb{C}$.

**1.** Utilizar el método de eliminación de Gauss para resolver, si es posible, los siguientes sistemas de ecuaciones lineales con coeficientes complejos:

*a)* $\left.\begin{aligned} x_1-x_2+ix_3&=1 \\ ix_2-x_3&=i \end{aligned}\right\}$ *b)* $\left.\begin{aligned} x_1-ix_2&=2 \\ x_2+ix_3&=1+2i \\ x_1+x_2&=1+i \end{aligned}\right\}$ *c)* $\left.\begin{aligned} x_1-x_2&=1 \\ ix_2+x_2&=0 \\ x_1+(1-i)x_2&=i \end{aligned}\right\}$

**2.** Determinar si los siguientes conjuntos de vectores son linealmente dependientes o independientes y en caso de que sean linealmente dependientes encontrar una combinación lineal entre ellos:

*a)* $\{(1, i), (-i, 1)\}$ *b)* $\{(2, i, 1+i), (i, -1, -2i)\}$

*c)* $\{(1, 1+i, 0, 1-i), (0, 1, i, 1+i), (1, 1, 1, 2-2i)\}$

**3.** Escribir las matrices de las siguientes aplicaciones lineales:

*a)* $f\colon \mathbb{C}^3 \mapsto \mathbb{C}^2$ tal que $f(z_1, z_2, z_3)=(iz_1+z_2, z_2+iz_3)$

*b)* $g\colon \mathbb{C}^3 \mapsto \mathbb{C}^3$ tal que $g(z_1, z_2, z_3)=(iz_3, iz_2, iz_1)$

*c)* $h\colon \mathbb{C}^2 \mapsto \mathbb{C}^2$ tal que $h(z_1, z_2)=(z_1-iz_2, z_1+iz_2)$

**4.** Si $f$, $g$ y $h$ son las aplicaciones del ejercicio 3, calcular:

*a)* $h\circ f$ *b)* $g\circ h$ *c)* $g^{-1}$ *d)* $g^{-1}\circ h$

**5.** Calcular los siguientes determinantes:

*a)* $\begin{vmatrix} i & 1 & 0 \\ 1 & -i & 1 \\ 0 & 1 & i \end{vmatrix}$ *b)* $\begin{vmatrix} e^{2i} & e^{3i} \\ 1 & e^{i} \end{vmatrix}$ *c)* $\begin{vmatrix} 1+i & i & 2i \\ 0 & 1 & 0 \\ 1 & 1-i & 1 \end{vmatrix}$

*d)* $\begin{vmatrix} i & 1 & 0 & 0 & 0 \\ 1 & i & 1 & 0 & 0 \\ 0 & 1 & i & 1 & 0 \\ 0 & 0 & 1 & i & 1 \\ 0 & 0 & 0 & 1 & i \end{vmatrix}$ *e)* $\begin{vmatrix} i & i & i & \cdots & i \\ i & i-z & i & \cdots & i \\ i & i & i-z & \cdots & i \\ \vdots & \vdots & \vdots & & \vdots \\ i & i & i & \cdots & i-z \end{vmatrix}_{n\times n}$, $z \in \mathbb{A}$

**6.** Hallar las inversas de las siguientes matrices, calculando primero la matriz de los cofactores:

*a*) $\begin{pmatrix} i & 1 \\ 1 & i \end{pmatrix}$ *b*) $\begin{pmatrix} 1+i & 3 & i \\ 1 & 2 & 1 \\ i & 1 & 0 \end{pmatrix}$ *c*) $\begin{pmatrix} 1 & i & -1 \\ 0 & 1 & i \\ 0 & 0 & 1 \end{pmatrix}$

**7.** Utilizar la regla de Cramer para resolver los siguientes sistemas de ecuaciones lineales con coeficientes complejos:

*a*) $\left.\begin{aligned} x_1+ix_2&=0 \\ -x_1+x_3&=-1 \\ x_2-ix_3&=i \end{aligned}\right\}$ *b*) $\left.\begin{aligned} x_1+x_2&=2 \\ x_1-x_2&=2i \end{aligned}\right\}$

**8.** Encontrar el rango de las siguientes matrices:

*a*) $\begin{pmatrix} 1 & i & 2i \\ 2 & -2i & 0 \\ 0 & 1+i & 1+i \end{pmatrix}$ *b*) $\begin{pmatrix} 2i & 1+i \\ 1 & 0 \\ -3 & 2-i \end{pmatrix}$ *c*) $\begin{pmatrix} i & -1 & -2i \\ -i & 1 & 2i \\ i & -1 & -2i \end{pmatrix}$

**9.** Encontrar el rango de la matriz

$$\begin{pmatrix} 1 & z & -1 \\ 2 & -1 & z \\ 1 & 10 & -6 \end{pmatrix}$$

para los distintos valores del número complejo $z$.

**10.** Estudiar la compatibilidad o incompatibilidad de los siguientes sistemas y resolverlos en el caso de que sean compatibles:

*a*) $\left.\begin{aligned} x_1-ix_2+x_3&=i \\ x_2-ix_3&=2 \\ ix_1+2x_2&=1 \end{aligned}\right\}$ *b*) $\left.\begin{aligned} x_1+x_2+x_3&=0 \\ 2x_2-3x_3&=5i \\ x_1+ix_2&=-1 \end{aligned}\right\}$ *c*) $\left.\begin{aligned} x_1-ix_2&=0 \\ x_1+x_3&=1 \\ x_2-ix_3&=2 \end{aligned}\right\}$

## CAPITULO 5

# ESPACIOS VECTORIALES



## 5.1. DEFINICION DE ESPACIO VECTORIAL. EJEMPLOS

El conjunto de los números reales y el conjunto de los números complejos, con los cuales se ha trabajado en los capítulos anteriores, tienen propiedades similares. En ambos pueden definirse dos operaciones $+$ y $\cdot$ que satisfacen ciertas propiedades; estas propiedades han sido enumeradas para los números complejos en la primera sección del capítulo 4. Un conjunto en el que se han definido dos operaciones $+$ y $\cdot$ que satisfacen las propiedades (S1) a (S4) y (M1) a (M4), enunciadas en el capítulo 4 para los números complejos, recibe el nombre de *cuerpo*. Observemos que todos los resultados obtenidos en los capítulos 1 y 2 son ciertos en cualquier cuerpo, con tal que el cuerpo posea infinitos elementos.

Todos los resultados de los tres capítulos siguientes son ciertos en cuerpos similares a $\mathbb{R}$, que denota el cuerpo de los números reales, y a $\mathbb{C}$, que denota el cuerpo de los números complejos. Sin embargo, nosotros trabajaremos fundamentalmente con $\mathbb{R}$ y $\mathbb{C}$, los cuales se denotarán por $\mathbb{K}$.

Al estudiar el conjunto de todos los *vectores en el espacio* hemos definido la suma de vectores y la multiplicación por un número real y enunciado las propiedades que satisfacían (véase sección 2 del capítulo 1).

Otros conjuntos a los cuales se les puede dotar de una estructura similar a la del conjunto de todos los vectores en el espacio son los conjuntos formados por *matrices*.

Si denotamos por $\mathcal{M}_{m\times n}(\mathbb{K})\equiv\mathcal{M}_{m\times n}$ al conjunto de todas las matrices de $m$ filas y $n$ columnas con elementos en el cuerpo $\mathbb{K}$, se ha demostrado en la sección 3 del capítulo 1, que satisface las mismas propiedades que los vectores, al menos cuando $\mathbb{K}$ es el cuerpo de los números reales. Iguales resultados se obtienen si $\mathbb{K}$ coincide con $\mathbb{C}$.

En la misma sección en que se estudiaron las propiedades de las matrices se definieron operaciones con *aplicaciones lineales de* $\mathbb{R}^m$ *en* $\mathbb{R}^n$ y se estudiaron sus propiedades. Estas son análogas a las propiedades de las matrices. Por tanto, el conjunto $L(\mathbb{R}^m, \mathbb{R}^n)$ de todas las aplicaciones lineales de $\mathbb{R}^m$ en $\mathbb{R}^n$ tiene la misma estructura que los vectores en el espacio y las matrices.

Cuando en varios conjuntos distintos aparecen estructuras similares es conveniente axiomatizar éstas y dar un nombre al ente resultante. La principal ventaja que se obtiene es que estudiando esta estructura, quedan estudiadas todas las estructuras particulares que en ella se encuadran. Cuando en un conjunto se da una estructura similar a la de los ejemplos anteriores, se dice que se tiene un *espacio vectorial.*

DEFINICIÓN *(Espacio vectorial)*

Un conjunto $V$, cuyos elementos se denotan mediante $\vec{u}, \vec{v}, \vec{w}, \ldots$, se dice que es un *espacio vectorial* sobre el cuerpo $\mathbb{K}$, si en él se han definido dos operaciones: la *suma*, $+$, de manera que a cada par de elementos $\vec{u}$ y $\vec{v}$ de $V$ se le hace corresponder el elemento $\vec{u}+\vec{v}$ de $V$, denominado suma de $\vec{u}$ y $\vec{v}$, y la *multiplicación por escalares*, de manera que a todo elemento $\vec{u}$ de $V$ y a todo elemento $a$ de $\mathbb{K}$ se le hace corresponder el elemento $a\vec{u}$ de $V$, que satisfacen las siguientes propiedades:

(S1) (Conmutativa) $\vec{u}+\vec{v}=\vec{v}+\vec{u}$ para todo $\vec{u}, \vec{v}$ de $V$.

(S2) (Asociativa) $\vec{u}+(\vec{v}+\vec{w})=(\vec{u}+\vec{v})+\vec{w}$ para todo $\vec{u},\vec{v}$ y $\vec{w}$ de $V$.

(S3) Existe un elemento de $V$, designado por $\vec{0}$ y denominado *neutro*, tal que $\vec{u}+\vec{0}=\vec{u}$ para todo $\vec{u}$ de $V$.

(S4) Para todo $\vec{u}$ de $V$ existe un elemento, designado por $-\vec{u}$ y denominado *opuesto* de $\vec{u}$, tal que $\vec{u}+(-\vec{u})=\vec{0}$.

(M1) $1\vec{u}=\vec{u}$ para todo $\vec{u}$ de $V$, donde 1 denota el elemento unidad de $\mathbb{K}$.

(M2) $a(b\vec{u})=(ab)\vec{u}$ para todo $\vec{u}$ de $V$ y todo $a, b$ de $\mathbb{K}$.

(M3) $(a+b)\vec{u}=a\vec{u}+b\vec{u}$ para todo $\vec{u}$ de $V$ y todo $a, b$ de $\mathbb{K}$.

(M4) $a(\vec{u}+\vec{v})=a\vec{u}+a\vec{v}$ para todo $\vec{u}, \vec{v}$ de $V$ y todo $a$ de $\mathbb{K}$.

*Notas.* 1) Los elementos de un espacio vectorial reciben el nombre genérico de vectores y en general se utiliza la notación indicada en la definición para denotarlos. Esto no es obstáculo para que en algunos casos particulares, por ejemplo, matrices o aplicaciones, se utilice la notación propia de cada caso.

2) Las propiedades de (S1) a (S4) se refieren a la suma, las propiedades (M1) y (M2) se refieren exclusivamente a la multiplicación por escalares y las propiedades (M3) y (M4) son las *distributivas* de una operación con respecto a la otra.

3) Si $\mathbb{K}$ es $\mathbb{R}$ se dice que $V$ es un *espacio vectorial real* y si $\mathbb{K}$ es $\mathbb{C}$ se dice que $V$ es un *espacio vectorial complejo.*



4) De (S3), (S4) y la propiedad conmutativa se deduce que $\vec{0}+\vec{u}=\vec{0}$ y $(-\vec{u})+\vec{u}=\vec{0}$

EJEMPLO A. 1) El conjunto de los vectores en el plano $V_2$ o el conjunto de los vectores en el espacio $V_3$ con las operaciones dadas en el capítulo 1 son espacios vectoriales reales.

2) El conjunto $\mathcal{M}_{m\times n}(\mathbb{K})\equiv\mathcal{M}_{m\times n}$ de todas las matrices de $m$ filas y $n$ columnas con elementos en $\mathbb{K}$ es un espacio vectorial sobre el cuerpo $\mathbb{K}$.

EJEMPLO B. El conjunto

$$\mathbb{K}^n=\{(x_1, x_2, \ldots, x_n)/x_j\in\mathbb{K}, j=1, 2, \ldots, n\}$$

con las operaciones

$$(x_1, x_2, \ldots, x_n)+(y_1, y_2, \ldots, y_n)=(x_1+y_1, x_2+y_2, \ldots, x_n+y_n)$$
$$a(x_1, x_2, \ldots, x_n)=(ax_1, ax_2, \ldots, ax_n)$$

es un espacio vectorial sobre $\mathbb{K}$, como fácilmente puede comprobarse.

En particular, $\mathbb{R}^n$ es un espacio vectorial real y $\mathbb{C}^n$ es un espacio vectorial complejo.

EJEMPLO C. Sea $\mathbb{R}^\infty$ el conjunto de todas las sucesiones de la forma $(a_n)_{n=1}^\infty=(a_1, a_2, \ldots)$, con $a_j\in\mathbb{R}$, con las operaciones

$$(a_n)_{n=1}^\infty+(b_n)_{n=1}^\infty=(a_n+b_n)_{n=1}^\infty \quad \text{y} \quad a(a_n)_{n=1}^\infty=(aa_n)_{n=1}^\infty.$$

Se deja para el lector la comprobación de que este conjunto es un espacio vectorial real con las operaciones que acabamos de definir.

EJEMPLO D. Sea $P_{\mathbb{K}}[x]$ el conjunto de todos los polinomios en la variable $x$ sobre el cuerpo $\mathbb{K}$, es decir, todos los elementos de la forma

$$p(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+a_nx^n=\sum_{j=0}^n a_jx^j, \quad a_j\in\mathbb{K}$$

En él definimos las siguientes operaciones: dados

$$p(x)=\sum_{j=0}^n a_jx^j \quad \text{y} \quad q(x)=\sum_{j=0}^m b_jx^j$$

con $m\geqq n$ su suma es

$$p(x)+q(x)=\sum_{j=0}^n (a_j+b_j)x^j+\sum_{j=n+1}^m b_jx^j.$$

Dado $a \in \mathbb{K}$, definimos

$$ap(x) = \sum_{j=0}^{n} (aa_j)x^j.$$

Con estas operaciones $P_{\mathbb{K}}[x]$ es un espacio vectorial sobre $\mathbb{K}$.

Si denotamos por $P_{\mathbb{K}}^{(n)}[x]$ al conjunto de todos los polinomios en la variable $x$, de grado menor o igual que $n$, tenemos de nuevo un espacio vectorial, con las mismas operaciones que en $P_{\mathbb{K}}[x]$.

EJEMPLO E. Sea $C([a, b])$ el conjunto de todas las funciones continuas definidas en el intervalo real $[a, b]$ con valores en $\mathbb{R}$; con las operaciones

$$(f+g)(x) = f(x) + g(x) \quad \text{y} \quad (af)(x) = a(f(x))$$

puede comprobarse que $C([a, b])$ es un espacio vectorial. El elemento neutro de este espacio vectorial es la aplicación nula.

EJEMPLO F. Sea $L(\mathbb{R}^m, \mathbb{R}^n)$ el conjunto de todas las aplicaciones lineales de $\mathbb{R}^m$ en $\mathbb{R}^n$; en la sección 3 del capítulo 1 se han definido operaciones en este conjunto y se ha visto que satisfacen las propiedades requeridas para ser un espacio vectorial real.

EJEMPLO G. Ha llegado el momento de mostrar algunas estructuras que no satisfacen todos los axiomas de un espacio vectorial. Si denotamos por $D$ al conjunto de todas las matrices cuadradas con determinante nulo y en él definimos las mismas operaciones que en $\mathcal{M}_{m \times n}$, $D$ *no* es un espacio vectorial; para ello basta observar que

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{y} \quad B = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

son matrices con determinante nulo, mientras que $A+B$ tiene determinante 1.

Tampoco el conjunto de todos los polinomios de grado *exactamente* $n$ es un espacio vectorial con las operaciones definidas en $P_{\mathbb{K}}[x]$.

EJEMPLO H. Un ejemplo muy importante es el conjunto $S(A)$ de soluciones del sistema homogéneo

$$A\vec{x} = \vec{0}, \quad A \in \mathcal{M}_{m \times n}(\mathbb{R}) \tag{I}$$

donde $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Ya sabemos que $\vec{0}$ es siempre solución de (I) y de los resultados de la sección 2 del capítulo 1 se deduce que si $\vec{x}$ e $\vec{y}$ son soluciones de (I) y $a \in \mathbb{R}$, $\vec{x}+\vec{y}$ y $a\vec{x}$ son soluciones de I.

* * *

Para terminar esta sección daremos algunos resultados que se deducen inmediatamente de las propiedades que definen un espacio vectorial.



PROPOSICIÓN 1

El elemento neutro de un espacio vectorial es único.

*Demostración.* Supongamos que $\vec{0}_1$ y $\vec{0}_2$ son dos elementos neutros de un espacio vectorial $V$; por ser $\vec{0}_1$ un elemento neutro se tiene que $\vec{0}_1+\vec{0}_2=\vec{0}_2$; por ser $\vec{0}_2$ un elemento neutro se tiene que $\vec{0}_2+\vec{0}_1=\vec{0}_1$; puesto que $\vec{0}_2+\vec{0}_1=\vec{0}_1+\vec{0}_2$ debido a la propiedad conmutativa, deducimos que $\vec{0}_2=\vec{0}_1$, mostrándose así la unicidad del elemento neutro. ■

PROPOSICIÓN 2

El opuesto de cada elemento en un espacio vectorial es único.

*Demostración.* Supongamos que el elemento $\vec{u}$ posee dos opuestos $\vec{v}_1$ y $\vec{v}_2$; de la definición de opuesto se deduce que $\vec{u}+\vec{v}_1=\vec{0}$. Sumando $\vec{v}_2$ a ambos lados de esta ecuación y utilizando las propiedades (S1), (S2) y (S3) de espacio vectorial se tiene que

$$\vec{v}_2+(\vec{u}+\vec{v}_1) = (\vec{v}_2+\vec{u})+\vec{v}_1 = \vec{0}+\vec{v}_1 = \vec{v}_1$$
$$\vec{v}_2+(\vec{u}+\vec{v}_1) = \vec{v}_2+\vec{0} = \vec{v}_2$$

De aquí deducimos que $\vec{v}_1=\vec{v}_2$. ■

PROPOSICIÓN 3

Para todo $\vec{u}$ de un espacio vectorial $V$, $0 \cdot \vec{u} = \vec{0}$.

*Demostración.* Tenemos la siguiente cadena de igualdades debida a varios de los axiomas de la definición de espacio vectorial

$$\vec{u} = 1\cdot\vec{u} = (0+1)\vec{u} = 0\cdot\vec{u} + 1\cdot\vec{u} = 0\cdot\vec{u}+\vec{u}$$

Esto implica que $0\cdot\vec{u}$ es el neutro del espacio vectorial $V$, y por tanto, $0\cdot\vec{u}=\vec{0}$. ■

PROPOSICIÓN 4

Para todo elemento $\vec{u}$ de un espacio vectorial $V$, $(-1)\vec{u}$ es su opuesto.

*Demostración.* Tenemos que

$$\vec{u}+(-1)\vec{u} = 1\cdot\vec{u}+(-1)\vec{u} = (1+(-1))\vec{u} = 0\cdot\vec{u} = \vec{0}$$

en donde se ha utilizado (M1), (M3) y la proposición 3. Puesto que el inverso es único, debido a la proposición 2, queda probado el resultado. ■

PROPOSICIÓN 5

En todo espacio vectorial $V$, $a\cdot\vec{0}=\vec{0}$, donde $a\in\mathbb{K}$ y $\vec{0}$ es el elemento neutro de $V$.

*Demostración.* $a\cdot\vec{0}=a(\vec{0}+\vec{0})=a\cdot\vec{0}+\vec{a}\cdot\vec{0}$; sumando el opuesto de $a\cdot\vec{0}$ en ambos lados se tiene $a\cdot\vec{0}=\vec{0}$, que era lo que queríamos demostrar. ■

•

## 5.2. BASE Y DIMENSION DE UN ESPACIO VECTORIAL

Sea $V$ un espacio vectorial sobre un cuerpo $\mathbb{K}$; un número finito de vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ se dice que son *linealmente dependientes* si existen $n$ elementos de $\mathbb{K}$, $a_1, a_2, ..., a_n$, no todos nulos, tal que

$$a_1\vec{v}_1+a_2\vec{v}_2+\cdots+a_n\vec{v}_n=\vec{0} \qquad (*)$$

Si los vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ *no* son linealmente dependientes, se dice que son *linealmente independientes;* por tanto, los vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ son linealmente independientes si cualquier igualdad como la que aparece en (*) implica que todos los elementos de $\mathbb{K}$, $a_1, a_2, ..., a_n$, son nulos.

Si en la igualdad (*) $a_n$ es no nulo, podemos escribir

$$\vec{v}_n=-\frac{a_1}{a_n}\vec{v}_1-\frac{a_2}{a_n}\vec{v}_2-\cdots-\frac{a_{n-1}}{a_n}\vec{v}_{n-1}$$

decimos entonces que $\vec{v}_n$ es una *combinación lineal* de los vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_{n-1}$. En general, diremos que $\vec{v}$ es *combinación lineal* de los vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ si existen $k$ números $a_1, a_2, ..., a_k$ de $\mathbb{K}$ tal que

$$\vec{v}=a_1\vec{v}_1+a_2\vec{v}_2+\cdots+a_k\vec{v}_k.$$

Un conjunto finito de vectores $\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_k\}$ de un espacio vectorial $V$ se dice que es un *sistema de generadores de* $V$ si todo elemento de $V$ puede escribirse como una combinación lineal de los vectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$.

Antes de exponer algunos ejemplos es conveniente realizar algunas observaciones.

En primer lugar, observamos que *todo conjunto finito de vectores que contiene al elemento neutro es linealmente dependiente;* basta observar que

$$a\cdot\vec{0}+0\cdot\vec{v}_2+\cdots+0\cdot\vec{v}_n=\vec{0}$$

para cualquier $a\in\mathbb{K}$.

En segundo lugar, observaremos que *todo conjunto finito de vectores linealmente independientes no puede contener un subconjunto de vectores que sean linealmente dependientes.* En efecto, si $\vec{v}_1, ..., \vec{v}_n$ son linealmente independientes y suponemos, por comodidad de notación, que $\vec{v}_1, ..., \vec{v}_k$, $k\leqq n$, son linealmente dependientes tendríamos



$$a_1\vec{v}_1+a_2\vec{v}_2+\cdots+a_k\vec{v}_k=\vec{0}$$

con no todos los $a_j$ igual a cero; basta observar que, entonces,

$$a_1\vec{v}_1+a_2\vec{v}_2+\cdots+a_k\vec{v}_k+0\vec{v}_{k+1}+\cdots+0\vec{v}_n=\vec{0},$$

con lo cual los vectores originales serían linealmente dependientes.

EJEMPLO A. Tres vectores no nulos en $V_2$ son siempre linealmente dependientes. Para demostrarlo tomar

$$\vec{u}=(u_1, u_2), \quad \vec{v}=(v_1, v_2), \quad \vec{w}=(w_1, w_2)$$

tres vectores cualesquiera de $V_2$. La igualdad

$$a_1\vec{u}+a_2\vec{v}+a_3\vec{w}=\vec{0}$$

se traduce en

$$\left.\begin{aligned} a_1u_1+a_2v_1+a_3w_1=0\\ a_1u_2+a_2v_2+a_3w_2=0\end{aligned}\right\}$$

que es un sistema de dos ecuaciones con tres incógnitas.

Puesto que el rango de la matriz de los coeficientes es inferior a 3, este sistema homogéneo posee infinitas soluciones. Basta elegir una no nula para tener el resultado deseado.

Este resultado puede demostrarse también geométricamente como puede apreciarse en la figura adjunta. La recta que contiene a $\vec{u}$ y la recta que es paralela a $\vec{v}$ y pasa por

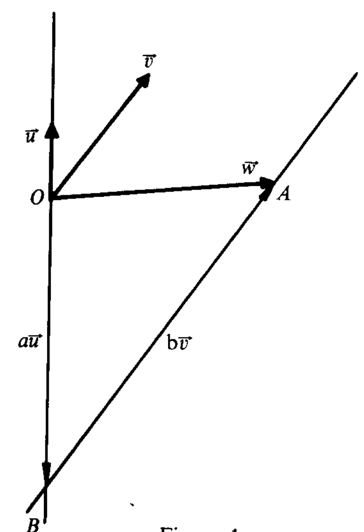



Figura 1

el extremo de $\vec{w}$ se cortan en un punto $B$. Observar que si estas rectas no se cortan, $\vec{u}$ y $\vec{v}$ son proporcionales y ya habríamos probado el resultado. En la figura se observa que $\overrightarrow{OB}+\overrightarrow{BA}=\vec{w}$; por tanto:

$$a\vec{u}+b\vec{v}=\vec{w}$$

lo cual prueba el resultado deseado.

Este mismo razonamiento muestra que dos vectores linealmente independientes cualesquiera de $V_2$ son siempre un sistema de generadores de este espacio vectorial.

EJEMPLO B. En general, $n+1$ vectores de $\mathbb{K}^n$ son linealmente dependientes; basta observar que esta afirmación se reduce a demostrar la existencia de solucuiones no nulas de un sistema homogéneo de $n$ ecuaciones con $n+1$ incógnitas (ver ejemplo A); este último resultado es cierto por el teorema de Rouché-Frobenius.

En $\mathbb{R}^3$ puede darse una demostración geométrica de este resultado. Supongamos que $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ son tres vectores de $\mathbb{R}^3$ linealmente independientes. Por el extremo de $\vec{u}_4$ se traza una paralela a $\vec{u}_3$ hasta que corte al plano que determinan $\vec{u}_1$ y $\vec{u}_2$. Tenemos que

$$u_4=\overrightarrow{OB}+\overrightarrow{BA}$$

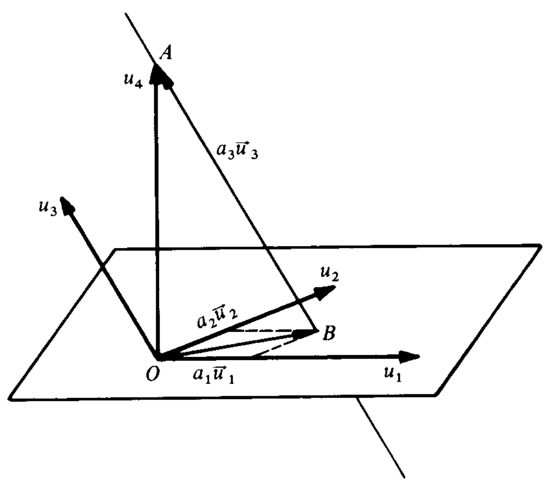



Figura 2

Puesto que $\overrightarrow{OB}$ es un vector del plano determinado por $\vec{u}_1$ y $\vec{u}_2$ se tiene que $\overrightarrow{OB}=a_1\vec{u}_1+a_2\vec{u}_2$. Utilizando este resultado junto con el hecho de que $\overrightarrow{BA}=a_3\vec{u}_3$ se tiene que

$$u_4=a_1\vec{u}_1+a_2\vec{u}_2+a_3\vec{u}_3.$$

Si los vectores $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$ fueran linealmente dependientes, se tendría una igualdad de la forma $a_1\vec{u}_1+a_2\vec{u}_2+a_3\vec{u}_3=\vec{0}$ y, por tanto, los cuatro vectores serían linealmente dependientes, ya que tendríamos la igualdad

$$a_1\vec{u}_1+a_2\vec{u}_2+a_3\vec{u}_3+0\cdot\vec{u}_4=\vec{0}.$$



Este razonamiento muestra, además, que tres vectores cualesquiera linealmente independientes de $\mathbb{R}^3$ son siempre un sistema de generadores de este espacio vectorial.

EJEMPLO C. 1) Las funciones $p_0(x)=1$, $p_1(x)=x$, $p_2(x)=x^2$, ..., $p_n(x)=x^n$ son linealmente independientes, ya que si tenemos la igualdad

$$a_0 1+a_1x+a_2x^2+\cdots+a_nx^n=0$$

para todo $x\in\mathbb{R}$ se ha de tener $a_0=a_1=a_2=\cdots=a_n=0$ (véase el problema 15 de la sección 4 del capítulo 2).

2) Las funciones $f(x)=\cos^2 x$, $g(x)=\operatorname{sen}^2 x$ y $h(x)=1$ son linealmente dependientes en $C([0, 2\pi])$, ya que

$$\cos^2 x+\operatorname{sen}^2 x=1$$

* * *

DEFINICIÓN 1

Un conjunto finito de vectores $\{\vec{e}_1,\vec{e}_2,...,\vec{e}_n\}$ se dice que es una *base* de un espacio vectorial $V$ si se cumplen las dos siguientes condiciones:

1) Los vectores $\vec{e}_1,\vec{e}_2,...,\vec{e}_n$ son linealmente independientes, y
2) Todo elemento de $V$ es una combinación lineal de los vectores $\vec{e}_1,\vec{e}_2,...,\vec{e}_n$.

Observar que la condición 2) de esta definición es equivalente al hecho de que el conjunto de vectores $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ sea un sistema de generadores de $V$. Sin embargo, no todo el sistema de generadores de un espacio vectorial $V$ es una base; convénzase el lector por su propia cuenta de que tres vectores en $V_2$, dos de los cuales son linealmente independientes, son un sistema de generadores de $V_2$; sin embargo, no pueden formar una base de $V_2$, ya que son linealmente dependientes, como se ha mostrado en el ejemplo A.

EJEMPLO D. Si $\vec{e}_j=(0, ..., 1, ..., 0)\in\mathbb{K}^n$, donde el 1 ocupa el lugar $j$, se tiene que $\{\vec{e}_1,\vec{e}_2,...,\vec{e}_n\}$ son linealmente independientes, ya que si

$$\sum_{j=1}^{n} a_j\vec{e}_j=(0, 0, ..., 0)$$

se tiene que $(a_1, a_2, ..., a_n)=(0, 0, ..., 0)$. Por tanto, $a_1=a_2=\cdots=a_n=0$. Además, si $\vec{x}=(x_1, x_2, ..., x_n)\in\mathbb{K}^n$ se tiene que

$$\vec{x}=\sum_{j=1}^{n} x_j\vec{e}_j$$

Por tanto, $\{\vec{e}_1,\vec{e}_2,...,\vec{e}_n\}$ es una base de $\mathbb{K}^n$, que recibe el nombre de *base canónica* de este espacio.

EJEMPLO E. 1) El conjunto $\{1, x, x^2, ..., x^n\}$ es una base de $P_{\mathbb{K}}^{(n)}[x]$, ya que son linealmente independientes de acuerdo con el resultado del ejemplo C, y todo polinomio $p$ de grado inferior o igual a $n$ puede escribirse de la forma

$$p(x)=a_0 1+a_1 x+a_2 x^2+\cdots+a_n x^n$$

2) El conjunto $\{1, x, x^2, ..., x^n\}$ no es una base de $P_{\mathbb{K}}[x]$, ya que el polinomio $p(x)=$ $=x^{n+1}$ no es combinación lineal de éstos.

EJEMPLO F. Las matrices

$$E_1=\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix},\quad E_2=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix},\quad E_3=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\quad \text{y}\quad E_4=\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}$$

son una base de $\mathscr{M}_{2\times 2}(\mathbb{K})$. Son linealmente independientes, ya que

$$a_1E_1+a_2E_2+a_3E_3+a_4E_4=\begin{pmatrix}a_1 & a_2\\ a_3 & a_4\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

implica $a_1=a_2=a_3=a_4=0$. Además, si

$$A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$$

podemos escribir

$$A=aE_1+bE_2+cE_3+dE_4$$

* * *

Si $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ es una base de un espacio vectorial $V$ y $\vec{v}$ es cualquier elemento de $V$ podemos escribir $\vec{v}$ como combinación lineal de $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ de la forma

$$\vec{v}=v_1\vec{e}_1+v_2\vec{e}_2+\cdots+v_n\vec{e}_n$$

con $v_j\in\mathbb{K}$. Los números $v_1, v_2, ..., v_n$ se denominan *componentes de* $\vec{v}$ con respecto a la base $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$. Las componentes de un vector $\vec{v}$ con respecto a una base son únicas, ya que si tenemos

$$\vec{v}=v_1\vec{e}_1+v_2\vec{e}_2+\cdots+v_n\vec{e}_n$$

y

$$\vec{v}=v_1'\vec{e}_1+v_2'\vec{e}_2+\cdots+v_n'\vec{e}_n$$

se tiene también que

$$\vec{0}=(v_1-v_1')\vec{e}_1+(v_2-v_2')\vec{e}_2+\cdots+(v_n-v_n')\vec{e}_n$$



Puesto que $\vec{e}_1, \vec{e}_2, ..., \vec{e}_n$ son linealmente independientes hemos de tener $v_1=v'_1$, $v_2=v'_2$, ..., $v_n=v'_n$, que era lo que queríamos demostrar.

* * *

Un mismo espacio vectorial puede poseer varias bases; nuestro próximo objetivo es demostrar que todas ellas han de poseer el mismo número de elementos.

PROPOSICIÓN 2

Si $V$ es un espacio vectorial que posee una base con $n$ elementos, cualesquiera $n+1$ vectores de $V$ son linealmente dependientes.

*Demostración.* Sea $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ una base de $V$ y sean $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n, \vec{x}_{n+1}$ cualesquiera $n+1$ vectores de $V$; podemos escribir

$$\vec{x}_1=\sum_{j=1}^{n} x_{j1}\vec{e}_j,\quad \vec{x}_2=\sum_{j=1}^{n} x_{j2}\vec{e}_j,\ ...,\ \vec{x}_n=\sum_{j=1}^{n} x_{jn}\vec{e}_j,\quad \vec{x}_{n+1}=\sum_{j=1}^{n} x_{j,n+1}\vec{e}_j$$

ya que todo elemento de $V$ es una combinación lineal de los vectores de la base.

Tratamos de demostrar que podemos encontrar $a_1, a_2, ..., a_n, a_{n+1}$, no todos nulos, tal que

$$a_1\vec{x}_1+a_2\vec{x}_2+\cdots+a_n\vec{x}_n+a_{n+1}\vec{x}_{n+1}=\vec{0} \tag{1}$$

Esto es equivalente a escribir

$$a_1\left(\sum_{j=1}^{n} x_{j1}\vec{e}_j\right)+a_2\left(\sum_{j=1}^{n} x_{j2}\vec{e}_j\right)+\cdots+a_n\left(\sum_{j=1}^{n} x_{jn}\vec{e}_j\right)+a_{n+1}\left(\sum_{j=1}^{n} x_{j,n+1}\vec{e}_j\right)=\vec{0}.$$

Igualando las componentes se tiene

$$a_1x_{j1}+a_2x_{j2}+\cdots+a_nx_{jn}+a_{n+1}x_{j,n+1}=0,\quad j=1, 2, ..., n$$

que es un sistema de $n$ ecuaciones con $n+1$ incógnitas; puesto que el sistema es homogéneo, siempre ha de poseer una solución no trivial; esto demuestra el resultado. ■

De esta proposición se deduce un resultado un poco más general: *en un espacio vectorial V que posee una base con n elementos, cualesquiera m vectores de V, con m > n, son linealmente dependientes.* Basta observar que $n+1$ de los $m$ vectores dados han de ser linealmente dependientes, debido a la proposición 2, y por tanto, todos ellos han de formar un conjunto de vectores linealmente dependiente.

Este resultado se aplica en la demostración del siguiente teorema:

TEOREMA 3

Todas las bases de un mismo espacio vectorial $V$ poseen el mismo número de elementos.

*Demostración.* Sean $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ y $\{\vec{e}'_1, \vec{e}'_2, \ldots, \vec{e}'_m\}$ dos bases de $V$; por el razonamiento anterior $m \leqq n$, ya que en caso contrario los vectores de la segunda base serían linealmente dependientes. Similarmente, $n \leqq m$, ya que en caso contrario los vectores de la primera base serían linealmente dependientes. Se tiene, por tanto, que $m = n$. ■

El número de elementos que posee una base cualquiera de un espacio vectorial $V$ recibe el nombre de *dimensión de* $V$; este número será designado mediante $\dim(V)$.

Si el espacio vectorial sólo contiene un elemento, que necesariamente ha de ser su elemento neutro, es decir, $V=\{\vec{0}\}$, diremos que $V$ tiene *dimensión cero.* Las razones para dar esta definición se verán claramente más adelante.

De los ejemplos que hemos realizado anteriormente podemos deducir los siguientes resultados:

1) la dimensión de $\mathbb{K}^n$ es $n$,
2) la dimensión de $P_{\mathbb{K}}^{(n)}[x]$ es $n+1$,
3) la dimensión de $\mathcal{M}_{2\times 2}(\mathbb{K})$ es 4.

En general, tenemos que la dimensión de $\mathcal{M}_{m\times n}(\mathbb{K})$ es $m \times n$.

PROPOSICIÓN 4

Sea $V$ un espacio vectorial de dimensión $n$. Todo conjunto de $n$ vectores de $V$ que sean linealmente independientes son una base de $V$.

*Demostración.* Sean $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ $n$ vectores linealmente independientes; si $\vec{v}$ es cualquier otro vector de $V$, los vectores $\vec{v}, \vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ son linealmente dependientes por la proposición 2; por tanto, existen $a_0, a_1, \ldots, a_n \in \mathbb{K}$ tal que

$$a_0\vec{v} + a_1\vec{u}_1 + a_2\vec{u}_2 + \cdots + a_n\vec{u}_n = \vec{0}$$

con algún $a_j \neq 0$. El número $a_0$ debe ser no nulo, ya que en caso contrario los vectores $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n$ serían linealmente dependientes. Por tanto:

$$\vec{v} = -\frac{a_1}{a_0}\vec{u}_1 - \frac{a_2}{a_0}\vec{u}_2 - \cdots - \frac{a_n}{a_0}\vec{u}_n$$

y queda probado que $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ es un sistema de generadores de $V$. ■



EJEMPLO G. Los vectores $\vec{u}_1=(1, 0, 1, 0)$, $\vec{u}_2=(0, 1, 0, 0)$, $\vec{u}_3=(1, -1, 0, 0)$ y $\vec{u}_4=(0, 0, 1, -1)$ son linealmente independientes en $\mathbb{R}^4$, ya que

$$r\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix} = 4$$

Por la proposición 4 son una base de $\mathbb{R}^4$, ya que $\dim(\mathbb{R}^4)=4$.

* * *

Una forma de encontrar una base de un espacio vectorial $V$ es «añadir» vectores a un conjunto de vectores linealmente independientes de $V$; la forma de «añadirlos» se explica en la demostración de la siguiente proposición:

PROPOSICIÓN 5

Sea $V$ un espacio vectorial de dimensión finita $n$; todo conjunto de vectores linealmente independientes de $V$ puede completarse para obtener una base, es decir, dados $k$ vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k$, $k<n$, de $V$, linealmente independientes, existen $n-k$ vectores $\vec{e}_{k+1}, \ldots, \vec{e}_n$ de $V$ tal que el conjunto $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k, \vec{e}_{k+1}, \ldots, \vec{e}_n\}$ es una base de $V$.

*Demostración.* Puesto que $k$ es inferior a $n$ podemos encontrar un elemento de $V$ linealmente independiente con $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k$; en caso contrario, los vectores $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_k$ formarían una base de $V$. Sea $\vec{e}_{k+1}$ este vector. El razonamiento se repite con el conjunto $\{\vec{e}_1, \ldots, \vec{e}_k, \vec{e}_{k+1}\}$ hasta encontrar $n$ vectores linealmente independientes, que han de ser necesariamente una base, por la proposición 4. ■

* * *

Supongamos que $V$ es un espacio vectorial de dimensión finita $n$; acabamos de probar que $k$ vectores linealmente independientes de $V$ pueden «completarse» para obtener una base. Demostraremos a continuación que si $S$ es un sistema de generadores de $V$, de él puede extraerse una base de $V$. Antes de demostrar este resultado necesitamos el siguiente lema:

LEMA 6

Supongamos que $\dim(V)=n$; sea $S$ un sistema de generadores de $V$ y supongamos que $S=S_1 \cup S_2$, donde $S_1$ y $S_2$ son disjuntos. Si los elementos de $S_2$ son combinación lineal de los elementos de $S_1$, $S_1$ es también un sistema de generadores de $V$.

*Demostración.* Supongamos, para evitar complicaciones de notación, que $S_1 = \{\vec{e}_1, ..., \vec{e}_l\}$ y $S_2 = \{\vec{e}_{l+1}, ..., \vec{e}_m\}$. Puesto que los elementos de $S_2$ son combinación lineal de los elementos de $S_1$ tenemos que

$$\vec{e}_k = \sum_{i=1}^{l} a_{ik}\vec{e}_i, \quad k = l+1, ..., m.$$

Si $\vec{x} \in V$, puesto que $S$ es un sistema de generadores de $V$, se tiene que

$$\vec{x} = \sum_{i=1}^{m} x_i\vec{e}_i.$$

Por tanto,

$$\vec{x} = \sum_{i=1}^{l} x_i\vec{e}_i + \sum_{k=l+1}^{m} x_k\vec{e}_k = \sum_{i=1}^{l} x_i\vec{e}_i + \sum_{k=l+1}^{m} x_k\left(\sum_{i=1}^{l} a_{ik}\vec{e}_i\right) =$$
$$= \sum_{i=1}^{l} x_i\vec{e}_i + \sum_{i=1}^{l}\left(\sum_{k=l+1}^{m} x_k a_{ik}\right)\vec{e}_i = \sum_{i=1}^{l}\left(x_i + \sum_{k=l+1}^{m} x_k a_{ik}\right)\vec{e}_i$$

Esto prueba que $S_1$ es un sistema de generadores de $V$. ■

PROPOSICIÓN 7

Sea $S$ un sistema de generadores de un espacio vectorial $V$ de dimensión $n$; podemos encontrar un subconjunto $S_1$ de $S$ que sea una base de $V$.

*Demostración.* Elegir $\vec{u}_1 \in S$ tal que $\vec{u}_1 \neq \vec{0}$; esto siempre puede hacerse salvo en el caso en que $V = \{\vec{0}\}$ (si $V = \{\vec{0}\}$ la proposición es una trivialidad). Elegir a continuación $\vec{u}_2 \in S$ tal que $\vec{u}_2$ es independiente con $\vec{u}_1$. Suponiendo que hemos elegido $\vec{u}_1, \vec{u}_2, ..., \vec{u}_k$, elegir $\vec{u}_{k+1} \in S$ que sea linealmente independiente con los anteriores. Este proceso se continúa hasta que es imposible encontrar vectores de $S$ linealmente independientes con los anteriores. Sea

$$S_1 = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_l\}$$

el subconjunto de $S$ obtenido de esta manera. Por el lema anterior, $S_1$ es un sistema de generadores de $V$. Debido a la elección realizada, los vectores de $S_1$ son linealmente independientes; por tanto, $S_1$ es una base de $V$. Se termina así la demostración de la proposición 7. ■

* * *

Terminamos esta sección haciendo unos comentarios acerca de la dependencia o independencia lineal de subconjuntos infinitos de un espacio vectorial.

Un *conjunto infinito* $S$ de elementos de un espacio vectorial $V$ se dice *linealmente independiente* si cualquier subconjunto finito de $S$ es linealmente independiente. En caso contrario, $S$ se dice *linealmente dependiente;* es decir, $S$ es linealmente dependiente si existe un subconjunto finito de él que es linealmente dependiente.



Un espacio vectorial $V$ en el que se puede encontrar un subconjunto $S$ linealmente independiente y con infinitos elementos, se dice que tiene *dimensión infinita.*

Los espacios vectoriales $P_{\mathbb{K}}[x]$ y $C([0, 2\pi])$, introducidos en la sección anterior, son espacios vectoriales de dimensión infinita. El conjunto $S = \{x^n/n \in \mathbb{N}\}$ es un conjunto linealmente independiente en $P_{\mathbb{K}}[x]$. El conjunto $S = \{\cos nx, \operatorname{sen} mx\}_{n,m\in\mathbb{N}}$ es un conjunto linealmente independiente en $C([0, 2\pi])$. De las dos últimas afirmaciones anteriores la primera es sencilla de demostrar; la segunda es más complicada y se pide su demostración por parte del lector en el ejercicio 13 de la siguiente sección; en este ejercicio se dan las sugerencias adecuadas para facilitar su resolución.

## 5.3. CAMBIO DE BASE

Supongamos que un espacio vectorial $V$ de dimensión finita posee dos bases

$$B = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\} \quad \text{y} \quad B' = \{\vec{e}_1', \vec{e}_2', ..., \vec{e}_n'\}.$$

Denominaremos *base antigua* a la primera de ellas y *base nueva* a la segunda. Como todo elemento de la base nueva se escribe como combinación lineal de los elementos de la base antigua podemos escribir:

$$\begin{aligned} \vec{e}_1' &= a_{11}\vec{e}_1 + a_{21}\vec{e}_2 + \cdots + a_{n1}\vec{e}_n \\ \vec{e}_2' &= a_{12}\vec{e}_1 + a_{22}\vec{e}_2 + \cdots + a_{n2}\vec{e}_n \\ &\ \vdots \\ \vec{e}_n' &= a_{1n}\vec{e}_1 + a_{2n}\vec{e}_2 + \cdots + a_{nn}\vec{e}_n. \end{aligned} \tag{1}$$

Puesto que es conveniente que el lector se familiarice con notaciones abreviadas escribiremos (1) de la forma

$$e_j' = \sum_{i=1}^{n} a_{ij}\vec{e}_i, \quad j = 1, 2, ..., n.$$

Diremos entonces que la nueva base se obtiene de la antigua mediante la matriz

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

donde la $j$-ésima columna de $A$ son las componentes del vector $\vec{e}_j'$ con respecto a la base antigua, $j = 1, 2, ..., n$. La matriz $A$ se denomina *matriz del cambio de base.*

Cuando, por necesidades de notación sea necesario hacer constar las bases $B$ y $B'$ escribiremos

$$A = M_B^{B'}$$

Si $B$ y $B'$ coinciden se tiene que $M_B^B = I_{n\times n}$. La matriz $I_{n\times n}$ es invertible; esto sucede para cualquier matriz de un cambio de base, como se demuestra en la siguiente proposición.

PROPOSICIÓN 1

La matriz $A$ del cambio de base de $B$ a $B'$ es invertible y su inversa es la matriz del cambio de base de $B'$ a $B$. Podemos, por tanto, escribir:

$$M_{B'}^B = (M_B^{B'})^{-1} = A^{-1}$$

*Demostración.* Para demostrar que $A$ es invertible es suficiente demostrar que su determinante es no nulo (véanse los resultados del capítulo 2). Ciertamente, el determinante de $A$ es no nulo, ya que de lo contrario sus columnas y, por consiguiente, los vectores $\vec{e}_1'$, $\vec{e}_2'$, ..., $\vec{e}_n'$, de la base $B'$, serían linealmente dependientes.

Sea $A' = (a'_{ij})_{i,j=1,\dots,n}$ la matriz del cambio de base de $B'$ a $B$; hemos de demostrar que $A' = A^{-1}$. Por definición de $A'$ tenemos

$$\vec{e}_j = \sum_{i=1}^n a'_{ij}\vec{e}_i', \quad j = 1, 2, \dots, n;$$

por definición de $A$ tenemos

$$\vec{e}_i' = \sum_{k=1}^n a_{ki}\vec{e}_k, \quad i = 1, 2, \dots, n.$$

Por tanto,

$$\vec{e}_j = \sum_{i=1}^n a'_{ij}\left(\sum_{k=1}^n a_{ki}\vec{e}_k\right) = \sum_{k=1}^n \left(\sum_{i=1}^n a_{ki}a'_{ij}\right)\vec{e}_k.$$

Puesto que $\sum_{i=1}^n a_{ki}a'_{ij}$ es el elemento que ocupa el lugar $(k, j)$ en el producto de las matrices $A$ y $A'$ se tiene que

$$I_{n\times n} = M_B^B = AA'.$$

Esto demuestra el resultado deseado. ■

* * *



Tratemos ahora de relacionar entre sí las coordenadas de un mismo vector en las bases nueva $B'$ y antigua $B$. Supongamos que $\vec{x} \in V$,

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n \tag{2}$$

y además:

$$\vec{x} = x_1'\vec{e}_1' + x_2'\vec{e}_2' + \cdots + x_n'\vec{e}_n' \tag{3}$$

Sustituyendo en (3) las expresiones dadas en (1) obtenemos

$$\begin{aligned}\vec{x} &= x_1'\left(\sum_{i=1}^n a_{i1}\vec{e}_i\right) + x_2'\left(\sum_{i=1}^n a_{i2}\vec{e}_i\right) + \cdots + x_n'\left(\sum_{i=1}^n a_{in}\vec{e}_i\right) = \\ &= (a_{11}x_1' + a_{12}x_2' + \cdots + a_{1n}x_n')\vec{e}_1 + (a_{21}x_1' + a_{22}x_2' + \cdots + a_{2n}x_n')\vec{e}_2 + \\ &\quad + \cdots + (a_{n1}x_1' + a_{n2}x_2' + \cdots + a_{nn}x_n')\vec{e}_n\end{aligned}$$

Comparando esta última igualdad con (2) podemos escribir

$$\begin{aligned} x_1 &= a_{11}x_1' + a_{12}x_2' + \cdots + a_{1n}x_n' \\ x_2 &= a_{21}x_1' + a_{22}x_2' + \cdots + a_{2n}x_n' \\ &\vdots \\ x_n &= a_{n1}x_1' + a_{n2}x_2' + \cdots + a_{nn}x_n'. \end{aligned} \tag{4}$$

Si convenimos en escribir $X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ y $X' = \begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}$ a las componentes del vector $\vec{x}$ en la antigua y en la nueva base, respectivamente, (4) se escribe brevemente de la forma

$$X = AX'$$

Esto nos permite obtener las coordenadas del vector $\vec{x}$ en la base antigua conociendo las coordenadas del mismo vector en la nueva base.

EJEMPLO A. Sea $\vec{e}_1 = (1, 0, 0)$, $\vec{e}_2 = (0, 1, 0)$, $\vec{e}_3 = (0, 0, 1)$ la base canónica de $\mathbb{R}^3$. Dadas las bases

$$\vec{u}_1 = \vec{e}_1 + \vec{e}_3, \quad u_2 = \vec{e}_2, \quad u_3 = \vec{e}_2 + \vec{e}_3$$

y

$$\vec{u}_1' = \vec{e}_1 - \vec{e}_2, \quad \vec{u}_2' = \vec{e}_2, \quad \vec{u}_3' = 2\vec{e}_1 + \vec{e}_3$$

encontrar las coordenadas del vector $\vec{x} = 3\vec{u}_1 + 2\vec{u}_2$ en la base $\{\vec{u}_1', \vec{u}_2', \vec{u}_3'\}$. Comenzar observando que ambos conjuntos son bases de $\mathbb{R}^3$.

Podemos escribir $\vec{x} = 3(\vec{e}_1 + \vec{e}_3) + 2\vec{e}_2 = 3\vec{e}_1 + 2\vec{e}_2 + 3\vec{e}_3$. El problema queda reducido a encontrar la matriz del cambio de base de $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ a $\{\vec{u}'_1, \vec{u}'_2, \vec{u}'_3\}$. De las expresiones dadas deducimos que la matriz del cambio de base es

$$A=\begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Escribiendo $\vec{x}=x'_1\vec{u}'_1+x'_2\vec{u}'_2+x'_3\vec{u}'_3$, por los resultados anteriores tenemos que

$$\begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}\begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}=\begin{pmatrix} 1 & 0 & -2 \\ 1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}=\begin{pmatrix} -3 \\ -1 \\ 3 \end{pmatrix}$$

Por tanto, $\vec{x}=-3\vec{u}'_1-\vec{u}'_2+3\vec{u}'_3$

EJEMPLO E. Sean $\vec{e}_1, \vec{e}_2$ dos vectores perpendiculares unitarios en $\mathbb{R}^2$ en la dirección de los ejes de coordenadas cartesianas. Girando los ejes de coordenadas un ángulo $\phi$ en sentido positivo (contrario a las agujas del reloj) se obtiene una nueva base $\vec{e}'_1, \vec{e}'_2$. ¿Cuál es la matriz del cambio de base?.

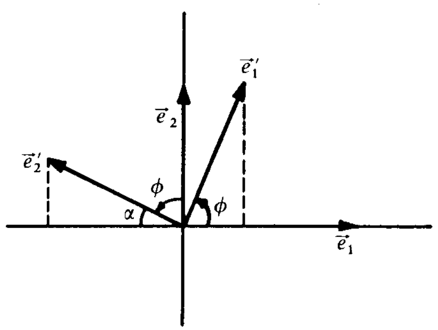
En el dibujo se observa que

$$\vec{e}'_1=(\cos\phi)\vec{e}_1+(\operatorname{sen}\phi)\vec{e}_2$$

$$\vec{e}'_2=\cos\left(\frac{\pi}{2}+\phi\right)\vec{e}_1+\operatorname{sen}\left(\frac{\pi}{2}+\phi\right)\vec{e}_2=-(\operatorname{sen}\phi)\,\vec{e}_1+(\cos\phi)\vec{e}_2$$

con lo que la matriz del cambio de base es

$$A=\begin{pmatrix} \cos\phi & -\operatorname{sen}\phi \\ \operatorname{sen}\phi & \cos\phi \end{pmatrix}$$






Las coordenadas antiguas a través de las nuevas vienen dadas por

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} \cos\phi & -\operatorname{sen}\phi \\ \operatorname{sen}\phi & \cos\phi \end{pmatrix}\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \Leftrightarrow \begin{matrix} x_1=x'_1\cos\phi-x'_2\operatorname{sen}\phi \\ x_2=x'_1\operatorname{sen}\phi+x'_2\cos\phi \end{matrix}$$

* * *

## EJERCICIOS (SECCIONES 5.1, 5.2 y 5.3)

**1.** *a*) Decir si el conjunto $F=\{f\in C([0, 1])/\ 2f(0)=f(1)\}$ es un espacio vectorial con las mismas operaciones que las definidas en $C([0, 1])$.

*b*) Decir si el conjunto

$$B=\left\{\begin{pmatrix} a & b \\ c & d \end{pmatrix}\in \mathcal{M}_{2\times 2}(\mathbb{R})/\ a-b+c+d=0\right\}$$

es un espacio vectorial con las mismas operaciones que las definidas en $\mathcal{M}_{2\times 2}(\mathbb{R})$.

**2.** Estudiar si las siguientes familias de vectores son linealmente dependientes o independientes:

*a*) $\{(1, 2), (0, 1)\}\subset\mathbb{R}^2$.

*b*) $\{(1, i), (i, -1)\}\subset\mathbb{C}^2$ (como espacio vectorial sobre $\mathbb{C}$).

*c*) $\{e^{i\pi t}, e^{i\pi(t-1)}, e^{i\pi(t+1)}\}\subset C_{\mathbb{C}}([0, 1])$, donde $C_{\mathbb{C}}([0, 1])$ denota las funciones continuas definidas en $[0, 1]$ con valores en $\mathbb{C}$.

*d*) $\left\{\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}\right\}\subset \mathcal{M}_{2\times 2}(\mathbb{R})$.

*e*) $\{\operatorname{sen}\pi t, \operatorname{sen}2\pi t\}\subset C_{\mathbb{R}}([0, 1])$.

**3.** Encontrar los valores de $z$ para los cuales los vectores $(z, 1, 0)$, $(1, z, 1)$ y $(0, 1, z)$ son linealmente dependientes en $\mathbb{C}^3$.

**4.** Demostrar que en $\mathbb{C}^2$ puede definirse una estructura de espacio vectorial real. ¿Cuál es la dimensión de este espacio?

**5.** Estudiar si los siguientes conjuntos son bases del espacio vectorial dado:

*a*) $\{1, x+3, (x+3)^2, (x+3)^3\}$ en $P_{\mathbb{R}}^{(3)}[x]$.

*b*) $\left\{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\right\}$ en $\mathcal{M}_{2\times 2}(\mathbb{R})$.

*c*) $\left\{\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\right\}$ en $\mathcal{M}_{2\times 2}(\mathbb{R})$.

**6.** Encontrar una base de $\mathbb{R}^4$ que contenga a los vectores $(0, 0, 1, 1)$ y $(1, 1, 0, 0)$.

**7.** Estudiar si las aplicaciones lineales $f_1, f_2, f_3$ y $f_4$ dadas por $f_1(x_1, x_2)=(x_1, -x_2)$, $f_2(x_1, x_2)=(x_1, x_2)$, $f_3(x_1, x_2)=(2x_1, x_2)$ y $f_4(x_1, x_2)=(x_1+2x_2, x_2)$ forman una base de $L(\mathbb{R}^2, \mathbb{R}^2)$.

**8.** Sean $p_1(x), p_2(x), ..., p_k(x)$ polinomios y supongamos que

$$\begin{vmatrix} p_1(a_1) & p_2(a_1) & \cdots & p_k(a_1) \\ p_1(a_2) & p_2(a_2) & \cdots & p_k(a_2) \\ \vdots & \vdots & & \vdots \\ p_1(a_k) & p_2(a_k) & \cdots & p_k(a_k) \end{vmatrix} \neq 0$$

para algunos números reales $a_j$. Demostrar que el conjunto $\{p_1(x), p_2(x), ..., p_k(x)\}$ es linealmente independiente en $P_{\mathbb{R}}[x]$.

**9.** Dada una matriz $A \in \mathcal{M}_{n \times n}(\mathbb{K})$ definimos su *traza* como la suma de los elementos de su diagonal principal; es decir, si $A=(a_{ij})_{i,j=1,...,n}$, *traza* $(A)=\sum_{i=1}^{n} a_{ii}$. Demostrar que

$$T=\{A \in \mathcal{M}_{n \times n}(\mathbb{K}) / \ \text{traza}\,(A)=0\}$$

es un espacio vectorial con las operaciones definidas en $\mathcal{M}_{n \times n}(\mathbb{K})$. ¿Cuál es la dimensión del espacio $T$?

**10.** *a*) Demostrar que los siguientes conjuntos son bases de $\mathbb{R}^4$:

$$B_1=\{\vec{v}_1=(1, 1, 0, 0), \vec{v}_2=(0, 1, 0, 0), \vec{v}_3=(0, 0, 1, 1), \vec{v}_4=(0, 0, 0, 1)\}$$
$$B_2=\{\vec{u}_1=(1, 2, 0, 0), \vec{u}_2=(0, 1, 2, -1), \vec{u}_3=(1, -1, -1, -1), \vec{u}_4=(0, 1, 1, 0)\}$$

*b*) Encontrar las componentes del vector $\vec{v}=3\vec{v}_1-\vec{v}_3+2\vec{v}_2$ con respecto a la base $B_2$.

**11.** Sea $\vec{e}_1, \vec{e}_2, ..., \vec{e}_k$ la base canónica de $\mathbb{R}^n$ y sea $\vec{u}_1=\vec{e}_2-\vec{e}_1, \vec{u}_2=\vec{e}_3-\vec{e}_2, ..., \vec{u}_{n-1}=\vec{e}_n-\vec{e}_{n-1}, \vec{u}_n=\vec{e}_n$.

*a*) Demostrar que $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ es una base de $\mathbb{R}^n$.
*b*) Expresar el vector $\vec{x}=\vec{e}_1+\vec{e}_2+\cdots+\vec{e}_n$ como una combinación lineal de los vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$.

**12.** Demostrar que el conjunto $S=\{(x+1), (x-1), x^2-1, x^2+1\}$ es un sistema de generadores de $P_{\mathbb{R}}^{(2)}[x]$. Encontrar un subconjunto $S_1$ de $S$ que sea una base de $P_{\mathbb{R}}^{(2)}[x]$.

**13.** Demostrar que el conjunto

$$S=\{\cos nx, \operatorname{sen} mx\}_{n, m \in \mathbb{N}}$$



es linealmente independiente en $C([0, 2\pi])$. [*Sugerencia:* observar que

$$\int_0^{2\pi} (\cos^2 nx)\,dx \neq 0, \quad \int_0^{2\pi} (\operatorname{sen}^2 mx)\,dx \neq 0, \quad n, m \in \mathbb{N}$$

y

$$\int_0^{2\pi} (\cos nx)(\cos mx)\,dx = \int_0^{2\pi} (\cos nx)(\operatorname{sen} mx)\,dx =$$

$$= \int_0^{2\pi} (\operatorname{sen} nx)(\operatorname{sen} mx)\,dx = 0$$

si $n \neq m$, $n, m \in \mathbb{N}$.]

**14.** Sea $V$ un espacio vectorial de dimensión $n$ y $B=\{\vec{e}_1, ..., \vec{e}_n\}$ una base de $V$. Sea $A=(a_{ij})_{i,j=1,...,n}$ una matriz de orden $n$ con determinante no nulo. Demostrar que

$$\vec{e}_j{}'=\sum_{i=1}^{n} a_{ij}\vec{e}_i, \quad j=1, 2, ..., n$$

es una base de $V$. (Observar que este resultado es el recíproco de la primera parte de la proposición 1 de la sección 5.3.)

**15.** Decir para qué valores de $a$ los vectores $\vec{u}_1=(a, 0, 1)$, $\vec{u}_2=(0, 1, 1)$ y $\vec{u}_3=(2, -1, a)$ forman una base de $\mathbb{R}^3$.

## 5.4. SUBESPACIOS VECTORIALES. INTERSECCION Y SUMA DE SUBESPACIOS VECTORIALES

Algunos subconjuntos de un espacio vectorial $V$ son a su vez espacios vectoriales con las operaciones definidas en $V$; estos subconjuntos especiales reciben el nombre de *subespacios vectoriales* de $V$.

DEFINICIÓN 1 *(Subespacio vectorial)*

Un *subespacio vectorial* de un espacio vectorial $V$ es un subconjunto $V_1$ de $V$, que a su vez es un espacio vectorial con las operaciones definidas en $V$.

Para demostrar que un subconjunto es un subespacio vectorial no es necesario comprobar de nuevo que satisface todas las propiedades de espacio vectorial. En realidad, es suficiente demostrar que la suma de dos elementos de $V_1$ es otro elemento de $V_1$ y que la multiplicación de un elemento de $V_1$ por un elemento del cuerpo $\mathbb{K}$ es otro elemento de $V_1$; es decir:

1) Si $\vec{u}, \vec{v} \in V_1$, $\vec{u}+\vec{v} \in V_1$
2) Si $a \in \mathbb{K}$ y $\vec{u} \in V_1$, $a\vec{u} \in V_1$.

En efecto, las propiedades (S1), (S2), (M1), (M2), (M3) y (M4) se cumplen en $V_1$ por cumplirse en el espacio más grande $V$; además, $0\cdot\vec{u}=\vec{0}\in V_1$ y $(-1)\vec{u}=-\vec{u}\in V_1$ debido a 2), lo cual prueba que las propiedades (S3) y (S4) son también ciertas en $V_1$.

EJEMPLO A. *a*) Todas las rectas y planos que pasan por el origen son subespacios vectoriales de $\mathbb{R}^3$.

*b*) $P_{\mathbb{R}}^{(n)}[x]$ es un subespacio vectorial de $P_{\mathbb{R}}[x]$; a su vez, $P_{\mathbb{R}}[x]$ es un subespacio vectorial de $C(\mathbb{R})$.

*c*) Los subconjuntos $\{\vec{0}\}$ y $V$ son siempre subespacios vectoriales de $V$; éstos reciben el nombre de *subespacios vectoriales impropios* de $V$ y el resto se denominan *propios.*

* * *

Además de ser intuitivamente claro no es difícil demostrar que si $V$ es un espacio vectorial de dimensión finita $n$, *cualquier subespacio $V_1$ de $V$ tiene dimensión que no supera a $n$.*

EJEMPLO B. Sea $S=\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ un conjunto de vectores de un espacio vectorial $V$. Definimos

$$L(S)=L(\vec{u}_1, \vec{u}_2, ..., \vec{u}_n)=\left\{\sum_{j=1}^{n} a_j\vec{u}_j/ \quad a_j\in\mathbb{K}, j=1, 2, ..., n\right\}.$$

El conjunto $L(S)$ es un subespacio vectorial de $V$, ya que

$$\sum_{j=1}^{n} a_j\vec{u}_j+\sum_{j=1}^{n} b_j\vec{u}_j=\sum_{j=1}^{n}(a_j+b_j)\vec{u}_j\in L(S)$$

y

$$a\sum_{j=1}^{n} a_j\vec{u}_j=\sum_{j=1}^{n}(aa_j)\vec{u}_j\in L(S).$$

Este subespacio vectorial recibe el nombre de *subespacio vectorial generado por S.*

Los vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ son un sistema de generadores de $L(S)$; de ellos se puede elegir unos cuantos que sean base de $L(S)$ (proposición 7, sección 5.2). Por tanto, la dimensión de $L(S)$ no supera a $n$.

Por ejemplo, los vectores $\vec{u}_1=(1, 0, 1)$, $\vec{u}_2=(-1, 1, 0)$ y $\vec{u}_3=(1, 1, 2)$ son linealmente dependientes en $\mathbb{R}^3$, ya que

$$\begin{vmatrix} 1 & -1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 2 \end{vmatrix} \overset{C_2+2C_1}{=\!=\!=} \begin{vmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 2 & 2 \end{vmatrix} \overset{C_2=C_3}{=\!=\!=} 0.$$

Por tanto, $L(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ es un subespacio vectorial de $\mathbb{R}^3$ de dimensión inferior a 3. Puesto que

$$\begin{vmatrix} 1 & -1 \\ 0 & 1 \end{vmatrix}=1$$



$\vec{u}_1$ y $\vec{u}_2$ son linealmente independientes, mientras que $\vec{u}_3$ depende linealmente de los dos primeros. Por tanto, $L(\vec{u}_1, \vec{u}_2)=L(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ y este subespacio vectorial tiene dimensión 2.

EJEMPLO C. Dada una matriz $A$ de $m$ filas y $n$ columnas y de rango $r$ todas las soluciones del sistema de ecuaciones homogéneo

$$A\vec{x}=\vec{0}, \quad \vec{x}\in\mathbb{R}^n \tag{I}$$

son de la forma

$$\vec{x}=a_1\vec{u}_1+a_2\vec{u}_2+\cdots+a_k\vec{u}_k$$

donde $k=n-r$, $a_j\in\mathbb{R}$ y los vectores $\vec{u}_j$ son *linealmente independientes* (véase proposición 3 de la sección 1.2 del capítulo 1). Por tanto, *las soluciones del sistema homogéneo* (I) *constituyen un subespacio vectorial de dimensión $k=n-r$ de* $\mathbb{R}^n$. Además, $\vec{u}_1, \vec{u}_2, ..., \vec{u}_k$ es una base de este subespacio.

Recíprocamente, todo subespacio vectorial de $\mathbb{R}^n$ puede determinarse por un sistema de ecuaciones lineales homogéneas. Sea $V_1$ un subespacio de $\mathbb{R}^n$ de dimensión $k$ y sea $\vec{e}_1', \vec{e}_2', ..., \vec{e}_k'$ una de sus bases. Completamos esta base hasta obtener una base

$$\vec{e}_1', \vec{e}_2', ..., \vec{e}_k', \vec{e}_{k+1}', ..., \vec{e}_n'$$

de $\mathbb{R}^n$. Es claro que si $\vec{x}=x_1'\vec{e}_1'+x_2'\vec{e}_2'+\cdots+x_n'\vec{e}_n'$ el vector $\vec{x}$ pertenece a $V_1$ si y sólo si

$$x_{k+1}'=0, ..., x_n'=0.$$

Además, si $\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2+\cdots+x_n\vec{e}_n$, donde $\{\vec{e}_1, ..., \vec{e}_n\}$ es la base canónica de $\mathbb{R}^n$, realizando el cambio de base adecuado, se tiene que

$$\begin{aligned} 0&=x_{k+1}'=a_{k+1,1}x_1+\cdots+a_{k+1,n}x_n \\ 0&=x_{k+2}'=a_{k+2,1}x_1+\cdots+a_{k+2,n}x_n \\ &\vdots \\ 0&=\ x_n'\ =\ a_{n1}x_1\ +\ \cdots\ +\ a_{nn}x_n \end{aligned}$$

que es un sistema homogéneo de $n-k$ ecuaciones con $n$ incógnitas que determina el subespacio $V_1$.

* * *

Dados dos subespacios vectoriales $V_1$ y $V_2$ de un espacio vectorial $V$ podemos definir su *intersección*

$$V_1\cap V_2=\{\vec{u}/\vec{u}\in V_1 \text{ y } \vec{u}\in V_2\}$$

y su *suma*

$$V_1+V_2=\{\vec{u}_1+\vec{u}_2/\vec{u}_1\in V_1, \vec{u}_2\in V_2\}.$$

Como puede comprobarse fácilmente estos dos nuevos subconjuntos son también subespacios vectoriales de $V$. La relación que existe entre las dimensiones de estos subespacios vectoriales y las dimensiones de los subespacios $V_1$ y $V_2$ queda plasmada en el siguiente resultado:

PROPOSICIÓN 2

$$\dim(V_1)+\dim(V_2)=\dim(V_1+V_2)+\dim(V_1\cap V_2)$$

para cualesquiera subespacios vectoriales $V_1$ y $V_2$ de un espacio vectorial $V$ de dimensión finita.

*Demostración.* Sea $\vec{e}_1, \ldots, \vec{e}_l$ una base de $V_1 \cap V_2$; completar esta base hasta obtener una base

$$\vec{e}_1, \ldots, \vec{e}_l, \vec{f}_{l+1}, \ldots, \vec{f}_k$$

de $V_1$ y una base

$$\vec{e}_1, \ldots, \vec{e}_l, \vec{g}_{l+1}, \ldots, \vec{g}_m$$

de $V_2$. Demostraremos que

$$S=\{\vec{e}_1, \ldots, \vec{e}_l, \vec{f}_{l+1}, \ldots, \vec{f}_k, \vec{g}_{l+1}, \ldots, \vec{g}_m\}$$

es una base de $V_1+V_2$. De aquí se deducirá que

$$\begin{aligned}\dim(V_1+V_2)&=l+(k-l)+(m-l)=k+m-l=\\&=\dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2)\end{aligned}$$

que era lo que queríamos demostrar.

Para demostrar que $S$ es una base de $V_1+V_2$ basta demostrar que $S$ es un conjunto de vectores linealmente independientes, ya que debido a la construcción realizada resulta claro que todo elemento de $V_1+V_2$ es una combinación lineal de elementos de $S$. Supongamos, por tanto, que tenemos una expresión de la forma

$$a_1\vec{e}_1+\cdots+a_l\vec{e}_l+b_{l+1}\vec{f}_{l+1}+\cdots+b_k\vec{f}_k+c_{l+1}\vec{g}_{l+1}+\cdots+c_m\vec{g}_m=\vec{0}.$$

El vector $\vec{v}=a_1\vec{e}_1+\cdots+a_l\vec{e}_l+b_{l+1}\vec{f}_{l+1}+\cdots+b_k\vec{f}_k\in V_1$ coincide con $-c_{l+1}\vec{g}_{l+1}-\cdots-c_m\vec{g}_m\in V_2$ y, por tanto, es un elemento de $V_1\cap V_2$. Como $V_1\cap V_2$ está generado por los vectores $\vec{e}_1, \ldots, \vec{e}_l$ hemos de tener $c_{l+1}=\cdots=c_m=0$. Por tanto, la igualdad anterior se transforma en

$$a_1\vec{e}_1+\cdots+a_l\vec{e}_l+b_{l+1}\vec{f}_{l+1}+\cdots+b_k\vec{f}_k=\vec{0}$$

Puesto que $\{\vec{e}_1, \ldots, \vec{e}_l, \vec{f}_{l+1}, \ldots, \vec{f}_k\}$ es una base de $V_1$ deducimos que $a_1=\cdots=a_l=b_{l+1}=\cdots=b_k=0$. Esto termina la demostración. ■



EJEMPLO D. Dos subespacios vectoriales $V_1$ y $V_2$ en $\mathbb{R}^3$, ambos de dimensión 2, que se cortan en una recta, han de tener una suma que coincida con todo $\mathbb{R}^3$, ya que

$$\dim(V_1+V_2)=\dim(V_1)+\dim(V_2)-\dim(V_1\cap V_2)=2+2-1=3$$

* * *

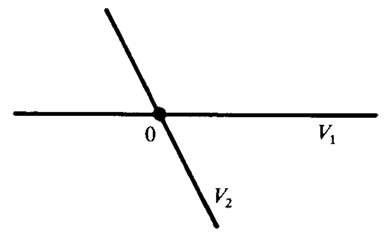
DEFINICIÓN 3

Un espacio vectorial $V$ es *suma directa* de dos de sus subespacios $V_1$ y $V_2$ si:

1) $V_1+V_2=V$.
2) $V_1\cap V_2=\{\vec{0}\}$.

Utilizaremos la notación $V=V_1\oplus V_2$ para indicar que $V$ es suma directa de los subespacios vectoriales $V_1$ y $V_2$.

*Observaciones.* 1) El plano $\mathbb{R}^2$ puede escribirse como suma directa de dos rectas no coincidentes que pasan por el origen.




2) El espacio $\mathbb{R}^3$ puede escribirse como suma directa de un plano que pasa por el origen y una recta que le corta en este punto.

3) De la proposición 2 deducimos que si $V=V_1\oplus V_2$ se tiene que

$$\dim(V)=\dim(V_1\oplus V_2)=\dim(V_1)+\dim(V_2)$$

ya que el subespacio vectorial impropio $\{\vec{0}\}$ tiene dimensión 0.

Si $V=V_1+V_2$ todo elemento $\vec{v}\in V$ puede escribirse de la forma $\vec{v}=\vec{v}_1+\vec{v}_2$ con $\vec{v}_1\in V_1$ y $\vec{v}_2\in V_2$. Si la suma es directa, esta descomposición es única:

PROPOSICIÓN 4

Sean $V_1$ y $V_2$ subespacios vectoriales de un espacio vectorial $V$. Las siguientes afirmaciones son equivalentes:

1) $V=V_1\oplus V_2$.
2) Para todo $\vec{v}\in V$ existe una descomposición *única* de la forma $\vec{v}=\vec{v}_1+\vec{v}_2$, con $\vec{v}_1\in V_1$ y $\vec{v}_2\in V_2$.

*Demostración.* Para demostrar que 1) implica 2) es suficiente probar la unicidad; supongamos que $\vec{v}=\vec{v}_1+\vec{v}_2$ y $\vec{v}=\vec{u}_1+\vec{u}_2$ con $\vec{v}_1, \vec{u}_1 \in V_1$ y $\vec{v}_2, \vec{v}_2 \in V_2$. Tenemos

$$\vec{v}_1+\vec{v}_2=\vec{u}_1+\vec{u}_2$$

o bien

$$\vec{v}_1-\vec{u}_1=\vec{u}_2-\vec{v}_2.$$

La parte izquierda de esta igualdad es un elemento de $V_1$, mientras que la parte derecha es un elemento de $V_2$. Puesto que $\vec{0}$ es el único vector que $V_1$ y $V_2$ tienen en común, deducimos que

$$\vec{v}_1-\vec{u}_1=\vec{u}_2-\vec{v}_2=\vec{0}$$

Por tanto:

$$\vec{u}_1=\vec{v}_1 \quad \text{y} \quad \vec{u}_2=\vec{v}_2$$

que era lo que queríamos demostrar.

Para demostrar que 2) implica 1) es suficiente demostrar que $V_1 \cap V_2=\{\vec{0}\}$. Si $\vec{v} \in V_1 \cap V_2$, podemos escribir

$$\vec{v}=\vec{v}+\vec{0}=\vec{0}+\vec{v}.$$

Puesto que la descomposición ha de ser única se tiene que $\vec{v}=\vec{0}$. Esto prueba el resultado deseado. ■

* * *

También podemos definir la intersección, la suma y la suma directa de más de dos subespacios vectoriales. En general, dados $n$ subespacios vectoriales $V_1, V_2, \ldots, V_n$ de un espacio vectorial $V$ definimos

$$\bigcap_{j=1}^{n} V_j=\{\vec{u} \in V/\vec{u} \in V_j \text{ para todo } j=1, \ldots, n\}$$

y

$$\sum_{j=1}^{n} V_j=\left\{\sum_{j=1}^{n} \vec{u}_j/\vec{u}_j \in V_j,\ j=1, \ldots, n\right\}$$

que reciben el nombre de *intersección* y *suma*, respectivamente, de los subespacios vectoriales dados. Estos dos nuevos subespacios son también subespacios vectoriales de $V$.

La definición de *suma directa* de varios subespacios vectoriales es un poco más complicada en general, que si solamente hay dos. La definición más razonable es la



que guarda similitud con la afirmación 2) de la proposición 4. Diremos que $V$ es *suma directa* de los subespacios vectoriales $V_1, V_2, \ldots, V_n$, y escribiremos

$$V=V_1 \oplus V_2 \oplus \cdots \oplus V_n$$

o

$$V=\bigoplus_{i=1}^{n} V_i$$

si todo vector $\vec{v}$ de $V$ tiene una descomposición *única* de la forma $\vec{v}=\sum_{i=1}^{n} \vec{v}_i$, con $\vec{v}_i \in V_i$, $i=1, 2, \ldots, n$.

Si $n=2$, la proposición 4 muestra que esta definición de suma directa y la dada anteriormente son equivalentes. Si $n \geqq 3$ se pide demostrar en el ejercicio 6 al final de esta sección que una condición equivalente de suma directa es la siguiente:

1) $\sum_{i=1}^{n} V_i=V$

2) $V_i \cap \left(\sum_{k \neq i} V_k\right)=\{\vec{0}\}$ para todo $i=1, 2, \ldots, n$

A partir de aquí puede demostrarse que

$$\dim\left(\bigoplus_{i=1}^{n} V_i\right)=\sum_{i=1}^{n} \dim(V_i),$$

lo cual se obtiene aplicando repetidas veces el resultado ya conocido cuando $n$ es igual a 2.

## EJERCICIOS 5.4

**1.** Los siguientes subconjuntos y familias de vectores de algunos espacios vectoriales son subespacios y bases de éstos. Verificar la verdad o falsedad de esto en los ejemplos siguientes:

*a)* $\{(a, b) \in \mathbb{R}^2/b=1\}$; $\{(2, 1)\}$

*b)* $\{p(x) \in P_{\mathbb{C}}^{(3)}[x]/(x-1) \text{ divide a } p(x)\}$; $\{x-1, x^2-1\}$

*c)* $\operatorname{com}(B)=\{A \in \mathcal{M}_{2\times 2}(\mathbb{R})/BA=AB\}$ con $B=\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$; $\left\{\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right\}$

**2.** Estudiar si los siguientes subconjuntos de $\mathcal{M}_{2\times 2}(\mathbb{R})$ son subespacios vectoriales:

*a)* $R=\{A \in \mathcal{M}_{2\times 2}(\mathbb{R})/r(A)=1\}$ donde $r(A)$ designa el rango de $A$.

*b)* $T=\{A \in \mathcal{M}_{2\times 2}(\mathbb{R})/\text{traza}(A)=0\}$, donde traza$(A)$ denota la suma de los elementos de la diagonal principal de $A$.

**3.** Encontrar la dimensión del subespacio generado por los siguientes conjuntos de vectores:

*a)* $\{(1, 2), (0, 1), (-1, 3)\}$ en $\mathbb{R}^2$.

*b)* $\left\{\begin{pmatrix} i & 1 \\ 0 & 1+i \end{pmatrix}, \begin{pmatrix} 1+i & 1-i \\ 0 & 2 \end{pmatrix}\right\}$ en $\mathcal{M}_{2\times 2}(\mathbb{C})$.

**4.** Hallar en cada uno de los ejemplos siguientes la suma y la intersección del par de subespacios dados, y comprobar que se verifica la ecuación

$$\dim(V_1)+\dim(V_2)=\dim(V_1+V_2)+\dim(V_1\cap V_2)$$

*a)* $V_1=\mathrm{com}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $V_2=\mathrm{com}\begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}$ (ver el ejercicio 1 para la definición de $\mathrm{com}(B)$).

*b)* $V_1=\{p(x)\in P_{\mathbb{C}}^{(3)}[x]/x+1$ divide a $p(x)\}$,
$V_2=\{p(x)\in P_{\mathbb{C}}^{(3)}[x]/x-1$ divide a $p(x)\}$

*c)* $L\{\operatorname{sen} t, \cos t\}$, $L(e^{it}, e^{-it})$ considerados ambos como subespacios del espacio vectorial *real* $C_{\mathbb{C}}([0, 1])$.

**5.** Decir cuáles de los siguientes pares de subespacios se suman directamente y cuál es la suma (directa o no) en cada caso:

*a)* $\mathcal{S}=\{A\in\mathcal{M}_{2\times 2}(\mathbb{C})/A=A^t\}$ (matrices simétricas)
$\mathcal{A}=\{A\in\mathcal{M}_{2\times 2}(\mathbb{C})/A=-A^t\}$ (matrices antisimétricas)

*b)* $P=\{f\colon[-1, 1]\mapsto\mathbb{R}/f(-t)=f(t)\}$ (funciones pares)
$I=\{f\colon[-1, 1]\mapsto\mathbb{R}/f(-t)=-f(t)\}$ (funciones impares)
[*Sugerencia:* $g(t)=(g(t)+g(-t))/2)+(g(t)-g(-t))/2)$]

*c)* $V_1=\{f\colon[-1, 1]\mapsto\mathbb{R}/f(t)=0,\ t\geqq 0\}$
$V_2=\{f\colon[-1, 1]\mapsto\mathbb{R}/f(t)=0,\ t\leqq 0\}$

**6.** Sean $V_1, V_2, \ldots, V_n$ subespacios vectoriales de un espacio vectorial $V$. Demostrar que si $n\geqq 3$ las siguientes afirmaciones son equivalentes:

*a)* $V=\bigoplus_{i=1}^{n} V_i$

*b)* $V=\sum_{i=1}^{n} V_i$ y $V_i\cap\left(\sum_{i\neq k} V_k\right)=\{\vec{0}\}$ para todo $i=1, \ldots, n$.

## 5.5. VARIEDADES LINEALES. ESPACIO AFIN

En el ejemplo C de la sección anterior se observó que si $A$ es una matriz de $m$ filas y $n$ columnas de rango $r$, las soluciones del sistema homogéneo

$$A\vec{x}=\vec{0},\quad \vec{x}\in\mathbb{R}^n \tag{I}$$



son un subespacio vectorial de $\mathbb{R}^n$ de dimensión $k=n-r$. Este subespacio vectorial, que denotamos por S(I), es un subespacio vectorial generado por $k$ vectores linealmente independientes $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k$ que sean solución de (I); es decir:

$$\mathrm{S(I)}=L(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k).$$

Si consideramos el sistema *no* homogéneo

$$A\vec{x}=\vec{b},\quad \vec{x}\in\mathbb{R}^n,\ \vec{b}\in\mathbb{R}^m \tag{II}$$

sabemos que sus soluciones pueden escribirse de la forma

$$\vec{x}=\vec{v}+c_1\vec{u}_1+\cdots+c_k\vec{u}_k,\quad c_j\in\mathbb{R}$$

(teorema 2, sección 1.2, capítulo 1). Por tanto, si denotamos por S(II) al conjunto de las soluciones de (II), podemos escribir

$$\mathrm{S(II)}=\vec{v}+\mathrm{S(I)}$$

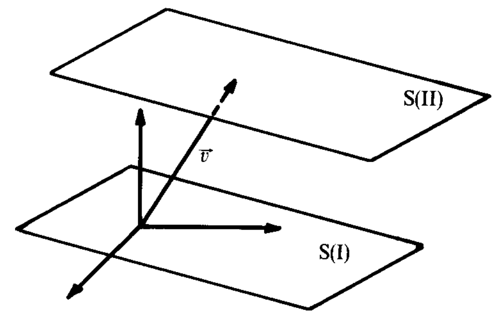



Figura 1

Este conjunto se obtiene trasladando $\mathrm{S(I)}=L(\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_k)$ mediante el vector $\vec{v}$. Si $k=1$ se obtiene una recta, si $k=2$ se obtiene un plano y, en general, si $k>2$ los objetos que se obtienen se denominan *k-planos*. Todos ellos reciben el nombre de *variedades lineales.*

Este tipo de construcción puede hacerse no solamente en $\mathbb{R}^n$, sino en otros espacios vectoriales.

Sea $\mathcal{P}$ un conjunto de elementos, que se denominarán *puntos,* y $V$ un espacio vectorial. El par $(\mathcal{P}, V)$ recibe el nombre de *espacio afín, A,* si todo par ordenado $M$, $N$ de puntos puede ponerse en correspondencia con un solo vector $\vec{v}$ de $V$, lo cual escribiremos de la forma $\overrightarrow{MN}=\vec{v}$, y que satisface las siguientes propiedades:

1) Para todo punto $M$ y para todo vector $\vec{v}$ de $V$, existe un solo punto $N$ tal que $\overrightarrow{MN}=\vec{v}$. Escribiremos entonces $N=M+\vec{v}$.
2) Para cada tres puntos $M$, $N$ y $P$ se tiene que $\overrightarrow{MN}+\overrightarrow{NP}=\overrightarrow{MP}$.


Figura 2

En muchas ocasiones el conjunto $\mathscr{P}$ de puntos y el espacio vectorial $V$ coinciden como conjuntos; éste es el caso de $\mathbb{R}^2$, $\mathbb{R}^3$, o en general, de $\mathbb{R}^n$ o $\mathbb{C}^n$.

El espacio afín $A$ se dice que tiene *dimensión n* si el espacio vectorial $V$ subyacente a $A$ es de dimensión $n$. Todo subconjunto de $A$ de la forma

$$P+V_1$$

donde $P$ es un punto y $V_1$ es un subespacio vectorial de $V$, recibe el nombre de *variedad lineal de A*. Como casos particulares de variedades lineales podemos citar las rectas, los planos y, en general, los hiperplanos en $\mathbb{R}^n$.

* * *

En un espacio afín $A$ un *sistema de referencia* está formado por un punto $O$, que se considera el origen, y una base $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ del espacio vectorial subyacente. Utilizaremos la notación

$$\mathscr{R}=\{O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$$

para designar este sistema de referencia.

Dado un punto $X$, existe un vector $\vec{x}$ de $V$ que está en correspondencia con el par de puntos $O$, $X$; si

$$\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2+\cdots+x_n\vec{e}_n$$

diremos que $(x_1, x_2, ..., x_n)$ son las *coordenadas del punto X con respecto al sistema de referencia $\mathscr{R}$*. Por abuso de notación escribiremos

$$\overrightarrow{OX}=x_1\vec{e}_1+x_2\vec{e}_2+\cdots+x_n\vec{e}_n.$$

Dado un nuevo sistema de referencia $\mathscr{R}''=\{P; \vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ en el espacio afín, las coordenadas del punto $X$ con respecto a este nuevo sistema de referencia serán distintas de las anteriores. Para encontrar estas nuevas coordenadas $(x_1'', x_2'', ..., x_n'')$ de $X$ con respecto a $\mathscr{R}''$ procedemos en dos etapas:

*a*) Cambiamos el sistema de referencia $\mathscr{R}$ a $\mathscr{R}'=\{P; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$, de manera que solamente hemos variado el origen.

*b*) Pasamos de $\mathscr{R}'$ a $\mathscr{R}''$ dejando fijo el origen y realizando el cambio de base en el espacio vectorial subyacente.



En la primera etapa el cambio de sistema de referencia se realiza de la siguiente manera: si $(x_1', x_2', ..., x_n')$ designan las coordenadas del punto $X$ con respecto al sistema de referencia $\mathscr{R}'$ y $(p_1, p_2, ..., p_n)$ las del punto $P$ con respecto a $\mathscr{R}$ se tiene que

$$\overrightarrow{OX}=\overrightarrow{OP}+\overrightarrow{PX} \Leftrightarrow (x_1, x_2, ..., x_n)=(p_1, p_2, ..., p_n)+(x_1', x_2', ..., x_n');$$

por tanto:

$$(x_1', x_2', ..., x_n')=(x_1-p_1, x_2-p_2, ..., x_n-p_n).$$

En la segunda etapa se trata de realizar únicamente un cambio de base en el espacio vectorial subyacente: si $(x_1'', x_2'', ..., x_n'')$ son las coordenadas del punto $X$ con respecto a $\mathscr{R}''$ y $A$ es la matriz del cambio de la base $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ a la base $\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ se tiene que

$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}=A\begin{pmatrix} x_1'' \\ x_2'' \\ \vdots \\ x_n'' \end{pmatrix}.$$

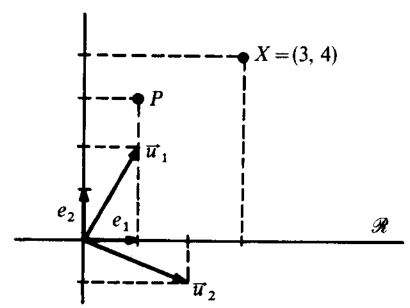
Combinando los resultados de estas dos etapas se tiene:

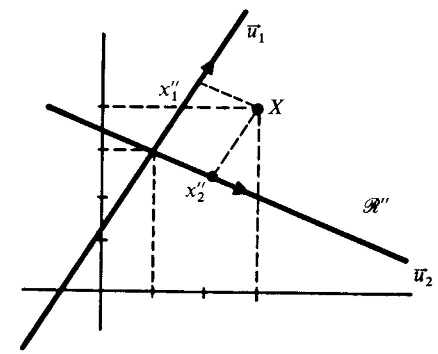
$$\begin{pmatrix} x_1'' \\ x_2'' \\ \vdots \\ x_n'' \end{pmatrix}=A^{-1}\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix}=A^{-1}\begin{pmatrix} x_1-p_1 \\ x_2-p_2 \\ \vdots \\ x_n-p_n \end{pmatrix}=-A^{-1}\begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}+A^{-1}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

que nos da las coordenadas de $X$ con respecto al sistema de referencia $\mathscr{R}''$.

EJEMPLO A. Sea $X$ un punto de coordenadas (3, 4) en un sistema de referencia $\mathscr{R}=\{O; \vec{e}_1, \vec{e}_2\}$ del espacio afín $\mathbb{R}^2$ y sea $\mathscr{R}''=\{P; \vec{u}_1, \vec{u}_2\}$ un nuevo sistema de referencia, donde $P$ tiene como coordenadas (1, 3) con respecto a $\mathscr{R}$ y $\vec{u}_1=\vec{e}_1+2\vec{e}_2$, $\vec{u}_2=2\vec{e}_1-\vec{e}_2$. Las coordenadas del punto $X$ con respecto a $\mathscr{R}''$ son

$$\begin{pmatrix} x_1'' \\ x_2'' \end{pmatrix}=\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}^{-1}\begin{pmatrix} 3-1 \\ 4-3 \end{pmatrix}=\frac{1}{5}\begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix}=\begin{pmatrix} 4/5 \\ 3/5 \end{pmatrix}$$



EJEMPLO B. Sea $\mathscr{R}=\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$ un sistema de referencia en $\mathbb{R}^3$ con respecto al cual el plano $\pi$ tiene de ecuación $x_1+x_2+x_3=2$. Sea $\mathscr{R}'=\{P; \vec{u}_1, \vec{u}_2, \vec{u}_3\}$ un nuevo sistema de referencia tal que $P=(1, 1, 0)$, $\vec{u}_1=-\vec{e}_1+\vec{e}_3$, $\vec{u}_2=-\vec{e}_2+\vec{e}_3$ y $\vec{u}_3=\vec{e}_1+\vec{e}_2+\vec{e}_3$.

Para encontrar la ecuación del plano $\pi$ con respecto a $\mathscr{R}'$ observamos que

$$(x_1, x_2, x_3)=(1, 1, 0)+t(-1, 0, 1)+s(0, -1, 1)$$

son coordenadas paramétricas del plano $\pi$; por tanto, las coordenadas $(x_1, x_2, x_3)$ de un punto $X$ de $\pi$ se transforman en $(x_1', x_2', x_3')$ de manera que

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}^{-1} \begin{pmatrix} x_1-1 \\ x_2-1 \\ x_3-0 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} -t \\ -s \\ t+s \end{pmatrix} =$$

$$= \frac{1}{3}\begin{pmatrix} -2t-s+t+s \\ -t+2s+t+s \\ -t-s+t+s \end{pmatrix} = \begin{pmatrix} -t \\ s \\ 0 \end{pmatrix}$$

Por tanto, $x_3'=0$ es la ecuación del plano $\pi$ con respecto al sistema de referencia $\mathscr{R}'$.

## EJERCICIOS 5.5

**1.** En el plano, y respecto a la referencia canónica $\mathscr{R}$, se dan los puntos $A=(1, 1)$ y $B=(-2, 0)$, los vectores $\vec{u}_1=(1, 2)$ y $\vec{u}_2=(-1, 1)$ y la recta $r \equiv x_1-x_2=1$. Hallar:

*a)* Las coordenadas de $B$ con respecto al sistema de referencia $\mathscr{R}'=\{A; \vec{u}_1,\vec{u}_2\}$.

*b)* La ecuación de la recta $r$ con respecto a $\mathscr{R}'$.

**2.** Sea $\mathscr{R}'$ el sistema de referencia en el plano que se obtiene girando un ángulo $\alpha$ en sentido positivo los vectores de un sistema de referencia dado $\mathscr{R}$. Si $C \equiv (x_1-1)^2+x_2^2=4$ es la ecuación de una circunferencia respecto a $\mathscr{R}$, encontrar la ecuación de $C$ respecto a $\mathscr{R}'$. ¿Cuál es el centro de la nueva circunferencia respecto a $\mathscr{R}'$?

**3.** Dado el plano de ecuación $x_1-x_2+x_3=3$ con respecto a un sistema de referencia $\mathscr{R}$ en $\mathbb{R}^3$, encontrar un sistema de referencia $\mathscr{R}'$ de $\mathbb{R}^3$ en el que el plano anterior tenga por ecuación $x_2'=0$.

**4.** Dada la circunferencia $C \equiv (x_1-1)^2+(x_2-1)^2=9$ en el plano, encontrar sus ecuaciones en el sistema de referencia $\mathscr{R}'=\{A; \vec{u}_1,\vec{u}_2\}$, donde $A=(3, 3)$, $\vec{u}_1=\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$, $\vec{u}_2=\left(\frac{2}{\sqrt{2}}, -\frac{2}{\sqrt{2}}\right)$. ¿Qué tipo de curva es $C$ en el sistema de referencia $\mathscr{R}'$?

CAPITULO 6

# APLICACIONES LINEALES ENTRE ESPACIOS VECTORIALES



## 6.1. DEFINICION DE APLICACION LINEAL. EJEMPLOS

En la sección 3 del capítulo 1 hemos definido el concepto de aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$. Recordamos que en ese contexto toda aplicación lineal quedaba determinada por una matriz de $m$ filas y $n$ columnas; además, las siguientes propiedades caracterizan una aplicación lineal $A$ entre estos espacios:

1) $A(\vec{x}+\vec{y})=A(\vec{x})+A(\vec{y})$ para todo $\vec{x}, \vec{y} \in \mathbb{R}^n$
2) $A(r\vec{x})=rA(\vec{x})$ para todo $\vec{x} \in \mathbb{R}^n$ y todo número real $r$.

(Ver los teoremas 1 y 2 de la sección antes citada.)

Utilizaremos esta caracterización para definir las aplicaciones lineales entre dos espacios vectoriales cualesquiera. En la siguiente sección mostraremos que, fijadas dos bases en los espacios vectoriales, toda aplicación lineal queda determinada por una matriz.

DEFINICIÓN 1 *(Aplicación lineal)*

Sean $V$ y $W$ espacios vectoriales; una *aplicación lineal* $A$ de $V$ en $W$ es una aplicación $A$: $V \mapsto W$ tal que:

1) $A(\vec{v}+\vec{w})=A(\vec{v})+A(\vec{w})$ para todo $\vec{v}, \vec{w} \in V$.
2) $A(a\vec{v})=aA(\vec{v})$ para todo $a \in \mathbb{K}$ y todo $\vec{v} \in V$.

*Nota.* Observar que ambos conjuntos $V$ y $W$ deben ser espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$.

En el capítulo 1 se han dado varios ejemplos de aplicaciones lineales entre los espacios vectoriales $\mathbb{R}^n$ y $\mathbb{R}^m$. Recordamos únicamente los giros en el plano y la simetría con respecto a un eje en $\mathbb{R}^2$.

EJEMPLO A. Dado un número real $a$, la aplicación que asocia a cada polinomio del conjunto $P_{\mathbb{R}}[x]$ su valor para $x=a$ es una aplicación lineal; esta aplicación queda definida mediante las siguientes expresiones:

$$A\colon P_{\mathbb{R}}[x] \mapsto \mathbb{R}/A(p(x)) = p(a).$$

El hecho de que $A$ es una aplicación lineal se deduce de las siguientes igualdades:

$$A(p(x)+q(x)) = A((p+q)(x)) = (p+q)(a) = p(a)+q(a) = \\ = A(p(x)) + A(q(x)$$

y

$$A(cp(x)) = A(cp)(x)) = (cp)(a) = c(p(a)) = cA(p(x))$$

para todo número real $c$.

EJEMPLO B. Dada la matriz

$$A = \begin{pmatrix} i & 1 & 0 \\ 1 & i & 1 \end{pmatrix}$$

de elementos complejos, podemos definir la aplicación

$$A\colon \mathbb{C}^3 \mapsto \mathbb{C}^2$$

dada por $A((z_1, z_2, z_3)) = (iz_1 + z_2, z_1 + iz_2 + z_3)$. Puesto que

$$A((z_1, z_2, z_3)) = A\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

donde en la parte derecha se realiza una multiplicación de matrices, es fácil demostrar que $A$ es una aplicación lineal.

En general, dada una matriz $A \in \mathcal{M}_{m\times n}(\mathbb{K})$, la aplicación

$$A\colon \mathbb{K}^n \mapsto \mathbb{K}^m$$



dada por

$$A((z_1, z_2, \ldots, z_n)) = A\begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix}$$

es una aplicación lineal, que recibe el nombre de *aplicación lineal asociada con la matriz A*.

EJEMPLO C. Sea $D$ la aplicación que a cada polinomio le hace corresponder su derivada, es decir:

$$D\colon P_{\mathbb{R}}[x] \mapsto P_{\mathbb{R}}[x]/D(a_0 + a_1x + \cdots + a_nx^n) = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

Esta es una aplicación lineal, ya que la derivada de una suma de funciones es la suma de las derivadas de cada una de las funciones y la derivada de un número real por una función coincide con el producto del número real por la derivada de la función.

Observar que si el conjunto inicial de esta aplicación es $P_{\mathbb{R}}^{(n)}[x]$, como conjunto final podemos tomar el mismo, o incluso $P_{\mathbb{R}}^{(n-1)}[x]$; esto es cierto ya que la derivada de un polinomio de grado no superior a $n$ es otro polinomio de grado no superior a $n-1$.

EJEMPLO D. La aplicación que a cada matriz cuadrada le hace corresponder su determinante no es una aplicación lineal, ya que, en general, el determinante de una suma de matrices no coincide con la suma de los determinantes de cada una de ellas. Baste como ejemplo el siguiente:

$$-1 = \det\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right) \neq \det\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \det\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = 0$$

* * *

Aplicando repetidas veces las propiedades 1) y 2) de la definición de aplicación lineal entre espacios vectoriales se consigue demostrar que la imagen, mediante una aplicación lineal, de una combinación lineal de vectores del espacio vectorial inicial es una combinación lineal de vectores del espacio vectorial final. De forma más precisa:

$$A\left(\sum_{j=1}^{n} c_j\overline{v}_j\right) = \sum_{j=1}^{n} c_jA(\overline{v}_j)$$

donde $c_j \in \mathbb{K}$ y $\overline{v}_j \in V$ para todo $j=1, 2, \ldots, n$.

Otras propiedades que se deducen inmediatamente de la definición de aplicación lineal son las siguientes:

PROPOSICIÓN 2

Sea $A$ una aplicación lineal entre los espacios vectoriales $V$ y $W$. Se tienen los siguientes resultados:

1) La imagen del elemento neutro de $V$ mediante $A$ es el elemento neutro de $W$, es decir, $A(\vec{0})=\vec{0}$.
2) La imagen mediante $A$ del opuesto de un elemento $\vec{v}$ de $V$ es el opuesto de $A(\vec{v})$, es decir, $A(-\vec{v})=-A(\vec{v})$.

*Demostración.* Para demostrar 1) escribimos

$$A(\vec{0})=A(\vec{0}+\vec{0})=A(\vec{0})+A(\vec{0})$$

y restamos $A(\vec{0})$ en ambos lados. Para demostrar 2) basta observar que

$$A(-\vec{v})=A((-1)\vec{v})=(-1)A(\vec{v})=-A(\vec{v})$$ ■

En el capítulo 1 demostramos que la imagen de una recta mediante una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ es otra recta o un punto. Para aplicaciones lineales entre espacios vectoriales se tiene el siguiente resultado:

PROPOSICIÓN 3

Sea $A: V \mapsto W$ una aplicación lineal entre espacios vectoriales. La imagen mediante $A$ de cualquier subespacio vectorial de $V$ es un subespacio vectorial de $W$.

*Demostración.* Sea $V_1$ un subespacio vectorial de $V$ y sea $W_1=A(V_1)$ la imagen de $V_1$ mediante $A$. Recordamos que

$$W_1=A(V_1)=\{A(\vec{v}_1)/\vec{v}_1 \in V\}.$$

Sean $\vec{w}_1, \vec{w}_1' \in W_1$; existen, por tanto, dos vectores $\vec{v}_1, \vec{v}_1' \in V_1$ tal que $A(\vec{v}_1)=\vec{w}_1$ y $A(\vec{v}_1')=\vec{w}_1'$. Entonces

$$\vec{w}_1+\vec{w}_1'=A(\vec{v}_1)+A(\vec{v}_1')=A(\vec{v}_1+\vec{v}_1')$$

y puesto que $\vec{v}_1+\vec{v}_1' \in V_1$, por ser $V_1$ un subespacio vectorial de $V$, se tiene que $\vec{w}_1+\vec{w}_1' \in W_1$.

Tomemos ahora $a \in \mathbb{K}$ y $w_1 \in W_1$; existe $v_1 \in V_1$ tal que $A(\vec{v}_1)=\vec{w}_1$. Por tanto,

$$a\vec{w}_1=aA(\vec{v}_1)=A(a\vec{v}_1)$$

y puesto que $a\vec{v}_1 \in V_1$, por ser $V_1$ un subespacio vectorial de $V$, se tiene que $a\vec{w}_1 \in W_1$.



Estos dos resultados son suficientes para probar que $W_1$ es un subespacio vectorial de $W$. ■

* * *

Si el subespacio $V_1$ tiene $\{\vec{v}_1, ..., \vec{v}_k\}$ como base, todo elemento $w$ de $W_1$ puede escribirse como combinación lineal de los vectores $A(\vec{v}_1), ..., A(\vec{v}_k)$. Esto es cierto ya que tomando $\vec{v} \in V_1$ tal que $A(\vec{v})=\vec{w}$ se tiene que

$$\vec{w}=A(\vec{v})=A\left(\sum_{j=1}^{k} c_j\vec{v}_j\right)=\sum_{j=1}^{k} c_jA(\vec{v}_j)$$

Por tanto, $W_1$ coincide con el subespacio generado por los vectores $A(\vec{v}_1), ..., A(\vec{v}_k)$, es decir, $W_1=L(A(\vec{v}_1), ..., A(\vec{v}_k))$. En consecuencia, la dimensión de $W_1$ no puede superar $k$.

Hemos probado el siguiente resultado:

PROPOSICIÓN 4

La imagen mediante una aplicación lineal de un subespacio vectorial de dimensión $k$ es un subespacio vectorial de dimensión no superior a $k$.

Para finalizar esta sección demostramos que una aplicación lineal queda determinada cuando se conocen las imágenes de los elementos de una base del espacio inicial. El enunciado preciso de esta afirmación se da en el siguiente teorema:

TEOREMA 5

Sea $B=\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ una base de un espacio vectorial $V$ y sean $\vec{w}_1, \vec{w}_2, ..., \vec{w}_n$ $n$ vectores cualesquiera de otro espacio vectorial $W$. En estas condiciones, existe una única aplicación lineal $A$ de $V$ en $W$ tal que

$$A(\vec{e}_j)=\vec{w}_j, \quad j=1, 2, ..., n$$

*Demostración.* Definimos $A$ de la siguiente manera: dado $\vec{v} \in V$ podemos escribir $\vec{v}=\sum_{j=1}^{n} v_j\vec{e}_j$, con $v_j \in \mathbb{K}$; entonces definimos

$$A(\vec{v})=\sum_{j=1}^{n} v_j\vec{w}_j.$$

A partir de aquí es un simple ejercicio comprobar que $A$ es la única aplicación lineal tal que $A(\vec{e}_j)=\vec{w}_j$, $j=1, 2, ..., n$. ■

## 6.2. MATRIZ DE UNA APLICACION LINEAL. OPERACIONES CON APLICACIONES LINEALES

En esta sección, y mientras no se indique lo contrario, $V$ y $W$ denotarán dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y $A$ una aplicación lineal de $V$ en $W$.

Sea $B=\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ una base de $V$ y $\bar{B}=\{\vec{f}_1, \vec{f}_2, ..., \vec{f}_m\}$ una base de $W$. El elemento $A(\vec{e}_1)$ es un vector de $W$ y, por tanto, podemos escribir

$$A(\vec{e}_1)=a_{11}\vec{f}_1+a_{21}\vec{f}_2+\cdots+a_{m1}\vec{f}_m.$$

Análogamente,

$$\begin{aligned} A(\vec{e}_2)&=a_{12}\vec{f}_1+a_{22}\vec{f}_2+\cdots+a_{m2}\vec{f}_m\\ &\ \ \vdots \\ A(\vec{e}_n)&=a_{1n}\vec{f}_1+a_{2n}\vec{f}_2+\cdots+a_{mn}\vec{f}_m \end{aligned}$$

Estas igualdades se escriben abreviadamente de la forma

$$A(\vec{e}_j)=\sum_{i=1}^{m} a_{ij}\vec{f}_i, \quad j=1, 2, ...., n.$$

En estas condiciones diremos que

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

es la *matriz de la aplicación* $A$ *con respecto a las bases* $B$ *y* $\bar{B}$.

Observar que la $j$-ésima columna de la matriz de la aplicación lineal $A$ son las componentes de $A(\vec{e}_j)$ con respecto a la base $\bar{B}$ de $W$.

Para ser precisos sería necesario designar la matriz $A$ con un símbolo que incluyera las bases $B$ y $\bar{B}$; este símbolo podría ser $M_{\bar{B}}^{B}(A)$. Nosotros preferimos denominar a la matriz con la misma letra que a la aplicación, siempre que esta economía en la notación no sea causa de incomprensión.

Dado $\vec{x}\in V$, podemos escribir

$$\vec{x}=\sum_{j=1}^{n} x_j\vec{e}_j \quad \text{e} \quad \vec{y}=A(\vec{x})=\sum_{i=1}^{m} y_i\vec{f}_i.$$

La relación entre las coordenadas $y_i$ y $x_j$ viene dada por la matriz $A$. En efecto, la igualdad

$$\begin{aligned}\sum_{i=1}^{m} y_i\vec{f}_i=A(\vec{x})=A\left(\sum_{j=1}^{n} x_j\vec{e}_j\right)=\sum_{j=1}^{n} x_jA(\vec{e}_j)=\\ =\sum_{j=1}^{n} x_j\left(\sum_{i=1}^{m} a_{ij}\vec{f}_i\right)=\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij}x_j\right)\vec{f}_i\end{aligned}$$



implica las siguientes igualdades:

$$y_i=\sum_{j=1}^{n} a_{ij}x_j, \quad i=1, 2, ..., m. \tag{1}$$

Con notación matricial, podemos escribir

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}=A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

No sólo toda aplicación lineal puede representarse mediante una matriz con respecto a dos bases dadas, sino que recíprocamente, *fijadas bases* $B=\{\vec{e}_1, ..., \vec{e}_n\}$ *y* $\bar{B}=\{\vec{f}_1, ..., \vec{f}_m\}$ *en los espacios inicial y final, respectivamente, y dada cualquier matriz de orden* $m\times n$

$$A=(a_{ij})_{\substack{i=1,...,m\\ j=1,...,n}}, \quad a_{ij}\in\mathbb{K}$$

*existe una única aplicación lineal que tiene a* $A$ *como matriz*: si $\vec{x}=\sum_{j=1}^{n} x_j\vec{e}_j$ definimos

$$A(\vec{x})=\sum_{i=1}^{m} y_i\vec{f}_i$$

donde $y_1, y_2, ..., y_m$ están dados por las relaciones que aparecen en (1).

*Nota.* Si los espacios vectoriales $V$ y $W$ coinciden y en ambos se toma la misma base $B$ para representar una aplicación $A$, su matriz se dice que está dada con respecto a la base $B$.

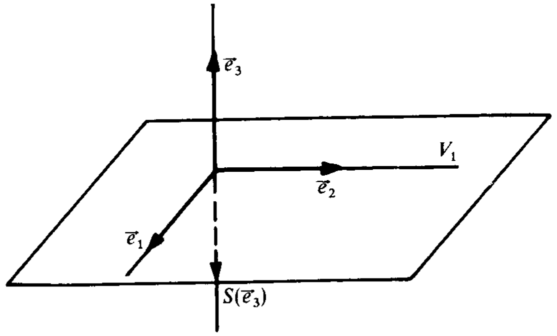
EJEMPLO A. Sea $R_\alpha$ la *rotación* de ángulo $\alpha$, en el plano, alrededor del origen en sentido positivo. $R_\alpha$ es una aplicación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ que, referida a la base canónica del plano, tiene como matriz

$$R_\alpha=\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

(Ver el ejemplo B de la sección 5.3.)

EJEMPLO B. En el espacio vectorial real $\mathbb{R}^3$ consideramos el subespacio $V_1=\{(x_1, x_2, 0)/x_1, x_2\in\mathbb{R}\}$. Si $B=\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es la base canónica de $\mathbb{R}^3$, y $S$ es la *simetría* con respecto al subespacio vectorial $V_1$, se tiene que

$$S(\vec{e}_1)=e_1, \quad S(\vec{e}_2)=\vec{e}_2, \quad S(\vec{e}_3)=-\vec{e}_3.$$



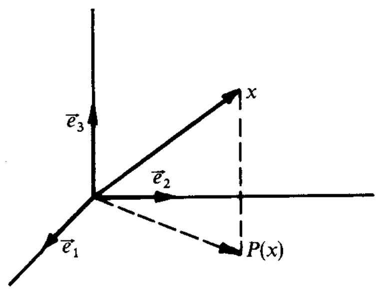

Por tanto, su matriz con respecto a la base canónica es

$$S=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Observar que las simetrías pueden definirse con respecto a cualquier subespacio vectorial de $\mathbb{R}^3$.

EJEMPLO C. Sea $P$ la *proyección ortogonal* de un vector de $\mathbb{R}^3$ sobre el plano $xOy$; $P$ es una aplicación lineal de $\mathbb{R}^3$ y puesto que

$$P(\vec{e}_1)=\vec{e}_1, \quad P(\vec{e}_2)=\vec{e}_2 \quad \text{y} \quad P(\vec{e}_3)=\vec{0}.$$




su matriz con respecto a la base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es

$$P=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

EJEMPLO D. Dados dos espacios vectoriales $V$ y $W$ la *aplicación nula* que envía todo vector de $V$ al elemento neutro de $W$ es una aplicación lineal, cuya matriz es la matriz nula con respecto a cualesquiera bases de $V$ y $W$.



En un mismo espacio vectorial $V$, la aplicación que lleva todo vector de $V$ en sí mismo se denomina *aplicación identidad*. Es fácil comprobar que, con respecto a cualquier base de $V$, la matriz de la aplicación identidad es

$$I=\begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & \ddots & \\ 0 & & & 1 \end{pmatrix}_n,$$

si el espacio vectorial tiene dimensión $n$.

EJEMPLO E. Sea $D$ la aplicación *derivación* del ejemplo C de la sección 6.1 y consideremos $D$ como una aplicación de $P_{\mathbb{R}}^{(n)}[x]$ en $P_{\mathbb{R}}^{(n-1)}[x]$. Sea $B=\{1, x, x^2, ..., x^n\}$ una base de $P_{\mathbb{R}}^{(n)}[x]$ y $\bar{B}=\{1, x, x^2, ..., x^{n-1}\}$ una base de $P^{(n-1)}[x]$. Tenemos

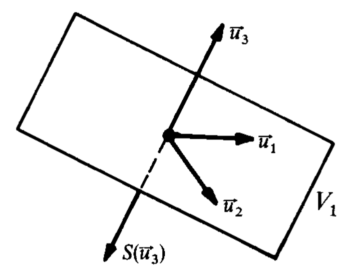
$$D(1)=0,\ D(x)=1,\ D(x^2)=2x,\ D(x^3)=3x^2,\ ...,\ D(x^n)=nx^{n-1}$$

Por tanto, la matriz de $D$ con respecto a las bases $B$ y $\bar{B}$ es la matriz de orden $n\times(n+1)$ dada por

$$D=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & n \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

EJEMPLO F. Tratamos de encontrar una base en $\mathbb{R}^3$ de manera que la matriz de la simetría con respecto al subespacio vectorial $V_1=\{(x, y, z)\in\mathbb{R}^3/x+y+z=0\}$ sea

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$



En el ejemplo B se ha resuelto el caso en que el subespacio vectorial coincida con el plano $xOy$. Un estudio detallado del ejemplo citado nos da la idea adecuada para resolver el problema planteado aquí: basta tomar dos vectores cualesquiera de $V_1$ que sean linealmente independientes y un tercero que sea perpendicular a $V_1$. Por ejemplo,

$$\vec{u}_1=(1,\,-1,\,0),\quad \vec{u}_2=(0,\,1,\,-1),\quad \vec{u}_3=(1,\,1,\,1)$$

* * *

Ya hemos mencionado anteriormente que si en los espacios vectoriales $V$ y $W$, de dimensión finita $n$ y $m$, respectivamente, se fijan bases, existe una correspondencia biunívoca entre las aplicaciones lineales de $V$ en $W$ y el conjunto de las matrices $\mathcal{M}_{m\times n}(\mathbb{K})$, de orden $m\times n$ sobre el cuerpo $\mathbb{K}$.

Puesto que el conjunto $\mathcal{M}_{m\times n}(\mathbb{K})$ posee una estructura de espacio vectorial, no es de extrañar que en el conjunto de todas las aplicaciones lineales entre dos espacios vectoriales puedan definirse dos operaciones que le den estructura de espacio vectorial.

Estas operaciones ya han sido definidas en el capítulo 1 cuando los espacios vectoriales son $\mathbb{R}^n$ y $\mathbb{R}^m$. Las operaciones que daremos aquí son una copia de aquellas.

Sean $V$ y $W$ dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$; el conjunto de todas las aplicaciones lineales de $V$ en $W$ se designará mediante el símbolo

$$L(V;\,W).$$

Si $A$ y $B$ son elementos de $L(V;\,W)$, definimos su suma mediante

$$(A+B)(\vec{v})=A(\vec{v})+B(\vec{v})\quad \text{para todo } \vec{v}\in V.$$

Si $A$ es un elemento de $L(V;\,W)$ y $c$ es un elemento de $\mathbb{K}$, definimos la *multiplicación de c por A* mediante

$$(cA)(\vec{v})=c(A(\vec{v}))\quad \text{para todo } \vec{v}\in V.$$

Las operaciones que acabamos de definir tienen ciertas propiedades que coinciden con las enumeradas para matrices en la sección 3 del capítulo 1. Una forma elegante y rápida de enumerar estas propiedades se da en el siguiente teorema:

TEOREMA 1

Sean $V$ y $W$ dos espacios vectoriales sobre un mismo cuerpo $\mathbb{K}$; el conjunto $L(V;\,W)$ de las aplicaciones lineales entre $V$ y $W$, con las operaciones anteriormente definidas, es un espacio vectorial sobre el cuerpo $\mathbb{K}$.

La demostración completa de este resultado se deja para el lector; nosotros observaremos que el elemento neutro de este espacio es la aplicación nula y el opuesto de una aplicación $A$ es la aplicación $-A$ definida por

$$(-A)(\vec{u})=-(A(\vec{u})).$$

Si $V$ y $W$ coinciden escribimos $L(V)$ en lugar de $L(V;\,V)$.



Puesto que toda aplicación lineal puede representarse mediante una matriz y recíprocamente, el siguiente resultado no debe extrañar al lector

TEOREMA 2

Sean $V$ y $W$ espacios vectoriales de dimensiones $n$ y $m$, respectivamente; entonces, el espacio vectorial $L(V,\,W)$ tiene dimensión $n\times m$.

*Demostración.* Hemos de exhibir $n\times m$ elementos de $L(V,\,W)$ que formen una base de este espacio. Fijadas

$$B=\{\vec{e}_1,\,\dots,\,\vec{e}_n\}\quad \text{y}\quad \bar{B}=\{\vec{f}_1,\,\dots,\vec{f}_m\}$$

bases de $V$ y $W$, respectivamente, definimos, para $j=1,\,\dots,\,n$, $i=1,\,\dots,\,m$,

$$E_{ji}(\vec{e}_k)=\begin{Bmatrix}\vec{f}_i & \text{si} & j=k\\ \vec{0} & \text{si} & j\neq k\end{Bmatrix},\quad k=1,\,2,\,\dots,\,n.$$

Esta definición se simplifica, en cuanto a su notación, si utilizamos un símbolo denominado *delta de Kronecker:*

$$\delta_{jk}=\begin{Bmatrix}1 & \text{si} & j=k\\ 0 & j\neq k & \end{Bmatrix},\quad \begin{matrix}k=1,\,2,\,\dots,\,n.\\ j=1,\,2,\,\dots,\,n.\end{matrix}$$

Entonces podemos escribir

$$E_{ji}(\vec{e}_k)=\delta_{jk}\vec{f}_i,\quad k=1,\,2,\,\dots,\,n.$$

Puesto que $E_{ji}$ ha sido definido para los elementos de una base de $V$, determina una única aplicación lineal de $V$ en $W$, que seguiremos denotando por $E_{ji}$; este resultado se ha demostrado en el teorema 5 de la sección anterior.

Para finalizar la demostración del teorema es necesario probar que

$$E=\{E_{ji}\}_{\substack{j=1,\dots,n\\ i=1,\dots,m}}$$

es una base de $L(V,\,W)$. Sea

$$\sum_{j=i}^{n}\sum_{i=1}^{m}a_{ji}E_{ji}=0$$

una combinación lineal nula de los elementos de $E$ (0 designa la aplicación nula de $V$ en $W$); para todo $k=1,\,2,\,\dots,\,n$

$$\sum_{j=i}^{n}\sum_{i=1}^{m}a_{ji}E_{ji}(\vec{e}_k)=\sum_{i=1}^{m}a_{ki}\vec{f}_i=\vec{0}\in W.$$

Puesto que $\bar{B}$ es una base de $W$ se deduce que $a_{ki}=0$, $i=1, 2, ..., m$. Esto prueba que $E$ es un conjunto de aplicaciones lineales linealmente independientes.

Falta probar que $E$ es un sistema de generadores de $L(V, W)$. Sea $A \in L(V, W)$ y sea $A=(a_{ij})$ su matriz con respecto a las bases $B$ y $\bar{B}$; para todo $\vec{e}_k \in V$ tenemos

$$A(\vec{e}_k)=\sum_{i=1}^{m} a_{ik}\vec{f}_i=\sum_{i=1}^{m} a_{ik}E_{ki}(\vec{e}_k)=\sum_{i=1}^{m} E_{ki}(a_{ik}\vec{e}_k).$$

Puesto que $E_{ji}(\vec{e}_k)=\vec{0}$ si $j\neq k$, podemos escribir

$$A(\vec{e}_k)=\sum_{i=1}^{m}\sum_{j=1}^{n} E_{ji}(a_{ij}\vec{e}_k)=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ji}(\vec{e}_k).$$

De aquí se deduce que

$$A=\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}E_{ji}$$

ya que ambas aplicaciones coinciden sobre los elementos de una base de $V$.

Esto termina la demostración del teorema 2. ■

*Nota.* Más adelante se dará otra demostración del teorema 2 estableciendo una correspondencia biyectiva entre $L(V, W)$ y $\mathcal{M}_{m\times n}(\mathbb{K})$ de manera que se conserve la dimensión. (Ver sección 6.4.)

* * *

Con la misma situación que al comienzo de esta sección, sean $A$ y $A'$ dos aplicaciones lineales de $V$ en $W$. Sean $M(A)$, $M(A')$, $M(A+A')$ y $M(cA)$ las matrices de las aplicaciones lineales $A$ y $A'$ con respecto a las bases $B$ y $\bar{B}$ de $V$ y $W$, respectivamente. En estas condiciones se tiene que

$$M(A+A')=M(A)+M(A')$$

y

$$M(cA)=cM(A), \quad c\in\mathbb{K}.$$

Estas igualdades expresan que la matriz de la aplicación lineal «suma» coincide con la suma de las matrices de cada una de las aplicaciones y que la matriz de la aplicación lineal $cA$ coincide con el producto de la matriz $A$ por el escalar $c$.

Las demostraciones de estos resultados son análogas a las realizadas en la sección 3 del capítulo 1 para aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$ y, por tanto, se dejan para el lector.

* * *

Con dos matrices se puede definir su producto siempre que concurran circunstancias favorables. El producto de matrices corresponde a la *composición de aplicaciones lineales*, que definimos a continuación.



Dadas $A\in L(V; W)$ y $B\in L(W; X)$ definimos la *composición de A y B* mediante

$$(B\circ A)(\vec{v})=B(A(\vec{v})).$$

Dadas las condiciones impuestas sobre $A$ y $B$ es fácil comprobar que $B\circ A\in L(V; X)$.

Si los espacios $V$, $W$ y $X$ tienen dimensión finita y si denotamos por $M(A)$, $M(B)$ y $M(B\circ A)$ las matrices de $A$, $B$ y $B\circ A$, respectivamente, con respecto a bases de antemano fijadas se tiene el siguiente resultado:

$$M(B\circ A)=M(B)\cdot M(A).$$

La demostración es análoga a la dada en la sección 1.3 para aplicaciones lineales entre espacios vectoriales de la forma $\mathbb{R}^n$. La repetimos aquí para conveniencia del lector.

Sean $\{\vec{e}_i\}_{i=1}^{n}$, $\{\vec{f}_j\}_{j=1}^{m}$, $\{\vec{g}_k\}_{k=1}^{p}$ bases de $V$, $W$ y $X$, respectivamente. La $i$-ésima columna de la matriz de $B\circ A$ son las componentes del vector $(B\circ A)(\vec{e}_i)$ con respecto a la base $\{\vec{g}_k\}_{k=1}^{p}$; por tanto, si escribimos

$$A=(a_{ji})_{i=1,j=i}^{m\ \ n}, \quad B=(b_{kj})_{k=1,j=1}^{p\ \ m}$$

se tiene

$$B\circ A(e_i)=B(A(\vec{e}_i))=B\left(\sum_{j=1}^{m} a_{ji}\vec{f}_j\right)=\sum_{j=1}^{m} a_{ji}B(\vec{f}_j)=$$
$$=\sum_{j=1}^{m} a_{ji}\left(\sum_{k=1}^{p} b_{kj}\vec{g}_k\right)=\sum_{k=1}^{p}\left(\sum_{j=1}^{m} b_{kj}a_{ji}\right)\vec{g}_k$$

Esto prueba que $\sum_{j=1}^{m} b_{kj}a_{ji}$ es el elemento que ocupa el lugar $(k, i)$ de la matriz $M(B\circ A)$; este valor coincide con el valor del elemento que ocupa el lugar $(k, i)$ en el producto de matrices $M(B)\cdot M(A)$.

## 6.3. CAMBIO DE BASE PARA APLICACIONES LINEALES

Sean $V$ y $W$ dos espacios vectoriales, sobre el mismo cuerpo, de dimensiones $n$ y $m$, respectivamente. Sea $A$ una aplicación lineal de $V$ en $W$ con matriz

$$A=(a_{ij})_{i=1,j=1}^{m\ \ n}$$

con respecto a las bases $B=\{\vec{e}_1, ..., \vec{e}_n\}$ y $\bar{B}=\{\vec{f}_1, ..., \vec{f}_m\}$ de $V$ y $W$, respectivamente.

Si deseamos conocer la matriz $A'=(a'_{ij})_{i=1,j=1}^{m\ \ n}$ de la misma aplicación $A$, con respecto a dos nuevas bases $B'=\{\vec{e}'_1, ..., \vec{e}'_n\}$ y $\bar{B}'=\{\vec{f}'_1, ..., \vec{f}'_m\}$ de $V$ y $W$, respectivamente, es necesario realizar los cambios de base adecuados en los espacios inicial y final.

Para que el lector siga con mayor facilidad el razonamiento, realizamos el siguiente diagrama:

$$\begin{tikzcd}
 & V & y=A(x) & W & \\
x_1\vec{e}_1+\cdots+x_n\vec{e}_n \arrow[d, phantom, "\|"] & B=\{\vec{e}_1, ..., \vec{e}_n\} \arrow[r, "A=(a_{ij})"] \arrow[dd, "C"] & & \bar{B}=\{\vec{f}_1, ..., \vec{f}_m\} \arrow[dd, "D"] & y_1\vec{f}_1+\cdots+y_m\vec{f}_m \arrow[d, phantom, "\|"] \\
x \arrow[d, phantom, "\|"] & & & & y \arrow[d, phantom, "\|"] \\
x'_1\vec{e}'_1+\cdots+x'_n\vec{e}'_n & B'=\{\vec{e}'_1, ..., \vec{e}'_n\} \arrow[r, "A'=(a'_{ij})"] & & \bar{B}'=\{\vec{f}'_1, ..., \vec{f}'_m\} & y'_1\vec{f}'_1+\cdots+y'_m\vec{f}'_m
\end{tikzcd}$$

Si escribimos $x \in V$ de la forma

$$x_1\vec{e}_1+\cdots+x_n\vec{e}_n=\vec{x}=x'_1e'_1+\cdots+x'_ne'_n$$

e $\vec{y}=A(\vec{x})$ de la forma

$$y_1\vec{f}_1+\cdots+y_m\vec{f}_m=\vec{y}=y'_1\vec{f}'_1+\cdots+y'_m\vec{f}'_m$$

los resultados de la sección anterior, junto con los relativos al cambio de base en un mismo espacio vectorial, que se estudiaron en el capítulo anterior, nos permiten escribir

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}=A\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}=C\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} \qquad \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix}=D\begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{pmatrix} \tag{2}$$

y

$$\begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{pmatrix}=A'\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix} \tag{3}$$

En las igualdades anteriores $C$ y $D$ son las matrices del cambio de base de $B$ a $B'$ y de $\bar{B}$ a $\bar{B}'$, respectivamente.



Sustituyendo (2) en (1) se obtiene

$$D\begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{pmatrix}=AC\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

Puesto que $D$ es una matriz de un cambio de base, posee inversa; por tanto:

$$\begin{pmatrix} y'_1 \\ y'_2 \\ \vdots \\ y'_m \end{pmatrix}=D^{-1}AC\begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$



Comparando esta expresión con (3) se obtiene

$$A'=D^{-1}AC \tag{4}$$

que nos permite calcular la matriz $A'$ de la aplicación $A$ con respecto a las bases $B'$ y $\bar{B}'$, conocida la matriz $A$ de la misma aplicación con respecto a las bases $B$ y $\bar{B}$ y las matrices $C$ y $D$ del cambio de base de $B$ a $B'$ y de $\bar{B}$ a $\bar{B}'$, respectivamente.

En muchos casos los espacios inicial y final de una aplicación coinciden. Si, además, $B$ y $\bar{B}$ coinciden y $B'$ y $\bar{B}'$ coinciden la fórmula del cambio de base es más sencilla.

Si $A$ es la matriz de la aplicación $A \in L(V)$ con respecto a una base $B$ de $V$, la matriz $A'$ de la misma aplicación con respecto a una base $B'$ de $V$ está dada por

$$A'=C^{-1}AC$$

donde $C$ es la matriz del cambio de base de $B$ a $B'$.

Observar que, en este caso, $|A'|=|A|$, ya que el determinante de un producto de matrices es el producto de los determinantes de cada una de ellas.

EJEMPLO A. Una aplicación lineal $A$ tiene, en una base $\{\vec{e}_1, \vec{e}_2\}$ de un espacio vectorial de dimensión 2, la matriz

$$A=\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}.$$

Queremos determinar la matriz de esta aplicación lineal en la base $\vec{e}'_1=\vec{e}_1+2\vec{e}_2$, $\vec{e}'_2=2\vec{e}_1+3\vec{e}_2$. Puesto que la matriz del cambio de base es

$$C=\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

la matriz $A'$ de esta aplicación en la base $\{\vec{e}_1', \vec{e}_2'\}$ es

$$A' = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}^{-1} \begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

EJEMPLO B. Sea $A: \mathbb{R}^3 \mapsto \mathbb{R}^3$ la aplicación dada por

$$A(x_1, x_2, x_3) = (x_1 + x_2, x_3, x_2 - x_3)$$

donde $(x_1, x_2, x_3)$ son las coordenadas de un punto $x \in \mathbb{R}^3$ con respecto a la base canónica en $\mathbb{R}^3$. Sea $B'$ la base $\{\vec{e}_3, \vec{e}_1, \vec{e}_2\}$ que se obtiene permutando los elementos de la base canónica.

Puesto que la matriz de $A$ en la base canónica es

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix},$$

y la matriz del cambio de base es

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

la matriz $A'$ de $A$ en la base $B'$ es

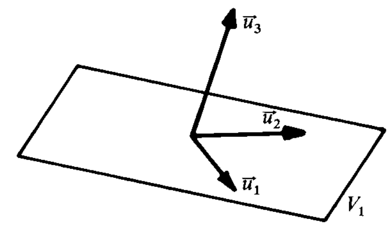
$$A' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} =$$

$$= \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

* * *

EJEMPLO C. Queremos encontrar las ecuaciones de la simetría $S$ con respecto al subespacio vectorial $V_1 = \{(x, y, z)\}/x - y + 2z = 0\}$, referida a la base canónica de $\mathbb{R}^3$. La estrategia consiste en encontrar en primer lugar una base de $\mathbb{R}^3$ en la cual la matriz de la simetría sea lo más sencilla posible y a continuación realizar un cambio de base para pasarla a la base canónica.






En la base $B' = \{\vec{u}_1 = (1, 1, 0), \vec{u}_2 = (0, 2, 1), \vec{u}_3 = (1, -1, 2)\}$ determinada por dos vectores de $V_1$ y un vector, $\vec{u}_3$, perpendicular a él, la matriz $S'$ de $S$ es

$$S' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

ya que $S(\vec{u}_1) = \vec{u}_1$, $S(\vec{u}_2) = \vec{u}_2$ y $S(\vec{u}_3) = -\vec{u}_3$. Puesto que la matriz del cambio de base de la base canónica a $B'$ es

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}$$

se tiene que

$$S' = C^{-1} S C \Leftrightarrow S = C S' C^{-1}.$$

Por tanto:

$$S = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 2 & -1 \\ 0 & 1 & 2 \end{pmatrix}^{-1} =$$

$$= \begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 0 & 1 & -2 \end{pmatrix} \left[ \frac{1}{6} \begin{pmatrix} 5 & 1 & -2 \\ -2 & 2 & 2 \\ 1 & -1 & 2 \end{pmatrix} \right] = \frac{1}{6} \begin{pmatrix} 4 & 2 & -4 \\ 2 & 4 & 4 \\ -4 & 4 & -2 \end{pmatrix}$$

Así pues,

$$S(x_1, x_2, x_3) = \left( \frac{2}{3}x_1 + \frac{1}{3}x_2 - \frac{2}{3}x_3,\ \frac{1}{3}x_1 + \frac{2}{3}x_2 + \frac{2}{3}x_3,\ -\frac{2}{3}x_1 + \frac{2}{3}x_2 - \frac{1}{3}x_3 \right)$$

donde $(x_1, x_2, x_3)$ son las coordenadas de un vector $\vec{x} \in \mathbb{R}^3$ con respecto a la base canónica.

EJERCICIOS (SECCIONES 6.1, 6.2 Y 6.3)

**1.** Estudiar cuáles de las siguientes aplicaciones son lineales entre los espacios vectoriales dados:

*a)* $M_B$: $\mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{M}_{2\times 1}(\mathbb{R})$ dada por $M_B(A) = AB$ con $B = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

*b)* $S_B$: $\mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{M}_{2\times 2}(\mathbb{R})$ dada por $S_B(A) = A + B$ con $B \in \mathcal{M}_{2\times 2}(\mathbb{R})$ fija.

*c)* $C_B$: $\mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{M}_{2\times 2}(\mathbb{R})$ dada por $C_B(A) = AB - BA$ con $B = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$.

*d)* $S$: $\mathcal{M}_{m\times n}(\mathbb{C}) \mapsto \mathcal{S}$ dada por $S(A) = \frac{1}{2}(A + A^t)$ donde $\mathcal{S} = \{A \in \mathcal{M}_{n\times n}(\mathbb{C}) / A = A^t\}$.

*e)* $R$: $\mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{S}_n(\mathbb{R})$ dada por $R(A) = AA^t$.

*f)* $A$: $P_{\mathbb{C}}^{(n)}[x] \mapsto P_{\mathbb{C}}^{(n)}[x]$ dada por $A(p(x)) = p(x+1)$.

*g)* $B$: $P_{\mathbb{C}}^{(n)}[x] \mapsto P_{\mathbb{C}}^{(n)}[x]$ dada por $A(p(x)) = p(x) + 1$.

**2.** Sea $A$: $\mathbb{R}^3 \mapsto \mathbb{R}^3$ dada por $A(x_1, x_2, x_3) = (x_1 + x_2, x_3, x_1 + x_3)$. Encontrar la matriz de $A$ con respecto a la base canónica. Hallar la imagen mediante $A$ de los siguientes subespacios vectoriales de $\mathbb{R}^3$.

*a)* $V_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 / x_1 + x_2 + x_3 = 0\}$.

*b)* $V_2 = \{(x_1, x_2, 0) / x_1, x_2 \in \mathbb{R}\}$.

*c)* $V_3 = \{(x_1, x_2, x_3) = t(1, -1, 1) / t \in \mathbb{R}\}$.

En cada caso indicar la dimensión del subespacio y la dimensión de su imagen mediante $A$.

**3.** Encontrar las matrices de las siguientes aplicaciones lineales con respecto a las bases canónicas de los espacios vectoriales dados:

*a)* $M_B$ y $C_B$ del ejercicio 1.

*b)* $A$: $P_{\mathbb{R}}^{(2)}[x] \mapsto \mathbb{R}^4 / A(p(x)) = (p(0), p(1), p(2), p(3))$.

*c)* $A$: $P_{\mathbb{C}}^{(3)}[x] \mapsto P_{\mathbb{C}}^{(3)}[x] / A(p(x)) = p(x+1)$.

**4.** Respecto de la base canónica en $\mathbb{R}^3$ hallar las matrices de las siguientes aplicaciones lineales:

*a)* Giro de $\alpha$ grados con respecto al eje $z$.

*b)* Simetría con respecto a la recta $x = 0$, $y = 0$.

*c)* Simetría con respecto a la recta $x = y$, $z = 0$.

*d)* Proyección sobre el plano $x - y + z = 0$.

*e)* Simetría con respecto a la recta $(x, y, z) = t(1, 1, 1)$.



**5.** La *traza* de una matriz cuadrada $A$ se define como la suma de los elementos de su diagonal principal y se designa por *traza*$(A)$. Demostrar que la aplicación $T$: $\mathcal{M}_{m\times n}(\mathbb{K}) \mapsto \mathbb{K}$ es lineal. Dar una base de $\mathcal{M}_{m\times n}(\mathbb{K})$ y encontrar la matriz de $T$ con respecto a esta base y a la base canónica de $\mathbb{K}$.

**6.** Dadas $A$: $\mathbb{R}^3 \mapsto \mathbb{R}^3$ mediante $A(x_1, x_2, x_3) = (x_2, x_3, 0)$ y $B$: $\mathbb{R}^3 \mapsto \mathbb{R}^2$ dada por $B(x_1, x_2, x_3) = (x_1 + x_3, x_2)$ calcular $A^n = A \circ \overset{n}{\cdots} \circ A$ para todo $n \in \mathbb{N}$ y $B \circ A$. [*Sugerencia*: encontrar las matrices de $A$ y $B$.]

**7.** Sabiendo que la aplicación $A$ lleva los vectores

$$\vec{u}_1 = (1, 0, 0), \quad \vec{u}_2 = (1, 1, 0) \quad \text{y} \quad \vec{u}_3 = (1, 1, 1)$$

de $\mathbb{R}^3$ en los vectores

$$\vec{w}_1 = (2, 1, 2), \quad \vec{w}_2 = (3, 1, 2) \quad \text{y} \quad \vec{w}_3 = (6, 2, 3)$$

respectivamente, encontrar la matriz de $A$ en las siguientes bases:

*a)* La base canónica de $\mathbb{R}^3$.

*b)* La base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$.

**8.** Encontrar las ecuaciones de las siguientes aplicaciones lineales realizando cambios de base adecuados:

*a)* Simetría con respecto a la recta $(x, y, z) = t(1, 1, 1)$.

*b)* Proyección sobre el plano $x - y + z = 0$.

*c)* Giro de 90° con respecto a la recta $x + y = 0$, $z = 0$.

**9.** Sea $T$: $P_{\mathbb{R}}^{(3)}[x] \mapsto P_{\mathbb{R}}^{(3)}[x]$ tal que $T(1) = x^2 + 1$, $T(x) = -x$, $T(x^2) = x^3$ y $T(x^3) = x^2 + x - 1$. Calcular $T(x^2 + 2x + 1)$ y $T((x-2)^2 + x^3)$. Encontrar la matriz de $T$ con respecto a la base $\{1, x, x^2, x^3\}$ de $P_{\mathbb{R}}^{(3)}[x]$.

**10.** Sea $A \in L(V; W)$; si $V_1$ es un subespacio vectorial de $V$, la aplicación $A$ puede considerarse como una aplicación lineal de $V_1$ en $W$; esta aplicación recibe el nombre de *restricción de $A$ a $V_1$*, y se designa con el símbolo $A|V_1$. En los siguientes casos encontrar una base del subespacio vectorial dado y la matriz de la aplicación dada restringida a este subespacio con respecto a esta base:

*a)* $A$: $\mathbb{R}^3 \mapsto \mathbb{R}^2 / A(x_1, x_2, x_3) = (x_1 + x_2, x_2 + x_3)$
$V_1 = \{(x_1, x_2, x_3) / x_1 - 2x_2 + 3x_3 = 0\}$

*b)* $A$: $\mathbb{C}^2 \mapsto \mathbb{C}^3 / A(z_1, z_2) = (z_1 + z_2, iz_2, iz_1)$
$V_1 = \{(z_1, z_2) / z_1 = iz_2\}$

**11.** Dada $A \in L(V)$ y $\lambda \in \mathbb{K}$, demostrar que el conjunto

$$E(\lambda) = \{\vec{x} \in V / A(\vec{x}) = \lambda \vec{x}\}$$

es un subespacio vectorial de $V$. En los siguientes casos encontrar una base de $E(\lambda)$:

*a)* $A$: $\mathbb{R}^3 \mapsto \mathbb{R}^3 / A(x_1, x_2, x_3) = (-2x_1, -x_2, -x_1 + x_3)$, $\lambda = -1$.

*b)* $A: \mathbb{R}^4 \mapsto \mathbb{R}^4$ tal que

$$A(x_1, x_2, x_2, x_4) = \begin{pmatrix} 1 & 0 & 2 & -6 \\ 0 & 1 & -1 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \quad ; \quad \lambda=2;\ \lambda=1;\ \lambda=0$$

*c)* $A: \mathbb{C}^3 \mapsto \mathbb{C}^3 / A(z_1, z_2, z_3) = (2z_1 + z_3, -z_2, -3z_1 - z_3)$; $\lambda = e^{\frac{\pi}{3} i}$.

**12.** Sea $\mathscr{S}_n(\mathbb{K})$ el espacio vectorial de las matrices simétricas de orden $n$ con elementos en el cuerpo $\mathbb{K}$. Definimos

$$T: \mathscr{M}_{n \times n}(\mathbb{K}) \mapsto \mathscr{S}_n(\mathbb{K})$$

mediante $T(A) = \frac{1}{2}(A + A^t)$.

*a)* Encontrar una base de $\mathscr{S}_n(\mathbb{K})$.

*b)* Encontrar la matriz de $T$ con respecto a la base canónica de $\mathscr{M}_{n \times n}(\mathbb{K})$ y a la base de $\mathscr{S}_n(\mathbb{K})$ encontrada en *a)*.

**13.** Sea $T$ una aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$; demostrar que $T$ transforma variedades $k$-dimensionales de $\mathbb{R}^n$ en variedades $l$-dimensionales de $\mathbb{R}^m$, con $l \leqq k$. En las siguientes aplicaciones lineales hallar la imagen de las variedades lineales dadas:

*a)* $T: \mathbb{R}^2 \mapsto \mathbb{R}^3$ dada por $T(x_1, x_2) = (x_1 + x_2, x_1, x_2)$ y

$$M = \{x_1 = x_2 / x_1, x_2 \in \mathbb{R}\}$$

*b)* $T: \mathbb{R}^3 \mapsto \mathbb{R}^3$ dada por la matriz

$$T = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

y $M = \{(x_1, x_2, x_3) / x_1 - 2x_2 + x_3 = 4\}$.

**14.** Sea $V$ el espacio vectorial de las funciones reales de variable real generado por sen $x$ y cos $x$. Calcular el determinante de la aplicación «derivación» definida en $V$.

## 6.4. APLICACIONES LINEALES INYECTIVAS Y SUPRAYECTIVAS. NUCLEO Y RANGO DE UNA APLICACION LINEAL

Sean $V$ y $W$ dos espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y $A$ una aplicación lineal de $V$ en $W$. Recordamos que $A$ es inyectiva si $A(\vec{x}) = A(\vec{y})$ implica $\vec{x} = \vec{y}$; la aplicación $A$ es suprayectiva si para todo $\vec{y} \in W$ existe $\vec{x} \in V$ tal que $A(\vec{x}) = \vec{y}$ o equivalentemente $A(V) = W$, donde $A(V)$ denota la imagen de $V$ mediante $A$; finalmente, recordamos que $A$ es *biyectiva* si es a la vez inyectiva y suprayectiva.



En el caso de aplicaciones *lineales* cada uno de los tipos anteriores recibe un nombre especial: una aplicación lineal *inyectiva* recibe el nombre de *monomorfismo;* si la aplicación lineal es *suprayectiva* se le da el nombre de *epimorfismo;* finalmente, si la aplicación lineal es *biyectiva* se dice que es un *isomorfismo.*

El objetivo de esta sección es encontrar condiciones sencillas que sirvan para determinar si una aplicación lineal es de cualquiera de los tipos anteriores.

Comenzamos con las aplicaciones lineales inyectivas. Para ello necesitamos definir el concepto de *núcleo* de una aplicación lineal; dada una aplicación lineal

$$A: V \mapsto W$$

definimos el *núcleo de A*, que se denota mediante *ker (A)*, como el conjunto de todos los $\vec{v} \in V$ tal que $A(\vec{v}) = \vec{0}$; es decir:

$$\ker(A) = \{\vec{v} \in V / A(\vec{v}) = \vec{0}\}.$$

*Nota.* En algunos textos se utiliza $N(A)$ en lugar de $\ker(A)$; la notación $\ker(A)$ es la que se utiliza en toda la literatura matemática inglesa; esta notación viene de la palabra inglesa *kernel,* que significa «núcleo».

El subconjunto $\ker(A)$ no es nunca vacío, ya que $\vec{0} \in \ker(A)$; esto se deduce de que $A(\vec{0}) = \vec{0}$, como se ha demostrado en la sección 6.1. El subconjunto $\ker(A)$ es, además, un subconjunto distinguido de $V$; se tiene el siguiente resultado:

**PROPOSICIÓN 1**

Si $A: V \mapsto W$ es una aplicación lineal entre espacios vectoriales, $\ker(A)$ es un subespacio vectorial de $V$.

*Demostración.* Si $\vec{x}$ e $\vec{y}$ son elementos de $\ker(A)$ se tiene que $A(\vec{x}) = \vec{0}$ y $A(\vec{y}) = \vec{0}$; por tanto:

$$A(\vec{x} + \vec{y}) = A(\vec{x}) + A(\vec{y}) = \vec{0} + \vec{0} = \vec{0}$$

de donde deducimos que $\vec{x} + \vec{y} \in \ker(A)$. Si $\vec{x} \in \ker(A)$ y $c \in \mathbb{K}$ se tiene que

$$A(c\vec{x}) = cA(\vec{x}) = c \cdot \vec{0} = \vec{0}$$

de donde se deduce que $c\vec{x} \in \ker(A)$. Estos dos resultados prueban la proposición. ■

**PROPOSICIÓN 2**

Una aplicación lineal $A: V \mapsto W$ es inyectiva si y sólo si $\ker(A) = \{\vec{0}\}$.

*Demostración.* Supongamos que $A$ es inyectiva; si $\vec{x} \in \ker(A)$ se tiene $A(\vec{x})=\vec{0}$; puesto que $A$ es lineal, $A(\vec{0})=\vec{0}$, y, por tanto, podemos escribir

$$A(\vec{x})=\vec{0}=A(\vec{0}).$$

Puesto que $A$ es inyectiva, $\vec{x}=\vec{0}$; esto prueba que

$$\ker(A)=\{\vec{0}\}$$

Supongamos ahora que $\ker(A)=\{\vec{0}\}$ y probemos que $A$ es inyectiva. Si $A(\vec{x})=A(\vec{y})$, se tiene que

$$A(\vec{x}-\vec{y})=A(\vec{x})-A(\vec{y})=\vec{0}$$

por tanto, $\vec{x}-\vec{y} \in \ker(A)$; como $\ker(A)=\{\vec{0}\}$ deducimos que $\vec{x}-\vec{y}=\vec{0}$, o equivalentemente, $\vec{x}=\vec{y}$, que era lo que deseábamos demostrar. ■

EJEMPLO A. Para estudiar si $A: \mathbb{R}^2 \mapsto \mathbb{R}^3$ dada por $A(x_1, x_2)=(2x_1+x_2, x_1+2x_2, 3x_1)$ es una aplicación lineal inyectiva hemos de encontrar $\ker(A)$. Si $(x_1, x_2) \in \ker(A)$ hemos de tener $A(x_1, x_2)=(0, 0, 0)$; esta igualdad se transforma en

$$\left.\begin{aligned} 2x_1+x_2&=0 \\ x_1+2x_2&=0 \\ 3x_1&=0 \end{aligned}\right\}$$

que es un sistema homogéneo de tres ecuaciones y dos incógnitas. Puesto que la matriz de sus coeficientes tiene rango 2, que coincide con el número de incógnitas, la única solución posible es la trivial $x_1=x_2=0$. Por tanto, $\ker(A)=\{(0, 0)\}$ y $A$ es inyectiva.

EJEMPLO B. Tratemos de encontrar el núcleo de la aplicación $T: \mathbb{R}^5 \mapsto \mathbb{R}^3$, cuya matriz es

$$T=\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & -1 & 1 \\ 1 & 0 & 1 & 2 & 0 \end{pmatrix}.$$

Se trata, por tanto, de encontrar todos los vectores $\vec{x}=(x_1, x_2, x_3, x_4, x_5)$ tal que

$$\left.\begin{aligned} x_1+x_2+x_3+x_4+x_5&=0 \\ x_2-x_4+x_5&=0 \\ x_1+x_3+2x_4&=0 \end{aligned}\right\}.$$



El rango de la matriz de los coeficientes de este sistema es 2 (observar que la tercera ecuación es combinación lineal de las dos primeras). Utilizando las dos primeras ecuaciones e introduciendo los parámetros $x_3=c_1$, $x_4=c_2$, $x_5=c_3$ tenemos que

$$x_2=c_2-c_3$$

y

$$x_1=-(c_2-c_3)-c_1-c_2-c_3=-c_1-2c_2$$

Así pues,

$$\begin{aligned}(x_1, x_2, x_3, x_4, x_5)&=(-c_1-2c_2, c_2-c_3, c_1, c_2, c_3)= \\ &=c_1(-1, 0, 1, 0, 0)+c_2(-2, 1, 0, 1, 0)+c_3(0, -1, 0, 0, 1)\end{aligned}$$

que son las ecuaciones paramétricas de $\ker(T)$. Observar que el núcleo de esta aplicación tiene dimensión 3.

* * *

Supongamos ahora que la aplicación lineal $A$ es suprayectiva, es decir, que $A$ es un epimorfismo; en estas condiciones se ha de cumplir que $A(V)=W$, donde $A(V)$ es la imagen de $V$ mediante la aplicación $A$.

El conjunto $A(V)$ recibe el nombre de *imagen de A,* y se designará, de ahora en adelante, mediante $\mathrm{img}(A)$. Por tanto, $A$ es suprayectiva si y sólo si

$$\mathrm{img}(A)=W$$

Por la proposición 3 de la sección 6.1 sabemos que la imagen de cualquier subespacio vectorial de $V$ es un subespacio vectorial de $W$; en particular, *img(A),* que es la imagen de $V$ mediante $A$, es un *subespacio vectorial de W.*

EJEMPLO C. Queremos encontrar la imagen de la aplicación $T$ del ejemplo B y decidir si es suprayectiva. La estrategia más sencilla es observar que la imagen de una base cualquiera de $\mathbb{R}^5$ es un sistema de generadores de $\mathrm{img}(T)$: si $B=\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4, \vec{e}_5\}$ es una base de $\mathbb{R}^5$ y $\vec{w} \in \mathrm{img}(T)$ existe $\vec{v}=\sum_{j=1}^{5} v_j\vec{e}_j \in \mathbb{R}^5$ tal que $T(\vec{v})=\vec{w}$; así pues,

$$\vec{w}=T\left(\sum_{j=1}^{5} v_j\vec{e}_j\right)=\sum_{j=1}^{5} v_jT(\vec{e}_j)$$

con lo cual queda probado el resultado. (Relacionado con este resultado, ver el problema 9 al final de esta sección.) Por tanto,

$$T(\vec{e}_1)=(1, 0, 1), \quad T(\vec{e}_2)=(1, 1, 0), \quad T(\vec{e}_3)=(1, 0, 1),$$
$$T(\vec{e}_4)=(1, -1, 2), \quad T(\vec{e}_5)=(1, 1, 0)$$

son un sistema de generadores de $\operatorname{img}(T)$; de este sistema de generadores es necesario extraer el mayor número de ellos linealmente independientes, y así podremos encontrar $\operatorname{img}(T)$. Puesto que la matriz que tiene a estos vectores por columnas es $T$ y ya sabemos que $T$ tiene rango 2, solamente es posible encontrar dos vectores linealmente independientes. Podemos tomar

$$\operatorname{img}(T)=L((1, 0, 1), (1, 1, 0)).$$

Debido a que $\operatorname{img}(T)$ tiene *dimensión 2*, llegamos a la conclusión de que $T$ no es suprayectiva.

Si queremos encontrar las ecuaciones cartesianas de $\operatorname{img}(T)$ hemos de eliminar los parámetros $c_1$ y $c_2$ de las ecuaciones

$$(y_1, y_2, y_3)=c_1(1, 0, 1)+c_2(1, 1, 0).$$

Así pues,

$$\left.\begin{aligned} c_1+c_2&=y_1\\ c_2&=y_2\\ c_1&=y_3 \end{aligned}\right\}$$

De aquí deducimos $y_2+y_3=y_1$ sin más que sustituir las dos últimas ecuaciones en la primera.

* * *

En los ejemplos $B$ y $C$ se observa la siguiente relación entre las dimensiones del núcleo de $T$ y de su imagen:

$$\dim(\ker(T))+\dim(\operatorname{img}(T))$$

coincide con la dimensión del espacio inicial de $T$.

Este resultado es cierto en un contexto más general; le enunciaremos y demostraremos a continuación. Después obtendremos algunas de sus consecuencias relacionadas con las aplicaciones lineales inyectivas y suprayectivas.

**TEOREMA 3**

Sean $V$ y $W$ dos espacios vectoriales de los cuales $V$ es de dimensión finita y sea $A\colon V\mapsto W$ una aplicación lineal. Entonces,

$$\dim(\ker(A))+\dim(\operatorname{img}(A))=\dim(V)$$

*Demostración.* Sea $k=\dim(\ker(A))$ y $n=\dim(V)$. Elegimos una base $\{\vec{v}_1, ..., \vec{v}_k\}$ de $\ker(A)$ y la completamos con vectores $\vec{v}_{k+1}, ..., \vec{v}_n$ hasta obtener una base de $V$. Es suficiente probar que $B=\{A(\vec{v}_{k+1}), ..., A(\vec{v}_n)\}$ es una base de $\operatorname{img}(A)$, ya que entonces

$$\dim(\ker(A))+\dim(\operatorname{img}(A))=k+(n-k)=n=\dim(V).$$



Probamos en primer lugar que $B$ es un sistema de generadores de $\operatorname{img}(A)$; si $\vec{w}\in\operatorname{img}(A)$, elegimos $\vec{v}\in V$ tal que $A(\vec{v})=\vec{w}$; puesto que

$$\vec{v}=\sum_{j=1}^{n} c_j\vec{v}_j$$

se tiene que

$$\vec{w}=A(\vec{v})=A\left(\sum_{j=1}^{n} c_j\vec{v}_j\right)=\sum_{j=1}^{n} c_jA(\vec{v}_j)$$

debido a que $\vec{v}_1, ..., \vec{v}_k$ son elementos de $\ker(A)$ se tiene que $A(\vec{v}_j)=\vec{0}$ si $1\leqq j\leqq k$; por tanto,

$$\vec{w}=\sum_{j=k+1}^{n} c_jA(\vec{v}_j)$$

y queda probado que $B$ es un sistema de generadores de $\operatorname{img}(A)$.

Finalmente, probamos que los elementos de $B$ son linealmente independientes. Supongamos que tenemos una relación de la forma

$$\sum_{j=k+1}^{n} c_jA(\vec{v}_j)=\vec{0}. \tag{1}$$

Utilizando la linealidad de $A$, la igualdad anterior se escribe de la forma

$$A\left(\sum_{j=k+1}^{n} c_j\vec{v}_j\right)=\vec{0}$$

Esto implica que el vector $\sum_{j=k+1}^{n} c_j\vec{v}_j\in\ker(A)$. Puesto que $\ker(A)$ está generado por los vectores $\vec{v}_1, ..., \vec{v}_k$ podemos escribir

$$c_{k+1}\vec{v}_{k+1}+\cdots+c_n\vec{v}_n=\sum_{j=k+1}^{n} c_j\vec{v}_j=c_1\vec{v}_1+\cdots+c_k\vec{v}_k.$$

Puesto que $\{\vec{v}_1, ..., \vec{v}_n\}$ es una base de $V$ la igualdad anterior solamente es posible si $c_1=\cdots=c_k=c_{k+1}=\cdots=c_n=0$. En particular, la igualdad (1) solamente es cierta cuando todos los $c_j$ son nulos. Se termina así la demostración del teorema. ■

Para deducir algunos corolarios de este teorema es necesario hacer uso del concepto de rango de una matriz estudiado en los capítulos 1 y 2.

Sea $A$ una aplicación lineal entre los espacios vectoriales $V$ y $W$, ambos de dimensión finita $n$ y $m$, respectivamente. Sea $A$ la matriz de la aplicación lineal $A$ en dos bases cualesquiera de $V$ y $W$; para encontrar el núcleo de $A$ es necesario resolver el sistema homogéneo

$$A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

donde $x_1, \ldots, x_n$ son las coordenadas de un vector de $V$ con respecto a la base elegida. Debido a la proposición 3 de la sección 1.2 (capítulo 1) este sistema posee $n - r(A)$ soluciones linealmente independientes que generan todas las restantes, donde $r(A)$ denota el rango de la matriz $A$.

Podemos, por tanto, concluir que

$$\dim(\ker(A)) = \dim(V) - r(A).$$

Comparando esta igualdad con el teorema 3 se deduce que

$$\dim(\operatorname{img}(A)) = r(A).$$

Puesto que $\operatorname{img}(A)$ no depende de las bases que se elijan en $V$ y $W$ para encontrar su matriz, de la igualdad anterior se deduce que todas las matrices de la aplicación $A$ tienen el mismo rango. Podemos, por tanto, definir el *rango de una aplicación lineal* $A$, que seguiremos escribiendo mediante $r(A)$, como el rango de una cualquiera de sus representaciones matriciales.

Una vez hechos estos comentarios el lector no tendrá ninguna dificultad en demostrar los siguientes resultados:

COROLARIO

Sean $V$ y $W$ espacios vectoriales de dimensión finita y $A: V \mapsto W$ una aplicación lineal. Entonces:

*a*) $A$ es inyectiva si y sólo si $r(A) = \dim(V)$.

*b*) $A$ es suprayectiva si y sólo si $r(A) = \dim(W)$.

EJEMPLO D. Supongamos que $V$ y $W$ son espacios vectoriales de dimensión finita y $A: V \mapsto W$ una aplicación lineal.

Si $A$ es *inyectiva*, $\ker(A) = \{\vec{0}\}$ y, por tanto, $\dim(V) = \dim(\operatorname{img}(A))$; como $\operatorname{img}(A)$ es un subespacio vectorial de $W$ concluimos que

$$\dim(V) \leqq \dim(W)$$

Si $A$ es *suprayectiva*, del teorema 3 se deduce que

$$\dim(V) \geqq \dim(\operatorname{img}(A)) = r(A) = \dim(W)$$



En particular, una aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^2$ *no* puede ser inyectiva y una aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^4$ *no* puede ser suprayectiva.

* * *

Para terminar esta sección estudiamos las aplicaciones lineales *biyectivas* o *isomorfismos* entre espacios vectoriales. Supongamos que $V$ y $W$ son espacios vectoriales de dimensión finita, y que $A$ es un isomorfismo entre ellos. Del corolario anterior deducimos que

$$\dim(V) = r(A) = \dim(W)$$

Otras consecuencias sencillas de algunos resultados anteriores se recogen en el siguiente teorema.

TEOREMA 4

Sean $V$ y $W$ espacios vectoriales de dimensión finita $n$ y sea $A: V \mapsto W$ una aplicación lineal. Las siguientes condiciones son equivalentes:

*a*) $A$ es biyectiva.

*b*) $A$ es inyectiva.

*c*) $\ker(A) = \{\vec{0}\}$.

*d*) $A$ es suprayectiva.

*e*) El rango de $A$ es $n$.

*Demostración.* Entre *b*), *c*), *d*) y *e*) se tienen las siguientes equivalencias:

$$\begin{array}{ccc} b) & \overset{\text{Prop. 2}}{\Leftrightarrow} & c) \\ \text{Cor.}\Updownarrow & & \\ e) & \underset{\text{Cor.}}{\Leftrightarrow} & d) \end{array}$$

Por tanto, todas ellas son equivalentes entre sí.

Finalmente, $a) \Rightarrow b)$, ya que toda aplicación biyectiva es inyectiva y puesto que *b*) y *d*) son equivalentes en este contexto y ambas implican *a*) se tiene que *a*) y *b*) son equivalentes. Esto termina la demostración. ■

Diremos que dos espacios vectoriales cualesquiera son *isomorfos* si podemos encontrar un isomorfismo entre ellos. Para que esto ocurra entre espacios vectoriales de dimensión finita ya sabemos que ambos han de ser de la misma dimensión. El recíproco también es cierto.

TEOREMA 5

Dado cualquier número natural $n$, todos los espacios vectoriales de dimensión $n$ sobre un mismo cuerpo son isomorfos.

*Demostración.* Sean $V$ y $W$ espacios vectoriales de dimensión $n$ sobre el mismo cuerpo y sean $B=\{\vec{v}_1, ..., \vec{v}_n\}$ y $\bar{B}=\{\vec{w}_1, ..., \vec{w}_n\}$ bases de $V$ y $W$, respectivamente.

Para definir el isomorfismo $A$ entre $V$ y $W$ basta definirlo sobre los elementos de la base $B$:

$$A(\vec{v}_j)=\vec{w}_j, \quad j=1, 2, ..., n$$

Por el teorema 5 de la sección 6.1 $A$ se extiende a una aplicación lineal de $V$ en $W$. Puesto que $r(A)=n$, ya que $\bar{B}$ es una base de $W$, del teorema anterior deducimos que $A$ es un isomorfismo. ■

EJEMPLO E. 1) Todos los espacios vectoriales *reales* de dimensión $n$ son isomorfos a $\mathbb{R}^n$ y todos los espacios vectoriales *complejos* de dimensión $n$ son isomorfos a $\mathbb{C}^n$. En general, todos los espacios vectoriales de dimensión $n$ sobre un cuerpo $\mathbb{K}$ son isomorfos a $\mathbb{K}^n$.

2) Los espacios vectoriales $\mathcal{M}_{2\times 2}(\mathbb{R})$, $P_{\mathbb{R}}^{(3)}[x]$ y $\mathbb{R}^4$ son isomorfos entre sí.

EJEMPLO F. Sean $V$ y $W$ espacios vectoriales de dimensión $n$ y $m$, respectivamente, sobre un mismo cuerpo $\mathbb{K}$. Deseamos probar que los espacios vectoriales

$$\mathcal{M}_{m\times n}(\mathbb{K}) \quad \text{y} \quad L(V; W)$$

son isomorfos. Puesto que $V$ es isomorfo a $\mathbb{K}^n$ y $W$ es isomorfo a $\mathbb{K}^m$ es suficiente con encontrar un isomorfismo entre

$$\mathcal{M}_{m\times n}(\mathbb{K}) \quad \text{y} \quad L(\mathbb{K}^n; \mathbb{K}^m)$$

(ver el problema 10 al final de esta sección). La elección más natural de este isomorfismo es la siguiente: dada $A\in\mathcal{M}_{m\times n}$ definimos

$$F(A)=A\begin{pmatrix} z_1\\ \vdots \\ z_n\end{pmatrix}\in\mathbb{K}^m$$

para todo $(z_1, z_2, ..., z_n)\in\mathbb{K}^n$. Observar que $F(A)$ es la aplicación que tiene como matriz $A$ en las bases canónicas de $\mathbb{K}^n$ y $\mathbb{K}^m$ y, por tanto, $F(A)\in L(\mathbb{K}^n; \mathbb{K}^m)$.

Es necesario probar que $F$ es un isomorfismo; como $F$ es claramente una aplicación lineal, hemos de probar que es biyectiva. Comenzamos encontrando su núcleo: si $F(A)=0$ se tiene que

$$A\begin{pmatrix} z_1\\ \vdots \\ z_n\end{pmatrix}=\begin{pmatrix} 0\\ \vdots \\ 0\end{pmatrix}\in\mathbb{K}^n$$

para todo $\vec{z}=(z_1, ..., z_n)\in\mathbb{K}^n$; eligiendo $\vec{z}=\vec{e}_j$, $j=1, 2, ..., n$, deducimos que $a_{ij}=0$ para todo $i=1, 2, ..., m$; por tanto, $A$ coincide con la matriz cero, que es por tanto el único elemento del núcleo de $F$.



Finalmente, hemos de probar que $F$ es suprayectiva; dada una aplicación lineal $A$ de $\mathbb{R}^m$, con matriz $M(A)$ con respecto a las bases canónicas, basta observar que

$$F(M(A))=A$$

Puesto que los espacios $\mathcal{M}_{m\times n}(\mathbb{K})$ y $L(V; W)$ son isomorfos y $\mathcal{M}_{m\times n}(\mathbb{K})$ tiene dimensión $m\times n$ deducimos que

$$\dim(L(V; W))=\dim(\mathcal{M}_{m\times n}(\mathbb{K}))=m\times n$$

Hemos dado otra demostración del teorema 2 de la sección 6.2.

* * *

*Observación.* Supongamos que $A$ es un isomorfismo entre dos espacios vectoriales $V$ y $W$ de dimensión $n$; debido al teorema 4 su rango es $n$ y, por tanto, la matriz $M(A)$ de $A$ en cualesquiera bases de $V$ y $W$ es *invertible.* La inversa de $M(A)$ es la matriz de la aplicación inversa de $A$.

## EJERCICIOS 6.4

**1.** Dadas las siguientes aplicaciones lineales encontrar las ecuaciones paramétricas del núcleo y la imagen comprobando en cada caso la ecuación

$$\dim(\ker)+\dim(\text{img})=\text{dimensión del espacio inicial}$$

e indicar si son inyectivas o suprayectivas:

*a)* $A$: $\mathbb{R}^3\mapsto\mathbb{R}^3/A(x_1, x_2, x_3)=(x_1+x_3, x_2, x_1+x_2+x_3)$.

*b)* $B$: $\mathbb{C}^2\mapsto\mathbb{C}^2/B(z_1, z_2)=(iz_1+z_2, z_1+iz_2)$.

*c)* $C$: $\mathbb{R}^2\mapsto\mathbb{R}^4/C(x_1, x_2)=(x_1, x_1+x_2, 2x_1, x_1+2x_2)$.

*d)* $D$: $\mathbb{C}^3\mapsto\mathbb{C}^3/D$ tiene como matriz

$$D=\begin{pmatrix} 1 & 1 & 0\\ 0 & 1 & 2\\ i & -1 & 0\end{pmatrix}$$

**2.** Describir el núcleo y la imagen de las siguientes aplicaciones lineales, indicando si son inyectivas, suprayectivas o biyectivas:

*a)* $M_B$: $\mathcal{M}_{2\times 2}(\mathbb{R})\mapsto\mathcal{M}_{2\times 1}(\mathbb{R})$ tal que $M_B(A)=AB$ con

$$B=\begin{pmatrix} 1\\ -1\end{pmatrix}.$$

*b)* $C_B$: $\mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{M}_{2\times 2}(\mathbb{R})$ tal que $C_B(A) = AB - BA$ con

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

*c)* $A$: $P_{\mathbb{R}}^{(3)}[x] \mapsto P_{\mathbb{R}}^{(3)}[x]$ tal que $A(1) = x^2 + 1$, $A(x) = x + 2$, $A(x^2) = x^3 - x$ y $A(x^3) = 1$.
*d)* La aplicación derivación de $P_{\mathbb{R}}^{(n)}[x]$ en $P_{\mathbb{R}}^{(n-1)}[x]$.

**3.** Describir el núcleo y la imagen de la aplicación lineal *traza* e indicar las dimensiones de cada uno de ellos. (La definición de la aplicación *traza* se ha dado en el problema 5 de la sección anterior.)

**4.** Dada $A \in L(V; W)$, demostrar los siguientes resultados:

*a)* $A^2 = A \circ A = 0$ si y sólo si $\mathrm{img}(A) \subset \ker(A)$.
*b)* $\ker(A) \subset \ker(A^2) \subset \ker(A^3) \subset \cdots \subset \ker(A^{n-1}) \subset \ker(A^n) \subset \cdots$
*c)* $\mathrm{img}(A) \supset \mathrm{img}(A^2) \supset \mathrm{img}(A^3) \supset \cdots \supset \mathrm{img}(A^{n-1}) \supset \mathrm{img}(A^n) \supset \cdots$

**5.** Construir aplicaciones que sean isomorfismos entre los siguientes pares de espacios vectoriales:

*a)* $\mathbb{C}^5$ y $\mathbb{R}^{10}$.
*b)* $P_{\mathbb{C}}^{(n)}[x]$ y $\mathbb{C}^{n+1}$.
*c)* $\mathcal{M}_{3\times 2}(\mathbb{R})$ y $\mathbb{R}^6$.

**6.** Sea $A$: $V \mapsto W$ una aplicación lineal y $V_1$ un subespacio vectorial de $V$; denotaremos por $A|_{V_1}$ la restricción de $A$ a $V_1$. Demostrar que:

*a)* $\ker(A|_{V_1}) = (\ker(A)) \cap V_1$.
*b)* $\mathrm{img}\, A = \mathrm{img}(A|_{V_1})$ si $V_1 + \ker(A) = V$.

**7.** Sean $A \in L(V; W)$ y $B \in L(W; X)$.

*a)* Demostrar que $\mathrm{img}(B \circ A) \subset \mathrm{img}(B)$ y que $\ker(B \circ A) \supset \ker(A)$.
*b)* Demostrar que $r(B \circ A) \leqq r(A)$ y $r(B \circ A) \leqq r(B)$.

**8.** Sea $V$ un espacio vectorial de dimensión $n$ y $W$ un espacio vectorial arbitrario. Dados dos subespacios vectoriales $V_1$ de $V$ y $W_1$ de $W$ tal que

$$\dim V_1 + \dim W_1 = n$$

demostrar que existe una aplicación lineal $A$: $V \mapsto W$ tal que $\ker(A) = V_1$ e $\mathrm{img}(A) = W_1$.

**9.** Sea $A$ un isomorfismo de los espacios vectoriales $V$ y $W$, ambos de dimensión finita. Demostrar que toda base de $V$ se transforma en una base de $W$.

**10.** Sean $V$ y $W$ espacios vectoriales de dimensiones $n$ y $m$, respectivamente. Demostrar que $L(V; W)$ es isomorfo a $L(\mathbb{K}^n; \mathbb{K}^m)$.

En los siguientes problemas $V$ y $W$ son espacios vectoriales de dimensión finita.



**11.** Sea $A \in L(V; W)$ una aplicación lineal inyectiva; demostrar que toda base de $V$ se transforma mediante $A$ en un conjunto de vectores linealmente independientes. Deducir de este resultado que si existe una aplicación lineal inyectiva entre dos espacios vectoriales $V$ y $W$ de dimensión finita, se ha de tener $\dim(V) \leqq \dim(W)$ (ver ejemplo D).

**12.** Sea $A' \in L(V; W)$ una aplicación lineal suprayectiva; demostrar que toda base de $V$ se transforma mediante $A'$ en un sistema de generadores de $W$. Deducir de este resultado que si existe una aplicación lineal suprayectiva entre dos espacios vectoriales $V$ y $W$ de dimensión finita, se ha de tener $\dim(V) \geqq \dim(W)$ (ver ejemplo D).

**13.** Dar un ejemplo de espacios vectoriales $V$ y $W$ y aplicaciones lineales $A$ y $A' \in L(V; W)$ y una base $B$ de $V$ tal que:

*a)* $A$ sea inyectiva y $A(B)$ no sea base de $W$.
*b)* $A'$ sea suprayectiva y $A'(B)$ no sea base de $W$. (Ver los problemas 11 y 12.)

**14.** Sea $A$ un isomorfismo de $V$ en $W$. *a)* Demostrar que toda base de $V$ se transforma en una base de $W$ mediante $A$. *b)* Demostrar que toda base de $W$ se transforma en una base de $V$ mediante $A^{-1}$. Deducir de *a)* o de *b)* que si $V$ y $W$ son isomorfos, $\dim V = \dim W$.

## 6.5. ESPACIO DUAL DE UN ESPACIO VECTORIAL

Dado un espacio vectorial $V$ sobre un cuerpo $\mathbb{K}$, podemos considerar el conjunto $L(V: \mathbb{K})$ de todas las aplicaciones lineales de $V$ en el espacio vectorial $\mathbb{K}$. Obsérvese que $\mathbb{K}$ es un espacio vectorial de dimensión 1 sobre $\mathbb{K}$.

Este espacio ha sido estudiado con mayor generalidad en la sección 6.2. En particular, el teorema 1 de la sección citada nos permite concluir que $L(V; \mathbb{K})$ es un espacio vectorial sobre $\mathbb{K}$. Este espacio vectorial recibe el nombre de *espacio dual* del espacio vectorial $V$ y se utiliza comúnmente el símbolo

$$V^*,$$

en lugar de $L(V; \mathbb{K})$, para indicarlo. Por tanto, $V^*$ es el espacio vectorial de todas las aplicaciones lineales de $V$ en $\mathbb{K}$.

Los elementos de $V^*$ son aplicaciones lineales y, por tanto, se indicarán con letras mayúsculas. Algunos autores prefieren utilizar letras minúsculas seguidas de una * en la parte superior derecha, tal como $v^*$, $e^*$, etc. La primera de estas notaciones será la que se utilice en el teorema que se enuncia en esta sección.

Si $V$ es un espacio vectorial de dimensión finita $n$, del teorema 2 de la sección 6.2 se deduce que el espacio dual $V^*$ tiene dimensión $n$. El siguiente teorema exhibe una base $B^*$ asociada de manera única y natural a una base fijada $B$ de $V$. Obsérvese que la demostración de un teorema más general, que incluye a éste, se ha dado en la sección

6.2 (teorema 2): incluimos la demostración porque este caso particular puede servir para aclarar el caso más general anteriormente descrito.

TEOREMA 1

Sea $V$ un espacio vectorial de dimensión finita $n$ y $B=\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ una base de $V$. Existe una única base

$$B^*=\{E_1, E_2, ..., E_n\}$$

de $V^*$ tal que $E_i(\vec{e}_i)=1$ para todo $i=1, 2, ..., n$, y $E_j(\vec{e}_i)=0$ si $i\neq j$. La base $B^*$ se denomina *base dual de B*.

*Nota.* La propiedad de $B^*$ enunciada en el teorema se escribe más fácilmente utilizando el simbolismo denominado de la *delta* ($\delta$) *de Kronecker* que se ha introducido en la demostración del teorema 2 de la sección 6.2. Recordemos que $\delta_{ji}$, $j$, $i\in\mathbb{N}$, es un número definido como sigue:

$$\delta_{ji}=\begin{cases}1 & \text{si} \quad i=j\\ 0 & \text{si} \quad i\neq j.\end{cases}$$

Con esta notación los elementos de la base dual de $B$ satisfacen

$$E_j(\vec{e}_i)=\delta_{ji}, \quad j, i=1, 2, ..., n.$$

*Demostración.* Para $j=1, 2, ..., n$ definimos $E_j$ de la siguiente manera: dado $\vec{v}=\sum_{i=1}^{n} v_j\vec{e}_j\in V$ tomamos

$$E_j(\vec{v})=v_j$$

es decir, la $j$-ésima coordenada de $\vec{v}$ en la base $B$.

Es fácil comprobar que $E_j\in V^*$ para todo $j=1, ..., n$, y que $E_j$ satisface la propiedad requerida en el teorema.

Unicamente falta demostrar que $B^*=\{E_1, E_2, ..., E_n\}$ es una base de $V^*$.

Si tenemos una igualdad de la forma

$$\sum_{j=1}^{n} c_jE_j \equiv 0$$

donde 0 representa la aplicación lineal nula, se tiene que

$$\sum_{j=1}^{n} c_jE_j(\vec{e}_i)=\vec{0}$$



para todo $i=1, 2, ..., n$. Por tanto, $c_i=0$, $i=1, 2, ..., n$, ya que $E_j(\vec{e}_i)$ es no nulo solamente si $j=i$. Esto prueba que $B^*$ es un conjunto de aplicaciones lineales linealmente independientes.

Finalmente, dada $A\in V^*$ se tiene que

$$A=\sum_{j=1}^{n} A(\vec{e}_j)E_j$$

ya que si $\vec{v}=\sum_{j=1}^{n} \vec{v}_j\vec{e}_j$ tenemos que

$$\sum_{j=1}^{n} A(\vec{e}_j)E_j(\vec{v})=\sum_{j=1}^{n} A(\vec{e}_j)v_j=\sum_{j=1}^{n} A(v_j\vec{e}_j)=A\left(\sum_{j=1}^{n} v_j\vec{e}_j\right)=A(\vec{v}).$$

Por tanto, $B^*$ es un sistema de generadores de $V^*$ y queda totalmente demostrado el teorema. ■

EJEMPLO A. Tratemos de encontrar la base dual de la base canónica $B=\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ de $\mathbb{R}^3$. La aplicación lineal $E_1$ satisface

$$E_1(\vec{e}_1)=1, \quad E_1(\vec{e}_2)=0, \quad E_1(\vec{e}_3)=0$$

Por tanto, $E_1(x_1, x_2, x_3)=E_1(x_1\vec{e}_1+x_2\vec{e}_2+x_3\vec{e}_3)=x_1$.

De manera similar se concluye que

$$E_2(x_1, x_2, x_3)=x_2 \quad \text{y} \quad E_3(x_1, x_2, x_3)=x_3$$

EJEMPLO B. Sea $B=\{\vec{u}_1=(1, 1, 0), \vec{u}_2=(0, 1, 1), \vec{u}_3=(1, 0, 1)\}$ una base de $\mathbb{R}^3$; queremos encontrar la base dual, $B^*$, de $B$. Denotemos por $U_1$, $U_2$, $U_3$ los elementos de $B^*$. Por definición de $B^*$ tenemos que

$$U_1(\vec{u}_1)=1, \quad U_1(\vec{u}_2)=0, \quad U_1(\vec{u}_3)=0.$$

Por tanto,

$$U_1(\vec{e}_1+\vec{e}_2)=1, \quad U_1(\vec{e}_2+\vec{e}_3)=0, \quad U_1(\vec{e}_1+\vec{e}_3)=0.$$

Debido a la linealidad de $U_1$ tenemos que

$$\left.\begin{array}{l}U_1(\vec{e}_1)+U_1(\vec{e}_2)=1\\ U_1(\vec{e}_2)+U_1(\vec{e}_3)=0\\ U_1(\vec{e}_1)+U_1(\vec{e}_3)=0\end{array}\right\}$$

Resolviendo este sistema se obtiene

$$U_1(\vec{e}_1)=U_1(\vec{e}_2)=\frac{1}{2}, \quad U_1(\vec{e}_3)=-\frac{1}{2}.$$

Por tanto,

$$U_1(x_1, x_2, x_3)=\frac{1}{2}x_1+\frac{1}{2}x_2-\frac{1}{2}x_3.$$

De manera similar se obtienen los otros dos elementos de la base $B^*$:

$$U_2(x_1, x_2, x_3)=-\frac{1}{2}x_1+\frac{1}{2}x_2+\frac{1}{2}x_3$$

$$U_3(x_1, x_2, x_3)=\frac{1}{2}x_1-\frac{1}{2}x_2+\frac{1}{2}x_3$$

* * *

EJERCICIOS 6.5

**1.** Encontrar la base dual de $B=\{\vec{u}_1=(1, 2, 1), \vec{u}_2=(0, 1, 1), \vec{u}_3=(1, 1, 1)\}$ en $\mathbb{R}^3$.

**2.** Encontrar la base dual de $B=\{1, x+1, x^2-2, x^3-x^2\}$ en el espacio vectorial $P_{\mathbb{R}}^{(3)}[x]$ de todos los polinomios de grado no superior a 3.

## CAPITULO 7

# VALORES Y VECTORES PROPIOS. FORMA DE JORDAN



## 7.1. INTRODUCCION

En el ejemplo A de la sección 6.3 se demostró que si la aplicación lineal $A$ tiene como matriz

$$A=\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$$

con respecto a una base $\{\vec{e}_1, \vec{e}_2\}$, la misma aplicación $A$ con respecto a la base $\vec{e}_1=\vec{z}_1+2\vec{z}_2$, $\vec{e}_2'=2\vec{z}_1+3\vec{z}_2$ tiene como matriz

$$A'=\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Tenemos, pues, que $A'=C^{-1}AC$ o equivalentemente $A=CA'C^{-1}$, donde $C$ es la matriz del cambio de base.

La matriz de la aplicación lineal $A$ en la base $\{\vec{e}'_1, \vec{e}'_2\}$ es mucho más sencilla que correspondiente en la base $\{\vec{e}_1, \vec{e}_2\}$. La matriz $A'$ es diagonal. A la vista de este ejemplo podemos preguntarnos si, dada una aplicación lineal $A$, siempre puede encontrarse un cambio de base de manera que la matriz de $A$ con respecto a la nueva base sea diagonal.

Un ejemplo muy sencillo nos hace perder la esperanza de resolver afirmativamente esta pregunta: la matriz

$$A=\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$$

no puede diagonalizarse.

En efecto, supongamos que existe una matriz $C=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ de cambio de base, con determinante $|C|$, tal que

$$C^{-1}AC=A'=\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix}.$$

Tenemos, pues, que

$$A=C\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix}C^{-1}=\begin{pmatrix}a & b\\ c & d\end{pmatrix}\begin{pmatrix}\alpha & 0\\ 0 & \beta\end{pmatrix}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}\cdot\frac{1}{|C|}=$$

$$=\begin{pmatrix}a\alpha & b\beta\\ c\alpha & d\beta\end{pmatrix}\frac{1}{|C|}\begin{pmatrix}d & -b\\ -c & a\end{pmatrix}=\frac{1}{|C|}\begin{pmatrix}ad\alpha-bc\beta & -ab\alpha+ab\beta\\ cd\alpha-cd\beta & -bc\alpha+ad\beta\end{pmatrix}.$$

Comparando la primera y la última de estas matrices se tiene:

$$1=\frac{1}{|C|}(ad\alpha-bc\beta) \qquad (1)$$

$$0=\frac{1}{|C|}cd(\alpha-\beta) \qquad (2)$$

$$1=\frac{1}{|C|}ab(-\alpha+\beta) \qquad (3)$$

$$1=\frac{1}{|C|}(-bc\alpha+ad\beta) \qquad (4)$$

De (2) se deduce que o bien $c=0$, o bien $d=0$, o bien $\alpha=\beta$. Si $c=0$, de (1) y (4) se deduce que $\alpha=|C|/ad=\beta$, y sustituyendo esto en (3) se obtiene la contradicción $1=0$. Si $d=0$ se obtiene la misma contradicción mediante un razonamiento semejante. Finalmente, si $\alpha=\beta$, (3) es claramente una contradicción.



A pesar de que no todas las matrices son diagonalizables, el objetivo de este capítulo es encontrar la forma «más sencilla» en que una matriz dada puede transformarse mediante un cambio de base; esta forma «más sencilla» se denominará *matriz de Jordan* de la matriz dada. (El nombre se debe al matemático Camille Jordan —Lión, 1838-Milán, 1922—, que ayudó a comprender la clasificación de las aplicaciones lineales.)

No sólo puede utilizarse la forma de Jordan de una matriz para clasificar las aplicaciones lineales en un espacio vectorial, sino que podemos utilizarla para realizar operaciones con matrices. Como ilustración tratemos de encontrar la sexta potencia de la matriz

$$A=\begin{pmatrix}6 & -2\\ 6 & -1\end{pmatrix}$$

Debido al ejemplo A de la sección 6.3 podemos escribir:

$$A=C\begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix}C^{-1} \quad \text{con} \quad C=\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix}$$

Entonces:

$$A^6=\left[C\begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix}C^{-1}\right]^6=C\begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix}^6C^{-1}=C\begin{pmatrix}2^6 & 0\\ 0 & 3^6\end{pmatrix}C^{-1}=$$

$$=\begin{pmatrix}1 & 2\\ 2 & 3\end{pmatrix}\begin{pmatrix}2^6 & 0\\ 0 & 3^6\end{pmatrix}\begin{pmatrix}-3 & 2\\ 2 & -1\end{pmatrix}=\begin{pmatrix}2^6 & 2\cdot 3^6\\ 2^7 & 3^7\end{pmatrix}\begin{pmatrix}-3 & 2\\ 2 & -1\end{pmatrix}$$

$$=\begin{pmatrix}-3\cdot 2^6+4\cdot 3^6 & 2^7-2\cdot 3^6\\ -3\cdot 2^7+2\cdot 3^7 & 2^8-3^7\end{pmatrix}$$

## 7.2. SUBESPACIOS INVARIANTES. VALORES Y VECTORES PROPIOS DE UNA APLICACION LINEAL

Dado un espacio vectorial $V$ y una aplicación lineal $A\colon V\mapsto V$, es decir, $A\in L(V)$, un subespacio vectorial $W$ de $V$ se llama *invariante* respecto a $A$ si $A(W)\subset W$, es decir, la imagen $A(\vec{x})$ de todo vector $\vec{x}\in W$ es un elemento de $W$.

EJEMPLO A. Sea $A\in L(\mathbb{R}^2)$ una aplicación lineal en $\mathbb{R}^2$ cuya matriz respecto de una base $\{\vec{e}_1, \vec{e}_2\}$ de $\mathbb{R}^2$ está dada por

$$A=\begin{pmatrix}2 & 0\\ 0 & 1\end{pmatrix}$$

Entonces, $W_1=\{x_1\vec{e}_1/x_1\in\mathbb{R}\}$ y $W_2=\{x_2\vec{e}_2/x_2\in\mathbb{R}\}$ son invariantes respecto de $A$. En efecto,

$$A(x_1\vec{e}_1)=x_1A(\vec{e}_1)=x_1(2\vec{e}_1)=(2x_1)\vec{e}_1\in W_1$$

y

$$A(x_2\vec{e}_2)=x_2A(\vec{e}_2)=x_2\vec{e}_2\in W_2.$$

EJEMPLO B. Sea $R_\alpha$ una rotación de ángulo $\alpha\neq 0$ en $\mathbb{R}^3$ con respecto al eje $Oz$. Geométricamente se observa que el plano $xOy$ y la recta $Oz$ son invariantes con respecto a esta aplicación. Para comprobar algebraicamente que el plano $xOy$ es invariante observar que la matriz de $R_\alpha$ con respecto a la base canónica de $\mathbb{R}^3$ es

$$R=\begin{pmatrix}\cos\alpha & -\operatorname{sen}\alpha & 0\\ \operatorname{sen}\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix}.$$

Si $\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2$ es un elemento del plano $xOy$ se tiene que

$$\begin{pmatrix}\cos\alpha & -\operatorname{sen}\alpha & 0\\ \operatorname{sen}\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ 0\end{pmatrix}=\begin{pmatrix}x_1\cos\alpha-x_2\operatorname{sen}\alpha\\ x_1\operatorname{sen}\alpha+x_2\cos\alpha\\ 0\end{pmatrix}$$

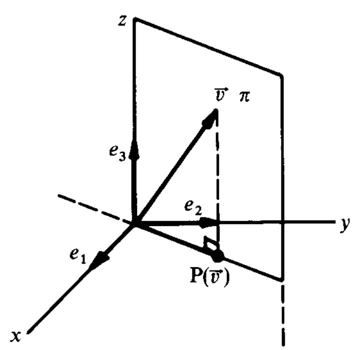
y, por tanto,

$$R_\alpha(\vec{x})=(x_1\cos\alpha-x_2\operatorname{sen}\alpha)\vec{e}_1+(x_1\operatorname{sen}\alpha+x_2\cos\alpha)\vec{e}_2$$

que es de nuevo un elemento del plano $xOy$.

EJEMPLO C. Sea $P$ la proyección ortogonal de $\mathbb{R}^3$ sobre el plano $xOy$; todo plano $\pi$ que contiene al eje $Oz$ es invariante. En efecto, como la matriz de $P$ es

$$P=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}$$






con respecto a la base canónica, y los elementos de $\pi$ son de la forma $\vec{v}=x_1\vec{e}_1+\lambda x_1\vec{e}_2+x_3\vec{e}_3$, $\lambda\in\mathbb{R}$, tenemos que

$$P(\vec{v})=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ \lambda x_1\\ x_3\end{pmatrix}=\begin{pmatrix}x_1\\ \lambda x_1\\ 0\end{pmatrix}$$

El elemento $x_1\vec{e}_1+\lambda x_1\vec{e}_2$ es de nuevo un elemento del plano $\pi$. Otros subespacios invariantes de esta proyección ortogonal son el plano $xOy$, el eje $Oz$ y cualquier recta del plano $xOy$ que pase por el origen de coordenadas.

* * *

Para cualquier aplicación lineal $A\in L(V)$ el subespacio $E=\{\vec{0}\}$, formado sólo por el elemento nulo, es invariante ya que $A(\vec{0})=\vec{0}$ y el propio espacio vectorial $V$ es también invariante.

PROPOSICIÓN 1

La intersección y la suma de subespacios invariantes respecto de una aplicación lineal $A\in L(V)$ son subespacios invariantes respecto de $A$.

*Demostración.* Sean $W_1$, $W_2$ subespacios vectoriales de $V$ que son invariantes respecto de $A$; si $\vec{x}\in W_1\cap W_2$ se tiene que $\vec{x}\in W_1$ y $\vec{x}\in W_2$; como $W_1$ y $W_2$ son invariantes tenemos que $A(\vec{x})\in W_1$ y $A(\vec{x})\in W_2$; así pues, $A(\vec{x})\in W_1\cap W_2$. Esto prueba que $W_1\cap W_2$ es invariante.

Sea ahora $\vec{x}=\vec{x}_1+\vec{x}_2\in W_1+W_2$; como $A$ es lineal y $W_1$, $W_2$ son invariantes respecto de $A$, se tiene que

$$A(\vec{x})=A(\vec{x}_1+\vec{x}_2)=A(\vec{x}_1)+A(\vec{x}_2)\in W_1+W_2$$

con lo que se demuestra que $W_1+W_2$ es también invariante. El razonamiento es similar si se consideran más de dos subespacios vectoriales. ■

Si $W$ es un subespacio vectorial de $V$, de dimensión 1, y es invariante respecto de una aplicación lineal $A\in L(V)$ tenemos que si $\vec{x}\in W$, no nulo, $A(\vec{x})=\lambda\vec{x}$; si $\vec{y}$ es otro vector de $W$, $\vec{y}=\alpha\vec{x}$ con $\alpha\in\mathbb{K}$; así pues,

$$A(\vec{y})=A(\alpha\vec{x})=\alpha A(\vec{x})=\alpha(\lambda\vec{x})=\lambda(\alpha\vec{x})=\lambda\vec{y}$$

y, por tanto, satisface la misma ecuación que $\vec{x}$.

Esto nos conduce a la siguiente definición:

DEFINICIÓN 2 *(Valores y vectores propios)*

Un vector $\vec{x}\neq\vec{0}$ de un espacio vectorial $V$ sobre $\mathbb{K}$ se llama *vector propio* de una aplicación lineal $A\in L(V)$ si existe un escalar $\lambda\in\mathbb{K}$ tal que $A(x)=\lambda x$; este número $\lambda$ se denomina *valor propio* de la aplicación $A$ correspondiente al vector $\vec{x}$.

*Nota.* Otra nomenclatura que se utiliza para vectores propios y valores propios es la de *autovectores* y *autovalores*, respectivamente.

Observar que, por el razonamiento realizado antes de la definición 2, si $\vec{x}$ es un vector propio de $A$ con autovalor $\lambda$, todo elemento no nulo del subespacio unodimensional generado por $\vec{x}$ es un autovector de $A$ con el mismo autovalor $\lambda$.

Supongamos que una aplicación lineal $A$ en un espacio $V$ de dimensión $n$ tiene $n$ vectores propios linealmente independientes $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ con valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$, respectivamente; tomando $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ como una base de $V$ se tiene que

$$A(\vec{e}_1)=\lambda_1\vec{e}_1,\ A(\vec{e}_2)=\lambda_2\vec{e}_2,\ \ldots,\ A(\vec{e}_n)=\lambda_n\vec{e}_n$$

y, por tanto, la matriz de $A$ con respecto a esta base es la matriz diagonal

$$A=\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

Recíprocamente, toda aplicación lineal que tiene una matriz diagonal en una cierta base tiene a los elementos de esta base como vectores propios. Si decimos que una aplicación lineal $A \in L(V)$ es *diagonizable* si existe una base de $V$ en la cual la matriz de $A$ es diagonal, hemos probado el siguiente resultado:

PROPOSICIÓN 3

Una aplicación lineal $A \in L(V)$ es diagonizable si y sólo si existe una base de $V$ formada por vectores propios.

DEFINICIÓN 3

Una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ se dice diagonalizable en $\mathbb{K}$ si la aplicación lineal $A\colon \mathbb{K}^n \mapsto \mathbb{K}^n$ que la tiene como matriz es diagonalizable.

De esta definición se deduce que $A \in \mathcal{M}_{n\times n}(\mathbb{K})$ es diagonalizable en $\mathbb{K}$ si existe una matriz $C \in \mathcal{M}_{n\times n}(\mathbb{K})$, con determinante no nulo, tal que $A'=C^{-1}AC$ es una matriz diagonal.

EJEMPLO D. Demostrar que $\vec{x}=\vec{e}_1+2\vec{e}_2$ e $\vec{y}=2\vec{e}_1+3\vec{e}_2$ son vectores propios de la aplicación lineal $A$ de matriz

$$A=\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}$$

con respecto a la base $\{\vec{e}_1, \vec{e}_2\}$ de $\mathbb{R}^2$ y encontrar sus autovalores. ¿Es esta matriz diagonalizable en $\mathbb{R}$?



Puesto que

$$\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix}=\begin{pmatrix} 2 \\ 4 \end{pmatrix}=2\begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

y

$$\begin{pmatrix} 6 & -2 \\ 6 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix}=\begin{pmatrix} 6 \\ 9 \end{pmatrix}=3\begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

se tiene que $A(\vec{x})=2\vec{x}$ y $A(\vec{y})=3\vec{y}$ con lo que $\vec{x}$ e $\vec{y}$ son vectores propios de $A$ con valores propios 2 y 3, respectivamente. Como $\{\vec{x}, \vec{y}\}$ forman una base de $\mathbb{R}^2$, ya que

$$\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}=3-4=-1\neq 0$$

la matriz $A$ es diagonalizable en $\mathbb{R}$ y su matriz diagonal asociada es

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

* * *

A continuación mostramos cómo se calculan autovalores y autovectores de una aplicación lineal.

Supongamos que $\vec{x}$ es un vector propio de una aplicación lineal $A$ en un espacio vectorial $V$ y que $\lambda$ es su autovalor, es decir, $A(\vec{x})=\lambda\vec{x}$. Sea $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ una base de $V$ y $\vec{x}=x_1\vec{e}_1+\cdots+x_n\vec{e}_n$. Si $A=(a_{ij})$ es la matriz de $A$ con respecto a esta base tenemos que

$$\sum_{j=1}^{n}\lambda x_j\vec{e}_j=\lambda\sum_{j=1}^{n}x_j\vec{e}_j=\lambda\vec{x}=A(\vec{x})=A\left(\sum_{j=1}^{n}x_j\vec{e}_j\right)=$$

$$=\sum_{j=1}^{n}x_jA(\vec{e}_j)=\sum_{j=1}^{n}x_j\left(\sum_{i=1}^{n}a_{ij}\vec{e}_i\right)=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}x_j\right)\vec{e}_i.$$

Puesto que $\{\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n\}$ es una base de $V$ hemos de tener

$$\left.\begin{aligned} (a_{11}-\lambda)x_1+a_{12}x_2+\cdots+a_{1n}x_n&=0 \\ a_{21}x_1+(a_{22}-\lambda)x_2+\cdots+a_{2n}x_n&=0 \\ a_{n1}x_1+a_{n2}x_2+\cdots+(a_{nn}-\lambda)x_n&=0 \end{aligned}\right\} \qquad (1)$$

Puesto que (1) es un sistema homogéneo, para que posea una solución no nula se ha de tener que el determinante de la matriz de sus coeficientes sea nulo, esto es:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = |A-\lambda I| = 0 \tag{2}$$

donde $I$ denota la matriz identidad. La igualdad (2) es una ecuación en $\lambda$ de grado $n$ y sus soluciones en $\mathbb{K}$ son los autovalores de $A$. A partir de ahora restringiremos nuestra atención a espacios vectoriales sobre el cuerpo $\mathbb{R}$ o sobre $\mathbb{C}$. Si $V$ es un espacio vectorial complejo, la ecuación (2) tiene $n$ soluciones complejas contando cada una con su multiplicidad, debido al teorema fundamental del álgebra (ver la sección 4 del capítulo 4). Si $V$ es un espacio vectorial real no podemos asegurar que la ecuación (2) tenga $n$ soluciones reales.

EJEMPLO A. Determinar los valores y vectores propios de la aplicación lineal de $\mathbb{R}^2$ en $\mathbb{R}^2$ que tiene como matriz

$$A=\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}.$$

Los valores propios se determinan resolviendo la ecuación

$$0=|A-\lambda I|=\left|\begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix}-\lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right|=\begin{vmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{vmatrix}=$$
$$=(1-\lambda)(4-\lambda)-10=\lambda^2-5\lambda-6.$$

Sus raíces son $\lambda_1=6$, $\lambda_2=-1$ y, por tanto, éstos son los autovalores de $A$. El subespacio invariante correspondiente a $\lambda_1=6$ satisface las ecuaciones

$$A\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=6\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{o equivalentemente} \quad (A-6I)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

esto es:

$$\begin{pmatrix} -5 & 2 \\ 5 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

y, por tanto, $-5x_1+2x_2=0$ (la otra ecuación es combinación lineal de ésta). Así pues los vectores propios correspondientes a $\lambda_1=6$ son de la forma $\alpha\begin{pmatrix} 2 \\ 5 \end{pmatrix}$. Para $\lambda_2=-1$ tenemos

$$\begin{pmatrix} 2 & 2 \\ 5 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix}$$



y, por tanto, $2x_1+2x_2=0$; así pues, los vectores propios correspondientes a $\lambda_2=-1$ son de la forma $\beta\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Como $\begin{pmatrix} 2 \\ 5 \end{pmatrix}$ y $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ forman una base de $\mathbb{R}^2$ la proposición 3 nos permite deducir que $A$ es diagonalizable, con matriz diagonal $\begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix}$ y el cambio de base viene dado por la matriz

$$C=\begin{pmatrix} 2 & 1 \\ -5 & -1 \end{pmatrix}$$

* * *

El polinomio (2) se denomina *polinomio característico* de la aplicación $A$ o de la matriz $A$. Para poder hablar propiamente de «polinomio característico» es necesario demostrar que el polinomio (2) no depende de la base elegida en $V$ para escribir su matriz. Para demostrar esto, sea $P_B(\lambda)=|A-\lambda I|$ el polinomio característico de la aplicación $A$ en la base $B=\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ y sea $P_{B'}(\lambda)=|A'-\lambda I|$ el polinomio característico de $A$ en la base $B'=\{\vec{e}_1', \vec{e}_2', ..., \vec{e}_n'\}$; si $C$ es la matriz del cambio de base sabemos que $A'=C^{-1}AC$ y, por tanto,

$$P_{B'}(\lambda)=|A'-\lambda I|=|C^{-1}AC-\lambda I|=|C^{-1}AC-C^{-1}\lambda IC|=$$
$$=|C^{-1}||A-\lambda I||C|=|A-\lambda I|=P_B(\lambda)$$

Esto prueba que el polinomio característico no depende de la base elegida en $V$ para representar $A$.

* * *

A continuación realizamos más ejemplos de cálculo de autovalores y autovectores.

EJEMPLO B. La notación $R_\alpha$ de ángulo $\alpha$ en el plano tiene como matriz

$$R_\alpha=\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

con respecto a la base canónica de $\mathbb{R}^2$.
Sus autovalores son las soluciones de la ecuación

$$0=|R_\alpha-\lambda I|=\begin{vmatrix} \cos\alpha-\lambda & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha-\lambda \end{vmatrix}=(\cos\alpha-\lambda)^2+\operatorname{sen}^2\alpha=$$
$$=\lambda^2-2(\cos\alpha)\lambda+1$$

Sus soluciones son $\lambda_1 = \cos\alpha + i \operatorname{sen}\alpha$, $\lambda_2 = \cos\alpha - i\operatorname{sen}\alpha$. Estas soluciones son números complejos excepto si $\alpha = 2k\pi$ o $\alpha = (2k+1)\pi$ con $k$ un número entero, en cuyo caso los valores propios son reales. Si $\alpha = 2k\pi$ se tiene que

$$R_\alpha = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

que es la aplicación identidad; en este caso todo vector de $\mathbb{R}^2$ es un vector propio con valor propio 1. Si $\alpha = (2k+1)\pi$ se tiene que

$$R_\alpha = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Esta es la matriz de una simetría respecto al origen de coordenadas; sus vectores propios se determinan de la ecuación $(R_\alpha - (-1)I)(\vec{x}) = 0$, esto es, $0 \cdot \vec{x} = 0$; por tanto, todos los vectores no nulos de $\mathbb{R}^2$ son vectores propios de esta simetría; su autovalor es $-1$.

EJEMPLO C. La rotación $R_\alpha$ de ángulo $\alpha$, en el espacio, alrededor del eje $0z$ tiene como matriz

$$R_\alpha = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Su polinomio característico es

$$0 = \begin{vmatrix} \cos\alpha - \lambda & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha - \lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 2(\cos\alpha)\lambda + 1)$$

cuyas soluciones son $\lambda_1 = 1$, $\lambda_2 = \cos\alpha + i\operatorname{sen}\alpha$, $\lambda_3 = \cos\alpha - i\operatorname{sen}\alpha$. Los vectores propios para $\lambda_1 = 1$ son las soluciones en $\mathbb{R}^3$ de las ecuaciones

$$0 = \begin{pmatrix} \cos\alpha - 1 & -\operatorname{sen}\alpha & 0 \\ \operatorname{sen}\alpha & \cos\alpha - 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

esto es:

$$\left.\begin{aligned} (\cos\alpha - 1)x_1 - (\operatorname{sen}\alpha)x_2 = 0 \\ (\operatorname{sen}\alpha)x_1 + (\cos\alpha - 1)x_2 = 0 \end{aligned}\right\} \tag{3}$$



Puesto que

$$\begin{vmatrix} \cos\alpha - 1 & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha - 1 \end{vmatrix} = (\cos\alpha - 1)^2 + \operatorname{sen}^2\alpha = 2 - 2\cos\alpha = 4\operatorname{sen}^2\frac{\alpha}{2}$$

el sistema (3) tiene únicamente la solución $x_1 = x_2 = 0$ si $\alpha \neq 2k\pi$ con $k$ entero. En este caso los vectores propios correspondientes a $\lambda_1 = 1$ son de la forma $(0, 0, x_3)$.

Si $\alpha = 2k\pi$ se trata de la aplicación identidad y en este caso todos los vectores de $\mathbb{R}^3$ son invariantes. Si $\alpha = (2k+1)\pi$, $\lambda_2 = \lambda_3 = -1$ y sus vectores propios son todos los del plano $x0y$; en este caso tenemos una simetría con respecto al eje $0z$.

* * *

La proposición 3 nos da una condición necesaria y suficiente para saber cuándo una aplicación lineal es diagonalizable, a saber, que exista una base del espacio vectorial $V$ formada por vectores propios; en algunos casos puede resultar laborioso el encontrar esta base. Una condición que es suficiente para poder asegurar la diagonalización de una matriz está contenida en la proposición siguiente:

PROPOSICIÓN 4

Los vectores propios de una aplicación $A$ correspondientes a valores propios distintos dos a dos, son linealmente independientes.

*Nota.* Daremos dos demostraciones de esta proposición; una de ellas a continuación y la otra al final de esta sección.

*Demostración.* Realizaremos la demostración por inducción según el número de autovalores. Si sólo hay un autovalor $\lambda_1$ y $\vec{x}_1$ es uno de sus autovectores, $\vec{x}_1$ es linealmente independiente puesto que $\vec{x}_1 \neq 0$ por definición de vector propio.

Supongamos que existen dos autovalores $\lambda_1, \lambda_2$ con vectores propios $\vec{x}_1, \vec{x}_2$, respectivamente, y que $\lambda_1 \neq \lambda_2$. Si $\vec{x}_1, \vec{x}_2$ fueran linealmente dependientes podríamos encontrar $\alpha_1, \alpha_2$ no nulos a la vez, tal que

$$\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 = \vec{0}. \tag{4}$$

Aplicando $A$ a ambos miembros de (4) tenemos

$$\alpha_1 A(\vec{x}_1) + \alpha_2 A(\vec{x}_2) = \alpha_1\lambda_1\vec{x}_1 + \alpha_2\lambda_2\vec{x}_2 = \vec{0} \tag{5}$$

y multiplicando (4) por $\lambda_2$ tenemos

$$\alpha_1\lambda_2\vec{x}_1 + \alpha_2\lambda_2\vec{x}_2 = \vec{0}. \tag{6}$$

Restando (6) de (5) obtenemos $\alpha_1(\lambda_1 - \lambda_2)\vec{x}_1 = \vec{0}$; como $\lambda_1 \neq \lambda_2$, hemos de tener $\alpha_1 = 0$; sustituyendo $\alpha_1$ en (4) obtenemos $\alpha_2 = 0$, lo cual es una contradicción.

Para demostrar la proposición por inducción supongamos que es válida para cualesquiera $k-1$ valores propios y que tenemos $k$ valores propios $\lambda_1, \lambda_2, ..., \lambda_k$ distintos dos a dos con vectores propios $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$, respectivamente. Supongamos que tenemos una combinación lineal de la forma

$$\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_k\vec{x}_k = \vec{0}. \tag{7}$$

Aplicando $A$ a ambos miembros de (7) tenemos

$$A\left(\sum_{j=1}^{k} \alpha_j\vec{x}_j\right) = \sum_{j=1}^{k} \alpha_j A(\vec{x}_j) = \sum_{j=1}^{k} \alpha_j\lambda_j\vec{x}_j = \vec{0}. \tag{8}$$

y multiplicando (7) por $\lambda_k$ tenemos

$$\sum_{j=1}^{k} \alpha_j\lambda_k\vec{x}_j = \vec{0}. \tag{9}$$

Restando (9) de (8) obtenemos

$$\sum_{j=1}^{k-1} \alpha_j(\lambda_j - \lambda_k)\vec{x}_j = \vec{0}.$$

Puesto que los $\lambda_j$ son distintos dos a dos, la hipótesis de inducción nos permite concluir que $\alpha_1, \alpha_2, ..., \alpha_{k-1}$ son cero; sustituyendo en (7) se obtiene $\alpha_k = 0$ y, por tanto, $\{\vec{x}_1, ..., \vec{x}_k\}$ son linealmente independientes. ■

EJEMPLO D. Estudiaremos si la matriz

$$A = \begin{pmatrix} 6 & -2 & 1 \\ 6 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

es diagonalizable.

Su polinomio característico es

$$0 = \begin{vmatrix} 6-\lambda & -2 & 1 \\ 6 & -1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)\begin{vmatrix} 6-\lambda & -2 \\ 6 & -1-\lambda \end{vmatrix} = (1-\lambda)(\lambda-2)(\lambda-3),$$

con lo que sus valores propios son $\lambda_1 = 3$, $\lambda_2 = 2$, $\lambda_3 = 1$; sean $\vec{e}_1, \vec{e}_2, \vec{e}_3$ vectores propios correspondientes a $\lambda_1$, $\lambda_2$ y $\lambda_3$, respectivamente; como los autovalores son distintos dos a dos, $\vec{e}_1, \vec{e}_2, \vec{e}_3$ son linealmente independientes; por estar en un espacio de dimensión 3 forman una base y por la proposición 3 la matriz $A$ es diagonalizable.

* * *



*Observación.* Una matriz $A$ puede ser diagonalizable y tener autovalores múltiples, por ejemplo, la matriz identidad es diagonalizable y tiene como único autovalor 1.

EJEMPLO E. Tratemos de encontrar los valores de $a \in \mathbb{R}$ para los que la matriz

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & a & 0 \\ a & 0 & 0 \end{pmatrix}$$

es diagonalizable (en $\mathbb{R}^3$). Su polinomio característico es

$$0 = \begin{vmatrix} -\lambda & 0 & 0 \\ 0 & a-\lambda & 0 \\ a & 0 & -\lambda \end{vmatrix} = (-\lambda)(a-\lambda)(-\lambda) = (a-\lambda)\lambda^2.$$

Los autovalores de $A$ son $\lambda = a$ (simple) y $\lambda = 0$ (doble). Si $a = 0$, $A$ es la matriz nula, que es diagonal y, por tanto, diagonalizable.

Si $a \neq 0$, hemos de estudiar si existe una base de autovectores para poder utilizar la proposición 3.

Los autovectores correspondientes a $\lambda = a$ satisfacen:

$$\begin{pmatrix} -a & 0 & 0 \\ 0 & 0 & 0 \\ a & 0 & -a \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} ax_1 = 0 \\ ax_1 - ax_3 = 0 \end{matrix} \Leftrightarrow x_1 = 0,\ x_3 = 0;$$

por tanto, son de la forma

$$\vec{x} = (0, x_2, 0), \quad x_2 \in \mathbb{R}. \tag{1}$$

Los autovectores correspondientes a $\lambda = 0$ satisfacen:

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & a & 0 \\ a & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} ax_2 = 0 \\ ax_1 = 0 \end{matrix} \Leftrightarrow x_1 = 0,\ x_2 = 0$$

por tanto, son de la forma

$$\vec{y} = (0, 0, x_3), \quad x_3 \in \mathbb{R}. \tag{2}$$

Con los vectores que aparecen en (1) y (2) no puede obtenerse una base de $\mathbb{R}^3$ y, en consecuencia, la matriz $A$ no es diagonalizable si $a \neq 0$.

En resumen, la matriz $A$ es diagonalizable si $a = 0$ y no es diagonalizable si $a \neq 0$.

* * *

Sea $A \in L(V)$, donde $V$ es un espacio vectorial de dimensión finita sobre $\mathbb{K}$ y sea $\lambda_0$ un autovalor de $A$, con $\lambda_0 \in \mathbb{K}$. Denominamos *subespacio propio correspondiente a* $\lambda_0$ al subconjunto

$$E(\lambda_0) = \ker(A - \lambda_0 I).$$

Observar que $E(\lambda_0)$ contiene todos los vectores propios correspondientes al valor propio $\lambda_0$ junto con el vector $\vec{0}$.

Puesto que el núcleo de cualquier aplicación lineal es un subespacio vectorial de $V$, tenemos que $E(\lambda_0)$ es un subespacio vectorial de $V$. Además, de los resultados de la sección 6.4, se deduce que

$$\begin{aligned}\dim(E(\lambda_0)) &= \dim(\ker(A-\lambda_0 I)) = \dim(V) - \dim(\operatorname{img}(A-\lambda_0 I)) = \\ &= \dim(V) - r(A - \lambda_0 I)\end{aligned}$$

El siguiente ejemplo queda como ejercicio para el lector.

EJEMPLO F. Encontrar los subespacios propios de la aplicación lineal $A \in L(\mathbb{R}^3)$, que tiene como matriz

$$A = \begin{pmatrix} 3 & 1 & 5 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

* * *

El siguiente ejemplo sirve para mostrar que una matriz con elementos reales puede *no* ser diagonalizable en $\mathbb{R}$ y, sin embargo, ser diagonalizable en $\mathbb{C}$.

EJEMPLO E. La matriz

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

es diagonalizable en $\mathbb{C}$, ya que sus autovalores son $\lambda_1 = i$ y $\lambda_2 - i$, que son distintos (ver proposición 4). Sin embargo, el lector puede comprobar directamente que la matriz $A$ *no* es diagonalizable en $\mathbb{R}$.

* * *

*Demostración alternativa de la proposición* 4. Supongamos que $A \in L(V)$, donde $V$ es un espacio vectorial de dimensión $n$, y sea $B = \{\vec{e}_1, \vec{e}_2, ..., \vec{e}_m\}$ una base de $V$. Sean $\lambda_1, \lambda_2, ..., \lambda_n$ autovalores de $A$ distintos dos a dos y sean $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ autovectores correspondientes a $\lambda_1, \lambda_2, ..., \lambda_n$, respectivamente.

Para demostrar que $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ son linealmente independientes hemos de demostrar que una relación de la forma

$$\sum_{i=1}^{n} \alpha_i \vec{v}_i = \vec{0}, \quad \alpha_i \in \mathbb{K} \tag{1}$$



implica $\alpha_i = 0$, $\forall i = 1, 2, ..., n$. Aplicando $A$ a ambos miembros de la igualdad (1) se deduce que

$$\sum_{i=1}^{n} \alpha_i \lambda_i \vec{v}_i = \vec{0} \tag{2}$$

De manera similar se obtienen las relaciones

$$\sum_{i=1}^{n} \alpha_i \lambda_i^2 \vec{v}_i = \vec{0}, ..., \sum_{i=1}^{n} \alpha_i \lambda_i^{n-1} \vec{v}_i = \vec{0}$$

Escribiendo $\vec{v}_i = \sum_{k=1}^{m} v_{ik} \vec{e}_k$, las relaciones anteriores implican

$$\sum_{i=1}^{n} \alpha_i v_{ik} = 0, \sum_{i=1}^{n} \alpha_i \lambda_i v_{ik} = 0, ..., \sum_{i=1}^{n} \alpha_i \lambda_i^{n-1} v_{ik} = 0$$

para todo $k = 1, 2, ..., m$. Para cada $k = 1, 2, ..., m$, las igualdades anteriores determinan un sistema de ecuaciones lineales y homogéneo en las incógnitas $(\alpha_1 v_{1k}, ..., \alpha_n v_{nk})$. Puesto que la matriz de los coeficientes de este sistema es

$$A_n = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

su determinante es el determinante de Vandermonde y, por tanto,

$$|A_n| = \prod_{1 \le j < k \le n} (\lambda_k - \lambda_j)$$

(ver la sección 2.3). Puesto que los $\lambda_j$ son distintos dos a dos, $|A_n| \neq 0$, y, por tanto, el sistema anterior sólo tiene la solución trivial. Así pues, $\alpha_i v_{ik} = 0$ para todo $k = 1, 2, ..., m$. De aquí deducimos que

$$\alpha_i \vec{v}_i = \alpha_i \sum_{k=1}^{m} v_{ik} \vec{e}_k = \vec{0}, \quad i = 1, 2, ..., n$$

Puesto que los $\vec{v}_i$ son no nulos, ya que son autovectores, concluimos que $\alpha_i = 0$ para todo $i = 1, 2, ..., n$. Esto prueba que los autovectores $\vec{v}_1, \vec{v}_2, ..., \vec{v}_n$ son linealmente independientes, que era lo que queríamos demostrar. ■

## EJERCICIOS 7.2

**1.** Hallar los autovalores reales y los autovectores de $\mathbb{R}^n$ de las siguientes matrices:

$a)\ \begin{pmatrix} 4 & 6 \\ -3 & -5 \end{pmatrix} \quad b)\ \begin{pmatrix} 5 & -1 \\ 4 & 1 \end{pmatrix} \quad c)\ \begin{pmatrix} 2 & -1 \\ 3 & 1 \end{pmatrix} \quad d)\ \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$

$e)\ \begin{pmatrix} 0 & 0 & -1 \\ 1 & -2 & -1 \\ -2 & 3 & 1 \end{pmatrix} \quad f)\ \begin{pmatrix} 2 & 2 & -1 \\ 0 & -2 & 1 \\ -1 & 0 & 0 \end{pmatrix} \quad g)\ \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 0 \\ -1 & 1 & 0 \end{pmatrix}$

$h)\ \begin{pmatrix} 0 & -1 & 2 \\ 0 & -1 & 0 \\ -1 & 1 & -3 \end{pmatrix} \quad i)\ \begin{pmatrix} 2 & 2 & -1 \\ 0 & 2 & 0 \\ 3 & 1 & 1 \end{pmatrix}$

[*Sol.*: *a*) 1, $-2$; *b*) 3; *c*) no hay; *d*) $-1$; *e*) $\pm 1$; *f*) 0, $\sqrt{5}$, $\sqrt{-5}$; *g*) 1; *h*) $-1$, $-2$; *i*) 2.]
ble en real.

**2.** En los casos del ejercicio anterior en los que sea posible, hallar una base de $\mathbb{R}^n$ formada por autovectores, y la matriz, en esa base, de la aplicación dada.

[*Sol.*: $a)\ \begin{pmatrix} 1 & \\ & -2 \end{pmatrix} \quad b)\ \begin{pmatrix} -1 & \\ & -1 \end{pmatrix} \quad c)\ \begin{pmatrix} 0 & & \\ & 1 & \\ & & -3 \end{pmatrix} \quad d)\begin{pmatrix} -1 & & \\ & -1 & \\ & & -2 \end{pmatrix}$.]

**3.** Demostrar que el subespacio generado por los vectores $\vec{e}_1 + 2\vec{e}_2$ y $\vec{e}_2 + \vec{e}_3 + 2\vec{e}_4$ es invariante mediante la aplicación que en la base $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4\}$ tiene como matriz

$$A = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & 4 & -2 \\ 2 & -1 & 0 & 1 \\ 2 & -1 & -1 & 2 \end{pmatrix}$$

**4.** Encontrar los autovalores y los autovectores de las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^n$ que están dadas por las siguientes matrices:

$a)\ \begin{pmatrix} 2 & -1 & 2 \\ 5 & -3 & 3 \\ -1 & 0 & -2 \end{pmatrix} \quad b)\ \begin{pmatrix} 4 & -5 & 2 \\ 5 & -7 & 3 \\ 6 & -9 & 4 \end{pmatrix} \quad c)\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$



**5.** Decir cuáles de las siguientes matrices pueden reducirse a una matriz diagonal y encontrar una matriz de cambio de base *P*:

$a)\ \begin{pmatrix} -1 & 3 & -1 \\ -3 & 5 & -1 \\ -3 & 3 & 1 \end{pmatrix} \quad b)\ \begin{pmatrix} 4 & -1 & -1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \quad c)\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$

**6.** Encontrar los autovalores y autovectores de la aplicación derivación, $D$, en $P_{\mathbb{R}}^{(3)}[x]$.

**7.** Determinar para qué valores de $a$, $b \in \mathbb{R}$ la matriz $A = \begin{pmatrix} a & b & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ es diagonalizable en $\mathbb{R}$.

**8.** Estudiar para qué valores reales de $\alpha$ la matriz $A = \begin{pmatrix} 1 & -2 & -2-\alpha \\ 0 & 1 & \alpha \\ 0 & 0 & 1 \end{pmatrix}$ es diagonalizable.

**9.** Dada la matriz

$$A = \begin{pmatrix} 2 & -2 & 6 \\ 0 & a & 4-a \\ 0 & a & -a \end{pmatrix}$$

probar que $A$ es diagonalizable en $\mathbb{C}^3$ para todo $a \in \mathbb{C} - \{0, 1\}$; probar que también es diagonalizable para $a = 1$ y que no lo es para $a = 0$.

**10.** Encontrar los valores de $a$, $b \in \mathbb{R}$ para que la matriz $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & b & 0 \\ a & 0 & 0 \end{pmatrix}$ sea diagonalizable, sobre los números reales.

**11.** Sea $A \in L(\mathbb{R}^2)$ (= aplicaciones lineales en $\mathbb{R}^2$) tal que existe una base de $\mathbb{R}^2$ en la que su matriz $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ es simétrica. Probar que los autovalores de $A$ son reales y distintos, a menos que $a = c$ y $b = 0$, en cuyo caso $A = aI$ y $a$ es el único autovalor. Encontrar los autovalores y autovectores de la matriz $A$ dada por $A = \begin{pmatrix} 4 & 2 \\ 2 & 1 \end{pmatrix}$ en la base canónica.

**12.** Sea $S=\begin{pmatrix}\alpha & \beta\\ \beta & -\alpha\end{pmatrix}$ con $\alpha^2+\beta^2=1$ la matriz de $\mathscr{S}$ en la base canónica. Hallar los autovectores y autovalores de $\mathscr{S}$. ¿Existe alguna interpretación geométrica de la aplicación $\mathscr{S}$?

**13.** Sea $P=\begin{pmatrix}\alpha & \beta\\ \beta & 1-\alpha\end{pmatrix}$ con $\alpha^2+\beta^2=\alpha$ la matriz de $P\in L(\mathbb{R}^2)$ en la base canónica. Hallar los autovalores y autovectores de $P$. ¿Existe alguna interpretación geométrica de la aplicación $P$?

**14.** Si $A$ es una matriz triangular de orden $n$ cuyos elementos de la diagonal principal son todos diferentes, probar que $A$ es diagonalizable.

**15.** Probar que si $A$ es una aplicación lineal de $\mathbb{R}^n$ con $n$ autovalores distintos, cualquier aplicación lineal $B$ de $\mathbb{R}^n$ que conmute con $A$ tiene una base de autovectores, y todo autovector de $A$ es un autovector de $B$.

**16.** Dada una suma directa $V=W\oplus Z$ podemos definir $E\in L(V)$ como $E(\vec{v})=\vec{z}$ si $\vec{v}=\vec{w}+\vec{z}$ con $\vec{w}\in W$, $\vec{z}\in Z$; equivalentemente $E(\vec{v})\in Z$ y $(I-E)(\vec{v})\in W$.

*a*) Probar que $E^2=E$, $W=\ker(E)$ y $Z=\operatorname{img}(E)$.

*b*) Encontrar los autovalores de la aplicación $E$.

## 7.3. FORMA DE JORDAN DE MATRICES DE ORDEN 2

Dada una matriz $A\in\mathscr{M}_{n\times n}(\mathbb{K})$, la «forma más sencilla» a la cual puede reducirse mediante un cambio de base recibe el nombre de *matriz de Jordan de A*. Ya hemos observado en la sección 7.1 que la «forma más sencilla» no es siempre la matriz diagonal.

A la matriz de Jordan de $A$ se le representa por $J$, o $J_A$ cuando sea necesario, y a la matriz del cambio de base de $A$ a $J$ se le representa por $P$; así pues, se tiene

$$J=P^{-1}AP$$

o equivalentemente

$$A=PJP^{-1}.$$

Es conveniente observar que, aun cuando los elementos de $A$ sean reales, las matrices $J$ y $P$ pueden ser complejas.



En esta sección encontraremos la forma de Jordan de aplicaciones lineales $A\in L(V)$ donde $V$ es un espacio vectorial de dimensión dos.

* * *

Sea $A\in L(V)$ y supongamos que su matriz es

$$A=\begin{pmatrix}a & c\\ b & d\end{pmatrix}$$

en una cierta base fijada anteriormente. Su polinomio característico es

$$\varphi\,(A)=\begin{vmatrix}a-\lambda & c\\ b & d-\lambda\end{vmatrix}=(a-\lambda)(d-\lambda)-bc$$

con lo que $\varphi(A)=0$ es una ecuación de grado 2 en la variable $\lambda$. Tendremos dos casos diferentes según que las dos soluciones de esta ecuación sean iguales o distintas.

CASO I. *Las raíces del polinomio característico son distintas:* $\lambda\neq\mu$.

En este caso la matriz $A$ es diagonalizable (proposiciones 4 y 3 de la sección 7.2), su forma de Jordan es

$$\begin{pmatrix}\lambda & 0\\ 0 & \mu\end{pmatrix}$$

y la matriz $P$ del cambio de base tiene como primera columna las coordenadas de un vector $\vec{x}$ que satisface $(A-\lambda I)\vec{x}=\vec{0}$ (esto es, $\vec{x}\in\ker(A-\lambda I)$) y como segunda columna las coordenadas de un vector $\vec{y}$ que satisface $(A-\mu I)\vec{y}=\vec{0}$ (esto es, $\vec{y}\in\ker(A-\mu I)$).

EJEMPLO A. Encontrar la forma de Jordan $J$ de $A=\begin{pmatrix}3 & -4\\ 2 & -3\end{pmatrix}$ y la matriz del cambio de base.

Como

$$0=\begin{vmatrix}3-\lambda & -4\\ 2 & -3-\lambda\end{vmatrix}=(3-\lambda)(-3-\lambda)+8=\lambda^2-1$$

sus autovalores son $\lambda_1=1$, $\lambda_2=-1$; su matriz de Jordan es

$$J=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Si $\lambda_1=1$, tenemos

$$(A-I)\vec{x}=\vec{0}\Leftrightarrow\begin{pmatrix}2 & -4\\ 2 & -4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}0\\ 0\end{pmatrix}\Leftrightarrow 2x_1-4x_2=0;$$

así pues,

$$\ker(A-I)=\left\{c\begin{pmatrix}2\\1\end{pmatrix}/\ c\in\mathbb{K}\right\}.$$

Para $\lambda_2=-1$ tenemos

$$\begin{pmatrix}4 & -4\\2 & -2\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix}\Leftrightarrow x_1=x_2;$$

así pues,

$$\ker(A+I)=\left\{c\begin{pmatrix}1\\1\end{pmatrix}/\ c\in\mathbb{K}\right\}.$$

Tomando $\begin{pmatrix}2\\1\end{pmatrix}$ como base de $\ker(A-I)$ y $\begin{pmatrix}1\\1\end{pmatrix}$ como base de $\ker(A+I)$ obtenemos

$$A=\begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}2 & 1\\1 & 1\end{pmatrix}^{-1}$$

CASO II. *Las raíces $\lambda$ y $\mu$ del polinomio característico coinciden: $\lambda=\mu$.*

En este caso calculamos $\ker(A-\lambda I)$. Si $\ker(A-\lambda I)$ tiene dimensión 2 podemos encontrar una base de autovectores de $A$ y por la proposición 3 de la sección 7.2 la matriz $A$ es diagonalizable; su forma de Jordan es

$$J=\begin{pmatrix}\lambda & 0\\0 & \lambda\end{pmatrix}$$

y la matriz del cambio de base queda determinada por cualquier base de $\ker(A-\lambda I)$.

Si, por el contrario, $\ker(A-\lambda I)$ tiene dimensión 1 no podemos encontrar una base de autovectores en $V$; en este caso observamos que se tiene el siguiente resultado:

LEMA 1

Si $A$ tiene dos autovalores iguales $\lambda$, $(A-\lambda I)^2=0$.

*Demostración.* La demostración es sólo un ejercicio de cálculo; si $A=\begin{pmatrix}a & c\\b & d\end{pmatrix}$, sus autovalores satisfacen la ecuación

$$0=\begin{vmatrix}a-\lambda & c\\b & d-\lambda\end{vmatrix}=(a-\lambda)(d-\lambda)-bc=\lambda^2-(a+d)\lambda+ad-bc.$$



Si ambas soluciones de esta ecuación coinciden se ha de tener

$$(a+d)^2-4(ad-bc)=0 \quad \text{y} \quad \lambda=\frac{a+d}{2}$$

o equivalentemente,

$$(a-d)^2+4bc=0 \quad \text{y} \quad \lambda=\frac{a+d}{2}. \qquad (1)$$

Al calcular $(A-\lambda I)^2$ encontramos

$$(A-\lambda I)^2=\begin{pmatrix}a-\lambda & c\\b & d-\lambda\end{pmatrix}^2=\begin{pmatrix}(a-\lambda)^2+cb & (a-\lambda)c+c(d-\lambda)\\b(a-\lambda)+b(d-\lambda) & bc+(d-\lambda)^2\end{pmatrix}$$

que coincide con la matriz cero si utilizamos las igualdades (1). ■

Sea $E_1=\ker(A-\lambda I)$ y $E_2=\ker(A-\lambda I)^2$; el lema 1 nos dice que $E_2$ coincide con el espacio vectorial $V$ que estamos considerando y que es de dimensión dos. Como $E_1$ tiene dimensión 1 podemos encontrar $\vec{u}_2\in E_2-E_1$; tomar $\vec{u}_1=(A-\lambda I)(\vec{u}_2)$. Los vectores $\vec{u}_1$, $\vec{u}_2$ son linealmente independientes puesto que $u_2\notin E_1$ y $u_1\in E_1$ (ya que $(A-\lambda I)(\vec{u}_1)=(A-\lambda I)^2(\vec{u}_2)=0(\vec{u}_2)=\vec{0}$) y ninguno de ellos es el vector nulo. Como $V$ es un espacio de dimensión 2, $\{\vec{u}_1, \vec{u}_2\}$ es una base de $V$. En esta base tenemos

$$(A-\lambda I)(\vec{u}_1)=\vec{0} \Leftrightarrow A(\vec{u}_1)=\lambda\vec{u}_1$$
$$(A-\lambda I)(\vec{u}_2)=\vec{u}_1 \Leftrightarrow A(\vec{u}_2)=\vec{u}_1+\lambda\vec{u}_2$$

con lo que la matriz de la transformación lineal en esta base es

$$\begin{pmatrix}\lambda & 1\\0 & \lambda\end{pmatrix}.$$

Esta no es una matriz diagonal, pero es casi diagonal, y es la forma de Jordan de la matriz $A$ en este caso.

Podemos resumir estos resultados en la siguiente proposición, en donde $\mathbb{K}$ designa el cuerpo de los números reales o el de los complejos.

PROPOSICIÓN 2

Dada una matriz $A\in\mathcal{M}_{2\times 2}(\mathbb{K})$ siempre puede encontrarse una matriz $J\in\mathcal{M}_{2\times 2}(\mathbb{K})$ de una cualquiera de las formas

$$\begin{pmatrix}\lambda & 0\\0 & \mu\end{pmatrix} \quad \text{ó} \quad \begin{pmatrix}\lambda & 1\\0 & \lambda\end{pmatrix}, \quad \lambda, \mu\in\mathbb{K}$$

y una matriz $P \in \mathscr{M}_{2\times 2}(\mathbb{K})$ de determinante no nulo, tal que

$$A = PJP^{-1}$$

La matriz $J$ se denomina *matriz de Jordan de A*.

EJEMPLO B. Reducir la matriz $A = \begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix}$ a su forma de Jordan.

El polinomio característico de la matriz $A$ es

$$\begin{vmatrix} -\lambda & 2 \\ -2 & 4-\lambda \end{vmatrix} = (4-\lambda)(-\lambda)+4 = \lambda^2 - 4\lambda + 4 = (\lambda-2)^2$$

que tiene como raíz (o solución) $\lambda = 2$. Con $\lambda = 2$ se tiene que

$$E_1 = \ker(A-2I) = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} / x_1 = x_2 \right\} = \left\{ c\begin{pmatrix} 1 \\ 1 \end{pmatrix} / c \in \mathbb{K} \right\}$$

ya que

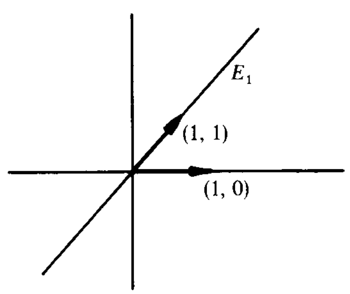
$$\begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 = x_2.$$

Sabemos, debido al lema 1, que $E_2 = \ker(A-2I)^2$ tiene dimensión 2; elegimos $\vec{u}_2$ de manera que $\vec{u}_2$ no esté en $E_1$; por ejemplo:

$$\vec{u}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Entonces,

$$\vec{u}_1 = (A-2I)(\vec{u}_2) = \begin{pmatrix} -2 & 2 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -2 \end{pmatrix}.$$






La matriz de Jordan es $J = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$ y la matriz del cambio de base es $P = \begin{pmatrix} -2 & 1 \\ -2 & 0 \end{pmatrix}$; así pues, hemos de tener

$$\begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -2 & 1 \\ -2 & 0 \end{pmatrix}^{-1}$$

lo cual puede comprobarse fácilmente.

* * *

EJEMPLO C. Encontrar todas las funciones $x(t)$ que coinciden con su derivada segunda, esto es, $x''(t) = x(t)$.

Si llamamos $y(t) = x'(t)$ podemos escribir

$$x'(t) = y(t)$$
$$y'(t) = x(t)$$

que en forma matricial se escribe como

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}.$$

Encontramos la forma de Jordan de $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$; como

$$\begin{vmatrix} -\lambda & 1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 - 1 = (\lambda+1)(\lambda-1)$$

la matriz $A$ es diagonalizable (sus autovalores son distintos) con matriz de Jordan

$$J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Para encontrar la matriz del cambio de base calculamos $E(1) = \ker(A-I)$ y $E(-1) = \ker(A+I)$:

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 = x_2 \Leftrightarrow E(-1) = \left\{ c\begin{pmatrix} 1 \\ 1 \end{pmatrix} / c \in \mathbb{K} \right\}$$

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1 = -x_2 \Leftrightarrow E(-1) = \left\{ c\begin{pmatrix} -1 \\ 1 \end{pmatrix} / c \in \mathbb{K} \right\}.$$

Podemos tomar $\vec{u}_1=\begin{pmatrix}1\\1\end{pmatrix}$ y $\vec{u}_2=\begin{pmatrix}-1\\1\end{pmatrix}$ como vectores de la nueva base. Tenemos entonces

$$\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}=\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}^{-1}=$$

$$=\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1/2 & 1/2\\-1/2 & 1/2\end{pmatrix}$$

Por tanto,

$$\begin{pmatrix}x'(t)\\y'(t)\end{pmatrix}=\begin{pmatrix}1 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1/2 & 1/2\\-1/2 & 1/2\end{pmatrix}\begin{pmatrix}x(t)\\y(t)\end{pmatrix}$$

o equivalentemente,

$$\begin{pmatrix}1/2 & 1/2\\-1/2 & 1/2\end{pmatrix}\begin{pmatrix}x'(t)\\y'(t)\end{pmatrix}=\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}\begin{pmatrix}1/2 & 1/2\\-1/2 & 1/2\end{pmatrix}\begin{pmatrix}x(t)\\y(t)\end{pmatrix}$$

Si hacemos $u(t)=\frac{1}{2}x(t)+\frac{1}{2}y(t)$ y $v(t)=-\frac{1}{2}x(t)+\frac{1}{2}y(t)$ tenemos $u'=u$ y $v'=-v$, de donde se deduce

$$u(t)=u_0e^t, \quad v(t)=v_0e^{-t}$$

Como $x(t)=u-v$ tenemos

$$x(t)=u_0e^t-v_0e^{-t}$$

que son todas las funciones buscadas, cuando $u_0$ y $v_0$ son dos constantes cualesquiera.

## EJERCICIOS 7.3

**1.** Encontrar la forma de Jordan de $A=\begin{pmatrix}1 & 1\\1 & 0\end{pmatrix}$ y calcular $A^5$.

**2.** Probar por inducción que

$$\begin{pmatrix}a & 1\\0 & a\end{pmatrix}^n=\begin{pmatrix}a^n & na^{n-1}\\0 & a^n\end{pmatrix}$$



y calcular

$$\begin{pmatrix}3 & 1\\-1 & 1\end{pmatrix}^{10}$$

**3.** Dado un polinomio $p(x)=\sum_{n=0}^{N} a_nx^n$, definimos

$$p(A)=\sum_{n=0}^{N} a_nA^n$$

Dada $A=\begin{pmatrix}-6 & 2\\-6 & 1\end{pmatrix}$ y $p(x)=x^2+5x+6$, calcular $p(A)$ haciendo uso de la forma de Jordan de $A$.

**4.** Si $p_A(x)$ es el polinomio característico de la matriz $A$, demostrar que $p_A(A)=0$, donde $A\in\mathcal{M}_{2\times 2}(\mathbb{C})$. (Este resultado se conoce con el nombre de *teorema de Cayley-Hamilton;* un resultado análogo es cierto para matrices de orden superior.)

**5.** Sea $A$ una matriz de orden 2 tal que $A^2=0$. Demostrar que para todo $\lambda$, $|A-\lambda I|=\lambda^2$.

## 7.4. FORMA DE JORDAN DE MATRICES DE ORDEN 3

En esta sección realizaremos varios ejemplos de diagonalización de aplicaciones lineales entre espacios vectoriales de dimensión 3, que nos servirán para comprender mejor los resultados teóricos necesarios para la obtención de la forma de Jordan de aplicaciones en espacios vectoriales de cualquier dimensión. Los resultados generales se expondrán en las secciones siguientes.

Al final de la sección resumiremos los procedimientos seguidos en estos ejemplos.

EJEMPLO A. Intentaremos reducir la matriz

$$A=\begin{pmatrix}0 & 3 & 1\\2 & -1 & -1\\-2 & -1 & -1\end{pmatrix}$$

a su forma de Jordan.

Sus autovalores satisfacen la ecuación

$$0=|A-\lambda I|=-\lambda^3-2\lambda^2+4\lambda+8=-(\lambda-2)(\lambda+2)^2,$$

y, por tanto, son

$$\lambda=2 \text{ (simple)} \quad \text{y} \quad \mu=-2 \text{ (doble)}.$$

Calculemos $E(2)=\ker(A-2I)$, que es el subespacio vectorial definido por las ecuaciones

$$\begin{pmatrix}-2 & 3 & 1\\ 2 & -3 & -1\\ -2 & -1 & -3\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\begin{pmatrix}0 & 0 & 0\\ 2 & -3 & -1\\ 0 & -4 & -4\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\begin{matrix}2x_1-3x_2-x_3=0\\ -4x_2-4x_3=0;\end{matrix}$$

tomando $x_3=-1$ se tiene $x_2=1$, $x_1=1$, con lo que $\vec{u}_1=\begin{pmatrix}1\\1\\-1\end{pmatrix}$ es un autovector correspondiente a $\lambda=2$; además, tenemos que $\dim(E(2))=1$.

Para $\mu=-2$, $E_1(-2)=\ker(A+2I)$ es el subespacio vectorial definido por las ecuaciones

$$\begin{pmatrix}2 & 3 & 1\\ 2 & 1 & -1\\ -2 & -1 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\begin{pmatrix}2 & 3 & 1\\ 2 & 1 & -1\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow$$

$$\Leftrightarrow\begin{pmatrix}2 & 3 & 1\\ 0 & -2 & -2\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}$$

con lo que

$$E_1(-2)=\left\{a\begin{pmatrix}1\\-1\\1\end{pmatrix},\ a\in\mathbb{K}\right\}$$

Como $E(2)+E_1(-2)$ no llena todo el espacio vectorial no podemos elegir una base de autovectores. Es necesario en este caso realizar un truco análogo al realizado en el caso de raíces iguales para matrices de orden 2. Calculamos $E_2(-2)=\ker(A+2I)^2$, que es el subespacio vectorial que tiene por ecuaciones

$$\begin{pmatrix}2 & 3 & 1\\ 2 & 1 & -1\\ -2 & -1 & 1\end{pmatrix}^2\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\begin{pmatrix}8 & 8 & 0\\ 8 & 8 & 0\\ -8 & -8 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow x_1=-x_2;$$

por tanto,

$$E_2(-2)=\left\{a\begin{pmatrix}1\\-1\\0\end{pmatrix}+b\begin{pmatrix}0\\0\\1\end{pmatrix}/a,\ b\in\mathbb{K}\right\}$$



que es un subespacio de dimensión 2. Como $E_1(2)+E_2(-2)$ llena todo el espacio, podemos elegir una base de $V$ de manera que la matriz de $A$ sea, en esta base, bastante sencilla. La forma de elegir esta base se muestra a continuación.

Sea $\vec{u}_3=\begin{pmatrix}0\\0\\1\end{pmatrix}$ un vector de $E_2(-2)$ que *no* está en $E_1(-2)$ y tomemos

$$u_2=(A+2I)\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}2 & 3 & 1\\ 2 & 1 & -1\\ -2 & -1 & 1\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}1\\-1\\1\end{pmatrix}.$$

Fácilmente se comprueba que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base y en esta base la expresión de la aplicación $A$ que tiene a $A$ como matriz es

$$A(\vec{u}_1)=2\vec{u}_1,\quad A(\vec{u}_2)=-2\vec{u}_2,\quad A(\vec{u}_3)=\vec{u}_2-2\vec{u}_3$$

con lo que

$$\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ -1 & 1 & 1\end{pmatrix}^{-1}A\begin{pmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ -1 & 1 & 1\end{pmatrix}=\begin{pmatrix}2 & 0 & 0\\ 0 & -2 & 1\\ 0 & 0 & -2\end{pmatrix}.$$

La matriz de la derecha de esta expresión es una matriz de Jordan de $A$.

* * *

EJEMPLO B. Reducir la matriz $A=\begin{pmatrix}1 & 0 & 0\\ 1 & 0 & -1\\ 1 & -1 & 0\end{pmatrix}$ a una forma de Jordan.

Como $|A-\lambda I|=(1-\lambda)(\lambda^2-1)=(1-\lambda)(\lambda+1)(\lambda-1)$, los autovalores de $A$ son $\lambda=1$ (doble), $\mu=-1$ (simple).

Para $\lambda=1$,

$$\begin{pmatrix}0 & 0 & 0\\ 1 & -1 & -1\\ 1 & -1 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow x_1=x_2+x_3$$

y, por tanto,

$$E_1(1)=\ker(A-I)=\left\{a\begin{pmatrix}1\\1\\0\end{pmatrix}+b\begin{pmatrix}1\\0\\1\end{pmatrix}/a,\ b \text{ escalares}\right\}.$$

Para $\lambda=-1$,

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & -1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & -1 \\ 2 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} x_1=0 \\ x_1+x_2=x_3 \end{matrix}$$

y, por tanto,

$$E_1(-1)=\ker(A+I)=\left\{a\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}/a \text{ escalar}\right\}.$$

Como $E_1(1)+E_1(-1)$ llena todo el espacio la matriz $A$ es diagonalizable y se tiene

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} A \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

* * *

EJEMPLO C. Tratemos de encontrar una forma de Jordan de la matriz

$$A=\begin{pmatrix} -2 & 1 & -1 \\ -1 & -1 & 0 \\ 0 & 1 & -3 \end{pmatrix}.$$

Puesto que $|A-\lambda I|=-\lambda^3-6\lambda^2-12\lambda-8=-(\lambda+2)^3$, los autovalores de $A$ son $\lambda=-2$ (triple).

Como

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1=x_2=x_3$$

se tiene que

$$E_1(-2)=\ker(A+2I)=\left\{a\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}/a \text{ escalar}\right\}.$$



Como no existe una base de autovectores procedemos a calcular $E_2(-2)=\ker(A+2I)^2$:

$$\begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}^2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow x_1=x_3$$

$$E_2(-2)=\left\{a\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}+b\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}/a,\ b \text{ escalares}\right\}$$

Puesto que con $E_2(-2)$ no se llena todo el espacio vectorial calculamos $E_3(-2)=\ker(A-\lambda I)^3$; como $(A-\lambda I)^3=0$, lo cual puede comprobarse fácilmente, formamos la cadena de subespacios.

$$E_1(-2) \subsetneqq E_2(-2) \subsetneqq E_3(-2)=V$$

y eleguimos $\vec{u}_3 \in V$ de manera que $\vec{u}_3 \notin E_2(-2)$; por ejemplo,

$$\vec{u}_3=\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

Tomamos $\vec{u}_2=(A+2I)\vec{u}_3$, $\vec{u}_1=(A+2I)\vec{u}_2$; tenemos

$$\vec{u}_2=\begin{pmatrix} 0 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix} \quad ; \quad \vec{u}_1=\begin{pmatrix} 0 & 1 & 1 \\ -1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}=\begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

con lo que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base; en esta base,

$$A(\vec{u}_1)=-2\vec{u}_1, \quad A(\vec{u}_2)=\vec{u}_1-2\vec{u}_2, \quad A(\vec{u}_3)=\vec{u}_2-2\vec{u}_3$$

con lo cual deducimos

$$\begin{pmatrix} -1 & 0 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}^{-1} A \begin{pmatrix} -1 & 0 & 1 \\ -1 & -1 & 0 \\ -1 & 0 & 0 \end{pmatrix}=\begin{pmatrix} -2 & 1 & 0 \\ 0 & -2 & 1 \\ 0 & 0 & -2 \end{pmatrix}$$

* * *

EJEMPLO D. Encontrar una forma de Jordan de la matriz

$$A=\begin{pmatrix}-2 & 0 & 1\\ 0 & -1 & 0\\ -1 & 0 & 0\end{pmatrix}$$

Sus autovalores son $\lambda=-1$ (triple) puesto que $|A-\lambda I|=-(\lambda+1)^3$. Las ecuaciones de $E_1(-1)=\ker(A+I)$ son

$$\begin{pmatrix}-1 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix} \Leftrightarrow x_1=x_3$$

por tanto,

$$E_1(-1)=\left\{a\begin{pmatrix}1\\ 0\\ 1\end{pmatrix}+b\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}/a,\ b \text{ escalares}\right\}$$

Puesto que $E_1(-1)$ no llena todo el espacio vectorial encontramos $E_2(-1)=\ker(A+I)^2$:

$$\begin{pmatrix}-1 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 1\end{pmatrix}^2\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix} \Leftrightarrow \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x_1\\ x_2\\ x_3\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}$$

Tenemos la cadena $E_1(-1) \subsetneq E_2(-1) = V$, con lo cual podemos elegir $\vec{u}_3 \in V-E_1(-1)$: por ejemplo, $\vec{u}_3=\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}$. Tomamos ahora $\vec{u}_2=(A+I)\vec{u}_3$:

$$u_2=\begin{pmatrix}-1 & 0 & 1\\ 0 & 0 & 0\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix}1\\ 0\\ 0\end{pmatrix}=\begin{pmatrix}-1\\ 0\\ -1\end{pmatrix}$$

Como $E_1(-1)$ es un espacio de dimensión 2 aún podemos elegir $\vec{u}_1 \in E_1(-1)$ de manera que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ sea una base de $V$; tomar, por ejemplo,

$$\vec{u}_1=\begin{pmatrix}0\\ 1\\ 0\end{pmatrix}.$$



En esta base,

$$A(\vec{u}_1)=-\vec{u}_1,\quad A(\vec{u}_2)=-\vec{u}_2,\quad A(\vec{u}_3)=\vec{u}_2-\vec{u}_3$$

y, por tanto,

$$\begin{pmatrix}0 & -1 & 1\\ 1 & 0 & 0\\ 0 & -1 & 0\end{pmatrix}^{-1} A\begin{pmatrix}0 & -1 & 1\\ 1 & 0 & 0\\ 0 & -1 & 0\end{pmatrix}=\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 1\\ 0 & 0 & -1\end{pmatrix}$$

* * *

Resumimos a continuación el procedimiento utilizado en la obtención de las matrices de Jordan de los ejemplos anteriores.

Una vez calculados los autovalores $\lambda$ se obtienen los núcleos de $A-\lambda I$, donde $A$ es la matriz dada; si estos núcleos llenan todo el espacio vectorial puede encontrarse una base de autovectores y la matriz es diagonalizable; si los núcleos no llenan todo el espacio vectorial es porque existen autovalores dobles o triples y en este caso es necesario calcular la cadena de los núcleos de $(A-\lambda I)^j$, $j=1, 2, 3$, Cuando se haya llenado todo el espacio vectorial procederemos a elegir una base adecuada como se ha mostrado en los ejemplos.

## EJERCICIOS 7.4

**1.** Encontrar una forma canónica de Jordan y el cambio de base correspondiente de las siguientes matrices:

$$A=\begin{pmatrix}3 & -1 & -1\\ 1 & 1 & -1\\ 1 & -1 & 1\end{pmatrix},\ B=\begin{pmatrix}1 & 2 & 1\\ 2 & 1 & 1\\ -1 & 1 & 0\end{pmatrix},\ C=\begin{pmatrix}4 & -1 & -2\\ 2 & 1 & -2\\ 1 & -1 & 1\end{pmatrix},\ D=\begin{pmatrix}3 & 2 & -2\\ 0 & 4 & -1\\ 0 & 1 & 2\end{pmatrix}$$

$$E=\begin{pmatrix}0 & -1 & -2\\ 1 & 3 & 1\\ 1 & 0 & 3\end{pmatrix},\ F=\begin{pmatrix}-2 & 0 & -1\\ -1 & -1 & -1\\ 1 & 0 & 0\end{pmatrix},\ G=\begin{pmatrix}3 & 2 & -3\\ 4 & 10 & -12\\ 3 & 6 & -7\end{pmatrix},\ H=\begin{pmatrix}0 & 1 & -1\\ 1 & 0 & 1\\ 1 & -1 & 2\end{pmatrix}$$

**2.** Hallar $C^6$.

**3.** Encontrar todos los subespacios invariantes de la aplicación $A$: $\mathbb{R}^3 \mapsto \mathbb{R}^3$ dada por la matriz

$$A=\begin{pmatrix}4 & -2 & 2\\ 2 & 0 & 2\\ -1 & 1 & 1\end{pmatrix}$$

4. Sea $A=\begin{pmatrix}\alpha & \lambda & \mu\\ \lambda & \beta & \nu\\ \mu & \nu & \gamma\end{pmatrix}$ la matriz de $A\in L(\mathbb{R}^3)$ en la base canónica; supongamos que $\operatorname{rango}(A)=1$ y $\operatorname{tr}(A)=\alpha+\beta+\gamma=1$. Hallar los autovalores y autovectores de $A$.

## 7.5. APLICACIONES LINEALES Y SUBESPACIOS INVARIANTES

En esta sección daremos dos propiedades de las aplicaciones lineales que nos ayudarán en la demostración del teorema de Jordan sobre la reducción de matrices a su forma «más sencilla».

Si $A$ es una aplicación definida en un espacio vectorial complejo, con matriz $A$ respecto de una base fijada, el polinomio característico

$$P_A(\lambda)=|A-\lambda I|$$

de esta aplicación tiene tantas soluciones complejas como el orden del polinomio (este resultado se conoce con el nombre de *teorema fundamental del álgebra* y ha sido enunciado con anterioridad en el capítulo 4). Si $\lambda_0$ es una solución de $P_A(\lambda)=0$ con autovector $\vec{v}_0$, el subespacio $L(\vec{v}_0)=\{a\vec{v}_0/a\in\mathbb{C}\}$ es invariante:

$$A(a\vec{v}_0)=aA(\vec{v}_0)=a\lambda_0\vec{v}_0\in L(v_0).$$

Hemos probado el siguiente resultado:

**PROPOSICIÓN 1**

Toda aplicación lineal en un espacio vectorial *complejo* posee un subespacio invariante de dimensión 1.

En el caso de que la aplicación lineal esté definida sobre un espacio vectorial *real*, su polinomio característico puede tener todas las soluciones complejas y en este caso no podemos realizar el razonamiento anterior para obtener un subespacio de dimensión 1; se tiene, sin embargo, el siguiente resultado:

**PROPOSICIÓN 2**

Toda aplicación lineal en un espacio vectorial *real* posee un subespacio invariante de una o dos dimensiones.

*Demostración.* Si el polinomio característico $P_A(\lambda)=|A-\lambda I|$ de la aplicación $A$ tiene al menos una raíz real se procede como en el caso complejo para encontrar un subespacio invariante de dimensión uno.



Supongamos, por el contrario, que $P_A(\lambda)$ no posee raíces reales y sea $\lambda=\alpha+i\beta$ una de sus raíces complejas.

Sea $A\in\mathcal{M}_{n\times n}(\mathbb{R})$ la matriz de $A$ en una base cualquiera del espacio vectorial. Los elementos $\vec{z}=(z_1,\dots,z_n)\in\mathbb{C}^n$ tal que $(A-\lambda I)\vec{z}=\vec{0}$ pueden escribirse de la forma

$$z_1=x_1+iy_1,\quad z_2=x_2+iy_2,\dots,\ z_n=x_n+iy_n$$

donde $x_j$, $y_j$ son números reales. Escribiendo todas las ecuaciones del sistema $(A-\lambda I)\vec{z}=\vec{0}$ tenemos

$$\begin{array}{l}
a_{11}(x_1+iy_1)+a_{12}(x_2+iy_2)+\cdots+a_{1n}(x_n+iy_n)=(\alpha+i\beta)(x_1+iy_1)\\
a_{21}(x_1+iy_1)+a_{22}(x_2+iy_2)+\cdots+a_{2n}(x_n+iy_n)=(\alpha+i\beta)(x_2+iy_2)\\
\vdots\\
a_{n1}(x_1+iy_1)+a_{n2}(x_2+iy_2)+\cdots+a_{nn}(x_n+iy_n)=(\alpha+i\beta)(x_n+iy_n)
\end{array}$$

Igualando las partes reales e imaginarias obtenemos los sistemas

$$\left\{\begin{array}{ccccccc}
a_{11}x_1 & + & a_{12}x_2 & +\cdots+ & a_{1n}x_n & = & \alpha x_1-\beta y_1\\
a_{21}x_1 & + & a_{22}x_2 & +\cdots+ & a_{2n}x_n & = & \alpha x_2-\beta y_2\\
\vdots & & \vdots & & \vdots & \vdots & \vdots\\
a_{n1}x_1 & + & a_{n2}x_2 & +\cdots+ & a_{nn}x_n & = & \alpha x_n-\beta y_n
\end{array}\right\} \qquad \text{(I)}$$

y

$$\left\{\begin{array}{ccccccc}
a_{11}y_1 & + & a_{12}y_2 & +\cdots+ & a_{1n}y_n & = & \alpha y_1+\beta x_1\\
a_{21}y_1 & + & a_{22}y_2 & +\cdots+ & a_{2n}y_n & = & \alpha y_2+\beta x_2\\
\vdots & & \vdots & & \vdots & \vdots & \vdots\\
a_{n1}y_1 & + & a_{n2}y_2 & +\cdots+ & a_{nn}y_n & = & \alpha y_n+\beta x_n
\end{array}\right\} \qquad \text{(II)}$$

Considerando los vectores reales

$$\vec{u}=\operatorname{real}\vec{z}=x_1\vec{e}_1+\cdots+x_n\vec{e}_n,\quad \vec{v}=\operatorname{img}\vec{z}=y_1\vec{e}_1+\cdots+y_n\vec{e}_n$$

los sistemas (I) y (II) nos permiten escribir

$$A(\vec{u})=\alpha\vec{u}-\beta\vec{v},\quad A(\vec{v})=\alpha\vec{v}+\beta\vec{u}. \qquad (1)$$

Por tanto, el subespacio $W$ generado por los vectores $\vec{u}$ y $\vec{v}$ es invariante respecto de $A$. Unicamente falta demostrar que este subespacio es de dimensión dos; esto es cierto porque si $\vec{u}$ y $\vec{v}$ fuesen linealmente dependientes tendríamos $\vec{v}=\gamma\vec{u}$ con $\gamma\in\mathbb{R}$ y, por tanto,

$$A(\vec{u})=\alpha\vec{u}-\beta\vec{v}=(\alpha-\beta\gamma)\vec{u}$$

con lo que $A$ tendría un autovalor real, en contra de lo supuesto. ■

*Nota.* Si se mira a la demostración anterior, se deduce de (1) que la matriz de $A$. restringida al plano $W$, con respecto a la base real $\{\vec{u}, \vec{v}\}$ tiene la forma

$$A|_W = \begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}$$

EJEMPLO A. Si existe un autovalor complejo $\lambda$ de la matriz

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 5 & 3 & 0 \\ 2 & 0 & -4 \end{pmatrix}$$

encontrar un plano invariante en $\mathbb{R}^3$ y la matriz respecto de la base $\{\text{real}\,\vec{z}, \text{img}\,\vec{z}\}$ de la restricción a este plano, donde $\vec{z}$ es un autovector correspondiente a $\lambda$.

Puesto que $|A-\lambda I| = (-4-\lambda)(\lambda^2-4\lambda+8)$ los autovalores de $A$ son $\lambda=-4$, $\mu=2+2i$, $\eta=2-2i$. Los autovectores (complejos) correspondientes a $\mu$ se obtienen del sistema

$$\begin{pmatrix} -1-2i & -1 & 0 \\ 5 & 1-2i & 0 \\ 2 & 0 & -6-2i \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \left\{\begin{matrix} (1+2i)z_1 + z_2 = 0 \\ 5z_1 + (1-2i)z_2 = 0 \\ 2z_1 - (6+2i)z_3 = 0 \end{matrix}\right\} \Leftrightarrow$$

$$\left\{\begin{matrix} z_2 = -(1+2i)z_1 \\ z_1 = (3+i)z_3 \end{matrix}\right\}$$

y, por tanto, el subespacio invariante (complejo) correspondiente a $\mu$ es $\ker(A-\mu I)$ $= \left\{ a\begin{pmatrix} 3+i \\ -1-7i \\ 1 \end{pmatrix} \Big/ a \in \mathbb{C} \right\}$. Tomar los vectores (reales)

$$\vec{u} = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 1 \\ -7 \\ 0 \end{pmatrix}$$

que son los que generan el plano invariante $W$ en $\mathbb{R}^3$ correspondiente a $\mu = 2+2i$; con respecto a esta base:

$$A|_W = \begin{pmatrix} 2 & 2 \\ -2 & 2 \end{pmatrix}. \tag{2}$$

Para el autovalor $\eta = 2-2i$, que es el conjugado de $\mu$, el sistema $(A-\eta I)\vec{z} = (A-\bar{\mu} I)\vec{z} = \vec{0}$ tiene las soluciones $\bar{z} = (\bar{z}_1, \bar{z}_2, \bar{z}_3)$ con $z_1, z_2, z_3$ como antes, con lo cual

$$\ker(A-\eta I) = \left\{ a\begin{pmatrix} 3-i \\ -1+7i \\ 1 \end{pmatrix} \Big/ a \in \mathbb{C} \right\}.$$



Los vectores (reales)

$$\vec{u}' = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix}, \quad \vec{v}' = \begin{pmatrix} -1 \\ 7 \\ 0 \end{pmatrix}$$

generan el mismo subespacio $W$ de antes (observar que $\vec{v}' = -\vec{v}$); por tanto, el plano invariante es el mismo y la matriz coincide con la dada en (2), si se toma la base $\{\vec{u}, \vec{v}\}$.

* * *

La otra observación que haremos en esta sección se refiere a la matriz de una aplicación lineal $A$ en un espacio vectorial $V$ el cual puede descomponerse en suma directa de dos o más subespacios invariantes respecto de $A$.

Supongamos que $V = W_1 \oplus W_2$ es suma directa de los subespacios $W_1$, $W_2$ que son invariantes respecto de $A$; recordemos que $V = W_1 \oplus W_2$ si y sólo si $V = W_1 + W_2$ y si $\vec{w}_1 + \vec{w}_2 = \vec{0}$ con $\vec{w}_j \in W_j$ se tiene $\vec{w}_j = \vec{0}$, $j=1, 2$. Sea $\{\vec{e}_1, ..., \vec{e}_r\}$ una base de $W_1$ y $\{\vec{e}_{r+1}, ..., \vec{e}_n\}$ una base de $W_2$; como $W_1$ y $W_2$ son invariantes tenemos:

$$\begin{aligned} A(\vec{e}_1) &= a_{11}\vec{e}_1 + \cdots + a_{r1}\vec{e}_1 \\ &\vdots \\ A(\vec{e}_r) &= a_{1r}\vec{e}_1 + \cdots + a_{rr}\vec{e}_r \\ A(\vec{e}_{r+1}) &= a_{r+1,r+1}\vec{e}_{r+1} + \cdots + a_{n,r+1}\vec{e}_n \\ &\vdots \\ A(\vec{e}_n) &= a_{r+1,n}\vec{e}_{r+1} + \cdots + a_{n,n}\vec{e}_n \end{aligned}$$

con lo que la matriz de $A$ respecto de la base $\{\vec{e}_1, ..., \vec{e}_r, \vec{e}_{r+1}, ..., \vec{e}_n\}$ tiene la forma

$$A = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$$

con $A_1$ matriz de dimensión $r$ y $A_2$ matriz de dimensión $n-r$.

Se tiene, pues, que, en este caso, la matriz de $A$ puede dividirse en «cajas» que corresponden a las matrices de $A$ restringidas a cada uno de los subespacios invariantes. Los elementos que están fuera de estas «cajas» son todos ceros.

Si se conocen $m$ subespacios invariantes de $A$, que llenan todo $V$ mediante una suma directa, la matriz de $A$ puede escribirse de la forma

$$A = \begin{pmatrix} A_1 & & & 0 \\ & A_2 & & \\ & & \ddots & \\ 0 & & & A_m \end{pmatrix}$$

en una base convenientemente elegida.

La demostración de este último resultado es muy similar a la demostración del resultado anterior; el lector no tendrá ninguna dificultad para realizarla por sí solo.

EJERCICIOS 7.5

**1.** En los casos siguientes, hallar los autovalores (reales o complejos) y los correspondientes autovectores de $\mathbb{C}^n$; y cuando haya un par de autovectores conjugados $u$, $\bar{u}$ hallar el correspondiente plano invariante en $\mathbb{R}^n$, y la matriz respecto de la base $\{\operatorname{real}\bar{u}, \operatorname{img}\bar{u}\}$ de la restricción a este plano.

*a*) $\begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix}$ *b*) $\begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}$ *c*) $\begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$ *d*) $\begin{pmatrix} 1 & 1+i \\ -i & 1-i \end{pmatrix}$

[*Sol.:* *a*) $1,\ \pm\sqrt{3}i;\ x_2=0;\ \begin{pmatrix} 0 & \sqrt{3} \\ -\sqrt{3} & 0 \end{pmatrix}$

*b*) $1\pm i\sqrt{6};\ \begin{pmatrix} 1 & \sqrt{6} \\ -\sqrt{6} & 1 \end{pmatrix}$

*c*) $\pm i;\ \{x_2=x_4=0\};\ \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix};\ \{x_1=-x_2,\ x_3=0\}$

*d*) $0,\ 2-i$

**2.** Sea $A\in L(V)$ una aplicación lineal invertible y $W$ un subespacio vectorial de $V$ invariante por $A$. Demostrar que $W$ es también invariante por $A^{-1}$.

## 7.6. TEOREMA DE CLASIFICACION DE JORDAN

Con la experiencia acumulada en los ejemplos que hemos realizado en las secciones precedentes, nos disponemos a demostrar a continuación el teorema de clasificación de matrices, que se atribuye a C. Jordan, y que es uno de los teoremas más profundos del álgebra lineal.

El lector que desee comenzar aprendiendo la forma de obtener matrices de Jordan puede pasar directamente a la sigiente sección, en donde se hace un resumen de cómo se calculan y se dan varios ejemplos. Es conveniente, sin embargo, que se lea el enunciado del teorema de clasificación que a continuación se expone.

Dos matrices $A$, $B\in\mathcal{M}_n(\mathbb{C})$ se dicen *equivalentes* si existe un cambio de base $P$ tal que

$$P^{-1}AP=B.$$



Denominaremos *matriz elemental de Jordan* de orden $k$ y autovalor $\lambda\in\mathbb{C}$ a la matriz $J_k(\lambda)$ de orden $k$ cuyos elementos son todos nulos, excepto los de la diagonal principal, que valen $\lambda$, y los situados inmediatamente encima de la diagonal principal, que son unos. Por ejemplo:

$$J_1(\lambda)=(\lambda),\quad J_2(\lambda)=\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix},\quad J_3(\lambda)=\begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix},\quad J_4(\lambda)=\begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}$$

y así sucesivamente.

Llamaremos *matriz de Jordan* a cualquier matriz cuadrada formada por yuxtaposición de matrices elementales de Jordan a lo largo de la diagonal, de la forma

$$\begin{pmatrix} \boxed{J_{k_1}(\lambda_1)} & & & 0 \\ & \boxed{J_{k_2}(\lambda_2)} & & \\ & & \ddots & \\ 0 & & & \boxed{J_{k_r}(\lambda_r)} \end{pmatrix}$$

TEOREMA DE CLASIFICACIÓN DE JORDAN

Toda matriz cuadrada (real o compleja) es equivalente a una matriz de Jordan (compleja), determinada de manera única salvo por permutaciones de las matrices de Jordan elementales que la componen.

Existen varias demostraciones de este teorema; la que nosotros daremos se basa en la obtención práctica de la matriz de Jordan de una dada. La demostración será larga y en ella habrá varios lemas que servirán para que el lector comprenda mejor el resultado final.

Sea $\lambda\in\mathbb{C}$ un autovalor de una matriz $A$ de orden $n$ de la cual queremos obtener su forma de Jordan. Formamos la cadena creciente de subespacios vectoriales

$$E_1(\lambda)\subset E_2(\lambda)\subset\cdots\subset E_j(\lambda)\subset\cdots$$

donde $E_j(\lambda)=\ker(A-\lambda I)^j$. Como están contenidas en un espacio de dimensión finita existe al menos un número natural $m\in\mathbb{N}$ tal que $E_m(\lambda)=E_{m+1}(\lambda)$; sea $p\in\mathbb{N}$ el menor de estos números naturales $m$; este $p$ tiene la propiedad de que a a partir de él todos los subespacios $E_{p+1}(\lambda)$, $E_{p+2}(\lambda)$, ... coinciden con $E_p(\lambda)$. Esto está contenido en el siguiente lema:

LEMA 1

Si $E_p(\lambda)=E_{p+1}(\lambda)$, entonces $E_p(\lambda)=E_q(\lambda)$ para todo $q>p$.

*Demostración.* La demostración la realizamos por inducción en $r$, donde $q=p+r$. Si $r=1$ la conclusión coincide con la hipótesis y no es necesario demostrar nada. Supongamos que el lema es cierto para $q=p+r$ y demostrémoslo para $q=p+r+1$. Sea $\vec{x}\in E_{p+r+1}(\lambda)$, es decir,

$$(A-\lambda I)^{p+r}(A-\lambda I)\vec{x}=(A-\lambda I)^{p+r+1}\vec{x}=\vec{0}.$$

Por tanto, $(A-\lambda I)\vec{x}\in E_{p+r}(\lambda)=E_p(\lambda)$ (por la hipótesis de inducción); entonces $\vec{0}=(A-\lambda I)^p(A-\lambda I)\vec{x}=(A-\lambda I)^{p+1}\vec{x}$ y, por tanto, $\vec{x}\in E_{p+1}(\lambda)=E_p(\lambda)$. Hemos probado que $E_{p+r+1}(\lambda)\subset E_p(\lambda)$, de donde se deduce la igualdad puesto que la otra inclusión es trivial de verificar. ■

Sea $F_p=\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_r\}$ un conjunto de vectores linealmente independientes de $E_p(\lambda)-E_{p-1}(\lambda)$ de manera que $L(\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_r\})\oplus E_{p-1}(\lambda)=E_p(\lambda)$. (Observar que si $d_j=\dim(E_j(\lambda))$, $r=d_p-d_{p-1}$). Probaremos a continuación que las imágenes sucesivas de los vectores $\vec{u}_1, \vec{u}_2, ..., \vec{u}_n$ mediante potencias $(A-\lambda I)^k$, $1\leqq k\leqq p-1$, forman un conjunto de vectores linealmente independientes.

LEMA 2

Con las condiciones anteriores, $\{(A-\lambda I)^k\vec{u}_1, (A-\lambda I)^k\vec{u}_2, ..., (A-\lambda I)^k\vec{u}_r\}\subset E_{p-k}(\lambda)$ son linealmente independientes y

$$L(\{(A-\lambda I)^k\vec{u}_1, ...., (A-\lambda I)^k\vec{u}_r\})\bigcap E_{p-k-1}(\lambda)=\{\vec{0}\}$$

*Demostración.* Puesto que $(A-\lambda I)^{p-k}(A-\lambda I)^k\vec{u}_j=(A-\lambda I)^p\vec{u}_j=\vec{0}$, hemos probado que $(A-\lambda I)^k\vec{u}_j\in E_{p-k}(\lambda)$ para todo $j=1, ..., r$.

Supongamos que tenemos una combinación lineal de la forma $\sum_{j=1}^{r}\alpha_j(A-\lambda I)^k\vec{u}_j=\vec{0}$; como $(A-\lambda I)^k$ es lineal tenemos

$$(A-\lambda I)^k\left(\sum_{j=1}^{r}\alpha_j\vec{u}_j\right)=\vec{0}.$$

Si probamos que $(A-\lambda I)^k$ es inyectiva en $L(F_p)$ (=espacio generado por $F_p$), de la anterior igualdad podemos deducir $\sum_{j=1}^{r}\alpha_j\vec{u}_j=\vec{0}$, lo cual implica $\alpha_j=0$, para todo $j=1, ..., r$ puesto que los $\vec{u}_j$ son linealmente independientes.

Para probar que $(A-\lambda I)^k$ es inyectiva en $L(F_p)$ basta observar que

$$\ker(A-\lambda I)^k\cap L(F_p)=E_k(\lambda)\cap L(F_p)\subset E_{p-1}\cap L(F_p)=\{\vec{0}\}$$

Finalmente, sea $\vec{v}\in E_{p-k-1}(\lambda)$ y tal que

$$\vec{v}=\sum_{j=1}^{r}\alpha_j(A-\lambda I)^k\vec{u}_j;$$



entonces,

$$\vec{0}=(A-\lambda I)^{p-k-1}(\vec{v})=(A-\lambda I)^{p-1}\left(\sum_{j=1}^{r}\alpha_j\vec{u}_j\right)$$

de aquí deducimos que $\sum_{j=1}^{r}\alpha_j\vec{u}_j\in E_{p-1}(\lambda)$. Puesto que los $\vec{u}_j$ se han elegido de $E_p(\lambda)-E_{p-1}(\lambda)$ hemos de tener

$$\sum_{j=1}^{r}\alpha_j\vec{u}_j=\vec{0}\Leftrightarrow\alpha_1=\alpha_2=\cdots=\alpha_r=0$$

y, por tanto, $\vec{v}=0$. Esto prueba la segunda parte del lema 2 y termina su demostración. ■

El lema 2 muestra que la familia de vectores

$$\begin{aligned}B_p=\{&\vec{u}_1^{p-1}=(A-\lambda I)^{p-1}\vec{u}_1, ..., \vec{u}_1^1=(A-\lambda I)\vec{u}_1, \vec{u}_1,\\ &\vec{u}_2^{p-1}=(A-\lambda I)^{p-1}\vec{u}_2, ..., \vec{u}_2^1=(A-\lambda I)\vec{u}_2, \vec{u}_2,\\ &..., \vec{u}_r^{p-1}=(A-\lambda I)^{p-1}\vec{u}_r, ..., \vec{u}_r^1=(A-\lambda I)\vec{u}_r, \vec{u}_r\}\end{aligned}$$

de las potencias de todos los vectores de $F_p$ son linealmente independientes; además,

$$\begin{array}{rcl}
\vec{u}_1^{p-1}\in E_1(\lambda) & \Rightarrow & A(\vec{u}_1^{p-1})=\lambda\vec{u}_1^{p-1}\\
(A-\lambda I)(A-\lambda I)^{p-2}\vec{u}_1=\vec{u}_1^{p-1} & \Rightarrow & A(\vec{u}_1^{p-2})=\vec{u}_1^{p-1}+\lambda\vec{u}_1^{p-2}\\
 & \vdots & \\
(A-\lambda I)\vec{u}_1=\vec{u}_1^1 & \Rightarrow & A(\vec{u}_1)=\vec{u}_1^1+\lambda\vec{u}_1\\
\vec{u}_2^{p-1}\in E_1(\lambda) & \Rightarrow & A(\vec{u}_2^{p-1})=\lambda\vec{u}_2^{p-1}\\
(A-\lambda I)(A-\lambda I)^{p-2}\vec{u}_2=\vec{u}_2^{p-1} & \Rightarrow & A(\vec{u}_2^{p-2})=\vec{u}_2^{p-1}+\lambda\vec{u}_2^{p-2}\\
 & \vdots & \\
(A-\lambda I)\vec{u}_2=\vec{u}_2^1 & \Rightarrow & A(u_2)=\vec{u}_2^2+\lambda\vec{u}_2\\
 & \vdots &
\end{array}$$

Entonces, la matriz de $A$ restringida al subespacio vectorial generado por $B_p$, que es invariante, tiene la forma

$$\left(\begin{array}{ccc|ccc|ccc}
\lambda & 1 & & & & & & & \\
 & \lambda & 1 & & & & & 0 & \\
 & & \ddots\ \lambda & & & & & & \\
\hline
 & & & \lambda & 1 & & & & \\
 & & & & \lambda & 1 & & & \\
 & & & & & \ddots\ \lambda & & & \\
\hline
 & & & & & & \lambda & 1 & \\
 & 0 & & & & & & \lambda & 1\\
 & & & & & & & & \ddots\ \lambda
\end{array}\right)=\left(\begin{array}{ccc}
\boxed{J_p(\lambda)} & & 0\\
 & \boxed{J_p(\lambda)} & \\
0 & & \ddots\ \boxed{J_p(\lambda)}
\end{array}\right)$$

es decir, está formada por la yuxtaposición de $r$ matrices elementales de Jordan de orden $p$ todas ellas colocadas sobre la diagonal principal.

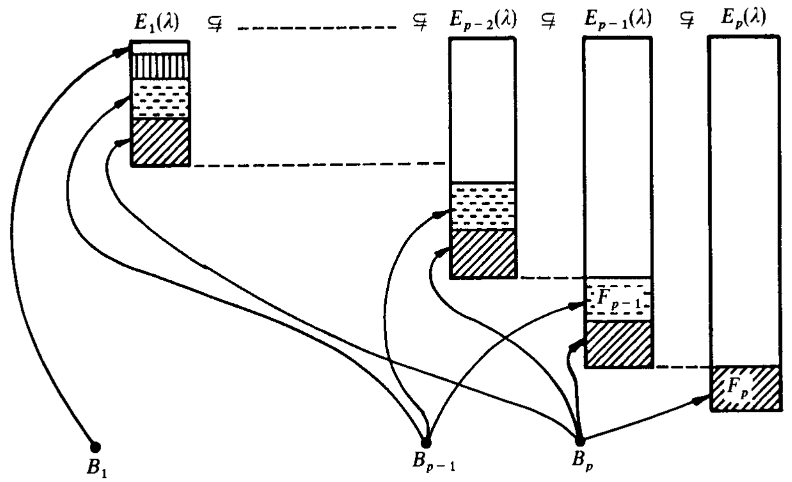



Figura 1

Es conveniente ahora observar la figura 1 para aclarar la situación; los elementos de $B_p$ que están en $E_{p-1}(\lambda)$ puede que no llenen todo el espacio entre $E_{p-1}(\lambda)$ y $E_{p-2}(\lambda)$; en este caso se eligen $\{\vec{u}_{r+1}, ..., \vec{u}_l\}$ linealmente independientes de $E_{p-1}(\lambda) - E_{p-2}(\lambda)$ y se procede como en el caso anterior (las mismas demostraciones sirven) para encontrar una colección $B_{p-1}$ de vectores linealmente independientes con respecto a la cual la matriz de $A$ restringida a $L(B_{p-1})$ tiene la forma

$$\begin{pmatrix} J_{p-1}(\lambda) & & & 0 \\ & J_{p-1}(\lambda) & & \\ & & \ddots & \\ 0 & & & J_{p-1}(\lambda) \end{pmatrix}$$

es decir, está formada por la yuxtaposición de $l-r$ matrices elementales de Jordan de orden $p-1$.

Este proceso se repite hasta llegar a elegir un conjunto de vectores $B_1$ de $E_1(\lambda) - L(B_p) \cup \cdots \cup L(B_2)$ que sean linealmente independientes.

Por la forma en que se han ido eligiendo los vectores, la colección

$$B(\lambda) = B_p \cup B_{p-1} \cup \cdots \cup B_2 \cup B_1$$

forman una base de $E_p(\lambda) = \ker(A - \lambda I)^p$.

En esta base la matriz de $A$ restringida a $E_p(\lambda)$ es una yuxtaposición de matrices elementales de Jordan de diversos órdenes asociadas con el autovalor $\lambda$.

**LEMA 3**

El subespacio máximo $E_p(\lambda)$, asociado con un autovalor $\lambda$, es invariante.



*Demostración.* Sea $\vec{v} \in E_p(\lambda)$; entonces $A\vec{v} = (A - \lambda I)\vec{v} + \lambda\vec{v}$ y, por tanto, basta demostrar que $(A - \lambda I)\vec{v} \in E_p(\lambda)$. Pero

$$(A - \lambda I)^p (A - \lambda I)\vec{v} = (A - \lambda I)^{p+1}\vec{v} = (A - \lambda I)(A - \lambda I)^p \vec{v} = \vec{0},$$

lo cual prueba el resultado deseado. ■

Hasta ahora hemos conseguido obtener la forma de Jordan asociada con la matriz $A$ restringida a los subespacios máximos $E_p(\lambda)$ de cada autovalor. El último paso es «pegar» adecuadamente las matrices de todos los autovalores.

Sean $\lambda_1, \lambda_2, ..., \lambda_m$ $(m \leqslant n)$ todos los autovalores distintos de $A$ y sean $B(\lambda_i) \equiv B^i$, $i = 1, ..., m$, las bases de los autoespacios máximos $E_{p_i}(\lambda_i)$ correspondientes a $\lambda_i$.

Tomar

$$B = B^1 \cup B^2 \cup \cdots \cup B^m$$

Si demostramos que

$$V = E_{p_1}(\lambda_1) \oplus E_{p_2}(\lambda_2) \oplus \cdots \oplus E_{p_m}(\lambda_m), \tag{1}$$

donde $V$ es el espacio vectorial en el cual estamos trabajando, la última observación de la sección 7.5 nos permite escribir la matriz de $A$ *respecto de la base* $B$ como

$$J = \begin{pmatrix} A|_{E_{p_1}(\lambda_1)} & & & 0 \\ & A|_{E_{p_2}(\lambda_2)} & & \\ & & \ddots & \\ 0 & & & A|_{E_{p_m}\lambda_m)} \end{pmatrix}$$

donde cada una de las matrices $A|_{E_{p_j}(\lambda_j)}$ es una yuxtaposición de matrices elementales de Jordan. Esto probaría la primera parte del teorema.

La demostración de (1) se da en los dos lemas siguientes.

**LEMA 4**

Si $\vec{v}_j \in E_{p_j}(\lambda_j)$, $j = 1, 2, ..., s$ y $\vec{v}_1 + \cdots + \vec{v}_s = \vec{0}$, se tiene que $\vec{v}_1 = \vec{v}_2 = \cdots = \vec{v}_s = \vec{0}$.

*Demostración.* La demostración se realiza por inducción sobre $s$. Si $s = 1$ el resultado es trivial. Supongamos que se cumple para $s$ vectores $\vec{v}_1, ..., \vec{v}_s$ y sea $\vec{v}_{s+1} \in E_{p_{s+1}}(\lambda_{s+1})$ $(s + 1 \leqq m)$; si

$$\vec{v}_1 + \cdots + \vec{v}_s + \vec{v}_{s+1} = \vec{0}$$

se tiene que

$$(A - \lambda_{s+1} I)^{p_{s+1}} \vec{v}_1 + \cdots + (A - \lambda_{s+1} I)^{p_{s+1}} \vec{v}_s + (A - \lambda_{s+1} I)^{p_{s+1}} \vec{v}_{s+1} = \vec{0}$$

o equivalentemente

$$(A-\lambda_{s+1}I)^{p_{s+1}}\vec{v}_1+\cdots+(A-\lambda_{s+1}I)^{p_{s+1}}\vec{v}_s=\vec{0}.$$

Como $(A-\lambda_{s+1}I)$ *es inyectiva en* $E_{p_j}(\lambda_j)$ para $j\neq s+1$ (¿demostración?) se tiene que

$$\vec{v}_1+\cdots+\vec{v}_s=\vec{0}.$$

Por la hipótesis de inducción, $\vec{v}_1=\cdots=\vec{v}_s=\vec{0}$ y consecuentemente $\vec{v}_{s+1}=0$. ■

LEMA 5

$$V=E_{p_1}(\lambda_1)+E_{p_2}(\lambda_2)+\cdots+E_{p_m}(\lambda_m)$$

*Demostración.* Supongamos, en contra de lo que queremos demostrar, que $B=\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_N\}$ con $N<n$, y consideremos $\bar{B}=\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$ una base de $V$ que se obtiene ampliando $B$. Sea $H=L(\bar{B}-B)$. Si para cada $j=1, ..., n$, $A\vec{v}_j=\sum_{k=1}^{n}\alpha_j^k\vec{v}_k$ definimos una aplicación lineal $A_0$ de $H$ en $H$ mediante

$$A_0\vec{v}_j=\sum_{k=N+1}^{n}\alpha_j^k\vec{v}_k.\quad j=N+1, ..., n.$$

$A_0$ posee un autovalor $\lambda_0$ y un autovector $\vec{v}_0\in H$. Como $A\vec{v}_0-A_0\vec{v}_0\in L(B)$, existen $\vec{v}_1\in E_{p_1}(\lambda_1), ..., \vec{v}_m\in E_{p_m}(\lambda_m)$ tal que

$$A\vec{v}_0=\lambda_0\vec{v}_0+\vec{v}_1+\cdots+\vec{v}_m. \tag{2}$$

Si $\lambda_0=\lambda_1$, aplicando $(A-\lambda_1I)^n$ a ambos miembros tenemos

$$(A-\lambda_1I)^{n+1}\vec{v}_0=(A-\lambda_1I)^n\vec{v}_2+\cdots+(A-\lambda_1I)^n\vec{v}_m.$$

Como $(A-\lambda_1I)$ es un automorfismo de $E_{p_j}(\lambda_j)$, $j=2, ..., m$ (inténtese demostrar; ver ejercicio 8), existen $\vec{w}_1\in E_{p_1}(\lambda_1), ..., \vec{w}_m\in E_{p_m}(\lambda_m)$ tal que $(A-\lambda_1I)\vec{w}_j=\vec{v}_j$, $j=2, ..., m$, con lo cual se tiene

$$(A-\lambda_1I)^{n+1}\vec{v}_0=(A-\lambda_1I)^{n+1}\vec{w}_2+\cdots+(A-\lambda_1I)^{n+1}\vec{w}_m.$$

Puesto que $p_1\leqq n+1$ tenemos que $\vec{v}_0-\vec{w}_2-\cdots-\vec{w}_m\in E_{p_1}(\lambda_1)$, o lo que es lo mismo, $\vec{v}_0\in E_{p_1}(\lambda_1)+\cdots+E_{p_m}(\lambda_m)=L(B)$, lo cual es absurdo puesto que $\vec{v}_0\in H$, $\vec{v}_0\neq 0$.

Si $\lambda_0$ coincidiera con cualquier otro de los $\lambda_j$ se realizaría un argumento similar.

Por último, si $\lambda_0$ no fuera ningún autovalor de $A$, $(A-\lambda_0I)$ sería un automorfismo de $E_{p_j}(\lambda_j)$, $j=1, 2, ..., m$ (ver ejercicio 8); existirían, por tanto, $\vec{w}_j\in E_{p_j}(\lambda_j)$, $j=1, 2, ..., m$, tal que $(A-\lambda_0I)\vec{w}_j=\vec{v}_j$; sustituyendo en (2) tendríamos

$$(A-\lambda_0I)\vec{v}_0=(A-\lambda_0I)\vec{v}_1+\cdots+(A-\lambda_0I)\vec{v}_m.$$



Por tanto, $\vec{v}_0-\vec{v}_1-\cdots-\vec{v}_m\in\ker(A-\lambda_0I)$; puesto que $\lambda_0$ no es autovalor de $A$, $\ker(A-\lambda_0I)=\{\vec{0}\}$ y entonces $\vec{v}_0=\vec{v}_1+\cdots+\vec{v}_m\in L(B)$, lo cual es también una contradicción. ■

Esto concluye la demostración de la primera parte del teorema de clasificación de Jordan. El hecho de que la matriz de Jordan es única salvo permutaciones de las matrices de Jordan elementales que la componen se deduce de que éstas quedan determinadas por las dimensiones de los núcleos

$$\ker(A-\lambda_jI)^k$$

$j=1, 2, ..., m$, $k=1, 2, ..., p_j$. ■

*Nota.* La demostración que acaba de darse está tomada de unas notas de R. Moriyón: *Clasificación de Jordan de matrices*, U.A.M.

* * *

La obtención del autoespacio máximo $E_p(\lambda)$ asociado con un autovalor $\lambda$ mediante la estabilización de la cadena

$$E_1(\lambda)\subset E_2(\lambda)\subset\cdots\subset E_j(\lambda)\subset\cdots$$

puede aliviarse si se conoce la siguiente propiedad:

$$\dim E_p(\lambda)=\text{multiplicidad de }\lambda. \tag{3}$$

El resto de esta sección lo dedicaremos a demostrar esta propiedad. Sean $\lambda_1, ..., \lambda_m$ los autovalores de $A$ con multiplicidad $r_1, ..., r_m$, respectivamente: esto es,

$$|A-\lambda I|=(\lambda-\lambda_1)^{r_1}\times\cdots\times(\lambda-\lambda_m)^{r_m}. \tag{4}$$

LEMA 6

$$\dim E_{p_i}(\lambda_i)\geqq r_i,\quad i=1, 2, ..., m$$

*Demostración.* Como $V=E_{p_1}(\lambda_1)\oplus\cdots\oplus E_{p_m}(\lambda_m)$ deducimos de la última observación de la sección anterior que la matriz de $A$ en una base adecuada puede escribirse de la forma

$$\begin{pmatrix} \boxed{A_1} & & & \\ & \boxed{A_2} & & \\ & & \ddots & \\ & & & \boxed{A_m} \end{pmatrix}$$

donde $A_j=A|_{E_{p_j}(\lambda_j)}$. Por tanto,

$$|A-\lambda I|=|A_1-\lambda I|\times\cdots\times|A_m-\lambda I|$$

Como $\lambda-\lambda_i$ no divide a $|A_j-\lambda\cdot I|$, $j\neq i$, $j=1, ..., m$, se tiene que $(\lambda-\lambda_i)^{r_i}$ divide a $|A_i-\lambda I|$; por tanto, $|A_i-\lambda I|$ es un polinomio de grado $\geqq r_i$; por otro lado, el grado de $|A_i-\lambda I|$ coincide con $\dim E_{p_i}(\lambda_i)$. Esto prueba el resultado deseado. ■

Si $n$ es la dimensión del espacio vectorial $V$, de (1) se deduce que

$$n=\dim E_{p_1}(\lambda_1)+\cdots+\dim E_{p_m}(\lambda_m)$$

y de (4) se deduce que

$$n=r_1+\cdots+r_m$$

Del lema 6 y estas dos expresiones se deduce que

$$r_j=\dim E_{p_j}(\lambda_j)$$

$j=1, 2, ..., m$, que es lo que se afirma en (3).

## EJERCICIOS 7.6

**1.** Escribir todos los tipos posibles de forma de Jordan de matrices de orden menor o igual que cuatro.

**2.** Sea $A$ una matriz real de orden 2 y $C$ la matriz de un cambio de base; demostrar directamente que $\text{traza}(A)=\text{traza}\ (C^{-1}AC)$, donde la traza de una matriz se define como la suma de los elementos de su diagonal principal.

**3.** *a)* Si $A$ es una matriz de orden 3 y $C$ la matriz de un cambio de base, demostrar que $\text{traza}(A)=\text{traza}(C^{-1}AC)$. [*Sugerencia:* utilizar el hecho de que el polinomio característico no depende de la base elegida.]
*b)* Generalizar el resultado anterior para matrices de orden $n$.

**4.** *a)* Demostrar que $T^k-\lambda^k I=(T-\lambda I)(T^{k-1}+\lambda T^{k-2}+\cdots+\lambda^{k-1}I)$.
*b)* Si $q(\lambda)=a_0+a_1\lambda+\cdots+a_m\lambda^m$ y $q(T)=a_0I+a_1\ T+\cdots+a_mT^m$, demostrar que: $\lambda$ autovalor de $T\Rightarrow q(\lambda)$ autovalor de $q(T)$. (Este problema nos indica entre qué valores buscar los autovalores para casos sencillos.)
*c)* Qué valores posibles de autovalores puede tener $T$ si: 1) $T^2=T$; 2) $T^2=I$ es una involución; 3) $T^n=I$, $n\geqslant 3$; 4) $T^2=-I$.
*d)* Si $T$ es invertible y $\lambda$ es autovalor de $T$, demostrar que $\lambda^{-1}$ es autovalor de $T^{-1}$.

**5.** *a)* Supongamos que $T$ y $S$ conmutan, esto es, $TS=ST$. Demostrar que $\ker(S)$ e $\text{img}(S)$ son invariantes por $T$.
*b)* Probar que si $p$ y $q$ son polinomios, entonces $p(T)q(T)=q(T)p(T)$. En consecuencia, $\ker(q(T))$ e $\text{img}(q(T))$ son invariantes por $p(T)$.

**6.** Dados $T$, $S\in L(V)$, $\dim V=n$, tales que ambos poseen $n$ autovalores distintos dos a dos (no necesariamente iguales), probar que $TS=ST\Leftrightarrow T$ y $S$ tienen los mismos autovectores.



**7.** Sea $E_p(\lambda)$ el autoespacio máximo correspondiente a un autovalor $\lambda$ de la aplicación $T\in L(V)$ y sea

$$d_j=\dim(E_j(\lambda))-\dim(E_{j-1}(\lambda))$$

$j=1, 2, ..., p$. Demostrar que $d_1\geqq d_2\geqq\cdots\geqq d_p\geqq 1$.

**8.** Si $\lambda_0$ no es un autovalor de $\mathscr{A}$, demostrar que $\mathscr{A}-\lambda_0 I$ es un automorfismo de $E_p(\lambda)$, donde $\lambda$ es un autovalor de $\mathscr{A}$ y $E_p(\lambda)$ es su autoespacio máximo. (Ver lema 5 de la demostración del teorema de Jordan.)

* * *

Una aplicación lineal $T$ se dice *nilpotente* si existe $k\in\mathbb{N}$ tal que $T^k=0$; si, además, $T^{k-1}\neq 0$, a $k$ se le llama *orden de nilpotencia de T.*
Dado $T\in L(V)$ nilpotente, se dice que es nilcíclico si existe $\vec{u}\in V$ tal que $V=L(\vec{u}, T\vec{u}, T^2\vec{u}, ..., T^{k-1}\vec{u})$, con $T^{k-1}\vec{u}\neq\vec{0}$.

**9.** Si $T$ es nilcíclico con orden de nilpotencia $k$ y $\vec{u}\in V$ es tal que $V=L(\vec{u}, T\vec{u}, ..., T^{k-1}\vec{u})$, demostrar que $\{\vec{u}, T\vec{u}, ..., T^{k-1}\vec{u}\}$ es una base de $V$.

**10.** Si $\vec{u}_1=T^{k-1}\vec{u}$, $\vec{u}_2=T^{k-2}\vec{u}$, ..., $\vec{u}_{k-1}=T\vec{u}$, $\vec{u}_k=\vec{u}$, con $\vec{u}$ y $T$ como en el ejercicio 9, encontrar la matriz de $T$ en esta base.

**11.** Si $T=\begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}$, calcular $T^n$ para todo $n\in\mathbb{N}$. Demostrar que $\lambda=0$ es el único autovalor de un operador nilpotente (ver ejercicio 4).

## 7.7. OBTENCION DE LA FORMA DE JORDAN DE UNA MATRIZ

En esta sección daremos algunos ejemplos de obtención de la forma de Jordan de una matriz dada; el algoritmo que utilizaremos se basa en la demostración dada en la sección anterior. Para aquellas personas que no hayan leído la sección anterior realizamos a continuación un resumen del «algoritmo» que de ella se sigue.

1) Dada una matriz $A$ se comienza calculando los autovalores de $A$; sean éstos $\lambda_1, ..., \lambda_m$ con multiplicidades $r_1, ..., r_m$, respectivamente, tal que $r_1+\cdots+r_m=n$.

2) Para cada autovalor $\lambda$ se calcula la cadena de subespacios

$$E_1(\lambda)=\ker(A-\lambda I)\subsetneqq E_2(\lambda)=\ker(A-\lambda I)^2\subsetneqq$$
$$\subsetneqq E_p(\lambda)=\ker(A-\lambda I)^p=E_{p+1}(\lambda)=\ker(A-\lambda I)^{p+1}=\cdots$$

formada por los núcleos de las potencias sucesivas de $(A-\lambda I)$; la cadena anterior se estabiliza después de un cierto número de pasos $p$ (el subespacio $E_p(\lambda)$ se denomina *autoespacio máximo* asociado a $\lambda$). El valor de $p$ puede obtenerse sabiendo que $\dim E_p(\lambda)=$ multiplicidad de $\lambda$.

*Nota.* Es conveniente ahora observar la figura 1 de la sección 7.6.

3) Tomar el mayor número posible de vectores linealmente independientes de $E_p(\lambda)-E_{p-1}(\lambda)$; sean éstos

$$F_p=\{\vec{u}_1, ..., \vec{u}_r\}.$$

Las imágenes sucesivas mediante $(A-\lambda I)$ de *cada uno* de los vectores de $F_p$ forman una matriz elemental de Jordan: $J_p(\lambda)$. En el subespacio generado por $F_p$ y las potencias sucesivas mediante $(A-\lambda I)$ de los elementos de $F_p$ la matriz de $A$ respecto a la base

$$B_p(\lambda)=(A-\lambda I)^{p-1}(F_p)\cup\cdots\cup(A-\lambda I)(F_p)\cup F_p$$

está formada por la yuxtaposición de $r$ matrices elementales de Jordan de orden $p$ dispuestas sobre la diagonal principal.

4) Se elige a continuación el mayor número posibles de vectores de $E_{p-1}(\lambda)-E_{p-2}(\lambda)$ que sean linealmente independientes entre sí y linealmente independientes con la colección $B_p(\lambda)$; sean éstos

$$F_{p-1}=\{\vec{u}_{r+1}, ..., \vec{u}_l\}.$$

Las imágenes sucesivas de cada uno de estos vectores mediante $(A-\lambda I)$ nos dan un conjunto de vectores linealmente independientes $B_{p-1}(\lambda)$ con respecto a los cuales la matriz de $A$ es una yuxtaposición en la diagonal principal de $r-l$ matrices de Jordan elementales de orden $p-1$.

5) Se continúa realizando el procedimiento descrito en 4) hasta llegar a $E_1(\lambda)$ en donde se eligen, si es posible, los vectores necesarios para que junto con todos los vectores anteriormente elegidos se tenga una base $B(\lambda)$ del autoespacio máximo $E_p(\lambda)$.

6) Los pasos desde el 2) hasta el 5) se repiten con cada uno de los autovalores $\lambda_j$; como los autoespacios máximos son invariantes respecto de $A$ (lema 3, sección 7.6) la matriz de Jordan de $A$ es la yuxtaposición de las matrices de Jordan de $A$ restringida a $E_{p_i}(\lambda_i)$. El cambio de base viene dado por todos los vectores anteriormente elegidos.



EJEMPLO A. Tratemos de encontrar una forma de Jordan $J$ de

$$A=\begin{pmatrix}1&0&2&-6\\0&1&-1&3\\0&0&1&3\\0&0&0&2\end{pmatrix}$$

y una matriz $P$ tal que $AP=PJ$.

Puesto que $|A-\lambda I|=(\lambda-1)^3(\lambda-2)$, los autovalores de $A$ son $\lambda=1$ (triple) y $\mu=2$ (simple). Calculamos $E_1(1)=\ker(A-I)$:

$$\begin{pmatrix}0&0&2&-6\\0&0&-1&3\\0&0&0&3\\0&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow\begin{matrix}x_3=0\\x_4=0\end{matrix}$$

$$E_1(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix}\right\}$$

Calculamos $E_2(1)=\ker(A-I)^2$:

$$\begin{pmatrix}0&0&2&-6\\0&0&-1&3\\0&0&0&3\\0&0&0&1\end{pmatrix}^2\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow\begin{pmatrix}0&0&0&0\\0&0&0&0\\0&0&0&3\\0&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow x_4=0$$

$$E_2(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\right\}$$

Como $(A-I)^3=(A-I)^2$, $E_2(1)$ es el autoespacio máximo correspondiente a $\lambda=1$.

Tomamos

$$\vec{u}_2=\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\in E_2(1)-E_1(1)$$

$\vec{u}_1=(A-I)\vec{u}_2=\begin{pmatrix}2\\-1\\0\\0\end{pmatrix}$ ; elegimos $\vec{u}_3=\begin{pmatrix}1\\0\\0\\0\end{pmatrix}$, que es de $E_1(1)$ y es linealmente independiente con $\vec{u}_1$ y $\vec{u}_2$.

Finalmente calculamos $E_1(2)=\ker(A-2I)$:

$$\begin{pmatrix}-1&0&2&-6\\0&-1&-1&3\\0&0&-1&3\\0&0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow\left\{\begin{array}{r}-x_1+2x_3-6x_4=0\\-x_2-x_3+3x_4=0\\x_3=3x_4\end{array}\right\}\quad\begin{array}{r}x_1=0\\x_2=0\\x_3=3x_4\end{array}$$

$$E_1(2)=L\left\{\begin{pmatrix}0\\0\\3\\1\end{pmatrix}\right\}$$

Tomando $\vec{u}_4=\begin{pmatrix}0\\0\\3\\1\end{pmatrix}$ se tiene una base del espacio vectorial; por tanto:

$$P=\begin{pmatrix}2&0&1&0\\-1&0&0&0\\0&1&0&3\\0&0&0&1\end{pmatrix}\quad;\quad J=\left(\begin{array}{cc|c|c}1&1&0&0\\0&1&0&0\\\hline 0&0&1&0\\\hline 0&0&0&2\end{array}\right)$$

EJEMPLO B. Reducir la matriz $A=\begin{pmatrix}1&2&3&0&0\\0&1&2&0&0\\0&0&1&2&0\\0&0&0&2&1\\0&0&0&0&1\end{pmatrix}$ a su forma de Jordan y encontrar $P$ tal que $AP=PJ$.

Puesto que $|A-\lambda I|=(\lambda-1)^4(\lambda-2)$, los autovalores de $A$ son $\lambda=1$ (cuádruple), $\mu=2$.



Calculamos $E_1(1)=\ker(A-I)$:

$$\left\{\begin{array}{r}2x_2+3x_3=0\\2x_3=0\\2x_4=0\\x_4+x_5=0\end{array}\right\}\Leftrightarrow x_2=x_3=x_4=x_5=0$$

$$E_1(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix}\right\}$$

Calculamos $E_2(1)=\ker(A-I)^2$:

$$(A-I)^2=\begin{pmatrix}0&0&4&6&0\\0&0&0&4&0\\0&0&0&2&2\\0&0&0&1&1\\0&0&0&0&0\end{pmatrix}\quad;\quad\begin{array}{r}4x_3+6x_4=0\\x_4=0\\x_5=0\end{array}$$

$$E_2(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix}\right\}.$$

Calculamos $E_3(1)=\ker(A-I)^3$:

$$(A-I)^3=\begin{pmatrix}0&0&0&14&6\\0&0&0&4&4\\0&0&0&2&2\\0&0&0&1&1\\0&0&0&0&0\end{pmatrix},\ x_4=x_5=0$$

$$E_3(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix}\right\}.$$

Calculamos $E_4(1)=\ker(A-I)^4$:

$$(A-I)^4=\begin{pmatrix}0&0&0&14&14\\0&0&0&4&4\\0&0&0&2&2\\0&0&0&1&1\\0&0&0&0&0\end{pmatrix},\ x_4=-x_5$$

$$E_4(1)=L\left\{\begin{pmatrix}1\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\1\\-1\end{pmatrix}\right\}.$$

Como $E_5(1)=E_4(1)$, ya que $(A-I)^5=(A-I)^4$, $E_4(1)$ es el autoespacio máximo. Esto tiene que ser así porque el subespacio $\ker(A-2I)$ es al menos de dimensión 1.

Tomamos

$$\vec{u}_4=\begin{pmatrix}0\\0\\0\\1\\-1\end{pmatrix}\ ;\ \vec{u}_3=(A-I)\vec{u}_4=\begin{pmatrix}0\\0\\2\\0\\0\end{pmatrix}\ ;\ \vec{u}_2=(A-I)\vec{u}_3=\begin{pmatrix}6\\4\\0\\0\\0\end{pmatrix}\ ;\ \vec{u}_1=(A-I)\vec{u}_2=\begin{pmatrix}8\\0\\0\\0\\0\end{pmatrix}$$

Calculamos $\ker(A-2I)=E_1(2)$:

$$\begin{pmatrix}-1&2&3&0&0\\0&-1&2&0&0\\0&0&-1&2&0\\0&0&0&0&1\\0&0&0&0&-1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\\0\end{pmatrix}\Leftrightarrow\begin{array}{r}-x_1+2x_2+3x_3=0\\-x_2+2x_3=0\\-x_3+2x_4=0\\x_5=0\end{array}$$

$$E_1(2)=L\left\{\begin{pmatrix}14\\4\\2\\1\\0\end{pmatrix}\right\}.$$



Por tanto,

$$J=\left(\begin{array}{cccc|c}1&1&&&\\&1&1&&\\&&1&1&\\&&&1&\\\hline&&&&2\end{array}\right),\quad P=\begin{pmatrix}8&6&0&0&14\\0&4&0&0&4\\0&0&2&0&2\\0&0&0&1&1\\0&0&0&-1&0\end{pmatrix}$$

* * *

EJEMPLO C. Tratemos de hallar la forma de Jordan de la matriz

$$A=\begin{pmatrix}0&1&-1&0\\0&-1&0&1\\1&1&0&0\\0&-2&0&1\end{pmatrix}.$$

Sus autovalores satisfacen la ecuación $(\lambda^2+1)(\lambda^2+1)=0$ y, por tanto, son $\lambda=i$, $\mu=-i$ (ambos dobles).

Calculamos $E_1(i)=\ker(A-iI)$:

$$\begin{pmatrix}-i&1&-1&0\\0&-1-i&0&1\\1&1&-i&0\\0&-2&0&1-i\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\\z_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow\begin{array}{r}-iz_1+z_2-z_3=0\\z_1+z_2-iz_3=0\\z_4=(1+i)z_2\end{array}$$

$$E_1(i)=L\left\{\begin{pmatrix}i\\0\\1\\0\end{pmatrix}\right\}.$$

Calculamos $E_2(i)=\ker(A-iI)^2$:

$$(A-\lambda I)^2=\begin{pmatrix}-2&-2-2i&2i&1\\0&-2+2i&0&-2i\\-2i&-2i&-2&1\\0&4i&0&-2-2i\end{pmatrix}\begin{pmatrix}z_1\\z_2\\z_3\\z_4\end{pmatrix}=\begin{pmatrix}0\\0\\0\\0\end{pmatrix}\Leftrightarrow\begin{array}{r}(-1+i)z_2=iz_4\\2z_1+(1+i)z_2-2iz_3=0\end{array}$$

$$E_2(i)=L\left\{\begin{pmatrix}i\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\1+i\\1\\2i\end{pmatrix}\right\}.$$

Puesto que $\dim E_2(i)=2=$multiplicidad de $i$, aquí se «estabiliza» la cadena de subespacios. Tomamos

$$\vec{u}_2=\begin{pmatrix}0\\1+i\\1\\2i\end{pmatrix}, \quad \vec{u}_1=(A-iI)\vec{u}_2=\begin{pmatrix}-i&1&-1&0\\0&-1-i&0&1\\1&1&-i&0\\0&-2&0&1-i\end{pmatrix}\begin{pmatrix}0\\1+i\\1\\2i\end{pmatrix}=\begin{pmatrix}i\\0\\1\\0\end{pmatrix}.$$

Para calcular $E_1(-i)=\ker(A+iI)$ observamos que $A+iI=\overline{A-iI}$ y, por tanto, el sistema $(A+iI)\vec{z}=\vec{0}$ es equivalente al sistema $(A-iI)\overline{\vec{z}}=\vec{0}$, que coincide con el sistema utilizado para calcular $\ker(A-iI)$ excepto que $\vec{z}$ es sustituido por $\overline{\vec{z}}$. Entonces

$$E_1(-i)=L\left\{\begin{pmatrix}-i\\0\\1\\0\end{pmatrix}\right\}.$$

De manera similar se tiene

$$E_2(-i)=L\left\{\begin{pmatrix}-i\\0\\1\\0\end{pmatrix}, \begin{pmatrix}0\\1-i\\1\\-2i\end{pmatrix}\right\},$$

y podemos elegir $\vec{u}_4$ y $\vec{u}_3$ como los conjugados de $\vec{u}_2$ y $\vec{u}_1$, respectivamente; es decir,

$$\vec{u}_4=\begin{pmatrix}0\\1-i\\1\\-2i\end{pmatrix}, \quad \vec{u}_3=\begin{pmatrix}-i\\0\\1\\0\end{pmatrix}.$$

Puesto que $A(\vec{u}_1)=i\vec{u}_1$, $A(\vec{u}_2)=\vec{u}_1+i\vec{u}_2$, $A(\vec{u}_3)=-i\vec{u}_3$, $A(\vec{u}_4)=\vec{u}_3-i\vec{u}_4$, la matriz de Jordan de $A$ es

$$J=\left(\begin{array}{cc|cc}i&1&0&0\\0&i&0&0\\\hline 0&0&-i&1\\0&0&0&-i\end{array}\right)$$



## EJERCICIOS 7.7

En los ejercicios siguientes hallar la forma de Jordan de la matriz dada y la matriz de paso.

**1.** $A=\begin{pmatrix}6&-9&5&4\\7&-13&8&7\\8&-17&11&8\\1&-2&1&3\end{pmatrix}.$

**2.** $B=\begin{pmatrix}3&1&-1&-1\\0&2&0&1\\1&1&1&0\\0&0&0&2\end{pmatrix}$

**3.** $C=\begin{pmatrix}2&0&0&-2\\5&2&-2&2\\3&1&0&-3\\1&0&0&0\end{pmatrix}$

## 7.8. FORMA DE JORDAN REAL DE MATRICES REALES CON AUTOVALORES COMPLEJOS

Aunque una aplicación lineal esté definida en un espacio vectorial real, su forma de Jordan, obtenida como en las secciones 6 y 7, puede ser compleja. Una muestra se tiene en el ejemplo C de la sección anterior. En esta sección trataremos de dar una «forma de Jordan» real de toda aplicación lineal definida en un espacio vectorial real.

Si la forma de Jordan de una matriz $A$ obtenida mediante el procedimiento de las secciones 6 y 7 es real, a ésta se le llama *forma de Jordan real de A;* si, por el contrario, alguno de los autovalores de $A$ es complejo se dará en esta sección su *forma de Jordan real.*

Comencemos con el ejemplo C de la sección anterior. Sus autovalores eran $\lambda=i$, $\mu=-i$ (dobles); la forma de Jordan de $A$ es

$$J=\left(\begin{array}{cc|cc}i&1&0&0\\0&i&0&0\\\hline 0&0&-i&1\\0&0&0&-i\end{array}\right)$$

y una base $B$ es $\{\vec{u}_1, \vec{u}_2, \bar{\vec{u}}_1, \bar{\vec{u}}_2\}$ (base del espacio vectorial complejo $V_{\mathbb{C}}$ obtenido a partir del espacio vectorial real $V$ (ver ejercicio 4 al final de esta sección)), donde

$$\vec{u}_1=\begin{pmatrix} i\\ 0\\ 1\\ 0\end{pmatrix}, \quad \vec{u}_2=\begin{pmatrix} 0\\ 1+i\\ 1\\ 2i\end{pmatrix}.$$

Tomar

$$\vec{v}_1=\text{real } \vec{u}_1=\begin{pmatrix} 0\\ 0\\ 1\\ 0\end{pmatrix}, \quad \vec{w}_1=\text{img } \vec{u}_1=\begin{pmatrix} 1\\ 0\\ 0\\ 0\end{pmatrix}$$

$$\vec{v}_2=\text{real } \vec{u}_2=\begin{pmatrix} 0\\ 1\\ 1\\ 0\end{pmatrix}, \quad \vec{w}_2=\text{img } \vec{u}_2=\begin{pmatrix} 0\\ 1\\ 0\\ 2\end{pmatrix};$$

se observa que $\{\vec{v}_1, \vec{w}_1, \vec{v}_2, \vec{w}_2\}$ es una base (real) de $V$; en esta base se tiene

$$A\vec{v}_1=\begin{pmatrix} -1\\ 0\\ 0\\ 0\end{pmatrix}=-\vec{w}_2 \quad ; \quad A\vec{w}_1=\begin{pmatrix} 0\\ 0\\ 1\\ 0\end{pmatrix}=\vec{v}_1$$

$$A\vec{v}_2=\begin{pmatrix} 0\\ -1\\ 1\\ -2\end{pmatrix}=\vec{v}_1-\vec{w}_2 \quad ; \quad A\vec{w}_2=\begin{pmatrix} 1\\ 1\\ 1\\ 0\end{pmatrix}=\vec{w}_1+\vec{v}_2.$$

Por tanto,

$$\begin{pmatrix} 0&1&0&0\\ 0&0&1&1\\ 1&0&1&0\\ 0&0&0&2\end{pmatrix}^{-1} A \begin{pmatrix} 0&1&0&0\\ 0&0&1&1\\ 1&0&1&0\\ 0&0&0&2\end{pmatrix}=\left(\begin{array}{cc|cc} 0&1&1&0\\ -1&0&0&1\\ \hline 0&0&0&1\\ 0&0&-1&0\end{array}\right)=J$$

que es una forma de Jordan real de $A$.



Observar que $J$ está formada por dos matrices iguales $B$ yuxtapuestas sobre la diagonal principal y la matriz identidad encima de ellas, esto es:

$$J=\begin{pmatrix} B & I\\ O & B\end{pmatrix};$$

además, $B$ es la matriz del subespacio invariante real de dos dimensiones asociado con el autovalor $\lambda=i$ como en la proposición 2 de la sección 7.5.

* * *

Sea $A$ una aplicación lineal en un espacio vectorial *real* $V$ con al menos un autovalor complejo. Definiendo en $V\times V$ la suma

$$(x_1, y_1)+(x_2, y_2)=(x_1+x_2, y_1+y_2)$$

y la multiplicación por números complejos

$$(a+ib)(x, y)=(ax-by, ay+bx),$$

$V\times V$ se convierte en un espacio vectorial complejo, que se denota por $V_{\mathbb{C}}$ (ver ejercicio 4). Los elementos de este espacio vectorial complejo se denotan mediante

$$(x, y)\equiv \vec{x}+i\vec{y}.$$

Definiendo $A_{\mathbb{C}}(\vec{x}+i\vec{y})=A(\vec{x})+iA(\vec{y})$, $A_{\mathbb{C}}$ es una aplicación lineal en $V_{\mathbb{C}}$.

Toda base de $V$ es una base de $V_{\mathbb{C}}$ y, por tanto, $A_{\mathbb{C}}$ posee la misma matriz que $A$ en esta base. Así pues, la matriz $A$ es equivalente en $V_{\mathbb{C}}$ a una matriz de Jordan $J$ y $V_{\mathbb{C}}$ puede escribirse como una suma directa de autoespacios máximos

$$V_{\mathbb{C}}=E_{p_1}(\lambda_1)\oplus\cdots\oplus E_{p_m}(\lambda_m). \qquad (1)$$

Si $\lambda=\alpha+i\beta$ es un autovalor complejo de $A$, con $\beta\neq 0$, $\bar{\lambda}=\alpha-i\beta$ es también un autovalor de $A$ puesto que $A$ es real: en efecto, si $|A-\lambda I|=0$ se tiene que

$$0=\overline{|A-\lambda I|}=|\bar{A}-\bar{\lambda}I|=|A-\bar{\lambda}I|.$$

Además, las ecuaciones $(A-\lambda I)^k\vec{z}=0$ son equivalentes a $(A-\bar{\lambda}I)^k\bar{\vec{z}}=0$ y, por tanto, las bases de los autoespacios máximos $E_p(\lambda)$ y $E_p(\bar{\lambda})$ pueden elegirse conjugadas.

Sean

$$B=\{\vec{u}_k=(A-\lambda I)\vec{u}_{k-1}, \ldots, \vec{u}_2=(A-\lambda I)\vec{u}_1, \vec{u}_1\}$$

y

$$\bar{B}=\{\bar{\vec{u}}_k, \ldots, \bar{\vec{u}}_2, \bar{\vec{u}}_1\}$$

dos colecciones de vectores linealmente independientes que originan dos matrices elementales de Jordan:

$$J_k(\lambda), \quad J_k(\bar{\lambda})$$

de orden $k$. Debido a (1), basta encontrar una «forma de Jordan real» que agrupe estas matrices para tener una «forma de Jordan real» de $A$.

Tomemos los vectores reales que son las partes real e imaginaria de los elementos de $B$, es decir,

$$\begin{array}{ll} \vec{v}_k = \text{real } \vec{u}_k \quad , & \vec{w}_k = \text{img } \vec{u}_k \\ \vec{v}_{k-1} = \text{real } \vec{u}_{k-1} \ , & \vec{w}_{k-1} = \text{img } \vec{u}_{k-1} \\ \quad \vdots & \quad \vdots \\ \vec{v}_2 = \text{real } \vec{u}_2 \quad , & \vec{w}_2 = \text{img } \vec{u}_2 \\ \vec{v}_1 = \text{real } \vec{u}_1 \quad , & \vec{w}_1 = \text{img } \vec{u}_1 \end{array} \tag{2}$$

El conjunto de vectores de (2) es linealmente independiente; en efecto, si tenemos

$$\sum_{j=1}^{k} \alpha_j \vec{v}_j + \sum_{j=1}^{k} \beta_j \vec{w}_j = \vec{0},$$

puesto que $\vec{v}_j = \dfrac{\vec{u}_j + \bar{\vec{u}}_j}{2}$ y $\vec{w}_j = \dfrac{\vec{u}_j - \bar{\vec{u}}_j}{2i}$ se tiene que

$$\begin{aligned} \vec{0} &= \sum_{j=1}^{k} \frac{\alpha_j}{2} (\vec{u}_j + \bar{\vec{u}}_j) + \sum_{j=1}^{k} \frac{\beta_j}{2i} (\vec{u}_j - \bar{\vec{u}}_j) = \\ &= \sum_{j=1}^{k} \frac{\alpha_j - i\beta_j}{2} \vec{u}_j + \sum_{j=1}^{k} \frac{\alpha_j + i\beta_j}{2} \bar{\vec{u}}_j \end{aligned}$$

como los elementos de $B \cup \bar{B}$ son linealmente independientes (¿por qué?) se tiene que $\alpha_j - i\beta_j = 0$, $\alpha_j + i\beta_j = 0$, de donde se deduce que $\alpha_j = \beta_j = 0$, $j = 1, \ldots, k$.

Encontraremos ahora la matriz de $A$ restringida al espacio vectorial generado por los elementos de (2). Puesto que $A_{\mathbb{C}}(\vec{u}_k) = \lambda \vec{u}_k$ tenemos

$$\begin{aligned} A(\vec{v}_k) + iA(\vec{w}_k) &= A_{\mathbb{C}}(\vec{u}_k) = \lambda \vec{u}_k = (\alpha + i\beta)(\vec{v}_k + i\vec{w}_k) = \\ &= (\alpha \vec{v}_k - \beta \vec{w}_k) + i(\beta \vec{v}_k + \alpha \vec{w}_k) \end{aligned}$$

e igualando las partes reales e imaginarias obtenemos

$$A(\vec{v}_k) = \alpha \vec{v}_k - \beta \vec{w}_k \quad ; \quad A(\vec{w}_k) = \beta \vec{v}_k + \alpha \vec{w}_k \tag{3}$$



Para $\vec{u}_{k-1}$ tenemos que $A_{\mathbb{C}}(\vec{u}_{k-1}) = \vec{u}_k + \lambda \vec{u}_{k-1}$ y, por tanto,

$$\begin{aligned} A(\vec{v}_{k-1}) + iA(\vec{w}_{k-1}) &= A_{\mathbb{C}}(\vec{u}_{k-1}) = \vec{u}_k + \lambda \vec{u}_{k-1} = \\ &= (\vec{v}_k + i\vec{w}_k) + (\alpha + i\beta)(\vec{v}_{k-1} + i\vec{w}_{k-1}) = \\ &= (\vec{v}_k + \alpha \vec{v}_{k-1} - \beta \vec{w}_{k-1}) + i(\vec{w}_k + \beta \vec{v}_{k-1} + \alpha \vec{w}_{k-1}), \end{aligned}$$

de donde se deduce que

$$A(\vec{v}_{k-1}) = \vec{v}_k + \alpha \vec{v}_{k-1} - \beta \vec{w}_{k-1} \quad ; \quad A(\vec{w}_{k-1}) = \vec{w}_k + \beta \vec{v}_{k-1} + \alpha \vec{w}_{k-1}. \tag{4}$$

Por el mismo razonamiento se obtienen expresiones análogas a estas últimas para $A(\vec{v}_j)$ y $A(\vec{w}_j)$ con $j = 1, 2, \ldots, k-2$, es decir,

$$A(\vec{v}_j) = \vec{v}_{j-1} + \alpha \vec{v}_j - \beta \vec{w}_j \quad ; \quad A(\vec{w}_j) = \vec{w}_{j-1} + \beta \vec{v}_j + \alpha \vec{w}_j. \tag{5}$$

Por tanto, la matriz de $A$ restringida al espacio vectorial generado por (2), en la base (2), ordenada de izquierda a derecha y de arriba a abajo, es

$$\left(\begin{array}{cc|cc|cc|cc} \alpha & \beta & 1 & 0 & & & & \\ -\beta & \alpha & 0 & 1 & & & & \\ \hline & & \alpha & \beta & 1 & 0 & & \\ & & -\beta & \alpha & 0 & 1 & & \\ \hline & & & & \ddots & & 1 & 0 \\ & & & & & & 0 & 1 \\ \hline & & & & & & \alpha & \beta \\ & & & & & & -\beta & \alpha \end{array}\right).$$

Esta matriz está formada por la yuxtaposición sobre su diagonal principal de las matrices del subespacio invariante *real* de $\lambda$ con matrices identidad de orden 2 encima de la diagonal principal, y ceros en el resto.

La matriz de *Jordan real* de $A$ se obtiene mediante yuxtaposición de matrices del tipo anterior y matrices elementales de Jordan correspondientes a autovalores reales.

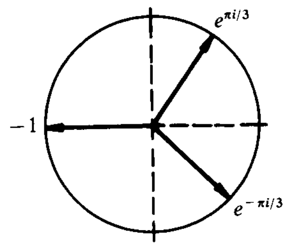
EJEMPLO A. Encontrar la forma de Jordan real $J$ de

$$A = \begin{pmatrix} 2 & 0 & -3 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

y una matriz real $R$ tal que $AR = RJ$.

La ecuación característica de $A$ es $-\lambda^3-1=0$, con lo que sus autovalores son las soluciones de $\lambda^3=-1$, es decir,

$$\lambda_1=-1, \quad \lambda_2=e^{\pi i/3}=\frac{1+i\sqrt{3}}{2}, \quad \bar{\lambda}=e^{-\pi i/3}=\frac{1-i\sqrt{3}}{2}$$




A $\lambda_1=-1$ le corresponde la matriz elemental de Jordan $J_1(-1)=(-1)$ y al par de autovalores conjugados $\lambda_2$ y $\bar{\lambda}_2$ le corresponde la matriz

$$\begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix},$$

con lo que

$$J=\left(\begin{array}{c|cc} -1 & 0 & 0 \\ \hline 0 & \frac{1}{2} & \frac{\sqrt{3}}{2} \\ 0 & \frac{-\sqrt{3}}{2} & \frac{1}{2} \end{array}\right).$$

Para encontrar $R$ calculamos $E_1(-1)=\ker(A+I)$ y $E_1\left(\frac{1+i\sqrt{3}}{2}\right)=$ $=\ker\left(A-\left(\frac{1}{2}+\frac{i\sqrt{3}}{2}\right)I\right)$:

$$\begin{pmatrix} 3 & 0 & -3 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} 3x_1=3x_3 \\ x_1=0 \end{matrix} \Rightarrow E_1(-1)=L\left\{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\}$$



$$\begin{pmatrix} \frac{3}{2}-\frac{\sqrt{3}}{2}i & 0 & -3 \\ 0 & -\frac{3}{2}-\frac{\sqrt{3}}{2}i & 0 \\ 1 & 0 & -\frac{3}{2}-\frac{\sqrt{3}}{2}i \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{matrix} z_2=0 \\ z_1=\left(\frac{3}{2}+\frac{\sqrt{3}}{2}i\right)z_3 \end{matrix} \Rightarrow$$

$$E_1\left(\frac{1+i\sqrt{3}}{2}\right)=L\left\{\begin{pmatrix} \frac{3}{2}+\frac{\sqrt{3}}{2}i \\ 0 \\ 1 \end{pmatrix}\right\}.$$

Tomando

$$\vec{u}_1=\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_1=\begin{pmatrix} 3/2 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{w}_1=\begin{pmatrix} \sqrt{3}/2 \\ 0 \\ 0 \end{pmatrix}$$

se tiene

$$R=\begin{pmatrix} 0 & \frac{3}{2} & \frac{\sqrt{3}}{2} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

El lector puede comprobar ahora que en la base $\{\vec{u}_1, \vec{v}_1, \vec{w}_1\}$ la aplicación dada por $A$ tiene como matriz $J$.

## EJERCICIOS 7.8

**1.** Encontrar la forma de Jordan real de las siguientes matrices:

*a)* $\begin{pmatrix} 0 & 1 & -1 \\ 0 & 1 & 0 \\ 3 & 1 & 0 \end{pmatrix}$ *b)* $\begin{pmatrix} 1 & -2 \\ 3 & 1 \end{pmatrix}$ *c)* $\begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$

**2.** Sea $G=\begin{pmatrix} \alpha & -\beta & 0 \\ \beta & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$ la matriz de $G\in L(\mathbb{R}^3)$ en la base canónica, con $\alpha^2+\beta^2=1$.

a) Hallar los subespacios invariantes de $G$, así como los autovalores.

**3.** Sea $G=\begin{pmatrix}\alpha & -\beta\\ \beta & \alpha\end{pmatrix}$ con $\alpha^2+\beta^2=1$ la matriz de $G\in L(\mathbb{R}^2)$ en la base canónica.

a) Hallar los autovalores y autovectores de $G$.
b) Probar que $G$ es un giro de ángulo $\varphi$ determinado por $\alpha=\cos\varphi$, $\beta=\operatorname{sen}\varphi$.
c) Si $z=u+iv\in\mathbb{C}^2$ es autovector de $G$, probar que la matriz de $G$ en la base $\{\overline{u},\overline{v}\}$ es $A^t$ o $A$ dependiendo de que el autovalor correspondiente sea $\lambda=\alpha+i\beta$ o $\overline{\lambda}=\alpha-i\beta$.

**4.** a) Sea $V$ un espacio vectorial real y sea $V_{\mathbb{C}}=\{x_1+ix_2/x_1,\ x_2\in V\}$ con las operaciones

$$(x_1+ix_2)+(y_1+iy_2)=(x_1+y_1)+i(x_2+y_2)$$
$$(a+ib)(x_1+ix_2)=(ax_1-bx_2)+i(bx_1+ax_2)$$

Demostrar que $V_{\mathbb{C}}$ es un espacio vectorial complejo.
b) Si $\mathscr{A}\in L(V)$, demostrar que $\mathscr{A}_{\mathbb{C}}$ definida en $V_{\mathbb{C}}$ por $\mathscr{A}_{\mathbb{C}}(x_1+ix_2)=\mathscr{A}(x_1)+i\mathscr{A}(x_2)$ es una aplicación lineal en $V_{\mathbb{C}}$.

## 7.9. EL TEOREMA DE CAYLEY-HAMILTON

En 1858 Arthur Cayley (1821-1895) enunció un resultado que más tarde sería conocido como el teorema de Cayley-Hamilton, y que él demostró solamente para matrices de orden 2.

Dado un polinomio $p(x)$ con coeficientes complejos,

$$p(x) = \sum_{k=0}^{n} a_k x^k = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

y una matriz $A$ de orden $n$, definimos

$$p(A) = \sum_{k=0}^{n} a_k A^k = a_n A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 A^0$$

donde $A^0 = I$. Es decir, se reemplaza $x$ por la matriz $A$ teniendo en cuenta que $A^0 = I$. Si $p(A)$ es la matriz nula decimos que $A$ satisface el polinomio $p(x)$.

Ejemplo A. La matriz $A=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$ satisface el polinomio $p(x) = x^2 + 1$, ya que

$$\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}^2 + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix} + \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$



Observar que el polinomio característico de la matriz $A$ del ejemplo anterior es $p(\lambda) = \lambda^2 + 1$ y, por tanto, la matriz $A$ satisface su polinomio característico. Este resultado, para todas las matrices de orden 2, es el que demostró Cayley.

Teorema de Cayley-Hamilton para matrices de orden 2

Toda matriz de orden 2 satisface su polinomio característico.

*Demostración.* Si $A=\begin{pmatrix}a & b\\ c & d\end{pmatrix}$ su polinomio característico es

$$p(x) = |A - xI| = \begin{vmatrix}a-x & b\\ c & d-x\end{vmatrix} = x^2 - (a+d)x + (ad - cb).$$

Como

$$A^2 - (a+d)A + (ad-cb)I =$$
$$= \begin{pmatrix}a^2+bc & ab+bd\\ ca+cd & cb+d^2\end{pmatrix} - \begin{pmatrix}a^2+ad & ab+bd\\ ac+db & ad+d^2\end{pmatrix} + (ad-cb)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

se obtiene el resultado deseado.

El teorema de Cayley-Hamilton para matrices de orden $n$ es similar al enunciado anteriormente y fue demostrado por Hamilton. La idea de la demostración es conseguir demostrar el teorema para matrices elementales de Jordan y a partir de aquí llegar al resultado para una matriz general haciendo uso del teorema de clasificación de la sección 7.6. Antes de enunciar y demostrar el resultado concreto daremos varios lemas.

Lema 1.

Si $A$ es una matriz de orden $h$ cuyos elementos son todos nulos excepto los que están encima de la diagonal principal que son unos se tiene que $A^h = 0$.

*Demostración.* Si $h = 2$,

$$\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}^2 = \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0\\ 0 & 0\end{pmatrix}$$

Si $h = 3$,

$$\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix}^3 = \begin{pmatrix}0 & 0 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & 0 & 0\end{pmatrix} = \begin{pmatrix}0 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix}$$

Observar que el efecto que se produce al multiplicar una matriz $A$ por sí misma es desplazar la línea de unos hacia arriba.

Para realizar la demostración procedemos por inducción en el orden de $A$. Si el resultado se cumple para toda matriz de orden $h$:

$$\left(\begin{array}{c|ccccc} 0 & 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{array}\right)^{h+1}_{(h+1)} = \left(\begin{array}{c|ccccc} 0 & 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{array}\right)^{h} \left(\begin{array}{c|ccccc} 0 & 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{array}\right) =$$

$$= \left(\begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 1 \\ \hline 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{array}\right) \left(\begin{array}{c|ccccc} 0 & 1 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{array}\right) = \left(\begin{array}{c|cccc} 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \end{array}\right)$$

en donde la penúltima igualdad ha hecho uso de la hipótesis de inducción.

**Lema 2.**

> Si $J_k(\lambda)$ es una matriz elemental de Jordan, $J_k(\lambda)$ satisface su polinomio característico.

*Demostración.* El polinomio característico de $J_k(\lambda)$ es $P_{J_k(\lambda)}(x) = (\lambda - x)^k$. Entonces

$$P_{J_k(\lambda)}(J_k(\lambda)) = (\lambda I - J_k(\lambda)) = \left[\begin{pmatrix} \lambda & & 0 \\ & \lambda & \\ 0 & & \ddots \; \lambda \end{pmatrix} - \begin{pmatrix} \lambda & 1 & 0 \\ & \ddots & 1 \\ 0 & & \lambda \end{pmatrix}\right] =$$

$$= \begin{pmatrix} 0 & -1 & \cdots & 0 \\ & 0 & -1 & \vdots \\ & 0 & \ddots & -1 \\ & & & 0 \end{pmatrix}^k = (-1)^k \begin{pmatrix} 0 & 1 & \cdots & 0 \\ & 0 & \ddots & \vdots \\ & 0 & \ddots & 1 \\ & & & 0 \end{pmatrix}$$

donde la última igualdad es debida al resultado probado en el lema 1. ■



**Lema 3.**

> Si $J$ es una matriz de Jordan, $J$ satisface su polinomio característico.

*Demostración.* Una matriz de Jordan es de la forma

$$J = \begin{pmatrix} J_{k_1}(\lambda_1) & & & 0 \\ & J_{k_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{k_l}(\lambda_2) \end{pmatrix}$$

y, por tanto, un polinomio característico es $P_J(x) = \prod_{i=1}^{l} (x - \lambda_i)^{k_i}$. Como $(J - \lambda_i I)^{k_i}$ tiene la caja $i$ nula debido al lema 2, al calcular el producto $P_J(J) = \prod_{i=1}^{l} (J - \lambda_i I)^{k_i}$ el resultado es nulo ya que las $l$ matrices que se multiplican tienen al menos una caja nula en filas y columnas diferentes. ■

Estamos ya preparados para demostrar el teorema de Cayley-Hamilton en su versión general.

**Teorema de Cayley-Hamilton**

> Toda matriz $A$ satisface su polinomio característico.

*Demostración.* Sea $P_A(x) = \sum_{l=0}^{n} a_l x^l$ el polinomio característico de la matriz $A$. Por el teorema de clasificación de Jordan demostrado en la sección 7.6, $A$ es equivalente a su matriz de Jordan, es decir, existe una matriz $P$ tal que $J = P^{-1}AP$.

Observar que $P_A(x) = P_J(x)$ ya que $A$ y $J$ son equivalentes:

$$P_A(x) = |A - xI| = |PJP^{-1} - P(xI)P^{-1}| = |P||J - xI||P^{-1}| =$$
$$= |J - xI| = P_J(x).$$

Entonces,

$$P_A = \sum_{l=0}^{n} a_l A^l = \sum_{l=0}^{n} a_l (PJP^{-1})^l = P\left(\sum_{l=0}^{n} a_l J^l\right)P^{-1} = Pp_A(J)P^{-1} = Pp_J(J)P^{-1} = 0$$

donde la última igualdad se debe al lema 3 ya que $J$ satisface su polinomio característico $P_J(x)$. ■

EJERCICIOS 7.9

**1.** Demostrar directamente que las siguientes matrices satisfacen su polinomio característico:

*a)* $$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$ *b)* $$B = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 2 & 3 & 0 & 0 \\ 0 & 0 & 4 & 5 \\ 0 & 0 & 6 & 7 \end{pmatrix}$$ *c)* $$C = \begin{pmatrix} 0 & 3 & 1 \\ 2 & -1 & -1 \\ -2 & -1 & -1 \end{pmatrix}.$$

**2.** Determinar todas las matrices $A$ tal que $A^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

**3.** Determinar todas las matrices $A$ tal que $A^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

BIOGRAFIA

**Camille Jordan** nació el 5 de enero de 1838 en Lión y murió el 20 de enero de 1922 en Milán. Fue un matemático cuyos trabajos en grupos de permutaciones y en la teoría de ecuaciones hizo posible un perfecto entendimiento de las teorías del eminente matemático francés Évariste Galois.

Jordan comenzó trabajando en geometría. Su *Traité des substitutions et des equations algébriques* (1870) (*Tratado sobre las permutaciones y ecuaciones algebraicas*), por el cual obtuvo el premio Poncelet de la Academia de Ciencias de Francia, fue fundamental en el entendimiento de la teoría de Galois sobre los grupos de permutaciones, lo cual fue aplicado para la resolución de ecuaciones algebraicas. También resolvió un problema propuesto por Niels Henrick Abel acerca de la no solubilidad de una cierta ecuación algebraica mediante radicales. Jordan publicó sus lecciones y sus investigaciones en análisis en *Cours d'analyse de l'École Polytechnique* (3 vols., 1882) (*Curso de análisis de la Escuela Politécnica*).

En la tercera edición (1909-15) de su trabajo en análisis, la cual contiene muchas más investigaciones de Jordan que la primera, trató la teoría de funciones desde un punto de vista moderno, trabajando con funciones de variación acotada; sus trabajos en este campo fueron aplicados a la curva que hoy se conoce con el nombre de «curva de Jordan». Las álgebras de Jordan se llaman así en su honor.

En topología, Jordan probó (1887) que un arco simple no divide al plano y que una curva cerrada simple divide a un plano en dos partes. A pesar de que estas afirmaciones son intuitivamente obvias, su demostración requiere sofisticados métodos.

C. Jordan fue profesor de matemáticas en la École Polytechnique (París, 1876-1912). También fue editor de la revista matemática *Journal de Mathématiques Pures et Appliquées* (1885-1922).

# EJERCICIOS DE REPASO: CAPITULOS 1 A 7



A continuación presentamos varios ejercicios que pueden servir para repasar los conceptos introducidos anteriormente. Los ejercicios del 1 al 37 están ordenados de acuerdo con el orden de los capítulos de este libro. El resto son problemas que se han propuesto en varias convocatorias a los alumnos de primer curso de las licenciaturas de Matemáticas y Físicas de la Universidad Autónoma de Madrid.

**1.** Utilizar el método de eliminación de Gauss para encontrar todas las soluciones del sistema:

$$\left.\begin{aligned} x_1+x_2-2x_3+x_4&=1\\ 2x_1-x_2+3x_4&=-2\\ 4x_1+x_2-4x_3+6x_4&=0 \end{aligned}\right\}$$

**2.** Sea $r$ el mayor número de soluciones linealmente independientes del sistema:

$$\left.\begin{aligned} x_1+2x_2+4x_3-3x_4&=0\\ 3x_1+5x_2+6x_3-4x_4&=0\\ 4x_1+5x_2-2x_3+3x_4&=0\\ 3x_1+8x_2+24x_3-19x_4&=0 \end{aligned}\right\}$$

Hallar $r$ y encontrar $r$ soluciones linealmente independientes de este sistema.

**3.** Resolver el siguiente sistema:

$$\left.\begin{aligned} x_1+x_2+x_3+x_4+x_5&=1\\ x_1\phantom{+x_2}+x_3+x_4+x_5&=2\\ x_1+x_2\phantom{+x_3}+x_4+x_5&=3\\ x_1+x_2+x_3\phantom{+x_4}+x_5&=4\\ x_1+x_2+x_3+x_4\phantom{+x_5}&=1 \end{aligned}\right\}$$

**4.** Estudiar la compatibilidad o incompatibilidad del sistema

$$\left.\begin{aligned} x+y+az&=1\\ x+by+z&=1\\ 2x+y+z&=1\end{aligned}\right\}$$

según los valores reales de $a$ y $b$ y encontrar sus soluciones en los casos en que sea compatible.

**5.** Sean $a$ y $b$ dos números reales, no nulos y distintos. Estudiar el rango de la siguiente matriz para los distintos valores de $x\in\mathbb{R}$:

$$A=\begin{pmatrix} x & b & a & b\\ b & x & b & a\\ a & b & x & b\\ b & a & b & x\end{pmatrix}$$

**6.** Hallar la inversa de la siguiente matriz siempre que sea posible:

$$A=\begin{pmatrix} x & 1 & 0 & 0\\ 0 & x & 1 & 0\\ 0 & 0 & x & 1\\ 0 & 0 & 0 & x\end{pmatrix}$$

**7.** Encontrar todos los vectores $\vec{x}\in\mathbb{R}^3$ tal que $T(\vec{x})=-\vec{x}$, donde $T(x_1, x_2, x_3)=(-2x_1+x_3, -x_2, -x_1)$.

**8.** *a*) Sean $\vec{u}_1=(1, 0, 1)$, $\vec{u}_2=(0, 1, 1)$ y $\vec{u}_3=(3, -1, a)$. Decir para qué valores de $a\in\mathbb{R}$ el conjunto $B=\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base de $\mathbb{R}^3$.

*b*) Si $B$ es la base anterior con $a=1$ y $(x_1', x_2', x_3')$ son las coordenadas del vector $(x_1, x_2, x_3)\in\mathbb{R}^3$ respecto de la base $B$, hallar la ecuación en coordenadas $x_1'$, $x_2'$, $x_3'$ del plano

$$\pi: 2x_1-x_2+3x_3=1$$

**9.** Calcular el siguiente determinante de orden $n+1$:

$$\begin{vmatrix} 1 & 1 & 1 & \cdots & 1\\ 1 & 1+a_1 & 1 & \cdots & 1\\ 1 & 1 & 1+a_2 & \cdots & 1\\ \vdots & \vdots & \vdots & & \vdots\\ 1 & 1 & 1 & \cdots & 1+a_n\end{vmatrix}$$



**10.** Probar que

$$\begin{vmatrix} a_1-b_1 & a_1-b_2 & a_1-b_3 & a_1-b_4\\ a_2-b_1 & a_2-b_2 & a_2-b_3 & a_2-b_4\\ a_3-b_1 & a_3-b_2 & a_3-b_3 & a_3-b_4\\ a_4-b_1 & a_4-b_2 & a_4-b_3 & a_4-b_4\end{vmatrix}=0$$

**11.** Si $A\in\mathcal{M}_{n\times n}(\mathbb{R})$ no es invertible, demostrar que existe $B\in\mathcal{M}_{n\times n}(\mathbb{R})$, $B\neq 0$, tal que $AB=0$.

**12.** *a*) Encontrar la ecuación del haz de rectas que pasa por el punto $P=(3, 1)$.

*b*) De todas las rectas que pasan por el punto $P$ del apartado anterior, determinar la ecuación cartesiana de aquellas cuya distancia al origen sea 1.

**13.** *a*) Encontrar el punto de intersección de las rectas

$$(x, y, z)=(0, 1, 2)+t(1, -1, 3)$$
$$(x, y, z)=(3, 0, 4)+s(1, 1, -4)$$

*b*) Hallar la ecuación cartesiana de la recta perpendicular a las dos anteriores que pasa por su punto de intersección.

**14.** Encontrar la ecuación cartesiana de la recta que pasa por $P=(1, 1, 1)$ y es paralela a los planos

$$\pi_1: x+y+z=1$$
$$\pi_2: (x, y, z)=t(1, 1, 2)+s(1, -1, 1)$$

**15.** Sean $P_1$ y $P_2$ las proyecciones ortogonales del punto $P=(-1, 2, 0)$ sobre los planos de ecuaciones

$$\pi_1: x-y-z=2$$
$$\pi_2: (x, y, z)=(1, -1, 0)+t(1, 0, 1)+s(1, 2, 0)$$

Encontrar $P_1$, $P_2$ y el área del triángulo $P_1PP_2$.

**16.** Hallar la distancia entre las rectas

$$r_1: \left.\begin{aligned} x-y+z&=1\\ 2x+y&=3\end{aligned}\right\} \quad \text{y} \quad r_2: \left.\begin{aligned} y+z&=0\\ x-2y+z&=1\end{aligned}\right\}$$

**17.** Dado el plano $\pi$: $5x-3y+4z=1$, hallar la recta paralela al vector $(5, 3, -4)$, contenida en $\pi$ y que esté más próxima al punto $(1, 0, 0)$.

**18.** Dados los puntos $A=(2, -1, 1)$ y $B=(0, 1, 2)$, la recta $r$ de ecuación $(x, y, z)=(1, 0, 0)+t(3, -2, 1)$ y el plano $\pi$: $2x+z-1=0$, hay exactamente dos paralelogramos, uno de cuyos lados es el segmento $\overline{AB}$ y cuyos dos restantes vértices están, respectivamente, en $r$ y $\pi$. Hallar los vértices de estos paralelogramos y el volumen del paralelepípedo que los tiene por caras.

**19.** Hallar la ecuación de un plano que contenga a la recta $x-1=y=z$ y diste lo mismo del origen que del punto $P=(2, 4, -4)$.

**20.** Encontrar la recta $r$ que corta y es perpendicular a las rectas $r_1$ y $r_2$ dadas por

$$r_1: \begin{cases} x=-t \\ y=2+2t \\ z=-2 \end{cases}, \quad r_2: \begin{cases} x=1-t \\ y=3-t \\ z=-3+t \end{cases}$$

**21.** Hallar la longitud de la proyección ortogonal del segmento $\overline{PQ}$, donde $P=(3, 1, 2)$ y $Q=(-5, 0, 1)$, sobre el plano $3x+2z=0$.

**22.** En los siguientes casos hallar los valores de $x$ para los cuales el determinante de la matriz dada es nulo:

*a)* $A=\begin{pmatrix} 3-x & -1 & 0 \\ 6 & -3-x & 2 \\ 8 & -6 & 5-x \end{pmatrix}$ *b)* $B=\begin{pmatrix} 4-x & -5 & 7 \\ 1 & -4-x & 9 \\ -4 & 0 & 5-x \end{pmatrix}$

*c)* $C=\begin{pmatrix} 3-x & -4 & 0 & 2 \\ 4 & -5-x & -2 & 4 \\ 0 & 0 & 3-x & -2 \\ 0 & 0 & 2 & -1-x \end{pmatrix}$

**23.** *a)* Encontrar la matriz del cambio de base de $B=\{1, x, x^2, x^3, x^4\}$ a $B'=\{1, x-1, (x-1)^2, (x-1)^3, (x-1)^4\}$ en $P_{\mathbb{R}}^{(4)}[x]$.

*b)* Encontrar las coordenadas del polinomio

$$f(x)=5+4x+3x^2+2x^3+x^4$$

con respecto a la base $B'$.

**24.** Encontrar la dimensión y una base del subespacio vectorial generado por los vectores

$$\vec{u}_1=(1, 0, 0, -1), \quad \vec{u}_2=(2, 1, 1, 0), \quad \vec{u}_3=(1, 1, 1, 1)$$
$$\vec{u}_4=(1, 2, 3, 4), \quad \vec{u}_5=(0, 1, 2, 3)$$

Encontrar la ecuaciones cartesianas de este subespacio.

**25.** Dados los vectores $\vec{u}_1=(2, 1, -1)$, $\vec{u}_2=(3, 0, 1)$ y $\vec{u}_3=(1, 2, -3)$, determinar el subespacio que generan y hallar la intersección de dicho subespacio con el generado por $\vec{u}_1=(5, 1, 0)$, $\vec{v}_2=(1, -1, 1)$.



**26.** *a)* Encontrar las ecuaciones de la simetría $S$ en $\mathbb{R}^3$ con respecto a la recta $(x, y, z)=t(-1, 1, 2)$.

*b)* Hallar las ecuaciones de la proyección $P$ en $\mathbb{R}^3$ sobre el plano $x+y-z=3$.

**27.** Si $T, S: V\mapsto W$ son aplicaciones lineales, demostrar que $\ker(T)\cap\ker(S)\subset\ker(T+S)$.

**28.** Sean $V \overset{A}{\mapsto} W \overset{B}{\mapsto} X$ aplicaciones lineales. Si $B\circ A$ es suprayectiva y $\dim(W)=\dim(X)$, probar que $A$ y $B$ son suprayectivas.

**29.** *a)* Sea $A\in L(\mathbb{R}^{2n+1})$; demostrar que $\ker(A)\neq \mathrm{img}(A)$.

*b)* Sea $A\in L(\mathbb{R}^{2n})$; demostrar que $\ker(A) = \mathrm{img}(A)$ si y sólo si $A^2 = A\circ A = 0$ y $\dim(\mathrm{img}(A))=n$.

**30.** *a)* Hallar las ecuaciones paramétricas de $\ker(T)$ y de $\mathrm{img}(T)$, si $T:\mathbb{R}^3\mapsto\mathbb{R}^3$ está dada por

$$T(x_1, x_2, x_3)=(x_1+2x_2+x_3, \ -x_1+2x_2, \ 2x_2+x_3)$$

*b)* Encontrar la intersección de $\ker(T)$ e $\mathrm{img}(T)$.

**31.** Encontrar la forma de Jordan de las matrices

$$A=\begin{pmatrix} -1 & 2 & -3 \\ 0 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix} \quad \text{y} \quad B=\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 1 & -1 & 2 \end{pmatrix}$$

indicando el cambio de base.

**32.** Encontrar la forma de Jordan de $A=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ y calcular $A^3$.

**33.** Decir si la matriz $A=\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 2 \end{pmatrix}$ es diagonalizable, justificando la respuesta.

**34.** Encontrar la forma de Jordan de las matrices

$$A=\begin{pmatrix} 1 & -1 & & 0 \\ & 1 & -1 & \\ & & \ddots & \ddots & -1 \\ 0 & & & 1 \end{pmatrix}, \quad B=\begin{pmatrix} 1 & 2 & 3 & \cdots & n \\ & 1 & 2 & \cdots & n-1 \\ & & 1 & \cdots & n-2 \\ 0 & & & \ddots & \vdots \\ & & & & 1 \end{pmatrix}$$

ambas de orden $n$.

**35.** Dada la matriz $A=\begin{pmatrix} 2 & 0 & 0 \\ -1 & 1 & 15 \\ 3 & 0 & 1 \end{pmatrix}$, calcular:

*a)* El polinomio característico y la forma de Jordan de $A$.

*b)* $-A^3+4A^2-5A+2$.

*c)* $P(A)$, donde $P(x)=-x^6+4x^5-5x^4+2x^3-3I$.

**36.** Sea $A_n=\begin{pmatrix} n & 1 & 1 \\ 1 & n & 1 \\ 1 & 1 & n \end{pmatrix}$ con $n\in\mathbb{R}$.

*a)* Demostrar que $A_n$ es diagonalizable para todo $n$.

*b)* Calcular $A_n^{-1}$ cuando exista.

*c)* Si $A_n^{-1}$ no existe, encontrar $\ker(A_n)$ e $\operatorname{img}(A_n)$.

**37.** Encontrar (razonadamente) los autovalores de $A^{10}$ si $A$ es la matriz

$$A=\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

* * *

**38.** Discutir, según los valores de $a$ y $b$, el sistema

$$\left.\begin{aligned} ax-y+2z&=1 \\ x+ay-4z&=0 \\ -x+y-2z&=b \end{aligned}\right\}$$

y resolverlo cuando sea compatible indeterminado.

**39.** Se consideran los puntos $P=(1, 2, -1)$, $Q=(0, 1, 2)$ y los planos $\pi$ y $\pi'$ descritos por las condiciones siguientes: $\pi$ pasa por $P$ y es perpendicular a $(1, 1, 1)$ y $\pi'$ pasa por $Q$ y es perpendicular a $\overrightarrow{PQ}$. Se designa por $r$ la recta intersección de $\pi$ y $\pi'$. Determinar las ecuaciones cartesianas de la recta que pasa por $P$ y es perpendicular a $r$.

**40.** Considerar los puntos $A=(1, 1, 0)$ y $B=(0, 1, 1)$.

*a)* Determinar un punto $C$ sobre la recta $r$: $(0, 1, 1)+s(1, 0, 1)$, cuya distancia a la recta determinada por $A$ y $B$ sea $2\sqrt{2}$.

*b)* Determinar un punto $D$ sobre la recta $r'$: $(1, 1, 1)+t(0, 1, 0)$ de tal manera que el volumen del paralelepípedo determinado por los puntos *A*, *B*, *C* y *D* sea 4, donde $C$ es la solución del apartado *a)* que tiene todas sus coordenadas positivas.



**41.** Encontrar bases del núcleo y de la imagen de la aplicación lineal $L$: $P_{\mathbb{C}}^{(3)}[x]\mapsto P_{\mathbb{C}}^{(2)}[x]$ dada por

$$L(ax^3+bx^2+cx+d)=(a-c+2d)x^2+(-2a+b)x+(b-2c+4d)$$

**42.** *a)* Estudiar si la aplicación $T$: $\mathbb{R}^3\mapsto\mathbb{R}^3$ dada por

$$T(x_1, x_2, x_3)=\frac{1}{6}(-4x_1+4x_2,\ 5x_1+4x_2,\ -x_1-2x_2-6x_3)$$

es diagonalizable.

*b)* Encontrar los subespacios invariantes de dimensión 1 de la aplicación $T$.

**43.** La suma de tres números positivos no nulos es $M$; el primero, más el doble del segundo, menos el triple del tercero es 5 y el primero, más el triple del segundo, menos 7 veces el tercero es 0. Encontrar estos tres números.

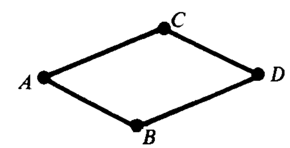
**44.** Dados los puntos $P=(1, 2, -1)$ y $Q=(1, 3, 2)$ y la recta $r$ de ecuación

$$r\colon \begin{cases} 2x+y-z=6 \\ x-y+z=1 \end{cases}$$

Hallar: *a)* la proyección ortogonal de $P$ sobre $r$; *b)* un punto situado en el plano perpendicular a $r$ por $P$ y que esté a la menor distancia posible de $Q$.

**45.** *a)* Sean $A=(1, 0, 1)$, $B=(2, 1, 1)$ y $C=(-1,0, 1)$ tres puntos en el espacio; encontrar el cuarto vértice $D$ del paralelogramo de la figura adjunta.

*b)* Encontrar los vértices $E$ de las pirámides rectas que tienen como base $ABCD$ y cuyo volumen es 2 (hay 2 vértices).




**46.** Dados los conjuntos de vectores $T_1=\{(1, 1, 0), (1, 0, 1)\}$ y $T_2=\{(1, 1, 1), (0, 1, 1)\}$ en $\mathbb{R}^3$, determinar:

*a)* La relación de inclusión entre $L(T_1)$ y $L(T_2)$.

*b)* $L(T_1)\cap L(T_2)$.

*c)* $L(T_1)+L(T_2)$.

**47.** Encontrar la matriz, en la base canónica, de la transformación lineal $A$ de $\mathbb{R}^3$ en $\mathbb{R}^3$ que lleva $\vec{u}_j$ en $\vec{v}_j$, donde

$$\vec{u}_1=(2, 3, 5),\quad \vec{u}_2=(0, 1, 2),\quad \vec{u}_3=(1, 0, 0)$$
$$\vec{v}_1=(1, 1, 1),\quad \vec{v}_2=(1, 1, -1),\quad \vec{v}_3=(2, 1, 2)$$

**48.** Estudiar para qué valores del parámetro $a \in \mathbb{R}$, los vectores $\vec{u}_1 = (1, -1, 1, -1)$, $\vec{u}_2 = (a, 1, 1, a)$, $\vec{u}_3 = (2, 0, 2, 0)$ y $\vec{u}_4 = (-1, a, 1, a)$ son linealmente dependientes y encontrar la relación que existe entre estos vectores para los valores de $a$ que los hagan linealmente dependientes.

**49.** Encontrar las ecuaciones paramétricas y cartesianas de la proyección de la recta $(2, 1, 1) + t(-1, 0, 2)$ sobre el plano $2x + y - z = 0$.

**50.** Los puntos $A = (0, 0, 0)$, $B = (4, 3, 0)$ y $C = (0, 0, 2)$ son vértices de una cara de un paralelogramo de volumen 20. Hallar la ecuación del plano que contiene a la cara opuesta a la dada.

**51.** Se considera la base de $\mathbb{C}^3$ siguiente:

$$B = \{(1, 1, 1), (1, 1, 0), (0, -1, 2)\}$$

y la aplicación lineal $A$, de $\mathbb{C}^3$ en $\mathbb{C}^2$, que cumple:

$$A(1, 1, 1) = (i, 1), \quad A(1, 1, 0) = (1, -i), \quad A(0, -1, 2) = (2 + i, 1 - 2i)$$

*a*) Determinar la matriz de $A$ en las bases canónicas de $\mathbb{C}^3$ y $\mathbb{C}^2$.

*b*) Determinar bases del núcleo y de la imagen de la aplicación $A$ del problema anterior.

**52.** Estudiar la posibilidad de diagonalizar la matriz $\begin{pmatrix} a & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 2 & 2 \end{pmatrix}$ según los distintos valores de $a$.

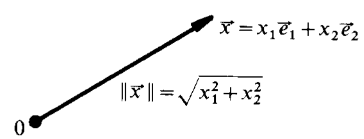
# CAPITULO 8

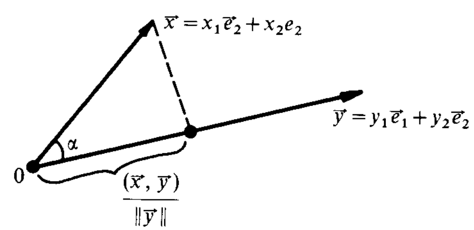
# ESPACIOS EUCLIDEOS



## 8.1. DEFINICION DE ESPACIO EUCLIDEO. EJEMPLOS

Una gran variedad de hechos geométricos se basan principalmente en la posibilidad de medir las longitudes de segmentos y los ángulos entre ellos. Obsérvese que esto no es posible en un espacio vectorial. Para introducir el concepto de longitud de un vector y de ángulo entre dos vectores fijémonos en el espacio vectorial $V_2$ de los vectores en el plano.



La idea intuitiva que poseemos de longitud de un vector, que designamos por $\|\vec{x}\|$, es $\|\vec{x}\| = \sqrt{x_1^2+x_2^2}$, mientras que para el coseno del ángulo $\alpha$, que forman $\vec{x}$ e $\vec{y}$, se tiene

$$\cos\alpha = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\|\|\vec{y}\|} \tag{1}$$

en donde $(\vec{x}, \vec{y})$ denota el *producto escalar* de dos vectores $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2$ e $\vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2$. Recordar que el producto escalar de dos vectores en el plano ha sido introducido en la sección 3.3 (capítulo 3) y que la fórmula (1) ha sido demostrada en la misma sección.

La fórmula (1) relaciona ángulos con medidas y con productos escalares y es la base para introducir axiomas en un espacio vectorial, de manera que en los nuevos espacios se puedan estudiar propiedades geométricas.

Nuestra táctica es definir estos espacios, que serán los *espacios euclídeos*, a partir de una definición axiomática de «producto escalar».

Daremos en esta sección la definición de estos nuevos espacios, así como algunos ejemplos. En secciones posteriores estudiaremos los conceptos de medida, ángulo y ortogonalidad en estos espacios; finalmente estudiaremos diversos tipos de aplicaciones lineales entre ellos.

DEFINICIÓN 1 *(Espacio euclídeo)*

Un espacio vectorial *real* $E$ se dice *euclídeo* si hay una regla que asigne a cada par de vectores $\vec{x}, \vec{y} \in E$ un número real llamado *producto escalar* de los vectores $\vec{x}$ e $\vec{y}$, designado por $(\vec{x}, \vec{y})$, de manera que se cumplen las siguientes propiedades:

*a)* Simétrica: $(\vec{x}, \vec{y}) = (\vec{y}, \vec{x})$ para todo $\vec{x}, \vec{y} \in E$.
*b)* Distributiva: $(\vec{x}, \vec{y}+\vec{z}) = (\vec{x}, \vec{y}) + (\vec{x}, \vec{z})$ para todo $\vec{x}, \vec{y}, \vec{z} \in E$.
*c)* $(\lambda\vec{x}, \vec{y}) = \lambda(\vec{x}, \vec{y})$ para todo $\vec{x}, \vec{y} \in E$ y todo $\lambda \in \mathbb{R}$.
*d)* $(\vec{x}, \vec{x}) > 0$ para todo $\vec{x} \neq \vec{0}$.

De *b)* y *a)* se deduce:

*b')* $(\vec{x}+\vec{z}, \vec{y}) = (\vec{x}, \vec{y}) + (\vec{z}, \vec{y})$ para todo $\vec{x}, \vec{y}, \vec{z} \in E$.

De *c)* y *a)* se deduce:

*c')* $(\vec{x}, \lambda\vec{y}) = \lambda(\vec{x}, \vec{y})$ para todo $\vec{x}, \vec{y} \in E$ y todo $\lambda \in \mathbb{R}$.

De *b)* se deduce que $(\vec{0}+\vec{y}, \vec{x}) = (\vec{y}, \vec{x}) = (\vec{0}, \vec{x}) + (\vec{y}, \vec{x}) \Leftrightarrow (\vec{0}, \vec{x}) = 0$.
Por la propiedad simétrica $(\vec{x}, \vec{0}) = 0$. En particular, $(\vec{x}, \vec{x}) = 0$ si $\vec{x} = \vec{0}$.

EJEMPLO A. El espaco vectorial $\mathbb{R}^2$ con

$$(\vec{x}, \vec{y}) = x_1y_1 + x_2y_2$$



donde $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2)$ y en general el espacio vectorial $\mathbb{R}^n$ con

$$(\vec{x}, \vec{y}) = x_1y_1 + \cdots + x_ny_n$$

donde $\vec{x} = (x_1, ..., x_n)$ e $\vec{y} = (y_1, ..., y_n)$. Las propiedades *a)* a *d)* son fáciles de verificar y se dejan como ejercicio para el lector.

EJEMPLO B. Dados $\vec{x} = (x_1, x_2)$ e $\vec{y} = (y_1, y_2)$ vectores en $\mathbb{R}^2$, definimos

$$(\vec{x}, \vec{y}) = (x_1, x_2)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

La aplicación que acabamos de definir de $\mathbb{R}^2 \times \mathbb{R}^2$ en $\mathbb{R}$ es un producto escalar: las propiedades *b)* y *c)* se deducen de las propiedades de la multiplicación de matrices; para probar *a)* y *d)* observamos que

$$(\vec{x}, \vec{y}) = x_1y_1 + x_2y_1 + y_2x_1 + 2y_2x_2;$$

de aquí se deduce la propiedad conmutativa; para probar la propiedad *d)* observar que

$$(\vec{x}, \vec{x}) = x_1^2 + 2x_1x_2 + 2x_2^2 = (x_1+x_2)^2 + x_2^2.$$

Por tanto, tenemos un nuevo espacio euclídeo en el espacio vectorial $\mathbb{R}^2$.

EJEMPLO C. Si $\vec{x} = (x_1, x_2)$, $\vec{y} = (y_1, y_2) \in \mathbb{R}^2$ y definimos $(\vec{x}, \vec{y}) = x_1y_1$, esto no es un producto escalar en $\mathbb{R}^2$ puesto que no se cumple la propiedd *d)*; en efecto, si $\vec{x} = (0, x_2) \neq \vec{0}$ se tiene $(\vec{x}, \vec{x}) = 0$.

EJEMPLO D. En el espacio $\mathscr{C}([a, b])$ de las funciones continuas reales definidas en el intervalo $[a, b]$, el producto escalar de las funciones $f(t)$ y $g(t)$ se define como

$$(f, g) = \int_a^b f(t)g(t)\,dt,$$

obteniéndose un espacio euclídeo. El mismo producto escalar sirve para convertir el espacio vectorial $\mathscr{P}([a, b])$, de los polinomios en el intervalo $[a, b]$, en un espacio euclídeo.

* * *

Sea $E$ un espacio euclídeo de dimensión finita $n$ (es decir, el espacio vectorial asociado a $E$ es de dimensión finita $n$) y sea $B = \{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ una base de $E$. En esta base, si $\vec{x} \in E$ e $\vec{y} \in E$, podemos escribir:

$$\vec{x} = \sum_{i=1}^n x_i\vec{u}_i, \quad \vec{y} = \sum_{j=1}^n y_j\vec{u}_j.$$

Utilizando las propiedades *b*) y *c*) del producto escalar se tiene que

$$(\vec{x}, \vec{y}) = \left( \sum_{i=1}^{n} x_i \vec{u}_i, \sum_{j=1}^{n} y_j \vec{u}_j \right) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\vec{u}_i, \vec{u}_j) \tag{1}$$

La matriz

$$P \equiv P_B = \begin{pmatrix} (\vec{u}_1, \vec{u}_1) & (\vec{u}_2, \vec{u}_1) & \cdots & (\vec{u}_n, \vec{u}_1) \\ (\vec{u}_1, \vec{u}_2) & (\vec{u}_2, \vec{u}_2) & \cdots & (\vec{u}_n, \vec{u}_2) \\ \vdots & \vdots & & \vdots \\ (\vec{u}_1, \vec{u}_n) & (\vec{u}_2, \vec{u}_n) & \cdots & (\vec{u}_n, \vec{u}_n) \end{pmatrix}$$

recibe el nombre de *matriz del producto escalar con respecto a la base B*. Con esta notación la expresión (1) puede escribirse de la forma

$$(\vec{x}, \vec{y}) = (x_1, \ldots, x_n) P_B \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$$

La matriz $P_B$ de un producto escalar es siempre simétrica debido a la propiedad *a*) y los elementos de su diagonal principal son todos positivos debido a la propiedad *d*).

Podemos prguntarnos si estas dos condiciones son suficientes para que una matriz $P$ defina un producto escalar; la respuesta es negativa, pues el lector puede encontrar un contraejemplo en el ejercicio 1 propuesto al final de la siguiente sección. En el capítulo 12, dedicado a las formas bilineales y cuadráticas, se dará una condición necesaria y suficiente para que una matriz defina una producto escalar en un espacio vectorial.

## 8.2. LONGITUDES, ANGULOS Y ORTOGONALIDAD

La *longitud* (*o norma*) de un vector $\vec{x}$ en un espacio euclídeo se define como

$$\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})}, \quad \vec{x} \in E.$$

Obsérvese que $\sqrt{(\vec{x}, \vec{x})}$ tiene sentido debido a la propiedad *d*) del producto escalar. En el ejemplo A de la sección anterior se tiene:

$$\|\vec{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}, \quad \vec{x} = (x_1, \ldots, x_n).$$

En el ejemplo B de la misma sección se tiene que

$$\|\vec{x}\| = \sqrt{(\vec{x}, \vec{x})} = \sqrt{(x_1 + x_2)^2 + x_2^2}, \quad \vec{x} = (x_1, x_2).$$



Para el ejemplo D de las funciones continuas reales definidas en el intervalo $[a, b]$ se tiene

$$\|f\| = \sqrt{\int_a^b |f(t)|^2 \, dt}, \quad f \in \mathscr{C}([a, b]).$$

La longitud de un vector $\vec{x} \in E$ multiplicado por un número real $\lambda$ coincide con $|\lambda|$ veces la longitud de $\vec{x}$, ya que

$$|\lambda \vec{x}| = \sqrt{(\lambda \vec{x}, \lambda \vec{x})} = \sqrt{\lambda^2 (\vec{x}, \vec{x})} = |\lambda| \sqrt{(\vec{x}, \vec{x})} = |\lambda| \|\vec{x}\|.$$

Todo vector de longitud 1 se dice *unitario;* todo vector $\vec{x}$ no nulo de un espacio euclídeo puede *normalizarse*, es decir, hacerle unitario, multiplicándole por $1/\|\vec{x}\|$. La *bola unidad* es el conjunto de todos los $\vec{x} \in E$ tal que $\|\vec{x}\| \leq 1$, mientras que *la esfera unidad* es el conjunto de todos los $\vec{x} \in E$ tal que $\|\vec{x}\| = 1$.

* * *

Dados dos vectores $\vec{x}, \vec{y}$ de un espacio euclídeo definimos el *coseno del ángulo* $\alpha$ que forman entre ellos como

$$\cos \alpha = \frac{(\vec{x}, \vec{y})}{\|\vec{x}\| \|\vec{y}\|} \tag{2}$$

en analogía con la fórmula (1) de la sección anterior. Para que (2) tenga sentido es necesario demostrar que el valor absoluto del cociente $(\vec{x}, \vec{y})/\|\vec{x}\| \|\vec{y}\|$ es menor o igual que 1.

PROPOSICIÓN 1 *(Desigualdad de Schwarz)*

En todo espacio euclídeo $E$,

$$|(\vec{x}, \vec{y})| \leq \|\vec{x}\| \|\vec{y}\|$$

para todo $\vec{x}, \vec{y} \in E$.

*Demostración.* Por la propiedad *d*) del producto escalar se tiene

$$(\lambda \vec{x} - \vec{y}, \lambda \vec{x} - \vec{y}) \geq 0$$

para todo número real $\lambda$; de las propiedades *a*), *b*) y *c*) del producto escalar se deduce que

$$\lambda^2 \|\vec{x}\|^2 - 2\lambda (\vec{x}, \vec{y}) + \|\vec{y}\|^2 \geq 0.$$

La parte izquierda de esta ecuación es una expresión cuadrática en $\lambda$ que no puede tener raíces reales distintas, y, por tanto, su discriminante ha de ser $\leq 0$; por tanto:

$$(\vec{x}, \vec{y})^2 - \|\vec{x}\|^2\|\vec{y}\|^2 \leq 0.$$

Tomando raíces cuadradas se deduce $|(\vec{x}, \vec{y})| \leq \|\vec{x}\|\|\vec{y}\|$, que es el resultado que deseábamos probar. ■

En $\mathbb{R}^n$ con el producto escalar «usual» dado en el ejemplo A la desigualdad de Schwarz se escribe de la forma

$$\left|\sum_{j=1}^{n} x_j y_j\right| \leq \sqrt{\sum_{j=1}^{n} x_j^2}\ \sqrt{\sum_{j=1}^{n} y_j^2}, \tag{3}$$

y en el espacio $\mathscr{C}([a, b])$ es

$$\left|\int_a^b f(t)\,g(t)\,dt\right| \leq \sqrt{\int_a^b [f(t)]^2 dt}\ \sqrt{\int_a^b [g(t)]^2 dt} \tag{4}$$

La desigualdad (3) se atribuye al matemático francés Cauchy, mientras que (4) es atribuida al matemático ruso Buniakovski.

* * *

Finalmente introducimos el concepto de *ortogonalidad* o *perpendicularidad* entre dos vectores de un espacio euclídeo: dos vectores $\vec{x}, \vec{y} \in E$ se dicen *ortogonales* si

$$(\vec{x}, \vec{y}) = 0$$

En el espacio vectorial $V_2$ de los vectores en el plano esta definición coincide con la idea intuitiva de perpendicularidad: es decir, los vectores forman un ángulo de 90°.

EJEMPLO A. Los vectores

$$\vec{e}_1 = (1, 0, ..., 0),\ \vec{e}_2 = (0, 1, 0, ..., 0),\ ...,\ \vec{e}_n = (0, 0, ..., 1)$$

son ortogonales dos a dos en el espacio euclídeo $\mathbb{R}^n$ con el producto escalar usual

EJEMPLO B. En $\mathscr{C}([-\pi, \pi])$ los vectores del sistema trigonométrico

$$1, \cos t, \operatorname{sen} t, \cos 2t, \operatorname{sen} 2t, ..., \cos nt, \operatorname{sen} nt, ...$$

son ortogonales dos a dos; por ejemplo, si $n \neq m$:

$$\int_{-\pi}^{\pi} (\cos nt \operatorname{sen} mt)\,dt = \frac{1}{2}\int_{-\pi}^{\pi} \operatorname{sen}(n+m)t\,dt - \frac{1}{2}\int_{-\pi}^{\pi} \operatorname{sen}(n-m)t\,dt = 0$$

puesto que $n \neq m$. El resto se comprueba de manera similar.



Una proposición simple, asociada con el concepto de ortogonalidad, es la siguiente:

* * *

LEMA 2

Si los vectores no nulos $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n$ son ortogonales entre sí en un espacio euclídeo, también son linealmente independientes.

*Demostración.* Supongamos que $\alpha_1\vec{x}_1 + \alpha_2\vec{x}_2 + \cdots + \alpha_n\vec{x}_n = \vec{0}$; hallando el producto escalar con $\vec{x}_1$ se tiene

$$\alpha_1(\vec{x}_1, \vec{x}_1) + \alpha_2(\vec{x}_1, \vec{x}_2) + \cdots + \alpha_n(\vec{x}_1, \vec{x}_n) = (\vec{x}_1, \vec{0}) = \vec{0}$$

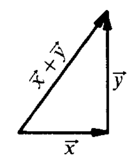
o equivalentemente $\alpha_1\|\vec{x}_1\| = \vec{0}$, ya que los vectores son ortogonales entre sí; como $\|\vec{x}_1\| \neq 0$, se tiene $\alpha_1 = 0$. De forma similar se deduce $\alpha_2 = \cdots = \alpha_n = 0$. ■

* * *

Podemos ahora trasladar teoremas de la geometría elemental a este nuevo marco de la «geometría» en los espacios euclídeos. Comencemos con el *teorema de Pitágoras*. Por analogía con los vectores en el plano podemos considerar $\vec{x} + \vec{y}$ como la hipotenusa de un triángulo rectángulo determinado por los vectores ortogonales $\vec{x}$ e $\vec{y}$. Por la definición de producto escalar y la ortogonalidad de $\vec{x}$ e $\vec{y}$ se tiene:

$$\|\vec{x} + \vec{y}\|^2 = (\vec{x} + \vec{y}, \vec{x} + \vec{y}) = \|\vec{x}\|^2 + 2(\vec{x}, \vec{y}) + \|\vec{y}\|^2 = \|\vec{x}\|^2 + \|\vec{y}\|^2$$

Esta igualdad representa el teorema de Pitágoras en los espacios euclídeos.



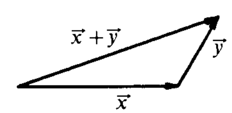

Otra propiedad geométrica que puede demostrarse en los espacios euclídeos es la *desigualdad triangular*, es decir, en todo triángulo la longitud $\|\vec{x} + \vec{y}\|$ de uno de sus lados es menor o igual que la suma de las longitudes de los otros dos. Para demostrar esto, se utiliza la desigualdad de Schwarz para obtener

$$\|\vec{x} + \vec{y}\|^2 = (\vec{x} + \vec{y}, \vec{x} + \vec{y}) = \|\vec{x}\|^2 + 2(\vec{x}, \vec{y}) + \|\vec{y}\|^2 \leq$$
$$\leq \|\vec{x}\|^2 + 2\|\vec{x}\|\|\vec{y}\| + \|\vec{y}\|^2 = (\|\vec{x}\| + \|\vec{y}\|)^2$$

por tanto:

$$\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$$

que es el resultado deseado.



## EJERCICIOS (SECCIONES 8.1 Y 8.2)

**1.** ¿Cuáles de las siguientes funciones ( , ): $\mathbb{R}^2 \times \mathbb{R}^2 \mapsto \mathbb{R}$ definen un producto escalar?:

*a)* $(\vec{x}, \vec{y}) = x_1y_1 + 2x_2y_2 + 2y_1x_2 - 3x_1y_2$
*b)* $(\vec{x}, \vec{y}) = x_1y_1 - x_2y_2$
*c)* $(\vec{x}, \vec{y}) = x_1y_1 + 2x_1y_2 + 3y_1x_2 + 7x_2y_2$
*d)* $(\vec{x}, \vec{y}) = x_1y_1 + 2x_2y_2 + 3x_1y_2 + 3x_2y_1$
*e)* $(\vec{x}, \vec{y}) = 2x_1y_1 + 3x_2y_2 + 2x_1y_2 + 2x_2y_1$

Justificar la respuesta. Encontrar la matriz, en la base canónica, de aquellas funciones que sean producto escalar.

**2.** Encontrar la expresión analítica del módulo de un vector y del coseno del ángulo de dos vectores en $\mathbb{R}^2$ para aquellas de las funciones en 1, que sean producto escalar.

**3.** Encontrar el ángulo entre las aristas opuestas de un tetraedro regular.

**4.** Encontrar los cosenos de los ángulos entre la recta $x_1 = x_2 = \cdots = x_n$ y los ejes coordenados en $\mathbb{R}^n$, con el producto escalar usual.

**5.** En el espacio vectorial $\mathcal{M}_{3\times 3}(\mathbb{R})$ de las matrices reales de orden 3, demostrar que $\langle A, B\rangle = \text{traza}(AB^t)$ es un producto escalar.

**6.** Sea $(V, ( \, , \, ))$ un espacio euclídeo. Demostrar:

*a)* $2(\|\vec{u}\|^2 + \|\vec{v}\|^2) = \|\vec{u}+\vec{v}\|^2 + \|\vec{u}-\vec{v}\|^2$, $\forall \vec{u}, \vec{v} \in V$ (ley del paralelogramo).
*b)* $4(\vec{u}, \vec{v}) = \|\vec{u}+\vec{v}\|^2 - \|\vec{u}-\vec{v}\|^2$, $\forall \vec{u}, \vec{v} \in V$ (polarización)
*c)* $2(\vec{u}, \vec{v}) = \|\vec{u}+\vec{v}\|^2 - \|\vec{u}\|^2 - \|\vec{v}\|^2$, $\forall \vec{u}, \vec{v} \in V$.

**7.** Sea $V$ un espacio vectorial. Una norma en $V$ es una función $\| \; \|: V \mapsto \mathbb{R}$ que satisface:

i) $\|\vec{u}\| \geqq 0$, $\forall \vec{u} \in V$ y $\|\vec{u}\| = 0 \Leftrightarrow \vec{u} = 0$
ii) $\|\lambda\vec{u}\| = |\lambda|\|\vec{u}\|$, $\forall \vec{u} \in V$, $\lambda \in \mathbb{K}$ $(=\mathbb{R}$ o $\mathbb{C})$
iii) $\|\vec{u}+\vec{v}\| \leqq \|\vec{u}\| + \|\vec{v}\|$ (desigualdad triangular).

*a)* Si $V$ es un espacio vectorial euclídeo con producto escalar ( , ) el módulo de un vector es una norma que satisface la ley del paralelogramo. Probar que si $V$ es un espacio vectorial con norma $\| \; \|$ que satisface la ley del paralelogramo, entonces existe un producto escalar ( , ) en $V$ tal que $\|\vec{u}\| = (\vec{u}, \vec{u})^{1/2}$. En consecuencia, la condición necesaria y suficiente para que una norma sea el módulo de un producto escalar es que satisfaga la ley del paralelogramo. [*Sugerencia*: comenzar demostrando que $(\vec{u}_1 + \vec{u}_2, \vec{v}) = (\vec{u}_1, \vec{v}) + (\vec{u}_2, \vec{v})$.] (ver problema 6).

*b)* Sea $\|\vec{x}\| = |x_1| + |x_2|$ definida en $\mathbb{R}^2$. Demostrar que $\| \; \|$ es una norma en $\mathbb{R}^2$, pero no es el módulo de ningún producto escalar en $\mathbb{R}^2$.



**8.** Sea $(E, ( \, , \, ))$ un espacio euclídeo y $\{\vec{u}_1, ..., \vec{u}_n\}$ un conjunto de vectores ortogonales de $E$. Demostrar:

*a)* $\displaystyle\sum_{i=1}^{n} \frac{|(\vec{v}, \vec{u}_i)|^2}{\|\vec{u}_i\|^2} \leqq \|\vec{v}\|^2$, $\forall \vec{v} \in E$ (desigualdad de Bessel).

*b)* Si los $\{\vec{u}_i\}$ son una base ortogonal de $E$

$$(\vec{v}, \vec{w}) = \sum_{i=1}^{n} \frac{(\vec{v}, \vec{u}_i)(\vec{u}_i, \vec{w})}{\|u_i\|^2}, \quad \forall \vec{v}, \vec{w} \in E \text{ (identidad de Parseval)}$$

**9.** Demostrar que la desigualdad de Schwarz es una igualdad si y sólo si los vectores son proporcionales.

**10.** Demostrar la siguiente generalización del teorema de Pitágoras:

$$\|\vec{x}_1 + \vec{x}_2 + \cdots + \vec{x}_n\|^2 = \|\vec{x}_1\|^2 + \|\vec{x}_2\|^2 + \cdots + \|\vec{x}_n\|^2$$

si los vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_n \in E$ son ortogonales dos a dos.

**11.** Si $\vec{x}$ e $\vec{y}$ son dos vectores cualesquiera en un espacio euclídeo $E$, demostrar que $\|\vec{x}-\vec{y}\| \geqq |\|\vec{x}\| - \|\vec{y}\||$.

## 8.3. BASES ORTONORMALES EN UN ESPACIO EUCLIDEO

En un espacio euclídeo $E$ una base se dice *ortogonal* si sus elementos son ortogonales dos a dos; si, además, los elementos de una base ortogonal son de longitud 1, la base se llama *ortonormal.*

Ya hemos demostrado en la sección 8.2, lema 2, que todo conjunto de vectores, en un espacio euclídeo, ortogonales dos a dos, son linealmente independientes. Demostraremos a continuación que en todo espacio euclídeo existen bases ortogonales; en el proceso de la demostración daremos un procedimiento, denominado *método de ortogonalización,* que permite obtener una base ortogonal a partir de cualquier base del espacio vectorial. En la literatura matemática también se le llama *método de Gram-Schmidt.*

TEOREMA 1 (*Método de ortogonalización*)

Sea $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k, ...$ una sucesión finita o infinita de vectores linealmente independientes en un espacio euclídeo $E$ y sea $L_k = L(\vec{x}_1, ..., \vec{x}_k)$, el espacio vectorial generado por los $k$ primeros vectores. Entonces, existe un conjunto de vectores $\vec{y}_1, \vec{y}_2, ..., \vec{y}_k, ...$ tal que:

*a)* El espacio vectorial $L_k = L(\vec{y}_1, ..., \vec{y}_k)$ coincide con $L_k$ para todo entero positivo $k$.
*b)* El vector $\vec{y}_{k+1}$ es ortogonal a $L_k$ para todo entero positivo $k$.

*Nota.* Un vector $\vec{y}$ en un espacio euclídeo se dice *ortogonal a un subespacio vectorial* $L$ si es ortogonal a todos los vectores de $L$; si $L$ está generado por los vectores $\vec{x}_1, \vec{x}_2, ..., \vec{x}_k$, basta probar que $\vec{y}$ es ortogonal a cada uno de los vectores $\vec{x}_j$, $j=1, 2, ..., k$; en efecto, si $\vec{x} = \sum_{j=1}^{k} \alpha_j \vec{x}_j$ se tiene que

$$(\vec{x}, \vec{y}) = \left(\sum_{j=1}^{k} \alpha_j \vec{x}_j, \vec{y}\right) = \sum_{j=1}^{k} \alpha_j (\vec{x}_j, \vec{y}) = 0$$

TEOREMA 2

En todo espacio euclídeo $E$ de dimensión finita existen bases ortonormales.

*Demostración.* Sea $\{\vec{x}_1, \vec{x}_2, ..., \vec{x}_n\}$ una base cualquiera de $E$; tomamos $\{\vec{y}_1, \vec{y}_2, ..., \vec{y}_n\}$ como en el método de ortogonalización del teorema 1; esta nueva colección genera todo el espacio vectorial debido a *a*) y son ortogonales debido a *b*). Para encontrar una base ortonormal tomar $\vec{e}_j = \vec{y}_j / \|\vec{y}_j\|$, $j=1, ..., n$. ■

*Demostración del teorema* 1. Tomar $\vec{y}_1 = \vec{x}_1$ e $\vec{y}_2 = \vec{x}_2 + \alpha \vec{y}_1$; elegimos $\alpha$ de manera que $\vec{y}_1$ e $\vec{y}_2$ sean ortogonales:

$$0 = (\vec{y}_1, \vec{y}_2) = (\vec{y}_1, \vec{x}_2 + \alpha \vec{y}_1) = (\vec{y}_1, \vec{x}_2) + \alpha (\vec{y}_1, \vec{y}_1)$$

de donde se deduce que

$$\alpha = -\frac{(\vec{y}_1, \vec{x}_2)}{\|\vec{y}_1\|^2}$$

El denominador de esta fracción es no nulo, puesto que $\vec{y}_1 \neq 0$; de la construcción realizada se deducen inmediatamente las propiedades *a*) y *b*) para $k=2$.

Procedemos ahora por inducción; supongamos que hemos elegido $\vec{y}_1, \vec{y}_2, ..., \vec{y}_k$, de manera que $L_k = L'_k$ y son ortogonales dos a dos. Tomemos

$$\vec{y}_{k+1} = \vec{x}_{k+1} + \lambda_1 \vec{y}_1 + \lambda_2 \vec{y}_2 + \cdots + \lambda_k \vec{y}_k$$

y escojamos $\lambda_1, \lambda_2, ..., \lambda_k$ de modo que $\vec{y}_{k+1}$ sea ortogonal a $\vec{y}_1, \vec{y}_2, ..., \vec{y}_k$. Para ello tenemos

$$0 = (\vec{y}_{k+1}, \vec{y}_j) = (\vec{x}_{k+1}, \vec{y}_j) + \lambda_j (\vec{y}_j, \vec{y}_j), \quad j=1, 2, ..., k$$

de donde deducimos que

$$\lambda_j = -\frac{(\vec{x}_{k+1}, \vec{y}_j)}{\|\vec{y}_j\|^2}$$

Como antes, el denominador de esta fracción es no nulo y $L'_{k+1} = L_{k+1}$. ■



**EJEMPLO A.** **Encontremos una base ortogonal en el espacio $P_{\mathbb{R}}^{(3)}([-1, 1])$ de los polinomios de grado menor o igual que 3 en el intervalo $[-1, 1]$ con el producto escalar del ejemplo D (sección 8.1).**

**Tomemos como base inicial $x_1 = 1$, $x_2 = t$, $x_3 = t^2$, $x_4 = t^3$. Elegimos $y_1 = x_1 = 1$, $y_2 = t + \alpha \cdot 1$ y calculamos $\alpha$ para que $(y_1, y_2) = 0$:**

$$0 = (y_1, y_2) = (1, t+\alpha) = (1, t) + (1, \alpha) = \int_{-1}^{1} t\,dt + \alpha \int_{-1}^{1} dt = 0 + 2\alpha;$$

por tanto, $y_2 = t$. Tomamos $y_3 = t^2 + \lambda_1 \cdot 1 + \lambda_2 t$:

$$0 = (y_1, y_3) = \int_{-1}^{1} t^2\,dt + \lambda_1 \int_{-1}^{1} dt + \lambda_2 \int_{-1}^{1} t\,dt = \frac{2}{3} + 2\lambda_1 \Rightarrow$$

$$\lambda_1 = -\frac{1}{3};$$

$$0 = (y_2, y_3) = \int_{-1}^{1} t^3\,dt + \lambda_1 \int_{-1}^{1} t\,dt + \lambda_2 \int_{-1}^{1} t^2\,dt = \lambda_2 \frac{2}{3} \Rightarrow \lambda_2 = 0;$$

por tanto:

$$y_3 = t^2 - \frac{1}{3}$$

Tomamos $y_4 = t^3 + \lambda_1 \cdot 1 + \lambda_2 t + \lambda_3 \left(t^2 - \frac{1}{3}\right)$:

$$0 = (y_1, y_4) = \int_{-1}^{1} t^3\,dt + \lambda_1 \int_{-1}^{1} dt + \lambda_2 \int_{-1}^{1} t\,dt + \lambda_3 \int_{-1}^{1} \left(t^2 - \frac{1}{3}\right) dt = 2\lambda_1 \Rightarrow$$

$$\lambda_1 = 0;$$

$$0 = (y_2, y_4) = \int_{-1}^{1} t^4\,dt + \lambda_1 \int_{-1}^{1} t\,dt + \lambda_2 \int_{-1}^{1} t^2\,dt + \lambda_3 \int_{-1}^{1} t\left(t^2 - \frac{1}{3}\right) dt =$$

$$= \frac{2}{5} + \frac{2}{3}\lambda_2 \Rightarrow$$

$$\lambda_2 = -\frac{3}{5};$$

finalmente,

$$0=(y_3, y_4)=(y_3, t^3)+\lambda_1(y_3, 1)+\lambda_2(y_3, t)+\lambda_3(y_3, y_3)=$$
$$=\int_{-1}^{1}\left(t^2-\frac{1}{3}\right)t^3\,dt+0+0+\lambda_3\int_{-1}^{1}\left(t^2-\frac{1}{3}\right)^2 dt=0+\lambda_3\|y_3\|^2 \Rightarrow$$
$$\lambda_3=0;$$

por tanto,

$$y_4=t^3-\frac{3}{5}t;$$

Resumiendo,

$$\left\{y_1=1,\ y_2=t,\ y_3=t^2-\frac{1}{3},\ y_4=t^3-\frac{3}{5}t\right\}$$

es una base ortogonal de este espacio.

* * *

Sea $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ una base ortonormal en un espacio euclídeo $E$. Puesto que $(\vec{e}_i, \vec{e}_j)=\delta_{ij}$, $i, j=1, 2, ..., n$ (¿recuerda la $\delta$ de Krönecker introducida en el capítulo 6?), la matriz del producto escalar en esta base es

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}=I.$$

Por tanto, si $\vec{x}=\sum_{i=1}^{n} x_i\vec{e}_i\in E$ e $\vec{y}=\sum_{j=1}^{n} \vec{y}_j\vec{e}_j\in E$ se tiene que

$$(\vec{x}, \vec{y})=x_1y_1+x_2y_2+\cdots+x_ny_n \tag{3}$$

que nos da una forma sencilla de expresar en coordenadas el producto escalar de un espacio euclídeo, cuando en él se ha elegido una base ortonormal.

Recíprocamente, si un producto escalar tiene una expresión como en (3) en una cierta base, esta base es ortonormal; basta observar que de (3) se deduce que $(\vec{e}_i, \vec{e}_j)=0$ si $i\neq j$ y $(\vec{e}_i, \vec{e}_i)=1\cdot 1=1$, $i, j=1, 2, ..., n$.

*Ejercicio.* Los polinomios $y_1=1$, $y_2=t$, $y_3=t^2-\frac{1}{3}$, $y_4=t^3-\frac{3}{5}t$ encontrados en el ejemplo A reciben el nombre de polinomios de Legendre, porque fue este matemático francés quien los introdujo en 1785. Demostrar que los polinomios $y_n(t)$ coinciden, salvo constantes numéricas, con los polinomios dados por la fórmula



$$p_n(t)=\frac{d^n}{dt^n}[(t^2-1)^n], \quad n=0, 1, 2, 3$$

Este fórmula fue encontrada en 1814 por el matemático portugués Rodrigues.

*Ejercicio.* En el espacio vectorial $\mathscr{C}([-\infty, \infty])$ de las funciones continuas en la recta real, demostrar que $(f, g)=\int_{-\infty}^{\infty} f(x)g(x)e^{-x^2}$ es un producto escalar y encontrar los tres primeros términos de la ortogonalización de la familia de funciones $1, x, x^2, x^3$,

*Nota.* Es necesario saber que

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx=\sqrt{\pi}$$

## EJERCICIOS 8.3

**1.** Sean $\vec{u}_1=(-2, -1, 1)$, $\vec{u}_2=(0, -1, 0)$ y $\vec{u}_3=(1, -1, 0)$ tres vectores linealmente independientes de $\mathbb{R}^3$. Definimos un producto escalar en $\mathbb{R}^3$ afirmando que $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es una base ortonormal. Encontrar la expresión analítica de este producto escalar en la base canónica de $\mathbb{R}^3$.

**2.** Dados $\vec{x}_1=(1, 0, 0)$, $\vec{x}_2=(4, -2, 0)$ y $\vec{x}_3=(1, 1, 5)$ en $\mathbb{R}^3$, construir los vectores $\vec{y}_1$, $\vec{y}_2$ e $\vec{y}_3$ del teorema de ortogonalización. ($\mathbb{R}^3$ tiene el producto escalar usual.)

**3.** Sean $\vec{u}_1=(1, 2, 0, 0)$, $\vec{u}_2=(1, 0, 0, 2)$ y $\vec{u}_3=(1, -1, -1, 2)$ tres vectores de $\mathbb{R}^4$. Sea $W=L(\vec{u}_1, \vec{u}_2, \vec{u}_3)$ el subespacio engendrado por $\vec{u}_1, \vec{u}_2, \vec{u}_3$.

*a)* Encontrar una base ortonormal $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ de $W$ usando el método de ortonormalización.

*b)* Extender la base encontrada en *a)* a una base ortonormal de $\mathbb{R}^4$.

**4.** Utilizar el método del teorema de ortogonalización para construir una base ortogonal en el subespacio tridimensional de $\mathbb{R}^4$ generado por los vectores $\vec{u}_1=(1, 2, 1, 3)$, $\vec{u}_2=(4, 1, 1, 1)$ y $\vec{u}_3=(3, 1, 1, 0)$.

**5.** Sea $(\vec{x}, \vec{y})=x_1y_1+(x_1+x_2)(y_1+y_2)+(x_1+x_2+x_3)(y_1+y_2+y_3)$ un producto escalar en $\mathbb{R}^3$. Encontrar una base ortogonal de $M\subset\mathbb{R}$ si:

*a)* $M=\{x\in\mathbb{R}^3 \mid x_1=x_2=x_3\}$

*b)* $M=\{x\in\mathbb{R}^3 \mid x_1+2x_2+3x_3=0\}$

**6.** En $\mathbb{R}^5$ con el producto escalar usual, ortonormalizar los vectores $\vec{u}_1=(1, 0, 1, 0, 0)$, $\vec{u}_2=(0, 1, 0, 1, 0)$ y $\vec{u}_3=(0, 0, 1, 0, 1)$.

**7.** Sean $\vec{u}=(1, 2, 3, 4)$, $\vec{v}=(0, 3, -2, 1)$ y $\vec{w}=(1, 1, 0, 0)\in\mathbb{R}^4$. Encontrar todos los vectores de la forma $\vec{w}+\alpha\vec{u}+\beta\vec{v}$, $\alpha, \beta\in\mathbb{R}$ que son ortogonales a $\vec{u}$ y $\vec{v}$ simultáneamente.

**8.** Sean $\vec{a}=(1, 2, 0, -1, 0)$, $\vec{b}=(0, 1, -1, 1, 0)$ y $\vec{c}=(0, 0, 1, 2, 1)\in\mathbb{R}^5$. Descomponer $\vec{c}$ en dos vectores, uno de ellos combinación lineal de $\vec{a}$ y $\vec{b}$ y el otro ortogonal al anterior.

**9.** Sea $E$ un espacio euclídeo. Demostrar que $\vec{u}+\vec{v}$ y $\vec{u}-\vec{v}$ son ortogonales si y sólo si $\|\vec{u}\|=\|\vec{v}\|$.

**10.** Encontrar una base ortonormal para cada uno de los subespacios siguientes:

*a)* $\{(x, y, z)\in\mathbb{R}^3 \,/\, x+y-z=0\}$ ($\mathbb{R}^3$ tiene el producto escalar usual).

*b)* $\left\{p(x)\in P_{\mathbb{R}}^{(3)}[-1, 1] \,/\, x\dfrac{dp(x)}{dx}=p(x)\right\}$ (en $P_{\mathbb{R}}^{(3)}[-1, 1]$ se considera el producto escalar dado por $\displaystyle\int_{-1}^{1} p(x)\,q(x)\,dx$).

*c)* $\{A\in\mathcal{M}_{3\times 3}(\mathbb{R}) \,/\, \text{traza}(A)=0\}$ (en $\mathcal{M}_{3\times 3}(\mathbb{R})$ se considera el producto escalar dado en el problema 5 de la sección 8.2).

**11.** Demostrar que si $(\vec{x}, \vec{x})=x_1^2+x_2^2+\cdots+x_n^2$ en una base de un espacio euclídeo, esta base es ortonormal.

## 8.4. COMPLEMENTO ORTOGONAL. PROYECCIONES

DEFINICIÓN 1 (*Subespacios ortogonales*)

> Dos subespacios $W_1$ y $W_2$ de un espacio euclídeo $E$ se dicen *ortogonales*, y se escribe $W_1 \perp W_2$, si todos los vectores de $W_1$ son ortogonales a todos los vectores de $W_2$.

En $\mathbb{R}^3$ con el producto escalar usual, un plano y una recta perpendicular a él, que pasen por el origen, son ortogonales (figura 1); sin embargo, dos planos que son perpendiculares en el sentido de la geometría de $\mathbb{R}^3$ no son ortogonales, ya que los vectores $\vec{x}_1$, $\vec{x}_2$ de la figura 2 no son ortogonales.

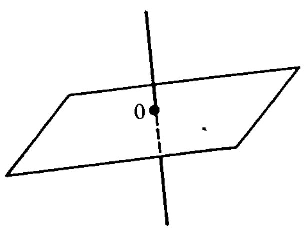


Figura 1

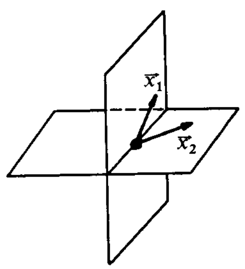


Figura 2



Se tienen las siguientes *propiedades* de los subespacios ortogonales.

1) *$W_1$ y $W_2$ son ortogonales si y sólo si todos los vectores de una base de $W_1$ son ortogonales a cada uno de los vectores de una base de $W_2$.*

En efecto, si $\{\vec{u}_1, ..., \vec{u}_k\}$ es una base de $W_1$ y $\{\vec{v}_1, ..., \vec{v}_j\}$ es una base de $W_2$ y si $\vec{x}\in W_1$, $\vec{y}\in W_2$ podemos escribir $\vec{x}=x_1\vec{u}_1+\cdots+x_k\vec{u}_k$, $\vec{y}=y_1\vec{v}_1+\cdots+y_j\vec{v}_j$; por tanto,

$$(\vec{x}, \vec{y})=\sum_{i=1}^{k}\sum_{l=1}^{j} x_i y_l(\vec{u}_i, \vec{v}_l)=0$$

ya que $(\vec{u}_i, \vec{v}_l)=0$ para todo $i=1, ..., k$, $l=1, ..., j$.

El recíproco es obvio.

2) *Si $W_1 \perp W_2$ se tiene que $W_1\cap W_2=\{\vec{0}\}$.*

Para demostrar esta afirmación suponer que $\vec{x}\in W_1\cap W_2$; entonces $\vec{x}\in W_1$ y $\vec{x}\in W_2$; como $W_1\perp W_2$ se tiene que $(\vec{x}, \vec{x})=0$, y de la propiedad *d)* del producto escalar se deduce que $\vec{x}=\vec{0}$.

Dado un subespacio vectorial $W$ de un espacio euclídeo $E$ el conjunto

$$W^{\perp}=\{\vec{y}\in E \,/\, \vec{y}\perp\vec{x} \text{ para todo } \vec{x}\in W\}$$

es un subespacio vectorial de $E$, que recibe el nombre de *complemento ortogonal de W.* Para demostrar que $W^{\perp}$ es un subespacio vectorial de $E$ basta observar que si $\vec{y}_1, \vec{y}_2\in W^{\perp}$ y $a, b\in\mathbb{R}$,

$$(a\vec{y}_1+b\vec{y}_2, \vec{x})=a(\vec{y}_1, \vec{x})+b(\vec{y}_2, \vec{x})=0$$

y, por tanto, $a\vec{y}_1+b\vec{y}_2\in W^{\perp}$.

EJEMPLO A. En $\mathbb{R}^3$, con el producto escalar usual, queremos encontrar el complemento ortogonal de $W=\{(x_1, x_2, x_3) \,/\, x_1=x_2,\ x_1, x_2, x_3\in\mathbb{R}\}$.

Una base de $W$ es $\vec{u}_1=(1, 1, 0)$, $\vec{u}_2=(0, 0, 1)$; un elemento $\vec{y}$ pertenece a $W^{\perp}$ si y sólo si $\vec{y}$ es perpendicular a $\vec{u}_1$ y a $\vec{u}_2$. Si tomamos $\vec{y}=y_1\vec{e}_1+y_2\vec{e}_2+y_3\vec{e}_3$ se tiene que

$$0=(\vec{y}, \vec{u}_1)=(y_1\vec{e}_1+y_2\vec{e}_2+y_3\vec{e}_3, \vec{e}_1+\vec{e}_2)=y_1+y_2$$
$$0=(\vec{y}, \vec{u}_2)=(y_1\vec{e}_1+y_2\vec{e}_2+y_3\vec{e}_3, \vec{e}_3)=y_3$$

Este sistema es un sistema homogéneo con tres incógnitas, cuya matriz de los coeficientes tiene rango 2; por tanto, determina un subespacio de dimensión 1; este subespacio tiene por ecuaciones $y_3=0$, $y_1=-y_2$ y, por tanto, está generado por el vector $\vec{v}=(-1, 1, 0)$. Así pues,

$$W^{\perp}=L\{-1, 1, 0)\}$$

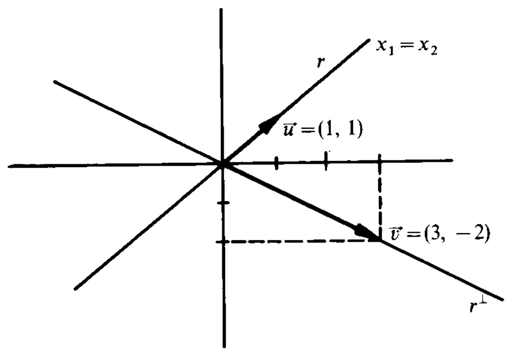
EJEMPLO B. En $\mathbb{R}^2$, con el producto escalar dado en el ejemplo B de la sección 8.1 queremos hallar el complemento ortogonal de la recta $r$: $x_1 = x_2$.

Puesto que $r$ está generada por $\vec{u} = (1, 1)$, tenemos que $\vec{y} = (y_1, y_2) \in r^\perp$ si y sólo si

$$0 = (\vec{y}, \vec{u}) = (y_1, y_2)\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = (y_1 + y_2, y_1 + 2y_2)\begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2y_1 + 3y_2$$

Por tanto, $r^\perp$ es una recta de ecuación $2y_1 = -3y_2$, es decir, generada por

$$\vec{v} = (3, -2)$$




En este ejemplo se observa que el complemento ortogonal de un subespacio vectorial depende en gran medida del producto escalar utilizado. Observar que en este ejemplo los subespacios $r$ y $r^\perp$ no son perpendiculares en el sentido de la geometría clásica estudiada en el capítulo 3.

* * *

El procedimiento seguido en los dos ejemplos anteriores para encontrar el complemento ortogonal de un subespacio vectorial se generaliza a continuación. Sea $E$ un espacio euclídeo con base $B = \{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_n\}$ y sea $W$ un subespacio vectorial de $E$ generado por los $r$ vectores linealmente independientes

$$\vec{v}_1 = \sum_{j=1}^{n} a_{1j}\vec{u}_j,\ \vec{v}_2 = \sum_{j=1}^{n} a_{2j}\vec{u}_j,\ \dots,\ \vec{v}_r = \sum_{j=1}^{n} a_{rj}\vec{u}_j.$$

Los elementos $\vec{x} \in W^\perp$ quedan caracterizados por las relaciones

$$(\vec{v}_1, \vec{x}) = 0,\ (\vec{v}_2, \vec{x}) = 0,\ \dots,\ (\vec{v}_r, \vec{x}) = 0$$

ya que esto es suficiente para demostrar que $\vec{x}$ es perpendicular a todos los elementos de $W$.



Si $\vec{x} = \sum_{i=1}^{n} x_i\vec{u}_i$, las igualdades anteriores se transforman en

$$\left.\begin{aligned}
0 &= \left(\sum_{j=1}^{n} a_{1j}\vec{u}_j, \sum_{i=1}^{n} x_i\vec{u}_i\right) = \sum_{i=1}^{n}\left[\sum_{j=1}^{n} a_{1j}(\vec{u}_j, \vec{u}_i)\right]x_i \\
0 &= \left(\sum_{j=1}^{n} a_{2j}\vec{u}_j, \sum_{i=1}^{n} x_i\vec{u}_i\right) = \sum_{i=1}^{n}\left[\sum_{j=1}^{n} a_{2j}(\vec{u}_j, \vec{u}_i)\right]x_i \\
&\ \vdots \qquad\qquad\qquad \vdots \\
0 &= \left(\sum_{j=1}^{n} a_{rj}\vec{u}_j, \sum_{i=1}^{n} x_i\vec{u}_i\right) = \sum_{i=1}^{n}\left[\sum_{j=1}^{n} a_{rj}(\vec{u}_j, \vec{u}_i)\right]x_i
\end{aligned}\right\}$$

que es un sistema homogéneo de $r$ ecuaciones con $n$ incógnitas $x_1, x_2, \dots, x_n$. Las soluciones de este sistema forman el subespacio ortogonal a $W$.

Si la base $B$ es ortonormal el sistema anterior se transforma en

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
a_{r1}x_1 + a_{r2}x_2 + \cdots + a_{rn}x_n &= 0
\end{aligned} \qquad \text{(I)}$$

*Ejercicio.* Si $W$ es un subespacio vectorial de un espacio euclídeo $E$, demostrar que $(W^\perp)^\perp = W$.

PROPOSICIÓN 2

Sea $W$ un subespacio vectorial de un espacio euclídeo $E$, entonces $E$ es suma directa de $W$ y su complemento ortogonal $W^\perp$.

*Demostración.* Si $\vec{x} \in W \cap W^\perp$ se tiene que $\vec{x} \in W$ y $\vec{x} \in W^\perp$; por tanto, $(\vec{x}, \vec{x}) = 0$, de donde deducimos que $\vec{x} = 0$. Así se comprueba que $W \cap W^\perp = \{\vec{0}\}$.

Falta probar que $W + W^\perp = E$. Para ello sea $B = \{\vec{e}_1, \dots, \vec{e}_r\}$ una base ortogonal de W y completémosla hasta obtener una base ortogonal $\{\vec{e}_1, \dots, \vec{e}_r, \vec{e}_{r+1}, \dots, \vec{e}_n\}$ de $E$; si $\vec{z} \in E$ podemos escribir

$$z = \sum_{i=1}^{n} z_i\vec{e}_i = \sum_{i=1}^{r} z_i\vec{e}^i + \sum_{i=r+1}^{n} z_i\vec{e}_i = \vec{x} + \vec{y}.$$

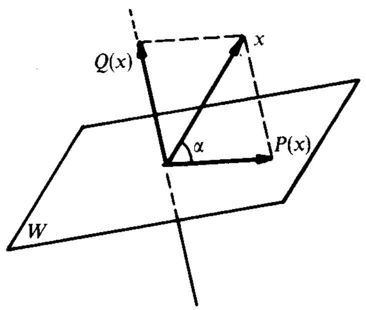
Claramente $\vec{x} \in W$; puesto que, además, $\vec{y} \in W^\perp$ ya que para todo $j = 1, \dots, r$

$$(\vec{e}_j, \vec{y}) = \left(\vec{e}_j, \sum_{i=r+1}^{n} z_i\vec{e}_i\right) = 0$$

se tiene el resultado deseado. ∎

*Observación.* De esta proposición y de la proposición 2 de la sección 5.4 se deduce que $\dim(W)+\dim(W^{\perp})=\dim(E)$.

Sea $W$ un subespacio vectorial de un espacio euclídeo $E$. Puesto que $E=W\oplus W^{\perp}$, de la caracterización de suma directa (ver proposición 4 de la sección 5.4) se deduce que todo $\vec{x}\in E$ posee una descomposición *única* de la forma $\vec{x}=\vec{y}+\vec{z}$ con $\vec{y}\in W$, $\vec{z}\in W^{\perp}$. El vector $\vec{y}$ recibe el nombre de *proyección ortogonal* de $\vec{x}$ sobre $W$ y se escribe $\vec{y}=P_W(\vec{x})\equiv P(\vec{x})$, mientras que $\vec{z}$ es la proyección ortogonal de $\vec{x}$ sobre $W^{\perp}$ y se escribe $\vec{z}=Q_W(\vec{x})\equiv Q(\vec{x})$.




El *ángulo que forma un vector* $\vec{x}\in E$ *con un subespacio vectorial* $W$ se define como el ángulo que forma $\vec{x}$ con $P(\vec{x})$; por tanto,

$$\cos \sphericalangle(\vec{x}, W)=\cos \sphericalangle(\vec{x}, P(\vec{x}))=\frac{(x, P(\vec{x}))}{\|\vec{x}\|\,\|P(\vec{x})\|}=\frac{(\vec{y}+\vec{z}, \vec{y})}{\|\vec{x}\|\,\|\vec{y}\|}=\frac{\|\vec{y}\|^2}{\|\vec{x}\|\,\|\vec{y}\|}=\frac{\|P(\vec{x})\|}{\|\vec{x}\|}$$

En las aplicaciones es conveniente conocer cómo se resuelve el problema de encontrar la proyección $P(\vec{x})$ de un vector $\vec{x}$ sobre un subespacio $W$ que tiene $\{\vec{u}_1, \ldots, \vec{u}_r\}$ como base (no necesariamente ortogonal). La solución de este problema se encuentra de la siguiente manera: como $P(\vec{x})\in W$ hemos de tener

$$P(x)=\lambda_1\vec{u}_1+\cdots+\lambda_r\vec{u}_r;$$

además, $Q(\vec{x})\in W^{\perp}$ y, por tanto, $(Q(\vec{x}), u_j)=0$ para todo $j=1, 2, \ldots, r$. Como $\vec{x}=P(\vec{x})+Q(\vec{x})$ se tiene que $Q(\vec{x})=\vec{x}-P(\vec{x})$ y, por tanto,

$$\left.\begin{array}{c}(\vec{x}, \vec{u}_1)-\lambda_1(\vec{u}_1, \vec{u}_1)-\cdots-\lambda_r(\vec{u}_r, \vec{u}_1)=0\\ \vdots \qquad \vdots \qquad\qquad \vdots \quad \vdots\\ (\vec{x}, \vec{u}_r)-\lambda_1(\vec{u}_1, \vec{u}_r)-\cdots-\lambda_r(\vec{u}_r, \vec{u}_r)=0\end{array}\right\} \qquad \text{(II)}$$



que es un sistema de $r$ ecuaciones con $r$ incógnitas $\lambda_1, \ldots, \lambda_r$; puesto que la proyección $P(\vec{x})$ de un vector $\vec{x}\in E$ es única, este sistema debe tener una única solución; por tanto, su determinante

$$D_r=\begin{vmatrix}(\vec{u}_1, \vec{u}_1) & \cdots & (\vec{u}_r, \vec{u}_1)\\ \vdots & & \vdots\\ (\vec{u}_1, \vec{u}_r) & \cdots & (\vec{u}_r, \vec{u}_r)\end{vmatrix}$$

debe de ser no nulo y las soluciones $j$ vienen dadas por

$$\lambda_j=\frac{1}{D_r}\begin{array}{c} \text{columna } j \\ \begin{vmatrix}(\vec{u}_1, \vec{u}_1), & \ldots, & (\vec{x}, \vec{u}_1), & \ldots, & (\vec{u}_r, \vec{u}_1)\\ \vdots & & \vdots & & \vdots\\ (\vec{u}_1, \vec{u}_r), & \ldots, & (\vec{x}, \vec{u}_r), & \ldots, & (\vec{u}_r, \vec{u}_r)\end{vmatrix}\end{array}, \quad j=1, 2, \ldots, r$$

de acuerdo con la regla de Cramer.

EJEMPLO C. Tratemos de encontrar el ángulo que forma el vector $p(t)=t^2$ con el plano generado por los vectores $p_0(t)=1$, $p_1(t)=t$ en el espacio $P_{\mathbb{R}}^{(3)}([-1, 1])$ de los polinomios de orden menor o igual que tres con el producto escalar de la integral dado en el ejemplo D de la sección 8.1.

Sea $W=L(1, t)$; como $\cos \sphericalangle(t^2, W)=\cos \sphericalangle(t^2, P(t^2))$ comenzamos calculando $P(t^2)$; tenemos que

$$P(t^2)=\lambda_1\cdot 1+\lambda_2 t,$$

de manera que $t^2-P(t^2)$ es perpendicular a $W$; por tanto,

$$0=(1, t^2-P(t^2))=\int_{-1}^{1} t^2\,dt-\int_{-1}^{1}\lambda_1\,dt-\int_{-1}^{1}\lambda_2 t\,dt=\frac{2}{3}-2\lambda_1$$

$$0=(t, t^2-P(t^2))=\int_{-1}^{1} t^3\,dt-\int_{-1}^{1}\lambda_1 t\,dt-\int_{-1}^{1}\lambda_2 t^2\,dt=-\frac{2}{3}\lambda_2$$

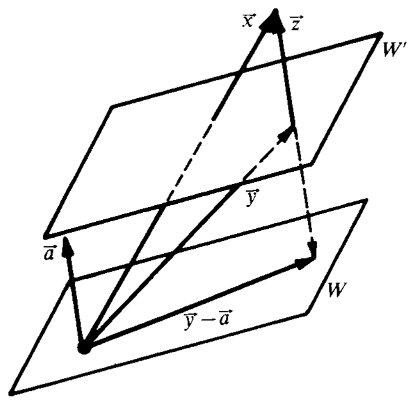
de donde se deduce que $P(t^2)=1/3$.

Una vez calculada la proyección se tiene que

$$\cos \sphericalangle(t^2, W)=\cos \sphericalangle(t^2, 1/3)=\frac{\|1/3\|}{\|t^2\|}=\frac{\sqrt{2}/3}{\sqrt{2}/\sqrt{5}}=\frac{\sqrt{5}}{3}$$

* * *

Ejemplo D. Consideremos el hiperplano $W'$ de $\mathbb{R}^3$ (considerado como un espacio euclídeo con el producto escalar usual) que se obtiene trasladando el subespacio $W$ de ecuaciones $x_1 = x_3$ mediante el vector $\vec{a} = (0, 1, 1)$. Dado $\vec{x} = (2, 1, 2)$ queremos hallar $\vec{y}$, $\vec{z}$ de manera que $\vec{x} = \vec{y} + \vec{z}$ y $\vec{z}$ sea ortogonal a $W$ (ver figura adjunta) e $\vec{y} \in W'$.




Si restamos $\vec{a}$ de la igualdad $\vec{x} = \vec{y} + \vec{z}$ se tiene

$$\vec{x} - \vec{a} = (\vec{y} - \vec{a}) + \vec{z};$$

donde $\vec{y} - \vec{a} \in W$ y $\vec{z} \in W^{\perp}$. Por tanto, $\vec{z}$ coincide con la proyección $Q(\vec{x} - \vec{a})$ de $\vec{x} - \vec{a}$ sobre $W^{\perp}$. Hallamos primero $P(\vec{x} - \vec{a})$; como $W$ está generado por $\vec{e}_1 = (1, 0, 1)$ y $\vec{e}_2 = (0, 1, 0)$ se tiene que

$$P(\vec{x} - \vec{a}) = \lambda_1 \vec{e}_1 + \lambda_2 \vec{e}_2;$$

además:

$$0 = (\vec{e}_1, (\vec{x} - \vec{a}) - P(\vec{x} - \vec{a})) = (2 + 0 + 1) - \lambda_1 2 \Rightarrow \lambda_1 = 3/2$$
$$0 = (\vec{e}_2, (\vec{x} - \vec{a}) - P(\vec{x} - \vec{a})) = 0 - \lambda_2 \cdot 1 \Rightarrow \lambda_2 = 0.$$

Por tanto, $P(\vec{x} - \vec{a}) = \frac{3}{2}\vec{e}_1$, con lo cual

$$\vec{z} = Q(\vec{x} - \vec{a}) = \vec{x} - \vec{a} - P(\vec{x} - \vec{a}) = (2, 0, 1) - \left(\frac{3}{2}, 0, \frac{3}{2}\right) = \left(\frac{1}{2}, 0, -\frac{1}{2}\right).$$

Ahora es sencillo calcular $\vec{y}$ puesto que

$$\vec{y} = \vec{x} - \vec{z} = (2, 1, 2) - \left(\frac{1}{2}, 0, -\frac{1}{2}\right) = \left(\frac{3}{2}, 1, \frac{5}{2}\right)$$

* * *



Observar que el ejemplo D podía haberse realizado siguiendo los procedimientos de geometría clásica en el espacio estudiados en el capítulo 3. La definición axiomática de espacio euclídeo nos da entonces otro procedimiento para resolver el problema de las proyecciones.

**Lema 3**

Sea $W$ un subespacio vectorial de un espacio euclídeo $E$ y sea $\vec{x} \in E$. Si $\vec{x} = \vec{y} + \vec{z}$ con $\vec{y} \in W$, se tiene que

$$\|\vec{z}\| \geqq \|Q(\vec{x})\|$$

*Demostración.* Basta aplicar el teorema de Pitágoras de la siguiente manera:

$$\|\vec{z}\|^2 = \|\vec{x} - \vec{y}\|^2 = \|P(\vec{x}) + Q(\vec{x}) - \vec{y}\|^2 = \|P(\vec{x}) - \vec{y} + Q(\vec{x})\|^2 =$$
$$= \|P(\vec{x}) - \vec{y}\|^2 + \|Q(\vec{x})\|^2 \geqq \|Q(\vec{x})\|^2$$

donde la última igualdad se debe a que $P(\vec{x}) - \vec{y} \in W$ y $Q(\vec{x}) \in W^{\perp}$ ∎

El lema 3 demuestra que si $x = P(\vec{x}) + Q(\vec{x})$ con $P(\vec{x}) \in W$, y $Q(\vec{x}) \in W^1$, $P(\vec{x})$ es la mejor aproximación a $\vec{x}$ de $W$. Para calcular $P(x)$ puede procederse de la manera que se explica a continuación. Tomando una base ortogonal $\{\vec{v}_1, \ldots, \vec{v}_r\}$ de $W$ podemos escribir $P(\vec{x}) = \sum_{i=1}^{r} e_i \vec{v}_i$. Como $Q(\vec{x})$ es perpendicular a $v_j$.

$$(\vec{x}, \vec{v}_j) = (P(\vec{x}) + Q(\vec{x}), \vec{v}_j) = (P(\vec{x}), \vec{v}_j) = C_j \|\vec{v}_j\|^2$$

y, por tanto, los coeficientes de $P(\vec{x})$ se calculan con la fórmula

$$C_j = \frac{(\vec{x}, \vec{v})}{\|v_j\|^2} \qquad j = 1, 2, \ldots, r$$

que se llaman componentes de $\vec{x}$ en la dirección de $\vec{v}_j$, $j = 1, \ldots, r$.

Ejemplo E. Tratemos de encontrar la función de la forma $a_1 + a_2$ sen $x$ + $+ a_3$ cos $x + a_4$ sen $2x + a_5$ cos $2x$ que mejor aproxime a la función

$$f(x) = \begin{cases} -1 & \text{si} \quad -\pi \leqq x < 0 \\ 1 & \text{si} \quad 0 \leqq x \leqq \pi \end{cases} \qquad x \in [-\pi, \pi]$$

según el producto escalar. La mejor aproximación es aquella en que los $a_j$ son las componentes de $f$ en la dirección de 1, sen $x$, cos $x$, sen $2x$ y cos $2x$. En este caso las componentes $a_j$ reciben el nombre de coeficientes de Fourier de $f$.

Como

$$(f, 1) = \int_{-\pi}^{\pi} f(x)\, dx = \int_{-\pi}^{0} -1\, dx + \int_{0}^{\pi} 1\, dx = -\pi + \pi = 0$$

el primer coeficiente $a_1$ es nulo.
Como

$$\|\text{sen } x\|^2 = \int_{-\pi}^{\pi} \text{sen}^2 x\, dx = \int_{-\pi}^{\pi} \frac{1 - \cos 2x}{2}\, dx = \pi$$

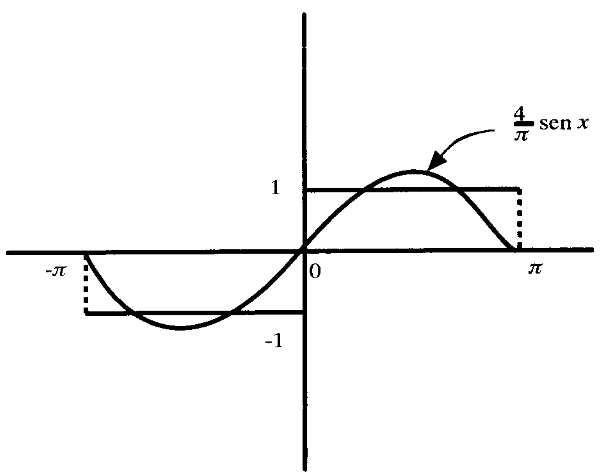
el segundo coeficiente es:

$$a_2 = \frac{1}{\pi}(f, \text{sen } x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x) \text{ sen } x\, dx = \frac{1}{\pi}\left[\int_{\pi}^{0} -\text{sen } x\, dx + \int_{0}^{\pi} \text{sen } x\, dx\right] =$$

$$= \frac{1}{\pi}[2 + 2] = \frac{4}{\pi}$$

De manera similar el lector puede calcular el resto de los coeficientes, que serán todos nulos. Por tanto, la mejor aproximación a $f$ en el espacio generado por $\{1, \text{sen } x, \cos x, \text{sen } 2x, \cos 2x\}$ es $\frac{4}{\pi}$ sen $x$.

En el gráfico adjunto se dibujan la función $f$ y su aproximación $\frac{4}{\pi}$ sen $x$.




Añadiendo funciones trigonométricas pueden lograrse mejores «aproximaciones» de esta función.



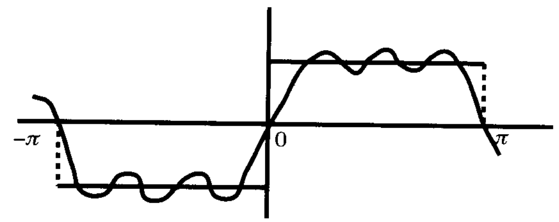
Con una cantidad infinita de funciones trigonométricas se consigue la aproximación

$$f(x) \sim \frac{4}{\pi}\sum_{n=0}^{\infty} \frac{1}{2n + 1} \text{ sen } (2n + 1)x$$

Pueden dibujarse algunas de las primeras sumas parciales de esta serie para comprobar que cada vez se aproximan más a la función $f$. En la figura siguiente se dibuja

$$\frac{4}{\pi} \text{ sen } x + \frac{4}{3\pi} \text{ sen } 3x + \frac{4}{5\pi} \text{ sen } 5x$$

que consta de los tres primeros sumandos de la serie anterior.




Poniendo $x = \frac{\pi}{2}$ se obtiene la aproximación

$$\frac{\pi}{4} \sim 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} + \frac{1}{11} + \cdots$$

en la que los denominadores son todos los números impares.

EJEMPLO F. El lema 3 se utilizará en este ejemplo. Dada una función $f(t)$ definida en el intervalo $[a, b]$, y $t_1, t_2, \ldots, t_n \in [a, b]$, se trata de encontrar un polinomio $p(t)$ de grado $k < n$ y tal que la *desviación cuadrática media* de la función $f$, definida por

$$\sigma(f; p)^2 = \sum_{j=1}^{n} [f(t_j) - p(t_j)]^2$$

sea mínima.
Si en el espacio de todas las funciones definidas en los puntos $t_1, t_2, \ldots, t_n$, definimos

$$(f, g) = \sum_{j=1}^{n} f(t_j)g(t_j), \quad f, g \text{ funciones}$$

se tiene un espacio euclídeo de dimensión $n$. En este espacio euclídeo el problema planteado se reduce a encontrar la proyección $P(f)$ de $f$ sobre el subespacio de todos los polinomios de grado no superior a $k$, ya que entonces $\|Q(f)\|^2=\sigma(f, P(f))^2$ es mínimo, de acuerdo con el lema 3.

Este problema ya sabemos resolverlo: se tendría

$$P(f)=\lambda_0+\lambda_1 t+\cdots+\lambda_k t^k$$

donde los $\lambda_k$ son soluciones de un sistema semejante a (II).

Realizamos un ejemplo sencillo para ilustrar las afirmaciones anteriores. Supongamos que queremos encontrar un polinomio de grado 2 de manera que la desviación cuadrática media de la función $f(t)=\cos \pi t$ en los puntos $t_1=0$, $t_2=1$, $t_3=2$ sea mínima para este polinomio en el intervalo $[0, 2]$.

El polinomio mínimo $P(f)$ será de la forma

$$P(f)=\lambda_0+\lambda_1 t+\lambda_2 t^2$$

de manera que

$$\begin{aligned}&(f, 1)-\lambda_0(1, 1)-\lambda_1(t, 1)-\lambda_2(t^2, 1)=0\\&(f, t)-\lambda_0(1, t)-\lambda_1(t, t)-\lambda_2(t^2, t)=0\\&(f, t^2)-\lambda_0(1, t^2)-\lambda_1(t, t^2)-\lambda_2(t^2, t^2)=0\end{aligned}$$

Utilizando el producto escalar anteriormente definido obtenemos

$$\left.\begin{aligned}1-\lambda_0 3-\lambda_1 3-\lambda_2 5&=0\\1-\lambda_0 3-\lambda_1 5-\lambda_2 9&=0\\3-\lambda_0 5-\lambda_1 9-\lambda_2 17&=0\end{aligned}\right\}.$$

Puesto que

$$\begin{vmatrix}3&3&5\\3&5&9\\5&9&17\end{vmatrix}=\begin{vmatrix}3&3&5\\0&2&4\\2&6&12\end{vmatrix}=3\begin{vmatrix}2&4\\6&12\end{vmatrix}+2\begin{vmatrix}3&5\\2&4\end{vmatrix}=4$$

tenemos:

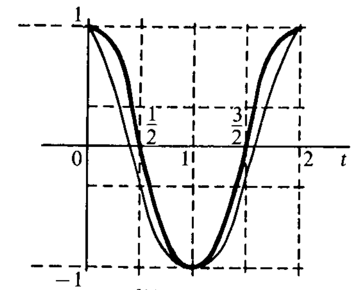
$$\lambda_0=\frac{1}{4}\begin{vmatrix}1&3&5\\1&5&9\\3&9&17\end{vmatrix}=\frac{1}{4}\begin{vmatrix}1&3&5\\0&2&4\\2&6&12\end{vmatrix}=\frac{1}{4}\cdot 2[12-10]=1$$

$$\lambda_1=\frac{1}{4}\begin{vmatrix}3&1&5\\3&1&9\\5&3&17\end{vmatrix}=\frac{1}{4}\begin{vmatrix}3&1&5\\0&0&4\\5&3&17\end{vmatrix}=\frac{-1}{4}\cdot 4\cdot 4=-4$$

$$\lambda_2=\frac{1}{4}\begin{vmatrix}3&3&1\\3&5&1\\5&9&3\end{vmatrix}=\frac{1}{4}\begin{vmatrix}3&3&1\\0&2&0\\5&9&3\end{vmatrix}=\frac{1}{4}\cdot 2\cdot 4=2$$



Por tanto, $P(f)=1-4t+2t^2$, que ha sido dibujado con trazo fino en la figura adjunta.




$f(t)=\cos \pi t$

## EJERCICIOS 8.4

**1.** Encontrar el complemento ortogonal del subespacio $W$ de $E$ cuando:

*a)* $E=\mathbb{R}^3$, $W=L(\vec{u}, \vec{v})$ con $\vec{u}=(1, 0, 1)$, $\vec{v}=(2, -1, 1)$.

*b)* $E=\mathbb{R}^4$, $W=\{x\in\mathbb{R}^4 \,/\, x_1+x_2+x_3+x_4=0 \text{ y } 2x_1-x_2=0\}$.

En ambos espacios se considera el producto escalar usual.

**2.** Sea $W$ como en 1.*b)* y $W(3, 2)=\{x\in\mathbb{R}^4/x_1+x_2+x_3+x_4=3,\ 2x_1-x_2=2\}$ una subvariedad afín trasladada de $W$.

*a)* Encontrar la expresión analítica, en la base canónica de $\mathbb{R}^4$, de la proyección ortogonal sobre $W$.

*b)* Similarmente sobre $W(3, 2)$.

**3.** Sea $E$ un espacio euclídeo, $A$ y $B$ subconjuntos de $E$ y $W$ y $Z$ subespacios vectoriales de $E$. Demostrar:

*a)* $A^\perp=\{\vec{u}\in E \,/\, (\vec{u}, \vec{v})=0,\ \forall \vec{v}\in A\}$ es un subespacio vectorial de $E$:

*a)* $A^{\perp\perp}=(A^\perp)^\perp=L(A)$, es decir, el subespacio engendrado por $A$.

*b)* $A\subset B \Rightarrow B^\perp\subset A^\perp$.

*c)* $(W+Z)^\perp=W^\perp\cap Z^\perp$.

*d)* $(W\cap Z)^\perp=W^\perp+Z^\perp$.

**4.** En $\mathbb{R}^4$ descomponer $\vec{f}=(5, 2, -2, 2)$ en una suma de dos vectores, un vector $\vec{g}$ perteneciente al subespacio vectorial generado por $\vec{b}_1=(2, 1, 1, -1)$ y $\vec{b}_2=(1, 1, 3, 0)$ y un vector $\vec{h}$ ortogonal a este subespacio.

**5.** Encontrar el complemento ortogonal de cada uno de los subespacios del ejercicio 10 de la sección 8.3. Además

*a)* En $\mathscr{M}_{3\times 3}(\mathbb{R})$ con el producto escalar $\langle A, B\rangle$ = traza$(AB)^t$, hallar el complemento ortogonal del subespacio de las matrices diagonales de $\mathscr{M}_{3\times 3}(\mathbb{R})$.

*b)* Determinar la proyección ortogonal y la mínima distancia de una matriz $X=(x_{ij})\in\mathscr{M}_{3\times 3}(\mathbb{R})$ al subespacio de las matrices diagonales.

**6.** En $\mathbb{R}^4$ con el producto escalar usual determinar las ecuaciones de la proyección ortogonal sobre el subespacio generado por $(1, 1, -1, 0)$ y $(0, 0, 2, 1)$.

**7.** Sea $E$ un espacio euclídeo y $W$ un subespacio vectorial de $E$; designar por $P_W$ la proyección ortogonal sobre $W$. Demostrar que $(P_W(\vec{x}), \vec{y}) = (\vec{x}, P_W(\vec{y}))$ para todo $\vec{x}, \vec{y} \in E$.

**8.** Sea $E$ un espacio euclídeo; $E^* = L(E, \mathbb{R}) = \{f\colon E \mapsto \mathbb{R} \mid f \text{ lineal}\}$ se denomina el espacio dual de $E$.

> *Teorema de representación de Riesz.* Para cada $f \in E^*$ existe un único $\vec{u}_f \in E$ tal que $f(\vec{v}) = (\vec{u}_f, \vec{v})$, $\forall \vec{v} \in E$. La aplicación $f \mapsto \vec{u}_f$ de $E^* \mapsto E$ es biyectiva y satisface $\vec{u}_{f+g} = \vec{u}_f + \vec{u}_g$ y $\vec{u}_{\lambda f} = \lambda \vec{u}_f$ (luego es lineal).

Sea $E = \mathbb{R}^3$ y $f \in E^*$ definido por $f(\vec{e}_1) = 2$, $f(\vec{e}_2) = 1$, $f(\vec{e}_3) = -1$. Encontrar $\vec{u}_f$.

**9.** Sea $E$ un espacio euclídeo, $W \subset E$ un subespacio, $P\colon E \mapsto E$ la proyección ortogonal sobre $W$. Demostrar:

*a)* $\forall \vec{u} \in E$, $\|\vec{u} - P(\vec{u})\| \leq \|\vec{u} - \vec{w}\|$, $\forall \vec{w} \in W$ ($P(\vec{u})$ es el punto de $W$ a distancia mínima de $\vec{u}$).

**10.** En lo que sigue se supone: (a) que no se conoce la descomposición $E = W \oplus W^{\perp}$ y (b) que para cada $\vec{u} \in E$ existe al menos un $\vec{z} \in W$ tal que

$$\|\vec{u} - \vec{z}\| \leq \|\vec{u} - \vec{w}\|, \quad \forall \vec{w} \in W \qquad (*)$$

*a)* Probar que para cada $\vec{u}$ existe un único $\vec{z} \in W$ satisfaciendo (*).

*b)* Sea $Q\colon E \mapsto E$ definido por $Q(\vec{u})$ como el único $\vec{z} \in W$ satisfaciendo (*). Probar que $(\vec{u} - Q(\vec{u}), \vec{w}) = 0$, $\forall \vec{w} \in W$.

*c)* Probar que $Q^2 = Q$ y $Q \in L(E)$ (aplicaciones lineales de $E$ en $E$) y, por tanto, es una proyección. Además, $\operatorname{img} Q = W$. *Nota*: debido a *c)*, $Q = P$ es la proyección ortogonal sobre $W$.

## 8.5. ADJUNTA DE UNA APLICACION

DEFINICIÓN 1

> Sea $E$ un espacio euclídeo y $A$ una aplicación lineal de $E$ en $E$, es decir, $A \in L(E)$; una aplicación $A' \in L(E)$ se dice que es *adjunta* de $A$ si para todo $\vec{x}, \vec{y} \in E$ se tiene
>
> $$(A\vec{x}, \vec{y}) = (\vec{x}, A'\vec{y})$$



La definición anterior plantea dos preguntas inmediatas:

1.ª ¿Existe siempre una aplicación adjunta de una dada?
2.ª Si existe una aplicación adjunta de $A$, ¿es única?

Comenzaremos contestando a la segunda de manera afirmativa; para probar que la adjunta de $A$ es única supongamos que existen aplicaciones lineales $B$ y $C$ tales que

$$(A\vec{x}, \vec{y}) = (\vec{x}, B(\vec{y})), \quad (A(\vec{x}), \vec{y}) = (\vec{x}, C(\vec{y}));$$

entonces tenemos $(\vec{x}, B(\vec{y})) = (\vec{x}, C(\vec{y}))$, o equivalentemente $(\vec{x}, B(\vec{y}) - C(\vec{y})) = 0$ para todo $\vec{x}, \vec{y} \in E$; tomando $\vec{x} = (B - C)(\vec{y})$ se tiene

$$((B-C)(\vec{y}), (B-C)\vec{y}) = 0$$

y por la propiedad *d)* del producto escalar se tiene $(B-C)(\vec{y}) = 0 \Rightarrow B(\vec{y}) = C(\vec{y})$, lo cual prueba que $B$ y $C$ coinciden.

En el razonamiento anterior se ha obtenido el siguiente resultado, que merece la pena destacar para su uso futuro:

LEMA 2

> Si $B, C \in L(E)$ y son tales que $(\vec{x}, B(\vec{y})) = (\vec{x}, C(\vec{y}))$ para todo $\vec{x}, \vec{y} \in E$, se tiene $B = C$.

La pregunta primera tiene también una respuesta afirmativa si $\dim(E) = n < \infty$. Para demostrarlo, sea $\{\vec{e}_1, \ldots, \vec{e}_n\}$ una base ortonormal de $E$ y $A = (a_{ij})_{j=1,\ldots,n}^{i=1,\ldots,n}$ la matriz de $A$ en esta base; sea $A^t$ la traspuesta de la matriz $A$ y sea $A'$ la aplicación lineal de matriz $A^t$ en esta misma base; demostraremos que $A'$ es la aplicación adjunta de $A$.

Tenemos que

$$(A(\vec{e}_i), \vec{e}_k) = (a_{1i}\vec{e}_1 + \cdots + a_{ni}\vec{e}_n, \vec{e}_k) = a_{ki}$$

y

$$(\vec{e}_i, A'(\vec{e}_k)) = (\vec{e}_i, a_{k1}\vec{e}_1 + \cdots + a_{kn}\vec{e}_n) = a_{ki}$$

con lo cual tenemos $(A(\vec{e}_i), \vec{e}_k) = (\vec{e}_i, A'(\vec{e}_k))$ para todo $i, k = 1, 2, \ldots, n$. Como todo vector de $E$ es combinación lineal de los vectores de la base, de la igualdad que acabamos de probar se deduce que $A'$ es la adjunta de $A$; en efecto, si

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i, \quad \vec{y} = \sum_{j=1}^{n} y_j \vec{e}_j$$

tenemos que

$$(A(\vec{x}), \vec{y}) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (A(\vec{e}_i), \vec{e}_j) = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\vec{e}_i, A'(\vec{e}_j)) = (x, A'(\vec{y}))$$

lo cual prueba el resultado deseado.

Hemos demostrado el siguiente resultado:

**PROPOSICIÓN 3**

Para toda aplicación lineal $A \in L(E)$, donde $E$ es espacio euclídeo de dimensión finita, existe una única aplicación adjunta $A'$ de $A$, cuya matriz en una base ortonormal es la transpuesta de la matriz de $A$.

* * *

A continuación damos algunas propiedades de la aplicación adjunta de una dada.

a) $I' = I$, donde $I$ denota la aplicación identidad (basta observar que $(I(\vec{x}), \vec{y}) = (\vec{x}, \vec{y}) = (\vec{x}, I(\vec{y}))$ para todo $\vec{x}, \vec{y} \in E$).

b) $(A')' = A$, ya que

$$(A(\vec{x}), \vec{y}) = (\vec{x}, A'(\vec{y})) = (A'(\vec{y}), \vec{x}) = (\vec{y}, (A')'(\vec{x})) = ((A')'(\vec{x}), \vec{y}) \text{ para todo } \vec{x}, \vec{y} \in E.$$

c) $(A+B)' = A' + B'$, ya que

$$(\vec{x}, (A+B)'\vec{y}) = ((A+B)\vec{x}, \vec{y}) = (A\vec{x}, \vec{y}) + (B\vec{x}, \vec{y}) = (\vec{x}, A'(\vec{y})) + (\vec{x}, B'\vec{y}) = (\vec{x}, (A'+B')\vec{y})$$

para todo $\vec{x}, \vec{y} \in E$.

d) $(A \circ B)' = B' \circ A'$, ya que

$$(\vec{x}, (A \circ B)'\vec{y}) = (A \circ B(\vec{x}), \vec{y}) = (B\vec{x}, A'(\vec{y})) = (\vec{x}, B' \circ A'(\vec{y}))$$

para todo $\vec{x}, \vec{y} \in E$.

e) *Si $A$ posee inversa, se tiene* $(A^{-1})' = (A')^{-1}$.
Como $A \circ A^{-1} = I$ se deduce de d) y a) que $(A \circ A^{-1})' = (A^{-1})' \circ A' = I' = I$, lo cual prueba que la invesa de $A'$ es la adjunta de la inversa.

*Nota.* Las propiedades anteriores tienen su traducción para matrices transpuestas; por ejemplo, d) se escribe como $(AB)^t = B^t A^t$.

* * *



## 8.6. APLICACIONES AUTOADJUNTAS

**DEFINICIÓN 1** *(Aplicación autoadjunta)*

Sea $E$ un espacio euclídeo y $A \in L(E)$; la aplicación $A$ se dice *autoadjunta* si su adjunta $A'$ coincide con $A$, esto es, $A' = A$. De forma equivalente,

$$(A\vec{x}, \vec{y}) = (\vec{x}, A\vec{y})$$

para todo $\vec{x}, \vec{y} \in E$.

Si $A$ es la matriz de una aplicación $A \in L(E)$ en una base ortonormal, sabemos que $A^t$ es la matriz de $A'$ en esta misma base; si $A$ es autoadjunta se ha de tener $A = A^t$. Por tanto, la matriz de una aplicación autoadjunta en una base ortonormal es simétrica.

En la sección anterior se ha demostrado que la aplicación adjunta de la identidad $I \in L(E)$ es ella misma; tenemos en $I$ el primer ejemplo de una aplicación autoadjunta.

Las aplicaciones autoadjuntas tienen las siguientes propiedades:

1. La suma de aplicaciones autoadjuntas es autoadjunta, ya que $(A+B)' = A' + B' = A + B$.
2. La aplicación inversa de una aplicación autoadjunta invertible es autoadjunta, ya que $(A^{-1})' = (A')^{-1} = A^{-1}$.
3. La composición de dos aplicaciones autoadjuntas es autoadjunta si y sólo si las aplicaciones conmutan, ya que $(AB)' = B'A' = BA$ y, por tanto, $(AB') = AB$ si y sólo si $BA = AB$.

En el ejercicio 11 de la sección 7.2 se pide demostrar que los autovalores de toda matriz *simétrica* de orden 2 son reales; además, del mismo problema se deduce que toda matriz real simétrica de orden 2 es diagonalizable. Nos preguntamos si toda matriz simétrica (o equivalentemente toda aplicación autoadjunta) es diagonalizable en los números reales con independencia de su orden. En esta sección daremos una respuesta afirmativa a tal pregunta y demostraremos, además, que los vectores de la nueva base pueden tomarse ortonormales.

En lo que sigue, y mientras no se indique lo contrario, $A$ denota una aplicación lineal de un espacio euclídeo $E$ en sí mismo.

**TEOREMA 1**

Si $W$ es un subespacio vectorial de $E$ invariante con respecto a $A \in L(E)$, su complemento ortogonal $W^{\perp}$ es invariante respecto de la aplicación adjunta $A'$ de $A$.

*Demostración.* Sea $\vec{y} \in W^{\perp}$; queremos demostrar que $A'(\vec{y}) \in W^{\perp}$. Para todo $\vec{x} \in W$ se tiene que $(\vec{x}, A'(\vec{y})) = (A(\vec{x}), \vec{y}) = 0$, ya que $\vec{y} \in W^{\perp}$ y $A(\vec{x}) \in W$ por ser $W$ invariante con respecto a $A$; la igualdad $(x, A'(\vec{y})) = 0$ para todo $\vec{x} \in W$ implica que $A'(\vec{y}) \in W^{\perp}$, que era lo que queríamos demostrar. ■

Si $A$ es una aplicación autoadjunta, se tiene $A'=A$, y del teorema 1 se deduce inmediatamente el siguiente corolario:

COROLARIO 2

Si $A$ es autoadjunta y $W$ es un subespacio vectorial de $E$ invariante con respecto a $A$, $W^{\perp}$ es invariante con respecto a $A$.

TEOREMA 3

Toda aplicación autoadjunta $A$ en un espacio euclídeo de dimensión finita tiene todos sus autovalores reales.

*Demostración.* Sea $A$ la matriz de la aplicación $A$. Supongamos que esta matriz posee al menos un autovalor complejo, que denotaremos por $\alpha+i\beta$ con $\beta\neq 0$. En la sección 7.5 del capítulo 7 se demostró que existen vectores $\vec{u}$, $\vec{v}$, no nulos simultáneamente, tal que

$$A(\vec{u})=\alpha\vec{u}-\beta\vec{v}, \quad A(\vec{v})=\beta\vec{u}+\alpha\vec{v}.$$

De aquí deducimos

$$(A\vec{u}, \vec{v})=\alpha(\vec{u}, \vec{v})-\beta(\vec{v}, \vec{v})$$
$$(\vec{u}, A\vec{v})=\beta(\vec{u}, \vec{u})+\alpha(\vec{u}, \vec{v})$$

Puesto que $A$ es autoadjunta se tiene que $(A\vec{u}, \vec{v})=(\vec{u}, A\vec{v})$ y, por tanto, $0=\beta(\|\vec{u}\|^2+\|\vec{v}\|^2)$. De aquí deducimos que $\beta=0$, que contradice el hecho de que $\beta$ es no nulo. Esta contradicción prueba el resultado deseado. ■

TEOREMA 4 *(Diagonalización de matrices simétricas)*

La matriz de una aplicación autoadjunta puede ser reducida en una base ortonormal determinada a una matriz diagonal.

*Observación.* Del teorema 4 se deduce que toda matriz simétrica es diagonalizable.

*Demostración.* Sea $A$ una aplicación autoadjunta y $\lambda_1$ un autovalor de $A$; por el teorema 3, $\lambda_1$ es real. Sea $\vec{e}_1$ un autovector de $A$ correspondiente a $\lambda_1$ de manera que sea unitario.

Sea $W_1=L(\vec{e}_1)$ el subespacio generado por $\vec{e}_1$; como $A$ es autoadjunta, del corolario 2 se deduce que $W_1^{\perp}$ es invariante respecto de $A$.

La aplicación $A$ restringida a $W_1^{\perp}$ sigue siendo autoadjunta y, por tanto, podemos elegir $\vec{e}_2\in W_1^{\perp}$ tal que $\vec{e}_2$ sea un autovector unitario de autovalor real $\lambda_2$ de $A$.

Tomamos $W_2=L(\vec{e}_1, \vec{e}_2)$ como el subespacio generado por $\vec{e}_1$ y $\vec{e}_2$; $W_2^{\perp}$ es invariante respecto de $A$ y el proceso puede repetirse para obtener $\lambda_3$, autovalor real, y $\vec{e}_3$, autovector unitario, ortogonal a $\vec{e}_1$ y $\vec{e}_2$.

Después de repetir este proceso tantas veces como la dimensión del espacio



euclídeo se tiene una base de autovectores ortonormales, y en esta base la matriz de $A$ tiene la forma

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots \\ & & & \lambda_n \end{pmatrix}$$ ■

EJEMPLO A. Reducir a su forma diagonal mediante una base ortonormal la matriz simétrica

$$A=\begin{pmatrix} 1 & \sqrt{8} \\ \sqrt{8} & 3 \end{pmatrix}$$

Sus autovalores son las soluciones de $0=|A-\lambda I|=(1-\lambda)(3-\lambda)-8=\lambda^2-4\lambda-5$; por tanto, $\lambda_1=5$, $\lambda_2=-1$. Puesto que

$$\begin{pmatrix} -4 & \sqrt{8} \\ \sqrt{8} & -2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow -4x+\sqrt{8}y=0 \Rightarrow \ker(A-5I)=L\left\{\vec{u}_1=\begin{pmatrix} 2 \\ \sqrt{8} \end{pmatrix}\right\}$$

tomamos $\vec{e}_1\in L(\vec{u}_1)$ unitario:

$$\vec{e}_1=\frac{\vec{u}_1}{\|\vec{u}_1\|}=\frac{1}{\sqrt{12}}\begin{pmatrix} 2 \\ \sqrt{8} \end{pmatrix}=\begin{pmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{pmatrix}$$

Puesto que

$$\begin{pmatrix} 2 & \sqrt{8} \\ \sqrt{8} & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow 2x+\sqrt{8}y=0 \Rightarrow \ker(A+I)=L\left\{\vec{u}_2=\begin{pmatrix} -4 \\ \sqrt{8} \end{pmatrix}\right\}$$

tomamos $\vec{e}_2\in L(\vec{u}_2)$ unitario:

$$\vec{e}_2=\frac{\vec{u}_2}{\|\vec{u}_2\|}=\frac{1}{\sqrt{24}}\begin{pmatrix} -4 \\ \sqrt{8} \end{pmatrix}=\begin{pmatrix} -2/\sqrt{6} \\ \sqrt{2}/\sqrt{6} \end{pmatrix}$$

Observar que $\vec{e}_1$ y $\vec{e}_2$ son ortonormales. La matriz de $A$ en la base $\{\vec{e}_1, \vec{e}_2\}$ es

$$J=\begin{pmatrix} 5 & 0 \\ 0 & -1 \end{pmatrix}$$

Por tanto, en la base $\{\vec{e}_1, \vec{e}_2\}$ $A$ produce una «dilatación» en el «eje» $\vec{e}_1$ que se obtiene multiplicando por 5, y una «simetría» con respecto a la recta $L\{\vec{e}_1\}$ (ver figuras 1 y 2).

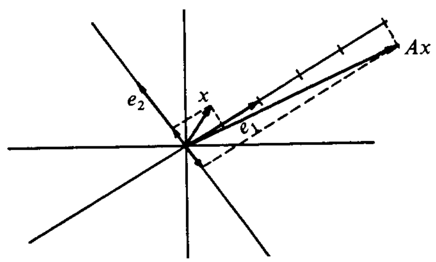



Figura 1

La imagen de la figura $OABC$ es $OA'B'BC'$ marcada en trazo grueso (figura 2).

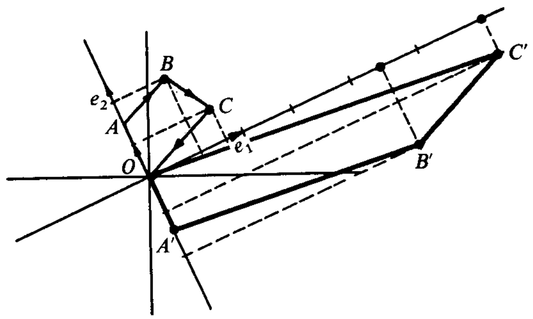



Figura 2

Podemos ahora dar la interpretación geométrica de las aplicaciones autoadjuntas; en la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$ del teorema 4, si $x = x_1\vec{e}_1 + \cdots + x_n\vec{e}_n$:

$$A\vec{x} = x_1\lambda_1\vec{e}_1 + \cdots + x_n\lambda_n\vec{e}_n$$

con lo cual en esta base $A$ se reduce a $n$ «dilataciones» a lo largo de los nuevos ejes; las dilataciones pueden o no cambiar el sentido según que el autovalor sea positivo o negativo.

## EJERCICIOS 8.5 Y 8.6

**1.** Encontrar la aplicación adjunta de las siguientes aplicaciones:

*a)* $T\colon \mathbb{R}^3 \mapsto \mathbb{R}^3$, $T(x, y, z) = (x+y+z,\ x+2y+2z,\ x+2y+3z)$

*b)* $A\colon \mathbb{R}^3 \mapsto \mathbb{R}^3$, $A(x, y, z) = (-y+z,\ -x+2z,\ x+2y)$

*c)* $A\colon P_{\mathbb{R}}^{(2)}[x] \mapsto P_{\mathbb{R}}^{(2)}[x]$, $A(p(x)) = x\dfrac{d}{dx}p(x) - \dfrac{d}{dx}(xp(x))$

*d)* $T\colon \mathcal{M}_{3\times 3}(\mathbb{R}) \mapsto \mathcal{M}_{3\times 3}(\mathbb{R})$, $T(A) = A^t + A$

*Nota.* Los productos escalares de $\mathbb{R}^3$, $P_{\mathbb{R}}^{(2)}[x]$ y $\mathcal{M}_{3\times 3}(\mathbb{R})$ son como en el ejercicio 10 de la sección 8.3.



**2.** Sean $\vec{u}_1 = (1, 1, 0)$, $\vec{u}_2 = (1, 0, 1)$ y $\vec{u}_3 = (1, 2, 0)$ los vectores de una base de $\mathbb{R}^3$. Estudiar si la aplicación $A \in L(\mathbb{R}^3)$ es o no es autoadjunta, cuando su matriz asociada en la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ es

*a)* $A = \begin{pmatrix} 1 & -4 & 2 \\ 2 & 2 & 3 \\ 0 & 3 & 0 \end{pmatrix}$ *b)* $\begin{pmatrix} -4 & -5 & -6 \\ 4 & 2 & 7 \\ 3 & 5 & 4 \end{pmatrix}$

**3.** Diagonalizar en una base ortonormal cada una de las siguientes aplicaciones, demostrando en primer lugar que son autoadjuntas:

*a)* $A\colon \mathbb{R}^2 \mapsto \mathbb{R}^2$, $A(x, y) = (2x+y,\ 2y+x)$

*b)* $A\colon \mathbb{R}^3 \mapsto \mathbb{R}^3$, $A(x, y, z) = (y+z,\ x+z,\ x+y)$

*c)* $T\colon \mathcal{M}_{2\times 2}(\mathbb{R}) \mapsto \mathcal{M}_{2\times 2}(\mathbb{R})$, $T(A) = A^t$

**4.** Sea $T$ una aplicación lineal de un espacio euclídeo $E$ en sí mismo. Demostrar que $T + T'$ es autoadjunta.

**5.** Con $T$ como en el problema 4, demostrar que

$$\ker(T') = [\operatorname{img}(T)]^{\perp} \quad \text{e} \quad \operatorname{img}(T') = [\ker(T)]^{\perp}$$

**6.** Demostrar que toda matriz simétrica $S \in \mathcal{M}_{n\times n}(\mathbb{R})$ determina una aplicación lineal autoadjunta en todo espacio euclídeo de dimensión $n$.

## 8.7. APLICACIONES ORTOGONALES: PARTE I

Sabemos que en todo espacio euclídeo $E$ existe el concepto de longitud de un vector; es natural, por tanto, preguntarse cuáles son aquellas aplicaciones que conservan la longitud de los vectores de $E$. A las aplicaciones que conservan la longitud de los vectores de $E$ las llamaremos *ortogonales*; su estudio, así como su significado geométrico, serán el objetivo principal de las dos próximas secciones.

En lugar de dar la definición utilizando la longitud de los vectores, la daremos utilizando la noción de producto escalar. Como mostraremos más adelante, ambas definiciones son equivalentes.

En lo que sigue $E$ se utiliza para designar un espacio euclídeo y $A$ una aplicación lineal de $E$ en $E$; además, cuando sea necesario utilizar la matriz de la aplicación $A$ se entiende que $E$ es de dimensión finita.

DEFINICIÓN 1

Una aplicación lineal $A$ de un espacio euclídeo $E$ se llama *ortogonal* si $(A(\vec{x}), A(\vec{y})) = (\vec{x}, \vec{y})$ para todo $\vec{x}, \vec{y} \in E$ (es decir, $A$ conserva el producto escalar).

Puesto que si $A$ es ortogonal,

$$\|\vec{x}\|^2=(\vec{x}, \vec{x})=(A(\vec{x}), A(\vec{x}))=\|A(\vec{x})\|^2$$

y

$$\cos \measuredangle (\vec{x}, \vec{y})=\frac{(\vec{x}, \vec{y})}{\|\vec{x}\|\|\vec{y}\|}=\frac{(A(\vec{x}), A(\vec{y}))}{\|A(\vec{x})\|\,\|A(\vec{y})\|}=\cos \measuredangle (A(\vec{x}), A(\vec{y}))$$

deducimos que toda aplicación ortogonal conserva las longitudes de los vectores y los ángulos entre ellos. Se tiene, además, la siguiente proposición:

PROPOSICIÓN 2

Toda aplicación lineal $A$ en un espacio euclídeo $E$ que conserve la longitud de los vectores es una aplicación ortogonal.

*Demostración.* Para todo $\vec{x}, \vec{y} \in A$ se tiene

$$\|A(\vec{x}+\vec{y})\|^2=(A(\vec{x})+A(\vec{y}), A(\vec{x})+A(\vec{y}))=\|A(\vec{x})\|^2+2(A(\vec{x}), A(\vec{y}))+\|A(\vec{y})\|^2$$

y

$$\|\vec{x}+\vec{y}\|^2=(\vec{x}+\vec{y}, \vec{x}+\vec{y})=\|\vec{x}\|^2+2(\vec{x}, \vec{y})+\|\vec{y}\|^2.$$

Teniendo en cuenta que $A$ conserva la longitud de los vectores y, por tanto, $\|A(\vec{x}+\vec{y})\| = \|\vec{x}+\vec{y}\|$, $\|A(\vec{x})\|=\|\vec{x}\|$ y $\|A(\vec{y})\|=\|\vec{y}\|$ se deduce que

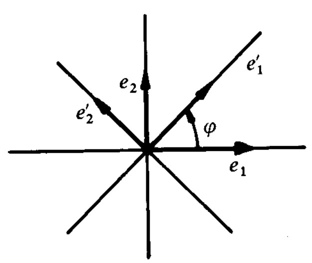
$$(A(\vec{x}), A(\vec{y}))=(\vec{x}, \vec{y})$$

lo cual prueba que $A$ es ortogonal. ■

Ha llegado el momento de dar algunos ejemplos de aplicaciones ortogonales, aparte del ejemplo trivial de la aplicación identidad.

EJEMPLO A. Cualquier *giro* en $\mathbb{R}^2$ (con centro el origen de coordenadas) es una aplicación ortogonal. Este resultado, que es geométricamente intuitivo, puede demostrarse algebraicamente. Para ello suponer que el giro es un giro de ángulo $\varphi$ en el sentido positivo (contrario a las agujas de un reloj) y que, por tanto, su matriz viene dada (respecto de una base ortonormal) por

$$G_\varphi=\begin{pmatrix}\cos\varphi & -\text{sen}\,\varphi\\ \text{sen}\,\varphi & \cos\varphi\end{pmatrix}.$$






Para todo $\vec{x}=(x_1, x_2)\in\mathbb{R}^2$ se tiene

$$\|G_\varphi(\vec{x})\|^2=\left\|\begin{pmatrix}\cos\varphi & -\text{sen}\,\varphi\\ \text{sen}\,\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\right\|^2=\left\|\begin{pmatrix}x_1\cos\varphi - x_2\,\text{sen}\,\varphi\\ x_1\,\text{sen}\,\varphi + x_2\cos\varphi\end{pmatrix}\right\|^2=$$
$$=(x_1\cos\varphi-x_2\,\text{sen}\,\varphi)^2+(x_1\,\text{sen}\,\varphi+x_2\cos\varphi)^2=x_1^2+x_2^2=\|\vec{x}\|^2.$$

Esto prueba que $G_\varphi$ conserva la longitud de los vectores y por la proposición 2 debe de ser ortogonal.

EJEMPLO B. En $\mathbb{R}^2$ cualquier *simetría* con respecto a un subespacio vectorial unidimensional es una aplicación ortogonal.

Encontremos primero las ecuaciones de la simetría $S_{\vec{u}}$ con respecto a $W=L(\vec{u})$. Dado $\vec{x}\in\mathbb{R}^2$,

$$\vec{x}+\vec{y}=P_W(\vec{x})$$

donde $P_W(\vec{x})$ es la proyección de $\vec{x}$ sobre $W$ e $\vec{y}=P_W(\vec{x})-\vec{x}$; además,

$$\vec{x}+2\vec{y}=S_{\vec{u}}(\vec{x})$$

como puede apreciarse en la figura 1. Estas igualdades producen

$$S_{\vec{u}}(\vec{x})=2P_W(\vec{x})-\vec{x}=(2P_W-I)(\vec{x})$$

Para encontrar $P_W(\vec{x})$ escribimos $P_W(\vec{x})=\lambda\vec{u}$ e imponemos la condición de que $\vec{y}=P_W(\vec{x})-\vec{x}=\lambda\vec{u}-\vec{x}$ sea perpendicular a $\vec{u}$:

$$0=(\vec{u}, \lambda\vec{u}-\vec{x})=\lambda\|\vec{u}\|^2-(\vec{u}, \vec{x})$$

Por tanto,

$$P_W(\vec{x})=\frac{(\vec{u}, \vec{x})}{\|\vec{u}\|^2}\vec{u}$$

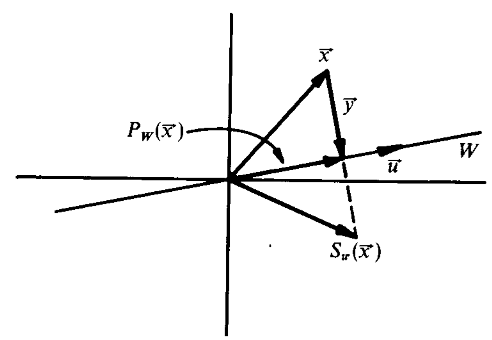



Figura 1

De aquí deducimos que la matriz de $P_W$ con respecto a la base canónica $\{\vec{e}_1, \vec{e}_2\}$ de $\mathbb{R}^2$ es

$$\begin{pmatrix} \dfrac{u_1^2}{\|\vec{u}\|^2} & \dfrac{u_1u_2}{\|\vec{u}\|^2} \\ \dfrac{u_1u_2}{\|\vec{u}\|^2} & \dfrac{u_2^2}{\|\vec{u}\|^2} \end{pmatrix}$$

donde $\vec{u} = u_1\vec{e}_1 + u_2\vec{e}_2$.

Puesto que $S_{\vec{u}}(\vec{x}) = (2P_W - I)(\vec{x})$ la matriz de $S_{\vec{u}}$ con respecto a la base canónica es

$$\begin{pmatrix} \dfrac{2u_1^2}{\|\vec{u}\|^2} - 1 & \dfrac{2u_1u_2}{\|\vec{u}\|^2} \\ \dfrac{2u_1u_2}{\|\vec{u}\|^2} & \dfrac{2u_2^2}{\|\vec{u}\|^2} - 1 \end{pmatrix} = \frac{1}{\|\vec{u}\|^2}\begin{pmatrix} u_1^2 - u_2^2 & 2u_1u_2 \\ 2u_1u_2 & u_2^2 - u_1^2 \end{pmatrix}$$

Podemos comprobar ahora que $S_{\vec{u}}$ es una aplicación ortogonal puesto que si $\vec{x} = (x_1, x_2) \in \mathbb{R}^2$,

$$S_{\vec{u}}(\vec{x}) = \frac{1}{\|\vec{u}\|^2}\begin{pmatrix} u_1^2 - u_2^2 & 2u_1u_2 \\ 2u_1u_2 & u_2^2 - u_1^2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} (u_1^2 - u_2^2)x_1 + 2u_1u_2x_2 \\ 2u_1u_2x_1 - (u_2^2 - u_1^2)x_2 \end{pmatrix}\frac{1}{\|\vec{u}\|^2}$$

implica que

$$\|S_{\vec{u}}(\vec{x})\|^2 = \frac{1}{\|\vec{u}\|^4}[(u_1^2 - u_2^2)^2x_1^2 + 4u_1^2u_2^2x_1^2 + 4u_1^2u_2^2x_2^2 + (u_2^2 - u_1^2)x_2^2] =$$

$$= \frac{1}{\|\vec{u}\|^4}[(u_1^2 + u_2^2)^2x_1^2 + (u_1^2 + u_2^2)^2x_2^2] = x_1^2 + x_2^2 = \|\vec{x}\|^2$$

*Nota.* Si hubiéramos elegido $\{\vec{u}, \vec{u}^{\perp}\}$ como la base en $\mathbb{R}^2$, la matriz de $S_{\vec{u}}$ con respecto a esta base sería

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

La matriz de $S_{\vec{u}}$ con respecto a la base canónica en $\mathbb{R}^2$ puede encontrarse a partir de aquí realizando un cambio de base.

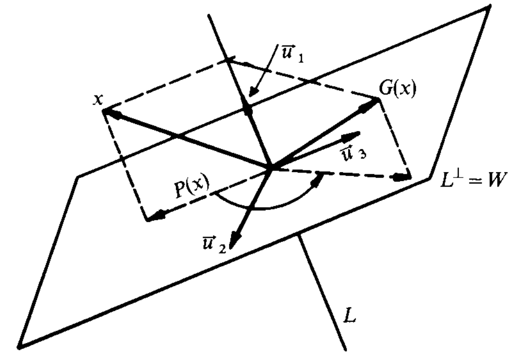
Si el lector intenta encontrar más aplicaciones ortogonales en $\mathbb{R}^2$ tendrá serias dificultades. Quizá puede pensar en la simetría con respecto al origen de coordenadas, pero puede convencerse fácilmente que esta simetría es un giro de 180° y, por tanto, ha sido incluida en el ejemplo *A*. De hecho demostraremos en la próxima sección *que las únicas aplicaciones ortogonales de* $\mathbb{R}^2$ *son los giros y las simetrías con respecto a una recta (o simetrías axiales).* (Así se explica el lector las dificultades anteriores.)

* * *



En $\mathbb{R}^3$ las posibilidades de aplicaciones ortogonales son mayores; a continuación daremos algunos ejemplos.

EJEMPLO C. En $\mathbb{R}^3$ el giro $G$ con respecto a una «recta» (subespacio vectorial de dimensión 1) de amplitud $\varphi$ es una aplicación ortogonal.

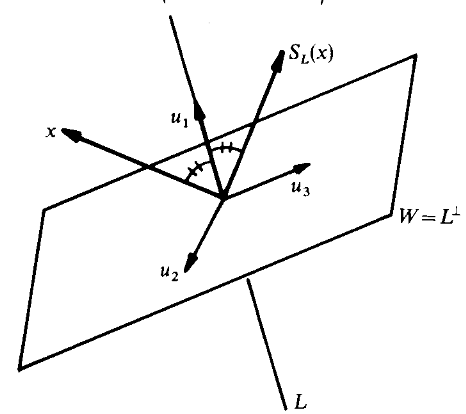



Supongamos que $L = L(\overline{u}_1)$ y que $W = L^\perp = L(\overline{u}_2, \overline{u}_3)$, de manera que $\overline{u}_1$, $\overline{u}_2$ y $\overline{u}_3$ son ortonormales; entonces $\{\overline{u}_1, \overline{u}_2, \overline{u}_3\}$ es una base ortonormal de $\mathbb{R}^3$ y en esta base

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}$$

EJEMPLO D. La simetría $S_L$ con respecto a una recta $L$, en $\mathbb{R}^3$, es una aplicación ortogonal. En la base $\{\overline{u}_1, \overline{u}_2. \overline{u}_3\}$ elegida como en la figura adjunta se tiene:

$$S_L = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$



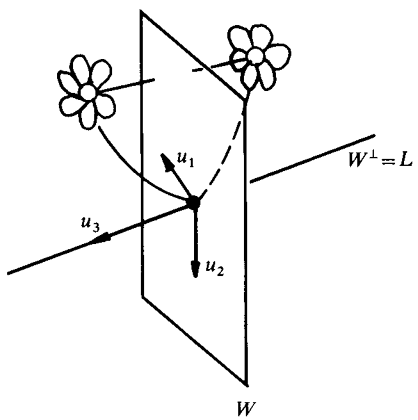

$\{\overline{u}_1, \overline{u}_2, \overline{u}_3\}$ base ortonormal

Observar que esta simetría es en realidad un giro de 180°, y, por tanto, es un caso particular del ejemplo C.

EJEMPLO E. La imagen tridimensional que produce un espejo es una aplicación ortogonal; se trata simplemente de una simetría con respecto al plano del espejo.

Si $W$ es el plano, $S_W$ designa esta simetría, y elegimos una base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ como en la figura se tiene

$$S_W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$




* * *

¿Serán todas estas las posibles aplicaciones ortogonales de $\mathbb{R}^3$? Recomendamos al lector que trate de encontrar alguna más y le advertimos que la respuesta a la pregunta anterior la encontrará más adelante en la próxima sección. No sólo llegaremos a la determinación de las posibles aplicaciones ortogonales de $\mathbb{R}^3$ sino de cualquier espacio euclídeo de dimensión finita.

Este trabajo requiere el estudio de los autovalores de aplicaciones ortogonales. Antes de comenzar este estudio daremos algunas propiedades útiles de las aplicaciones ortogonales.

Es claro que toda aplicación ortogonal transforma toda base ortonormal en una base ortonormal; el recíproco también es cierto y está contenido en la proposición siguiente:

PROPOSICIÓN 3

Toda aplicación lineal $A$ que transforma al menos una base ortonormal en una base ortonormal es ortogonal.



*Demostración.* Sea $\{\vec{e}_1, ..., \vec{e}_n\}$ una base ortonormal tal que $\{\vec{e}'_1, ..., \vec{e}'_n\}$, donde $\vec{e}'_j = A(\vec{e}_j)$, $j = 1, 2, ..., n$, es también ortonormal. Para todo

$$\vec{x} = \sum_{i=1}^{n} x_i \vec{e}_i, \quad \vec{y} = \sum_{j=1}^{n} y_j \vec{e}_j$$

del espacio euclídeo se tiene

$$(A(\vec{x}), A(\vec{y})) = \sum_{i,j=1}^{n} x_i y_j (A(\vec{e}_i), A(\vec{e}_j)) = \sum_{i,j=1}^{n} x_i y_j (\vec{e}'_i, \vec{e}'_j) = \sum_{i=1}^{n} x_i y_i$$

mientras que

$$(x, y) = \sum_{i,j=1}^{n} x_i y_j (\vec{e}_i, \vec{e}_j) = \sum_{i=1}^{n} x_i y_i$$

y, por tanto, $A$ conserva el producto escalar. ■

Si $A$ es una aplicación ortogonal y $A'$ denota su adjunta (ver sección 8.5) se tiene que

$$(x, y) = (A(\vec{x}), A(\vec{y})) = (x, A' \circ A(\vec{y}))$$

para todo $x$, $y$. Del lema 2 de la sección 8.5 se deduce que $A' \circ A = I$. Recíprocamente, si $A' \circ A = I$, $A$ es una aplicación ortogonal. De aquí se deduce que toda aplicación ortogonal tiene una inversa, a saber: $A^{-1} = A'$ y, por tanto, es no degenerada.

La inversa de una aplicación ortogonal es también ortogonal puesto que

$$A^{-1} \circ (A^{-1})' = A^{-1} \circ (A')^{-1} = (A' \circ A)^{-1} = I$$

y la composición o producto de dos aplicaciones ortogonales $A$ y $B$ es también ortogonal puesto que

$$(A \circ B)'(A \circ B) = B' \circ (A' \circ A) \circ B = B' \circ I \circ B = I$$

Sin embargo, la suma de dos aplicaciones ortogonales no es, en general, una aplicación ortogonal (dar un ejemplo).

Sea $A$ la matriz de $A$ en una base ortonormal; puesto que $A' \circ A = I$ deducimos que $A^t A = I$ o equivalentemente $A^t = A^{-1}$. Toda matriz $A$ que satisface una cualquiera de estas dos ecuaciones se llama *ortogonal*. El determinante de una matriz ortogonal es igual a 1 o a $-1$, ya que

$$|A|^2 = |A||A| = |A^t||A| = |A^t A| = |I| = 1.$$

El conjunto de las matrices ortogonales de orden $n$ con elementos reales se designa por $O(n; \mathbb{R})$ y, por tanto,

$$O(n; \mathbb{R}) = \{A \in \mathcal{M}_{n\times n}(\mathbb{R}) \quad / \quad A^t A = I\}$$

En el conjunto $O(n; \mathbb{R})$ la multiplicación de matrices es una operación interna, ya que el producto de dos matrices ortogonales es otra matriz ortogonal: basta observar que este resultado ha sido demostrado para aplicaciones ortogonales. Esta operación es asociativa, puesto que lo es la multiplicación de matrices (ver capítulo 1); la matriz identidad es ortogonal, ya que $I^t I = I \cdot I = I$ y la inversa de una matriz ortogonal es también una matriz ortogonal, ya que anteriormente se ha demostrado para aplicaciones ortogonales.

Por poseer estas propiedades, el conjunto $O(n; \mathbb{R})$ se dice que es un *grupo* con respecto a la multiplicación de matrices.

Ejemplos de grupos han aparecido implícitamente en los capítulos anteriores. Se invita al lector a localizarlos.

* * *

Procedemos ahora al estudio de los autovalores de aplicaciones ortogonales.

TEOREMA 4

Los autovalores reales de una aplicación ortogonal son iguales a 1 o a $-1$.

*Demostración.* Si $\lambda$ es un autovalor real de una aplicación ortogonal $A$ con autovector $\vec{x}$ se tiene

$$(\vec{x}, \vec{x}) = (A(\vec{x}), A(\vec{x})) = (\lambda\vec{x}, \lambda\vec{x}) = \lambda^2(\vec{x}, \vec{x})$$

Por tanto, $\lambda^2 = 1$ y $\lambda = \pm 1$. ■

*Observación.* La matriz de una aplicación ortogonal puede tener autovalores complejos; por ejemplo, el giro del ejemplo A tiene como autovalores $\lambda = \cos\varphi \pm i \operatorname{sen}\varphi$. Si dispusiéramos de un «producto escalar» en un espacio vectorial complejo podríamos demostrar que los autovalores complejos de toda matriz ortogonal tienen módulo 1. Esto se hará en el capítulo siguiente.

TEOREMA 5

Si un subespacio vectorial $W$ de un espacio euclídeo $E$ es invariante respecto a una aplicación ortogonal $A$, $W^\perp$ es también invariante respecto de $A$.

*Demostración.* Si $\vec{y} \in W^\perp$ y $\vec{x} \in W$ se tiene

$$(\vec{x}, A(\vec{y})) = (A \circ A^{-1}(\vec{x}), A(\vec{y})) = (A^{-1}(\vec{x}), \vec{y}).$$



Puesto que $A$ es invertible, $\dim(A^{-1}(W)) = \dim W$; además, $W \subset A^{-1}(W)$, ya que $W$ es invariante respecto de $A$. De aquí deducimos que $A^{-1}(W) = W$. Por tanto, $A^{-1}(\vec{x}) \in W$ y se tiene

$$(\vec{x}, A(\vec{y})) = (A^{-1}(\vec{x}), \vec{y}) = 0$$

Así pues, $A(\vec{y}) \in W^\perp$. ■

## 8.8. APLICACIONES ORTOGONALES: PARTE II

En esta sección clasificaremos las aplicaciones ortogonales en un espacio euclídeo $E$ de dimensión 2 ó 3. La clasificación consistirá en encontrar todos los posibles tipos de aplicaciones ortogonales en $E$, de manera que toda aplicación ortogonal sea de uno de estos tipos.

Comenzaremos clasificando las aplicaciones ortogonales en un espacio euclídeo de dimensión 2 y deduciendo de aquí algunos resultados de la geometría plana. A continuación daremos un teorema general sobre la forma de Jordan real de aplicaciones ortogonales en un espacio euclídeo de dimensión $n$, del cual deduciremos la clasificación de las aplicaciones ortogonales en un espacio euclídeo de dimensión 3.

Sea $A$ una aplicación ortogonal en un espacio euclídeo $E$ de dimensión 2 (en particular $\mathbb{R}^2$), que tiene como matriz

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

en una base ortonormal $B = \{\vec{e}_1, \vec{e}_2\}$ de $E$. Puesto que $A^t A = I$ se tiene que

$$A^t A = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Deducimos de aquí las igualdades

$$a_{11}^2 + a_{21}^2 = 1$$
$$a_{12}^2 + a_{22}^2 = 1$$
$$a_{11}a_{12} + a_{21}a_{22} = 0$$

Podemos elegir $\varphi$ y $\psi$ de manera que las dos primeras igualdades sean ciertas, es decir,

$$a_{11} = \cos\varphi, \quad a_{21} = \operatorname{sen}\varphi, \quad a_{12} = \cos\psi, \quad a_{22} = \operatorname{sen}\psi.$$

Sustituyendo en la tercera se obtiene

$$\cos\varphi\cos\psi + \operatorname{sen}\varphi\operatorname{sen}\psi = \cos(\psi - \varphi) = 0$$

Por tanto, $\psi-\varphi=\frac{\pi}{2}$ o $\psi-\varphi=\frac{3\pi}{2}$. Si $\psi-\varphi=\frac{\pi}{2}$

$$A=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}$$

y, por tanto, $A$ es la *rotación de ángulo* $\varphi$ *alrededor del origen de coordenadas.* (En particular, si $\varphi=0$ se tiene la aplicación identidad y si $\varphi=\pi$ se tiene la simetría respecto del origen de coordenadas.)

En el segundo caso, esto es, $\psi-\varphi=\frac{3\pi}{2}$,

$$A=\begin{pmatrix}\cos\varphi & \operatorname{sen}\varphi\\ \operatorname{sen}\varphi & -\cos\varphi\end{pmatrix}$$

$A$ es una matriz simétrica y, por tanto, corresponde a una aplicación autoadjunta; en la sección 8.6 se demostró que $A$ tiene como matriz

$$J=\begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}\quad,\quad \lambda_1.\ \lambda_2\in\mathbb{R}$$

en una (nueva) base ortonormal; por el teorema 4 de la sección anterior, $\lambda_1=1$ o $-1$ y $\lambda_2=1$ o $-1$. Además,

$$\lambda_1\lambda_2=|J|=|A|=\begin{vmatrix}\cos\varphi & \operatorname{sen}\varphi\\ \operatorname{sen}\varphi & -\cos\varphi\end{vmatrix}=-1$$

y, por tanto, $\lambda_1$ y $\lambda_2$ tienen signo opuesto. En consecuencia, en una (posiblemente nueva) base ortonormal $\{\vec{e}_1', \vec{e}_2'\}$ $A$ tiene como matriz

$$\bar{A}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

que representa una *simetría con respecto al «eje»* $\vec{e}_1'$.

En resumen, tenemos el siguiente resultado:

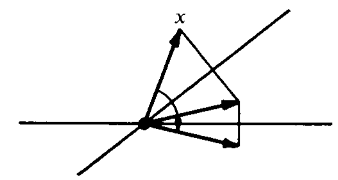
TEOREMA 1

Toda aplicación ortogonal de $\mathbb{R}^2$ es una *rotación* de ángulo $\varphi$ alrededor del origen de coordenadas (el determinante de esta aplicación es 1) o bien una *simetría axial* (con determinante $-1$).

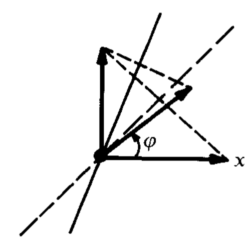
De este teorema se deducen resultados de la geometría elemental plana que nos serán útiles en el capítulo 10, que estará dedicado al estudio de los movimientos en el plano y en el espacio. Enunciamos alguno de ellos dejando la demostración para el lector:



1) La composición de dos simetrías axiales es una rotación de ángulo el doble del ángulo comprendido entre los ejes de simetría (observar que la composición de dos simetrías tiene determinante 1).




2) La composición de una rotación con respecto al origen y una simetría axial es de nuevo una simetría axial (observar que su determinante es $-1$).




3) La composición de dos giros con respecto al origen es otro giro con el mismo centro.

* * *

Pasamos ahora a clasificar las aplicaciones ortogonales en espacios euclídeos de cualquier dimensión finita.

TEOREMA 2

La matriz de una aplicación ortogonal se reduce en una base ortonormal determinada a la forma

$$A=\begin{pmatrix}
\begin{matrix}1 & & \\ & \ddots & \\ & & 1\end{matrix} & & & & 0\\
 & \begin{matrix}-1 & & \\ & \ddots & \\ & & -1\end{matrix} & & & \\
 & & \begin{matrix}\cos\varphi_1, & -\operatorname{sen}\varphi_1\\ \operatorname{sen}\varphi_1, & \cos\varphi_1\end{matrix} & & \\
 & & & \ddots & \\
0 & & & & \begin{matrix}\cos\varphi_k, & -\operatorname{sen}\varphi_k\\ \operatorname{sen}\varphi_k, & \cos\varphi_k\end{matrix}
\end{pmatrix}$$

*Demostración.* Procedemos por inducción en la dimensión $n$ del espacio euclídeo. Si $n=1$ es claro, puesto que sólo podemos tener $\lambda=\pm 1$. Para $n=2$ el teorema 2 coincide con el teorema 1, que ya hemos demostrado.

Sea ahora $E$ un espacio euclídeo de dimensión $n$ y $A$ una aplicación ortogonal. Pueden ocurrir dos casos:

1. *La aplicación $A$ tiene un valor propio real $\lambda=\pm 1$* (esto ocurre necesariamente si $n$ es impar). Si $\vec{e}_1$ es un autovector unitario correspondiente a este autovalor tomamos $W=L(\vec{e}_1)$; como $W^\perp$ es invariante por $A$ (ver sección 8.7) la hipótesis de inducción nos dice que la matriz de $A|_{W^\perp}$ es de la forma del teorema 2. Añadiéndole el vector $\vec{e}_1$ se obtiene una base de $E$ en la cual $A$ tiene la matriz deseada.
2. *La aplicación $A$ no tiene valores propios reales. $A$* debe de poseer al menos un autovalor complejo (y su conjugado) y en consecuencia tiene también un subespacio $W$ de dimensión 2 invariante. Por el teorema 2,

$$A|_W=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix} \quad \text{o} \quad A|_W=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

en una base ortonormal $\{\vec{e}_1, \vec{e}_2\}$ adecuada. El segundo caso no puede darse, puesto que ello implicaría la existencia de autovalores reales. Por tanto,

$$A|_W=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}$$

Como $W^\perp$ es invariante (sección 8.7) podemos aplicarle la hipótesis de inducción para obtener una base de $W^\perp$ en la cual $A|_{W^\perp}$ es como en el teorema 2. Añadiéndole $\{\vec{e}_1, \vec{e}_2\}$ adecuadamente se obtiene el resultado deseado. ■

* * *

Podemos ahora estudiar las posibles aplicaciones ortogonales en $\mathbb{R}^3$. Del teorema 2 se desprende que sus matrices en una base ortonormal $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ adecuada han de tener una de las siguientes formas:

$$a)\ \begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix} \quad b)\ \begin{pmatrix}1 & & \\ & 1 & \\ & & -1\end{pmatrix} \quad c)\ \begin{pmatrix}1 & & \\ & -1 & \\ & & -1\end{pmatrix} \quad d)\ \begin{pmatrix}-1 & & \\ & -1 & \\ & & -1\end{pmatrix}$$

$$e)\ \begin{pmatrix}1 & 0 & 0\\ 0 & \cos\varphi & -\operatorname{sen}\varphi\\ 0 & \operatorname{sen}\varphi & \cos\varphi\end{pmatrix} \quad f)\ \begin{pmatrix}-1 & 0 & 0\\ 0 & \cos\varphi & -\operatorname{sen}\varphi\\ 0 & \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}$$

Estas formas tienen la siguiente interpretación: *a)* es la identidad; *b)* es una simetría con respecto al plano determinado por $\{\vec{e}_1, \vec{e}_2\}$; *e)* es un giro de ángulo $\varphi$ alrededor del



eje $\vec{e}_1$ (en particular *c)* es un giro de ángulo 180° alrededor de $\vec{e}_1$); *f)* es un giro de ángulo $\varphi$ alrededor de $\vec{e}_1$ seguido de una simetría con respecto al plano determinado por $\{\vec{e}_2, \vec{e}_3\}$ (es decir, composición de *e)* y *b)*); *d)* es un caso particular de *f)* con $\varphi=180°$ y corresponde a una simetría con respecto al origen de coordenadas.

¿Había deducido el lector todos los casos posibles?

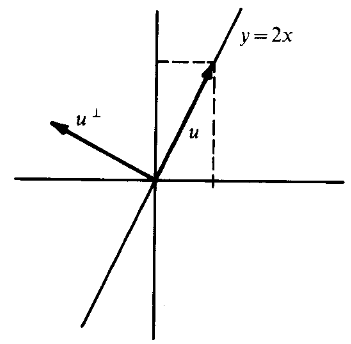
* * *

El lector que ha llegado hasta aquí debe probar los conocimientos adquiridos intentando por su cuenta y riesgo los ejercicios que a continuación realizamos.

EJERCICIO A. Encontrar en la base canónica de $\mathbb{R}^2$ la matriz de la simetría (ortogonal) con respecto a la recta $2x-y=0$.

*Solución.* La recta está generada por el vector $\vec{u}=\begin{pmatrix}1\\2\end{pmatrix}$; podríamos aplicar el resultado del ejemplo B, sección 8.7, pero preferimos hacerlo de nuevo. En la base ortonormal

$$\left\{\vec{v}_1=\frac{\vec{u}}{\|\vec{u}\|},\ \vec{v}_2=\frac{\vec{u}^\perp}{\|\vec{u}\|}\right\}$$




donde $\vec{u}^\perp=\begin{pmatrix}-2\\1\end{pmatrix}$ la matriz de esta simetría es

$$S'=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$$

Puesto que

$$\vec{v}_1=\begin{pmatrix}1/\sqrt{5}\\ 2/\sqrt{5}\end{pmatrix} \quad , \quad \vec{v}_2=\begin{pmatrix}-2/\sqrt{5}\\ 1/\sqrt{5}\end{pmatrix}$$

la matriz de la simetría en la base canónica (la antigua) es

$$S=\begin{pmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{pmatrix}^{-1}=$$

$$=\begin{pmatrix}1/\sqrt{5} & -2/\sqrt{5}\\ 2/\sqrt{5} & 1/\sqrt{5}\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}1/\sqrt{5} & 2/\sqrt{5}\\ -2/\sqrt{5} & 1/\sqrt{5}\end{pmatrix}=\frac{1}{5}\begin{pmatrix}-3 & 4\\ 4 & 3\end{pmatrix}$$

EJERCICIO B. Demostrar que la matriz

$$A=\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}$$

es ortogonal; interpretar geométricamente la aplicación $A$ que tiene a $A$ como matriz en la base canónica de $\mathbb{R}^3$.

*Solución.* El cálculo

$$A^tA=\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}=I$$

demuestra que $A$ es ortogonal. Su interpretación geométrica se deduce del estudio de sus autovalores. Como $A$ es simétrica sus autovalores han de ser reales y, por tanto, iguales a $\pm 1$ (sección 8.7). Un cálculo adecuado muestra que $\lambda_1=1$, $\lambda_2=-1$ (doble) son sus autovalores.

Para $\lambda_1=1$,

$$\begin{pmatrix}-1 & 1 & 0\\ 1 & -1 & 0\\ 0 & 0 & -2\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}\Rightarrow \left.\begin{matrix}x=y\\ z=0\end{matrix}\right\}\Rightarrow \ker(A-I)=L\left\{\vec{u}_1=\begin{pmatrix}1\\ 1\\ 0\end{pmatrix}\frac{1}{\sqrt{2}}\right\}$$

Para $\lambda_2=-1$:

$$\begin{pmatrix}1 & 1 & 0\\ 1 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 0\end{pmatrix}\Rightarrow x=-y\Rightarrow \ker(A+I)=L\left\{\vec{u}_2=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix},\ \vec{u}_3=\begin{pmatrix}1\\ -1\\ 0\end{pmatrix}\frac{1}{\sqrt{2}}\right\}$$



Como

$$J=\begin{pmatrix}1 & & \\ & -1 & \\ & & -1\end{pmatrix}$$

se trata de un giro de 180° alrededor de $L(\vec{u}_1)$ o, equivalentemente, una simetría respecto a la recta $L(\vec{u}_1)$.

* * *

Terminamos esta sección con una observación. Dada una matriz ortonormal $O\in O(n;\ \mathbb{R})$, del teorema 2 de esta misma sección se deduce que existe una matriz $C$ tal que

$$O=CAC^{-1}$$

donde $A$ es como en el teorema 2, es decir, una forma de Jordan real de $O$, y $C$ es la matriz del cambio de base de la base canónica de $\mathbb{R}^n$ a una base ortonormal $B=\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ de $\mathbb{R}^n$. Por tanto, la aplicación lineal en $\mathbb{R}^n$ que tiene como matriz $C$ transforma una base ortonormal en otra base ortonormal; esto es suficiente para asegurar que $C$ es una matriz ortogonal.

Debido al teorema 4 de la sección 8.6 el mismo razonamiento anterior muestra que si $S$ es una matriz simétrica con valores reales, existe una matriz diagonal $D$, y una matriz ortogonal $C$ tal que

$$S=CDC^{-1}$$

Si $C$ es ortogonal se tiene, además, que $C^{-1}=C^t$. Se tiene así el siguiente resultado:

PROPOSICIÓN 3

1) Dada $O\in O(n;\ \mathbb{R})$ existen $C\in O(n;\ \mathbb{R})$ y $A$ como en el teorema 2 tal que $O=CAC^t$.
2) Dada $S\in \mathcal{M}_{n\times n}(\mathbb{R})$ y tal que $S=S^t$, existen $C\in O(n;\ \mathbb{R})$ y $D\in\mathcal{M}_{n\times n}(\mathbb{R})$ diagonal, tal que $S=CDC^t$.

## EJERCICIOS 8.7 Y 8.8

**1.** Sean $\vec{u}_1=(1, 1, 0)$, $\vec{u}_2=(1, 0, 1)$ y $\vec{u}_3=(1, 2, 0)$ los vectores de una base de $\mathbb{R}^3$. Estudiar si $A\in L(\mathbb{R}^3)$ es o no es ortogonal cuando su matriz en la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ está dada por :

a) $\begin{pmatrix} 4 & 3 & 6 \\ -1 & -1 & -1 \\ -2 & -2 & -3 \end{pmatrix}$ b) $\begin{pmatrix} 0 & -5 & 1 \\ 0 & 1 & 0 \\ 1 & 3 & 1 \end{pmatrix}$ c) $\begin{pmatrix} 1 & 0 & -1 \\ 1 & 2 & 3 \\ 0 & 1 & 1 \end{pmatrix}$

**2.** Sea $T \in L(\mathbb{R}^3)$ una aplicación lineal cuya matriz en la base canónica de $\mathbb{R}^3$ es

$$T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}$$

Demostrar que $T$ es ortogonal. Encontrar una base ortonormal de $\mathbb{R}^3$ en la que $T$ tenga como matriz la forma canónica de Jordan (real).

**3.** En un espacio vectorial euclídeo $E$ de dimensión 3 se considera la base ortonormal $B = \{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ y $f: E \mapsto E$ lineal, definida por

$$3f(\vec{e}_1) = 2\vec{e}_1 - 2\vec{e}_2 + \vec{e}_3, \quad 3f(\vec{e}_2) = \alpha\vec{e}_1 + \vec{e}_2 - 2\vec{e}_3, \quad 3f(\vec{e}_3) = \beta\vec{e}_1 + \gamma\vec{e}_2 + 2\vec{e}_3.$$

Hallar $\alpha$, $\beta$ y $\gamma$ de manera que $f$ sea una aplicación ortogonal y encontrar la matriz de $f$ en la base $B$.

**4.** Sea $T \in L(E)$ una involución, es decir, $T^2 = I$. Demostrar que $T$ es una simetría ortogonal si y sólo si $T = T'$.

**5.** Sea $V$ un espacio vectorial real y $A$ una aplicación ortogonal. Probar que si $\lambda$ es una raíz del polinomio característico de $A$, entonces $|\lambda| = 1$. (Demostrar este resultado sin utilizar la complexificación de $V$, es decir, utilizando únicamente métodos reales: si $\lambda = \alpha + i\beta$, existen $\vec{u}, \vec{v} \in V$ tal que $A\vec{u} = \alpha\vec{u} - \beta\vec{v}$ y $A\vec{v} = \beta\vec{u} + \alpha\vec{v}$.)

**6.** Sea $E$ un espacio euclídeo y $T: E \mapsto E$ una aplicación que conserva el producto escalar, es decir,

$$(T(\vec{x}), T(\vec{y})) = (\vec{x}, \vec{y})$$

para todo $\vec{x}, \vec{y} \in E$. Demostrar que $T$ es lineal. [*Sugerencia:* calcular $\|T(\vec{x}, \vec{y}) - T(\vec{x}) - T(\vec{y})\|^2$ y $\|T(a\vec{x}) - aT(\vec{x})\|^2$.)

**7.** Sea $E$ un espacio euclídeo y $T: E \mapsto E$ una aplicación. Demostrar que las condiciones siguientes son equivalentes:

a) $T$ es una aplicación ortogonal.

b) $(T(\vec{u}), T(\vec{v})) = (\vec{u}, \vec{v})$ para todo $\vec{u}, \vec{v} \in E$.

c) $\|T(\vec{u}) - T(\vec{v})\| = \|\vec{u} - \vec{v}\|$ para todo $\vec{u}, \vec{v} \in E$ y $T(\vec{0}) = \vec{0}$.

(*Notas:* condiciones *b*) y *c*) no suponen que $T$ sea lineal. Hacer la demostración siguiendo el plan *c*) $\Rightarrow$ *b*) $\Rightarrow$ *a*) $\Rightarrow$ *c*). La mayor dificultad está en probar que $T$ es lineal en el paso *b*) $\Rightarrow$ *a*); para ello hemos propuesto el problema 6.)



**8.** Sea $T = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix}$ con $\alpha^2 + \beta^2 = 1$ la matriz de $T$ en la base canónica, de $\mathbb{R}^2$.

a) Demostrar que $T$ es la simetría (ortogonal) con respecto a la recta $ax + by = 0$, donde

$$\alpha = \frac{b^2 - a^2}{a^2 + b^2}, \quad \beta = \frac{-2ab}{a^2 + b^2}$$

b) Encontrar la recta de simetría para $T = \begin{pmatrix} 3/5 & -4/5 \\ -4/5 & -3/5 \end{pmatrix}$.

c) Encontrar $T$ para la recta $y = 0$.

**9.** Sea $P = \begin{pmatrix} \alpha & \beta \\ \beta & 1-\alpha \end{pmatrix}$ con $\alpha^2 + \beta^2 = \alpha$ la matriz de $P$ en la base canónica.

a) Demostrar que $P$ es la proyección ortogonal sobre la recta $ax + by = 0$, donde $\alpha = b^2/(a^2 + b^2)$, $\beta = -ab/(a^2 + b^2)$.

b) Si $P$ es la proyección sobre la recta $ax + by = 0$ y $T$ es la simetría (ortogonal) respecto de la perpendicular a $ax + by = 0$, demostrar que $T = I - 2P$.

c) Encontrar la recta de proyección para $P = \begin{pmatrix} 1/10 & 3/10 \\ 3/10 & 9/10 \end{pmatrix}$.

d) Encontrar $P$ para la recta $y = 0$.

**10.** Encontrar las ecuaciones de la simetría con respecto al plano $2x + y + z = 0$.

**11.** Sean $(E, (\ ,\ ))$ y $(F, \langle\ ,\ \rangle)$ dos espacios euclídeos; una aplicación lineal $T: E \mapsto F$ se dice que es una *isometría* si $T$ satisface:

$$\langle T(\vec{u}), T(\vec{v}) \rangle = (\vec{u}, \vec{v})$$

para todo $\vec{u}, \vec{v} \in E$. decir, razonadamente, cuáles de las siguientes aplicaciones son isometrías:

a) $T: \mathbb{R}^2 \mapsto \mathbb{R}^3$, $T(x, y) = (x, 0, y)$

b) $T: P_{\mathbb{R}}^{(2)}[x] \mapsto P_{\mathbb{R}}^{(3)}[x]$, $T(p(x)) = xp(x)$

c) $T: \mathbb{R}^3 \mapsto \mathcal{M}_{3\times 3}(\mathbb{R})$, $T(x, y, z) = \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix}$

*Nota.* Los productos escalares de los espacios anteriores son los ya utilizados varias veces en los problemas anteriores.

## 8.9. ESTRUCTURA DE LAS APLICACIONES LINEALES NO SINGULARES

En esta sección demostraremos que toda aplicación lineal no singular, es decir, biyectiva, de un espacio euclídeo $E$, de dimensión finita, en sí mismo es la composición de una aplicación autoadjunta y una ortonormal de $E$. Por tanto, el estudio de las aplicaciones lineales no singulares en un espacio euclídeo queda reducido al estudio realizado en las secciones anteriores.

TEOREMA 1

Toda aplicación lineal no singular $A$ en un espacio euclídeo $E$ de dimensión finita es de la forma

$$A = O \circ S$$

donde $S$ es una aplicación autoadjunta y $O$ es ortogonal.

Para demostrar este teorema necesitamos el siguiente lema:

LEMA 2

Sea $A$ una aplicación lineal en un espacio euclídeo $E$ de dimensión finita; la aplicación $B = A' \circ A$ es autoadjunta y los autovalores de $B$ son no negativos.

*Demostración.* La cadena de igualdades

$$B' = (A' \circ A)' = A' \circ (A')' = A' \circ A = B$$

muestra que $B$ es autoadjunta. Sea $\lambda$ un autovalor de $B$ con autovector $\vec{u}$. Debido al teorema 4 de la sección 8.6, $\lambda \in \mathbb{R}$. Además,

$$0 \leqq (A(\vec{u}), A(\vec{u})) = (\vec{u}, A' \circ A(\vec{u})) = (\vec{u}, B(\vec{u})) = (\vec{u}, \lambda\vec{u}) = \lambda \|\vec{u}\|^2$$

De aquí deducimos que $\lambda \geqq 0$, ya que $\|\vec{u}\| > 0$. Esto termina la demostración del lema 2. ■

*Demostración del teorema 1.* Por el teorema 4 de la sección 8.6 existe una base ortonormal $U$ de $E$ en la cual la matriz de la aplicación autoadjunta $B = A' \circ A$ es de la forma

$$M(B) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$



con los $\lambda_i \geqq 0$ (ver el lema anterior). Sea $S$ la aplicación autoadjunta que tiene como matriz

$$M(S) = \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}$$

en la base $U$ (observar que $S$ es autoadjunta puesto que su matriz es simétrica en la base ortonormal $U$).

Puesto que $[M(S)]^2 = M(B)$ tenemos que $S^2 = S \circ S = B$; además,

$$|M(S)|^2 = |M(B)| = |M(A')| \cdot |M(A)| = |M(A)|^2 \neq 0$$

Por tanto, $|M(S)| \neq 0$ y $S$ posee una inversa.

Para concluir la demostración del teorema tomamos

$$O = A \circ S^{-1}$$

con lo cual tenemos la igualdad $A = O \circ S$.

Solamente falta probar que $O$ es ortogonal:

$$O' \circ O = (A \circ S^{-1})' \circ (A \circ S^{-1}) = (S^{-1})' \circ (A' \circ A) \circ S^{-1} = (S^{-1})' \circ B \circ S^{-1} =$$
$$= (S^{-1}) \circ S^2 \circ S^{-1} = I$$ ■



Este teorema, junto con el teorema 2 de la sección 8.8 (teorema de clasificación de las aplicaciones ortogonales) y el teorema 4 de la sección 8.6 (diagonalización de aplicaciones autoadjuntas) nos permite interpretar geométricamente las aplicaciones lineales no singulares en un espacio euclídeo de dimensión finita: toda aplicación de este tipo se reduce a varias rotaciones alrededor de ciertos ejes, a varias simetrías alrededor de hiperplanos y a varias dilataciones (positivas) a lo largo de ejes ortogonales dos a dos.

El lector no se sentirá extrañado si decimos que el teorema 1 tiene una interpretación para matrices reales cuadradas no singulares, es decir, de determinante no nulo. Tenemos el siguiente resultado:

TEOREMA 3

Sea $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ con $|A| \neq 0$; existen dos matrices $O \in O(n; \mathbb{R})$ y $S$ simétrica, tal que $A = O \cdot S$.

*Demostración.* Considerar $A$ como una aplicación lineal en el espacio euclídeo $\mathbb{R}^n$ con el producto escalar usual. Por el teorema 1,

$$A = O \circ S$$

con $O$ ortogonal y $S$ autoadjunta. Si denotamos por $A$, $O$ y $S$ las matrices de las aplicaciones $A$, $O$ y $S$, respectivamente, en la base canónica de $\mathbb{R}^n$, se tiene el resultado deseado. ■

Se invita al lector a demostrar el teorema 3 directamente, siguiendo los pasos de la demostración del teorema 1; de esta forma obtendrá un método para calcular las matrices $O$ y $S$ del teorema 3. Algunos ejercicios relativos a este teorema se encuentran al final de esta sección.

## EJERCICIOS 8.9

**1.** Dada $A \in \mathscr{M}_{n\times n}(\mathbb{R})$ con $|A| \neq 0$, demostrar directamente que existen dos matrices, $O \in O(n; \mathbb{R})$ y $S$ simétrica, tal que $A = O \cdot S$.

**2.** Para las siguientes matrices, encontrar la descomposición $A = O \cdot S$ del problema 1.

*a)* $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ *b)* $A = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$

[*Sol.: a)* $O = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$, $S = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix}$.]

**3.** Dada una matriz $B \in \mathscr{M}_{n\times n}(\mathbb{R})$, decimos que $C \in \mathscr{M}_{n\times n}(\mathbb{R})$ es una *raíz cuadrada de B* si $C^2 = B$.

*a)* Demostrar que si $B$ es autoadjunta y todos sus autovalores son no negativos, siempre existe una raíz cuadrada de $B$.

*b)* Encontrar una raíz cuadrada de

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

## BIOGRAFIA

**Euclides** vivió alrededor del año 300 antes de Cristo, pero no se conocen ni la fecha ni el lugar exactos de su nacimiento. De su vida sólo se conoce que enseñó y fundó una escuela en Alejandría en la época de Ptolomeo I que reinó desde el año 323 hasta el año 285 antes de Cristo, aproximadamente.

Euclides es el matemático más prominente de la antigüedad, conocido por su libro *Los elementos.* Este libro ejerció una influencia enorme en el pensamiento matemático hasta el siglo XIX, en el que nuevas formas de geometría no euclideana fueron introducidas. Parte de *Los elementos* es una recopilación de trabajos de otros matemáticos puestos por primera vez juntos mediante un razonamiento lógico.

# CAPITULO 9

# ESPACIOS HERMITICOS



## 9.1. PRODUCTO HERMITICO

En el capítulo anterior hemos introducido la noción de producto escalar en un espacio vectorial *real*, obteniendo así los espacios euclídeos, en los cuales se pueden definir las nociones de longitud de un vector y de perpendicularidad de dos vectores. En este capítulo dotaremos a los espacios vectoriales *complejos* de una estructura adecuada para poder definir las nociones de longitud de un vector y de perpendicularidad de dos vectores en el espacio vectorial complejo. El nuevo concepto que introduciremos en un espacio vectorial complejo recibirá el nombre de *producto hermítico* y los nuevos espacios así obtenidos se llamarán *espacios hermíticos.*

Advertimos al lector que muchos de los resultados en espacios hermíticos son similares a los resultados en espacios euclídeos y que las demostraciones de aquellos son similares a las de éstos. Debido a esto, muchas de las demostraciones serán omitidas en este capítulo.

DEFINICIÓN 1 *(Producto hermítico)*

Sea $H$ un espacio vectorial complejo. Una aplicación $(\ ,\ )$: $H \times H \mapsto \mathbb{C}$ se dice que es un *producto hermítico* si satisface las siguientes propiedades:

*a)* $(\vec{z}, \vec{w}) = \overline{(\vec{w}, \vec{z})}$ para todo $\vec{z}, \vec{w} \in H$.

*b)* $(\vec{z}, \vec{w}+\vec{v}) = (\vec{z}, \vec{w}) + (\vec{z}, \vec{v})$ para todo $\vec{z}, \vec{w}, \vec{v} \in H$.

*c)* $(\lambda\vec{z}, \vec{w}) = \lambda(\vec{z}, \vec{w})$ para todo $\vec{z}, \vec{w} \in H$ y $\lambda \in \mathbb{C}$.

*d)* $(\vec{z}, \vec{z}) > 0$ para todo $\vec{z} \neq \vec{0}$ y $(\vec{z}, \vec{z}) = 0 \Leftrightarrow \vec{z} = \vec{0}$.

*Nota.* Para todo número complejo $\lambda=a+ib$, $\bar{\lambda}=a-ib$ recibe el nombre de *conjugado de* $\lambda$. El estudio de los números complejos se realizó en el capítulo 4. Para nuestros más inmediatos propósitos en este capítulo recordaremos que $a$ se denomina la *parte real de* $\lambda$ y se designa por $\mathrm{re}(\lambda)$, y $b$ se denomina la *parte imaginaria de* $\lambda$, y se designa por $\mathrm{im}(\lambda)$. Además,

$$\lambda\bar{\lambda}=(a+ib)(a-ib)=a^2+b^2=|\lambda|^2$$

donde $|\lambda|$ se utiliza para representar el módulo de $\lambda$.

De las propiedades *b*) y *a*) del producto hermítico se deduce

$$b') \quad (\vec{z}+\vec{w}, \vec{v})=(\vec{z}, \vec{v})+(\vec{w}, \vec{v}) \quad \text{para todo } \vec{z}, \vec{w}, \vec{v}\in H$$

De las propiedades *c*) y *a*) del producto hermítico se deduce:

$$c') \quad (\vec{z}, \lambda\vec{w})=\overline{(\lambda\vec{w}, \vec{z})}=\overline{\lambda(\vec{w}, \vec{z})}=\bar{\lambda}\overline{(\vec{w}, \vec{z})}=\bar{\lambda}(\vec{z}, \vec{w})$$

para todo $\vec{z}, \vec{w}\in H$ y $\lambda\in\mathbb{C}$. Observar que esta propiedad es diferente de la correspondiente para el producto escalar.

Sean $\vec{z}_1, \vec{z}_2, ..., \vec{z}_n, \vec{w}_1, \vec{w}_2, ..., \vec{w}_n\in H$ y $\alpha_1, \alpha_2, ..., \alpha_n, \beta_1, \beta_2, ..., \beta_n\in\mathbb{C}$; combinando las propiedades *b*) y *c*) junto con *b'*) y *c'*) se deduce:

$$\left(\sum_{j=1}^{n}\alpha_j\vec{z}_j, \sum_{k=1}^{m}\beta_k\vec{w}_k\right)=\sum_{j=1}^{n}\sum_{k=1}^{m}\alpha_j\bar{\beta}_k(\vec{z}_j, \vec{w}_k) \tag{1}$$

* * *

DEFINICIÓN 2 *(Espacio hermítico)*

Un espacio vectorial complejo $H$ que posee un producto hermítico recibe el nombre de *espacio hermítico.*

Un ejemplo de espacio hermítico es $\mathbb{C}^n$ con el producto hermítico dado por

$$(\vec{z}, \vec{w})=z_1\bar{w}_1+z_2\bar{w}_2+\cdots+z_n\bar{w}_n$$

donde $\vec{z}=(z_1, z_2, ..., z_n)$ y $\vec{w}=(w_1, w_2, .., w_n)$.

Otro ejemplo de espacio hermítico es el conjunto $C_{\mathbb{C}}([a, b])$ de las funciones continuas con valores complejos definidas en el intervalo $[a, b]\subset\mathbb{R}$, con el producto hermítico dado por

$$(f, g)=\int_a^b f(x)\overline{g(x)}\,dx$$

para toda $f, g\in C_{\mathbb{C}}([a, b])$. En particular, los conjuntos $P_{\mathbb{C}}([a, b])$ de todos los polinomios con coeficientes complejos definidos en el intervalo $[a, b]$ y $P_{\mathbb{C}}^{(n)}([a, b])$ de todos



los polinomios del conjunto anterior de grado no superior a $n$ son espacios hermíticos con el mismo producto hermítico que el definido en $C_{\mathbb{C}}([a, b])$.

La comprobación de las afirmaciones anteriores se deja como ejercicio.

Supongamos que el espacio hermítico $H$ es de dimensión finita $n$, es decir, el espacio vectorial complejo $H$ tiene dimensión $n$, y sea $B=\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$ una base de $H$. Si $\vec{z}, \vec{w}\in H$ podemos escribir

$$\vec{z}=\sum_{i=1}^{n}z_i\vec{u}_i, \quad \vec{w}=\sum_{j=1}^{n}w_j\vec{u}_j$$

con $z_i, w_i\in\mathbb{C}$ para todo $i=1, ..., n$.

Utilizando (1) se tiene el siguiente resultado:

$$(\vec{z}, \vec{w})=\left(\sum_{i=1}^{n}z_i\vec{u}_i, \sum_{j=1}^{n}w_j\vec{u}_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}z_i\bar{w}_j(\vec{u}_i, \vec{u}_j) \tag{2}$$

La matriz

$$P\equiv P_B=\begin{pmatrix}(\vec{u}_1, \vec{u}_1) & (\vec{u}_2, \vec{u}_1) & \cdots & (\vec{u}_n, \vec{u}_1)\\ (\vec{u}_1, \vec{u}_2) & (\vec{u}_2, \vec{u}_2) & \cdots & (\vec{u}_n, \vec{u}_2)\\ \vdots & \vdots & & \vdots\\ (\vec{u}_1, \vec{u}_n) & (\vec{u}_2, \vec{u}_n) & \cdots & (\vec{u}_n, \vec{u}_n)\end{pmatrix}$$

recibe el nombre de *matriz del producto hermítico con respecto a la base B de H*. Con esta notación la igualdad (2) puede escribirse de la forma

$$(\vec{z}, \vec{w})=(z_1, ..., z_n)P_B\begin{pmatrix}\bar{w}_1\\ \vdots\\ \bar{w}_n\end{pmatrix}$$

Debido a la propiedad *a*) del producto hermítico $(\vec{u}_i, \vec{u}_j)=\overline{(\vec{u}_j, \vec{u}_i)}$ para todo $i\neq j$, la matriz $P_B$ no es, en general, simétrica, como ocurría en el producto escalar, sino que es necesario hallar el conjugado de cada uno de los elementos de la matriz.

Dada una matriz $A=(a_{ij})_{j=1,...,m}^{i=1,...,n}$ con $a_{ij}\in\mathbb{C}$, se denomina *conjugada de A* a la matriz $\bar{A}\equiv(\overline{a_{ij}})_{j=1,...,m}^{i=1,...,n}$, que se obtiene hallando el conjugado de cada uno de los elementos de $A$.

Una matriz cuadrada $A=(a_{ij})_{j=1,...,n}^{i=1,...,n}$ se dice *simétrica conjugada* si

$$\bar{A}^t=A,$$

donde $\bar{A}^t$ representa la matriz traspuesta de $\bar{A}$.

Si $A$ es simétrica-conjugada, $a_{ii}=\overline{a_{ii}}$ para todo $i=1, ..., n$; puesto que un número complejo es igual a su conjugado si y sólo si este número es real, deducimos que los elementos de la diagonal principal de una matriz simétrica-conjugada son números reales. Además, $a_{ij}=\overline{a_{ji}}$, $i\neq j$, y, por tanto, los elementos simétricos con respecto a la diagonal principal son conjugados entre sí. Así pues,

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \overline{a_{12}} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \overline{a_{1n}} & \overline{a_{2n}} & \cdots & a_{nn} \end{pmatrix}, \quad a_{ii}\in\mathbb{R}$$

EJEMPLO A. La matriz $A=\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ es simétrica-conjugada.

*Nota.* El lector habrá quedado convencido de que la matriz $P_B$ del producto hermítico con respecto a una base $B$ es siempre una matriz simétrica-conjugada.

* * *

En un espacio hermítico $H$ la *longitud* o *norma* de un vector $\vec{z}\in H$ se define mediante

$$\|\vec{z}\|=\sqrt{(\vec{z},\vec{z})}$$

donde la raíz cuadrada se toma siempre positiva. Observar que la definición tiene sentido ya que $(\vec{z},\vec{z})\geq 0$ debido a la propiedad *d)* del producto hermítico. Para todo $\lambda\in\mathbb{C}$ se tiene la igualdad:

$$\|\lambda\vec{z}\|=\sqrt{(\lambda\vec{z},\lambda\vec{z})}=\sqrt{\lambda\bar{\lambda}(\vec{z},\vec{z})}=\sqrt{|\lambda|^2(\vec{z},\vec{z})}=|\lambda|\,\|\vec{z}\|$$

$\vec{z}\in H$. Un vector de norma 1 se dice *unitario*. Todo vector no nulo $\vec{z}\in H$ puede ser *normalizado*, es decir, multiplicado por un número complejo $\lambda$ de manera que $\lambda\vec{z}$ sea unitario: basta tomar

$$\lambda=\frac{1}{\|\vec{z}\|}.$$

En un espacio hermítico se cumple también la desigualdad de Schwarz:

$$|(\vec{w},\vec{z})|\leq\|\vec{w}\|\,\|\vec{z}\|,\qquad \vec{w},\vec{z}\in H$$

La demostración es similar a la del caso euclídeo, excepto que debe tenerse cuidado en las operaciones con números complejos (para el caso euclídeo ver la sección 8.2 del capítulo 8). La demostración procede de la siguiente manera: si $\lambda\in\mathbb{C}$,

$$0\leq(\vec{z}+\lambda\vec{w},\vec{z}+\lambda\vec{w})=\|\vec{z}\|^2+\bar{\lambda}(\vec{z},\vec{w})+\lambda(\vec{w},\vec{z})+|\lambda|^2\|\vec{w}\|^2$$



los números $\lambda(\vec{w},\vec{z})$ y $\bar{\lambda}(\vec{z},\vec{w})$ son conjugados, ya que $\bar{\lambda}(\vec{z},\vec{w})=\overline{\lambda(\vec{w},\vec{z})}=\overline{\lambda(\vec{w},\vec{z})}$ y, por tanto, su suma coincide con el doble de la parte real de $\lambda(\vec{w},\vec{z})$; tenemos, pues,

$$0\leq\|\vec{z}\|^2+2\,\mathrm{real}\,(\lambda(\vec{w},\vec{z}))+\|w\|^2|\lambda|^2$$

Puesto que la parte real de todo número complejo nunca supera a su módulo tenemos la desigualdad

$$0\leq\|\vec{w}\|^2|\lambda|^2+2|\lambda||(\vec{w},\vec{z})|+\|\vec{z}\|^2$$

Esta ecuación cuadrática en $|\lambda|\in\mathbb{R}$ no puede tener raíces reales distintas (el razonamiento es el mismo que en el caso euclídeo) y, por tanto, su discriminante ha de ser negativo o nulo:

$$|(\vec{w},\vec{z})|^2-\|z\|^2\|\vec{w}\|^2\leq 0$$

de donde deducimos el resultado deseado.

A partir de la desigualdad de Schwarz pueden demostrarse las *desigualdades triangulares:*

$$\|\vec{z}+\vec{w}\|\leq\|\vec{z}\|+\|\vec{w}\|,\quad \big|\|\vec{z}\|-\|\vec{w}\|\big|\leq\|\vec{z}-\vec{w}\|$$

A pesar de que no puede definirse el concepto de ángulo en un espacio hermítico (¿por qué?), podemos dar la noción de *ortogonalidad:* $\vec{z}\in H$ es *ortogonal* a $\vec{w}\in H$ si

$$(\vec{z},\vec{w})=0$$

Podemos definir, por tanto, el *complemento ortogonal* $W^{\perp}$, de un subespacio vectorial $W$ de $H$:

$$W^{\perp}=\{\vec{z}\in H\;/\;(\vec{z},\vec{w})=0 \text{ para todo } \vec{w}\in W\}$$

En un espacio hermítico el teorema de Pitágoras es también cierto: si $\vec{z}$ es ortogonal a $\vec{w}$,

$$\|\vec{w}+\vec{z}\|^2=\|\vec{w}\|^2+\|\vec{z}\|^2$$

En los espacios unitarios puede realizarse el *proceso de ortogonalización* de un conjunto de vectores linealmente independientes; el proceso es exactamente igual que en los espacios euclídeos, por lo cual se omite aquí. En particular, todo espacio hermítico de dimensión finita posee una base ortonormal, es decir, una base formada por vectores unitarios mutuamente ortogonales.

EJEMPLO B. Queremos aplicar el proceso de ortogonalización a los vectores $\vec{u}_1=(i,0,0)$, $\vec{u}_2=(0,i,1)$ y $\vec{u}_3=(1,0,1)$ del espacio hermítico $\mathbb{C}^3$ con el producto hermítico usual. Tomamos $\vec{e}_1=\vec{u}_1$ y $\vec{e}_2=\vec{u}_2+\lambda\vec{e}_1$ con $\lambda\in\mathbb{C}$; calculamos $\lambda$ imponiendo la condición $(\vec{e}_2, \vec{e}_1)=0$:

$$0=(\vec{e}_2, \vec{e}_1)=(\vec{u}_2, \vec{e}_1)+\lambda(\vec{e}_1, \vec{e}_1)=0+\lambda\cdot 1.$$

Por tanto, $\vec{e}_2=\vec{u}_2$. Finalmente, tomamos

$$\vec{e}_3=\vec{u}_3+\alpha\vec{e}_1+\beta\vec{e}_2 \quad , \quad \alpha, \beta\in\mathbb{C}$$

y calculamos $\alpha$ y $\beta$ con las condiciones $(\vec{e}_1, \vec{e}_3)=0$ y $(\vec{e}_2, \vec{e}_3)=0$.

$$0=(\vec{e}_1, \vec{e}_3)=(\vec{e}_1, \vec{u}_3)+\bar{\alpha}(\vec{e}_1, \vec{e}_1)=i+\bar{\alpha}\cdot 1$$
$$0=(\vec{e}_2, \vec{e}_3)=(\vec{e}_2, \vec{u}_3)+\bar{\beta}(\vec{e}_2, \vec{e}_2)=1+\bar{\beta}[1+1]$$

Por tanto, $\bar{\alpha}=-i$ o, equivalentemente, $\alpha=i$, y $\bar{\beta}=-1/2$ o, equivalentemente, $\beta=-1/2$. El tercer vector es

$$\vec{e}_3=\vec{u}_3+i\vec{e}_1-\frac{1}{2}\vec{e}_2=(1, 0, 1)+(-1, 0, 0)-\frac{1}{2}(0, i, 1)=\left(0, -\frac{i}{2}, \frac{1}{2}\right)$$

## 9.2. APLICACIONES ENTRE ESPACIOS HERMÍTICOS

En esta sección estudiaremos algunos tipos particulares de aplicaciones de un espacio hermítico en sí mismo, así como la diagonalización (compleja) de algunos de estos tipos.

Dado un espacio hermítico $H$ y una aplicación $A\in L(H)$, se denomina *adjunta de A* y la llamaremos $A^*$ a toda aplicación lineal que satisface

$$(A(\vec{z}), \vec{w})=(\vec{z}, A^*(\vec{w})) \quad , \quad \vec{z}, \vec{w}\in H.$$

Si el espacio hermítico $H$ es de dimensión finita, lo cual siempre se supondrá cierto de aquí en adelante, la existencia y unicidad de la aplicación adjunta $A^*$ de $A$ se demuestran como en el caso de espacios euclídeos. Es conveniente demostrar la existencia para encontrar la relación que existe entre las matrices de $A$ y $A^*$ en una base ortonormal de $H$.

Sea $A=(a_{ij})$ la matriz de $A\in L(H)$ en una base ortonormal $B=\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$; por tanto,

$$A(\vec{e}_j)=\sum_{i=1}^{n} a_{ij}\vec{e}_i.$$

Si $C=(c_{ij})$ es la matriz de $A^*\in L(H)$ se tiene que

$$A^*(\vec{e}_i)=\sum_{j=1}^{n} c_{ji}\vec{e}_j.$$



Por tanto,

$$a_{ij}=\left(\sum_{i=1}^{n} a_{ij}\vec{e}_i, \vec{e}_i\right)=(A(\vec{e}_j), \vec{e}_i)=(\vec{e}_j, A^*(\vec{e}_i))=$$
$$=\left(\vec{e}_j, \sum_{j=1}^{n} c_{ji}\vec{e}_j\right)=\overline{c_{ji}}$$

Hemos llegado a la conclusión de que la matriz de $A^*$ en la base ortonormal $B$ es la *conjugada-traspuesta* de $A$, es decir, $\bar{A}^t$. Esta matriz se suele denominar también *adjunta* de $A$ y se denota con el símbolo $A^*$.

Las propiedades de la aplicación adjunta $A^*$ son las mismas que las dadas en la sección 8.5 para la aplicación adjunta en espacios euclídeos.

* * *

Sea $A\in L(H)$; la aplicación $A$ se dice *autoadjunta* si $A^*=A$. Por tanto, si $A$ es autoadjunta en $H$, la matriz de $A$ en una base ortonormal ha de satisfacer

$$A=\bar{A}^t$$

es decir, $A$ debe ser una matriz *simétrica-conjugada;* a estas matrices se les da también el nombre de *hermíticas.*

* * *

Las aplicaciones de un espacio hermítico $H$ en sí mismo que conservan el producto hermítico reciben el nombre de *unitarias.* Por tanto, $A\in L(H)$ es *unitaria* si y sólo si

$$(A(\vec{z}), A(\vec{w}))=(\vec{z}, \vec{w}) \quad , \quad \vec{z}, \vec{w}\in H.$$

Si $A$ es unitaria se ha de tener

$$(\vec{z}, A^*\circ A(\vec{w}))=(\vec{z}, \vec{w})$$

para todo $\vec{z}, \vec{w}\in H$, y por tanto, $A^*\circ A=I$.

Si $A$ y $A^*$ representan las matrices de $A$ y $A^*$ en una base ortonormal de $H$, la igualdad anterior se escribe de la forma

$$A^*\cdot A=I$$

o, equivalentemente,

$$A^{-1}=A^*.$$

Las matrices que satisfacen una cualquiera de las dos desigualdades anteriores reciben el nombre de *matrices unitarias.* El conjunto de todas las matrices unitarias de orden $n$ se denota por $U(n)$ y, por tanto,

$$U(n)=\{A\in\mathcal{M}_{n\times n}(\mathbb{C}) \;/\; \bar{A}^t\cdot A=I\}.$$

Puesto que la composición de aplicaciones unitarias es otra aplicación unitaria y la inversa de una aplicación unitaria siempre existe y es una aplicación unitaria, el conjunto $U(n)$ es un grupo con respecto a la multiplicación de matrices. Se invita al lector a comprobar con detalle este resultado.

*Los autovalores de una aplicación unitaria deben ser números complejos de modulo 1 y los autovalores de una aplicación autoadjunta deben ser reales.* En efecto, si $A \in L(H)$ es una aplicación unitaria y $\lambda$ es un autovalor de $A$ con autovector $\vec{z} \in H$, se tiene

$$(\vec{z}, \vec{z}) = (A(\vec{z}), A(\vec{z})) = (\lambda\vec{z}, \lambda\vec{z}) = \lambda\bar{\lambda}(\vec{z}, \vec{z}) = |\lambda|^2 \|\vec{z}\|^2$$

de donde resulta $|\lambda| = 1$.

Si $A \in L(H)$ es una aplicación autoadjunta y $\lambda$ un autovalor de $A$ con autovector $\vec{z} \in H$, se tiene

$$\lambda\|\vec{z}\|^2 = \lambda(\vec{z}, \vec{z}) = (\lambda\vec{z}, \vec{z}) = (A(\vec{z}), \vec{z}) = (\vec{z}, A(\vec{z})) = (\vec{z}, \lambda\vec{z}) = \bar{\lambda}\|\vec{z}\|^2$$

de donde resulta que $\lambda = \bar{\lambda}$ y, por tanto, $\lambda \in \mathbb{R}$.

* * *

A continuación nos dedicaremos a demostrar que tanto las aplicaciones autoadjuntas como las aplicaciones unitarias en espacios hermíticos pueden diagonalizarse. (Comparar este resultado con los resultados análogos para aplicaciones autoadjuntas y aplicaciones ortogonales en espacios euclídeos obtenidos en el capítulo anterior.)

Nuestra forma de proceder es demostrar que una clase de aplicaciones, que definiremos a continuación, y que engloba a las anteriores, es diagonalizable.

Una aplicación lineal $A \in L(H)$ se dice *normal* si conmuta con su adjunta, es decir,

$$A^* \circ A = A \circ A^*$$

EJEMPLO A. 1. Si $A$ es autoadjunta, $A$ es normal, ya que $A^* \circ A = A \circ A^* = A \circ A$.

2. Si $A$ es unitaria, $A$ es normal, ya que $A^* \circ A = A^{-1} \circ A = I$ y $A \circ A^* = A \circ A^{-1} = I$.

3. La aplicación cuya matriz es

$$N = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}, \quad \lambda_j \in C$$

en una base ortonormal, es también normal ya que

$$N^* \cdot N = \begin{pmatrix} |\lambda_1|^2 & & & 0 \\ & |\lambda_2|^2 & & \\ & & \ddots & \\ 0 & & & |\lambda_n|^2 \end{pmatrix} = N \cdot N^*$$



Para reducir una aplicación normal a su forma más sencilla es necesario tener conocimiento de sus posibles autovalores y de sus subespacios invariantes.

PROPOSICIÓN 1

Sea $A \in L(H)$ una aplicación normal. Todo autovector $\vec{z} \in H$ de $A$, con autovalor $\lambda \in \mathbb{C}$ es un autovector de $A^*$, con autovalor $\bar{\lambda}$.

*Demostración.* El subespacio $P(\lambda)$ de todos los autovectores de $A$ que tienen el mismo autovalor $\lambda$ es invariante por $A^*$, ya que si $\vec{z} \in P(\lambda)$,

$$A \circ A^*(\vec{z}) = A^* \circ A(\vec{z}) = A^*(\lambda\vec{z}) = \lambda A^*(z)$$

y, por tanto, $A^*(z) \in P(\lambda)$. Además, si $\vec{z}, \vec{w} \in P(\lambda)$,

$$(A^*(\vec{z}), \vec{w}) = (\vec{z}, A(\vec{w})) = (\vec{z}, \lambda\vec{w}) = \bar{\lambda}(\vec{z}, \vec{w})$$

Tomando en particular $w = A^*(z) - \bar{\lambda}\vec{z} \in P(\lambda)$ se tiene que

$$(A^*(\vec{z}) - \bar{\lambda}\vec{z}, A^*(z) - \bar{\lambda}\vec{z}) = 0$$

de donde se deduce $A^*(\vec{z}) = \bar{\lambda}\vec{z}$. ■

PROPOSICIÓN 2

Sea $A \in L(H)$ una aplicación normal y sea $P(\lambda)$ el subespacio vectorial de todos los autovectores de $A$ con el mismo autovalor $\lambda \in \mathbb{C}$. Entonces, $Q(\lambda) = P(\lambda)^\perp$ es invariante por $A$.

*Demostración.* Si $\vec{w} \in Q(\lambda)$ y $\vec{z} \in P(\lambda)$ se tiene que $A^*(\vec{z}) = \bar{\lambda}\vec{z}$ (proposición 1) y, además,

$$(A(\vec{w}), \vec{z}) = (\vec{w}, A^*(\vec{z})) = (\vec{w}, \bar{\lambda}\vec{z}) = \lambda(\vec{w}, \vec{z}) = 0$$

de donde se deduce que $A(\vec{w}) \in Q(\lambda)$. ■

El siguiente teorema nos da la diagonalización de toda aplicación normal en un espacio hermítico.

TEOREMA 3 (*Teorema espectral*)

La matriz de una aplicación normal $A$ en un espacio hermítico puede ser reducida en una base ortonormal determinada a una matriz diagonal.

COROLARIO

Las matrices simétricas conjugadas y las unitarias son diagonalizables.

*Demostración.* Realizamos la demostración por inducción en la dimensión $n$ del subespacio hermítico $H$. Si $n=1$ el teorema es cierto. Supongamos que $n>1$ y sea $P(\lambda_1)$ el subespacio de autovectores de $A$ con autovalor $\lambda_1$; si $P(\lambda_1)=H$, cualquier base ortonormal de $H$ produce una matriz diagonal con $\lambda_1$ en su diagonal principal. Si $P(\lambda_1) \subsetneq H$, $Q(\lambda_1)=P(\lambda_1)^{\perp}$ es de dimensión mayor o igual que 1 y menor que $n$. Por la hipótesis de inducción,

$$A|_{P(\lambda_1)}=\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_1\end{pmatrix}, \quad A|_{Q(\lambda_1)}=\begin{pmatrix}\lambda_2 & & 0\\ & \ddots & \\ 0 & & \lambda_k\end{pmatrix}.$$

Puesto que $H=P(\lambda_1)\oplus Q(\lambda_1)$, $A$ puede reducirse a una matriz diagonal (yuxtaposición de las dos anteriores) en una base ortonormal. ■

En el caso particular de que $A$ sea *unitaria*, sus autovalores han de ser de módulo 1, es decir, $\lambda=e^{i\varphi}$ y, por tanto, puede reducirse a

$$\begin{pmatrix}e^{i\varphi_1} & & 0\\ & \ddots & \\ 0 & & e^{i\varphi_n}\end{pmatrix}$$

Si $A$ es autoadjunta, sus autovalores son reales, lo cual se ha demostrado anteriormente, y, por tanto, puede reducirse a

$$\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}, \quad \lambda_j\in\mathbb{R}$$

* * *

Sea $A\in U(n)$; por el teorema espectral existe una base ortonormal $B$ de $\mathbb{C}^n$ de manera que

$$A=C\begin{pmatrix}e^{i\varphi_1} & & 0\\ & \ddots & \\ 0 & & e^{i\varphi_n}\end{pmatrix}C^{-1}$$

donde $C$ es la matriz del cambio de base de la base canónica a $B$. Por tanto, $C$ transforma una base ortonormal en otra base ortonormal y esto es suficiente para asegurar que $C\in U(n)$. Por tanto,

$$A=C\begin{pmatrix}e^{i\varphi_1} & & 0\\ & \ddots & \\ 0 & & e^{i\varphi_n}\end{pmatrix}\bar{C}^t$$

Un razonamiento similar produce

$$S=C\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\bar{C}^t, \quad \lambda_i\in\mathbb{C}$$

con $C\in U(n)$, si $S$ es simétrica-conjugada o hermítica.



Para futuras referencias enunciamos este resultado a continuación; observar que este resultado es el análogo para espacios hermíticos de la proposición 3 de la sección 8.8.

PROPOSICIÓN 4

1) Dada $A\in U(n)$, existe $C\in U(n)$ tal que

$$A=C\begin{pmatrix}e^{\varphi_1} & & 0\\ & \ddots & \\ 0 & & e^{\varphi_n}\end{pmatrix}\bar{C}^t$$

2) Dada $S\in\mathcal{M}_{n\times n}(\mathbb{C})$ y tal que $S=\bar{S}^t$, existe $C\in U(n)$ tal que

$$S=C\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix}\bar{C}^t, \quad \lambda_j\in\mathbb{R}$$

* * *

En espacios hermíticos existe un resultado análogo al enunciado en el teorema 1 de la sección 8.9, capítulo 8, en espacios euclídeos. Se tiene entonces que *toda aplicación lineal no singular $A$ en un espacio hermítico puede escribirse de la forma*

$$A=U\circ S$$

*donde $S$ es autoadjunta y $U$ es unitaria.* La demostración de este resultado es análoga a la dada para el teorema 1 de la sección 8.9, capítulo 8.

En el caso de espacios hermíticos este resultado tiene una interpretación elegante; si $A$ estuviera definida en el espacio hermítico $\mathbb{C}$, la matriz de $A$ sería un número complejo $a+bi$ no nulo, la matriz de $S$ sería un número real $r$ y la matriz de $U$ sería un número complejo de módulo 1, $e^{i\varphi}$; por tanto,

$$a+bi=e^{i\varphi}r$$

que es la forma polar de un número complejo. Entonces, la fórmula

$$A=U\circ S$$

puede interpretarse como la «forma polar» de una aplicación no singular y de hecho recibe el nombre de *descomposición polar de $A$*.

Evidentemente, también existe una descomposición polar de toda matriz $A\in\mathcal{M}_{n\times n}(\mathbb{C})$ no singular, lo cual puede deducirse de la descomposición polar para aplicaciones. Puesto que no lo hicimos en el capítulo anterior, daremos aquí una demostración directa de este hecho.

PROPOSICIÓN 5

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{C})$ con $|A| \neq 0$; existen dos matrices $U \in U(n)$ y $S$ simétrica-conjugada, tal que $A = U \cdot S$.

*Demostración.* La aplicación $B = A^*A$ es simétrico-conjugada, ya que

$$B^* = \bar{B}^t = (\overline{\bar{A}^t A})^t = (A^t \bar{A})^t = \bar{A}^t A = B$$

Por la proposición 4 tenemos que existe $C \in U(n)$ tal que

$$B = C\begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_n \end{pmatrix}\bar{C}^t \quad , \quad \lambda_j \in \mathbb{R}$$

De la definición de $B$ se deduce que los $\lambda_j > 0$. Tomar

$$S = C\begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t$$

y observar que $S$ es invertible, ya que $|S|^2 = \lambda_1 \times \cdots \times \lambda_n = |B| = |A|^2 \neq 0$. Tomando $U = AS^{-1}$ se tiene que $A = US$ y $S$ es simétrica conjugada, ya que

$$S^* = \bar{S}^t = \overline{C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t}^{\,t} = \bar{\bar{C}}^t\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t = C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t = S$$

Unicamente falta demostrar que $U$ es unitaria:

$$U^*U = \bar{U}^t U = (\overline{A \cdot S^{-1}})^t (A \cdot S^{-1}) = (\overline{S^{-1}})^t \cdot \bar{A}^t \cdot A \cdot S^{-1} =$$

$$\overline{\left[C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t\right]}^{-1t} \cdot B \cdot \left[C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}\bar{C}^t\right]^{-1} =$$

$$= C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}^{-1}\bar{C}^t \cdot B \cdot C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}^{-1}\bar{C}^t =$$

$$= C\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}^{-1}\begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}\begin{pmatrix} \sqrt{\lambda_1} & & \\ & \ddots & \\ & & \sqrt{\lambda_n} \end{pmatrix}^{-1}\bar{C}^t = C \cdot I\bar{C}^t = I$$

en donde se ha utilizado repetidamente el hecho de que $C^{-1} = \bar{C}^t$. ■



Finalmente, queremos resaltar que todo espacio euclídeo puede convertirse en un espacio hermítico. Ya sabemos (sección 7.8) que todo espacio vectorial *real* $E$ puede convertirse en un espacio vectorial *complejo* $E_{\mathbb{C}}$; podemos tomar $H = E_{\mathbb{C}}$ y definir

$$(\vec{u} + i\vec{v}, \vec{z} + i\vec{w})_{\mathbb{C}} = [(\vec{u}, \vec{z}) + (\vec{v}, \vec{w})] + i[-(\vec{v}, \vec{z}) + (\vec{u}, \vec{w})]$$

donde ( , ) denota el producto escalar en $E$, y $\vec{u} + i\vec{v}$, $\vec{z} + i\vec{w} \in H$. El lector puede comprobar que $( \, , \, )_{\mathbb{C}}$ es un producto hermítico en $H$ (ver el ejercicio 6 al final de esta sección).

Las aplicaciones autoadjuntas en $E$ son también autoadjuntas en $H$ y las aplicaciones ortogonales en $E$ son unitarias en $H$; consecuentemente, los autovalores de aplicaciones ortogonales son números complejos de módulo 1 (comparar este resultado con el teorema 4 de la sección 8.7, capítulo 8).

## EJERCICIOS (CAPITULO 9)

**1.** En el espacio hermítico $\mathbb{C}^3$ con el producto hermítico usual encontrar el complemento ortogonal del subespacio vectorial generado por el vector $\vec{z} = (i, 0, 1)$.

**2.** Sea $\mathbb{C}^2$ con su producto hermítico usual. Estudiar si son unitarias las siguientes matrices:

*a)* $\begin{pmatrix} 3+2i & 5-i \\ 2+i & 1+i \end{pmatrix}$ *b)* $\begin{pmatrix} \dfrac{1+i\sqrt{2}}{2} & \dfrac{i}{2} \\ \dfrac{i}{2} & \dfrac{1-i\sqrt{2}}{2} \end{pmatrix}$ *c)* $\begin{pmatrix} 1+i & 2+i \\ -1-i & -1 \end{pmatrix}$

**3.** Encontrar las matrices de Jordan de las dos últimas matrices del ejercicio 2, así como matrices de transición, siendo estas últimas unitarias si la matriz de partida es unitaria.

**4.** Demostrar que $A = \begin{pmatrix} 0 & i & 0 \\ i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ es una matriz normal y encontrar una base ortonormal en $\mathbb{C}^3$ en la que $A$ tenga una forma diagonal.

**5.** Sea $A = \begin{pmatrix} 1+i & 1 \\ 1 & 1-i \end{pmatrix} \in \mathcal{M}_{\times 2}(\mathbb{C})$. Encontrar una descomposición polar de $A$, es decir, una matriz unitaria $U$ y una matriz simétrica conjugada, tal que $A = US$.

**6.** Sea $E$ un espacio euclídeo; en $H = E_{\mathbb{C}} = \{\vec{u} + i\vec{v} / \vec{u}, \vec{v} \in E\}$ definir

$$(\vec{u} + i\vec{v}, \vec{z} + i\vec{w})_{\mathbb{C}} = [(\vec{u}, \vec{z}) + (\vec{v}, \vec{w})] + i[-(\vec{v}, \vec{z}) + (\vec{u}, \vec{w})]$$

donde ( , ) describe el producto escalar en $E$ y $\vec{u} + i\vec{v}$, $\vec{z} + i\vec{w} \in H$. Demostrar que $( \, , \, )_{\mathbb{C}}$ es un producto hermítico en $H$.

BIOGRAFIA

**Charles Hermite** nació el 24 de diciembre de 1822 en Dieuze (Francia) y murió el 14 de enero de 1901 en París. A pesar de que a los 20 años ya había demostrado su capacidad matemática, sus dificultades con los exámenes le llevaron a dedicar cinco de los años más productivos de su vida a preparar los exámenes para la obtención de la licenciatura, que consiguió en 1848.

Enseñó en la Escuela Politécnica de París y después en el Colegio de Francia, en la misma ciudad. No fue hasta 1869 que fue nombrado profesor en la Escuela Normal de París, para pasar un año más tarde a ser profesor de Algebra en la Sorbonne.

En 1873 Hermite publicó la primera demostración de que el número $e$ es trascendente, es decir, no es raíz de ninguna ecuación algebraica con coeficientes racionales. Contribuyó decisivamente al desarrollo de la teoría de las aplicaciones y de las formas cuadráticas, así como a la solución de la ecuación general de quinto grado utilizando las funciones elípticas.



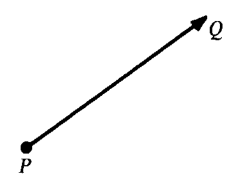
## CAPITULO 10

# MOVIMIENTOS EN UN ESPACIO AFIN EUCLIDEO. MOVIMIENTOS EN $\mathbb{R}^2$ Y $\mathbb{R}^3$



En este capítulo vamos a estudiar algunos tipos de aplicaciones entre espacios afines cuyo espacio vectorial subyacente es un espacio euclídeo (la definición de espacio afín ha sido dada en la sección 5.5 del capítulo 5). Para nuestros propósitos podemos siempre pensar que el espacio afín es $\mathbb{R}^n$ y que está formado por puntos, los cuales serán designados por letras mayúsculas como $P$, $Q$, $R$, ... Dados dos puntos $P$ y $Q$ de $\mathbb{R}^n$ definimos la distancia de $P$ a $Q$ mediante

$$d(P, Q) = \|\overrightarrow{PQ}\| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$



donde $(x_1, x_2, ..., x_n)$ e $(y_1, y_2, ..., y_n)$ son las coordenadas de $P$ y $Q$, respectivamente, con respecto a un sistema de referencia fijado en $\mathbb{R}^n$. En este capítulo vamos a estudiar detenidamente aquellos tipos de aplicaciones en $\mathbb{R}^2$ y $\mathbb{R}^3$ que conservan la distancia (las definiciones precisas se darán en la sección primera de este capítulo).

El lector habrá podido adivinar que debe existir cierta relación entre las aplicaciones en $\mathbb{R}^2$ y $\mathbb{R}^3$ que conservan las distancias y aquellas aplicaciones que conservan el producto escalar usual en los mismos espacios considerados como espacios euclídeos. Tales aplicaciones han sido estudiadas en el capítulo 8, y son las aplicaciones ortogonales. Debido a esto serán utilizados en este capítulo algunos resultados del capítulo 8.

## 10.1. TRANSFORMACIONES AFINES. EJEMPLOS

Recordemos que un espacio afín $A$ está formado por un conjunto $\mathscr{P}$ de puntos y un espacio vectorial $V$ de manera que a cada par de puntos $P, Q \in \mathscr{P}$ le corresponde un vector $\overrightarrow{PQ} \in V$ y tal que para cualesquiera tres puntos $P, Q, R \in \mathscr{P}$ se verifica $\overrightarrow{PR} = \overrightarrow{PQ} + \overrightarrow{QR}$.

Un sistema de referencia $\mathscr{R}$ en $A$ está formado por un punto $O \in \mathscr{P}$ y una base $\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$ de $V$. Dado un punto $X \in \mathscr{P}$ el vector $\overrightarrow{OX}$ puede escribirse de manera única como

$$\overrightarrow{OX} = x_1\vec{e}_1 + x_2\vec{e}_2 + \cdots + x_n\vec{e}_n$$

Los elementos $(x_1, x_2, ..., x_n)$ reciben el nombre de coordenadas del punto $X$ con respecto al sistema de referencia $\mathscr{R} = \{O; \vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}$. Ejemplos de cambios de sistema de referencia se han dado en la sección 5.5 del capítulo 5.

Un espacio afín se dice *euclídeo* si el espacio vectorial $V$ es euclídeo. En este caso definimos

$$d(P, Q) = \|\overrightarrow{PQ}\|$$

para cualesquiera par de puntos $P, Q \in \mathscr{P}$.

DEFINICIÓN 1

Una aplicación $T: A \mapsto A$ se dice *afín* si existe una aplicación lineal $\mathscr{T}: V \mapsto V$ tal que para todo par de puntos $P, Q \in \mathscr{P}$

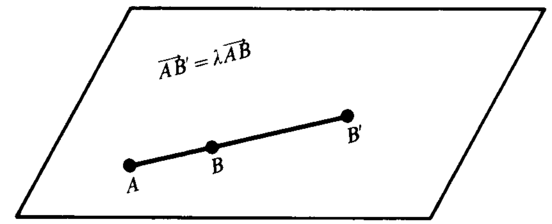
$$\overrightarrow{T(P)T(Q)} = \mathscr{T}(\overrightarrow{PQ})$$

Las aplicaciones afines también reciben el nombre de *transformaciones afines.*



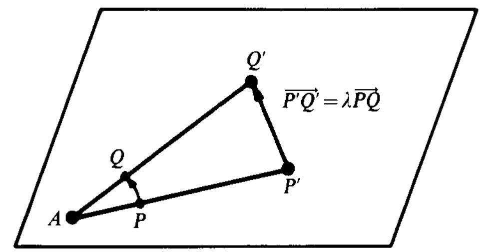
EJEMPLO A. En el espacio afín $\mathbb{R}^2$ definimos la *homotecia* de centro $A$ y razón $\lambda$ como la aplicación $H \equiv H_{A;\lambda}$ tal que si $B \in \mathbb{R}^2$, la imagen $B'$ de $B$ mediante $H$ satisface

$$\overrightarrow{AB'} = \lambda\overrightarrow{AB}$$




La aplicación lineal $\mathscr{H}$ asociada a $H$ es $\mathscr{H}(\vec{v}) = \lambda\vec{v}$, $\vec{v} \in \mathbb{R}^2$. En efecto, si $P, Q \in \mathbb{R}^2$,

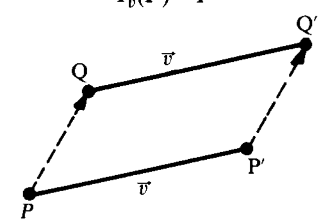
$$\overrightarrow{H(P)H(Q)} = \overrightarrow{P'Q'} = \overrightarrow{P'A} + \overrightarrow{AQ'} = -\overrightarrow{AP'} + \overrightarrow{AQ'} = -\lambda\overrightarrow{AP} + \lambda\overrightarrow{AQ} =$$
$$= \lambda(-\overrightarrow{AP} + \overrightarrow{AQ}) = \lambda\overrightarrow{PQ} = \mathscr{H}(\overrightarrow{PQ})$$




Cuando $\lambda = 1$ se tiene la transformación identidad. Observar que el mismo ejemplo es válido en el espacio afín $\mathbb{R}^n$.

EJEMPLO B. Dado un vector $\vec{v} \in \mathbb{R}^n$, la *traslación* de vector $\vec{v}$, que designamos por $T_{\vec{v}}$, es también una aplicación afín. En este caso:

$$T_{\vec{v}}(P) = P'$$




si $\overrightarrow{PP'} = \vec{v}$. Es fácil comprobar que $T_{\vec{v}}$ es una transformación afín, ya que

$$\overrightarrow{T_{\vec{v}}(P)T_{\vec{v}}(Q)} = \overrightarrow{P'Q'} = \overrightarrow{PQ}$$

y, por tanto, basta tomar la aplicación lineal asociada a $T_{\vec{v}}$ como la identidad en $\mathbb{R}^2$.

* * *

Tratamos ahora de encontrar la expresión analítica de una transformación afín en un sistema de referencia. Dado $\mathscr{R}=\{O; \vec{e}_1, ..., \vec{e}_n\}$ un sistema de referencia en un espacio afín $A=(\mathscr{P}, V)$ y $T$ una aplicación afín en $A$ tenemos que

$$\overrightarrow{OX'}=\overrightarrow{OT(X)}=\overrightarrow{OT(O)}+\overrightarrow{T(O)T(X)}=\overrightarrow{OT(O)}+\mathscr{T}(\overrightarrow{OX})$$

para todo $X\in\mathscr{P}$, donde $X'=T(X)$. Si $(x_1', ..., x_n')$ son las coordenadas de $X'$ con respecto a $\mathscr{R}$, $(x_1, ..., x_n)$ son las coordenadas de $X$ con respecto a $\mathscr{R}$, $(a_1, ..., a_n)$ son las coordenadas de $T(O)$ con respecto al mismo sistema de referencia y $T$ es la matriz de $\mathscr{T}$ con respecto a la base $\{\vec{e}_1, ..., \vec{e}_n\}$, la igualdad anterior se escribe en coordenadas de la forma

$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}=\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}+T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \tag{1}$$

El lector puede convencerse de que la expresión (1) sirve para definir una aplicación afín que tiene a (1) como expresión analítica en el sistema de referencia fijado.

Si $T(O)=O$ se tiene que $(a_1, ..., a_n)=(0, ..., 0)$ y la expresión (1) se reduce a

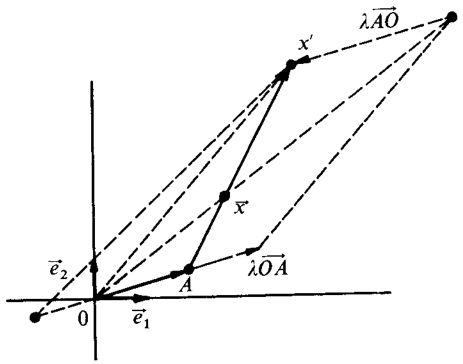
$$\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}=T\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

EJEMPLO C. Tratemos de encontrar la expresión analítica de la homotecia de centro $A$ y razón $\lambda$ del ejemplo A. Supongamos que $R=\{O; \vec{e}_1, \vec{e}_2\}$, donde $O=(0, 0)$, $\vec{e}_1=(1, 0)$ y $\vec{e}_2=(0, 1)$ en un sistema de referencia en $\mathbb{R}^2$ con respecto al cual $A=(a_1, a_2)$. Tenemos que

$$\overrightarrow{OX'}=\overrightarrow{OH(X)}=\overrightarrow{OH(O)}+\mathscr{H}(\overrightarrow{OX})=\overrightarrow{OA}+\overrightarrow{AH(O)}+\mathscr{H}(\overrightarrow{OX})=\overrightarrow{OA}+\lambda\overrightarrow{AO}+\lambda\overrightarrow{OX}$$

Por tanto,

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}+\lambda\begin{pmatrix} -a_1 \\ -a_2 \end{pmatrix}+\lambda\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}+\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$




* * *



DEFINICIÓN 2

Dado un espacio afín euclídeo $A=(\mathscr{P}, V)$ una aplicación afín $T$ en $A$ se llama *movimiento* si conserva las distancias entre puntos, es decir,

$$d(T(P), T(Q))=d(P, Q)$$

para todo $P, Q\in\mathscr{P}$.

*Nota.* Los movimientos en el espacio afín $\mathbb{R}^n$ reciben también el nombre de *isometrías;* la palabra «isometría» proviene del griego y significa «igual medida».

Sea $T$ un movimiento en un espacio afín $A$, $P$ y $Q$ dos puntos cualesquiera de $\mathscr{P}$ y $P'=T(P)$, $Q'=T(Q)$ las imágenes de $P$ y $Q$, respectivamente, mediante $T$. Entonces

$$\|\mathscr{T}(\overrightarrow{PQ})\|=\|\overrightarrow{T(P)T(Q)}\|=d(P', Q')=d(P, Q)=\|\overrightarrow{PQ}\|$$

Por tanto, $\mathscr{T}$ es una aplicación lineal del espacio euclídeo $V$ que conserva la longitud de los vectores. Tales aplicaciones son las aplicaciones ortogonales en $V$ que han sido estudiadas en el capítulo 8. Si $T$ es la matriz de $\mathscr{T}$ con una cierta base se tiene que $|T|=\pm 1$. Si $|T|=1$ el movimiento se llama *directo* y si $|T|=-1$ el movimiento se llama *inverso.* Observar que estos conceptos no dependen de la base elegida para escribir la matriz de $\mathscr{T}$.

Es claro que los resultados del capítulo 8 sobre aplicaciones ortogonales nos pueden ayudar ahora a clasificar los movimientos en $\mathbb{R}^n$. Es evidente que, al menos en $\mathbb{R}^2$, puede hacerse un estudio más geométrico de los movimientos.

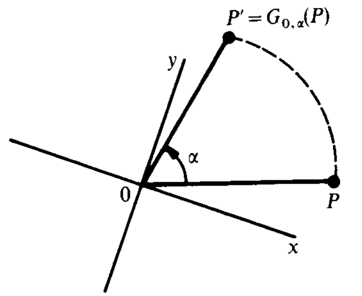
En lo que resta de este capítulo procederemos de la siguiente manera: comenzaremos haciendo un estudio geométrico de los movimientos en el plano, para realizar a continuación su estudio analítico; en el espacio, el estudio geométrico de los movimientos resulta más complicado y, por tanto, preferimos utilizar los resultados del capítulo 8 sobre aplicaciones ortogonales, lo cual nos dará directamente su expresión analítica y su clasificación.

* * *

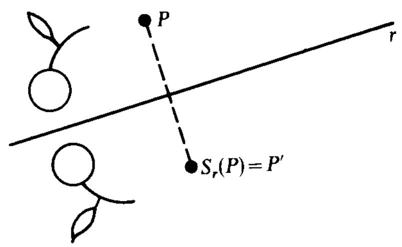
Además de la transformación identidad, damos a continuación varios ejemplos de movimientos.

EJEMPLO D. En un plano, un giro o rotación de ángulo $\alpha$ alrededor de un punto $O$ (no necesariamente el origen de coordenadas) es un movimiento, que se denotará por $G_{o,\alpha}$. Eligiendo un sistema de coordenadas como en la figura adjunta

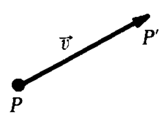
$$T(P)=\begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}P$$




EJEMPLO E. Una *simetría* con respecto a una recta $r$, que denotaremos por $S_r$, es un movimiento. Recordamos que el simétrico de un punto $P$ con respecto a $r$ es un punto $P'$ tal que el segmento $\overline{PP'}$ es perpendicular a $r$ y $r$ corta a $PP'$ en su punto medio. Con la simetría respecto de una recta también se asocia a veces el nombre de reflexión.

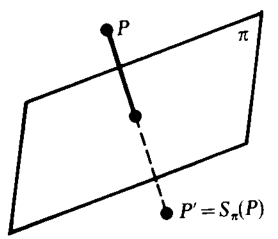



EJEMPLO F. Una traslación mediante un cierto vector $\vec{v}$, que denotaremos por $T_{\vec{v}}$, es también un movimiento.



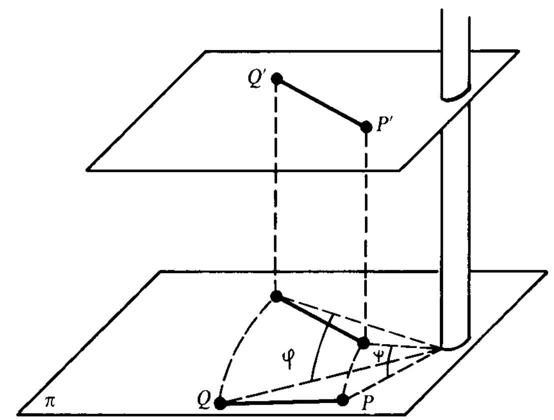

En este movimiento se produce un desplazamiento paralelo de todos los puntos.

EJEMPLO G. En el espacio, una *simetría* con respecto a un plano $\pi$, $S_\pi$, es un movimiento. El plano $\pi$ divide al segmento $PP'$, donde $P'$ es la imagen de $P$, en dos partes iguales, siendo $PP'$ perpendicular a $\pi$.






EJEMPLO H. El movimiento que se produce en el espacio al subir por una escalera de caracol se denomina un *movimiento helicoidal.* El efecto de un movimiento helicoidal está producido por un giro en un plano $\pi$ seguido de una traslación con vector perpendicular a $\pi$ de manera que los puntos del plano (o peldaño) $\pi$ se transforman en puntos de un plano paralelo a $\pi$ (o peldaño superior) (ver figura adjunta).




* * *

En el estudio de los movimientos es interesante considerar un tipo de puntos especiales que se denominan *puntos fijos:* un punto $P$ es un punto fijo de un movimiento $T$ si $T(P)=P$.

Todos los puntos son fijos para el movimiento identidad, mientras que las traslaciones no tienen ningún punto fijo. Un giro de centro $O$ tiene únicamente este punto como fijo (en el plano) y una simetría con respecto a una recta $r$ deja a todos los puntos de $r$ fijos.

En el estudio geométrico de los movimientos que haremos en la siguiente sección es conveniente considerar la composición de dos de ellos.

La composición de dos aplicaciones afines $T_1$ y $T_2$ es otra aplicación afín; en efecto, si $\mathscr{T}_1$ y $\mathscr{T}_2$ son las aplicaciones lineales asociadas a $T_1$ y $T_2$, $\mathscr{T}_2 \circ \mathscr{T}_1$ es la aplicación lineal asociada con $T_2 \circ T_1$, ya que

$$\overrightarrow{T_2 \circ T_1(P) T_2 \circ T_1(Q)} = \mathscr{T}_2(\overrightarrow{T_1(P)T_1(Q)}) = \mathscr{T}_2 \circ \mathscr{T}_1(\overrightarrow{PQ})$$

para todo par de puntos $P$ y $Q$. Si $T_2 \circ T_1 = I$ se dice que $T_1$ y $T_2$ son inversos una de otra.

Si $T_1$ y $T_2$ son movimientos, su composición es también un movimiento, ya que

$$d(T_2 \circ T_1(P), T_2 \circ T_1(Q)) = d(T_1(P), T_1(Q)) = d(P, Q)$$

para todo par de puntos $P$ y $Q$.

EJERCICIOS 10.1

**1.** Demostrar que una simetría en el plano con respecto a una recta es una aplicación afín. Encontrar la expresión analítica de la simetría con respecto a la recta $x=1$ en un sistema de referencia cartesiano.

**2.** Encontrar la expresión analítica (en un sistema de referencia cartesiano) de los siguientes movimientos en $\mathbb{R}^3$:

*a*) Giro $G$ de 90° con respecto al eje $x=1$, $y=1$.

*b*) Traslación $T$ de vector $\vec{v}=(0, 1, 1)$.

*c*) $T \circ G$ y $G \circ T$ que se obtienen como composición de los anteriores.

**3.** Determinar los puntos fijos, si existen, de la aplicación afín en $\mathbb{R}^3$ dada por

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 4 \\ 2 \\ 2 \end{pmatrix} + \frac{1}{3}\begin{pmatrix} -1 & -2 & -2 \\ -2 & 2 & -1 \\ -2 & -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

en un sistema de referencia cartesiano.

## 10.2. MOVIMIENTOS EN EL PLANO

Como ya hemos anunciado en la sección anterior comenzamos haciendo un estudio geométrico de los movimientos en el plano. Algunos de estos movimientos se han dado en los ejemplos *D*, *E* y *F* de la sección anterior: ¿serán todos estos los movimientos posibles en el plano? Para responder a esta pregunta es necesario estudiar la composición de dos movimientos fijados de antemano.

Comenzamos con un lema.

LEMA 1

Si un movimiento en un plano tiene más de un punto fijo, debe ser la identidad o una simetría con respecto a una recta que pasa por los puntos fijos.

*Demostración.* Sea $T$ un movimiento con dos puntos fijos distintos $A$ y $B$. Demostraremos primero que todos los puntos de la recta $r$ que pasa por $A$ y $B$ son puntos fijos de $T$. Sea $P \in r$ y sea $P'=T(P)$. Como $T$ es un movimiento $d(A, P)=d(A, P')$ y $d(B, P)=d(B, P')$. Por tanto, $P$ y $P'$ deben estar en las circunferencias de centro $A$ y radio $d(A, P)$ y de centro $B$ y radio $d(B, P)$; como estas circunferencias se cortan únicamente en $P$ se ha de tener $P'=P$. Esto prueba que todos los puntos de $r$ son fijos para la aplicación $T$.



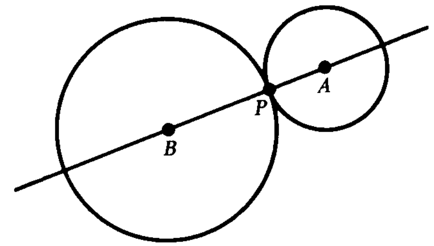



Figura 1

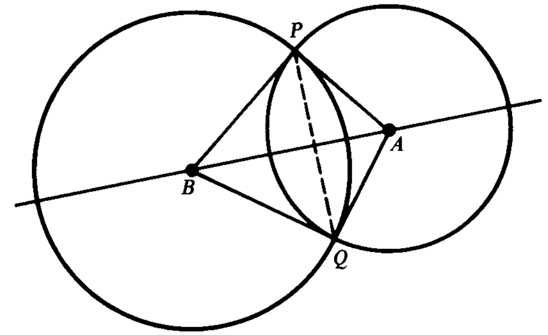



Figura 2

Sea ahora $P$ un punto que no está en $r$; el mismo razonamiento anterior muestra que $P'=T(P)$ debe pertenecer a la vez a las circunferencias de centro $A$ y radio $d(A, P)$ y de centro $B$ y radio $d(B, P)$. Por tanto, $P'=P$ o $P'=Q$ (ver figura 2). En el primer caso demostraremos que $T$ es la identidad y en el segundo que $T$ es una simetría con respecto a la recta $r$.

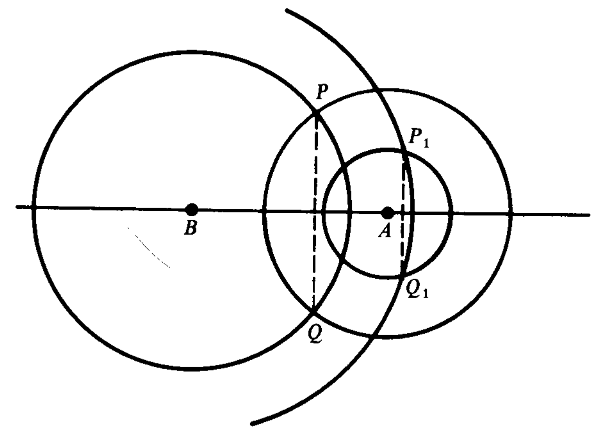



Figura 3

Supongamos que $P'=P$ y sea $P_1$ otro punto distinto de $P$ con $P_1 \notin r$; basta demostrar que $P_1=P'_1$, donde $P'_1=T(P_1)$. Por el mismo razonamiento anterior se tendría que $P'_1=P_1$ o $P'_1=Q_1$ (ver figura 3). Si fuera $P'_1=Q_1$ se tendría que

$$d(P, P_1)=d(P', P'_1)=d(P, Q_1)$$

de donde se deduciría que $P$ está en la recta $r$, en contra de lo supuesto. Por tanto, $P'_1=P_1$ y $T$ es la identidad.

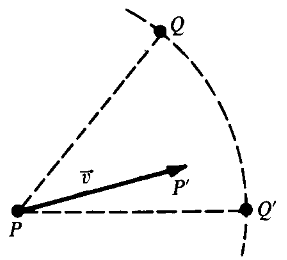
Supongamos ahora que $P'=Q$ y que $P'_1=P_1$. Se tendría entonces que

$$d(Q, P_1)=d(P', P'_1)=d(P, P_1)$$

ya que $T$ es un movimiento; de aquí deducimos que $P_1$ tiene que estar en la mediatriz de $\overline{PQ}$, que es la recta $r$. Como esto es una contradicción, ya que $P_1 \notin r$, se tiene que $P'_1 = Q_1$ y $T$ es, por tanto, una simetría con respecto a la recta $r$. ■

Sea $M$ un movimiento cualquiera en el plano, $P$ un punto del plano y $P' = M(P)$. Si $\vec{v} = \overrightarrow{PP'}$, el movimiento $T_{\vec{v}}^{-1} \circ M$ deja fijo el punto $P$ ($T_{\vec{v}}^{-1}$ es la inversa de la traslación de vector $\vec{v}$, es decir, $T_{-\vec{v}}$):

$$T_{\vec{v}}^{-1} \circ M(P) = T_{\vec{v}}^{-1}(P') = T_{-\vec{v}}(P') = P$$




Para cualquier otro punto $Q$ del plano sea $Q' = T_{\vec{v}}^{-1} \circ M(Q)$. Como $T_{\vec{v}}^{-1} \circ M$ es un movimiento $d(P, Q) = d(P, Q')$ y, por tanto, existe un giro $G_{P,\alpha}$ de centro $P$ y ángulo $\alpha$ tal que $G_{P,\alpha}(Q') = Q$. El movimiento $G_{P,\alpha} \circ T_{\vec{v}}^{-1} \circ M$ tiene $P$ y $Q$ como puntos fijos, ya que

$$G_{P,\alpha} \circ T_{\vec{v}}^{-1} \circ M(P) = G_{P,\alpha}(P) = P$$

y

$$G_{P,\alpha} \circ T_{\vec{v}}^{-1} \circ M(Q) = G_{P,\alpha}(Q') = Q$$

Por el lema 1, $G_{P,\alpha} \circ T_{\vec{v}}^{-1} \circ M$ es una simetría $S_r$, donde $r$ es una recta que pasa por $P$ o la identidad; por tanto,

$$M = T_{\vec{v}} \circ G_{P,-\alpha} \circ S_r \quad \text{o} \quad M = T_{\vec{v}} \circ G_{P,-\alpha}$$

Hemos demostrado el siguiente resultado:

PROPOSICIÓN 2

Todo movimiento en un plano es o bien la composición de un giro y una traslación o bien la composición de una simetría con respecto a una recta, un giro y una traslación; el eje de simetría pasa por el centro del giro.

*Observación.* El orden de composición en la proposición anterior es muy importante.



Una vez establecida la proposición 2 la determinación de todos los movimientos en el plano requiere el estudio de las posibles composiciones de los movimientos que allí intervienen. Antes de realizar este estudio analizamos las condiciones en que se obtiene la identidad o una simetría en el lema 1. Para ello observamos en la figura 4 que las simetrías cambian la orientación de la figura, mientras que la identidad evidentemente no la cambia.

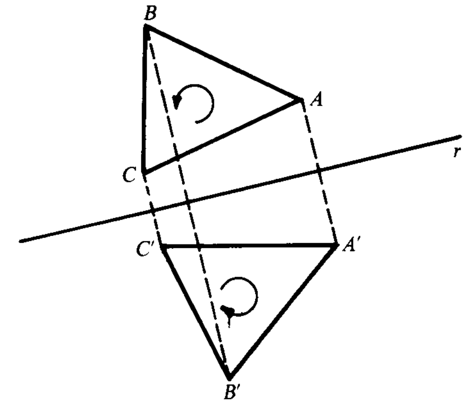



Figura 4

LEMA 3

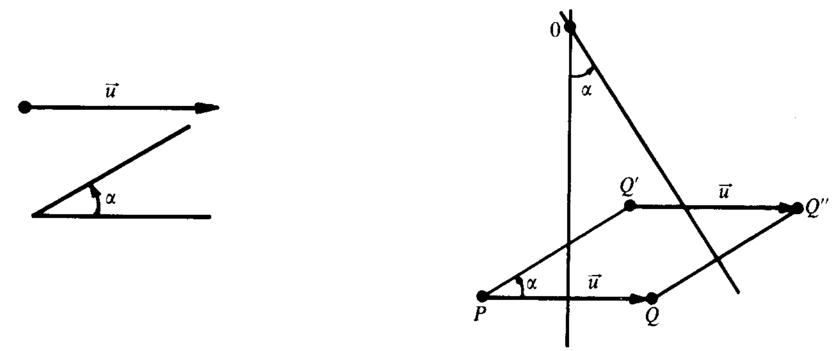
Si un movimiento en un plano tiene más de un punto fijo y no cambia la orientación de las figuras es la identidad y si la cambia es una simetría axial.

PROPOSICIÓN 4

La composición de un giro y una traslación es otro giro del mismo ángulo.

*Demostración.* Sean $G_{P,\alpha}$ y $T_{\vec{u}}$ el giro y la traslación dados (ver figura adjunta). Sea

$$Q = T_{\vec{u}}(P)$$
$$Q' = G_{P,\alpha}(Q)$$
$$Q'' = T_{\vec{u}}(Q')$$



Sea $O$ el punto de intersección de las mediatrices de los segmentos $\overline{PQ}$ y $\overline{QQ''}$. El movimiento $G_{P,\alpha}^{-1}\circ T_{\vec u}^{-1}\circ G_{0,\alpha}$ tiene $P$ y $Q$ como puntos fijos:

$$G_{P,\alpha}^{-1}\circ T_{\vec u}^{-1}\circ G_{0,\alpha}(P)=G_{P,\alpha}^{-1}(T_{\vec u}^{-1}(Q))=G_{P,\alpha}^{-1}(P)=P$$

y

$$G_{P,\alpha}^{-1}\circ T_{\vec u}^{-1}\circ G_{0,\alpha}(Q)=G_{P,\alpha}^{-1}(T_{\vec v}^{-1}(Q''))=G_{P,\alpha}^{-1}(Q')=Q$$

Puesto que este movimiento conserva la orientación (ya que las traslaciones y los giros la conservan) se tiene que $G_{P,\alpha}^{-1}\circ T_{\vec u}^{-1}\circ G_{0,\alpha}=I$, de donde deducimos que $G_{0,\alpha}=T_{\vec u}\circ G_{P,\alpha}$. Esto concluye la demostración de la proposición 4 excepto en el caso en que $Q''$, $Q$ y $P$ están en una misma recta; en este caso se ha de tener necesariamente $\alpha=180°$ y $Q''=P$. En esta circunstancia basta tomar $O$ como el punto medio del segmento $\overline{PQ}$. ■

*Observación.* El centro del nuevo giro que se obtiene en la proposición 4 es el punto de intersección de las mediatrices de los segmentos $\overline{PQ}$ y $\overline{QQ''}$, donde $Q=T_{\vec u}$ y $Q''=T_{\vec u}\circ G_{P,\alpha}\circ T_{\vec u}(P)$, siempre que $\alpha\neq 180°$. Si $\alpha=180°$, el centro es el punto medio de $\overline{PQ}$.

PROPOSICIÓN 5

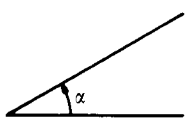
La composición de una simetría y un giro de centro perteneciente al eje de simetría es otra simetría.

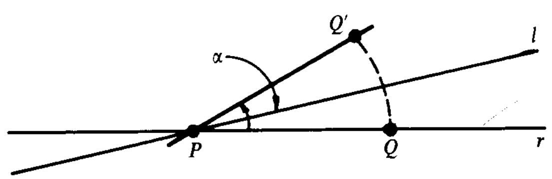
*Demostración.* Sean $S_r$ y $G_{P,\alpha}$, con $P\in r$, la simetría y el giro dados. Sea $Q$ otro punto de $r$ distinto de $P$ y $Q'=G_{P,\alpha}(Q)$. Sea $l$ la bisectriz del ángulo $QPQ'$. El movimiento

$$S_r^{-1}\circ G_{P,\alpha}^{-1}\circ S_l$$

tiene $P$ y $Q$ como puntos fijos:

$$S_r^{-1}\circ G_{P,\alpha}^{-1}\circ S_l(P)=P \quad \text{y} \quad S_r^{-1}\circ G_{P,\alpha}^{-1}\circ S_l(Q)=S_r^{-1}(G_{P,\alpha}^{-1}(Q'))=S_r^{-1}(Q)=Q$$



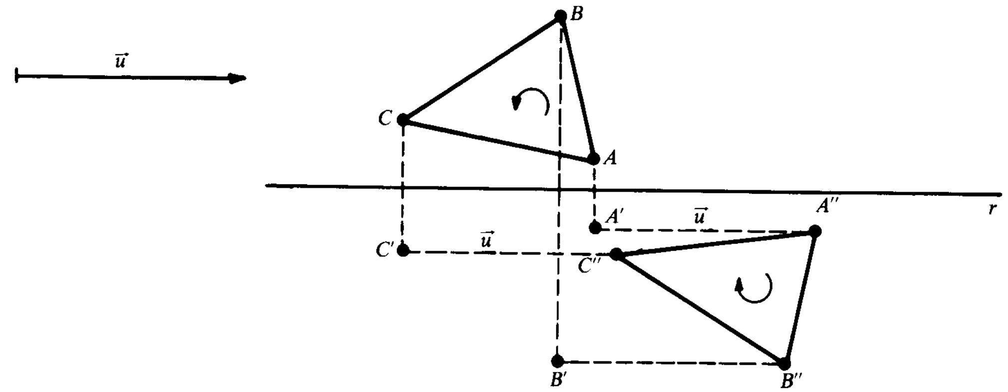




Como este movimiento no cambia la orientación (ya que las dos simetrías anulan el cambio de sentido que cada una de ellas produce) del lema 3 se deduce que $S_r^{-1}\circ G_{P,\alpha}^{-1}\circ S_l=I$ y, por tanto, $S_l=G_{P,\alpha}\circ S_r$. ■

*Observación.* La recta de la nueva simetría axial en la proposición anterior es la bisectriz del ángulo $QPQ'$, con $Q'=G_{P,\alpha}(Q)$ y $Q\in r$, $Q\neq P$.



De todos los casos posibles que pueden ocurrir en la proposición 2 nos queda por estudiar la composición de una simetría y una traslación. Comenzamos suponiendo que el vector de traslación $\vec u$ es *paralelo* a la recta de simetría $r$.

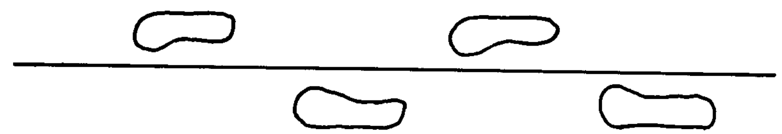



En la figura se observa que este movimiento cambia la orientación y, sin embargo, no es una simetría axial. Estamos ante un nuevo movimiento al cual damos el nombre de *simetría deslizante*.

DEFINICIÓN

A la composición de una simetría de recta $r$ y una traslación de vector $\vec u$ paralelo a $r$ se le denomina *simetría deslizante* y se representa por $S_{r,\vec u}$.

El lector puede visualizar una simetría deslizante como las huellas que una persona deja en la nieve al caminar en línea recta.



Supongamos ahora que el vector de traslación $\vec v$ es perpendicular a la recta $r$. En este caso demostraremos el siguiente resultado:

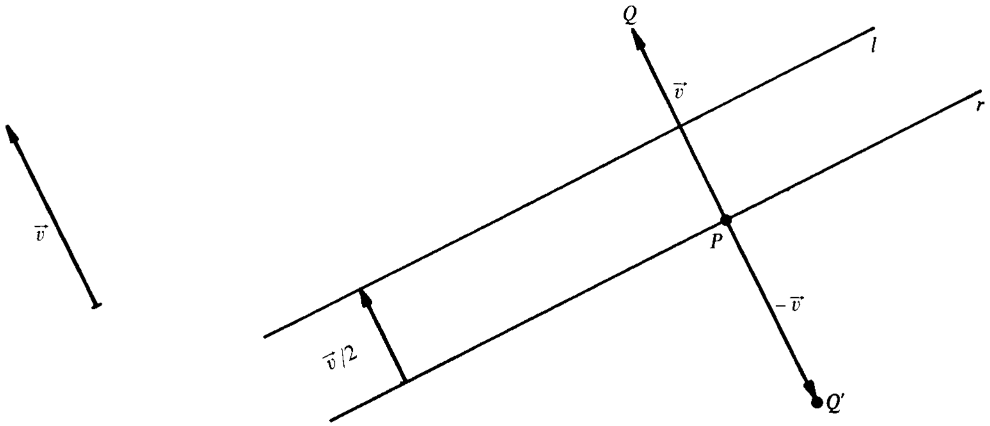
(*) La composición de una simetría $S_r$ y una traslación de vector $\vec v$ perpendicular a $r$ es una simetría con respecto a la recta $l$ trasladada de $r$ mediante el vector $\vec v/2$.

Basta observar que $S_r^{-1}\circ T_{\vec v}^{-1}\circ S_l$ es un movimiento que tiene $P\in r$ y $Q=T_{\vec v}(P)$ como puntos fijos:

$$S_r^{-1}\circ T_{\vec v}^{-1}\circ S_l(P)=S_r^{-1}\circ T_{\vec v}^{-1}(Q)=S_r^{-1}(P)=P$$

y

$$S_r^{-1}\circ T_{\vec v}^{-1}\circ S_l(Q)=S_r^{-1}\circ T_{\vec v}^{-1}(P)=S_r^{-1}(Q')=Q.$$




Puesto que este movimiento conserva la orientación, del lema 3 se deduce que $S_r^{-1}\circ T_{\vec{v}}^{-1}\circ S_l=I$, y, por tanto, $S_l=T_{\vec{v}}\circ S_r$, que es la afirmación (*).

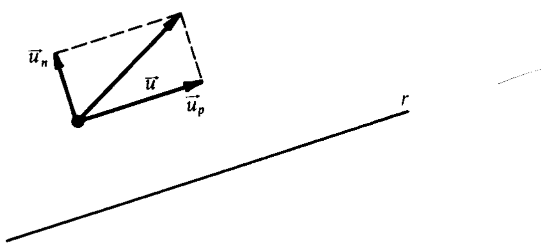
Puesto que conocemos el resultado de componer una simetría con una traslación de vector paralelo al eje de simetría y con una traslación de vector perpendicular al eje de simetría, podemos encontrar el resultado de componer una simetría con un vector cualquiera $\vec{u}$, sin más que descomponer $\vec{u}$ en sus componentes paralela y normal al eje de simetría. Se obtiene así el siguiente resultado.

PROPOSICIÓN 6

La composición de una simetría y una traslación es o bien una simetría deslizante o bien una simetría.

*Demostración.* Sea $S_r$ la simetría y $T_{\vec{u}}$ la traslación. Sea $\vec{u}=\vec{u}_p+\vec{u}_n$ la descomposición de $\vec{u}$ en un vector $\vec{u}_p$ paralelo a $r$ y un vector $\vec{u}_n$ perpendicular (o normal) a $r$. Observar que $T_{\vec{u}}=T_{\vec{u}_p}\circ T_{\vec{u}_n}$, de manera que

$$T_{\vec{u}}\circ S_r=T_{\vec{u}_p}\circ T_{\vec{u}_n}\circ S_r.$$




Si $\vec{u}_p=\vec{0}$, $T_{\vec{u}_n}\circ S_r$ es una simetría, mientras que si $\vec{u}_p\neq\vec{0}$, $T_{\vec{u}_p}\circ T_{\vec{u}_n}\circ S_r$ es una simetría deslizante.



Juntando las proposiciones 4, 5 y 6 junto con la proposición 2 se obtiene el siguiente resultado, que nos dice con exactitud el tipo de movimientos que pueden existir en un plano.

TEOREMA 7 *(Descripción de los movimientos en un plano)*

Todo movimiento en un plano es o bien la *identidad*, o una *traslación* o una *rotación* (movimientos *directos*, es decir, que no cambian la orientación) o bien una *simetría* o una *simetría deslizante* (movimientos *inversos*, es decir, que cambian la orientación).

El lector puede ahora preguntarse cuál es el resultado de componer dos movimientos cualesquiera que no coinciden con los realizados en las proposiciones anteriores. Por ejemplo: ¿cuál es el resultado de componer dos giros? Es conveniente «jugar» un poco con este tipo de composiciones. Algunos de estos «juegos» se proponen como ejercicios al final de esta sección.

## EJERCICIOS 10.2

**1.** Demostrar que la composición de dos simetrías en un plano, cuyos ejes se cortan en un punto es una rotación. ¿Cuáles son el centro y el ángulo de esta rotación?

**2.** Estudiar la composición de dos simetrías de ejes paralelos.

**3.** Demostrar que la composición de una simetría y un giro en un plano es una simetría o una simetría deslizante. ¿Cuál es el eje y cuál es el vector de la simetría deslizante?

**4.** *a)* Demostrar que la composición de dos giros de 90° en el mismo sentido alrededor de dos puntos distintos $C$ y $B$ es un giro de 180° alrededor del centro de uno de los cuadrados que tiene $\overline{CB}$ como lado.

*b)* Estudiar la composición de dos giros cualesquiera en el mismo sentido.

**5.** Demostrar que la composición de tres simetrías de ejes concurrentes es otra simetría. ¿Cuál es el eje de esta nueva simetría?

**6.** Estudiar la composición de tres simetrías de ejes paralelos.

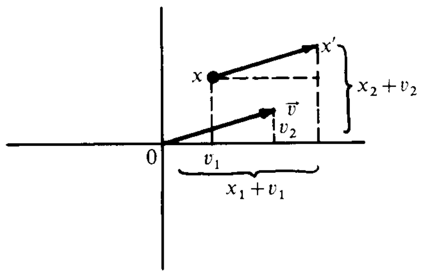
**7.** Un giro de 180° alrededor de un punto $P$ se dice también una simetría de centro $P$. Demostrar que la composición de dos simetrías de centros $P_1$ y $P_2$, respectivamente, es una traslación de vector $2\overrightarrow{P_1P_2}$.

## 10.3. ESTUDIO ANALITICO DE LOS MOVIMIENTOS EN $\mathbb{R}^2$

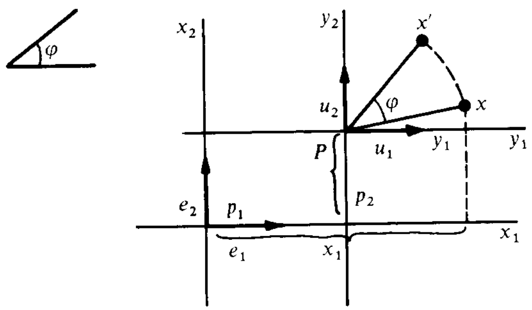
Comenzamos esta sección encontrando las ecuaciones de todos los movimientos en el plano, cuya descripción se ha dado en el teorema 7 de la sección anterior.

EJEMPLO A. *Traslación de vector* $\vec{v}=(v_1, v_2)(T_{\vec{v}})$: Cualquier punto $X=(x_1, x_2)$ se transforma en otro punto $T_{\vec{v}}(X)=X'=(x'_1, x'_2)$ tal que

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}=\begin{pmatrix} x_1+v_1 \\ x_2+v_2 \end{pmatrix}=\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}+\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$




EJEMPLO B. Giro o rotación de centro $P=(p_1, p_2)$ y ángulo $\varphi(G_{p,\varphi})$:




Si elegimos el sistema de referencia $\{P; \vec{u}_1, \vec{u}_2\}$ como muestra la figura adjunta, la imagen de un punto $X$ de coordenadas $(y_1, y_2)$ en este sistema de referencia es un punto $X'=(y'_1, y'_2)$ tal que

$$\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix}=\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

Si $X=(x_1, x_2)$ y $X'=(x'_1, x'_2)$ en el sistema de referencia canónico $\{O; \vec{e}_1, \vec{e}_2\}$ se tiene que

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} x_1-p_1 \\ x_2-p_2 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix}=\begin{pmatrix} x'_1-p_1 \\ x'_2-p_2 \end{pmatrix}$$



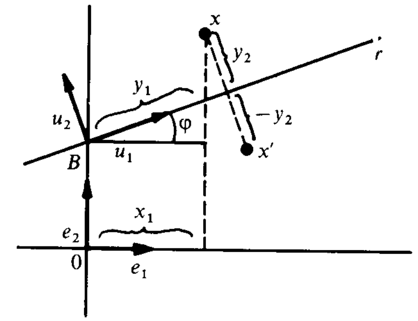
que sustituido en la expresión anterior nos permite obtener:

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}=\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}+\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}-\begin{pmatrix} p_1 \\ p_2 \end{pmatrix}\right]=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}+\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

* * *

EJEMPLO C. *Simetría con respecto a una recta* $r(S_r)$: Supongamos que $r$ tiene por ecuación $y=mx+b$; si elegimos el sistema de referencia $\{B; \vec{u}_1, \vec{u}_2\}$, donde $B=(0, b)\in r$, $\vec{u}_1$ es un vector director unitario de $r$ y $u_2$ es un vector unitario perpendicular a $r$, de manera que $\{\vec{u}_1, \vec{u}_2\}$ y $\{\vec{e}_1, \vec{e}_2\}$ tengan la misma orientación [(1)], la imagen de un punto $X$ de coordenadas $(y_1, y_2)$ en este sistema de referencia es $X'$ de coordenadas $(y'_1, y'_2)$ en este mismo sistema tal que

$$S_r\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix}=\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \qquad (*)$$




Para pasar al sistema de referencia $\{O; \vec{e}_1, \vec{e}_2\}$ realizamos un giro de ángulo $(-\varphi)$, donde $\operatorname{tg}\varphi=m$, y una traslación de vector $(0, -b)$, de manera que si $(x_1, x_2)$ son las coordenadas de $X$ en este sistema de referencia se tiene

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}+\begin{pmatrix} 0 \\ b \end{pmatrix};$$

de la misma manera, si $(x'_1, x'_2)$ son las coordenadas de $X'$ en el sistema de referencia canónico se tiene

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}=\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix}+\begin{pmatrix} 0 \\ b \end{pmatrix}$$

(1) Dos bases de un mismo espacio vectorial tienen la misma orientación si la matriz del cambio de base tiene determinante positivo.

Sustituyendo en (*) encontramos que

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi&\cos\varphi\end{pmatrix}^{-1}\left[\begin{pmatrix}x_1\\x_2\end{pmatrix}-\begin{pmatrix}0\\b\end{pmatrix}\right]+\begin{pmatrix}0\\b\end{pmatrix}=$$

$$=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}1&0\\0&-1\end{pmatrix}\begin{pmatrix}\cos\varphi & \operatorname{sen}\varphi\\ -\operatorname{sen}\varphi&\cos\varphi\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=$$

$$=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\cos2\varphi & \operatorname{sen}2\varphi\\ \operatorname{sen}2\varphi & -\cos2\varphi\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

* * *

EJEMPLO D. *Simetría deslizante de eje r y vector* $\overline{v}(S_{r,\overline{v}})$: Si $r$: $y=mx+b$, utilizando el resultado obtenido en el ejemplo C y el hecho de que $S_{r,\overline{v}}=T_{\overline{v}}\circ S_r$ se obtiene que

$$S_{r,\overline{v}}\begin{pmatrix}x_1\\x_2\end{pmatrix}=T_{\overline{v}}\circ S_r\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}v_1\\v_2\end{pmatrix}+\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\cos2\varphi & \operatorname{sen}2\varphi\\ \operatorname{sen}2\varphi & -\cos2\varphi\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=$$

$$=\begin{pmatrix}b_1\\b_2\end{pmatrix}+\begin{pmatrix}\cos2\varphi & \operatorname{sen}2\varphi\\ \operatorname{sen}2\varphi & -\cos2\varphi\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

donde $\overline{v}=(v_1, v_2)$. Recordar que $\overline{v}$ es paralelo a la recta $r$.

* * *

A continuación damos algunos ejemplos de casos particulares de los ejemplos anteriores para que el lector se familiarice con la obtención de las ecuaciones de movimientos. Una forma de realizar estos ejercicios es seguir en cada caso el razonamiento de los ejemplos anteriores; otra posibilidad es escribir la ecuación del movimiento en su forma general y encontrar los parámetros desconocidos; aquí seguiremos este segundo punto de vista.

EJEMPLO E. *Encontrar las ecuaciones del giro de centro* $P=(1, 2)$ *y ángulo* $\varphi=\pi/4$. Por tratarse de un giro de ángulo $\pi/4$ tendremos:

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\cos\frac{\pi}{4} & -\operatorname{sen}\frac{\pi}{4}\\ \operatorname{sen}\frac{\pi}{4} & \cos\frac{\pi}{4}\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\frac{\sqrt2}{2} & -\frac{\sqrt2}{2}\\ \frac{\sqrt2}{2} & \frac{\sqrt2}{2}\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$



A continuación calculamos $\begin{pmatrix}a_1\\a_2\end{pmatrix}$ imponiendo que $P=(1, 2)$ sea un punto fijo de este movimiento:

$$\begin{pmatrix}1\\2\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}\frac{\sqrt2}{2} & -\frac{\sqrt2}{2}\\ \frac{\sqrt2}{2} & \frac{\sqrt2}{2}\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix}\Rightarrow$$

$$\Rightarrow\begin{pmatrix}a_1\\a_2\end{pmatrix}=\begin{pmatrix}1\\2\end{pmatrix}-\begin{pmatrix}\frac{\sqrt2}{2}-\sqrt2\\ \frac{\sqrt2}{2}+\sqrt2\end{pmatrix}=\begin{pmatrix}1+\frac{\sqrt2}{2}\\ 2-\frac{3\sqrt2}{2}\end{pmatrix}$$

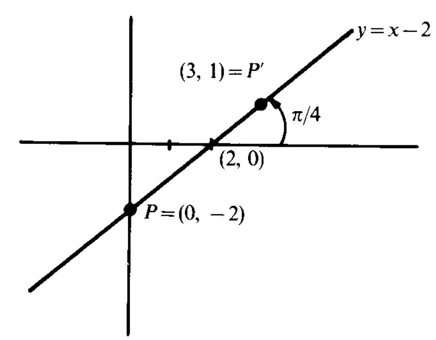
Por tanto,

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}1+\frac{\sqrt2}{2}\\ 2-\frac{3\sqrt2}{2}\end{pmatrix}+\begin{pmatrix}\frac{\sqrt2}{2} & -\frac{\sqrt2}{2}\\ \frac{\sqrt2}{2} & \frac{\sqrt2}{2}\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

* * *

EJEMPLO F. *Encontrar la expresión analítica de la simetría deslizante de eje* $y=x-2$ *y de vector* $\overline{v}=(3, 3)$. Como $\varphi=\pi/4$ del ejemplo D se deduce que las ecuaciones buscadas son de la forma

$$\begin{pmatrix}x_1'\\x_2\end{pmatrix}=\begin{pmatrix}b_1\\b_2\end{pmatrix}+\begin{pmatrix}0&1\\1&0\end{pmatrix}\begin{pmatrix}x_1'\\x_2\end{pmatrix}.$$



Calcularemos $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ si conocemos la imagen de un punto $P$; si tomamos $P=(0, -2)$,

$$S_{r,\vec{v}}\begin{pmatrix} 0 \\ -2 \end{pmatrix}=\begin{pmatrix} 0 \\ -2 \end{pmatrix}+\begin{pmatrix} 3 \\ 3 \end{pmatrix}=\begin{pmatrix} 3 \\ 1 \end{pmatrix},$$

de manera que

$$\begin{pmatrix} 3 \\ 1 \end{pmatrix}=\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}+\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ -2 \end{pmatrix} \Rightarrow \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}=\begin{pmatrix} 3 \\ 1 \end{pmatrix}-\begin{pmatrix} -2 \\ 0 \end{pmatrix}=\begin{pmatrix} 5 \\ 1 \end{pmatrix}.$$

Por tanto, la expresión analítica de esta simetría deslizante es

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}=\begin{pmatrix} 5 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

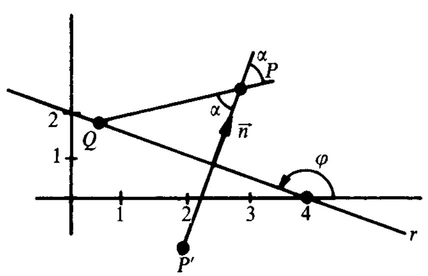
* * *

EJEMPLO G. *Encontrar la expresión analítica de la simetría con respecto a la recta r de ecuación* $x+2y=4$. Una forma de hacer este ejercicio es como en el ejemplo F, teniendo en cuenta que todos los puntos del eje deben quedar fijos. Dejamos al lector esta tarea. Nosotros mostraremos aquí una forma más geométrica de resolver este problema. La imagen de un punto $P$ es un punto $P'$ tal que

$$P'=P-\lambda\vec{n}$$

donde $\vec{n}$ es un vector unitario normal a $r$ (como muestra la figura) y $\lambda=\|\overrightarrow{PP'}\|=2d(P, r)$. Para calcular $\lambda$ tomamos un punto cualquiera $Q$ de la recta $r$ y observamos que

$$d(P, r)=\|\overrightarrow{QP}\|\cos\alpha=\|\overrightarrow{QP}\|\frac{(\overrightarrow{QP}, \vec{n})}{\|\overrightarrow{QP}\|\,\|\vec{n}\|}=(\overrightarrow{QP}, \vec{n})$$




La ecuación de la simetría es

$$P'=P-2(\overrightarrow{QP}, \vec{n})\vec{n}$$



En nuestro ejemplo, si $P=(x_1, x_2)$, $P'=(x_1', x_2')$, tomamos $Q=(0, 2)$ y observamos que $\vec{n}=(1, 2)\frac{1}{\sqrt{5}}$; tenemos

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}-2\left(\begin{pmatrix} x_1-0 \\ x_2-2 \end{pmatrix}, \frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right)\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}=$$

$$=\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}-\frac{2}{5}\begin{pmatrix} x_1+2x_2-4 \\ 2x_1+4x_2-8 \end{pmatrix}=\begin{pmatrix} 8/5 \\ 16/5 \end{pmatrix}+\begin{pmatrix} \frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & -\frac{3}{5} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

* * *

Con los ejemplos anteriores hemos aprendido a encontrar las ecuaciones de los movimientos en el plano. A continuación nos planteamos el problema recíproco, es decir, dadas las ecuaciones de un movimiento en el plano determinar qué tipo de movimiento es. Si se resuelve este problema, haciendo uso de los resultados del comienzo de la sección 8.8 del capítulo 8 se obtiene la descripción de los movimientos en el plano, que fue dada en el teorema 7 de la sección anterior.

En la sección 10.1 se demostró que las ecuaciones de un movimiento $M$ en el plano son de la forma

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix}=M\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}+T\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \tag{1}$$

donde $T$ es una matriz ortogonal de orden 2. De los cuatro primeros ejemplos de esta sección, que nos dan todos los movimientos en el plano de acuerdo con la descripción dada en la sección anterior, deducimos que $T$ es de la forma

$$T=\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix} \quad \text{ó} \quad T=\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{pmatrix}$$

Observar que la traslación está incluida en el primer caso cuando $\varphi=0$. Además, en el segundo caso, $T$ es una matriz ortogonal y simétrica de determinante $-1$, y por tanto, la aplicación lineal que tiene a $T$ como matriz es diagonalizable en una base ortonormal y su matriz en esta base es

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Esto último se deduce de los resultados del capítulo 8.

El lector se habrá dado cuenta que la descripción de $T$ se ha encontrado anteriormente al comienzo de la sección 8.8 (ver la discusión que precede al teorema 1 de la citada sección). En resumen, los casos posibles para $T$ son

$$T=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}\qquad\text{(I)}$$

o, en una nueva base $\{\vec{u}_1, \vec{u}_2\}$ si es necesario,

$$T=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\qquad\text{(II)}$$

Para seguir adelante en el problema propuesto es necesario estudiar los puntos fijos del movimiento. Los puntos fijos $X=(x_1, x_2)$ de (1) satisfacen el sistema de ecuaciones

$$\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}a_1\\ a_2\end{pmatrix}+T\begin{pmatrix}x_1\\ x_2\end{pmatrix}$$

o bien

$$(I-T)\begin{pmatrix}x_1\\ x_2\end{pmatrix}=\begin{pmatrix}a_1\\ a_2\end{pmatrix}\qquad(2)$$

Debido al teorema de Rouche-Frobenius el estudio de los puntos fijos de (1) se reduce al estudio del rango de la matriz $I-T$ y de la matriz ampliada $(I-T\,|\,A)$, donde $A=\begin{pmatrix}a_1\\ a_2\end{pmatrix}$. Estos rangos se designarán mediante

$$r(I-T) \quad \text{y} \quad r(I-T\,|\,A)$$

En el caso (I):

$$|I-T|=\begin{vmatrix}1-\cos\varphi & \operatorname{sen}\varphi\\ -\operatorname{sen}\varphi & 1-\cos\varphi\end{vmatrix}=(1-\cos\varphi)^2+\operatorname{sen}^2\varphi=2(1-\cos\varphi)$$

Por tanto, $r(I-T)=2$ si y sólo si $\varphi\neq 0$; en este caso, $r(I-T\,|\,A)=2$ y, en consecuencia, el movimiento posee un único punto fijo $P$ que es la solución del sistema (2). Tomando $P$ como origen del sistema de referencia las ecuaciones del movimiento son de la forma

$$\begin{pmatrix}y_1'\\ y_2'\end{pmatrix}=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}$$

y, por tanto, el movimiento es un *giro de centro $P$ y ángulo $\varphi$*.

Si $\varphi=0$,

$$T=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$$



y, por tanto, $r(I-T)=0$. Si $A=\begin{pmatrix}0\\ 0\end{pmatrix}$, $r(I-T\,|\,A)=0$ y el movimiento es la *identidad*. Si $A\neq\begin{pmatrix}0\\ 0\end{pmatrix}$, $r(I-T\,|\,A)=1$ y el movimiento es una *traslación* de vector $(a_1, a_2)$. Observar que todos los movimientos del caso (I) son *directos*, ya que $|T|=1$.

En el caso (II), y en la base $\{\vec{u}_1, \vec{u}_2\}$,

$$I-T=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}=\begin{pmatrix}0 & 0\\ 0 & 2\end{pmatrix}$$

y, por tanto, $r(I-T)=1$. Si $r(I-T\,|\,A)=1$, el sistema (2) tiene una recta de soluciones y, por tanto, el movimiento tiene una recta de puntos fijos. Si $P$ es uno de estos puntos fijos, en el sistema de referencia $\{P; \vec{u}_1, \vec{u}_2\}$ el movimiento tiene como ecuaciones

$$\begin{pmatrix}y_1'\\ y_2'\end{pmatrix}=\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}y_1\\ y_2\end{pmatrix}$$

y se trata, por tanto, de una *simetría con respecto* a la *recta de puntos fijos*. Observar que en este caso se ha de tener $a_1=0$ en la ecuación (1).

Si $r(I-T\,|\,A)=2$ el sistema (2) no tiene soluciones y, por tanto, el movimiento no posee puntos fijos. En este caso podemos escribir el movimiento como

$$\begin{pmatrix}x_1'\\ x_2'\end{pmatrix}=\begin{pmatrix}a_1\\ a_0\end{pmatrix}+\left[\begin{pmatrix}0\\ a_2\end{pmatrix}+\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}x_1\\ x_2\end{pmatrix}\right]\qquad(3)$$

donde necesariamente $a_1\neq 0$, ya que $r(I-T\,|\,A)=2$. La expresión entre corchetes son las ecuaciones de una simetría con respecto a un eje que tiene como vector director $\vec{u}_1$. Por tanto, (3) es la composición de esta simetría con la traslación de vector $\vec{a}=(a_1, 0)$, que es paralelo a $\vec{u}_1$: se trata de una *simetría deslizante*.

La descomposición de una simetría deslizante como una simetría y una traslación no es única; únicamente existe unicidad si el vector de traslación es paralelo al eje de simetría y en este caso la descomposición de la simetría deslizante se dice *canónica*.

A continuación resumimos los resultados anteriores en el siguiente cuadro:

| $T$ | $r(I-T)$ | $r(I-T\,|\,A)$ | Movimiento $P'=A+T(P)$ |
|---|---|---|---|
| $I$<br>$\lvert T\rvert>0$ | 0 | 0<br>1 | *Identidad*<br>*Traslación* |
| Simetría respecto a una recta:<br>$\lvert T\rvert<0$ | 1 | 1<br>2 | *Simetría* respecto de la recta de puntos fijos.<br>*Simetría deslizante:* composición de simetría y traslación paralela al eje de simetría. |
| Giro o rotación:<br>$\lvert T\rvert>0$ | 2 | 2 | *Giro* en torno al punto fijo. |

* * *

Para finalizar esta sección daremos algunos ejemplos de determinación de movimientos en el plano. A la vez aprenderemos a encontrar los elementos geométricos de cada uno de ellos, es decir, el vector de traslación si se trata de una traslación, el centro de giro y su ángulo si se trata de un giro, la recta de simetría si se trata de una simetría, o la descomposición canónica si se trata de una simetría deslizante.

Para referencias futuras diremos que una subvariedad lineal $L$ en un espacio afín es invariante mediante una aplicación afín $M$ si $M(A) \subset A$. En el plano, y dejando a un lado el plano, las subvariedades invariantes pueden ser puntos o rectas. Observar que en el plano hay movimientos que no tienen puntos fijos, pero que poseen subvariedades invariantes: ¿cuáles?

EJEMPLO H. Estudiar el movimiento cuya expresión es

$$P' = \begin{pmatrix} 3 \\ 5 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} P$$

Puesto que $|I-T| = \begin{vmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{vmatrix} = \frac{1}{4} + \frac{3}{4} = 1 \neq 0$, $r(I-T)=2$ y se trata de un giro.

El ángulo de giro $\varphi$ satisface $\cos\varphi = \frac{1}{2}$, $\operatorname{sen}\varphi = -\frac{\sqrt{3}}{2}$ y tenemos entonces $\varphi = -60°$. El punto fijo $P$ satisface la ecuación

$$P = \begin{pmatrix} 3 \\ 5 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix} P$$

Despejando $P$ se obtiene

$$P = \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} \frac{3}{2} + \frac{5\sqrt{3}}{2} \\ -\frac{3\sqrt{3}}{2} + \frac{5}{2} \end{pmatrix}$$

* * *



EJEMPLO I. *Estudiar el movimiento cuya expresión es*

$$P' = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} P$$

Puesto que $I-T = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, $r(I-T)=1$ y puesto que $(I-T\,|\,A) = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}$, $r(I-T\,|\,A)=1$; se trata, entonces, de una *simetría*. El eje de simetría es el conjunto de los puntos fijos y, por tanto, tiene por ecuación

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \Leftrightarrow x = 1$$

* * *

EJEMPLO J. *Estudiar el movimiento cuya expresión es*

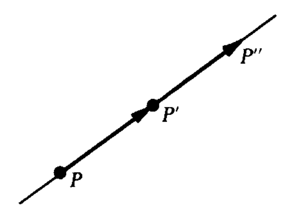
$$P' = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix} P$$

Puesto que $|I-T| = \begin{vmatrix} 1-\frac{\sqrt{3}}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1+\frac{\sqrt{3}}{2} \end{vmatrix} = 1 - \frac{3}{4} - \frac{1}{4} = 0$, $r(I-T)=1$ y puesto que

$$(I-T\,|\,A) = \begin{pmatrix} 1-\frac{\sqrt{3}}{2} & -\frac{1}{2} & 1 \\ -\frac{1}{2} & 1+\frac{\sqrt{3}}{2} & 0 \end{pmatrix}$$

$r(I-T\,|\,A)=2$. Se trata de un *simetría deslizante*. Para calcular el eje de simetría procedemos como sigue: si $P$ es un punto del eje de simetría, su imagen $P'$ pertenece también al eje de simetría y $P''$, la imagen de $P'$, pertenece también al eje de simetría de manera que

$$\overrightarrow{PP'} = \overrightarrow{P'P''} = A + T(P') - A - T(P) = T(\overrightarrow{PP'})$$



Por tanto, $\overrightarrow{PP'}$ es un autovector de $T$ de autovalor 1 y se tiene que

$$0=(T-I)(\overrightarrow{PP'})=(T-I)(A+T(P)-P)=(T-I)(A)+(T-I)^2(P)$$

que es la ecuación que satisface el eje de simetría.

En nuestro ejemplo:

$$0=\begin{pmatrix} \frac{\sqrt{3}}{2}-1 & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2}-1 \end{pmatrix}\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix} \frac{\sqrt{3}}{2}-1 & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2}-1 \end{pmatrix}^2\begin{pmatrix}x\\y\end{pmatrix} \Leftrightarrow 1-\frac{\sqrt{3}}{2}=(2-\sqrt{3})x-y$$

Las ecuaciones paramétricas del eje de simetría son

$$r: \left(\frac{1}{2},\ 0\right)+t(1,\ 2-\sqrt{3})$$

El vector de la traslación es $\vec{v}=\overrightarrow{PP'}$, donde $P$ es cualquier punto del eje de simetría; si tomamos $P=\left(\frac{1}{2},\ 0\right)$,

$$P'=\begin{pmatrix}1\\0\end{pmatrix}+\begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{\sqrt{3}}{2} \end{pmatrix}\begin{pmatrix}\frac{1}{2}\\0\end{pmatrix}=\begin{pmatrix}1+\frac{\sqrt{3}}{4}\\ \frac{1}{4}\end{pmatrix}.$$

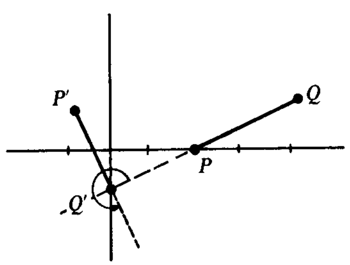
Por tanto,

$$\vec{v}=\overrightarrow{PP'}=\left(\frac{1}{2}+\frac{\sqrt{3}}{4},\ \frac{1}{4}\right).$$

* * *

EJEMPLO K. *Encontrar, si existe, un giro que lleve* $P=(2,0)$ *en* $P'=(-1,1)$ *y* $Q=(4,1)$ *en* $Q'=(0,-1)$. Puesto que

$$\cos \sphericalangle (\overrightarrow{PQ},\ \overrightarrow{P'Q'})=\frac{\left(\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}1\\-2\end{pmatrix}\right)}{5}=0$$






del dibujo deducimos que

$$\sphericalangle (\overrightarrow{PQ},\ \overrightarrow{P'Q'})=\frac{3\pi}{2}$$

La ecuación del giro toma la forma

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

Si existe un giro satisfaciendo las condiciones dadas se ha de tener

$$\begin{pmatrix}-1\\1\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\begin{pmatrix}2\\0\end{pmatrix}$$

y

$$\begin{pmatrix}0\\-1\end{pmatrix}=\begin{pmatrix}a_1\\a_2\end{pmatrix}+\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\begin{pmatrix}4\\1\end{pmatrix}$$

Si estas dos ecuaciones producen el mismo valor de $\begin{pmatrix}a_1\\a_2\end{pmatrix}$ existe el giro buscado; así es en este caso y se tiene $\begin{pmatrix}a_1\\a_2\end{pmatrix}=\begin{pmatrix}-1\\3\end{pmatrix}$. Por tanto,

$$\begin{pmatrix}x_1'\\x_2'\end{pmatrix}=\begin{pmatrix}-1\\3\end{pmatrix}+\begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix}$$

*Nota.* Este problema también puede resolverse imponiendo

$$P'=A+\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}P$$

y

$$Q'=A+\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}Q$$

y calculando $A$ y $\varphi$ si las cuatro ecuaciones anteriores son compatibles.

## EJERCICIOS 10.3.

**1.** Encontrar la expresión analítica de los siguientes movimientos:

*a)* Giro de centro $(1, 0)$ y ángulo $3\pi/4$.

*b)* Simetría deslizante de eje paralelo a la recta $2x+y=3$ y que transforma $(2, 1)$ en $(1, 0)$.

c) Giro de ángulo $\pi/3$ que lleve (2, 1) en (1, 0).

d) La composición de los movimientos de a) y b).

**2.** Estudiar los siguientes movimientos del plano, hallando su tipo, subespacios invariantes y elementos geométricos.

a) $M_1\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

b) $M_2\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

c) $M_3 = M_2 \circ M_1$ y $M_4 = M_1 \circ M_2$

**3.** Determinar todos los movimientos que transforman (1, 0) en (2, 1) y (0, 1) en (1, 0), indicando sus elementos geométricos.

**4.** Averiguar todos los movimientos del plano que conmutan con la simetría de eje $x-2y=1$.

**5.** Encontrar todos los movimientos del plano que conmutan con:

a) Traslación de vector $\vec{u}_0$.

b) Giro de ángulo $\pi$ y centro $P_0$.

**6.** Probar que todo movimiento del plano es la composición de tres simetrías si $|T| < 0$ y de dos simetrías si $|T| > 0$.

## 10.4. MOVIMIENTOS EN EL ESPACIO

En esta sección determinaremos todos los posibles movimientos en el espacio. Para llevar a cabo esta determinación puede realizarse un estudio geométrico similar al que se hizo en la sección 10.2 con los movimientos en el plano. Es éste un buen tema de trabajo para el lector o estudiante, ya que de esta manera puede ejercitarse en la intuición geométrica espacial. Sin embargo, también puede realizarse esta determinación de un modo analítico, basado en la clasificación de las aplicaciones ortogonales en espacios euclídeos (teorema 2 de la sección 8.8 del capítulo 8). Este segundo punto de vista será el que seguiremos aquí.



De los resultados de la sección 10.1 se deduce que las ecuaciones de los movimientos $M$ en $\mathbb{R}^3$ son de la forma

$$M(X) = X' = \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + T\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = A + T \cdot X$$

donde $T$ es una matriz ortogonal de orden 3. Del teorema de clasificación de las aplicaciones ortogonales en espacios euclídeos (teorema 2 de la sección 8.8 del capítulo 8) deducimos que, después de un cambio de base ortonormal si es necesario, $T$ puede ser de una de las siguientes formas:

$$\text{(I)}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix} \qquad \text{(II)}\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}$$

$$\text{(III)}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \qquad \text{(IV)}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Observamos que las matrices anteriores son las matrices de una aplicación ortogonal en una base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ que puede ser diferente en cada caso.

Observamos también que la identidad está incluida en el caso (I) con $\varphi = 0$ y que la *simetría central* de matriz

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

es un caso particular de (II) con $\varphi = \pi$.

En los casos (I) y (IV) los movimientos son directos, mientras que en los casos (II) y (III) son inversos.

Para determinar todos los movimientos en $\mathbb{R}^3$ es necesario estudiar sus puntos fijos; como ya se dijo en la sección 10.3 esto equivale a estudiar $r(I-T)$ y $r(I-T \mid A)$. Estudiamos cada uno de los casos por separado.

(I)

$$T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}$$

En este caso:

$$I-T=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1-\cos\varphi & \operatorname{sen}\varphi \\ 0 & -\operatorname{sen}\varphi & 1-\cos\varphi \end{pmatrix}$$

y

$$\begin{vmatrix} 1-\cos\varphi & \operatorname{sen}\varphi \\ -\operatorname{sen}\varphi & 1-\cos\varphi \end{vmatrix}=2(1-\cos\varphi)$$

Si $\varphi=0$, $r(I-T)=0$ y $T$ es la matriz identidad. Si $r(I-T\,|\,A)=0$, $A=(0,0,0)$ y el movimiento es la *identidad;* si $r(I-T\,|\,A)=1$, $A=(a_1, a_2, a_3)\neq(0,0,0)$ y el movimiento es una *traslación* de vector $\vec{a}=(a_1, a_2, a_3)$.

Cuando $\varphi\neq 0$ tenemos que $r(I-T)=2$. Si $r(I-T\,|\,A)=2$ el movimiento $M$ tiene una recta de puntos fijos; si $P$ es un punto de esta recta, las ecuaciones de $M$ en el sistema de referencia $\{P; \vec{u}_1, \vec{u}_2, \vec{u}_3\}$ son

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

y se trata, por tanto, de un *giro de ángulo $\varphi$ alrededor de un eje;* el eje tiene como vector director $\vec{u}_1$ y pasa por el punto $P$ y coincide con la recta de puntos fijos del giro. Observar que en este caso $r(I-T\,|\,A)=2$ implica $a_1=0$.

Si $r(I-T\,|\,A)=3$ el movimiento no posee ningún punto fijo; en este caso se ha de tener necesariamente $a_1\neq 0$ y podemos escribir las ecuaciones de $M$ en la forma

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}+\left[\begin{pmatrix} 0 \\ a_2 \\ a_3 \end{pmatrix}+\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right]$$

con $\vec{a}=(a_1, 0, 0)\neq(0,0,0)$. La expresión entre corchetes son las ecuaciones de un giro, $G_{\varphi,r}$, de ángulo $\varphi$ con respecto a una recta $r$ que tiene a $\vec{u}_1$ como vector director; si escribimos $T_{\vec{a}}$ para denotar la traslación de vector $\vec{a}\neq\vec{0}$ las ecuaciones anteriores nos dicen que

$$M=T_{\vec{a}}\circ G_{\varphi,r}$$

con la particularidad de que $\vec{a}=a_1\vec{u}_1$ es un vector paralelo a la recta $r$. La composición de un giro y una traslación en estas condiciones es lo que hemos llamado un *movimiento helicoidal* (ver el ejemplo H de la sección 10.1).



La descomposición de un movimiento helicoidal como composición de un giro y una traslación no es única; podríamos haber tomado

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} a_1 \\ b \\ c \end{pmatrix}+\left[\begin{pmatrix} 0 \\ a_2-b \\ a_3-c \end{pmatrix}+\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & \operatorname{sen}\varphi \\ 0 & -\operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right],\ b, c\neq 0$$

Sin embargo, existe unicidad si la traslación es paralela al eje de giro y en este caso la descomposición del movimiento helicoidal se dice *canónica.*

(IV)

$$T=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Este es un caso particular de (I) que se obtiene permutando los dos primeros elementos de la base y tomando $\varphi=\pi$. Se trata, por tanto, de un giro de 180° alrededor de un eje o de un movimiento helicoidal. El giro de 180° alrededor de un eje $r$ recibe el nombre de *simetría axial* de eje $r$, ya que todo punto se transforma en su simétrico con respecto a $r$ (ver figura 3).

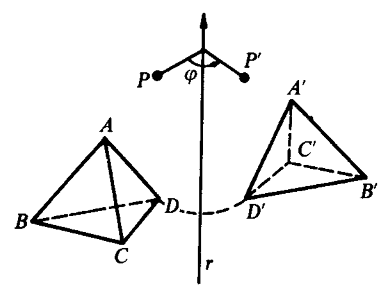



Figura 1. Giro de ángulo $\varphi$ alrededor de $r$ ($G_{\varphi,r}$)

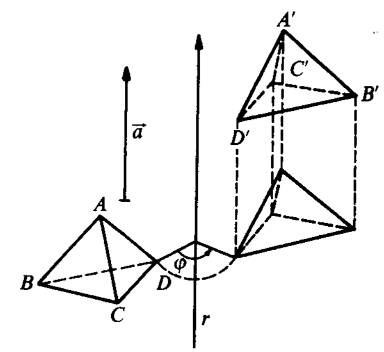



Figura 2. Movimiento helicoidal ($H_{r,\vec{a}}$)

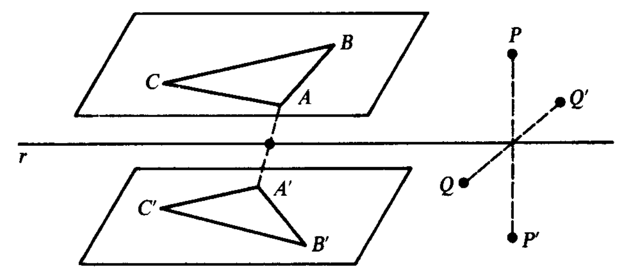



Figura 3. Simetría axial de eje $r$ ($S_r$)

(II)

$$T=\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}$$

En este caso tenemos

$$|I-T|=\begin{vmatrix} 2 & 0 & 0 \\ 0 & 1-\cos\varphi & \operatorname{sen}\varphi \\ 0 & -\operatorname{sen}\varphi & 1-\cos\varphi \end{vmatrix}=4(1-\cos\varphi)$$

Si $\varphi=0$, $r(I-T)=1$ y

$$T=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si $r(I-T\mid A)=1$, el movimiento posee un plano de puntos fijos $\pi$ cuyo vector característico es $\vec{u}_1$. Si tomamos $P$ perteneciente a $\pi$, las ecuaciones de $M$ en el sistema de referencia $\{P; \vec{u}_1, \vec{u}_2, \vec{u}_3\}$ son

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

por tanto, $M$ es una *simetría con respecto a un plano* que pasa por un punto fijo de $M$ y tiene como vector característico $\vec{u}_1$: este plano es el plano de puntos fijos de $M$.

Observar que en este caso $r(I-T\mid A)=1$ implica $a_2=0$, $a_3=0$.

Si $r(I-T\mid A)=2$ (el lector se convencerá fácilmente de que $r(I-T\mid A)\neq 3$ cuando $r(I-T)=1$) se ha de tener $a_2\cdot a_3\neq 0$; por tanto, las ecuaciones de $M$ pueden escribirse de la forma

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \end{pmatrix}=\begin{pmatrix} 0 \\ a_2 \\ a_3 \end{pmatrix}+\left[\begin{pmatrix} a_1 \\ 0 \\ 0 \end{pmatrix}+\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right]$$

con $\vec{a}=a_2\vec{u}_2+a_3\vec{u}_3\neq\vec{0}$. Así pues, $M$ es la composición de una simetría, $S_\pi$, con respecto a un plano $\pi$ de vector característico $\vec{u}_1$, y una traslación de vector $\vec{a}\neq\vec{0}$; es decir,

$$M=T_{\vec{a}}\circ S_\pi$$



Observar que $\vec{a}$ es un vector paralelo a $\pi$. Por analogía con el caso del plano, este movimiento recibe el nombre de *simetría deslizante*. La descomposición $M=T_{\vec{a}}\circ S_\pi$ es única si $\vec{a}$ es paralelo a $\pi$ y recibe el nombre de *descomposición canónica* de la simetría deslizante.

Cuando $\varphi\neq 0$, $r(I-T)=3$ y, por tanto, $r(I-T\mid A)=3$. En este caso el movimiento posee un punto fijo $P$. En el sistema de referencia $\{P; \vec{u}_1, \vec{u}_2, \vec{u}_3\}$ las ecuaciones de $M$ son

$$\begin{pmatrix} y'_1 \\ y'_2 \\ y'_3 \end{pmatrix}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

Esto es la composición de un giro con respecto al eje $\vec{u}_1$ y una simetría con respecto al plano generado por $\vec{u}_2$ y $\vec{u}_3$, ya que

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\varphi & -\operatorname{sen}\varphi \\ 0 & \operatorname{sen} & \cos\varphi \end{pmatrix}$$

Por tanto, $M$ es la composición de un giro con respecto a un eje $r$ que pasa por $P$ y una simetría con respecto a un plano perpendicular a $r$ que pasa por $P$. Este movimiento no recibe ningún nombre especial en general; solamente en el caso $\varphi=180°$, para el cual

$$T=\begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

recibe el nombre de *simetría central* o *simetría con respecto a un punto P* ($P$ recibe el nombre de *centro de la simetría*) ya que todo punto se transforma en su simétrico con respecto a $P$ (ver figura 6). El centro de simetría es el punto fijo del movimiento.

(III)

$$T=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Este es un caso particular de (II) que se obtiene permutando los elementos primero y tercero de la base y tomando $\varphi=0$. Por tanto, el movimiento es una simetría con respecto a un plano o una simetría deslizante.

Esto concluye el estudio de todos los movimientos en el espacio. Los resultados obtenidos se enuncian en el siguiente teorema:

> TEOREMA 1 *(Descripción de todos los movimientos en el espacio)*
>
> Todo movimiento en $\mathbb{R}^3$ es o bien la *identidad*, o una *traslación* o una *rotación alrededor de un eje* o un *movimiento helicoidal* (en estos casos $|T|>0$, el movimiento no cambia la orientación y se dice *directo*) o bien una *simetría con respecto a un plano, una simetría deslizante* o *la composición de un giro y una simetría con respecto a un plano perpendicular al eje de giro* (en estos casos $|T|<0$, el movimiento cambia la orientación y se dice *inverso*).

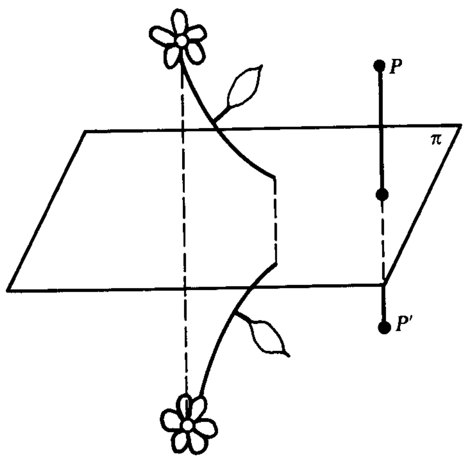



Figura 4. Simetría con respecto a un plano $\pi$ ($S_\pi$)

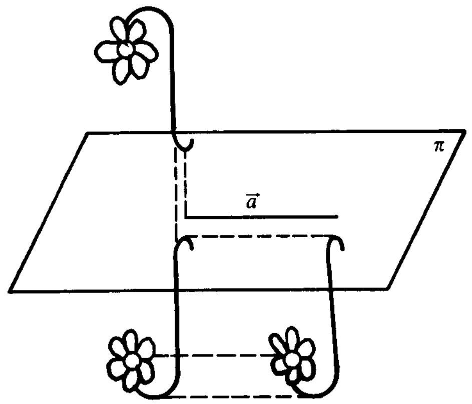



Figura 5. Simetría deslizante ($S_{\pi,\vec{a}}$)

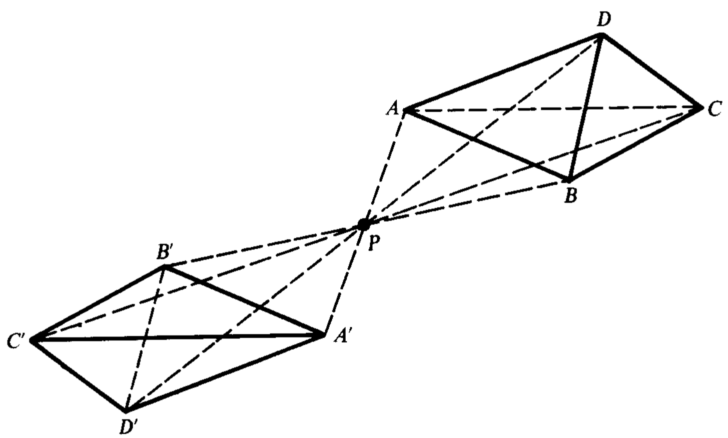



Figura 6. Simetría con respecto al punto $P$



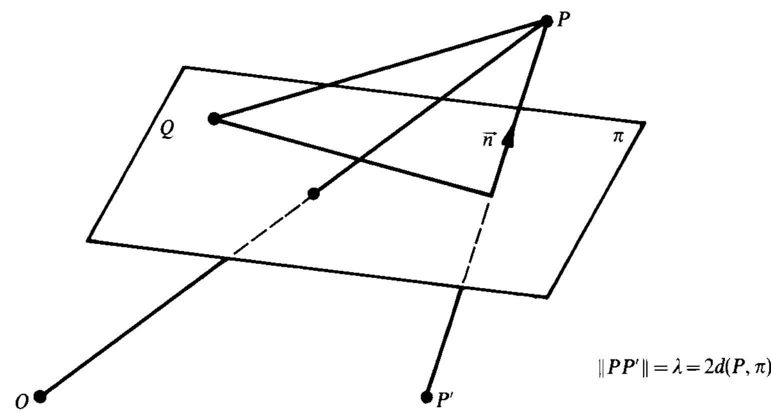
Es conveniente resumir los resultados anteriores en el siguiente cuadro.

| $T$ | $r(I-T)$ | $r(I-T\mid A)$ | Movimiento: $P'=A+T(P)$ |
|---|---|---|---|
| $I$<br>$\|T\|>0$ | 0 | 0<br>1 | Identidad<br>Traslación |
| Simetría respecto de un plano: $\|T\|<0$ | 1 | 1<br>2 | Simetría respecto de un plano de puntos fijos.<br>*Simetría deslizante:* composición de simetría y traslación paralela al plano de simetría. |
| Giro o rotación: $\|T\|>0$ | 2 | 2<br>3 | Giro en torno a a un eje de puntos fijos.<br>*Movimiento helicoidal:* composición de giro y traslación paralela al eje de giro. |
| Simetría ∘ Giro: $\|T\|<0$ | 3 | 3 | Composición de un giro y una simetría; al eje de giro y el plano de simetría son perpendiculares. |

## 10.5. MOVIMIENTOS EN $\mathbb{R}^3$. EJEMPLOS

El objetivo de esta sección es aprender a encontrar las ecuaciones de los movimientos en $\mathbb{R}^3$ así como a determinar los elementos geométricos de un movimiento a partir de sus ecuaciones. En lugar de dar una lista exhaustiva de las ecuaciones de los movimientos y de la forma de determinar sus elementos geométricos daremos varios ejemplos, con la esperanza de que resulten instructivos para el lector.

EJEMPLO A. *Ecuaciones de la simetría respecto de un plano.*



Sea $\pi$ un plano de ecuación $ax+by+cz=d$ y $\vec{n}$ un vector unitario normal a $\pi$, es decir, $\vec{n}=\dfrac{1}{\sqrt{a^2+b^2+c^2}}(a, b, c)$.

La imagen, $P'$, de un punto $P$ es de la forma $P'=P-\lambda\vec{n}$, donde $\lambda=2d(P, r)$. Para calcular $d(P, r)$ tomamos $Q$ un punto cualquiera de $\pi$ y observamos que

$$\cos \sphericalangle (\overrightarrow{QP}, \vec{n})=\frac{d(P, \pi)}{\|\overrightarrow{QP}\|}$$

por tanto,

$$d(P, \pi)=\|\overrightarrow{QP}\| \cos \sphericalangle (\overrightarrow{QP}, \vec{n})=(\overrightarrow{QP}, \vec{n})$$

Así pues, las ecuaciones de la simetría están dadas por

$$P'=P-2(\overrightarrow{QP}, \vec{n})\vec{n}$$

Para escribir esto en coordenadas cartesianas tomamos

$$P=\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \quad P'=\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}, \quad Q=\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix}\in\pi$$

Un sencillo cálculo permite llegar a la conclusión siguiente:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\frac{2d}{a^2+b^2+c^2}\begin{pmatrix} a \\ b \\ c \end{pmatrix}-\frac{2}{a^2+b^2+c^2}\begin{pmatrix} \dfrac{a^2-b^2-c^2}{2} & ba & ca \\ ab & \dfrac{b^2-a^2-c^2}{2} & cb \\ ac & bc & \dfrac{c^2-a^2-b^2}{2} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

que es la expresión en coordenadas cartesianas de la simetría con respecto al plano $\pi$. de la simetría con respecto al plano $\pi$.

* * *



EJEMPLO B. Encontrar las ecuaciones de la simetría con respecto al plano que pasa por (0, 1, 1) *y es perpendicular al vector* (2, 1, 1).

Si $P=\begin{pmatrix} x \\ y \\ z \end{pmatrix}$, $\vec{n}=\dfrac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}$ y $Q=\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$, puesto que $P'=P-2(\overrightarrow{QP}, \vec{n})\,\vec{n}$ tenemos:

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix}=\begin{pmatrix} x \\ y \\ z \end{pmatrix}-2\left(\begin{pmatrix} x \\ y-1 \\ z-1 \end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}\right)\frac{1}{\sqrt{6}}\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}=$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}-\frac{2}{6}[2x+y-1+z-1]\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}=\begin{pmatrix} x-\frac{4}{3}x-\frac{2}{3}y-\frac{2}{3}z+\frac{4}{3} \\ y-\frac{2}{3}x-\frac{1}{3}y-\frac{1}{3}z+\frac{2}{3} \\ z-\frac{2}{3}x-\frac{1}{3}y-\frac{1}{3}z+\frac{2}{3} \end{pmatrix}=$$

$$=\begin{pmatrix} \frac{4}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix}+\begin{pmatrix} -\frac{1}{3} & -\frac{2}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{2}{3} & -\frac{1}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

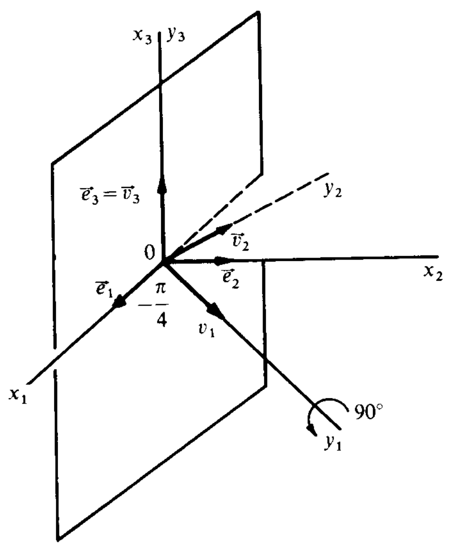
* * *

EJEMPLO C. Determinar las ecuaciones de un giro de ángulo $\pi/2$ con respecto a la recta $r$ que tiene $\vec{v}=(1, 1, 0)$ como vector director y pasa por (0, 0, 0).

Procedemos de la siguiente manera para resolver este problema: encontramos primero un sistema de referencia $\{O; \vec{v}_1, \vec{v}_2, \vec{v}_3\}$, positivamente orientado y ortonormal, con respecto al cual $G_{\pi/2,r}$ tenga una matriz conocida; después se realiza un cambio de sistema de referencia para llevarlo a $\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$.

Tomamos $\vec{v}_1=\dfrac{1}{\sqrt{2}}(1,1,0)$ y $\vec{v}_3=(0, 0, 1)$; elegimos $\vec{v}_2$ de manera que $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ sea positivamente orientada:

$$\vec{v}_2=\vec{v}_3\times\vec{v}_1=\begin{vmatrix} i & j & k \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{vmatrix}\frac{1}{\sqrt{2}}=\frac{1}{\sqrt{2}}(-1, 1, 0)$$




Con respecto al sistema de referencia $\{O; \vec{v}_1, \vec{v}_2, \vec{v}_3\}$ las ecuaciones de $G_{r,\pi/2}$ son

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \quad (*)$$

A continuación hacemos un giro de $-45°$ de manera que el sistema de referencia $\{O; \vec{v}_1, \vec{v}_2, \vec{v}_3\}$ se transforme en el sistema de referencia $\{O; \vec{e}_1, \vec{e}_2, \vec{e}_3\}$: si $\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$ se tiene

$$x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3 = y_1\vec{v}_1 + y_2\vec{v}_2 + y_3\vec{v}_3 = y_1\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + y_2\frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + y_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} =$$

$$= \left(\frac{y_1}{\sqrt{2}} - \frac{y_2}{\sqrt{2}}\right)\vec{e}_1 + \left(\frac{y_1}{\sqrt{2}} + \frac{y_2}{\sqrt{2}}\right)\vec{e}_2 + y_3\vec{e}_3.$$

Por tanto,

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \Rightarrow \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



Similarmente,

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}$$

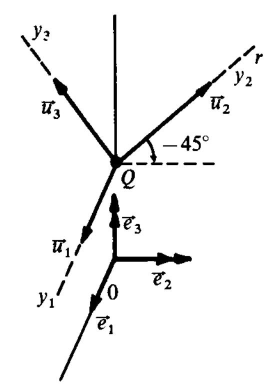
Sustituyendo estos resultados en (*) y realizando el cálculo correspondiente se tiene:

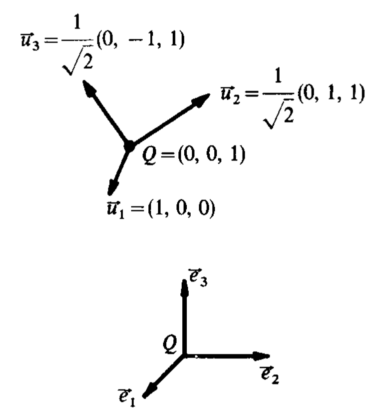
$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

* * *

EJEMPLO D. Hallar las ecuaciones de la composición del giro de ángulo $\pi$ con respecto a la recta $r$: $(0, 0, 1) + t(0, 1, 1)$ con la traslación de vector $\vec{v} = (1, 1, 0)$ y determinar de qué tipo de movimiento se trata.



Eligiendo $\{Q; \overrightarrow{u}_1, \overrightarrow{u}_2, \overrightarrow{u}_3\}$, donde $Q=(0, 0, 1)$, $\overrightarrow{u}_1=(1, 0, 0)$, $\overrightarrow{u}_2=\frac{1}{\sqrt{2}}(0, 1, 1)$, $\overrightarrow{u}_3=u_1 \times u_2=\frac{1}{\sqrt{2}}(0, -1, 1)$ las ecuaciones del giro de ángulo $\pi$ en torno a $r$ son

$$\begin{pmatrix} y_1' \\ y_2' \\ y_3' \end{pmatrix}=G_{\pi,r}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$$

A continuación pasamos al sistema de referencia $\{Q; \overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3\}$ haciendo un giro de $-45°$; si $\overrightarrow{x}=x_1\overrightarrow{e}_1+x_2\overrightarrow{e}_2+x_3\overrightarrow{e}_3$ y $\overrightarrow{x}=y_1\overrightarrow{u}_1+y_2\overrightarrow{u}_2+y_3\overrightarrow{u}_3$ se tiene que

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Por tanto, en el sistema de referencia $\{Q; \overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3\}$ se tiene

$$\begin{pmatrix} z_1' \\ z_2' \\ z_3' \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}=$$

$$=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix}$$

Para pasar ahora al sistema de referencia $\{O; \overrightarrow{e}_1, \overrightarrow{e}_2, \overrightarrow{e}_3\}$ trasladamos $X=(x_1, x_2, x_3)$ y $X'=(x_1', x_2', x_3')$ mediante el vector $(0, 0, 1)$ para obtener

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}+\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3-1 \end{pmatrix}=$$

$$=\begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}+\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$



Al hacer la composición con $T_{\overrightarrow{v}}$, donde $\overrightarrow{v}=(1, 1, 0)$ se tiene

$$M\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}+\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad M=T_{\overrightarrow{v}} \circ G_{\pi,r} \quad (**)$$

que son las ecuaciones buscadas.

*Observación.* Al mismo resultado se habría llegado si observamos que $T_{\overrightarrow{v}} \circ G_{\pi,r}(P)=A+T(P)$, donde

$$T=\begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

ya que $T(\overrightarrow{e}_1)=-\overrightarrow{e}_1$, $T(\overrightarrow{e}_2)=\overrightarrow{e}_3$ y, $T(\overrightarrow{e}_3)=\overrightarrow{e}_2$ y calculamos $A$ teniendo en cuenta que

$$T_{\overrightarrow{v}} \circ G_{\pi,r}(Q)=T_{\overrightarrow{v}} \circ G_{\pi,r}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}=T_{\overrightarrow{v}}\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

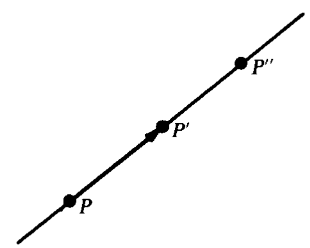
Finalmente, hemos de determinar qué tipo de movimiento es el obtenido en (**). Como

$$I-T=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

$r(I-T)=2$, y como

$$(I-T \mid A)=\begin{pmatrix} 2 & 0 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 1 \end{pmatrix},$$

$r(I-T \mid A)=3$. Del cuadro resumen de la sección 10.4 deducimos que (**) es un *movimiento helicoidal.* Hemos de determinar el eje de giro, el ángulo de giro y el vector de traslación (paralelo al eje de giro).



Si $P$ es un punto del eje de giro, $P'$ es la imagen de $P$ y $P''$ es la imagen de $P'$ se tiene que

$$\overrightarrow{P'P''}=\overrightarrow{M(P)M(P')}=\overrightarrow{PP'}$$

y por tanto:

$$\overrightarrow{PP'}=A+T(P')-A-T(P)=T(\overrightarrow{PP'}) \Rightarrow (I-T)(\overrightarrow{PP'})=0$$

Como $P'=A+T(P)$ se tiene

$$0=(T-I)(P'-P)=(T-I)(A)+(T-I)^2(P).$$

Entonces

$$\begin{pmatrix}0\\0\\0\end{pmatrix}=\begin{pmatrix}-2&0&0\\0&-1&1\\0&1&-1\end{pmatrix}\begin{pmatrix}1\\0\\1\end{pmatrix}+\begin{pmatrix}-2&0&0\\0&-1&1\\0&1&-1\end{pmatrix}^2\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}=$$

$$\begin{pmatrix}-2\\1\\-1\end{pmatrix}+\begin{pmatrix}4&0&0\\0&2&-2\\0&-2&2\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$$

o equivalentemente

$$x_1=\frac{1}{2}\ ,\ 2x_2-2x_3=-1$$

Estas son las ecuaciones del eje de giro. Un punto $P_0$ de este eje de giro es $P_0=\left(\frac{1}{2}, 0, \frac{1}{2}\right)$ y puesto que

$$P_0'=M(P_0)=\begin{pmatrix}1\\0\\1\end{pmatrix}+\begin{pmatrix}-1&0&0\\0&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1/2\\0\\1/2\end{pmatrix}=$$

$$=\begin{pmatrix}1\\0\\1\end{pmatrix}+\begin{pmatrix}-1/2\\1/2\\0\end{pmatrix}=\begin{pmatrix}1/2\\1/2\\1\end{pmatrix}$$

el vector de traslación de este movimiento helicoidal es

$$\overrightarrow{P_0P_0'}=\begin{pmatrix}0\\1/2\\1/2\end{pmatrix}$$



*Observación.* El eje de giro de este movimiento es paralelo al eje dado, ya que sus ecuaciones pueden escribirse de la forma

$$\begin{aligned}x_1&=\frac{1}{2}\\ x_2&=t\\ x_3&=\frac{1}{2}+t\end{aligned} \quad \text{o bien} \quad \begin{pmatrix}1/2\\0\\1/2\end{pmatrix}+t\begin{pmatrix}0\\1\\1\end{pmatrix}$$

Para encontrar el ángulo de giro es necesario utilizar el siguiente resultado:

**LEMA 1**

Si $M=A+T(P)$ es un giro o un movimiento helicoidal, el ángulo de giro $\varphi$ satisface $1+2\cos\varphi=\text{traza}\, T$ (=suma de los elementos en la diagonal principal de $T$).

*Demostración.* Puesto que estamos en el caso (I) de la sección 10.4, podemos escribir

$$C^{-1}TC=\begin{pmatrix}1&0&0\\0&\cos\varphi&\operatorname{sen}\varphi\\0&-\operatorname{sen}\varphi&\cos\varphi\end{pmatrix}$$

donde $C$ es un cambio de base adecuado. Utilizando el resultado del ejercicio 7 al final de esta sección deducimos la igualdad $\text{traza}(T)=\text{traza}(C^{-1}TC)$. De aquí se obtiene el resultado deseado. ■

Aplicando el lema 1 se tiene que $1+2\cos\varphi=-1$, y por tanto, $2\cos\varphi=-2$. De aquí se deduce que $\varphi=180°$.

* * *

EJEMPLO E. Estudiar el movimiento

$$P'=\begin{pmatrix}2\\2\\2\end{pmatrix}+\begin{pmatrix}0&0&-1\\0&-1&0\\1&0&0\end{pmatrix}P$$

Como

$$I-T=\begin{pmatrix}1&0&1\\0&2&0\\-1&0&1\end{pmatrix}\ ,\quad r(I-T)=3$$

y, por tanto, se trata de la *composición de un giro y una simetría* con plano de simetría perpendicular al eje de giro. El punto de intersección del plano de simetría y del eje de giro es el punto fijo de esta transformación. En nuestro caso el punto fijp es

$$P_0=\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

donde

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}+\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix};$$

por tanto,

$$\left.\begin{aligned} x&=2-z \\ y&=2-y \\ z&=2+x \end{aligned}\right\}$$

de donde deducimos que $P_0=(0, 1, 2)$.

Cualquier punto $P=(x, y, z)$ del eje de giro satisface

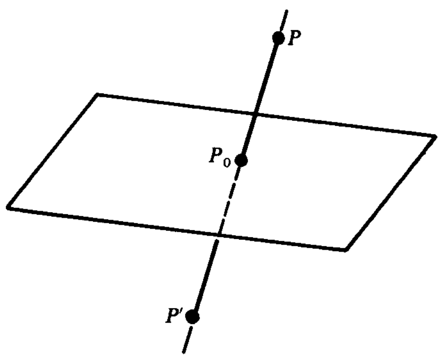
$$P''=P$$

donde $P''$ es la imagen de $P'$ y $P'$ es la imagen de $P$. Por tanto,

$$\overrightarrow{PP'}=-\overrightarrow{P'P''}=-A-T(P')+A+T(P)=-T(\overrightarrow{PP'})$$

Entonces,

$$0=(I+T)(\overrightarrow{PP'})=(I+T)(A+T(P)-P)=(I+T)(A)+$$
$$+(I+T)(T-I)(P)=(I+T)(A)+(T^2-I)(P).$$






En nuestro caso tenemos:

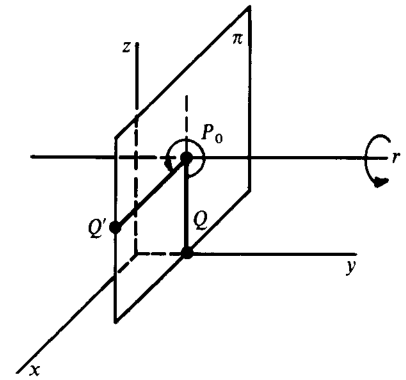
$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}=\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}+\left[\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}^2-\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right]\begin{pmatrix} x \\ y \\ z \end{pmatrix}=$$

$$=\begin{pmatrix} 0 \\ 0 \\ 4 \end{pmatrix}+\begin{pmatrix} -2x \\ 0 \\ -2z \end{pmatrix}\Rightarrow \begin{aligned} x&=0 \\ z&=2 \end{aligned}$$

La ecuaciones paramétricas del eje de giro son

$$r\colon (0, 0, 2)+t(0, 1, 0)$$

Como el plano de simetría ha de ser perpendicular al eje de giro y pasar por el punto fijo $P_0=(0, 1, 2)$, su ecuación es $y=1$.




Con un razonamiento similar al del lema 1 puede obtenerse que el ángulo de giro $\varphi$ satisface

$$-1+2\cos\varphi=\text{traza}\,(T)=-1$$

Por tanto, $\varphi=\pi/2$ o $3\pi/2$. En este caso esta condición no es suficiente para determinar el ángulo de giro. Una forma de determinarlo es tomar un punto $Q$ en el plano de simetría $\pi$, y hallar su imagen $Q'$; si $Q=(0, 1, 0)\in\pi$,

$$Q'=\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}+\begin{pmatrix} 0 & 0 & -1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}=\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}+\begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}=\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

y del gráfico se deduce que $\varphi=\dfrac{3\pi}{2}$.

* * *

EJEMPLO F. Estudiar el movimiento de ecuaciones

$$P'=\begin{pmatrix}-2\\0\\-2\end{pmatrix}+\begin{pmatrix}0&0&-1\\0&1&0\\-1&0&0\end{pmatrix}P$$

Como

$$I-T=\begin{pmatrix}1&0&1\\0&0&0\\1&0&1\end{pmatrix}\quad,\quad r(I-T)=1$$

como

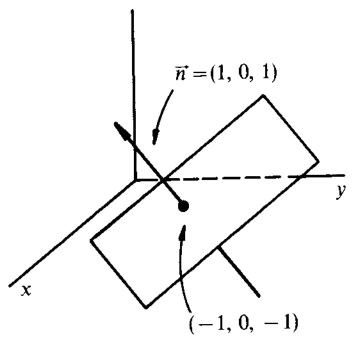
$$(I-T\,|\,A)=\begin{pmatrix}1&0&1&-2\\0&0&0&0\\1&0&1&-2\end{pmatrix}\quad,\quad r(I-T\,|\,A)=1$$

por tanto, es una *simetría*.

El plano de simetría es el conjunto de puntos fijos y, por tanto, satisfacen

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-2\\0\\-2\end{pmatrix}+\begin{pmatrix}0&0&-1\\0&1&0\\-1&0&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-2\\0\\-2\end{pmatrix}+\begin{pmatrix}-z\\+y\\-x\end{pmatrix}$$

Se trata del plano $x+z=-2$, que ha sido dibujado en la figura adjunta.




Todas las rectas perpendiculares al plano de simetría son rectas invariantes; estas rectas tienen por ecuaciones

$$(c_1, c_2, c_3)+t(1, 0, 1)\ ,\ t\in\mathbb{R}$$

con $(c_1, c_2, c_3)$ un punto cualquiera de $\pi$.

* * *



EJEMPLO G. Estudiar el movimiento $M$ de ecuaciones

$$\begin{pmatrix}x'\\y'\\z'\end{pmatrix}=\begin{pmatrix}1\\1\\0\end{pmatrix}-\frac{2}{3}\begin{pmatrix}-\frac{1}{2}&1&1\\1&-\frac{1}{2}&1\\1&1&-\frac{1}{2}\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

Puesto que

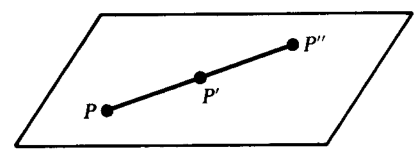
$$I-T=\begin{pmatrix}2/3&2/3&2/3\\2/3&2/3&2/3\\2/3&2/3&2/3\end{pmatrix}\quad\text{y}\quad(I-T\,|\,A)=\begin{pmatrix}2/3&2/3&2/3&1\\2/3&2/3&2/3&1\\2/3&2/3&2/3&0\end{pmatrix}$$

$r(I-T)=1$ y $r(I-T\,|\,A)=2$. El movimiento es una *simetría deslizante*. Los puntos $P$ del plano de simetría satisfacen $\overrightarrow{PP'}=\overrightarrow{P'P''}$, donde $P'=M(P)$ y $P''=M(P')$. Por tanto,

$$\overrightarrow{P'P''}=A+T(P')-A-T(P)=T(\overrightarrow{PP'}).$$

De aquí deducimos que $P$ satisface

$$0=(I-T)(\overrightarrow{PP'})=(I-T)(A+T(P)-P)=$$
$$=(I-T)(A)-(I-T)^2(P)$$






Si $P=(x, y, z)$ esta ecuación se transforma en

$$\begin{pmatrix}0\\0\\0\end{pmatrix}=\begin{pmatrix}2/3&2/3&2/3\\2/3&2/3&2/3\\2/3&2/3&2/3\end{pmatrix}\begin{pmatrix}1\\1\\0\end{pmatrix}-\begin{pmatrix}2/3&2/3&2/3\\2/3&2/3&2/3\\2/3&2/3&2/3\end{pmatrix}^2\begin{pmatrix}x\\y\\z\end{pmatrix}=\frac{4}{3}\begin{pmatrix}1\\1\\1\end{pmatrix}-\frac{4}{3}\begin{pmatrix}x+y+z\\x+y+z\\x+y+z\end{pmatrix}$$

Por tanto, $x+y+z=1$ es la ecuación del plano de simetría $\pi$. Para encontrar el vector de traslación tomamos $P=(1, 1, -1)\in\pi$, con lo cual tenemos

$$P'=\begin{pmatrix}1\\1\\0\end{pmatrix}-\frac{2}{3}\begin{pmatrix}-\frac{1}{2} & 1 & 1\\ 1 & -\frac{1}{2} & 1\\ 1 & 1 & -\frac{1}{2}\end{pmatrix}\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}1\\1\\0\end{pmatrix}-\frac{2}{3}\begin{pmatrix}-\frac{1}{2}\\-\frac{1}{2}\\\frac{5}{2}\end{pmatrix}=\begin{pmatrix}\frac{4}{3}\\\frac{4}{3}\\-\frac{5}{3}\end{pmatrix}$$

El vector de traslación $\vec{a}$ es

$$\vec{a}=\overrightarrow{PP'}=\begin{pmatrix}\frac{4}{3}\\\frac{4}{3}\\-\frac{5}{3}\end{pmatrix}-\begin{pmatrix}1\\1\\-1\end{pmatrix}=\begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\end{pmatrix}=\frac{1}{3}\begin{pmatrix}1\\1\\-2\end{pmatrix}$$

El plano $\pi$ es una subvariedad invariante de este movimiento. Las subvariedades invariantes de dimensión 1 de este movimiento son todas las rectas contenidas en el plano $\pi$ que tienen $\vec{a}=\frac{1}{3}(1, 1, -2)$ como vector director.

## EJERCICIOS (SECCIONES 10.4 Y 10.5)

**1.** Encontrar la expresión analítica de los siguientes movimientos en el espacio:

*a)* Simetría respecto al plano $3x-y+2z=1$.

*b)* Movimiento helicoidal respecto al eje $\{\lambda(1, -1, 0)\}$ con un giro de 180° y vector de traslación $\vec{v}=(2, -2, 0)$.

*c)* Giro cuyo eje pasa por $(1, 1, 0)$ y $(0, 0, 1)$ y que envía $(0, 1, 0)$ en $(1, 1, 1)$.

*d)* Composición de los movimientos *a)* y *b)*.

**2.** Estudiar los siguientes movimientos en el espacio, hallando su tipo, subvariedades invariantes y elementos geométricos:

*a)* $$M_1\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1\\-1\\0\end{pmatrix}+\begin{pmatrix}\frac{1}{2} & \frac{\sqrt{2}}{2} & \frac{1}{2}\\ \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2}\\ \frac{1}{2} & -\frac{\sqrt{2}}{2} & \frac{1}{2}\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$



*b)* $$M_2\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1\\1\\0\end{pmatrix}+\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & -1\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

*c)* $$M_3\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2\\1\\0\end{pmatrix}+\begin{pmatrix}0 & 0 & -1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

*d)* La composición $M_3\circ M_2$ y $M_2\circ M_3$.

**3.** Encontrar los movimientos del espacio que conmutan con la simétrica respecto del plano $z=0$.

**4.** Dado el movimiento helicoidal

$$M\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}2\\1\\0\end{pmatrix}+\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & -1\\ -1 & 0 & 0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}$$

encontrar su descomposición canónica.

**5.** *a)* Dados los planos paralelos $\pi_1$ y $\pi_2$ de ecuaciones $\pi_1$: $x+y=1$ y $\pi_2$: $x+y=4$, encontrar las ecuaciones del movimiento $M=S_{\pi_2}\circ S_{\pi_1}$. Demostrar que $M$ es una traslación y encontrar su vector de traslación.

*b)* Demostrar que para cualesquiera dos planos paralelos $\pi_1$ y $\pi_2$ en $\mathbb{R}^3$, $M=S_{\pi_2}\circ S_{\pi_1}$ es una traslación; encontrar el vector de traslación.

**6.** *a)* Dados los planos $\pi_1$ y $\pi_2$ de ecuaciones $\pi_1$: $z=1$ y $\pi_2$: $x+y-z=2$, encontrar las ecuaciones de $M=S_{\pi_1}\circ S_{\pi_2}$. Demostrar que $M$ es un giro con respecto a un eje y encontrar el eje y el ángulo de giro.

*b)* ¿Qué relación existe entre el eje del giro $M$ y los planos $\pi_1$ y $\pi_2$?

*c)* ¿Qué relación existe entre el ángulo de giro de $M$ y el ángulo que forman $\pi_1$ y $\pi_2$?

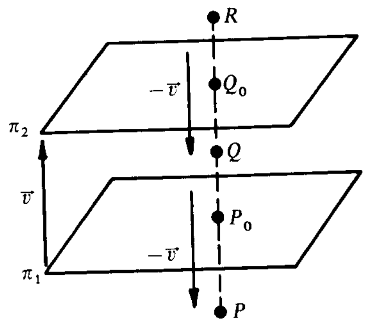
**7.** Sean $T$, $C\in\mathcal{M}_{3\times 3}(\mathbb{K})$ con $|C|\neq 0$. Demostrar que si $J=C^{-1}TC$, traza$(J)=$traza$(T)$. [*Sugerencia:* si $P_T(\lambda)=-\lambda^3+a_2\lambda^2-a_1\lambda+a_0$ es el polinomio característico de $T$, demostrar que $a_2=$traza$(T)$; utilizar a continuación que $P_T(\lambda)$ es invariante mediante un cambio de base (capítulo 7)].

* * *

El resto de los problemas de esta sección están dedicados a la composición de movimientos en el espacio. Están pensados para que se resuelvan utilizando razonamientos geométricos en lugar de razonamientos analíticos. Se da la solución del primero de ellos para que el lector se familiarice con la forma de demostración.

**8.** Demostrar que la composición de dos simetrías con respecto a dos planos paralelos es una traslación. [*Sol.:* Sea $\vec{v}$ un vector perpendicular a uno cualquiera de los planos $\pi_1$ o $\pi_2$ que nos da la distancia entre ellos como muestra la figura adjunta. Basta demostrar que

$$T_{2\vec{v}}^{-1}\circ S_{\pi_2}\circ S_{\pi_1}=I$$




Para todo punto $P$, $T_{2\vec{v}}^{-1}\circ S_{\pi_2}\circ S_{\pi_1}(P)=T_{2\vec{v}}^{-1}(R)$, donde $\vec{PR}\perp\pi_1$ y $\|\vec{PR}\|=\|\vec{PQ}\|+\|\vec{QR}\|$ $=2\|P_0Q\|+2\|QQ_0\|=2\|P_0Q_0\|=2d(\pi_1,\pi_2)=2\|\vec{v}\|$. Entonces

$$T_{2\vec{v}}^{-1}\circ S_{\pi_2}\circ S_{\pi_1}(P)=T_{\vec{v}}^{-1}(R)=P]$$

**9.** Demostrar que la composición de dos simetrías con respecto a planos secantes es un giro alrededor de la recta común de los dos planos y de ángulo doble del que forman los planos.

**10.** Demostrar que la composición de dos simetrías centrales es una traslación.

**11.** Sea $S_r$ una simetría axial de eje $r$ y $T_{\vec{a}}$ una traslación de vector $\vec{a}$ tal que $\vec{a}$ es perpendicular a $r$.

*a*) Demostrar que $T_{\vec{a}}\circ S_r$ es una simetría axial de eje $l$ que es la trasladada de $r$ mediante el vector $\vec{a}/2$.

*b*) Demostrar que $S_r\circ T_{\vec{a}}$ es una simetría axial de eje $l'$ simétrico de $l$ con respecto a $r$.

**12.** *a*) Demostrar que la composición de dos simetrías axiales respecto de ejes secantes es un giro.

*b*) Demostrar que la composición de dos simetrías axiales respecto de ejes que se cruzan puede escribirse como la composición de un giro y una traslación.

## CAPITULO 11

# SECCIONES CONICAS



### 11.1. DEFINICIONES

Un doble cono recto es la figura que engendra una recta $g$ al girar alrededor de una recta $h$ que la corta. La recta $h$ se denomina *eje del cono* y las distintas posiciones de $g$, *generatrices del cono*; el punto de intersección del eje con una cualquiera de las generatrices del cono se denomina *vértice*.

Toda figura (plana) que se obtiene como intersección de un doble cono recto con un plano que le corta se denomina una *sección cónica*.

Según las distintas posiciones del plano de corte las secciones cónicas, o simplemente *cónicas*, reciben nombres diferentes, que se dan a continuación (ver figura 2).

*a*) Si el plano es perpendicular al eje del cono, y no pasa por el vértice, la cónica se denomina una *circunferencia*. En el caso especial de que el plano pase por el vértice se obtiene *un punto*.

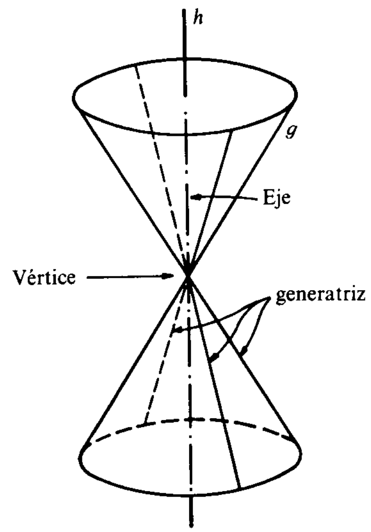
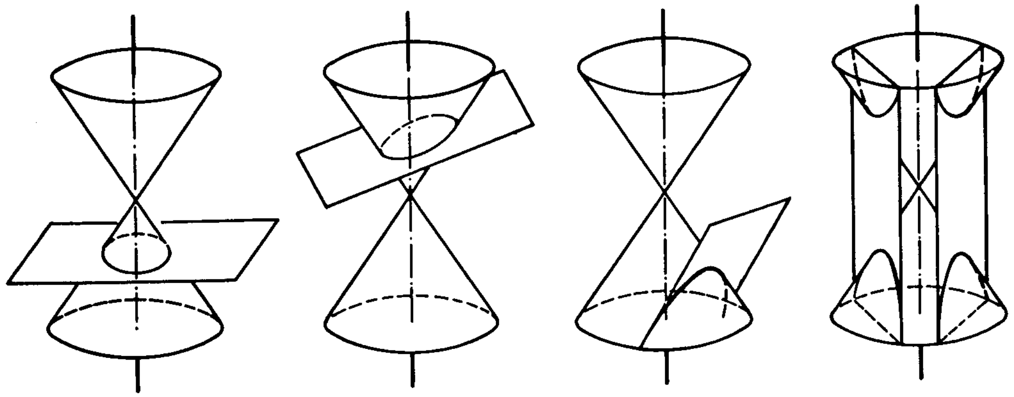



Figura 1



*a*) Circunferencia *b*) Elipse *c*) Parábola *d*) Hipérbola

Figura 2

*b*) Si el plano no es perpendicular al eje del cono y forman entre sí un ángulo superior al ángulo que forman el eje del cono y una cualquiera de las generatrices, la cónica resultante se denomina *elipse*, salvo en el caso especial en que el plano pase por el vértice, en cuyo caso se obtiene un *punto*.

*c*) Si el plano es paralelo a una cualquiera de las generatrices la cónica se denomina *parábola*, excepto si el plano pasa por el vértice, en cuyo caso se obtiene una *recta*.



*d*) Si el ángulo que forman el plano y el eje de giro es inferior al ángulo que forman el eje y una cualquiera de las generatrices, la cónica se denomina *hipérbola*, salvo en el caso especial en que el plano pase por el vértice, en cuyo caso se obtienen *dos rectas que se cortan*.

Los casos especiales que aparecen en los casos anteriores se denominan *secciones cónicas degeneradas* y son puntos, o rectas o pares de rectas que se cortan. De ellos *no* nos ocuparemos aquí por haber sido estudiados anteriormente.

La primera persona que estudió las secciones cónicas fue Menaechmus de Grecia, como consecuencia de su interés en el problema de construir con regla y compás un cubo de volumen doble al de un cierto cubo dado; esto sucedió en el siglo IV antes de Cristo. En el mismo siglo el geómetra Euclides escribió cuatro libros sobre las secciones cónicas, de los cuales ninguno se conserva actualmente.

El primer tratado escrito que se conserva sobre las secciones cónicas es debido a Apolonio de Perga; en sus ocho libros, de los cuales se conservan sólo los siete primeros, Apolonio fue el primero en estudiar las propiedades geométricas de las secciones cónicas. Varias de estas propiedades geométricas serán estudiadas en las secciones siguientes.

## 11.2. LA CIRCUNFERENCIA Y ALGUNA DE SUS PROPIEDADES

En la circunferencia la distancia del vértice $V$ a un punto cualquiera $P$ de la circunferencia es constante; como $\overrightarrow{VC}$ es fijo, donde $C$ es el punto de intersección del plano y el eje del cono (figura 3), se tiene que

$$\|\overrightarrow{CP}\| = \sqrt{\|VP\|^2 - \|VC\|^2}$$

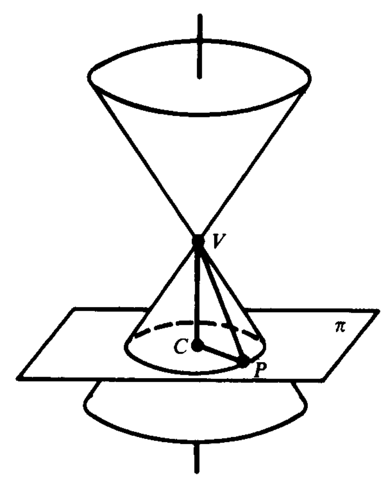



Figura 3

es constante para todo $P$ de la circunferencia. Tenemos entonces

PROPOSICIÓN 1

Una *circunferencia* es el conjunto de puntos $P$ de un plano $\pi$ que satisfacen que su distancia a un punto fijo $C$, llamado *centro*, es constante. Esta constante se denomina *radio de la circunferencia.*

PROPOSICIÓN 2

Sean $A$ y $B$ dos puntos de un plano. Un punto $P$ del plano pertenece a la circunferencia en la que $A$ y $B$ son diametralmente opuestos si y sólo si los vectores $\overrightarrow{AP}$ y $\overrightarrow{BP}$ son perpendiculares.

*Demostración.* Sea $r$ el radio de la circunferencia y $C$ su centro. El plano en el que estamos trabajando es «isomorfo» a un plano euclídeo y, por tanto, podemos suponer que en él se tiene el producto escalar usual, denotado por ( , ). Entonces tenemos (figura 4):

$$\begin{aligned} r^2 &= \|\overrightarrow{CP}\|^2=(\overrightarrow{CP}, \overrightarrow{CP})=(\overrightarrow{CA}+\overrightarrow{AP}, \overrightarrow{CB}+\overrightarrow{BP})= \\ &=(\overrightarrow{CA}, \overrightarrow{CB})+(\overrightarrow{CA}, \overrightarrow{BP})+(\overrightarrow{AP}, \overrightarrow{CB})+(\overrightarrow{AP}, \overrightarrow{BP})= \\ &= -r^2+(\overrightarrow{CA}, \overrightarrow{BP})-(\overrightarrow{CA}, \overrightarrow{AP})+(\overrightarrow{AP}, \overrightarrow{BP})= \\ &= -r^2+(\overrightarrow{CA}, \overrightarrow{BP}-\overrightarrow{AP})+(\overrightarrow{AP}, \overrightarrow{BP})= \\ &= -r^2+(\overrightarrow{CA}, \overrightarrow{BA})+(\overrightarrow{AP}, \overrightarrow{BP})= \\ &= -r^2+2r^2+(\overrightarrow{AP}, \overrightarrow{BP})=r^2+(\overrightarrow{AP}, \overrightarrow{BP}) \end{aligned}$$

Esto es equivalente a $(\overrightarrow{AP}, \overrightarrow{BP})=0$, lo cual demuestra el resultado. ■

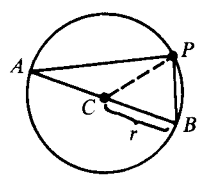
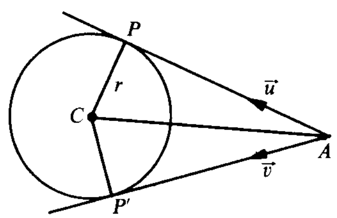



Figura 4

PROPOSICIÓN 3

Consideremos una circunferencia de centro $C$ y un punto $A$ exterior a ella. Si $P$ y $P'$ son los puntos de intersección de las tangentes a la circunferencia que pasan por $A$ se tiene que $\|\overrightarrow{AP}\|=\|\overrightarrow{AP'}\|$. Además, $\overrightarrow{CP}$ es perpendicular a $\overrightarrow{AP}$.






*Demostración.* Sea $r$ el radio de la circunferencia y $\vec{u}$ un vector unitario en la dirección de $\overrightarrow{AP}$. Para $\lambda=\|\overrightarrow{AP}\|$ se tiene que

$$\begin{aligned} r^2 &= \|\overrightarrow{CP}\|^2=\|\overrightarrow{CA}+\overrightarrow{AP}\|^2=\|\overrightarrow{CA}+\lambda\vec{u}\|^2= \\ &=\lambda^2+2\lambda(\overrightarrow{CA}, \vec{u})+\|\overrightarrow{CA}\|^2. \end{aligned}$$

Si consideramos la ecuación en $\lambda$, $\lambda^2+2\lambda(\overrightarrow{CA}, \vec{u})+\|\overrightarrow{CA}\|^2-r^2=0$, ésta ha de tener una solución únicamente, ya que $\overrightarrow{AP}$ es tangente a la circunferencia, y esta solución es $\lambda=\|\overrightarrow{AP}\|$. La ecuación de segundo grado tiene solución única si y sólo si

$$(\overrightarrow{CA}, \vec{u})^2=\|\overrightarrow{CA}\|^2-r^2 \tag{1}$$

y esta solución es

$$\|\overrightarrow{AP}\|=\lambda=-(\overrightarrow{CA}, \vec{u}) \tag{2}$$

Si $\vec{v}$ es un vector unitario en la dirección de $\overrightarrow{AP'}$ el mismo razonamiento anterior produce

$$(\overrightarrow{CA},\vec{v})^2=\|\overrightarrow{CA}\|^2-r^2 \tag{3}$$

y

$$\|\overrightarrow{AP'}\|=-(\overrightarrow{CA}, \vec{v}) \tag{4}$$

De (1), (2), (3) y (4) se deduce que $\|\overrightarrow{AP}\|=\|\overrightarrow{AP'}\|$. Además, de (1) y (2) obtenemos $\|\overrightarrow{AP}\|^2=\|\overrightarrow{CA}\|^2-\|\overrightarrow{CP}\|^2$ y, por tanto, el triángulo $CPA$ satisface el teorema de Pitágoras. Esto prueba que $CPA$ tiene un ángulo recto en $P$, lo cual demuestra la proposición. ■

## 11.3. LA ELIPSE Y LA HIPERBOLA

PROPOSICIÓN 1

Una elipse es el lugar geométrico de los puntos $A$ de un plano cuya suma de las distancias de $A$ a dos puntos fijos y distintos, llamados focos, $F_1$ y $F_2$, es constante ($\|\overrightarrow{AF}_1\|+\|\overrightarrow{AF}_2\|=$constante).

*Demostración.* Sea $\pi$ el plano dado que forma con el eje un ángulo mayor que $\alpha$, sin ser 90°, Considerar las dos esferas tangentes al cono y al plano como muestra la figura 5. Estas esferas tocan al plano $\pi$ en los puntos $F_1$ y $F_2$. Por la proposición 3 de la sección 11.2:

$$\|\overrightarrow{AF}_2\| = \|\overrightarrow{AN}\|$$

y

$$\|\overrightarrow{AF}_1\| = \|\overrightarrow{AM}\|$$

Como $\|\overrightarrow{AN}\| + \|\overrightarrow{AM}\| = \|\overrightarrow{JL}\|$, que es fija, se tiene que

$$\|\overrightarrow{AF}_1\| + \|\overrightarrow{AF}_2\| = \|\overrightarrow{JL}\| = \text{constante}$$ ■

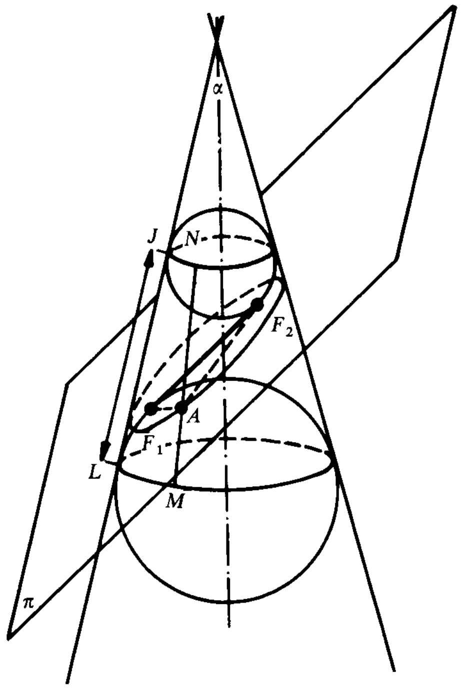


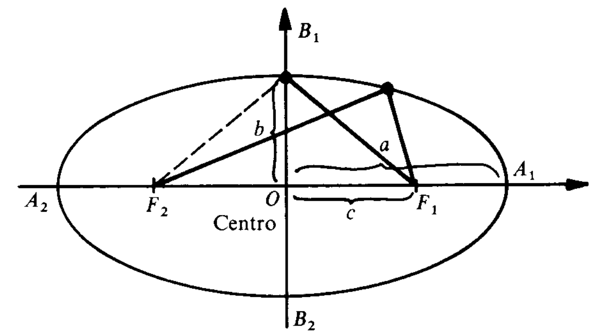

Figura 5

Si dibujamos la elipse en un sistema de coordenadas cartesianas, como en la figura 6, el eje que contiene a los focos se denomina *eje principal* y el eje perpendicular al eje mayor que pasa por el punto medio del segmento $\overline{F_1F_2}$ se denomina *eje secundario*.

Si $a$, $b$ y $c$ son las distancias indicadas en la figura 6, $2c$ se denomina *distancia focal*, $a$ se denomina *semieje principal* y $b$ se denomina *semieje secundario*. Al cociente $\varepsilon = c/a$, entre $c$ y $a$, se le denomina *excentricidad* y como $a > c$, se tiene que $0 < \varepsilon < 1$ en el caso de la elipse. Entre $a$, $b$ y $\varepsilon$ existe una relación que puede encontrarse aplicando el teorema de Pitágoras a $F_1OB_1$:

$$a^2 = \|B_1F_1\|^2 = b^2 + c^2 \Rightarrow b = \sqrt{a^2 - c^2} = \sqrt{-\varepsilon^2 a^2 + a^2} = a\sqrt{1-\varepsilon^2}$$






$A_1$, $A_2$, $B_1$, $B_2$ son vértices

Figura 6

Quizá es conveniente explicar la igualdad $a = \|B_1F_1\|$; esta igualdad es consecuencia del siguiente lema cuando $A = B_1$.

LEMA 2

Para todo punto $A$ de una elipse se tiene que $\|\overrightarrow{AF}_1\| + \|\overrightarrow{AF}_2\| = 2a$.

*Demostración.* Utilizando la definición de elipse $\|\overrightarrow{AF}_1\| + \|\overrightarrow{AF}_2\| = \text{constante}$; si en lugar de $A$ ponemos $A_1$ y $A_2$, se obtiene

$$\|\overrightarrow{A_1F}_1\| + \|\overrightarrow{A_1F}_2\| = \text{constante} = \|\overrightarrow{A_2F}_1\| + \|\overrightarrow{A_2F}_2\|.$$

Por tanto,

$$\|\overrightarrow{A_1F}_1\| + \|\overrightarrow{A_1F}_1\| + \|\overrightarrow{F_1F}_2\| = \|\overrightarrow{A_2F}_2\| + \|\overrightarrow{F_2F}_1\| + \|\overrightarrow{A_2F}_2\|.$$

de aquí deducimos $2\|\overrightarrow{A_1F}_1\| = 2\|\overrightarrow{A_2F}_2\|$ o equivalentemente, $\|\overrightarrow{A_1F}_1\| = \|\overrightarrow{A_2F}_2\|$. Utilizando este resultado se tiene

$$\|\overrightarrow{AF}_1\| + \|\overrightarrow{AF}_2\| = \text{constante} = \|\overrightarrow{A_1F}_1\| + \|\overrightarrow{A_1F}_2\| =$$
$$= \|\overrightarrow{A_2F}_2\| + \|\overrightarrow{A_1F}_2\| = \|\overrightarrow{A_1A}_2\| = 2a$$ ■

* * *

PROPOSICIÓN 3

Una hipérbola es el lugar geométrico de los puntos $P$ de un plano cuya diferencia de distancias a dos puntos fijos y distintos, $F_1$ y $F_2$, llamados focos, es constante en valor absoluto ($\|\overrightarrow{PF}_1\| - \|\overrightarrow{PF}_2\| = \pm 2a$, con $a$ constante).

*Demostración.* Considerar las dos esferas tangentes al cono y al plano $\pi$ como muestra la figura; sean $F_1$ y $F_2$ los puntos de contacto del plano y las esferas. Por la proposición 3 de la sección 11.2

$$\|\overrightarrow{PF}_1\| = \|\overrightarrow{PP'}\|$$

y

$$\|\overrightarrow{PF}_2\| = \|\overrightarrow{PP''}\|$$

Como $\|\overrightarrow{PP''}\| - \|\overrightarrow{PP'}\| = \|\overrightarrow{P'P''}\| = \|\overrightarrow{JL}\| = \text{constante}$, se tiene que

$$\|\overrightarrow{PF}_2\| - \|\overrightarrow{PF}_1\| = \text{constante} \qquad \blacksquare$$

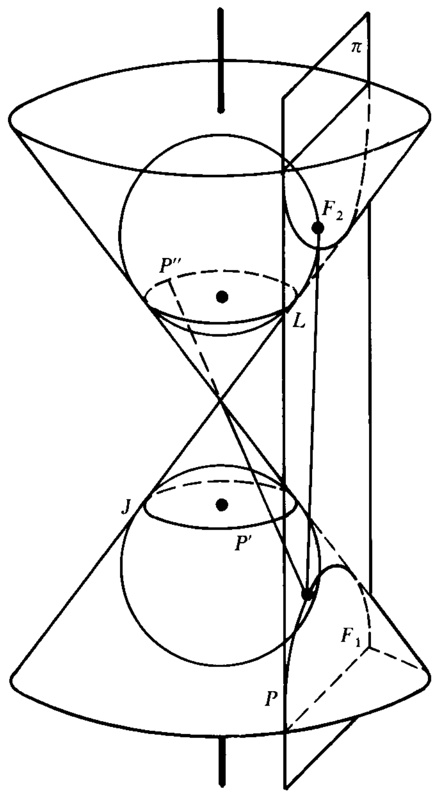



Figura 7

En la figura 8 se ha dibujado una hipérbola en un sistema de coordenadas cartesianas de manera que el *eje principal* contiene a los focos $F_1$, $F_2$ y el *eje secundario* es perpendicular al eje principal y divide al segmento $\overline{F_1F_2}$ en dos partes iguales. Los demás elementos de la hipérbola reciben los mismos nombres que en el caso de la elipse y están señalados en la figura 8.



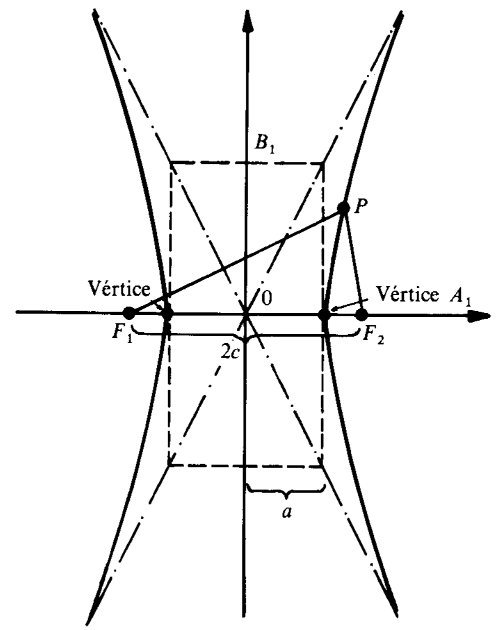
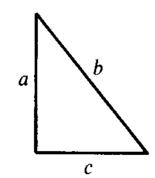



Figura 8

Una demostración análoga a la del lema 2 prueba que $\|\overrightarrow{PF}_1\| - \|\overrightarrow{PF}_2\| = \pm a$ ($+a$ si $P$ está en la rama de la derecha y $-a$ si $P$ está en la rama de la izquierda).

Para una hipérbola la excentricidad es $\varepsilon = c/a$ con $c > a$, y, por tanto, $\varepsilon > 1$. Por analogía con el caso de la elipse definimos $b$ como

$$b = \sqrt{c^2 - a^2}$$




es decir, el cateto de un triángulo rectángulo que tiene a $c$ como la hipotenusa y a $a$ como el otro cateto. Entonces

$$b = \sqrt{a^2\varepsilon^2 - a^2} = a\sqrt{\varepsilon^2 - 1}$$

El parámetro $b$ tiene un significado geométrico que está relacionado con las *asíntotas* a cada una de las ramas de la hipérbola. Recordamos que se denominan *asíntotas* a la hipérbola (en general a una curva plana) a las rectas que se acercan a la hipérbola (curva) en los puntos del infinito. Demostraremos que:

*las rectas que pasan por el centro y tienen $(a, b)$ o $(a, -b)$ como vector director son asíntotas de la hipérbola.*

La demostración que damos de este hecho es sólo aproximada, por cuanto no utilizaremos la «correcta» definición de límite.

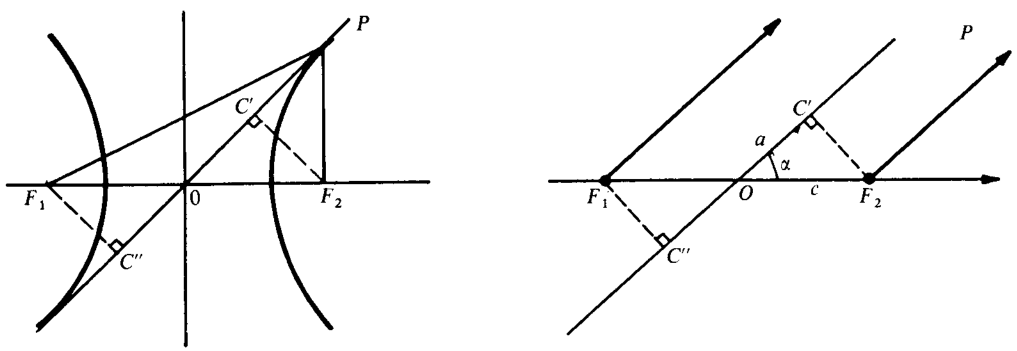



Figura 9

Si $C'$ y $C''$ son como en la figura 9 (derecha) se tiene que cuando $P$ va hacia el infinito,

$$\|C'C''\| \sim \|PF_1\| - \|PF_2\| = 2a$$

y, por tanto, $a \sim \|OC'\|$. En el triángulo rectángulo $OC'F_2$ se tiene $\cos\alpha = 1/\varepsilon$ y, por tanto, $\operatorname{tg}\alpha = \pm\sqrt{\sec^2\alpha - 1} = \pm\sqrt{\varepsilon^2 - 1} = \pm b/a$.

## 11.4. NUEVA DEFINICION DE LAS SECCIONES CANONICAS: LA ELIPSE, LA HIPERBOLA Y LA PARABOLA

PROPOSICIÓN 1

Dada una cónica no degenerada existe siempre un punto $F$, llamado *foco*, una recta $l$, llamada *directriz* (ambos en el plano de la cónica) y un número $\varepsilon > 0$ tal que todo punto $P$ de la cónica satisface

$$\|\vec{PF}\| = \varepsilon d(P, l)$$

Si $\varepsilon < 1$ se tiene una *elipse*, si $\varepsilon > 1$ se obtiene una *hipérbola* y si $\varepsilon = 1$ se tiene una *parábola*.

En la figura 10 se representan cónicas con el mismo foco $F$ y la misma directriz $l$, pero variando la excentricidad.



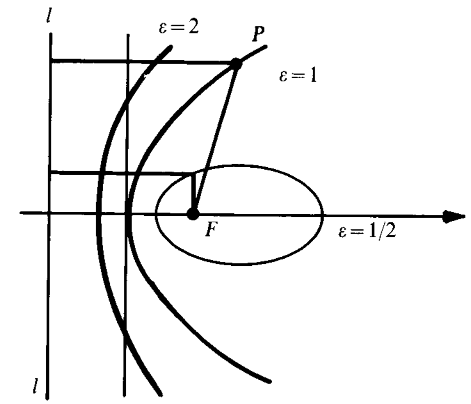



Figura 10

De acuerdo con la proposición 1 podemos decir que una *cónica de foco $F$, directriz $l$* y *excentricidad* $\varepsilon > 0$ es el conjunto de todos los puntos $P$ de un plano que verifican

$$\|\vec{FP}\| = \varepsilon d(P, l)$$

*Demostración.* Sea $\pi$ el plano cuya sección produce la cónica al cortarse con el doble cono recto y considerar la esfera inscrita en el cono y tangente al plano $\pi$ como en la figura 11.

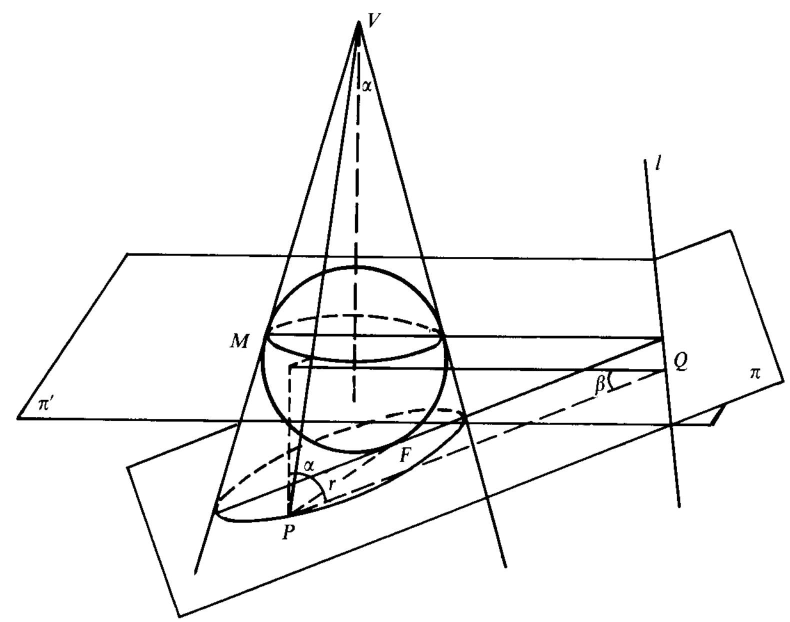



Figura 11

(Observar que en la figura 11 se ha representado el caso $\gamma > \alpha$ (o bien $\beta < \alpha$), que corresponde a una elipse, pero el razonamiento sería el mismo si $\gamma = \alpha = \beta$ —parábola— o $\gamma < \alpha < \beta$ —hipérbola—). Con los símbolos de la figura 11 se tiene que

$$\|\overrightarrow{PF}\| = \|\overrightarrow{PA}\|$$

debido a la proposición 3 de la sección 11.2. Entonces,

$$\frac{\|\overrightarrow{PF}\|}{\|\overrightarrow{PQ}\|} = \frac{\|\overrightarrow{PA}\|}{\|\overrightarrow{PQ}\|} = \frac{\|\overrightarrow{PM}\|/\cos\alpha}{\|\overrightarrow{PM}\|/\cos\gamma} = \frac{\cos\gamma}{\cos\alpha} = \varepsilon$$

y, por tanto, $\|\overrightarrow{PF}\| = \varepsilon\|\overrightarrow{PQ}\| = \varepsilon d(P, l)$.

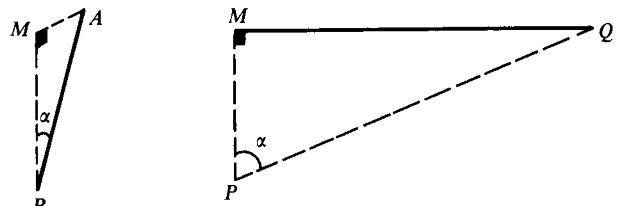



Figura 12

En el caso de la elipse, $\gamma > \alpha$ y, por tanto, $\cos\gamma < \cos\alpha$; se tiene, pues, que $\varepsilon < 1$. Si $\gamma = \alpha$, $\varepsilon = 1$ y se tiene una parábola. Si $\gamma < \alpha$, $\varepsilon = \cos\gamma/\cos\alpha > 1$ y se tiene una hipérbola. Observar que $\gamma$ y $\alpha$ son constantes. ■

* * *

Damos a continuación algunas propiedades más de las elipses, las hipérbolas y las parábolas.

**PROPOSICIÓN 2**

La tangente y la normal a una elipse en un punto $P$ de ella misma son las bisectrices de las rectas que unen el punto $P$ con los focos.

*Demostración.* Dados los focos $F_1$, $F_2$ y un punto $P$ de una elipse, sea $l$ la bisectriz de las rectas $PF_1$ y $PF_2$ como indica la figura 13.

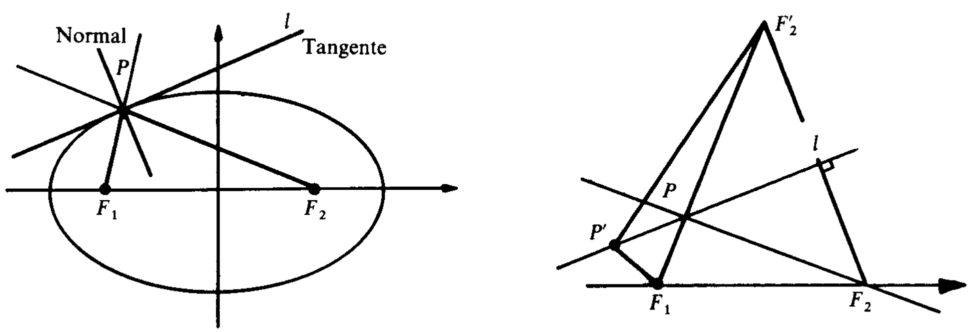



Figura 13

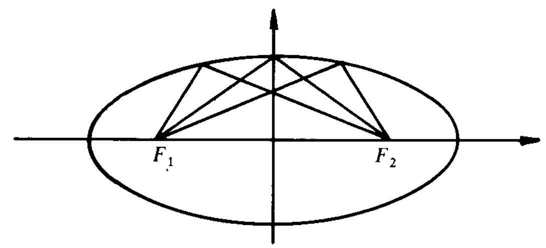


Basta demostrar que cualquier otro punto $P'$ de $l$ no puede estar en el interior de la elipse, puesto que entonces sólo toca a la elipse en el punto $P$ y es la tangente. De la figura 13 deducimos que

$$\|\overrightarrow{F_1P}\| + \|\overrightarrow{PF_2}\| = \|F_1P\| + \|PF_2'\| = \|F_1F_2'\| < \|F_1P'\| + \|P'F_2'\|,$$

en donde la última desigualdad es debida a la desigualdad triangular. Entonces $\|F_1P'\| + \|P'F_2\| > 2a$ y $P'$ está siempre fuera de la elipse, que era lo que queríamos demostrar. La otra bisectriz es precisamente la normal a la elipse en el punto $P$. ■

La proposición 2 expresa una propiedad de las elipses que puede considerarse como «el secreto del salón ovalado» (ver *Cuentos con cuentas*, M. de Guzmán, Editorial Labor). En un salón en forma de elipse, el sonido emitido por una persona colocada en uno de sus focos $F_1$ se refleja en sus paredes pasando siempre por el otro foco $F_2$, de manera que un sonido en $F_1$ puede ser perfectamente audible en $F_2$ e irreconocible en cualquier otra parte del salón.




Para la hipérbola se tiene una proposición semejante a la 2, que se deja como ejercicio.

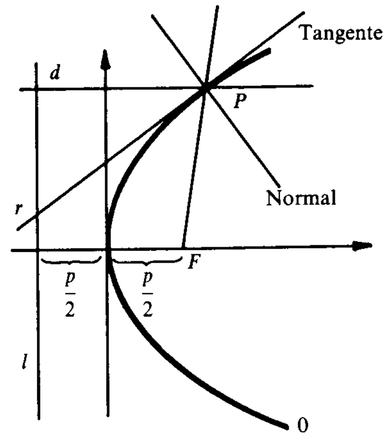
**PROPOSICIÓN 3**

La tangente y la normal a una hipérbola en un punto $P$ de ella misma son las bisectrices de las rectas que unen $P$ con cada uno de los focos.

Para la tangnte y la normal a una parábola se tiene la siguiente proposición.

**PROPOSICIÓN 4**

La tangente y la normal a una parábola en un punto $P$ son las bisectrices de la normal a la directriz por $P$ y la recta que une el punto con el foco $F$.




*Demostración.* Si $r$ es la bisectriz de las rectas $d$ y $PF$ y $P'$ es cualquier otro punto de $l$, de la figura 14 se deduce que

$$\|P'F\| = \|P'A\| > d(P', l)$$

y, por tanto, $P'$ está en el exterior de la parábola. ■

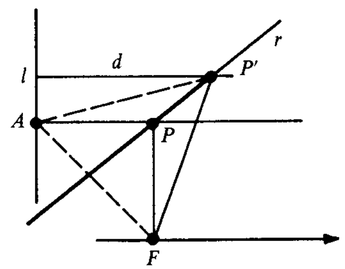
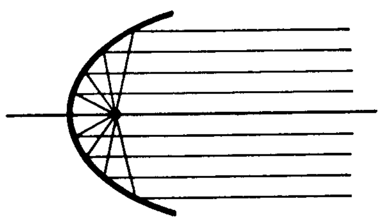



Figura 14

La proposición 4 expresa la propiedad de los espejos parabólicos de que todo rayo de luz procedente del infinito se refleja en su superficie interior pasando por el foco. Recíprocamente, cualquier foco de luz situado en el foco de un espejo parabólico (de revolución) produce un haz de rayos paralelos. Los faros de los coches tienen una forma que se aproxima a la de los espejos parabólicos.



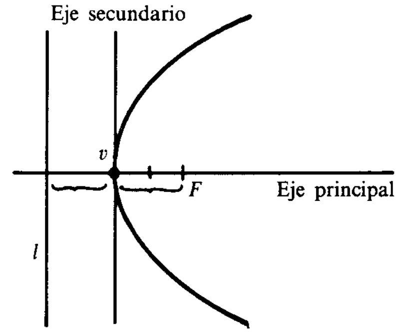


Los elementos principales de una parábola se describen en la figura adjunta. Observar que

$$2d(F, V) = d(F, l)$$

ya que para el punto $V$ se tiene que

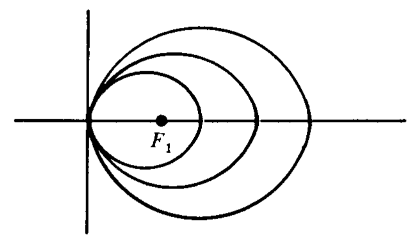
$$d(F, V) = d(V, l)$$



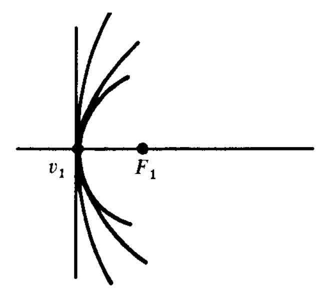

$v$ = vértice $F$ = foco $l$ = directriz

EJERCICIOS (SECCIONES 11.2, 11.3 Y 11.4)

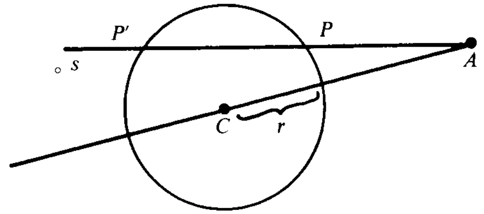
**1.** La figura muestra lo que ocurre si fijamos un vértice $V_1$, un foco contiguo $F_1$ de una elipse y variamos $\varepsilon$ (excentricidad). Hallar, en función de $\varepsilon$ ($0 \leqq \varepsilon < 1$), la posición del centro, del otro vértice y la longitud del semieje $b$.




**2.** La misma pregunta que en el problema anterior para una familia de hipérbolas, $1 < \varepsilon < \infty$, calculando la posición de las asíntotas.



**3.** Dada una circunferencia de centro $C$ y radio $r$ y un punto $A$ exterior a ella, se considera cualquier recta $s$ que pase por $A$ y corte a la circunferencia como en la figura. Demostrar que $\|\overrightarrow{AP}\|\,\|\overrightarrow{AP'}\| = \|\overrightarrow{AC}\| - r^2$.




*Nota.* $\|\overrightarrow{AP}\|\,\|\overrightarrow{AP'}\|$ se denomina *potencia* del punto $A$ respecto de la circunferencia dada. El ejercicio 1 demuestra que $\|\overrightarrow{AP}\|\,\|\overrightarrow{AP'}\|$ no depende de la recta $s$.

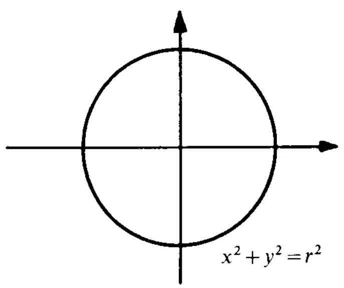
## 11.5. ECUACIONES DE LAS CONICAS EN UN SISTEMA DE COORDENADAS CARTESIANO

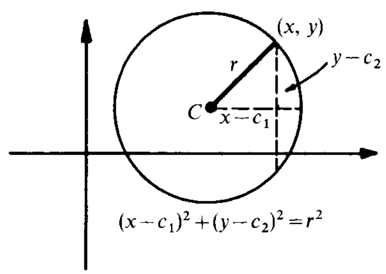
Es conocido que la ecuación de una circunferencia de centro el origen y radio $r$ es

$$x^2+y^2=r^2 \tag{1}$$

mientras que si su centro está en el punto $C=(c_1, c_2)$ su ecuación es

$$(x-c_1)^2+(y-c_2)^2=r^2$$

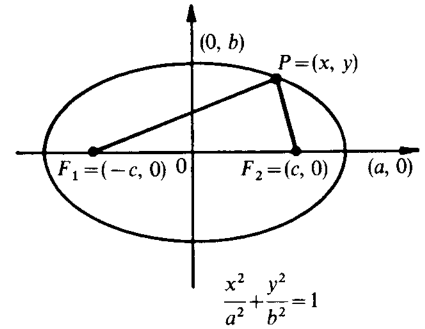



Tratemos de hallar ahora la ecuación de una elipse de semieje principal (o mayor) $a$ y los focos situados en los puntos $(c, 0)$ y $(-c, 0)$ con $a>c$. Utilizando la proposición 1 de la sección 11.3 tenemos

$$2a=\|\overrightarrow{PF}_1\|+\|\overrightarrow{PF}_2\|=\sqrt{(x+c)^2+y^2}+\sqrt{(x-c)^2+y^2}$$






Poniendo la primera de las raíces en el primer término y elevando el cuadrado se tiene

$$4a^2+x^2+2cx+c^2+y^2-4a\sqrt{(x+c)^2+y^2}=x^2-2cx+c^2+y^2$$

que una vez simplificado nos permite obtener

$$a^2+cx=a\sqrt{(x+c)^2+y^2}$$

Elevando de nuevo al cuadrado se obtiene

$$a^4+2cxa^2+c^2x^2=a^2x^2+2a^2cx+a^2c^2+a^2y^2.$$

Teniendo en cuenta que $b^2=a^2-c^2$ obtenemos

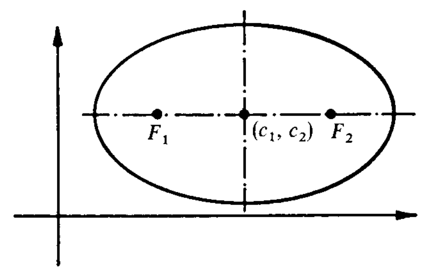
$$a^2b^2=b^2x^2+a^2y^2.$$

que puede escribirse de la forma

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \tag{2}$$

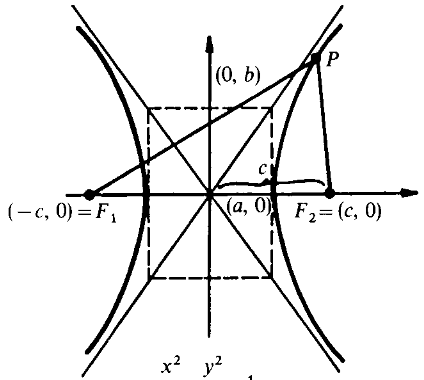
Si la elipse tiene su centro en el punto $(c_1, c_2)$ y sus ejes son paralelos a los ejes coordenados su ecuación es

$$\frac{(x-c_1)^2}{a^2}+\frac{(y-c_2)^2}{b^2}=1.$$



Para hallar la ecuación de una hipérbola de semieje mayor $a$ y focos en los puntos $(-c, 0)$ y $(c, 0)$ se procede como en el caso anterior, mediante la utilización de la proposición 2 de la sección 11.3:

$$\pm 2a = \|PF_2\| - \|PF_1\| = \sqrt{(x-c)^2+y^2} - \sqrt{(x+c)^2+y^2}$$




$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$$

Por tanto,

$$4a^2 \pm 4a\sqrt{(x+c)^2+y^2} + (x+c)^2+y^2 = (x-c)^2+y^2$$

que simplificando nos permite obtener

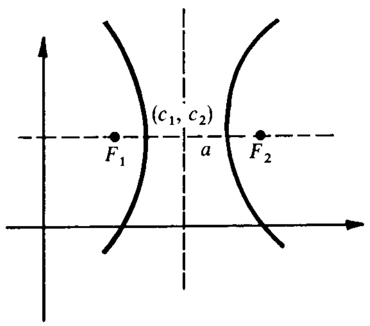
$$a^2+xc = \mp a\sqrt{(x+c)^2+y^2}.$$

Elevando de nuevo al cuadrado y teniendo en cuenta que $b^2=c^2-a^2$ se obtiene

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \tag{3}$$

Si la elipse tiene su centro en el punto $(c_1, c_2)$ y sus ejes son paralelos a los ejes coordenados su ecuación es

$$\frac{(x-c_1)^2}{a^2}+\frac{(y-c_2)^2}{b^2}=1$$




* * *

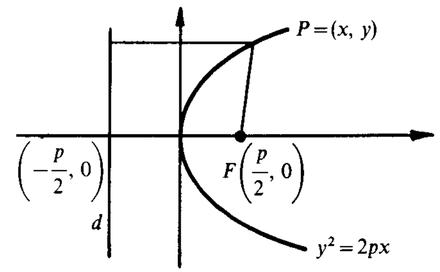


*Observación.* Si la elipse tiene los focos en el eje $OY$, su ecuación es $\dfrac{y^2}{a^2}+\dfrac{x^2}{b^2}=1$ con $a>b$, y si se trata de una hipérbola su ecuación es $\dfrac{y^2}{a^2}-\dfrac{x^2}{b^2}=1$.

* * *

La ecuación de una parábola de directriz $x=-\dfrac{p}{2}$ y foco en el punto $\left(\dfrac{p}{2}, 0\right)$ se obtiene de la proposición 1 de la sección 11.4:

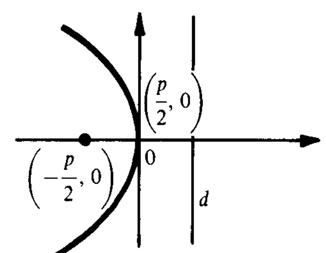
$$\|\vec{PF}\| = \sqrt{\left(x-\frac{p}{2}\right)^2+y^2} = d(P, d) = x+\frac{p}{2}$$



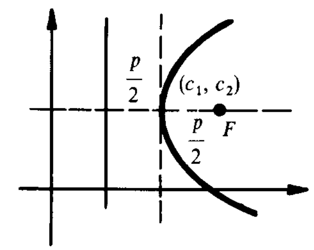

Elevando al cuadrado y simplificando se obtiene

$$y^2=2px \tag{4}$$

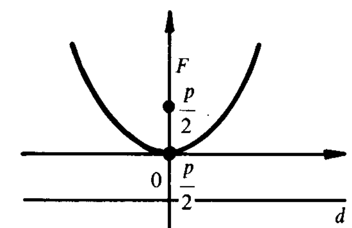
Algunos casos más, con sus ecuaciones, se dan en las figuras siguientes.




$y^2=-2px$




$(y-c_2)^2=-2p(x-c_1)$




$x^2=2py$

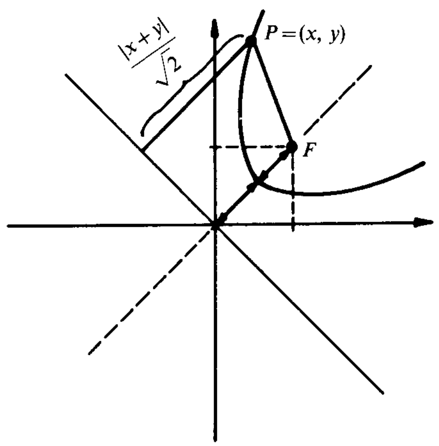
* * *

Las ecuaciones (1), (2), (3) y (4) reciben el nombre de *forma canónica* de la cónica correspondiente y es conveniente observar que en la forma canónica el eje principal coincide con el eje $OX$.

Casos más complicados se presentan cuando los ejes de la cónica no son paralelos a los ejes cartesianos; para familiarizar al lector con las ecuaciones de tales objetos realizamos algunos ejemplos, advirtiéndole que conviene intentar su realización antes de mirar la solución.

EJEMPLO A. *Encontrar la ecuación de la parábola de foco* $(\sqrt{2}, \sqrt{2})$ *y directriz* $x=-y$. Si $P=(x, y)$ es un punto de la parábola se ha de tener

$$\|\overrightarrow{PF}\| = \sqrt{(x-\sqrt{2})^2+(y-\sqrt{2})^2} = d(P, x=-y) = \frac{|x+y|}{\sqrt{2}}.$$




Elevando al cuadrado se obtiene

$$x^2-2x\sqrt{2}+2+y^2-2y\sqrt{2}+2=(x^2+2xy+y^2)/2$$

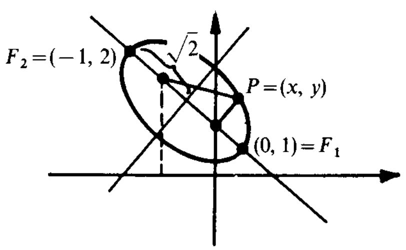
o bien

$$x^2+y^2-2xy-4\sqrt{2}x-4\sqrt{2}y+8=0$$

* * *

EJEMPLO B. *Encontrar la ecuación de la elipse de focos* $(0, 1)$ *y* $(-1, 2)$ *y semieje mayor* $a=\sqrt{2}$. Si $P=(x, y)$ es un punto de la elipse se tiene

$$2\sqrt{2}=\sqrt{(x+1)^2+(y-2)^2}+\sqrt{(x-0)^2+(y-1)^2}$$






Elevando al cuadrado y simplificando adecuadamente se obtiene

$$8+2x-4y+1+4-4\sqrt{2}\sqrt{(x+1)^2+(y-2)^2}=-2y+1$$

o bien

$$x-y+6=2\sqrt{2}\sqrt{(x+1)^2+(y-2)^2}.$$

Elevando de nuevo al cuadrado se obtiene

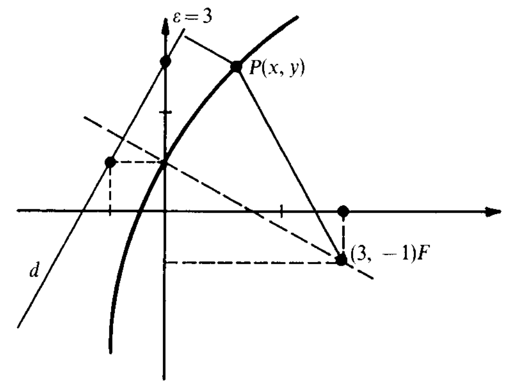
$$x^2+y^2+36-2xy+12x-12y=8[x^2+2x+1+y^2-4y+4]$$

que es equivalente a

$$7x^2+7y^2+2xy+4x-20y-4=0$$

* * *

EJEMPLO C. *Encontrar la ecuación de la hipérbola de directriz* $2x-y+3=0$, *foco en el punto* $(3, -1)$ *y excentricidad* $\varepsilon=3$.




$$\sqrt{(x-3)^2+(y+1)^2}=\|\overrightarrow{PF}\|=3d(P, d)=3\frac{|2x-y+3|}{\sqrt{5}}$$

es la ecuación pedida. Elevando al cuadrado obtenemos

$$x^2-6x+9+y^2+2y+1=\frac{9}{5}[4x^2+y^2+9-4xy+12x-6y]$$

$$31x^2+4y^2-36xy+138x-64y+31=0$$

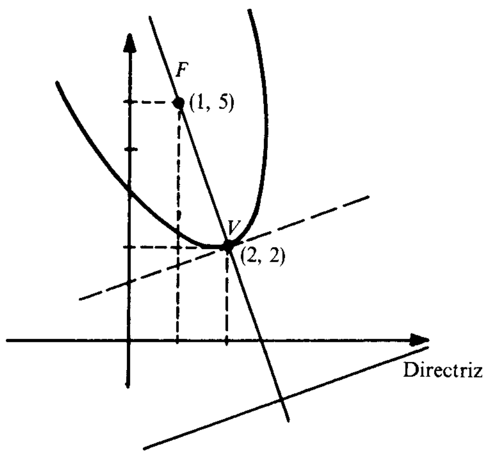
* * *

EJEMPLO D. *Encontrar la ecuación de la parábola de foco* $(1, 5)$ *y vértice* $(2, 2)$. La directriz es una recta perpendicular al eje de la parábola (este eje pasa por el foco y el vértice) y cuya distancia del vértice coincide con la distancia del vértice al foco. El eje de la parábola tiene por ecuacion

$$y-2=\frac{5-2}{1-2}(x-2)$$

o bien

$$3x+y-8=0$$




Una recta perpendicular a ésta tendrá por ecuación

$$d\colon\ x-3y=c;$$

determinamos $c$ imponiendo que

$$d(V, d)=\|\overrightarrow{FV}\|=\sqrt{1^2+3^2}=\sqrt{10}$$

por tanto,

$$\sqrt{10}=\frac{|2-3\cdot 2-c|}{\sqrt{10}} \quad ; \quad 10=|-4-c| \quad ; \quad c=-14,\ c=6$$

Las posibles directrices son $x-3y=-14$, $x-3y=6$; la primera de estas rectas pasa por el punto (1, 5) y, por tanto, no puede ser la directriz; concluimos que $x-3y=6$ es la directriz de la parábola. La ecuación de la parábola es, por tanto,

$$\|PF\|=\sqrt{(x-1)^2+(y-5)^2}=d(P, d)=\frac{|x-3y-6|}{\sqrt{10}}$$



Elevando al cuadrado y simplificando se obtiene

$$10[x^2-2x+1+y^2-10y+25]=x^2+9y^2+36-6xy-12x+36y$$

o bien

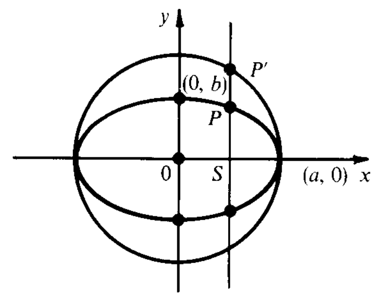
$$9x^2+y^2+6xy-8x-136y+224=0$$

* * *

Observar que cuando el eje de la sección cónica está girado con respecto a los ejes coordenados, en la ecuación de la cónica siempre aparece un término de la forma $xy$.

## EJERCICIOS 11.5

**1.** Considerar la elipse y la circunferencia que se indican en la figura adjunta. Si por cualquier punto $S$ del eje $Ox$ se traza una paralela al eje $Oy$ que corta a la elipse en el punto $P$ y a la circunferencia en $P'$, demostrar que $\dfrac{\|P'S\|}{\|PS\|}=\dfrac{a}{b}$




$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

**2.** Escribir las ecuaciones de las parábolas de foco (2, −1), que pasan por (2, 2) y tienen el eje en la dirección del eje $Ox$.

**3.** Escribir la ecuación de la hipérbola con un foco en (2, −1) y asíntotas $x=0$, $3x-4y=0$.

**4.** Escribir la ecuación de la elipse con vértices (−1, 2) y (−7, 2), y eje menor $2b=2$.

**5.** Escribir la ecuación de la hipérbola con asíntotas $y=\pm 2x-1$ y un foco en (3, −1).

**6.** Escribir la ecuación de la elipse con un vértice en (−1, 1), centro en (3, −1) y excentricidad 1/2.

## 11.6. DETERMINACION DE LAS CONICAS

Hemos visto en los ejemplos de la sección precedente que las ecuaciones de las cónicas que allí se obtuvieron vienen dadas mediante polinomios de segundo grado en dos variables. Esto es cierto para cualquier cónica y para demostrarlo utilizaremos el hecho de que toda cónica no degenerada es el lugar geométrico de los puntos $P$ que satisfacen $\|\vec{PF}\| = \varepsilon d(P, l)$, donde $F$ es el foco y $l$ la directriz (ver sección 11.4). Sea $F=(c_1, c_2)$ y $l$: $ax+by+c=0$. La igualdad $\|\vec{PF}\| = \varepsilon d(P, l)$ se transforma en

$$\sqrt{(x-c_1)^2+(y-c_2)^2} = \varepsilon \frac{|ax+by+c|}{\sqrt{a^2+b^2}}$$

donde $P=(x, y)$. Elevando al cuadrado y multiplicando por $a^2+b^2$ se obtiene

$$(a^2+b^2)(x-c_1)^2+(a^2+b^2)(y-c_2)^2 = \varepsilon^2[ax+by+c]^2.$$

Después de efectuar las operaciones indicadas en la fórmula anterior y de agrupar los términos adecuadamente se obtiene:

$$Ax^2+By^2+Cxy+Dx+Ey+F=0 \tag{1}$$

donde

$$A=a^2+b^2-\varepsilon^2a^2, \quad B=a^2+b^2-\varepsilon^2b^2, \quad C=-2\varepsilon^2ab$$
$$D=-2c_1(a^2+b^2)-2\varepsilon^2ac, \quad E=-2c_2(a^2+b^2)-2\varepsilon^2bc$$
$$F=(a^2+b^2)[c_1^2+c_2^2]-\varepsilon^2c^2$$

La igualdad (1) será denominada *ecuación general de una cónica* y muestra que toda cónica es una expresión polinómica de segundo grado en dos variables.

Para determinar una circunferencia es necesario conocer tres puntos de ella. Para determinar una cónica es necesario conocer más de tres puntos puesto que la ecuación (1) posee seis incógnitas $A$, $B$, $C$, $D$, $E$ y $F$. Dividiendo entre $A$ se obtiene una ecuación de la forma

$$x^2+B'y^2+C'xy+D'x+E'y+F'=0$$

y, por tanto, sólo tenemos cinco incógnitas. En efecto, éste es el número máximo de puntos que se necesitan para determinar una cónica no degenerada. El primer matemático que probó este resultado fue Blaise Pascal. En 1639, a sus diecisés años, B. Pascal (1623-1662) logró demostrar el siguiente teorema:

TEOREMA 1

Sean $l_1, l_2, l_3, l_4, l_5, l_6$ los lados de un «hexágono» cualquiera inscrito en una cónica arbitraria (una elipse en la figura 1 o una hipérbola en la figura 2). Entonces, los tres puntos de intersección de cada dos lados opuestos $l_1l_4$, $l_2l_5$ y $l_3l_6$ están situados sobre una recta, llamada *recta de Pascal.*



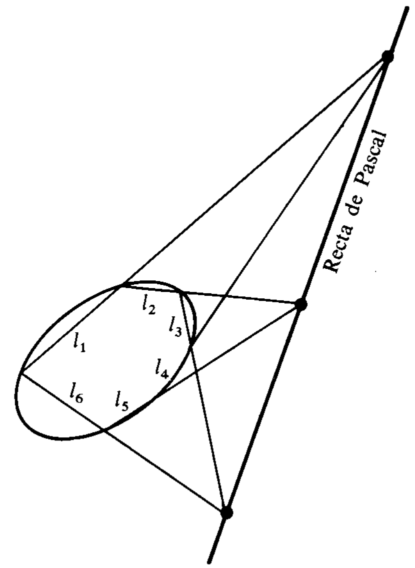



Figura 1

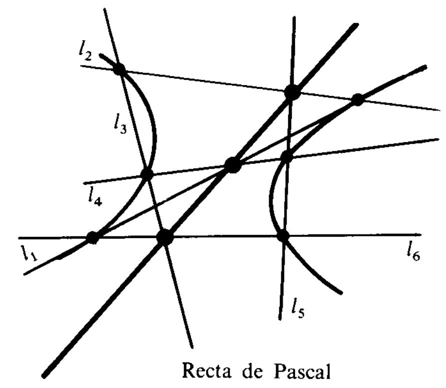



Figura 2

El teorema de Pascal puede utilizarse para demostrar que cinco puntos son suficientes para determinar una cónica. En efecto, supongamos que conocemos cinco puntos de un plano como en la figura 3. Trazamos las rectas que unen estos puntos en orden cíclico, como en la figura 3.

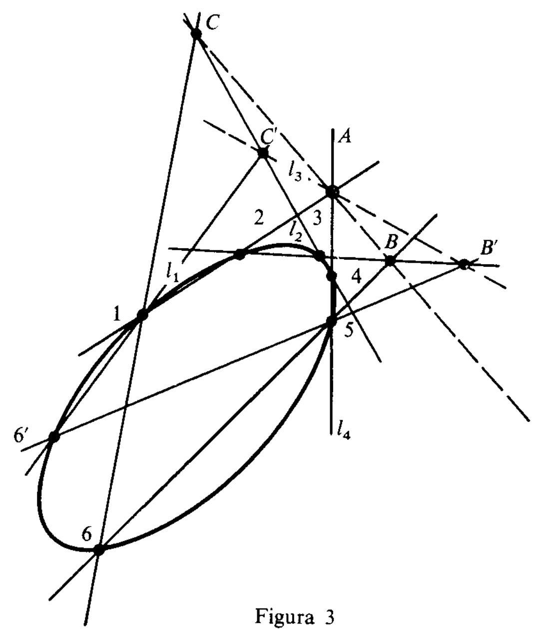



Figura 3

El punto $A$ de intersección de $l_1$ y $l_4$ es un punto de la recta de Pascal de esta cónica. Tomando una recta cualquiera que pase por $A$ se obtienen $B$ y $C$ como los puntos de intersección de esta recta con $l_2$ y $l_3$. Por $B$ se traza una recta que pasa por 5 y por $C$ se traza otra recta que pasa por 1; el punto de intersección de estas dos rectas es un punto (6) de la cónica. Nuevos puntos de la cónica se obtienen tomando diferentes rectas que pasen por $A$.

El teorema de Pascal puede demostrarse utilizando el denominado «principio de dualidad», que para que nos entendamos puede enunciarse en este contexto diciendo que si se intercambian «puntos» por «rectas», «cortar dos rectas» por «unir dos puntos» y «puntos de la cónica» por «tangentes a la cónica» todo teorema cierto se convierte en otro teorema cierto. Con la ayuda de este principio, Brianchon (1785-1864) demostró en 1806 el teorema de Pascal. La demostración de Brianchon se encuentra elegantemente expuesta en el libro *Mirar y ver*, de M. de Guzmán (Ed. Alhambra).

## 11.7. DETERMINACION DEL TIPO DE UNA CONICA

La ecuación general de una cónica, que se obtuvo en (1) de la sección 11.6, la escribimos de la forma

$$a_{11}x^2+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=0 \tag{1}$$

Nos preguntamos ahora si toda ecuación de la forma (1) representa una cónica y en caso de que la represente de qué tipo de cónica se trata. La observación fundamental se hizo ya en la sección 11.5 y es que las cónicas que tienen sus ejes paralelos a los ejes coordenados no poseen el término $xy$, y que los términos en $x$ y en $y$ pueden fácilmente eliminarse mediante traslaciones.

Los términos $a_{11}x^2+a_{12}xy+a_{22}y^2$ se denominan *parte principal de* (1) y pueden escribirse de la forma

$$a_{11}x^2+a_{12}xy+a_{22}y^2=(x,\ y)\begin{pmatrix} a_{11} & \dfrac{a_{12}}{2} \\ \dfrac{a_{12}}{2} & a_{22}\end{pmatrix}\begin{pmatrix} x \\ y\end{pmatrix}=(x,\ y)A\begin{pmatrix} x \\ y\end{pmatrix} \tag{2}$$

como puede comprobarse fácilmente.

Para eliminar el término $xy$ en (2) procedemos de la siguiente manera: la matriz $A$ es una matriz (real) simétrica; en el capítulo 8 se demostró que $A$ puede diagonalizarse mediante un cambio de base ortonormal y que sus autovalores son reales. Sea $C$ la matriz de cambio de base; puesto que $C$ transforma una base ortonormal (la canónica) en otra base ortonormal debe ser una matriz ortogonal; como las únicas matrices



ortogonales en $\mathbb{R}^2$ son los giros y las simetrías, si imponemos, además, que la nueva base conserve la orientación, $C$ debe ser un giro.

Si $(x_1,\ y_1)$ son las nuevas coordenadas se tiene

$$\begin{pmatrix} x \\ y\end{pmatrix}=C\begin{pmatrix} x_1 \\ y_1\end{pmatrix}$$

y puesto que $C^t=C^{-1}$ (puesto que $C$ es ortogonal) (2) puede escribirse de la forma

$$(x,\ y)A\begin{pmatrix} x \\ y\end{pmatrix}=(x_1,\ y_1)C^{-1}AC\begin{pmatrix} x_1 \\ y_1\end{pmatrix}=(x_1,\ y_1)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2\end{pmatrix}\begin{pmatrix} x_1 \\ y_1\end{pmatrix}=\lambda_1x_1^2+\lambda_2x_2^2$$

donde $\lambda_1$ y $\lambda_2$ son los autovalores de $A$. Hemos logrado, pues, eliminar el término $xy$ mediante un giro de los ejes coordenados.

La ecuación (1) se transforma ahora en

$$\lambda_1x_1^2+\lambda_2x_2^2+b_1x_1+b_2x_2+b=0 \tag{3}$$

Antes de realizar ejemplos pasamos a estudiar los posibles casos que pueden presentarse en (3).

CASO I. $\delta=|A|=\lambda_1\lambda_2>0$

En este caso los autovalores $\lambda_1$, $\lambda_2$ tienen el mismo signo y siempre podemos tomarlos positivos en (3) y $\lambda_1\leqslant\lambda_2$. Puesto que (3) puede escribirse de la forma

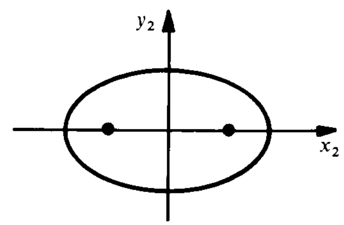
$$\lambda_1\left(x_1+\frac{b_1}{2\lambda_1}\right)^2+\lambda_2\left(y_1+\frac{b_2}{2\lambda_2}\right)^2-\frac{b_1^2}{4\lambda_1}-\frac{b_2^2}{4\lambda_2}+b=0,$$

la traslación dada por

$$\begin{cases} x_2=x_1+\dfrac{b_1}{2\lambda_1} \\ y_2=y_1+\dfrac{b_2}{2\lambda_2}\end{cases}$$

transforma la cónica en

$$\frac{x_2^2}{1/\lambda_1}+\frac{y_2^2}{1/\lambda_2}=c$$



donde $c=\frac{b_1^2}{4\lambda_1}+\frac{b_2^2}{4\lambda_2}-b$. Puesto que siempre podemos tomar $\lambda_2\geqslant\lambda_1>0$ (en caso de que sean negativos se cambian ambos de signo en (3)) hemos de tener $1/\lambda_1\geqslant 1/\lambda_2$ y, por tanto, la ecuación anterior es la de una elipse con los focos en el eje $Ox_2$ si $c>0$.

Si $c=0$ se obtiene un punto y si $c<0$ no se obtiene nada real puesto que la igualdad anterior es imposible.

En este caso la cónica se dice que es de tipo elíptico.

CASO II. $\delta=|A|=\lambda_1\lambda_2<0$

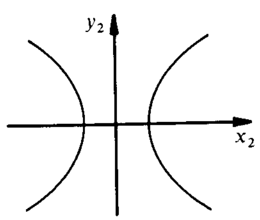
En este caso $\lambda_1$ y $\lambda_2$ tienen distinto signo y, por tanto, siempre podemos tomar $\lambda_1>0$, $\lambda_2<0$. Mediante una traslación como en el caso I, la ecuación (3) se transforma en

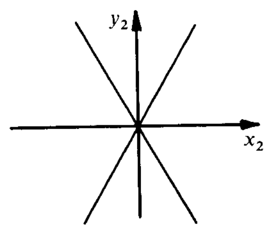
$$\lambda_1 x_2^2+\lambda_2 y_2^2=c$$

Si $c\neq 0$, se tiene una hipérbola; si $c=0$, se tiene

$$y_2=\pm\sqrt{\frac{-\lambda_1}{\lambda_2}}\,x_2$$

que representa un par de rectas de pendiente $\pm\sqrt{\frac{-\lambda_1}{\lambda_2}}$.




En este caso la cónica se dice que es de tipo *hiperbólico*.

CASO III. $\delta=|A|=\lambda_1\lambda_2=0$

En este caso siempre podemos tomar $\lambda_1=0$, $\lambda_2\neq 0$ (observad que los dos autovalores no pueden ser nulos a la vez, ya que entonces $A=0$). Por tanto, mediante un giro adecuado, (3) se transforma en

$$\lambda_2 y_1^2+b_1x_1+b_2y_1+b=0. \tag{4}$$

Si $b_1\neq 0$, (4) puede escribirse de la forma

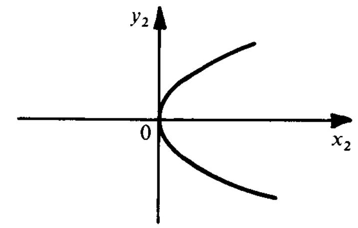
$$\lambda_2\left(y_1+\frac{b_2}{2\lambda_2}\right)^2+b_1\left(x_1+\frac{b}{b_1}-\frac{b_2^2}{4\lambda_2 b_1}\right)=0$$



y, por tanto, la traslación

$$\begin{cases} x_2=x_1+\dfrac{b}{b_1}-\dfrac{b_2^2}{4\lambda_2 b_1} \\ y_2=y_1+\dfrac{b_2}{2\lambda_2}\end{cases}$$

nos permite escribir $y_2^2=-\frac{b_1}{\lambda_2}x_2$, que representa a una parábola con el foco en la dirección de $x_2$.



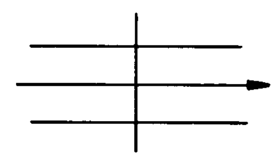

Si $b_1=0$, (4) puede escribise de la forma

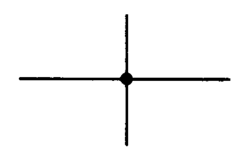
$$\lambda_2\left(y_1+\frac{b_2}{2\lambda_2}\right)^2+b-\frac{b_2^2}{4\lambda_2}=0$$

y, por tanto, la traslación

$$\begin{cases} x_2=x_1 \\ y_2=y_1+\dfrac{b_2}{2\lambda_2}\end{cases}$$

la transforma en $y_2^2=-\frac{c}{\lambda_2}$, donde $c=b-\frac{b_2^2}{4\lambda_2}$. Si $-\frac{c}{\lambda_2}>0$, $y_2=\pm\sqrt{-\frac{c}{\lambda_2}}$ representa dos rectas paralelas, si $-\frac{c}{\lambda_2}=0$ se obtiene un punto, y si $-\frac{c}{\lambda_2}<0$ se obtiene el conjunto vacío.





En este caso la cónica se dice de tipo *parabólico*.

* * *

A continuación realizamos varios ejemplos en los que dada una ecuación polinómica de segundo grado en dos variables, tratamos de identificarla y de encontrar sus elementos geométricos: ejes, focos, vértices, asíntotas, etc. De ahora en adelante cada vez que pidamos *estudiar una sección cónica*, nos referiremos a encontrar todos sus elementos geométricos, así como indicar su tipo y escribirla en forma canónica.

EJEMPLO A. *Estudiar la sección cónica* $3x^2-2xy+3y^2+2x-4y+1=0$. Puesto que $A=\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$, $|A|=8>0$ y, por tanto, es una sección de *tipo elíptico*. Los autovalores de $A$ son las soluciones de la ecuación

$$0=\begin{vmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix}=(3-\lambda)^2-1=\lambda^2-6\lambda+8=(\lambda-4)(\lambda-2);$$

podemos elegir $\lambda_1=2$, $\lambda_2=4$ para que tengamos $\lambda_1\leqslant\lambda_2$.

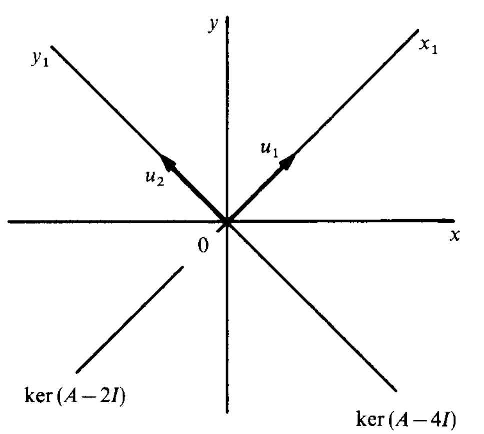
Tratamos ahora de encontrar los nuevos ejes de manera que el término $-2xy$ desaparezca.

$\ker(A-2I)$:

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x=y$$

$\ker(A-4I)$:

$$\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x=-y$$






Elegimos estos ejes como nuevo sistema de coordenadas, de manera que la matriz $C$ del cambio de base es

$$C=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}$$

ya que $\vec{u}_1=\left(\frac{1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$, $\vec{u}_2=\left(\frac{-1}{\sqrt{2}},\frac{1}{\sqrt{2}}\right)$. Así pues, tenemos

$$\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

o bien

$$x=\frac{1}{\sqrt{2}}(x_1-y_1),\quad y=\frac{1}{\sqrt{2}}(x_1+y_1)$$

Sustituyendo en la ecuación dada se obtiene

$$\frac{3}{2}(x_1-y_1)^2-(x_1-y_1)(x_1+y_1)+\frac{3}{2}(x_1+y_1)^2+$$
$$+2\frac{1}{\sqrt{2}}(x_1-y_1)-4\frac{1}{\sqrt{2}}(x_1+y_1)+1=0$$

que simplificando se transforma en

$$2x_1^2+4y_1^2-\frac{2}{\sqrt{2}}x_1-\frac{6}{\sqrt{2}}y_1+1=0.$$

Puesto que esta igualdad puede escribirse de la forma

$$2\left(x_1-\frac{1}{2\sqrt{2}}\right)^2+4\left(y_1-\frac{3}{4\sqrt{2}}\right)^2-\frac{2}{8}-\frac{9}{8}+1=0,$$

la traslación

$$\begin{cases} x_2 = x_1 - \dfrac{1}{2\sqrt{2}} \\ y_2 = y_1 - \dfrac{3}{4\sqrt{2}} \end{cases}$$

nos permite escribir $2x_2^2 + 4y_2^2 = \frac{3}{8}$, o bien

$$\frac{x_2^2}{3/16} + \frac{y_2^2}{3/32} = 1$$

que es la ecuación canónica de una *elipse*; para esta elipse se tiene $a = \frac{\sqrt{3}}{4}$, $b = \sqrt{\frac{3}{32}}$.

Si $C$ es el centro de la elipse, las coordenadas de $C$ en el sistema de referencia $\{0; \vec{u}_1, \vec{u}_2\}$ son

$$\left(\frac{1}{2\sqrt{2}}, \frac{3}{4\sqrt{2}}\right).$$

Por tanto, las coordenadas de $C$ en el sistema de referencia $\{0; \vec{e}_1, \vec{e}_2\}$ son

$$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} \dfrac{1}{2\sqrt{2}} \\ \dfrac{3}{4\sqrt{2}} \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} \dfrac{1}{2\sqrt{2}} - \dfrac{3}{4\sqrt{2}} \\ \dfrac{1}{2\sqrt{2}} + \dfrac{3}{4\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\dfrac{1}{8} \\ \dfrac{5}{8} \end{pmatrix}.$$

El eje principal es paralelo a $\ker(A-2I)$ y, por tanto, es

$$y - \frac{5}{8} = x + \frac{1}{8} \Leftrightarrow y = x + \frac{3}{4}$$

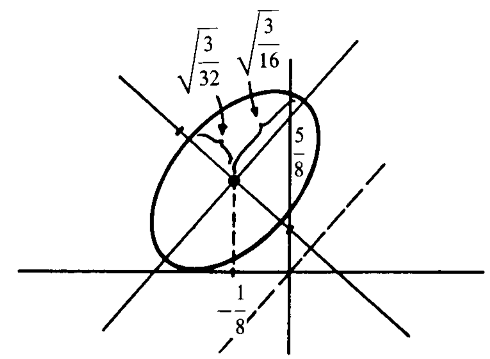
ya que pasa por $C$. El eje secundario es paralelo a $\ker(A-4I)$ y pasa por $C$; por tanto, su ecuación es

$$y - \frac{5}{8} = -\left(x + \frac{1}{8}\right) \Leftrightarrow y = -x + \frac{1}{2}$$



Finalmente,

$$c^2 = a^2 - b^2 = \frac{3}{16} - \frac{3}{32} = \frac{3}{32} \Leftrightarrow c = \sqrt{\frac{3}{32}}$$




* * *

EJEMPLO B. *Estudiar la cónica* $x^2 - 2xy + y^2 + 4x - 6y + 1 = 0$. Puesto que $A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $|A| = 1 - 1 = 0$ y, por tanto, es una cónica de tipo *parabólico*. Puesto que

$$\begin{vmatrix} 1-\lambda & -1 \\ -1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 1 = 1 - 2\lambda + \lambda^2 - 1 = \lambda^2 - 2\lambda$$

los autovalores de $A$ son $\lambda_1 = 0$, $\lambda_2 = 2$.

$\ker(A - 0I) = \ker(A)$:

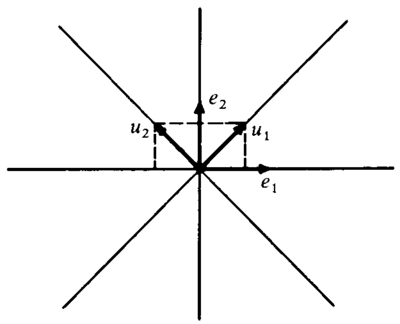
$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x = y. \quad \text{Tomar } \vec{u}_1 = \frac{1}{\sqrt{2}}(1, 1).$$

$\ker(A - 2I)$:

$$\begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow x = -y. \quad \text{Tomar } \vec{u}_2 = \frac{1}{\sqrt{2}}(-1, +1).$$

En este nuevo sistema de coordenadas (que está girado 45° con respecto al sistema canónico) tenemos que

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \dfrac{1}{\sqrt{2}} & -\dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$




o bien

$$x=\frac{1}{\sqrt{2}}(x_1-y_1) \quad , \quad y=\frac{1}{\sqrt{2}}(x_1+y_1)$$

Sustituyendo estos valores de $x$ e $y$ en la ecuación dada obtenemos

$$0\cdot x_1^2+2y_1^2+\frac{4}{\sqrt{2}}(x_1-y_1)-\frac{6}{\sqrt{2}}(x_1+y_1)+1=0$$

o bien

$$2y_1^2-\frac{2}{\sqrt{2}}x_1-\frac{10}{\sqrt{2}}y_1+1=0.$$

Completando cuadrados obtenemos

$$2\left(y_1-\frac{5}{2\sqrt{2}}\right)^2-\frac{2}{\sqrt{2}}\left(x_1-\frac{\sqrt{2}}{2}+\frac{25\sqrt{2}}{8}\right)=0,$$

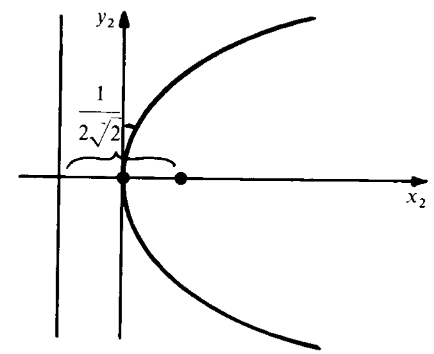
con lo cual la traslación

$$\left.\begin{aligned} x_2&=x_1-\frac{\sqrt{2}}{2}+\frac{25\sqrt{2}}{8}\\ y_2&=y_1-\frac{5}{2\sqrt{2}}\end{aligned}\right\}$$

la transforma en $2y_2^2-\frac{2}{\sqrt{2}}x_2=0$. Se trata, pues, de una parábola cuya forma canónica es

$$y_2^2=\frac{1}{\sqrt{2}}x_2$$






Tenemos que $2p=\frac{1}{\sqrt{2}} \Leftrightarrow p=\frac{1}{2\sqrt{2}}$.

El vértice de la parábola en el sistema $(x_1, y_1)$ es $\left(\frac{\sqrt{2}}{2}-\frac{25\sqrt{2}}{8}, \frac{5}{2\sqrt{2}}\right)$; en el sistema de referencia canónico su vértice es

$$V=\begin{pmatrix}x\\y\end{pmatrix}=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\frac{\sqrt{2}}{2}-\frac{25\sqrt{2}}{8}\\[2mm] \frac{5}{2\sqrt{2}}\end{pmatrix}=\begin{pmatrix}\frac{1}{2}-\frac{25}{8}-\frac{5}{4}\\[2mm] \frac{1}{2}-\frac{25}{8}+\frac{5}{4}\end{pmatrix}=\begin{pmatrix}-\frac{31}{8}\\[2mm] -\frac{11}{8}\end{pmatrix}$$

El eje principal pasa por el vértice y tiene $\vec{u}_1$ como vector director; por tanto,

$$y+\frac{11}{8}=x+\frac{31}{8} \Leftrightarrow y=x+\frac{5}{2}$$

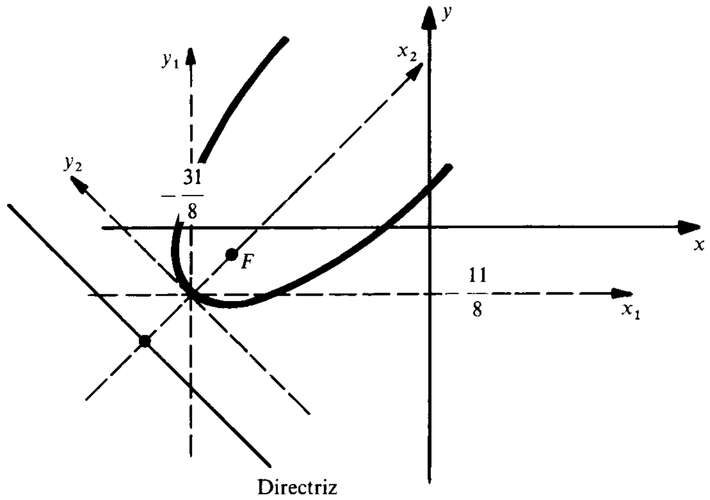
*El eje principal pasa también por el vértice y tiene* $\vec{u}_2$ como vector director; por tanto,

$$y+\frac{11}{8}=-\left(x+\frac{31}{8}\right) \Leftrightarrow y=-x-\frac{21}{4}$$

El foco está a distancia $\frac{1}{4\sqrt{2}}$ del vértice, en el eje principal; como el eje principal está inclinado 45° con respecto al eje $OX$ se tiene

$$F=\left(-\frac{31}{8}+\frac{1}{4\sqrt{2}}\cos 45^\circ, -\frac{11}{8}+\frac{1}{4\sqrt{2}}\operatorname{sen} 45^\circ\right)=$$

$$=\left(-\frac{31}{8}+\frac{1}{8}, -\frac{11}{8}+\frac{1}{8}\right)=\left(-\frac{30}{8}, -\frac{10}{8}\right)$$




Puesto que la directriz ha de ser paralela al eje secundario, su ecuación será de la forma $x+y+k=0$, donde

$$\frac{1}{4\sqrt{2}}=d(V, \text{directriz})=\frac{\left|-\frac{31}{8}-\frac{11}{8}+k\right|}{\sqrt{2}}$$

$$\frac{1}{4}=\left|-\frac{42}{8}+k\right| \Leftrightarrow \left\{\begin{array}{l} \frac{1}{4}=-\frac{21}{4}+k \\ \frac{1}{4}=\frac{21}{4}-k \end{array}\right\} \Leftrightarrow \begin{array}{l} k=\frac{22}{4}=\frac{11}{2} \\ k=\frac{20}{4}=5 \end{array}$$

La recta $x+y+\frac{11}{2}=0$ es la directriz, ya que la recta $x+y+5=0$ corta al eje principal en

$$\left.\begin{array}{l} x+y+5=0 \\ x-y+\frac{5}{2}=0 \end{array}\right\} \Leftrightarrow x=-\frac{15}{4}=-\frac{30}{8} \quad , \quad y=-\frac{10}{8}$$

y, por tanto, pasa por el foco.



## 11.8. INVARIANTES DE LAS CONICAS Y REDUCCION A SU FORMA CANONICA

Una ecuación de segundo grado en las variables $x$ e $y$ puede escribirse de la forma

$$a_{11}x^2+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=(x, y, 1)\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a \end{pmatrix}\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad (1)$$

como puede comprobarse fácilmente. Una expresión de este tipo igualada a cero será denominada una *curva de segundo grado.*

DEFINICIÓN 1

Se denominan *invariantes* de una curva de segundo grado a cualquier expresión formada por los coeficientes de su ecuación que no varía al cambiar de un sistema de coordenadas a otro mediante un giro o una traslación.

TEOREMA 2

Son invariantes de la curva dada en (1) los siguientes:

1. $s=a_{11}+a_{22}$ 2. $\delta=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22} \end{vmatrix}$ 3. $\Delta=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a \end{vmatrix}$

*Demostración.* Comenzamos demostrando que $s$, $\delta$ y $\Delta$ quedan invariantes mediante traslaciones; una traslación de la forma

$$\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} x' \\ y' \end{pmatrix}+\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

transforma (1) en

$$a_{11}(x'+\alpha)^2+a_{12}(x'+\alpha)(y'+\beta)+a_{22}(y'+\beta)^2+a_1(x'+\alpha)+a_2(y'+\beta)+a=$$
$$=a_{11}[x']^2+a_{12}x'y'+a_{22}[y']^2+(2a_{11}\alpha+a_{12}\beta+a_1)x'+(2a_{22}\beta+a_{12}\alpha+a_2)y'+$$
$$+(a_{11}\alpha^2+a_{12}\alpha\beta+a_{22}\beta^2+a_1\alpha+a_2\beta+a).$$

Como los coeficientes de los términos principales no han cambiado, $s$ y $\delta$ son invariantes mediante traslaciones. Observar que

$$2a_{11}\alpha+a_{12}\beta+a_1=\frac{\partial f}{\partial x}(\alpha,\,\beta)$$

$$2a_{22}\beta+a_{12}\alpha+a_2=\frac{\partial f}{\partial y}(\alpha,\,\beta)$$

$$a_{11}\alpha^2+a_{12}\alpha\beta+a_{22}\beta^2+a_1\alpha+a_2\beta+a=f(\alpha,\,\beta)$$

donde $f(x,\,y)$ es la expresión cuadrática dada en (1); por tanto, mediante la traslación dada, (1) se transforma en

$$a_{11}[x']^2+a_{12}x'y'+a_{22}[y']^2+\frac{\partial f}{\partial x}(\alpha,\,\beta)x'+\frac{\partial f}{\partial y}(\alpha,\,\beta)y'+f(\alpha,\,\beta). \tag{2}$$

Para (2) el determinante $\Delta$ es

$$\Delta=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & \frac{1}{2}\frac{\partial f}{\partial x}(\alpha,\,\beta) \\ \frac{a_{11}}{2} & a_{22} & \frac{1}{2}\frac{\partial f}{\partial y}(\alpha,\,\beta) \\ \frac{1}{2}\frac{\partial f}{\partial x}(\alpha,\,\beta) & \frac{1}{2}\frac{\partial P}{\partial x}(\alpha,\,\beta) & f(\alpha,\,\beta)\end{vmatrix}=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & a_{11}\alpha+\frac{a_{12}}{2}\beta+\frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & a_{22}\beta+\frac{a_{12}}{2}\alpha+\frac{a_2}{2} \\ a_{11}\alpha+\frac{a_{12}}{2}\beta+\frac{a_1}{2} & a_{22}\beta+\frac{a_{12}}{2}\alpha+\frac{a_2}{2} & f(\alpha,\,\beta)\end{vmatrix}$$

$$=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & a_{11}\alpha+\frac{a_{12}}{2}\beta+\frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & a_{22}\beta+\frac{a_{12}}{2}\alpha+\frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_1}{2}2+\frac{a_2}{2}\beta+a\end{vmatrix}=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a\end{vmatrix}$$

lo cual prueba el resultado deseado. (*Nota*: el lector no tendrá dificultad en encontrar los factores por los que se han multiplicado las filas y columnas de las matrices anteriores para llegar al resultado final.)

Supongamos ahora que hemos realizado un giro de la forma

$$\begin{pmatrix}x\\y\end{pmatrix}=C\begin{pmatrix}x'\\y'\end{pmatrix}=\begin{pmatrix}\cos\varphi & -\operatorname{sen}\varphi\\ \operatorname{sen}\varphi & \cos\varphi\end{pmatrix}\begin{pmatrix}x'\\y'\end{pmatrix}$$



la parte principal de la cónica se transforma en

$$(x,\,y)A\begin{pmatrix}x\\y\end{pmatrix}=(x',\,y')C^tAC\begin{pmatrix}x'\\y'\end{pmatrix}=(x',\,y')C^{-1}AC\begin{pmatrix}x'\\y'\end{pmatrix}=$$

$$a'_{11}[x']^2+a'_{12}x'y'+a'_{22}[y']^2$$

Por tanto, $s=a_{11}+a_{22}=\text{traza}(A)=\text{traza}(C^{-1}AC)=a'_{11}+a'_{22}$ (ver problema 2 de la sección 7.6 del capítulo 7) y

$$\delta=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22}\end{vmatrix}=|A|=|C^{-1}AC|=\begin{vmatrix} a'_{11} & \frac{a'_{12}}{2} \\ \frac{a'_{12}}{2} & a'_{22}\end{vmatrix}$$

lo cual prueba que $s$ y $\delta$ permanecen invariantes mediante un giro.

Para demostrar que $\Delta$ es invariante observamos que

$$(x,\;y,\;1)\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a\end{pmatrix}\begin{pmatrix}x\\y\\1\end{pmatrix}=(x,\;y,\;1)\left(\begin{array}{cc|c} \multicolumn{2}{c|}{A} & \begin{matrix}a_1/2\\ a_2/2\end{matrix} \\ \hline \frac{a_1}{2} & \frac{a_2}{2} & a\end{array}\right)\begin{pmatrix}x\\y\\1\end{pmatrix}$$

$$=\left((x,\,y)A+\left(\frac{a_1}{2},\frac{a_2}{2}\right),\;(x,\,y)\begin{pmatrix}a_{1/2}\\a_{2/2}\end{pmatrix}+a\right)\begin{pmatrix}x\\y\\1\end{pmatrix}=$$

$$=(x,\,y)A\begin{pmatrix}x\\y\end{pmatrix}+\left(\frac{a_1}{2},\frac{a_2}{2}\right)\begin{pmatrix}x\\y\end{pmatrix}+(x,\,y)\begin{pmatrix}a_{1/2}\\a_{2/2}\end{pmatrix}+a=$$

$$=(x',\,y')\,C^{-1}\,AC\begin{pmatrix}x'\\y'\end{pmatrix}+\left(\frac{a_1}{2},\frac{a_2}{2}\right)C\begin{pmatrix}x'\\y'\end{pmatrix}+(x',\,y')\,C^{-1}\begin{pmatrix}a_{1/2}\\a_{2/2}\end{pmatrix}+a=$$

$$=(x',\,y',\,1)\left(\begin{array}{c|c} C^{-1}AC & C^{-1}\begin{pmatrix}a_{1/2}\\a_{2/2}\end{pmatrix} \\ \hline \left(\frac{a_1}{2},\frac{a_2}{2}\right)C & a\end{array}\right)\begin{pmatrix}x'\\y'\\1\end{pmatrix}$$

la forma canónica de la hipérbola es $\sqrt{2}x_2-\sqrt{2}y_2=c$, donde $\Delta=-\delta c \Rightarrow c=-\frac{\Delta}{\delta}=-\frac{-1}{-2}=-\frac{1}{2}$. Por tanto, tenemos

$$\sqrt{2}x_2-\sqrt{2}y_2=-\frac{1}{2}$$

o bien

$$\frac{x_2}{-\frac{1}{2\sqrt{2}}}+\frac{y_2}{\frac{1}{2\sqrt{2}}}=1$$

Haciendo $x_2=Y$, $y_2=X$ se tiene

$$\frac{X}{\frac{1}{2\sqrt{2}}}-\frac{Y}{\frac{1}{2\sqrt{2}}}=1$$

y, por tanto, se trata de una hipérbola equilátera, es decir, $a=b=(\sqrt[4]{8})^{-1}$.

EJEMPLO B. *Reducir a su forma canónica (si es posible) la cónica* $x^2+3xy+2y^2+2x+5y-3=0$

$$\delta=\begin{vmatrix}1 & \frac{3}{2}\\ \frac{3}{2} & 2\end{vmatrix}=2-\frac{9}{4}=-\frac{1}{4} \quad ; \quad \Delta=\begin{vmatrix}1 & \frac{3}{2} & 1\\ \frac{3}{2} & 2 & \frac{5}{2}\\ 1 & \frac{5}{2} & -3\end{vmatrix}=\begin{vmatrix}1 & \frac{3}{2} & 1\\ 0 & -\frac{1}{4} & 1\\ 0 & 1 & -4\end{vmatrix}=0$$

Se trata de una cónica de tipo hiperbólico degenerada: dos rectas que se cortan. Para encontrar estas dos rectas consideramos $y$ fija y resolvemos la ecuación anterior para $x$:

$$x^2+(3y+2)x+[2y^2+5y-3]=0$$

$$x=\frac{-3y-2\pm\sqrt{9y^2+12y+4-8y^2-20y+12}}{2}=$$

$$=\frac{-3y-2\pm\sqrt{y^2-8y+16}}{2}=\frac{-3y-2\pm(y-4)}{2}=\begin{cases}\frac{-2y-6}{2}=-y-3\\ \frac{-4y+2}{2}=-2y+1\end{cases}$$



Por tanto,

$$x^2+3xy+2y^2+2x+5y-3=(x+y+3)(x+2y-1)$$

* * *

EJEMPLO C. *Determinar el tipo de la curva* $x^2+4xy+4y^2-2x-4y-3=0$ *y reducirla a su forma canónica si es posible.*

$$\delta=\begin{vmatrix}1 & 2\\ 2 & 4\end{vmatrix}=4-4=0 \quad ; \quad \Delta=\begin{vmatrix}1 & 2 & -1\\ 2 & 4 & -2\\ -1 & -2 & -3\end{vmatrix}=\begin{vmatrix}1 & 2 & -1\\ 0 & 0 & 0\\ -1 & -2 & -3\end{vmatrix}=0$$

Se trata de una curva de tipo parabólico degenerada: dos rectas paralelas, concurrentes o un punto. Procediendo como en el ejemplo B, se obtiene:

$$x^2+4xy+4y^2-2x-4y-3=(x+2y-3)(x+2y+1)$$

y se trata, por tanto, de dos rectas paralelas.

## 11.9. DETERMINACIÓN DEL CENTRO Y DE LOS EJES PRINCIPALES DE UNA CÓNICA CON CENTRO

Supongamos que

$$f(x, y)=a_{11}x^2+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=0 \tag{1}$$

representa una cónica no degenerada (es decir, $\Delta\neq 0$) y que tiene centro (es decir, es una elipse o una hipérbola y, por tanto $\delta\neq 0$). Las definiciones de $\Delta$ y $\delta$ se dieron en el teorema 2 de la sección 11.8.

En la sección 11.8 se ha descrito un método para obtener la forma canónica de una cónica sin necesidad de realizar un giro y una traslación; este método es rápido en cuanto a la obtención de la forma canónica, pero no permite calcular los ejes de la cónica. En esta sección expondremos un método para calculr los ejes de una cónica con centro.

Un vector director de los ejes es cualquier autovector de la matriz

$$A=\begin{pmatrix}a_{11} & \frac{a_{12}}{2}\\ \frac{a_{12}}{2} & a_{22}\end{pmatrix}$$

y son, por tanto, fácilmente obtenibles. Conocidos los vectores directores, los ejes quedan determinados si encontramos el *centro* de la cónica, por el cual necesariamente deben pasar los ejes. Supongamos que su centro es $(\alpha, \beta)$; entonces la traslación

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \tag{2}$$

debe transformar (1) en una cónica que no posee términos en $x'$ ni en $y'$. La traslación (2) sustituida en (1) produce

$$a_{11}[x']^2 + a_{12}x'y' + a_{22}[y']^2 + \frac{\partial f}{\partial x}(\alpha, \beta)x' + \frac{\partial f}{\partial y}(\alpha, \beta)y' + f(\alpha, \beta) = 0$$

(ver la demostración del teorema 2 de la sección 11.8). Por tanto

$$\frac{\partial f}{\partial x}(\alpha, \beta) = 0 \quad , \quad \frac{\partial f}{\partial y}(\alpha, \beta) = 0$$

Esto quiere decir que $(\alpha, \beta)$ son solución del sistema

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 2a_{11}x + a_{12}y + a_1 = 0 \\ \frac{\partial f}{\partial y} &= a_{12}x + 2a_{22}y + a_2 = 0 \end{aligned}\right\}$$

Observar que este sistema tiene una solución única, puesto que

$$\begin{vmatrix} 2a_{11} & a_{12} \\ a_{12} & 2a_{22} \end{vmatrix} = 4\delta \neq 0$$

* * *

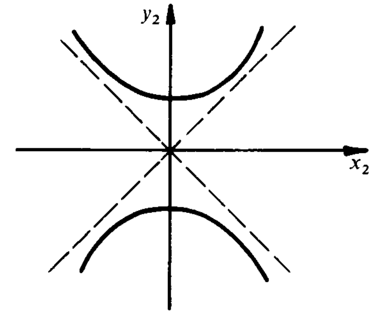
EJEMPLO C. *Determinar el centro, los ejes y las asíntotas de la hipérbola del ejemplo A de la sección* 11.8: $x^2 + 2xy - y^2 - 6x + 4y - 3 = 0$.

Sabemos que $\lambda_1 = \sqrt{2}$, $\lambda_2 = -\sqrt{2}$ y que la hipérbola puede escribirse de la forma

$$\frac{x_2}{-\frac{1}{2\sqrt{2}}} + \frac{y_2}{\frac{1}{2\sqrt{2}}} = 1$$

en un sistema de coordenadas en la dirección de los nuevos ejes y que se corten en su centro. Observar que $y_2$ es el eje principal, mientras que $x_2$ es el eje secundario.






Un autovector correspondiente a $\lambda_1 = \sqrt{2}$ es:

$$\begin{pmatrix} 1-\sqrt{2} & 1 \\ 1 & -1-\sqrt{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow (1-\sqrt{2})x + y = 0 \quad ; \quad \vec{u}_1 = (1, -1+\sqrt{2})$$

Un autovector correspondiente a $\lambda_2 = -\sqrt{2}$ es:

$$\begin{pmatrix} 1+\sqrt{2} & 1 \\ 1 & -1+\sqrt{2} \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow (1+\sqrt{2})x + y = 0 \quad ; \quad \vec{u}_2 = (1, -1-\sqrt{2})$$

La pendiente del eje principal es $k_2 = -1-\sqrt{2}$ y la pendiente del eje secundariao es $k_1 = -1+\sqrt{2}$.

El centro de la hipérbola es la solución del sistema

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 2x + 2y - 6 = 0 \\ \frac{\partial f}{\partial y} &= 2x - 2y + 4 = 0 \end{aligned}\right\}$$

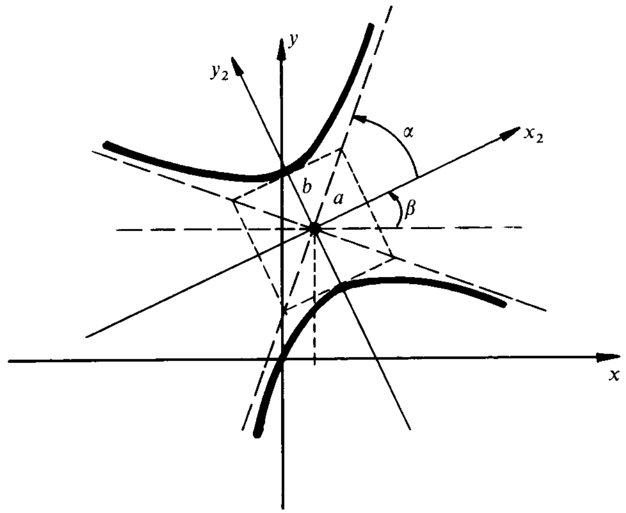
De aquí se deduce que $x = \frac{1}{2}$, $y = \frac{5}{2}$. *La ecuación del eje principal es, por tanto,*

$$y - \frac{5}{2} = -(1+\sqrt{2})\left(x - \frac{1}{2}\right)$$

y la ecuación del eje secundario es

$$y - \frac{5}{2} = (-1+\sqrt{2})\left(x - \frac{1}{2}\right)$$

La gráfica de esta hipérbola se representa en la figura adjunta.




Puesto que se trata de una hipérbola equilátera sus asíntotas son las bisectrices de sus ejes y, por tanto, forman un ángulo de 45° con ellos. Si $\alpha$ y $\beta$ son los ángulos de la figura se tiene que

$$\operatorname{tg}(\alpha+\beta)=\frac{\operatorname{tg}\alpha+\operatorname{tg}\beta}{1-\operatorname{tg}\alpha\operatorname{tg}\beta}=\frac{-\sqrt{2}}{2+\sqrt{2}}=\frac{-2\sqrt{2}+2}{2}=1-\sqrt{2}$$

que es la pendiente de una de las asíntotas; su ecuación es

$$y-\frac{5}{2}=(1-\sqrt{2})\left(x-\frac{1}{2}\right)$$

La otra asíntota es perpendicular a esta y, por tanto, su ecuación es

$$y-\frac{5}{2}=(1+\sqrt{2})\left(x-\frac{1}{2}\right)$$

*Observación.* Las asíntotas de una hipérbola también pueden hallarse mediante procedimientos de cálculo; basta observar que las asíntotas deben ser de la forma $y-\frac{5}{2}=m\left(x-\frac{1}{2}\right)$ (ya que pasan por el centro de la hipérbola) y que deben satisfacer

$$1=\lim_{x\to\infty}\frac{\dfrac{2x+4\pm\sqrt{4x^2+16-16x+4x^2-24x-12}}{2}}{mx-\dfrac{m}{2}+\dfrac{5}{2}}=\frac{1\pm\sqrt{2}}{m}$$



Por tanto, $m=1\pm\sqrt{2}$, que es el mismo resultado que habíamos obtenido anteriormente.

* * *

## 11.10. DETERMINACIÓN DEL VERTICE Y DEL EJE DE UNA PARÁBOLA

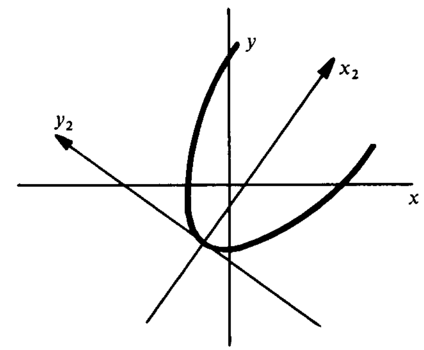
Supongamos que la ecuación

$$f(x,y)=a_{11}x^2+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=0 \tag{1}$$

determina una parábola, es decir $\Delta\neq 0$, $\delta=0$. El eje de la parábola tiene como vector director cualquier autovector correspondiente a $\lambda_1=0$; si $k_1$ es la pendiente de este eje, sólo es necesario conocer el vértice de la parábola para determinar su ecuación.

Para determinar el vértice observamos que el eje $y_2$ es tangente a la parábola en el vértice. La pendiente de la tangente a la parábola en un punto viene dada por

$$\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}y'=0.$$




Por tanto, el vértice ha de satisfacer

$$-\frac{1}{k_1}=-\frac{\dfrac{\partial f}{\partial x}}{\dfrac{\partial f}{\partial y}}. \tag{2}$$

Como el vértice debe estar en la parábola ha de satisfacer

$$f(x,y)=0 \tag{3}$$

Las ecuaciones (2) y (3) forman un sistema de dos ecuaciones con dos incógnitas que permite determinar el vértice.

EJEMPLO A. *Determinar el vértice y el eje de la parábola* $x^2-2xy+y^2+4x-6y+1=0$ (ejemplo B de la sección 11.7).

Puede comprobarse que $\delta=0$ y $\Delta\neq 0$ y los autovalores de $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ son $\lambda_1=0$, $\lambda_2=2$. Un autovector correspondiente a $\lambda_1=0$ es

$$\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad x=y \quad ; \quad \vec{u}_1=(1,\,1)$$

La pendiente de la tangente a la parábola en el vértice es $k_1=1$. Derivando implícitamente en la ecuación de la parábola obtenemos

$$(2x-2y+4)+(-2x+2y-6)y'=0$$

por tanto,

$$-\frac{2x-2y+4}{-2x+2y-6}=-1 \quad \Rightarrow \quad 2x-2y+4=-2x+2y-6$$

$$2x-2y+5=0. \tag{4}$$

El vértice es la intersección de esta recta con la parábola; como $x=y-\frac{5}{2}$, tenemos

$$\left(y-\frac{5}{2}\right)^2-2\left(y-\frac{5}{2}\right)y+y^2+4\left(y-\frac{5}{2}\right)-6y+1=0,$$

de donde se deduce que $y=-\frac{11}{8}$, $x=-\frac{31}{8}$. El eje pasa por $\left(-\frac{31}{8},-\frac{11}{8}\right)$ y tiene de pendiente 1: es la recta dada en (4).

## EJERCICIOS (SECCIONES 11.7, 11.8, 11.9 y 11.10)

En los siguientes ejercicios hacer un estudio lo más detallado posible (es decir, indicar el tipo, forma canónica, centro, ejes, ..., y hacer un esquema) de las siguientes cónicas:

**1.** $x^2+y^2+2xy-7x-5y+7=0$

**2.** $2x-2x^2+y^2+4xy-1=0$

**3.** $x^2+y^2+4x-4y-2xy=5$



**4.** $2x^2+y^2+2xy+2x+1=0$

**5.** $8x^2+17y^2+12xy-8x-16y=8$

**6.** $x^2+4y^2+4xy-6x-12y+9=0$

**7.** $x^2-y^2+2x+6y=13$

**8.** $x^2-2y^2-xy+2x+5y-3=0$

**9.** $x^2+y^2+x-1=0$

**10.** $5y^2+5x^2-1-8xy+4x-2y=0$

**11.** $y^2-4xy+4x+4x^2-2y+3=0$

**12.** Reducir la cónica $3x^2+3y^2-2xy+18x+10y+19=0$ a su forma canónica. Hallar los ejes, el centro y los focos de la cónica correspondiente y dibujarla.

**13.** Indicar el tipo y poner en forma canónica la cónica $x^2+y^2+2xy-7x-5y+7=0$.

**14.** Demostrar que las ecuaciones siguientes representan hipérbolas y encontrar sus asíntotas:

*a*) $x^2+4xy+y^2=7$

*b*) $11x^2-24xy+4y^2+6x+8y=-15$

**15.** Demostrar que las ecuaciones siguientes representan parábolas y encontrar su eje y su vértice:

*a*) $16x^2-24xy+9y^2-30x-40y=0$

*b*) $16x^2-24xy+9y^2-30x-40y=5$

**16.** Clasificar, para distintos valores de los parámetros $\alpha$, $\beta(\alpha;\beta\in\mathbb{R})$, las cónicas que admiten por ecuaciones

$$\alpha x^2+2\beta xy+\alpha y^2+(\alpha+\beta)(x+y)+1=0$$

## BIOGRAFIAS

**Apolonio de Perga** nació en Perga (actualmente en Turquía) en el año 262 a. de C. y murió en Alejandría (Egipto) alrededor del año 190 a. de C. Entre sus contemporáneos era conocido como «El Gran Geómetra», cuyo tratado *Cónicas* (sobre las secciones cónicas) es uno de los más impresionantes trabajos científicos de las civilizaciones antiguas. Muchos de sus escritos se han perdido, aunque algunos de sus títulos y un breve resumen de su contenido se conservan debido a otros escritores, especialmente Pappus de Alejandría.

En su juventud, Apolonio estudió en Alejandría, el centro del saber científico en su época, bajo discípulos de Euclides y más tarde enseñó en la universidad de la misma ciudad. Visitó Pérgamo, capital del reino Heleno en la provincia de Anatolia, con motivo de la apertura de una universidad y una biblioteca similares a las de Alejandría. Fue en Pérgamo donde escribió la primera edición de las *Cónicas*.

Los cuatro primeros libros de las *Cónicas* han llegado hasta nosotros en su lengua original (griego) y los tres siguientes traducidos al árabe. El libro octavo se ha perdido. El grado de originalidad de las *Cónicas* puede juzgarse por el prefacio del tratado. De acuerdo con Apolonio los libros del 1 al 4 contienen un desarrollo sistemático de las principales propiedades de las cónicas, muchas de las cuales eran conocidas por Euclides y Maneachmus. El resto de los libros conocidos son originales y se refieren a investigaciones sobre un problema particular.

Uno de los trabajos de Apolonio, *Sobre espejos*, cuya existencia es conocida a través de otros escritores de la antigüedad, demuestra que los rayos paralelos que se reflejan en un espejo esférico no pasan por el centro de la esfera, como se creía antes, y estudia las propiedades de los espejos parabólicos. En otro de sus trabajos, *El método rápido*, se asegura que Apolonio calculó límites más cercanos al valor de $\pi$ que $3\frac{1}{7}$ y $3\frac{10}{71}$, que eran los calculados por Arquímedes.

**Baise Pascal** nació el 19 de junio de 1623 en Clermont-Ferrand (Francia) donde su padre era magistrado. Su madre murió en 1626 y desde entonces su padre, que era también bastante conocido como matemático, se dedicó a la educación de sus hijos, Blaise y Jaqueline. Jaqueline mostró un talento precoz en el terreno literario y Blaise en el de las matemáticas. En 1640 publicó un trabajo sobre las secciones cónicas, *Essai pour les coniques* que despertó la envidia de personajes tan famosos como René Descartes (1596-1650), quien nunca pudo creer que aquel trabajo fuera obra de un joven.





Su genio brilló a gran altura en otros campos; asentó los fundamentos de la moderna teoría del cálculo de probabilidades, tal como se conoce actualmente; alrededor de 1646 comenzó a interesarse por los principios de la hidrostática, la presión atmosférica y el vacío, realizando experimentos que confirmaron las teorías de Galileo y Torricelli en estos campos.

En 1639 su padre fue nombrado administrador en Roven. Entre 1642 y 1644 B. Pascal pensó y construyó una máquina para ayudar a su padre en el cálculo de los impuestos. Esta máquina fue considerada por los contemporáneos de Pascal como su mayor legado para la posteridad; de hecho, esta máquina puede ser considerada como el primer calculador digital.

Sin embargo, Pascal se hizo también famoso como filósofo de la religión y como escritor, ocupando así un lugar único en la historia del pensamiento. En 1646 Pascal comenzó a interesarse por la religión, como consecuencia de una larga enfermedad de su padre. En 1654, después de una honda experiencia religiosa, que él llamó su «noche de fuego» decidió dedicarse enteramente a una vida de oración. Sus obras más importantes en este aspecto son *Las cortes provinciales*, de gran influencia en su tiempo, y sobre todo los fragmentos hallados a su muerte de la gran obra que planeaba en forma de una «Apología de la Religión Cristiana».

## CAPITULO 12

# FORMAS BILINEALES Y CUADRATICAS



### 12.1. DEFINICIONES

En el capítulo anterior hemos estudiado las curvas de segundo grado; la parte principal de estas cuevas puede escribirse de la forma

$$P(x, y)=(x, y)A\begin{pmatrix}x\\y\end{pmatrix}$$

donde $A$ es una matriz simétrica de orden 2; utilizando el producto escalar usual en $\mathbb{R}^2$, la expresión anterior puede escribirse de la forma

$$P(x, y)=\left(\begin{pmatrix}x\\y\end{pmatrix}, A\begin{pmatrix}x\\y\end{pmatrix}\right) \tag{1}$$

con $(x, y)\in\mathbb{R}^2$ La generalización de la expresión (1) a $\mathbb{R}^n$ puede escribirse de la forma

$$P(x)=\left(\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}, A\begin{pmatrix}x_1\\\vdots\\x_n\end{pmatrix}\right) \tag{2}$$

donde $x=\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ y $A$ es una matriz simétrica de orden $n$. Cuando $n=3$ se obtienen superficies de segundo grado, que serán estudiadas en el próximo capítulo.

Las expresiones (1) y (2) son ejemplos de *formas cuadráticas*; éstas, a su vez, son casos particulares de las *formas bilineales*.

DEFINICIÓN 1 (*Forma bilineal*)

Sea $V$ un espacio vectorial sobre un cuerpo $\mathbb{K}(=\mathbb{R}$ ó $\mathbb{C})$; una aplicación $A$: $V \times V \mapsto K$ tal que $A(\vec{x}, \vec{y})$ es lineal en la variable $\vec{x}$, dejando $\vec{y}$ fija, y $A(\vec{x}, \vec{y})$ es lineal en la segunda variable $\vec{y}$, dejando $\vec{x}$ fija, recibe el nombre de *forma bilineal*.

De la definición 1 se deduce que $A$ es una forma bilineal si y solo si se satisfacen las siguientes propiedades:

*a)* $A(\vec{x}_1+\vec{x}_2, \vec{y})=A(\vec{x}_1, \vec{y})+A(\vec{x}_2, \vec{y})$ , $\vec{x}_1, \vec{x}_2, \vec{y} \in V$
*b)* $A(\alpha\vec{x}, \vec{y})=\alpha A(\vec{x}, \vec{y})$ , $\alpha \in \mathbb{K}, \vec{x}, \vec{y} \in V$
*c)* $A(\vec{x}, \vec{y}_1+\vec{y}_2)=A(\vec{x}, \vec{y}_1)+A(\vec{x}, \vec{y}_2)$ , $\vec{x}, \vec{y}_1, \vec{y}_2 \in V$
*d)* $A(\vec{x}, \beta\vec{y})=\beta A(\vec{x}, \vec{y})$ , $\beta \in \mathbb{K}, \vec{x}, \vec{y} \in V$

EJEMPLO A. Si $V$ es un espacio vectorial de dimensión $n$ sobre un cuerpo $\mathbb{K}$ y $\{\vec{e}_1, ..., \vec{e}_n\}$ es una base de $V$, la aplicación

$$A(\vec{x}, \vec{y})=\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_iy_j$$

donde $a_{ij} \in \mathbb{K}$, i, j$=1, ..., n$, $\vec{x}=\sum_{j=1}^{n} x_j\vec{e}_j \in V$ e $y=\sum_{j=1}^{n} y_j\vec{e}_j \in V$, es una forma bilineal como puede comprobarse fácilmente.

EJEMPLO B. En un espacio euclídeo $E$ con producto escalar (,) la aplicación

$$A(\vec{x}, \vec{y})=(\vec{x}, \vec{y})$$

es una forma bilineal, ya que el producto escalar es lineal en las variables $\vec{x}$ e $\vec{y}$. De forma más general, si $\mathscr{A}$ es una aplicación lineal de $E$ en $E$, la aplicación

$$A(\vec{x}, \vec{y})=(\vec{x}, \mathscr{A}\vec{y})$$

es una forma bilineal.

* * *



Sea $A$ una forma bilineal en un espacio vectorial $V$ y sea $\{\vec{e}_1, ..., \vec{e}_n\}$ una base de $V$; si $\vec{x}=\sum_{i=1}^{n} x_i\vec{e}_j \in V$ e $y=\sum_{j=1}^{n} y_i\vec{e}_j \in V$ de la definición de forma bilineal se deduce que

$$A(\vec{x}, \vec{y})=A\left(\sum_{i=1}^{n} x_i\vec{e}_i, \sum_{j=1}^{n} y_j\vec{e}_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n} x_iy_jA(\vec{e}_i, \vec{e}_j)$$

Los «números» $a_{ij}=A(\vec{e}_i, \vec{e}_j)$ forman una matriz cuadrada

$$A=\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \tag{3}$$

que recibe el nombre de *matriz de la forma bilineal $A$ en la base* $\{\vec{e}_1,..., \vec{e}_n\}$.

Nos preguntamos ahora cómo varía la matriz de una forma bilineal al *cambiar de base*. Sea $A(a_{ij})_{i,j=1}^n$ la matriz de una forma bilineal $A$ con respecto a la base $\{\vec{e}_1, ...,\vec{e}_n\}$ y $B=(b_{ij})_{i,j=1}^n$ la matriz de $A$ respecto de la base $\{\vec{u}_1, ..., \vec{u}_n\}$; si $C=(C_{ij})_{i,j=1}^n$ es la matriz del cambio de base de la antigua a la nueva, es decir:

$$\vec{u}_i=c_{1i}\vec{e}_1+\cdots+c_{ni}\vec{e}_n$$

$i=1, 2, ..., n$, tenemos que

$$b_{ij}=A(\vec{u}_i, \vec{u}_j)=A(c_{1i}\vec{e}_1+\cdots+c_{ni}\vec{e}_n, c_{1j}\vec{e}_1+\cdots+c_{nj}\vec{e}_n)=$$
$$=\sum_{k,l=1}^{n} c_{ki}c_{lj}A(\vec{e}_k, \vec{e}_l)=\sum_{k,l=1}^{n} c_{ki}a_{kl}c_{lj}=\sum_{k=1}^{n} c_{ki}\left(\sum_{l=1}^{n} a_{kl}c_{lj}\right)$$

La expresión $\sum_{l=1}^{n} a_{kl}c_{lj}$ es el término que ocupa el lugar $(k, j)$ de la matriz $AC$ y, por tanto, la última expresión de la derecha es el término que ocupa el lugar $(i, j)$ de la matriz $C^t(AC)$. Por tanto,

$$B=C^tAC$$

que es la expresión del *cambio de base* buscada.

Puesto que $C$ es una matriz no degenerada, es decir, $\det C \neq 0$, el rango de la matriz $B$ coincide con el rango de la matriz $A$; a este rango se le denomina el *rango de la forma bilineal* y queda probado, por tanto, que el rango de una forma bilineal no depende de la base elegida para representar su matriz.

* * *

DEFINICIÓN 2 (*Forma bilineal simétrica*)

Una forma bilineal $A$ en un espacio vectorial $V$ se dice *simétrica* si $A(\vec{x}, \vec{y}) = A(\vec{y}, \vec{x})$ para todo $\vec{x}, \vec{y} \in V$

Si $A$ es una forma bilineal simétrica, $A(\vec{e}_i, \vec{e}_j) = A(\vec{e}_j, \vec{e}_i)$ y, por tanto, la matriz de $A$ en cualquier base $\{\vec{e}_1, ..., \vec{e}_n\}$ es una matriz simétrica. Recíprocamente, si la matriz de una forma bilineal es simétrica en alguna base $\{\vec{e}_1, ..., \vec{e}_n\}$, la forma bilineal es simétrica ya que

$$A(\vec{y}, \vec{x}) = \sum_{i,j=1}^{n} a_{ij} y_i x_j = \sum_{i,j=1}^{n} a_{ji} x_j y_i = A(\vec{x}, \vec{y}).$$

DEFINICIÓN 3 (*Forma bilineal antisimétrica*)

Una forma bilineal $A$ en un espacio vectorial $V$ se dice *antisimétrica* si $A(\vec{x}, \vec{y}) = -A(\vec{y}, \vec{x})$ para todo $\vec{x}, \vec{y} \in V$

Si $A$ es una forma bilineal antisimétrica, $A(\vec{e}_i, \vec{e}_j) = -A(\vec{e}_j, \vec{e}_i)$ y, por tanto, la matriz de $A$ en *cualquier base* $\{\vec{e}_1, ..., \vec{e}_n\}$ satisface $a_{ij} = -a_{ji}$; en particular $a_{ii} = -a_{ii}$, $i = 1, ..., n$ y, por tanto, los elementos de la diagonal de la matriz de una forma bilineal antisimétrica deben ser nulos. Por tanto, la matriz de $A$ se escribe de la forma

$$A = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & & & \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix}.$$

Las formas bilineales simétricas junto con las antisimétricas originan todas las formas bilineales posibles; el resultado se describe en la siguiente proposición:

PROPOSICIÓN 4

Toda forma bilineal puede ser representada como la suma de una forma bilineal simétrica y una forma bilineal antisimétrica.

*Demostración.* Sea $A(\vec{x}, \vec{y})$ la forma bilineal; tomar $B(\vec{x}, \vec{y}) = A(\vec{x}, \vec{y}) + A(\vec{y}, \vec{x})$ y $C(\vec{x}, \vec{y}) = A(\vec{x}, \vec{y}) - A(\vec{y}, \vec{x})$. Sumando ambas expresiones se tiene que

$$B(\vec{x}, \vec{y}) + C(\vec{x}, \vec{y}) = 2A(\vec{x}, \vec{y})$$

y, por tanto,

$$A(\vec{x}, \vec{y}) = \frac{1}{2}B(\vec{x}, \vec{y}) + \frac{1}{2}C(\vec{x}, \vec{y})$$



La demostración se concluye observando que $B$ es una forma bilineal simétrica y $C$ es una forma bilineal antisimétrica. ■

* * *

Una *forma cuadrática* en un espacio vectorial $V$ es toda aplicación $A(\vec{x}, \vec{x})$ de $V$ en $\mathbb{K}(\mathbb{R}$ ó $\mathbb{C})$ que se obtiene sustituyendo $\vec{y}$ por $\vec{x}$ en una forma bilineal $A(\vec{x}, \vec{y})$ sobre $V$.

EJEMPLO C. $a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$ es una forma cuadrática de $\mathbb{R}^2 (a_{ij} \in \mathbb{R})$ ya que si tomamos

$$A\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) = a_{11}x_1y_1 + \frac{1}{2}a_{12}x_1y_2 + \frac{1}{2}a_{12}x_2y_1 + a_{22}x_2y_2 \tag{3}$$

se tiene que

$$A\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}\right) = a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$$

EJEMPLO D. $a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{23}x_2x_3$ es una forma cuadrática en $\mathbb{C}^3$ $(a_{ij} \in \mathbb{R})$ ya que basta tomar

$$A\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}\right) = a_{11}x_1y_1 + a_{22}x_2y_2 + a_{33}x_3y_3 + \frac{1}{2}a_{12}x_1y_2 + \frac{1}{2}a_{12}x_2y_1 +$$

$$+ \frac{1}{2}a_{13}x_1y_3 + \frac{1}{2}a_{13}x_3y_1 + \frac{1}{2}a_{23}x_2y_3 + \frac{1}{2}a_{23}x_3y_2$$

* * *

En general toda expresión de la forma $\sum_{j=1}^{n} \sum_{i \le j} a_{ij}x_ix_j$ en un espacio vectorial define una forma cuadrática ya que basta tomar la forma bilineal

$$A(\vec{x}, \vec{y}) = \sum_{i=1}^{n} a_{ii}x_iy_i + \sum_{j=1}^{n} \sum_{i<j} \frac{a_{ij}}{2} x_iy_j + \sum_{j=1}^{n} \sum_{i<j} \frac{a_{ij}}{2} x_jy_i \tag{4}$$

donde $\vec{x} = \sum_{i=1}^{n} x_i\vec{e}_i$, $\vec{y} = \sum_{j=1}^{n} y_j\vec{e}_j$, para obtener que

$$A(\vec{x}, \vec{x}) = \sum_{j=1}^{n} \sum_{i \le j} a_{ij}x_ix_j$$

Recíprocamente, toda *forma cuadrática* tiene una expresión de la forma $\sum_{j=1}^{n} \sum_{i \leq j} b_{ij} x_i x_j$ en una cierta base. En efecto, si $A(\vec{x}, \vec{x})$ proviene de la forma bilineal $A(\vec{x}, \vec{y})$ de matriz $A=(a_{ij})_{i,j=1}^{n}$ en la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$ se tiene que

$$A(\vec{x}, \vec{x})=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j a_{ij} \quad , \quad x=\sum_{i=1}^{n} x_i \vec{e}_i$$

Agrupando los términos $x_i x_j$ y $x_j x_i$ se tiene que

$$A(\vec{x}, \vec{x})=\sum_{i=1}^{n} \sum_{i \leq j} (a_{ij}+a_{ji}) x_i x_j=\sum_{i=1}^{n} \sum_{i \leq j} b_{ij} x_i x_j$$

que era lo que queríamos demostrar.

Toda forma bilineal determina una forma cuadrática (lo cual se deduce de la misma definición), pero una forma cuadrática puede ser determinada por varias formas bilineales. En efecto, la forma bilineal

$$A_1\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right)=a_{11}x_1y_1+\frac{1}{3}a_{12}x_1y_2+\frac{2}{3}a_{12}x_2y_1+a_{22}x_2y_2$$

determina la forma cuadrática del ejemplo C, pero es diferente, en general, de la forma bilineal dada en (3).

Sin embargo, si la forma bilineal que genera la forma cuadrática se supone que es simétrica, la forma bilineal queda *unívocamente determinada.* Para probar este resultado observar que

$$A(\vec{x}+\vec{y}, \vec{x}+\vec{y})=A(\vec{x}, \vec{x})+A(\vec{x}, \vec{y})+A(\vec{y}, \vec{x})+A(\vec{y}, \vec{y})$$

por la definición de forma bilineal; si $A$ es simétrica tenemos

$$A(\vec{x}, \vec{y})=\frac{1}{2}[A(\vec{x}, \vec{y})+A(\vec{y}, \vec{x})]=\frac{1}{2}[A(\vec{x}+\vec{y}, \vec{x}+\vec{y})-A(\vec{x}, \vec{x})-A(\vec{y}, \vec{y})]$$

que prueba el resultado deseado.

Se denomina *matriz de una forma cuadrática* en una base $\{\vec{e}_1, \ldots, \vec{e}_n\}$ a la matriz de la forma bilineal simétrica que la determina, en la misma base $\{\vec{e}_1, \ldots, \vec{e}_n\}$. Como la forma bilineal dada en (4) es simétrica y determina la forma cuadrática

$$A(\vec{x}, \vec{x})=\sum_{j=1}^{n} \sum_{i \leq j} a_{ij} x_i x_j$$



deducimos del resultado sobre unicidad anteriormente demostrado que la matriz de la forma cuadrática $A(\vec{x}, \vec{x})$ en la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$ es

$$A=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \cdots & \frac{a_{1n}}{2} \\ \frac{a_{12}}{2} & a_{22} & \cdots & \frac{a_{2n}}{2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{a_{1n}}{2} & \frac{a_{2n}}{2} & \cdots & a_{nn} \end{pmatrix}$$

* * *

## 12.2. FORMAS BILINEALES Y CUADRÁTICAS EN UN ESPACIO EUCLIDEO

Hemos visto en el ejemplo B de la sección anterior que si $\mathscr{A}: E \mapsto E$ es una aplicación lineal en un espacio euclídeo,

$$A(\vec{x}, \vec{y})=(\vec{x}, \mathscr{A}\vec{y})$$

es una forma bilineal. De hecho, todas las formas bilineales en un espacio euclídeo son de esta forma, como se muestra en la siguiente proposición:

**PROPOSICIÓN 1**

Sea $A$: $E \times E \mapsto \mathbb{R}$ una forma bilineal en un espacio euclídeo $E$ con producto escalar ( , ). Existe una aplicación lineal $\mathscr{A}$: $E \mapsto E$ tal que $A(\vec{x}, \vec{y})=(\vec{x}, \mathscr{A}\vec{y})$. Además, las matrices de $A$ y $\mathscr{A}$ en una misma base ortonormal coinciden.

*Demostración.* Si $A=(a_{ij})_{i,j=1}^{n}$ es la matriz de la forma bilineal $A$ en una base ortonormal $\{\vec{e}_1, \ldots, \vec{e}_n\}$, podemos escribir

$$A(\vec{x}, \vec{y})=\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j \quad , \quad x=\sum_{i=1}^{n} x_i \vec{e}_i \quad , \quad y=\sum_{j=1}^{n} y_j \vec{e}_j$$

Sea $\mathscr{A}$ la aplicación lineal de $E$ en $E$ que tiene como matriz $A$ en la base $\{\vec{e}_1, \ldots, \vec{e}_n\}$, es decir, $\mathscr{A}(\vec{e}_j)=a_{1j}\vec{e}_1+\cdots+a_{nj}\vec{e}_n$. Se tiene que

$$\begin{aligned}(\vec{x}, \mathscr{A}(\vec{y}))&=\left(\sum_{i=1}^{n} x_i \vec{e}_i \quad , \quad \mathscr{A}\left(\sum_{j=1}^{n} y_j \vec{e}_j\right)\right)=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\vec{e}_i, \mathscr{A}(\vec{e}_j)) \\ &=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j (\vec{e}_i, a_{1j}\vec{e}_1+\cdots+a_{ij}\vec{e}_i+\cdots+a_{nj}\vec{e}_n)= \\ &=\sum_{i=1}^{n} \sum_{j=1}^{n} x_i y_j a_{ij}\end{aligned}$$

donde la última igualdad se debe a que $\{\vec{e}_1, ..., \vec{e}_n\}$ es una base ortonormal. Por tanto, $A(\vec{x}, \vec{y})=(\vec{x}, \mathscr{A}(\vec{y}))$ que era lo que queríamos probar. ■

TEOREMA 2 (*Reducción de una forma cuadrática en un espacio euclídeo a su forma canónica*)

> Dada una forma cuadrática $A(\vec{x}, \vec{x})$ en un espacio euclídeo $E$ existen números reales $\lambda_1, ..., \lambda_n$ y una base ortonormal $\{\vec{u}_1, ..., \vec{u}_n\}$ de $E$ tal que para todo
>
> $$\vec{x}=\sum_{i=1}^{n} x_i\vec{u}_i \in E$$
>
> se tiene que
>
> $$A(\vec{x}, \vec{x})=\lambda_1 x_1^2+\lambda_2 x_2^2+\cdots+\lambda_n x_n^2$$

*Demostración.* Según sabemos de la sección anterior $A(\vec{x}, \vec{x})$ determina unívocamente una forma bilineal simétrica $A(\vec{x}, \vec{y})$; por la proposición 1, $A(\vec{x}, \vec{y})=(\vec{x}, \mathscr{A}(\vec{y}))$, donde $\mathscr{A}$ es una aplicación lineal que tiene la misma matriz que $A(\vec{x}, \vec{y})$ en una base ortonormal; por tanto, la matriz de $\mathscr{A}$ es simétrica en esta base. Por los resultados del capítulo 8, sección 6, sabemos que toda matriz simétrica es diagonalizable en una base ortonormal y, por tanto, existe una base ortonormal $\{\vec{u}_1, ..., \vec{u}_n\}$ tal que

$$\mathscr{A}(\vec{x})=\mathscr{A}\left(\sum_{j=1}^{n} x_j\vec{u}_j\right)=\sum_{j=1}^{n} x_j\mathscr{A}(\vec{u}_j)=\sum_{i=1}^{n} x_j\lambda_j\vec{u}_j$$

donde $\lambda_j$, $j=1, ..., n$ son los autovalores de $\mathscr{A}$. Por tanto,

$$A(\vec{x}, \vec{x})=(\vec{x}, \mathscr{A}(\vec{x}))=\left(\sum_{i=1}^{n} x_i\vec{u}_i, \sum_{j=1}^{n} x_j\lambda_j\vec{u}_j\right)=$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n} x_i x_j\lambda_j(\vec{u}_i, \vec{u}_j)=\sum_{i=1}^{n} x_i^2\lambda_i=\lambda_1 x_1^2+\cdots+\lambda_n x_n^2$$ ■

*Observación.* Los números $\lambda_j$ son los autovalores de la matriz de $A(x, x)$ y los $\vec{u}_j$ son autovectores correspondientes a $\lambda_j$.

Supongamos que una forma cuadrática $A(\vec{x}, \vec{x})$ puede escribirse de la forma

$$A(\vec{x}, \vec{x})=\sum_{j=1}^{n} a_j x_j^2=a_1x_1^2+\cdots+a_nx_n^2 \tag{1}$$

en una cierta base del $\mathbb{K}$-espacio vectorial $V$ en que está definida, donde $a_j \in \mathbb{K}$, $j=1, 2, ..., n$.

La expresión (1) recibe el nombre de *forma canónica* de la forma cuadrática $A(\vec{x}, \vec{x})$.



En el teorema 2 que acabamos de demostrar se prueba que si una forma cuadrática está definida en un espacio euclídeo, puede reducirse a su forma canónica en una base ortonormal. Realizamos a continuación algunos ejemplos.

EJEMPLO A. *Reducir la forma cuadrática* $A(\vec{x}, \vec{x})=x_1^2+x_2^2+x_3^2-4x_1x_3$ *a una forma canónica (en* $\mathbb{R}^3$*) mediante una base ortonormal.*

La matriz de $A(\vec{x}, \vec{x})$ es

$$A=\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix};$$

sus autovalores son: $\lambda_1=1$, $\lambda_2=-1$, $\lambda_3=3$. Por tanto

$$A(\vec{x}, \vec{x})=x_1^2-x_2^2+3x_3^2$$

en una base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ formada por autovectores, donde $\vec{x}=\vec{x}_1\vec{u}_1+x_2\vec{u}_2+x_3\vec{u}_3$.

* * *

La forma de reducir una forma cuadrática a una expresión del tipo $a_1x_1^2+\cdots+a_nx_n^2$ no es única si se permite realizar cambios de base que no sean ortonormales. Como ilustración tomar el ejemplo A; completando cuadrados con los términos $x_1^2-4x_1x_3$ se tiene

$$A(\vec{x}, \vec{x})=(x_1-2x_3)^2+x_2^2-3x_3^2$$

y, por tanto, el cambio de base $y_1=x_1-2x_3$, $y_2=x_2$, $y_3=x_3$ reduce la forma cuadrática del ejemplo A a

$$y_1^2+y_2^2-3y_3^2$$

Observar que este cambio de base no transforma una base ortonormal en otra base ortonormal.

EJEMPLO B. *Reducir la forma cuadrática A dada por* $2xy-2xz$ *con respecto a la base canónica de* $\mathbb{R}^3$ *a una forma canónica.*

Escribiendo $x=x_1+y_1$, $y=x_1-y_1$, $z_1=z$, lo cual es un cambio de base ya que

$$\begin{vmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{vmatrix}=-2\neq 0$$

tenemos

$$2xy-2xz=2(x_1+y_1)(x_1-y_1)-2(x_1+y_1)z_1=2x_1^2-2y_1^2-$$
$$-2x_1z_1-2y_1z_1.$$

Completando cuadrados se tiene

$$2xy-2xz=2\left(x_1-\frac{z_1}{2}\right)-2\left(y_1+\frac{z_1}{2}\right)$$

con lo cual el cambio de coordenadas $x_2=x_1-\frac{z_1}{2}$, $y_2=y_1+\frac{z_1}{2}$, $z_2=z_1$ nos permite escribir

$$A=2x_2^2-2y_2^2$$

Si deseamos obtener de $A$ una suma de cuadrados mediante un cambio de base ortogonal no tenemos más que calcular los autovalores de la matriz

$$A=\begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix}$$

que son $\lambda_1=0$, $\lambda_2=\sqrt{2}$, $\lambda_3=-\sqrt{2}$; por tanto

$$A=\sqrt{2}y^2-\sqrt{2}z^2$$

en la base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ formada por autovectores de $A$. Estos autovectores pueden ser $\vec{u}_1=\frac{1}{\sqrt{2}}(0, 1, 1)$, $\vec{u}_2=\frac{1}{2}(\sqrt{2}, 1, -1)$, $\vec{u}_3=\frac{1}{2}(\sqrt{2}, -1, 1)$.

*Observación.* En los ejemplos anteriores se han utilizado dos métodos para reducir una forma cuadrática a una forma canónica; el primero de ellos está basado en el cálculo de autovalores de la matriz de $A$; sobre el segundo observaremos que se basa en los siguientes hechos:

*a*) Completar cuadrados si es posible (ver ejemplo A).

*b*) Si no existen cuadrados que se puedan completar (ver ejemplo B) es necesario hacer un cambio de la forma $y_j=x_j+x_k$, $y_k=x_j-x_k$ para introducirlos, y a continuación completar cuadrados.

Se deja para el lector la práctica de este segundo método.

## 12.3. LEY DE INERCIA DE LAS FORMAS CUADRÁTICAS

En la sección anterior hemos visto que toda forma cuadrática en un espacio euclídeo de dimensión finita puede escribirse de la forma

$$A(\vec{x}, \vec{x})=a_1x_1^2+\cdots+a_n\,x_n^2 \tag{1}$$



en una cierta base, donde $a_j\in\mathbb{R}$, $j=1, 2, ..., n$; la expresión (1) recibe el nombre de forma canónica de la forma cuadrática dada.

La forma canónica de una forma cuadrática no es única, puesto que depende de la forma de obtenerla, según se ha observado en la sección anterior. Sin embargo, el número de términos positivos $a_j$ que aparece en (1) y el número de términos negativos de la misma expresión permanecen invariantes para una misma forma cuadrática. Este hecho puede observarse en los ejemplos de la sección anterior. El enunciado preciso y la demostración de este resultado la damos a continuación.

TEOREMA (*Ley de inercia de las formas cuadráticas*)

Sea $A(\vec{x}, \vec{x})$ una forma cuadrática en un espacio euclídeo $E$ de dimensión $n$. Si $A(\vec{x}, \vec{x})$ se escribe de la forma

$$A(\vec{x}, \vec{x})=a_1x_1^2+a_2x_2^2+\cdots+a_mx_m^2 \tag{2}$$

en la base $F=\{\vec{f}_1, ..., \vec{f}_n\}$, con $a_j\in\mathbb{R}$, $a_j\neq 0$, $j=1, 2, ..., m$, y también se escribe de la forma

$$A(\vec{x}, \vec{x})=b_1y_1^2+b_2y_2^2+\cdots+b_qy_q^2 \tag{3}$$

en la base $G\{\vec{g}_1, ..., \vec{g}_n\}$, con $b_j\neq 0$, $b_j\in\mathbb{R}$, $j=1, 2, ..., q$, *se tiene que* $m=q$ $=\text{rango}(A)$ y el número de coeficientes positivos y negativos de (2) y (3) coinciden.

*Demostración.* Escribamos, más explícitamente,

$$A(\vec{x}, \vec{x})=\alpha_1x_1^2+\cdots+\alpha_kx_k^2-\alpha_{k+1}x_{k+1}^2-\cdots-\alpha_{mm}^2 \tag{4}$$

en la base $F$, con $\alpha_j>0$ , $j=1, 2, ..., m$, y

$$A(\vec{x}, \vec{x})=\beta_1y_1^2+\cdots+\beta_py_p^2-\beta_{p+1}y_{p+1}^2-\cdots-\beta_qy_q^2 \tag{5}$$

en la base $G$, con $\beta_j>0$, $j=1, 2, ..., q$. Hemos de demostrar que $k=p$ y $m=q$.

Sea $E_1$ el subespacio vectorial de $E$ generado por los vectores $\vec{f}_1, \vec{f}_2, ..., \vec{f}_k$ y $E_2$ el subespacio vectorial de $E$ generado por los vectores $\vec{g}_{p+1}, ..., \vec{g}_n$. Si suponemos que $k>p$ se tiene que

$$\dim(E_1)+\dim(E_2)=k+(n-p)>n=\dim(E)$$

Puesto que $\dim(E_1)+\dim(E_2)=\dim(E_1+E_2)+\dim(E_1\cap E_2)$ (proposición 2, sección 5.4, capítulo 5) hemos de tener $E_1\cap E_2\neq\phi$.

Sea $\vec{z} \in E_1 \cap E_2$ y $\vec{z} \neq \vec{0}$; puesto que $\vec{z} \in E_1 \cap E_2$ podemos escribir

$$\vec{z} = a_1\vec{f}_1 + \cdots + a_k\vec{f}_k \quad , \quad a_j \in \mathbb{R},\ j=1, \ldots, k$$

y

$$\vec{z} = b_{p+1}\vec{g}_{p+1} + \cdots + b_n\vec{g}_n \quad , \quad b_j \in \mathbb{R},\ j=p+1, \ldots, n$$

Para el vector $\vec{z}$ las expresiones (4) y (5) producen

$$A(\vec{z}, \vec{z}) = \alpha_1 a_1^2 + \cdots + \alpha_k a_k^2 > 0$$

ya que alguno de los $a_j$ debe ser no nulo, y

$$A(\vec{z}, \vec{z}) = -\beta_{p+1} b_{p+1}^2 - \cdots - \beta_q b_q^2 \leqq 0$$

(observar que la última expresión puede ser cero ya que $q$ puede ser menor que $n$).

En cualquier caso se llega a una contradición, que se ha producido por el hecho de suponer que $k > p$; por tanto, $k \leqq p$.

Puesto que el papel que juegan las expresiones (4) y (5) en el razonamiento anterior es simétrico, se deduce, de manera similar, la desigualdad $p \leqq k$, con lo cual se tiene $p = k$.

La igualdad $m = q = \text{rango}(A)$ se debe a que el rango de una forma cuadrática es invariante mediante cambios de base. ■

El número de términos que aparecen en la forma canónica de una forma cuadrática $A(\vec{x}, \vec{x})$ recibe el nombre de *índice de inercia* de la forma cuadrática; el número de términos positivos se llama *índice de inercia positivo* de la forma cuadrática y el número de términos negativos recibe el nombre de *índice de inercia negativo* de la forma cuadrática.

EJEMPLO A. Tratamos de encontrar los índices de inercia de la forma cuadrática dada por

$$A(\vec{x}, \vec{x}) = x_1x_2 + x_1x_3 + x_1x_4 + x_2x_3 + x_2x_4 + x_3x_4$$

El cambio de variable $x_1 = y_1 + y_2$, $x_2 = y_1 - y_2$, $x_3 = y_3 + y_4$, $x_4 = y_3 - y_4$ (¡comprobar que es un cambio de base!) transforma la expresión anterior en

$$y_1^2 - y_2^2 + (y_1+y_2)(y_3+y_4) + (y_1+y_2)(y_3-y_4) + (y_1-y_2)(y_3+y_4) + (y_1-y_2)(y_3-y_4) +$$
$$+ y_3^2 - y_4^2 = y_1^2 - y_2^2 + 4y_1y_3 + y_3^2 - y_4^2 = (y_1+2y_3)^2 - y_2^2 - 3y_3^2 - y_4^2$$

Escribiendo $z_1 = y_1 + 2y_3$, $z_2 = y_2$, $z_3 = y_3$, $z_4 = y_4$, que es de nuevo un cambio de base, se tiene:

$$A = z_1^2 - z_2^2 - 3z_3^2 - z_4^2$$

Por tanto, el índice de inercia positivo de $A$ es 1 y su índice de inercia negativo es 3.

* * *



Si una forma canónica de la forma cuadrática $A(\vec{x}, \vec{x})$ es

$$A(\vec{x}, \vec{x}) = a_1x_1^2 + \cdots + a_kx_k^2 - a_{k+1}x_{k+1}^2 - \cdots - a_mx_m^2$$

con $a_j \in \mathbb{R}$, $a_j > 0$, $j = 1, 2, \ldots, m$, el cambio de base dado por

$$y_1 = \sqrt{a_1}\,x_1, \ldots, y_k = \sqrt{a_k}\,x_k,\ y_{k+1} = \sqrt{a_{k+1}}\,x_{k+1}, \ldots, y_m = \sqrt{a_m}\,x_m,\ y_{m+1} = x_{m+1}, \ldots, y_n = x_n$$

nos permite escribir

$$A(\vec{x}, \vec{x}) = y_1^2 + \cdots + y_k^2 - y_{k+1}^2 - \cdots - y_m^2$$

que recibe el nombre de *forma normal* de la forma cuadrática dada. Observar que la forma normal de una forma cuadrática es una suma algebraica de cuadrados.

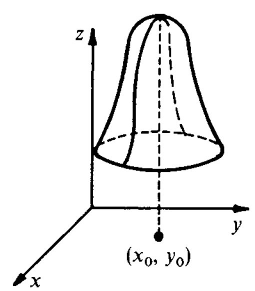
## 12.4. FORMAS CUADRATICAS DEFINIDAS. PUNTOS CRITICOS DE FUNCIONES DE VARIAS VARIABLES

La teoría de las formas cuadráticas será de gran ayuda para el estudio de las superficies de segundo grado; sin embargo, también posee otras aplicaciones, entre las que se cuentan la determinación de puntos críticos de funciones de varias variables.

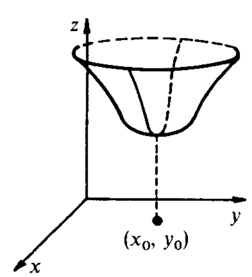
Para facilitar el razonamiento utilizaremos una función de dos variables $z = f(x, y)$, que representa una superficie en $\mathbb{R}^3$. Recordamos al lector que para una función $f$ suficientemente regular, ésta posee un *punto crítico* en $(x_0, y_0)$ si y sólo si

$$\frac{\partial f}{\partial x}(x_0, y_0) = 0 \quad , \quad \frac{\partial f}{\partial y}(x_0, y_0) = 0 \tag{1}$$

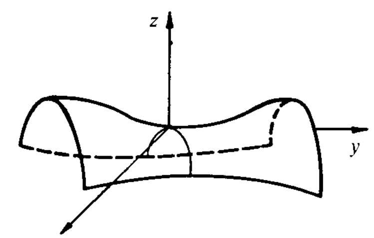
Ejemplos conocidos de puntos críticos son los *máximos, mínimos y los puntos de ensilladura* que se muestran geométricamente en los dibujos siguientes:



MAXIMO



MINIMO



PUNTO DE ENSILLADURA

Deseamos encontrar una condición que nos permita decidir entre todos los casos posibles de puntos críticos. Para ello escribimos la fórmula de Taylor de $f$ alrededor del punto $(x_0, y_0)$ hasta el término de orden 2 —estamos suponiendo que $f$ es lo suficientemente regular para que esto sea posible—:

$$f(x, y)=f(x_0, y_0)+\left[\frac{\partial f}{\partial x}(x_0, y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0, y_0)(y-y_0)\right]+$$
$$+\frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x-x_0)^2+\frac{\partial^2 f}{\partial x\partial y}(x_0, y_0)(x-x_0)(y-y_0)+\frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y-y_0)^2\right]$$
$$+R_3(x, y)$$

donde $R_3(x, y)$ es el término de error. Si $(x_0, y_0)$ es un punto crítico de $f$ la fórmula anterior puede escribirse de la forma

$$f(x, y)-f(x_0, y_0)=\frac{1}{2!}H(X, X)+R_3(x, y) \tag{2}$$

donde

$$H(X, X)=H\left(\begin{pmatrix}x-x_0\\ y-y_0\end{pmatrix}, \begin{pmatrix}x-x_0\\ y-y_0\end{pmatrix}\right)=\frac{\partial^2 f}{\partial x^2}(x_0, y_0)(x-x_0)^2$$
$$+2\frac{\partial^2 f}{\partial x\partial y}(x_0, y_0)(x-x_0)(y-y_0)+\frac{\partial^2 f}{\partial y^2}(x_0, y_0)(y-y_0)^2$$

es una forma cuadrática que posee como matriz

$$H=\begin{pmatrix}\frac{\partial^2 f}{\partial x^2}(x_0, y_0) & \frac{\partial^2 f}{\partial x\partial y}(x_0, y_0)\\ \frac{\partial^2 f}{\partial x\partial y}(x_0, y_0) & \frac{\partial^2 2}{\partial y^2}(x_0, y_0)\end{pmatrix},$$

que se denomina *la matriz hessiana* de $f$ en el punto $(x_0, y_0)$.

La función $f$ posee un máximo en $(x_0, y_0)$ si y sólo si $f(x, y)-f(x_0, y_0)\leqq 0$ para todo $(x, y)$ en un entorno de $(x_0, y_0)$; de la igualdad (2) deducimos que si $H(X, X)<0$, $(x_0, y_0)$ es un máximo ya que $R_3(x, y)$ puede hacerse tan pequeño como se quiera reduciendo el entorno alrededor de $(x_0, y_0)$. Esto sugiere la siguiente definición:

**DEFINICIÓN 1**

Una forma cuadrática $A(\vec{x}, \vec{x})$ se llama *definida negativa* si $A(\vec{x}, \vec{x})<0$ para todo $\vec{x}\neq 0$. Si $A(\vec{x}, \vec{x})\leqq 0$ para todo $\vec{x}$, A se dice *semidefinida negativa.*



Si $f$ posee un mínimo en $(x_0, y_0)$ se ha de tener $f(x, y)-f(x_0, y_0)\geqq 0$ para todo $(x, y)$ en un entorno de $(x_0, y_0)$; esto puede conseguirse en la igualdad (2) si se tiene $H(X, X)>0$ para todo $X\neq 0$. Esto sugiere la siguiente definición:

**DEFINICIÓN 2**

Una forma cuadrática $A(\vec{x}, \vec{x})$ se llama *definida positiva* si $A(\vec{x}, \vec{x})>0$ para todo $\vec{x}\neq 0$. Si $A(\vec{x}, \vec{x})\geqq 0$ para todo $\vec{x}$, A se dice *semidefinida positiva.*

Estos resultados quedan resumidos en el siguiente teorema:

**TEOREMA 1** (*Condición suficiente para la existencia de máximos y mínimos*)

Sea $f$: $\mathbb{R}^2\mapsto\mathbb{R}$ una función suficientemente regular que posee un punto crítico en $(x_0, y_0)$ *y sea* $H$ la forma cuadrática en $\mathbb{C}^2$ que tiene a

$$H=\begin{pmatrix}\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x\partial y}\\ \frac{\partial^2 f}{\partial x\partial y} & \frac{\partial^2 f}{\partial y^2}\end{pmatrix}(x_0, y_0)$$

como matriz. Entonces,

*a)* si $H$ es *definida negativa*, $f$ posee un *máximo* (relativo) en $(x_0, y_0)$, y
*b)* si $H$ es *definida positiva*, $f$ posee un *mínimo* (relativo) en $(x_0, y_0)$.

*Notas.* 1) En libros de cálculo se demuestra que si $H$ no es ni semidefinida positiva ni semidefinida negativa, $(x_0, y_0)$ es un punto de ensilladura de $f$.
2) Un criterio análogo al contenido en el teorema 1 se tiene para funciones de $n$ variables.

*Observaciones.* 1) Si $A(\vec{x}, \vec{x})=(\vec{x}, \vec{x})$, donde ( , ) representa el producto escalar en $E$, $A$ es una forma cuadrática definida positiva, ya que $(\vec{x}, \vec{x})\geqq 0$ para todo $\vec{x}\neq 0$.
2) Si una forma cuadrática $A(\vec{x}, \vec{x})$ se escribe de la forma $A(\vec{x}, \vec{x})=a_1x_1^2+\cdots+a_nx_n^2$ en una cierta base, $A$ es definida positiva si y sólo si todos los $a_j$ son positivos y es definida negativa si y sólo si todos los $a_j$ son negativos.

Para poder aplicar el teorema 1 es necesario poseer un criterio fácilmente aplicable que permita determinar si $H$ es definida positiva o es definida negativa; una posibilidad es hallar una forma canónica de $H$ y aplicar la observación 2); otra lo proporcionan los criterios de Sylvester que exponemos a continuación.

PROPOSICIÓN 2 *(Criterio de Sylvester para formas definidas positivas)*

Sea $A(\vec{x}, \vec{x})$ una forma cuadrática en un espacio euclídeo con matriz $A = (a_{ij})_{i,j=1}^{n}$ en una cierta base. Si los menores angulares de la matriz $A$ son positivos, es decir:

$$\Delta_1 = a_{11} > 0,\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{vmatrix} > 0,\ \Delta_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{vmatrix} > 0,\ \ldots,\ \Delta_n = |A| > 0$$

la forma cuadrática $A(\vec{x}, \vec{x})$ es definida positiva.

*Demostración.* Realizamos la demostración por inducción en la dimensión del espacio euclídeo $E$. Si $E$ tiene dimensión uno, la forma cuadrática puede escribirse de la forma $A(\vec{x}, \vec{x}) = a_{11}x^2$ un $a_{11} > 0$ y, por tanto, la proposición queda demostrada. Supongamos ahora que el resultado es cierto para cualquier forma cuadrática con todos sus menores angulares positivos en todo espacio euclídeo de dimensión inferior a $m$ y sea $A(\vec{x}, \vec{x})$ una forma cuadrática en un espacio euclídeo $E$ de dimensión $m$ con todos sus menores angulares positivos. Por ser los $m-1$ primeros menores angulares positivos, la hipótesis de inducción nos permite deducir que la forma cuadrática

$$B(\vec{x}', \vec{x}') = \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} a_{ij}x_i x_j$$

en el espacio euclídeo $E'$ generado por los $m-1$ primeros vectores de una base de $E$, es definida positiva; con un cambio de base adecuado en $E'$, $B(\vec{x}', \vec{x}')$ puede escribirse en forma canónica (teorema 2, sección 2) con coeficientes positivos:

$$B(\vec{x}', \vec{x}') = \lambda_1 x_1'^2 + \cdots + \lambda_{m-1} x_{m-1}'^2 \quad (\lambda_j > 0)$$

Realizando este cambio de base en $E$ y $x_m' = x_m$ la forma cuadrática $A(\vec{x}, \vec{x})$ se transforma en

$$A(\vec{x}, \vec{x}) = \lambda_1 x_1'^2 + \cdots + \lambda_{m-1} x_{m-1}'^2 + (b_{1m}x_1'x_m' + \cdots + b_{m-1,m}x_{m-1}'x_m') + a_{mm}x_m'^2$$

Completando cuadrados se obtiene:

$$A(\vec{x}, \vec{x}) = \lambda_1 \left( x_1' + \frac{b_{1m}}{2\lambda_1} x_m' \right)^2 + \cdots + \lambda_{m-1} \left( x_{m-1}' + \frac{b_{m-1,m}}{2\lambda_{m-1}} x_m' \right)^2 + b x_m'^2$$

donde

$$b = a_{mm} - \frac{b_{1m}^2}{4\lambda_1} - \cdots - \frac{b_{m-1,m}^2}{4\lambda_{m-1}}$$



Realizando el cambio de base:

$$y_1 = x_1' + \frac{b_{1m}}{2\lambda_1} x_m', \ldots, y_{m-1} = x_{m-1}' + \frac{b_{m-1,m}}{2\lambda_{m-1}} x_m', \quad y_m = x_m'$$

(observar que el determinante de esta transformación es 1) tenemos

$$A(\vec{x}, \vec{x}) = \lambda_1 y_1^2 + \cdots + \lambda_{m-1} y_{m-1}^2 + b y_m^2$$

Puesto que todos los $\lambda_j$ son positivos sólo falta demostrar que $b > 0$; para esto observar que si $C$ es la matriz del cambio de base de las coordenadas $x_j$ a las $y_j$ se tiene:

$$\begin{pmatrix} \lambda_1 & & & 0 \\ & \ddots & & \\ & & \lambda_{m-1} & \\ 0 & & & b \end{pmatrix} = C^t \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{1m} & \cdots & a_{mm} \end{pmatrix} C$$

(sección 1) y, por tanto, $\lambda_1 \times \cdots \times \lambda_{m-1} \times b = |C^t|\Delta_m|C| = |C|^2 \Delta_m > 0$, de donde se deduce que $b > 0$. ■

* * *

EJEMPLO A. Demostrar que la forma cuadrática en $\mathbb{R}^3$

$$A = 2x^2 - 2xy + y^2 + 4xz - 6yz + 11z^2$$

es definida positiva.

La matriz de $A$ es

$$A = \begin{pmatrix} 2 & -1 & 2 \\ -1 & 1 & -3 \\ 2 & -3 & 11 \end{pmatrix}$$

puesto que

$$\Delta_1 = 2 > 0 \quad \Delta_2 = \begin{vmatrix} 2 & -1 \\ -1 & 1 \end{vmatrix} = 1 > 0 \quad \text{y} \quad \Delta_3 = |A| = \begin{vmatrix} 0 & 1 & -4 \\ -1 & 1 & -3 \\ 0 & -1 & 5 \end{vmatrix} = 1 > 0$$

la forma cudrática dada es definida positiva.

* * *

El recíproco de la proposición anterior también es cierto.

PROPOSICIÓN 3

En las mismas condiciones que en la proposición 2, si $A(\vec{x}, \vec{x})$ es una forma cudrática definida positiva, todos los menores angulares de la matriz $A = (a_{ij})$ son positivos.

*Demostración.* Sea $A(\vec{x}, \vec{x}) = \sum_{i,j=1}^{n} a_{ij}x_ix_j$ la expresión de la forma cuadrática $A$, definida positiva, en una base $\{\vec{e}_1, ..., \vec{e}_n\}$ de $E$. Para todo $1 \leqq k \leqq n$ la forma cuadrática

$$A_k(\vec{x}, \vec{x}) = \sum_{i,j=1}^{k} a_{ij}x_ix_j$$

definida en el subespacio vectorial $E_k = L\{\vec{e}_1, ..., \vec{e}_k\}$ de $E$, es definida positiva. Un cambio de base adecuado, de matriz $C_k$, reduce $A_k$ a

$$A_k = a_1y_1^2 + \cdots + a_ky_k^2 \quad , \quad (a_j > 0)$$

y se tiene

$$\begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_k \end{pmatrix} = C_k^t \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} C_k;$$

por tanto,

$$0 < a_1 \times \cdots \times a_n = |C_k^t||\Delta_k||C_k| = |C_k|^2\Delta_k,$$

de donde se deduce que $\Delta_k > 0$, que era lo que queríamos demostrar. ■

Para formas definidas negativas el criterio de Sylvester puede deducirse de las proposiciones que acabamos de demostrar. Si $A(\vec{x}, \vec{x})$ es definida negativa, $\tilde{A}(\vec{x}, \vec{x}) = -A(\vec{x}, \vec{x})$ es definida positiva; si $A = (a_{ij})_{i,j=1}^n$ es la matriz de $A(\vec{x}, \vec{x})$, $\tilde{A} = (-a_{ij})_{i,j=1}^n$ es la matriz de $\tilde{A}(\vec{x}, \vec{x})$ (en la misma base) y, por tanto,

$$\tilde{\Delta}_1 = -a_{11} > 0 \quad ; \quad \tilde{\Delta}_2 = \begin{vmatrix} -a_{11} & -a_{12} \\ -a_{12} & -a_{22} \end{vmatrix} > 0 \quad ; \quad \cdots \quad ; \quad \tilde{\Delta}_n = |\tilde{A}| > 0$$

Puesto que

$$\tilde{\Delta}_k = \begin{vmatrix} -a_{11} & -a_{12} & \cdots & -a_{1k} \\ -a_{12} & -a_{22} & \cdots & -a_{2k} \\ \vdots & \vdots & & \vdots \\ -a_{1k} & -a_{2k} & \cdots & -a_{kk} \end{vmatrix} = (-1)^k\Delta_k$$

se tiene el siguiente resultado:

PROPOSICIÓN 4 (*Criterio de Sylvester para formas cuadráticas definidas negativas*)

Sea $A(\vec{x}, \vec{x})$ una forma cuadrática en un espacio euclídeo con matriz $A = (a_{ij})_{i,j=1}^n$ en una cierta base. $A(\vec{x}, \vec{x})$ es definida negativa si y sólo si los menores angulares de la matriz $A$ se alternan en signo y $a_{11} < 0$.



*Nota.* La demostración que se ha dado de los criterios de Sylvester es válida para espacios euclídeos puesto que se ha utilizado el teorema de reducción a su forma canónica (teorema 2, sección 2). Sin embargo, puede demostrarse que los criterios siguen siendo válidos para formas cuadráticas definidas en espacios vectoriales reales puesto que hemos comentado que podemos encontrar una forma canónica completando cuadrados.

EJEMPLO B. *Estudiar los puntos críticos de la función* $z = f(x, y) = 2x^2 + y^2 + xy - 6x - 5y$.

Puesto que la solución del sistema

$$\left.\begin{aligned} \frac{\partial f}{\partial x} &= 4x + y - 6 = 0 \\ \frac{\partial f}{\partial y} &= 2y + x - 5 = 0 \end{aligned}\right\}$$

es $x = 1$, $y = 2$, el punto $(1, 2)$ es un punto crítico de $f$; en este punto la matriz hessiana de $f$ es

$$H = \begin{pmatrix} 4 & 1 \\ 1 & 2 \end{pmatrix}$$

que es la matriz de una forma cuadrática definida positiva ya que $4 > 0$ y $|H| = 7 > 0$. $f$ posee un *mínimo* en $(1, 2)$.

* * *

En la sección primera del capítulo 8 se dio la definición de producto escalar en un espacio vectorial y se demostró que su matriz en cualquier base es simétrica con todos los elementos de la diagonal principal positivos. No toda matriz que satisface estas condiciones define un producto escalar, como ya se observó en la citada sección.

Una condición necesaria y suficiente para que una matriz simétrica determine un producto escalar es que la forma cuadrática asociada a la matriz sea definida positiva. Este resultado resulta evidente a la luz de las definiciones de producto escalar y de forma cuadrática definida positiva.

EJEMPLO C. Queremos encontrar los valores de $a$ para los cuales la forma cuadrática en $\mathbb{R}^3$ dada por

$$Q(\vec{x}, \vec{x}) = (x_1, x_2, x_3) \begin{pmatrix} a & 1 & 0 \\ 1 & a & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

define un producto escalar en $\mathbb{R}^3$.

Puesto que $Q$ ha de ser definida positiva hemos de tener

$$a>0 \quad , \quad \begin{vmatrix} a & 1 \\ 1 & a \end{vmatrix} > 0 \quad \text{y} \quad \begin{vmatrix} a & 1 & 0 \\ 1 & a & 1 \\ 0 & 1 & 1 \end{vmatrix} > 0$$

De estas tres condiciones deducimos

$$a>0 \quad , \quad a^2-1>0 \quad \text{y} \quad a^2-a-1>0$$

Puesto que $a^2-a-1=0$ tiene como soluciones $a=\frac{1\pm\sqrt{5}}{2}$ hemos de tener

$$a>0 \quad , \quad a>1 \quad , \quad a<-1 \quad , \quad a>\frac{1+\sqrt{5}}{2} \quad \text{y} \quad a<\frac{1-\sqrt{5}}{2}$$

De todas estas condiciones deducimos $a>\frac{1+\sqrt{5}}{2}$.

## 12.5. DIAGONALIZACION SIMULTANEA DE FORMAS CUADRÁTICAS

En algunos problemas físicos se presenta de manera natural el siguiente problema: dadas dos formas cuadráticas $A(\vec{x}, \vec{x})$ y $B(\vec{x}, \vec{x})$ en $\mathbb{R}^n$ encontrar una base en la cual ambas formas pueden reducirse a una suma de cuadrados. Dadas dos formas cuadráticas cualesquiera es posible que el problema no tenga solución; sin embargo, si una de ellas es definida positiva siempre puede encontrarse una base que diagonaliza a ambas formas cuadráticas. La elegante solución de este problema se expone a continuación.

Supongamos que $A=(a_{ij})_{i,j=1}^n$ y $B=(b_{ij})_{i,j=1}^n$ son las matrices de las formas cuadráticas $A(\vec{x}, \vec{x})$ y $B(\vec{x}, \vec{x})$, respectivamente, con respecto a la base $E=\{\vec{e}_1, ..., \vec{e}_n\}$ de $\mathbb{R}^n$ y que $B(\vec{x}, \vec{x})$ es definida positiva. En la sección 12.4 se ha observado que la aplicación $( \, , \, )_B$ de $\mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$ dada por

$$(\vec{x}, \vec{y})_B = B(\vec{x}, \vec{y}) \qquad \vec{x}, \vec{y} \in \mathbb{R}^n$$

define un producto escalar en $\mathbb{R}^n$. Sea $\mathbb{R}_B^n$ el espacio euclídeo $\mathbb{R}^n$ con el producto escalar $( \, , \, )_B$. Utilizando el proceso de ortogonalización de Gram-Smith (ver capítulo 8) puede encontrarse una base ortonormal

$$U=\{\vec{u}_1, \vec{u}_2, ..., \vec{u}_n\}$$



en $\mathbb{R}_B^n$ a partir de la base $E$. Puesto que $B(\vec{u}_i, \vec{u}_j)=(\vec{u}_i, \vec{u}_j)_B=\delta_{ij}$, $i, j=1, 2, ..., n$, la matriz de $B$ en la base $U$ es la identidad y si $C$ es la matriz del cambio de base de $E$ a $U$ se tiene que

$$I=C^tBC. \tag{1}$$

En esta nueva base $U$ la matriz de $A$ es

$$\bar{A}=C^tAC. \tag{2}$$

Puesto que $A(\vec{x}, \vec{x})$ es una forma cuadrática en el espacio euclídeo $\mathbb{R}_B^n$, el teorema 2 de la sección 12.4 nos asegura la existencia de una base ortonormal

$$V=\{\vec{v}_1, \vec{v}_2, ..., \vec{v}_n\}$$

en $\mathbb{R}_B^n$ tal que si $\vec{x}=\sum_{i=1}^n x_i\vec{v}_i$ se tiene que

$$A(\vec{x}, \vec{x})=\lambda_1x_1^2+\lambda_2x_2^2+\cdots+\lambda_nx_n^2 \tag{3}$$

con $\lambda_j \in \mathbb{R}$, $j=1, 2, ..., n$. Puesto que $V$ es una base ortonormal en $\mathbb{R}_B^n$ se tiene que

$$B(\vec{x}, \vec{x})=x_1^2+x_2^2+\cdots+x_n^2 \tag{4}$$

donde $\vec{x}=\sum_{i=1}^n x_i\vec{v}_i$. Las igualdades (3) y (4) resuelven el problema de la diagonalización simultánea de $A(\vec{x}, \vec{x})$ y $B(\vec{x}, \vec{x})$.

A continuación mostramos con un ejemplo como se realiza este proceso en un caso particular.

EJEMPLO A. Deseamos diagonalizar simultáneamente las formas cuadráticas en $\mathbb{R}^2$ dadas por

$$A(\vec{x}, \vec{x})=-4x_1x_2 \qquad \text{y} \qquad B(\vec{x}, \vec{x})=x_1^2-2x_1x_2+4x_2^2$$

donde $\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2$.

La matriz de $B(\vec{x}, \vec{x})$ en la base $\{\vec{e}_1, \vec{e}_2\}$ es

$$B=\begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}$$

y una sencilla aplicación del criterio de Sylvester (sección 12.3) nos permite deducir que $B(\vec{x}, \vec{x})$ es definida positiva. Comenzamos encontrando una base $U=\{\vec{u}_1, \vec{u}_2\}$ de $\mathbb{R}^2$

que sea ortonormal con respecto al producto escalar $(\vec{x}, \vec{y})_B = B(\vec{x}, \vec{y})$ determinado por $B$. Conviene observar que $B(\vec{x}, \vec{y}) = x_1y_1 - x_1y_2 - x_2y_1 + 4x_2y_2$, donde

$$\vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2 \quad \text{e} \quad \vec{y} = y_1\vec{e}_1 + y_2\vec{e}_2.$$

Puesto que

$$(\vec{e}_1, \vec{e}_1)_B = B(\vec{e}_1, \vec{e}_1) = 1$$

podemos tomar $\vec{u}_1 = \vec{e}_1$; además, $\vec{v}_2 = \vec{e}_2 + \alpha\vec{e}_1$, donde $(\vec{v}_2, \vec{u}_1)_B = 0$; por tanto,

$$0 = (\vec{v}_2, \vec{u}_1)_B = (\vec{e}_2, \vec{e}_1)_B + \alpha(\vec{e}_1, \vec{e}_1)_B = -1 + \alpha;$$

de aquí deducimos que $\vec{v}_2 = \vec{e}_2 + \vec{e}_1$. Puesto que

$$(\vec{v}_2, \vec{v}_2)_B = B(\vec{v}_2, \vec{v}_2) = B\left(\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}\right) = 1 - 2 + 4 = 3$$

tomamos

$$\vec{u}_2 = \frac{1}{\sqrt{3}}(\vec{e}_1 + \vec{e}_2)$$

La matriz $\bar{A}$ de $A(\vec{x}, \vec{x})$ en la base $U = \{\vec{u}_1, \vec{u}_2\}$ es

$$\bar{A} = \begin{pmatrix} 0 & -\dfrac{2}{\sqrt{3}} \\ -\dfrac{2}{\sqrt{3}} & -\dfrac{4}{3} \end{pmatrix}$$

ya que

$$A(\vec{u}_1, \vec{u}_1) = A\left(\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}\right) = 0$$

$$A(\vec{u}_1, \vec{u}_2) = (1, 0)\begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}\begin{pmatrix} \dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \end{pmatrix} = (0, -2)\begin{pmatrix} \dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \end{pmatrix} = -\frac{2}{\sqrt{3}}$$

y

$$A(\vec{u}_2, \vec{u}_2) = A\left(\begin{pmatrix} \dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \end{pmatrix}\right) = -\frac{4}{3}.$$



La forma cuadrática $A(\vec{x}, \vec{x})$ puede diagonalizarse por el procedimiento utilizado en la sección 12.2. Puesto que

$$\begin{vmatrix} -\lambda & -\dfrac{2}{\sqrt{3}} \\ -\dfrac{2}{\sqrt{3}} & -\dfrac{4}{3} - \lambda \end{vmatrix} = \lambda^2 + \frac{4}{3}\lambda - \frac{4}{3} = \frac{1}{3}(3\lambda^2 + 4\lambda - 4)$$

y la ecuación $3\lambda^2 + 4\lambda - 4 = 0$ tiene como soluciones

$$\lambda_1 = -2 \quad , \quad \lambda_2 = \frac{2}{3}$$

se tiene que

$$A(\vec{y}, \vec{y}) = -2y_1^2 + \frac{2}{3}y_2^2$$

$$B(\vec{y}, \vec{y}) = y_1^2 + y_2^2$$

donde $\vec{y} = y_1\vec{v}_1 + y_2\vec{v}_2$ y $V = \{\vec{v}_1, \vec{v}_2\}$ es una base ortonormal con respecto a $(\ ,\ )_B$.

Los vectores $\vec{v}_1$ y $\vec{v}_2$ se encuentran calculando los subespacios invariantes correspondientes a $\lambda_1$ y $\lambda_2$. Para $\lambda_1 = -2$ tenemos

$$\begin{pmatrix} 2 & -\dfrac{2}{\sqrt{3}} \\ -\dfrac{2}{\sqrt{3}} & -\dfrac{4}{3} + 2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad 2y_1 - \frac{2}{\sqrt{3}}y_2 = 0 \quad \Rightarrow \quad y_1 = 1,\ y_2 = \sqrt{3}$$

Tomar

$$\vec{g}_1 = 1 \cdot \vec{u}_1 + \sqrt{3}\vec{u}_2 = \vec{e}_1 + (\vec{e}_1 + \vec{e}_2).$$

Puesto que

$$(\vec{g}_1, \vec{g}_1)_B = B\left(\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}2\\1\end{pmatrix}\right) = 4 - 4 + 4 = 4$$

tomamos

$$\vec{v}_1 = \vec{e}_1 + \frac{1}{2}\vec{e}_2$$

Para $\lambda_2=\frac{2}{3}$ tenemos

$$\begin{pmatrix} -\frac{2}{3} & -\frac{2}{\sqrt{3}} \\ -\frac{2}{\sqrt{3}} & -\frac{4}{3}-\frac{2}{3} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad -\frac{2}{3}y_1-\frac{2}{\sqrt{3}}y_2=0 \quad \Rightarrow \quad y_1=\sqrt{3},\ y_2=-1.$$

Tomar

$$\vec{g}_2=\sqrt{3}\vec{u}_1-\vec{u}_2=\sqrt{3}\vec{e}_1-\frac{1}{\sqrt{3}}(\vec{e}_1+\vec{e}_2).$$

Puesto que

$$(\vec{g}_2, \vec{g}_2)_B=B\left(\begin{pmatrix} \sqrt{3}-\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix}, \begin{pmatrix} \sqrt{3}-\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \end{pmatrix}\right)=\frac{4}{3}+\frac{4}{3}+\frac{4}{3}=4$$

tomamos

$$\vec{v}_2=\frac{\sqrt{3}}{2}\vec{e}_1-\frac{\sqrt{3}}{6}(\vec{e}_1+\vec{e}_2)=\frac{\sqrt{3}}{3}\vec{e}_1-\frac{\sqrt{3}}{6}\vec{e}_2$$

Observar que la matriz del cambio de base que permite diagonalizar $A(\vec{x}, \vec{x})$ y $B(\vec{x}, \vec{x})$ simultáneamente es

$$\begin{pmatrix} 1 & \frac{\sqrt{3}}{3} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} \end{pmatrix}$$

y, por tanto, la transformación lineal que permite diagonalizar ambas formas cuadráticas es

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 1 & \frac{\sqrt{3}}{3} \\ \frac{1}{2} & -\frac{\sqrt{3}}{6} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

* * *



El lector habrá observado que el proceso de diagonalización simultánea seguido en el ejemplo anterior resulta bastante complicado. Los comentarios que siguen están encaminados a simplificar este proceso.

Es claro que si $B(\vec{x}, \vec{x})$ es la forma cuadrática definida positiva, tenemos

$$B(\vec{y}, \vec{y})=y_1^2+y_2^2+\cdots+y_n^2$$

en la base que diagonaliza a ambas formas cuadráticas a la vez. En esta misma base

$$A(\vec{y}, \vec{y})=\lambda_1 y_1^2+\lambda_2 y_2^2+\cdots+\lambda_n y_n^2$$

donde $\lambda_1, \lambda_2, ..., \lambda_n$ son los autovalores de $\bar{A}$. Utilizando las igualdades (1) y (2) tenemos

$$|\bar{A}-\lambda I|=|C^tAC-\lambda C^tBC|=|C^t(A-\lambda B)C|=|C^t||A-\lambda B||C|$$

de donde se deduce que las soluciones de $|\bar{A}-\lambda I|=0$ coinciden con las soluciones de $|A-\lambda B|=0$, ya que $|C|\neq 0$.

Así pues, la resolución de la ecuación $|A-\lambda B|=0$ nos da la diagonalización simultánea.

Si queremos encontrar la base en la cual ambas formas cuadráticas se diagonalizan, observamos que de $\bar{A}-\lambda I=C^t(A-\lambda B)C$ se obtiene que las soluciones de

$$(\bar{A}-\lambda_j I)C^{-1}\left(\sum_{k=1}^{n} y_k\vec{u}_k\right)=\vec{0} \qquad j=1, 2, ..., n$$

coinciden con las soluciones de

$$(A-\lambda_j B)\left(\sum_{k=1}^{n} y_k\vec{e}_k\right)=\vec{0} \qquad j=1, 2, ..., n \tag{5}$$

ya que $|C^t|=|C|\neq 0$. Por tanto basta con ortonormalizar los vectores que se encuentran al resolver (5).

En el ejemplo A tenemos

$$0=|A-\lambda B|=\left|\begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}-\lambda\begin{pmatrix} 1 & -1 \\ -1 & 4 \end{pmatrix}\right|=\begin{vmatrix} -\lambda & -2+\lambda \\ -2+\lambda & -4\lambda \end{vmatrix}=4\lambda^2-(\lambda-2)^2=$$
$$=3\lambda^2-4\lambda+4,$$

que produce las soluciones $\lambda_1=-2,\ \lambda_2=\frac{2}{3}$.

Para $\lambda_1=-2$,

$$\begin{pmatrix} 2 & -4 \\ -4 & 8 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad 2y_1-4y_2=0 \quad \Rightarrow \quad y_1=2,\ y_2=1$$

Puesto que

$$\left(\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}2\\1\end{pmatrix}\right)_B=B\left(\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}2\\1\end{pmatrix}\right)=4-4+4=4$$

tomamos

$$\vec{v}_1=\left(1,\frac{1}{2}\right)$$

Para $\lambda_2=\frac{2}{3}$,

$$\begin{pmatrix}-\frac{2}{3} & -\frac{4}{3}\\ -\frac{4}{3} & -\frac{8}{3}\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}=\begin{pmatrix}0\\0\end{pmatrix} \Rightarrow -2y_1-4y_2=0 \Rightarrow y_1=2,\ y_2=-1.$$

Puesto que

$$\left(\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}2\\-1\end{pmatrix}\right)_B=B\left(\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}2\\-1\end{pmatrix}\right)=4+4+4=12$$

tomamos

$$\vec{v}_2=\left(\frac{1}{\sqrt{3}},-\frac{1}{2\sqrt{3}}\right)$$

Observar que éstos son los mismos resultados que los obtenidos en el ejemplo A.

EJEMPLO B. Deseamos diagonalizar simultáneamente las formas cuadráticas en $\mathbb{R}^4$ dadas por

$$A(\vec{x},\vec{x})=2x_4^2+x_1x_2+x_1x_3-2x_2x_3+2x_2x_4$$

y

$$B(\vec{x},\vec{x})=\frac{1}{4}x_1^2+x_2^2+x_3^2+x_4^2+2x_2x_4$$

donde

$$x=x_1\vec{e}_1+x_2\vec{e}_2+x_3\vec{e}_3+x_4\vec{e}_4$$

Puesto que $B(\vec{x},\vec{x})$ es definida positiva tenemos

$$B(\vec{y},\vec{y})=y_1^2+y_2^2+y_3^2+y_4^2.$$



Además

$$|A-\lambda B|=\left|\begin{pmatrix}0 & \frac{1}{2} & \frac{1}{2} & 0\\ \frac{1}{2} & 0 & -1 & 1\\ \frac{1}{2} & -1 & 0 & 0\\ 0 & 1 & 0 & 2\end{pmatrix}-\lambda\begin{pmatrix}\frac{1}{4} & 0 & 0 & 0\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & 0\\ 0 & 1 & 0 & 2\end{pmatrix}\right|=(\lambda-1)^3(\lambda+3)$$

Por tanto,

$$A(\vec{y},\vec{y})=y_1^2+y_2^2+y_3^2-3y_4^2$$

Se deja como ejercicio encontrar el cambio de base que diagonaliza simultáneamente ambas formas cuadráticas. El resultado es

$$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 0 & -1\\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2}\end{pmatrix}\begin{pmatrix}y_1\\y_2\\y_3\\y_4\end{pmatrix}$$

* * *

*Observación.* Todos los resultados de esta sección se extienden fácilmente a espacios vectoriales de dimensión finita y formas cuadráticas definidas sobre ellos.

## EJERCICIOS (CAPITULO 12)

### SECCIÓN 12.1

**1.** Encontrar la matriz de las siguientes formas cuadráticas en $\mathbb{R}^n$ dadas con respecto a la base canónica:

*a)* $A(\vec{x},\vec{x})=2x_1^2-3x_1x_2+4x_2^2+6x_1x_3$ , $n=3$

*b)* $A(\vec{x},\vec{x})=\sum_{i=1}^{3}(x_i-s)^2$ , $s=\frac{1}{3}(x_1+x_2+x_3)$ , $n=3$

*c)* $A(\vec{x},\vec{x})=\sum_{i<j}^{n}(i-j)^2x_ix_j$

**2.** Sea $A(\vec{x}, \vec{y})$ una forma bilineal que tiene como expresión

$$A(\vec{x}, \vec{y}) = 2x_1y_1 + 3x_1y_2 - x_2y_1 + x_3y_3$$

con respecto a la base $\vec{u}_1 = (1, 0, 1)$, $\vec{u}_2 = (0, 1, 0)$, $\vec{u}_3 = (0, 1, -1)$ en $\mathbb{R}^3$, donde $x = x_1\vec{u}_1 + x_2\vec{u}_2 + x_3\vec{u}_3$, $y = y_1\vec{u}_1 + y_2\vec{u}_2 + y_3\vec{u}_3$.

*a)* Encontrar la expresión de $A(\vec{x}, \vec{y})$ con respecto a la base canónica.

*b)* Encontrar una forma bilineal simétrica $B$ y una forma bilineal antisimétrica $C$ tal que $A = B + C$, dando las expresiones de $B$ y $C$ con respecto a la base canónica de $\mathbb{R}^3$.

**3.** Sea $A$ una forma bilineal cuyo rango coincide con la dimensión del espacio vectorial $V$ en el cual está definida. Demostrar que para todo $\vec{x}_0 \in V$, $\vec{x}_0 \neq 0$ existe $\vec{y}_0 \in V$ tal que $A(\vec{x}_0, \vec{y}_0) \neq 0$.

## SECCIÓN 12.2

**4.** Encontrar una forma canónica de las siguientes formas cuadráticas:

*a)* $3x^2 + 2y^2 + z^2 - 6xy + 4xz$

*b)* $xy + 2xz$

*c)* $9x^2 - 3y^2 + 6xy + 18xz + 12yz$

**5.** Dar una forma canónica en una base ortonormal de las siguientes formas cuadráticas en $\mathbb{R}^n$:

*a)* $-2x^2 + y^2 + 4xy \quad , \quad n = 3$

*b)* $2x^2 - 6y^2 - 2z^2 - 2xz \quad , \quad n = 3$

*c)* $x^2 - 2xy - 2xz - 2yz \quad , \quad n = 3$

*d)* $x_1x_2 + x_2x_2 + x_3x_4 \quad , \quad n = 4$

**6.** Para cada una de las formas cuadráticas siguientes encontrar una base ortonormal que la reduzca a su forma canónica y escribir esta forma canónica:

*a)* $6x^2 + 5y^2 + 7z^2 - 4xy + 4xz$

*b)* $3y^2 + 3z^2 + 4xy + 4xz + 2yz$

*c)* $x^2 + y^2 - 2xz + 2yz$

## SECCIÓN 12.3

**7.** Encontrar la forma normal y los índices de inercia de las siguientes formas cuadráticas en $\mathbb{R}^3$:

*a)* $x^2 + y^2 + z^2 + 4xy + 2xz + 2yz$



*b)* $x^2 + 5y^2 - 2xy + 2xz$

*c)* $4x^2 + y^2 + z^2 - 4xy + 4xz - 3yz$

En el caso *b)* indicar también la transformación que lleva la forma cuadrática dada en la forma normal encontrada.

## SECCIÓN 12.4

**8.** Determinar los valores de $\alpha$ para los cuales la forma cuadrática en $\mathbb{R}^3$ dada por

$$x^2 + 4y^2 + 2\alpha xy + 2\alpha yz$$

es definida positiva

**9.** Determinar los valores de $\beta$ para los cuales la forma cuadrática en $\mathbb{R}^4$ dada por $\beta x^2 + 2xy + \beta y^2 + \beta z^2 + 2zt + \beta t^2$ es definida negativa.

**10.** Determinar la región tridimensional en la cual la matriz de las derivadas parciales segundas de la función $F(x, y, z) = x^2 + 3y^2 + y^2z^2$ determina una forma cuadrática definida positiva.

**11.** Encontrar los puntos de las superficies dadas en los cuales se alcanzan máximos o mínimos:

*a)* $f(x, y) = x^2 + y^2 + 2x + 4$

*b)* $g(x, y) = -x^4 + 8x^2 - y^2 - 4y$

*c)* $h(x, y) = x^3 - y^3 + 3xy$

**12.** Sea $A = a_{11}x^2 + a_{12}xy + a_{22}y^2$ una forma cuadrática en $\mathbb{R}^2$ tal que

$$\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22} \end{vmatrix} < 0$$

Demostrar que $A$ no es ni semidefinida positiva ni semidefinida negativa.

**13.** Determinar los valores de los parámetros $a$ y $b$ para los cuales la forma cuadrática cuya matriz se da a continuación define un producto escalar en $\mathbb{R}^3$:

$$a) \begin{pmatrix} 2 & 0 & a \\ 0 & a & 0 \\ a & 0 & 1 \end{pmatrix} \qquad b) \begin{pmatrix} a & 2b & 1 \\ 2 & 5 & 0 \\ a & 0 & 1 \end{pmatrix}$$

SECCIÓN 12.5

**14.** Diagonalizar simultáneamente las formas cuadráticas

$$A(\vec{x}, \vec{x}) = x_1^2 + 26x_2^2 + 10x_1x_2$$

y

$$B(\vec{x}, \vec{x}) = x_1^2 + 56x_2^2 + 16x_1x_2$$

## BIOGRAFÍA

**James Joseph Sylvester** (1814-1897) fue un matemático inglés que junto con Arthur Cayley fundó la teoría de los invariantes algebraicos, es decir, expresiones de los coeficientes de una expresión algebraica que quedan invariantes mediante rotaciones y traslaciones de los ejes coordenados.

En 1838 Sylvester consiguió un puesto como profesor de filosofía en University College (Londres). En 1841 aceptó un cargo de profesor de matemáticas en la Universidad de Virginia, Charlottesville (EE.UU.), en el que sólo estuvo tres meses. Cuatro años más tarde volvió a Londres como contable de una compañía aseguradora, manteniendo su interés en matemáticas a través de las clases privadas que impartía. Más tarde trabajó con abogado, comenzando en este período su relación matemática con Cayley.

Desde 1855 hasta 1870 Sylvester fue profesor de matemáticas en la Academia Real Militar en Woolwich (Inglaterra). Volvió a los Estados Unidos en 1876 para ser profesor de matemáticas en la Universidad Johns Hopkins en Baltimore (Maryland). Trabajando en esta Universidad fundó la revista matemática *American Journal of Mathematics*, introdujo los estudios de Doctorado en las Universidades de los Estados Unidos y contribuyó enormemente al desarrollo de las matemáticas en este país. En 1883 volvió a Inglaterra como profesor de geometría en la Universidad de Oxford.

Sylvester fue principalmente un algebrista. Realizó trabajos brillantes en teoría de números, principalmente en particiones (formas en que un número puede expresarse como suma de enteros positivos) y análisis diofántico (métodos para encontrar soluciones enteras de ecuaciones algebraicas con coefientes enteros).

Sylvester trabajaba por inspiración y es difícil encontrar en sus trabajos una prueba correcta para los moldes actuales. A pesar de que publicó cientos de artículos, solamente escribió un libro de matemáticas titulado *Treatise on Elliptic Functions* (1876). Sylvester publicó en 1870 un libro de versos.

# CAPITULO 13

# SUPERFICIES DE SEGUNDO GRADO



La expresión general de las cónicas es un polinomio de segundo grado en dos variables. La generalización a tres variables produce una expresión de la forma

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+a_{12}xy+a_{13}xz+a_{23}yz+a_1x+a_2y+a_3z+a=0$$

que recibe el nombre de superficie de segundo grado en $\mathbb{R}^3$ o cuádrica.

A la luz de los resultados obtenidos en el capítulo anterior sobre formas cuadráticas clasificaremos las superficies de segundo grado, de manera análoga a como se hizo con las cónicas.

En este capítulo trabajaremos en el espacio euclídeo $\mathbb{R}^3$ con el producto escalar usual y toda expresión estará dada con respecto a la base canónica si no se menciona explícitamente otra base.

## 13.1. CLASIFICACION DE LAS SUPERFICIES DE SEGUNDO GRADO EN $\mathbb{R}^3$

Una superficie de segundo grado en $\mathbb{R}^3$ es una expresión de la forma

$$f(x, y, z)=a_{11}x^2+a_{22}y^2+a_{33}z^2+a_{12}xy+a_{13}xz+ a_{23}yz+a_1x+a_2y+a_3z+a=0 \tag{1}$$

Los términos de orden dos determinan una forma cuadrática $A$, en $\mathbb{R}^3$, cuya matriz es

$$A=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} \end{pmatrix}.$$

Si escribimos $X=(x, y, z)$ y llamamos $L(X)=a_1x+a_2y+a_3z$ a la parte lineal, (1) puede escribirse de la forma

$$A(X, X)+L(X)+a=0 \tag{2}$$

Nuestro objetivo inmediato es reconocer todas las superficies que puedan aparecer en (1) al variar sus coeficientes. Puesto que $A(X, X)$ es una forma cuadrática en el espacio euclídeo $\mathbb{R}^3$, puede reducirse, en una base ortonormal $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$ a una suma de cuadrados

$$\lambda_1x_1^2+\lambda_2y_1^2+\lambda_3z_1^2$$

donde $\lambda_1$, $\lambda_2$, $\lambda_3$ son los autovalores de la matriz $A(\lambda_j\in\mathbb{R})$. Observar que la matriz del cambio de base es *ortogonal ya que transforma la base canónica de* $\mathbb{R}^3$ en una base ortonormal. En esta nueva base (1) se reduce a

$$\lambda_1x_1^2+\lambda_2y_2^2+\lambda_3z_1^2+b_1x_1+b_2y_1+b_3z_1+b=0 \tag{3}$$

Pueden presentarse los siguientes *casos*

I) Todos los $\lambda_i$ son diferentes de cero.
II) Uno de los $\lambda_i$ es igual a cero.
III) Dos de los $\lambda_i$ son iguales a cero.

CASO I. Como todos los $\lambda_i$ son distintos de cero, completando cuadrados en (3) obtenemos

$$\lambda_1\left(x_1+\frac{b_1}{2\lambda_1}\right)^2+\lambda_2\left(y_1+\frac{b_2}{2\lambda_2}\right)^2+\lambda_3\left(z_1+\frac{b_3}{2\lambda_3}\right)^2=C,$$



con lo cual la traslación

$$x_2=x_1+\frac{b_1}{2\lambda_1}$$
$$y_2=y_1+\frac{b_2}{2\lambda_2}$$
$$z_2=z_1+\frac{b_3}{2\lambda_3}$$

permite escribir (3) de la forma

$$\lambda_1x_2^2+\lambda_2y_2^2+\lambda_3z_2^2=C \tag{4}$$

que recibe el nombre de *forma canónica de una superficie con centro.* La razón de esta terminología es la siguiente: se denomina centro de una superficie al punto $C_0=(x_0, y_0, z_0)$ tal que si $(x_0+a_0, y_0+b_0, z_0+c_0)$ es un punto de la superficie, su simétrico respecto a $C_0$, esto es

$$(x_0-a_0, y_0-b_0, z_0-c_0)$$

está también en la superficie. Claramente, la superficie (4) tiene $C_0=(0, 0, 0)$ como centro.

* * *

Si $C\neq 0$, dividiendo por $C$, (4) se transforma en

$$\pm\frac{x_2^2}{a_2}\pm\frac{y_2^2}{b}\pm\frac{z_2^2}{c^2}=1 \tag{5}$$

donde

$$a=\sqrt{\frac{C}{|\lambda_1|}}\;, \quad b=\sqrt{\frac{C}{|\lambda_2|}} \quad \text{y} \quad c=\sqrt{\frac{C}{|\lambda_3|}}$$

se denominan *semiejes* de la superficie.

Excluyendo el caso en que todos los términos son negativos en (5), ya que en este caso (5) no tiene ninguna solución y reordenando, si es necesario, se tiene los siguientes casos

$$\text{I.}a)\quad \frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}+\frac{z_2^2}{c^2}=1$$

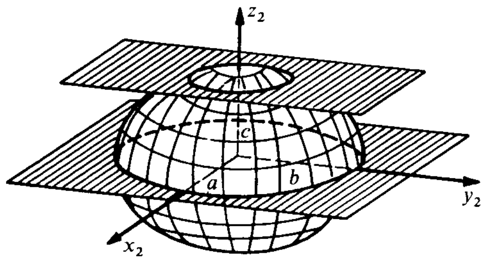
$$\text{I.}b)\quad \frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}=1$$

$$\text{I.}c)\quad \frac{x_2^2}{a^2}-\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}=1$$

I.*a*) Al cortar por planos paralelos al plano $x_2Oy_2$, es decir $z_2=d$ se obtienen elipses $\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1-\frac{d^2}{c^2}$ siempre que $c^2-d^2>0$ (i.e. $|d|<|c|$).

De manera similar al cortar por planos de la forma $y_2=d$(ó $x_2=d$) se obtienen elipses siempre que $|d|<|b|$ ($|d|<|a|$).

La superficie recibe el nombre de *elipsoide* y su gráfica se muestra en la figura 1.




$$\boxed{\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}+\frac{z_2^2}{c^2}=1}$$

Figura 1. *Elipsoide*

I.*b*) Al cortar por planos de la forma $z_2=d$ se obtienen elipses de ecuaciones

$$\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}=1+\frac{d^2}{c^2},$$

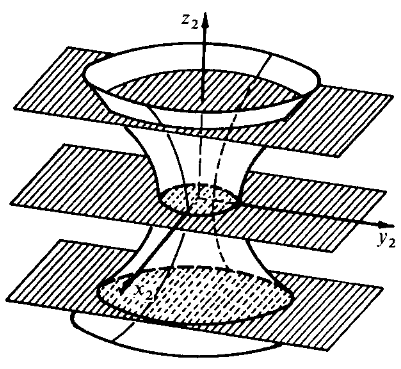
ecuación que posee soluciones para todo $d\in\mathbb{R}$. Si $y_2=0$ se obtiene la hipérbola

$$\frac{x_2^2}{a^2}-\frac{z_2^2}{c^2}=1$$

de semiejes $a$ y $c$ y si $x_2=0$ se obtiene la hipérbola

$$\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}=1$$

de semiejes $b$ y $c$. Esta superficie recibe el nombre de *hiperboloide de una hoja.*




$$\boxed{\frac{x_2^2}{a^2}+\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}=1}$$

Figura 2. *Hiperboloide de una hoja*



I.*c*) Al cortar por planos de la forma $z_2=d$ se obtienen las hipérbolas de ecuación

$$\frac{x_2^2}{a^2}-\frac{y_2^2}{b^2}=1+\frac{d^2}{c^2}$$

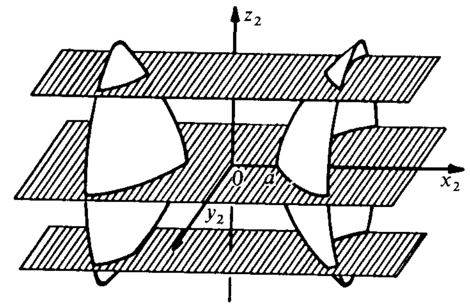
Cuando $y_2=0$ se obtiene la hipérbola

$$\frac{x_2^2}{a^2}-\frac{z_2^2}{c^2}=1$$

Finalmente, si $x_2=d$ se obtiene

$$\frac{y_2^2}{b^2}+\frac{z_2^2}{c^2}=-1+\frac{d^2}{a^2}$$

que representa una elipse siempre que $-a^2+d^2>0$, es decir, $|d|>|a|$. Esta superficie recibe el nombre de *hiperboloide de dos hojas.*




$$\boxed{\frac{x_2^2}{a^2}-\frac{y_2^2}{b^2}-\frac{z_2^2}{c^2}=1}$$

Figura 3. *Hiperboloide de dos hojas.*

Si $C=0$, (4) se escribe de la forma $\lambda_1x_2^2+\lambda_2y_2^2+\lambda_3z_2^2=0$; si todos los $\lambda_j$ son positivos o todos son negativos, esta ecuación solamente tiene como solución el punto $x_2=0$, $y_2=0$, $z_2=0$. En el resto de los casos siempre podemos suponer que dos de ellos son positivos y uno es negativo (si uno es positivo y el resto son negativos se cambia de signo a la ecuación). Suponiendo que $\lambda_1>0$, $\lambda_2>0$ y $\lambda_3<0$ y dividiendo por $|\lambda_3|$ se obtiene

$$\frac{\lambda_1}{|\lambda_3|}x_2^2+\frac{\lambda_2}{|\lambda_3|}y_2^2=z_2^2$$

En los planos de ecuación $z_2=d$ esta expresión produce elipses, salvo si $d=0$, en cuyo caso se obtiene el punto $x_2=0$, $y_2=0$. Si $y_2=0$ la ecuación se transforma en

$$\frac{\lambda_1}{|\lambda_3|}x_2^2=z_2^2 \quad\Rightarrow\quad z_2=\pm\sqrt{\frac{\lambda_1}{|\lambda_3|}}x_2$$

que representa dos rectas de pendiente $\pm\sqrt{\frac{\lambda_1}{|\lambda_3|}}$. De manera similar, si $x_2=0$ se obtienen dos rectas de pendiente

$$\pm\sqrt{\frac{\lambda_2}{|\lambda_3|}}.$$

La superficie que se obtiene en este caso se denomina *cono elíptico* (Fig. 4).

Debido a que cuando $C=0$, (4) produce superficies cónicas, en los casos no triviales, cuando $C=0$ en (4) la superficie recibe el nombre de *superficie cónica.*

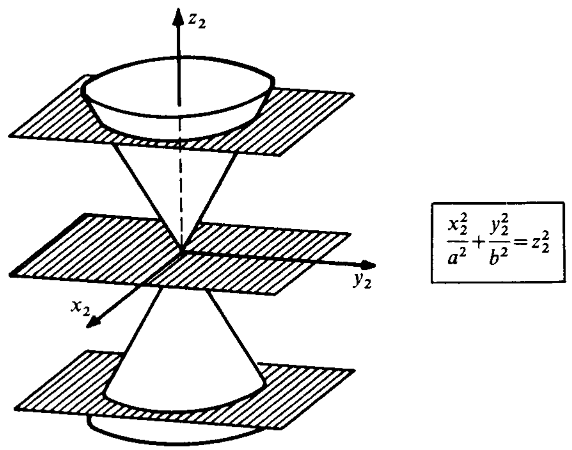



Figura 4. *Cono elíptico.*

* * *

CASO II. Podemos suponer, por ejemplo, que $\lambda_3=0$; podemos entonces completar cuadrados en $x_1$ e $y_1$ en la expresión (3) para obtener

$$\lambda_1\left(x_1+\frac{b_1}{2\lambda_1}\right)^2+\lambda_2\left(y_1+\frac{b_2}{2\lambda_2}\right)^2+b_3z_1=C.$$

La traslación

$$x_2=x_1+\frac{b_1}{2\lambda_1}$$
$$y_2=y_1+\frac{b_2}{2\lambda_2}$$
$$z_2=z_1$$



transforma la superficie anterior en

$$\lambda_1x_2^2+\lambda_2y_2^2+b_3z_2=C. \tag{6}$$

Pueden presentarse los siguientes casos:

II.*a*) $b_3\neq 0$
II.*b*) $b_3=0$

II.*a*) Como $b_3\neq 0$, la traslación $x_3=x_2$, $y_3=y_2$, $z_3=-z_2+\frac{C}{b_3}$ transforma (6) en

$$\lambda_1x_3^2+\lambda_2y_3^2=b_3z_3$$

o bien

$$\pm\frac{x_3^2}{a^2}\pm\frac{y_3^2}{b^2}=z_3$$

una vez que se ha dividido por $b_3$, donde $a=\sqrt{\frac{|b_3|}{|\lambda_1|}}$, y $b=\sqrt{\frac{|b_3|}{|\lambda_2|}}$. Esencialmente la ecuación anterior tiene sólo los siguientes casos:

$$a)\quad \frac{x_3^2}{a^2}+\frac{y_3^2}{b^2}=z_3 \quad ; \quad b)\quad \frac{x_3^2}{a^2}-\frac{y_3^2}{b^2}=z_3$$

(el caso en que ambos signos sean negativos es «equivalente» al caso *a*) si hacemos la simetría $z_3=-z_3$ y el caso en que el primero sea negativo y el segundo positivo es «equivalente» al caso *b*) realizando la misma simetría).

En el caso *a*) los planos $z_3=d$ producen una elipse siempre que $d>0$ y un punto si $d=0$; si $d<0$ no se produce ninguna intersección. Las elipses tienen por ecuación

$$\frac{x_3^2}{a^2}+\frac{y_3^2}{b^2}=d.$$

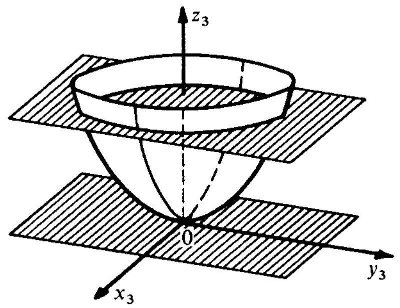
Si $y_3=0$ se obtiene la parábola

$$x_3^2=a^2z_3$$

(con su eje en el eje $0z_3$) y si $x_3=0$ se obtiene otra parábola con eje en $0z_3$, de ecuación

$$y_3^2=b^2z_3.$$

La superficie que se obtiene se denomina *paraboloide elíptico* y puede apreciarse en la figura 5.




$$\boxed{\frac{x_3^2}{a^2}+\frac{y_3^2}{b^2}=z_3}$$

Figura 5. *Paraboloide elíptico.*

En el caso *b*), el plano $z_3=0$ produce las rectas $x_3=\pm\frac{a}{b}y_3$, de pendiente $\frac{a}{b}$; si $z_3=d$, con $d>0$, se obtienen las hipérbolas

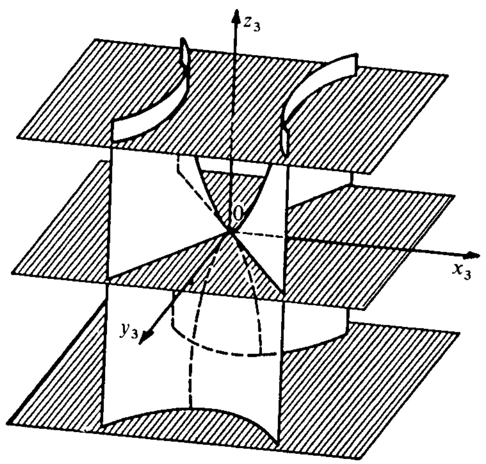
$$\frac{x_3^2}{a^2}-\frac{y_3^2}{b^2}=d$$

con eje principal en la dirección de $x_3$; finalmente, si $z_3=d$ con $d<0$, se obtienen las hipérbolas

$$-\frac{x_3^2}{a^2}+\frac{y_3^2}{b^2}=-d \qquad (-d>0)$$

cuyo eje principal está en la dirección de $y_3$.
El plano $y_3=0$ produce la parábola $x_3^2=a^2z_3$ que está dirigida en el sentido positivo de $z_3$. mientras que el plano $x_3=0$ produce la parábola $y_3^2=-b^2z_3$ que está dirigida en el sentido negativo de $z_3$.

La superficie que se obtiene se denomina *paraboloide hiperbólico* y se representa en la figura 6.




$$\boxed{\frac{x_3^2}{a^2}-\frac{y_3^2}{b^2}=z_3}$$

Figura 6. *Paraboloide hiperbólico.*



II.*b*) Si $b_3=0$, la ecuación (6) se escribe de la forma

$$\lambda_1x_2^2+\lambda_2y_2^2=C.$$

Puesto que le falta una de las variables, $z_3$, *decimos que la ecuación representa una superficie degenerada*; de hecho con cualquier plano $z_3=d$ posee como intersección la cónica $\lambda_1x_2^2+\lambda_2y_2^2=C$ y, por tanto, la superficie es una superficie *cilíndrica* que tiene a esta cónica como base: se tiene entonces un *cilindro elíptico, un cilindro hiperbólico, dos planos que se cortan o una recta* (que corresponden a las siguientes cónicas: elipse, hipérbola, dos rectas que se cortan y un punto). Ver figuras 7 y 8.

* * *

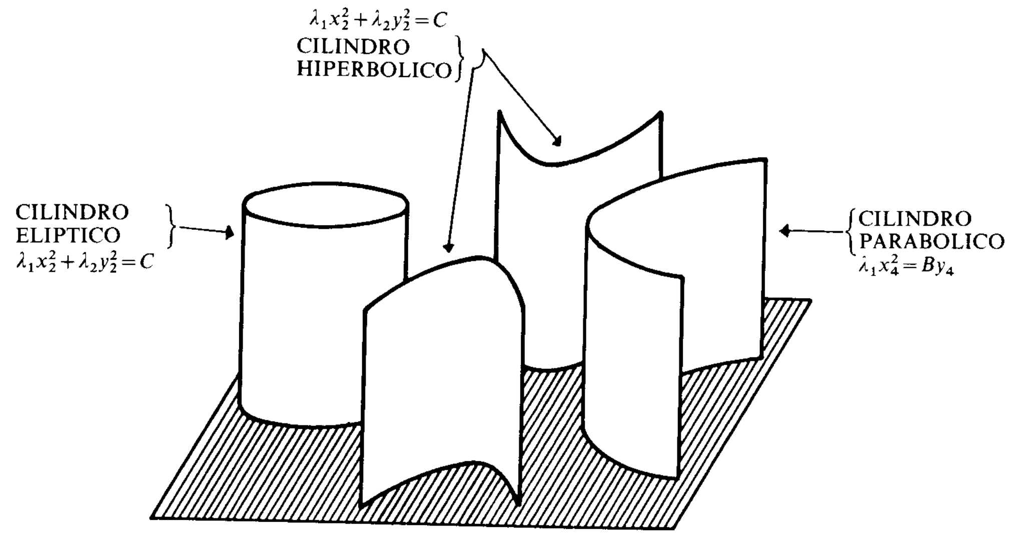



Figura 7

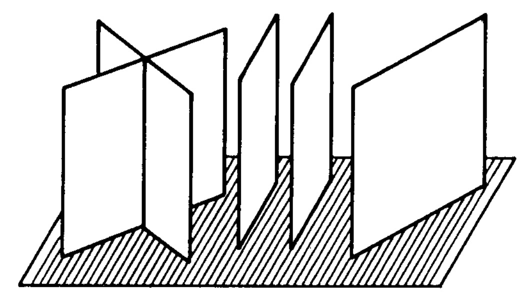


*Figura* 8

* * *

CASO III. Supongamos, por ejemplo, que $\lambda_2=\lambda_3=0$, con lo cual (3) se transforma en

$$\lambda_1 x_1^2+b_1x_1+b_2y_1+b_3z_1+b=0$$

La traslación $x_2=x_1+\frac{b_1}{2\lambda_1}$, $y_2=y_1$, $z_2=z_1$ la reduce a

$$\lambda_1 x_2^2+b_2y_2+b_3z_2+b'=0 \tag{7}$$

Si $b_2=b_3=0$ se obtienen *dos planos paralelos distintos* o *un solo plano*, o el conjunto vacío dependiendo de los valores y los signos de $\lambda_1$ y $b'$.

Si al menos uno de los coeficientes $b_2$ ó $b_3$ es no nulo realizamos la transformación

$$x_2=x_3$$
$$y_2=\frac{b_2y_3+b_3z_3}{\sqrt{b_2^2+b_3^2}}$$
$$z_2=\frac{b_3y_3-b_2z_3}{\sqrt{b_2^2+b_3^2}}$$

que es una transformación ortogonal (¿por qué?) de manera que (7) se escribe de la forma:

$$\lambda_1x_3^2+\sqrt{b_2^2+b_3^2}\,y_3+b'=0$$

Con la traslación

$$x_4=x_3\ ,\ y_4=y_3+\frac{b'}{\sqrt{b_2^2+b_3^2}}\ ,\ z_4=z_3$$

tenemos

$$\lambda_1x_4^2=By_4 \quad , \quad B=\sqrt{b_2^2+b_3^2}$$

que representa una parábola con su eje en la dirección de $y_4$; como estamos en $\mathbb{R}^3$ la superficie es un *cilindro parabólico*. (Ver figura 7).

*Notas*: 1. Las ecuaciones que aparecen al lado de las figuras reciben el nombre de *forma canónica* de la superficie que representan

2. Los paraboloides (Figs. 5 y 6) no poseen *centro*, pero poseen un *vértice* en el origen cuando su ecuación está dada en la forma canónica.

3. Aparte de los elementos geométricos que se citan en 2, las superficies no degeneradas poseen *ejes* que se han dibujado convenientemente en las figuras.



EJEMPLO A. Encontrar la forma canónica de la superficie de segundo grado en $\mathbb{R}^3$ dada por

$$3x^2+2y^2+3z^2-2xz-2x-4y-2z-5=0$$

indicando la correspondiente transformación geométrica.

SOLUCIÓN. La forma cuadrática $3x^2+2y^2+3z^2-2xz$ tiene como matriz

$$A=\begin{pmatrix}3&0&-1\\0&2&0\\-1&0&3\end{pmatrix}.$$

Los autovalores de $A$ son las soluciones de

$$0=\begin{vmatrix}3-\lambda&0&-1\\0&2-\lambda&0\\-1&0&3-\lambda\end{vmatrix}=(3-\lambda)(2-\lambda)(3-\lambda)-(2-\lambda)=(2-\lambda)[(3-\lambda)(3-\lambda)-1]$$
$$=(2-\lambda)[\lambda^2-6\lambda+8]=(2-\lambda)(\lambda-2)(\lambda-4)$$

y, por tanto, tenemos $\lambda_1=4$, $\lambda_2=2$, $\lambda_3=2$.

Para $\lambda_1=4$ un autovector satisface las ecuaciones

$$\begin{pmatrix}-1&0&-1\\0&-2&0\\-1&0&-1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \Leftrightarrow \begin{matrix}x=-z\\-2y=0\end{matrix} \Rightarrow \vec{u}_1=\frac{1}{\sqrt{2}}(-1,0,1)$$

Para $\lambda=2$ tenemos

$$\begin{pmatrix}1&0&-1\\0&0&0\\-1&0&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix} \Leftrightarrow x=z \Rightarrow \begin{matrix}\vec{u}_2=(0,1,0)\\ \vec{u}_3=\frac{1}{\sqrt{2}}(1,0,1)\end{matrix}$$

Con la transformación

$$\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}-\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\\0&1&0\\\frac{1}{\sqrt{2}}&0&\frac{1}{\sqrt{2}}\end{pmatrix}\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix}$$

la forma cuadrática se reduce a

$$4x_1^2+2y_1^2+2z_1^2-2\left(-\frac{x_1}{\sqrt{2}}+\frac{z_1}{\sqrt{2}}\right)-4y_1-2\left(\frac{x_1}{\sqrt{2}}+\frac{z_1}{\sqrt{2}}\right)-5=0,$$

o bien

$$4x_1^2+2y_1^2+2z_1^2-4y_1-4\frac{z_1}{\sqrt{2}}-5=0.$$

Completando cuadrados se obtiene

$$4x_1^2+2(y_1-1)^2+2\left(z_1-\frac{1}{\sqrt{2}}\right)^2-8=0$$

Haciendo la traslación

$$\begin{cases} x_2=x_1 \\ y_2=y_1-1 \\ z_2=z_1-\dfrac{1}{\sqrt{2}} \end{cases}$$

se obtiene

$$4x_2^2+2y_2^2+2z_2^2=8$$

o bien

$$\frac{x_2^2}{2}+\frac{y_2^2}{4}+\frac{z_2^2}{4}=1$$

que es la forma canónica de la superficie dada. Deducimos de esta expresión que la superficie dada es un *elipsoide* con semiejes $\sqrt{2}$, 2 y 2. Sus ejes se encuentran en las direcciones de $\vec{u}_1$, $\vec{u}_2$ y $\vec{u}_3$, respectivamente.

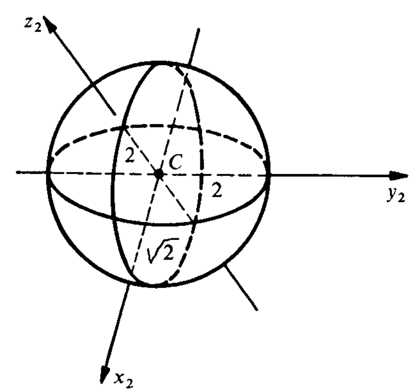
La transformación que transforma la superficie dada en su forma canónica es

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} \dfrac{-1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} x_2 \\ y_2+1 \\ z_2+\dfrac{1}{\sqrt{2}} \end{pmatrix}.$$

En el sistema de coordenadas $\{x_2, y_2, z_2\}$ el centro es el punto (0, 0, 0); por tanto, el centro del elipsoide es

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} \dfrac{-1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \dfrac{1}{\sqrt{2}} & 0 & \dfrac{1}{\sqrt{2}} \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ \dfrac{1}{\sqrt{2}} \end{pmatrix}=\begin{pmatrix} \dfrac{1}{2} \\ 1 \\ \dfrac{1}{2} \end{pmatrix}=C$$






## 13.2. INVARIANTES DE LAS SUPERFICIES DE SEGUNDO GRADO EN $\mathbb{R}^3$

El método descrito en la sección anterior para determinar las superficies de segundo grado puede resultar engorroso si los autovalores y autovectores no son sencillos. Al igual que en las cónicas, otra forma de estudiar las superficies de segundo grado es mediante la obtención de sus invariantes, permitiéndonos obtener su forma canónica sin necesidad de conocer la transformación que nos permite llegar a ella.

Para tratar de averiguar los invariantes de las superficies de segundo grado recordamos que los invariantes de la cónica $a_{11}x+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=0$ son

$$s=a_{11}+a_{22}\quad,\quad \delta=\begin{vmatrix} a_{11} & \dfrac{a_{12}}{2} \\ \dfrac{a_{12}}{2} & a_{22} \end{vmatrix}\quad \text{y}\quad \Delta=\begin{vmatrix} a_{11} & \dfrac{a_{12}}{2} & \dfrac{a_1}{2} \\ \dfrac{a_{12}}{2} & a_{22} & \dfrac{a_2}{2} \\ \dfrac{a_1}{2} & \dfrac{a_2}{2} & a \end{vmatrix}$$

y que los números $s$ y $\delta$ son, salvo el signo, los coeficientes del polinomio característico de la matriz

$$A=\begin{pmatrix} a_{11} & \dfrac{a_{12}}{2} \\ \dfrac{a_{12}}{2} & a_{22} \end{pmatrix}$$

(En efecto, $0=|A-\lambda I|=(a_{11}-\lambda)(a_{22}-\lambda)-\dfrac{a_{12}^2}{4}=\lambda_2-(a_{11}+a_{22})\lambda+|A|$.)

* * *

Sea $f=A(X, X)+L(X)+a=0$ una superficie de segundo grado en $\mathbb{R}^3$, donde $X=(x, y, z)$, $L(X)=a_1x+a_2y+a_3z$, $a\in\mathbb{R}$ y $A(X, X)$ *es la forma cuadrática que tiene como matriz*

$$A=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} \end{pmatrix}$$

Diremos que una expresión de los coeficientes de $f$ es un *invariante* si permanece fija mediante traslaciones y transformaciones ortogonales. (Observar que si $f$ está dada con respecto al sistema de referencia canónico, solamente se han utilizado transformaciones ortogonales y traslaciones en la sección 1 para reducirla a su forma canónica.)

PROPOSICIÓN 1

Son invariantes de una superficie de segundo grado en $\mathbb{R}^3$ los coeficientes del polinomio característico de la matriz $A$.

*Demostración.* Al realizar una transformación ortogonal de matriz $C$, la matriz de la forma cuadrática que aparece en $f(x, y, z)=0$ se transforma en $C^tAC$ (sección 1); como $C$ es ortogonal, $C^t=C^{-1}$ y, por tanto, $A$ se transforma en $C^{-1}AC$; por tanto, podemos decir que se ha realizado un cambio de base en la aplicación que tiene a $A$ como matriz en $\mathbb{R}^3$. Puesto que el polinomio característico de una matriz no depende de la base (capítulo 7, sección 2) deducimos que sus coeficientes quedan invariantes mediante transformaciones ortogonales.

Finalmente, cualquier expresión de los coeficientes de $A(X, X)$, y en particular los coeficientes de $|A-\lambda I|=0$, quedan invariantes mediante traslaciones. Para demostrarlo basta observar que la traslación

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}+\begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

transforma la superficie (1) de la sección 1 en

$$f=A(X_1, X_1)+\frac{\partial f}{\partial x}(\alpha, \beta, \gamma)x_1+\frac{\partial f}{\partial y}(\alpha, \beta, \gamma)y_1+\frac{\partial f}{\partial z}(\alpha, \beta, \gamma)z_1+f(\alpha, \beta, \gamma)=0$$

donde $X_1=(x_1, y_1, z_1)$, que deja invariante la forma cuadrática $A$. ■



Tenemos ahora que determinar los coeficientes del polinomio característico de la matriz $A$; un cálculo largo pero sencillo permite obtener

$$|A-\lambda I|=\begin{vmatrix} a_{11}-\lambda & \frac{a_{12}}{2} & \frac{a_{13}}{2} \\ \frac{a_{12}}{2} & a_{22}-\lambda & \frac{a_{23}}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33}-\lambda \end{vmatrix}=-\lambda^3+(a_{11}+a_{22}+a_{33})\lambda^2-$$

$$-\left(\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{13}}{2} \\ \frac{a_{13}}{2} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & \frac{a_{23}}{2} \\ \frac{a_{23}}{2} & a_{33} \end{vmatrix}\right)\lambda+|A|.$$

Por tanto, son invariantes de $f(x, y, z)=0$ los siguientes:

$$\begin{cases} s_1=a_{11}+a_{22}+a_{23} \\ s_2=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{13}}{2} \\ \frac{a_{13}}{2} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & \frac{a_{23}}{2} \\ \frac{a_{23}}{2} & a_{33} \end{vmatrix} \\ \delta=|A| \end{cases}$$

* * *

Para determinar un invariante más de las superficies de segundo grado en $\mathbb{R}^3$ escribimos

$$F(x, y, z, t)=a_{11}x^2+a_{22}y^2+a_{33}z^3+a_{12}xy+a_{13}xz+a_{23}yz$$
$$=a_1xt+a_2yt+a_3zt+at^2$$

y observamos que $f(x, y, z)=F(x, y, z, 1)$; por otro lado, $F$ es una forma cuadrática en $\mathbb{R}^4$ que tiene como matriz

$$\bar{A}=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & \frac{a_2}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_3}{2} & a \end{pmatrix},$$

a la cual se le denomina *matriz ampliada de f*.

PROPOSICIÓN 2

> $\Delta = |\bar{A}|$ es un invariante de las superficies de segundo grado en $\mathbb{R}^3$.

*Demostración.* Sea $C$ una transformación ortogonal en $\mathbb{R}^3$; la transformación de matriz

$$C_1 = \begin{pmatrix} C & 0 \\ 0 & 1 \end{pmatrix}$$

es ortogonal en $\mathbb{R}^4$ y transforma $\bar{A}$ en $C_1^t \bar{A} C_1 = C_1^{-1} \bar{A} C$; puesto que

$$|\bar{A} - \lambda I| = |C_1^{-1} A C_1 - \lambda I|$$

los coeficientes de $|\bar{A} - \lambda I|$ permanecen invariantes mediante $C_1$. El término independiente de $|\bar{A} - \lambda I|$ es $|\bar{A}|$ y, por tanto, éste es un invariante mediante transformaciones ortogonales de $\mathbb{R}^3$.

Si tenemos ahora la traslación

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix}$$

la matriz ampliada de $f$ en este nuevo sistema de referencia es (ver la demostración de la proposición 1):

$$\bar{A}_1 = \begin{pmatrix} a_{11} & \dfrac{a_{12}}{2} & \dfrac{a_{13}}{2} & \dfrac{1}{2}\dfrac{\partial f}{\partial x} \\ \dfrac{a_{12}}{2} & a_{22} & \dfrac{a_{23}}{2} & \dfrac{1}{2}\dfrac{\partial f}{\partial y} \\ \dfrac{a_{13}}{2} & \dfrac{a_{23}}{2} & a_{33} & \dfrac{1}{2}\dfrac{\partial f}{\partial z} \\ \dfrac{1}{2}\dfrac{\partial f}{\partial x} & \dfrac{1}{2}\dfrac{\partial f}{\partial y} & \dfrac{1}{2}\dfrac{\partial f}{\partial z} & f(\alpha, \beta, \gamma) \end{pmatrix}$$

donde las derivadas parciales están evaluadas en $(\alpha, \beta, \gamma)$. Un procedimiento análogo al utilizado en la demostración sobre los invariantes de las cónicas permite demostrar $|\bar{A}_1| = |\bar{A}|$, lo cual prueba que $|\bar{A}|$ es un invariante de $f$. ■

* * *



RESUMEN (*Invariantes de las superficies de segundo grado*)

> Son invariantes de la superficie de segundo grado
>
> $$f(x, y, z) = a_{11}x^2 + a_{22}y^2 + a_{33}z^2 + a_{12}xy + a_{13}xz + a_{12}yz + \\ + a_1x + a_2y + a_3z + a = 0 \qquad (1)$$
>
> los siguientes:
>
> $$s_1 = a_{11} + a_{22} + a_{33}\quad , \quad s_2 = \begin{vmatrix} a_{11} & \dfrac{a_{12}}{2} \\ \dfrac{a_{12}}{2} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & \dfrac{a_{13}}{2} \\ \dfrac{a_{13}}{2} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{22} & \dfrac{a_{23}}{2} \\ \dfrac{a_{23}}{2} & a_{33} \end{vmatrix}$$
>
> $$\delta = |A| = \begin{vmatrix} a_{11} & \dfrac{a_{12}}{2} & \dfrac{a_{13}}{2} \\ \dfrac{a_{12}}{2} & a_{22} & \dfrac{a_{23}}{2} \\ \dfrac{a_{13}}{2} & \dfrac{a_{23}}{2} & a_{33} \end{vmatrix} \quad \text{y} \quad \Delta = |\bar{A}| = \begin{vmatrix} a_{11} & \dfrac{a_{12}}{2} & \dfrac{a_{13}}{2} & \dfrac{a_1}{2} \\ \dfrac{a_{12}}{2} & a_{22} & \dfrac{a_{23}}{2} & \dfrac{a_2}{2} \\ \dfrac{a_{13}}{2} & \dfrac{a_{23}}{2} & a_{33} & \dfrac{a_3}{2} \\ \dfrac{a_1}{2} & \dfrac{a_2}{2} & \dfrac{a_3}{2} & a \end{vmatrix}$$

* * *

A continuación clasificamos las superficies de segundo grado en $\mathbb{R}^3$, también llamadas *cuádricas*, atendiendo a sus invariantes. Estudiamos por separado cada uno de los casos que aparecían en la sección 1.

En el *Caso I* la forma canónica es

$$\lambda_1 x_2^2 + \lambda_2 y_2^2 + \lambda_3 z_2^2 = C \qquad (\lambda_1\lambda_2\lambda_3 \neq 0).$$

Puesto que en esta ecuación

$$\Delta = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \lambda_3 & \\ 0 & & & -C \end{vmatrix} = -\lambda_1\lambda_2\lambda_3 C$$

se tiene que $C = -\Delta/\lambda_1\lambda_2\lambda_3$; además

$$\delta = \begin{vmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{vmatrix} = \lambda_1\lambda_2\lambda_3$$

y, por tanto, la forma canónica de las cuádricas del caso I es

$$\lambda_1 x_2^2 + \lambda_2 y_2^2 + \lambda_3 z_2^2 + \frac{\Delta}{\delta} = 0.$$

Si $\Delta \neq 0$ se obtienen *elipsoides, hiperboloides de una hoja* o *hiperboloides de dos hojas* dependiendo de los signos de $\lambda_1$, $\lambda_2$, $\lambda_3$, $\Delta$ y $\delta$.

Si $\Delta = 0$ se obtiene *una superficie cónica* o *un punto.*

* * *

En el *Caso II* la forma canónica es de la forma

$$\lambda_1 x_2^2 + \lambda_2 y_2^2 + b_3 z_2 = C \qquad (\lambda_1 \lambda_2 \neq 0)$$

Si $b_3 \neq 0$ esta puede escribirse de la forma

$$\lambda_1 x_3^2 + \lambda_2 y_3^2 + b_3 z_3 = 0.$$

Puesto que

$$\Delta = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & 0 & \frac{b_3}{2} \\ 0 & & \frac{b_3}{2} & 0 \end{vmatrix} = -\lambda_1 \lambda_2 \frac{b_3^2}{4}$$

y

$$s_2 = \begin{vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{vmatrix} + \begin{vmatrix} \lambda_1 & 0 \\ 0 & 0 \end{vmatrix} + \begin{vmatrix} \lambda_2 & 0 \\ 0 & 0 \end{vmatrix} = \lambda_1 \lambda_2$$

su forma canónica es

$$\lambda_1 x_3^2 + \lambda_2 y_3^2 \pm \sqrt{-4\frac{\Delta}{s_2}}\, z_3 = 0.$$

Puesto que $b_3 \neq 0$, $\Delta \neq 0$; como $\frac{\Delta}{s_2} = \frac{b_3^2}{4} < 0$ se obtienen *paraboloides elípticos o hiperbólicos* dependiendo de los signos de $\lambda_1$ y $\lambda_2$.

Si $b_3 = 0$ la forma canónica es

$$\lambda_1 x_2^2 + \lambda_2 y_2^2 = C;$$



por tanto,

$$s_2 = \lambda_1 \lambda_2 \neq 0 \quad , \quad \delta = \begin{vmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & 0 \end{vmatrix} = 0 \quad \text{y} \quad \Delta = \begin{vmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & 0 & \\ 0 & & & C \end{vmatrix} = 0$$

En este caso se obtienen *cilindros elípticos, hiperbólicos, dos planos que se cortan o una recta.* Su forma canónica no puede ser determinada con los invariantes de que disponemos.

* * *

En el *Caso III* la forma canónica es

$$\lambda_1 x_2^2 + b_2 y_2 + b_3 z_2 + b' = 0;$$

si $b_2 = b_3 = 0$ se deduce fácilmente que $s_2 = 0$, $\delta = 0$ y $\Delta = 0$ (en este caso se tienen *dos planos paralelos distintos o un solo plano*).

Si al menos uno de $b_2$ ó $b_3$ es no nulo, la forma canónica es

$$\lambda_1 x_4^2 - B y_4 = 0$$

y también se tiene $s_2 = 0$, $\delta = 0$ y $\Delta = 0$ (en este caso se tiene un *cilindro parabólico*).

* * *

Al final de este capítulo puede encontrarse un resumen de los resultados obtenidos sobre la clasificación de las cuádricas. Observar que la clasificación arriba realizada no es completa ya que no permite distinguir entre las superficies cilíndricas y los planos (ver notas adicionales en la última sección de este capítulo).

EJEMPLO A. *Reducir la superficie de segundo grado* $x^2 + y^2 + z^2 - 4xz - 4y + 2 = 0$ *a su forma canónica* y decir qué tipo de superficie representa.

Tenemos

$$\Delta = \begin{vmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 0 & -2 \\ -2 & 0 & 1 & 0 \\ 0 & -2 & 0 & 2 \end{vmatrix} = \begin{vmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 2 \end{vmatrix} - 2 \begin{vmatrix} 0 & -2 & 0 \\ 1 & 0 & -2 \\ -2 & 0 & 2 \end{vmatrix} =$$

$$-2 - 4[2 - 4] = 6 \neq 0$$

$$\delta = \begin{vmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{vmatrix} = 1 - 4 = -3 \neq 0.$$

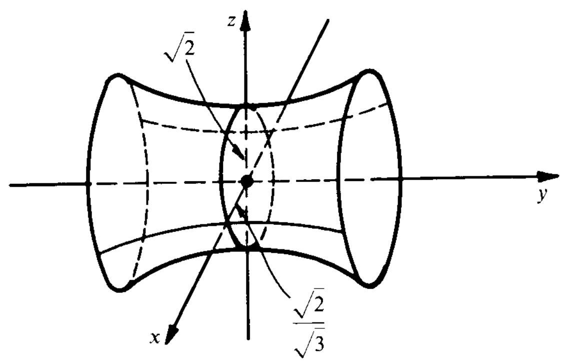
Puesto que los autovalores de la matriz

$$A=\begin{pmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}$$

son $\lambda_1=1$, $\lambda_2=-1$, $\lambda_3=3$ la forma canónica es

$$x^2-y^2+3z^2-2=0$$

o bien $\frac{x^2}{2}-\frac{y^2}{2}+\frac{z^2}{2/3}=1$ y se trata, por tanto, de un *hiperboloide de una hoja.*




EJEMPLO B. *Reducir la cuádrica $y^2+4xz+1=0$ a su forma canónica y decir qué tipo de superficie representa.*

Tenemos

$$\Delta=\begin{vmatrix} 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}=-4\neq 0 \quad , \quad \delta=\begin{vmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{vmatrix}=-4\neq 0$$

y los autovalores de

$$A=\begin{vmatrix} 0 & 0 & 2 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \end{vmatrix} \quad \text{son} \quad \lambda_1=1,\ \lambda_2=2,\ \lambda_3=-2.$$

La forma canónica de esta superficie es

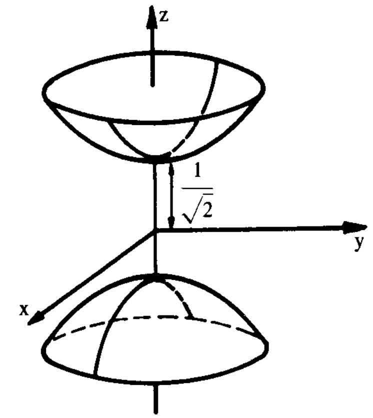
$$x^2+2y^2-2z^2+\frac{-4}{-4}=0$$



o bien

$$-x^2-\frac{y^2}{1/2}+\frac{z^2}{1/2}=1$$

y se trata, por tanto, de un *hiperboloide de dos hojas.*




* * *

## 13.3. DETERMINACION DE LOS ELEMENTOS GEOMETRICOS DE ALGUNAS CUADRICAS

Para las superficies de segundo grado en las que al menos uno de los números $\Delta$ ó $\delta$ es no nulo se ha dado un procedimiento para obtener su forma canónica en la sección anterior. A continuación describimos procedimientos para encontrar algunos elementos geométricos de estas superficies. Si $\delta\neq 0$ se trata de una cónica con *centro* y se dará un procedimiento para calcular éste; si $\delta=0$ se tienen los *paraboloides* (o cuádricas sin centro) y entonces se dará un procedimiento para calcular su *vértice.*

Sea

$$f(x, y, z)=A(X, X)+L(X)+a=0 \tag{1}$$

una superficie de segundo grado con $\Delta\neq 0$ ó $\delta\neq 0$. La dirección de sus ejes está dada por los autovectores de la matriz de $A$.

Sea $(\alpha, \beta, \gamma)$ el centro de una superficie de segundo grado con centro ($\delta\neq 0$); la traslación

$$x=x_1+\alpha \quad , \quad y=y_1+\beta \quad , \quad z=z_1+\gamma$$

transforma $(\alpha, \beta, \gamma)$ en $(0, 0, 0)$; por tanto $(0, 0, 0)$ es el centro de la superficie

$$f(x_1, y_1, z_1) = A(X_1, X_1) + \frac{\partial f}{\partial x}(\alpha, \beta, \gamma)x_1 + \frac{\partial f}{\partial y}(\alpha, \beta, \gamma)y_1 + \frac{\partial f}{\partial z}(\alpha, \beta, \gamma)z_1 + f(\alpha, \beta, \gamma) = 0 \tag{2}$$

que es la ecuación transformada de (1) mediante la traslación considerada; por tanto,

$$\frac{\partial f}{\partial x}(\alpha, \beta, \gamma) = 0 \quad , \quad \frac{\partial f}{\partial y}(\alpha, \beta, \gamma) = 0 \quad , \quad \frac{\partial f}{\partial z}(\alpha, \beta, \gamma) = 0$$

ya que en caso contrario $(0, 0, 0)$ no sería el centro de (2) (¿por qué?). Así pues, el centro de (1) satisface el sistema

$$\frac{\partial f}{\partial x} = 2a_{11}x + a_{12}y + a_{13}z + a_1 = 0$$
$$\frac{\partial f}{\partial y} = a_{12}x + 2a_{22}y + a_{23}z + a_2 = 0$$
$$\frac{\partial f}{\partial z} = a_{13}x + a_{23}y + 2a_{33}z + a_3 = 0$$

que posee una única solución ya que su matriz es $2A$ y $|A| = \delta \neq 0$.

EJEMPLO A. *Determinar el centro y los ejes del hiperboloide de una hoja del ejemplo A de la sección* 13.2: $x^2 + y^2 + z^2 - 4xz - 4y + 2 = 0$

Su centro satisface

$$\frac{\partial f}{\partial x} = 2x - 4z = 0$$
$$\frac{\partial f}{\partial y} = 2y - 4 = 0$$
$$\frac{\partial f}{\partial z} = 2z - 4x = 0$$

de donde deducimos $y = 2$, $x = 0$, $z = 0$. Por tanto, $C = (0, 2, 0)$.



Los autovalores de $A$ son $\lambda_1 = 1$, $\lambda_2 = -1$, $\lambda_3 = 3$ y las direcciones de sus ejes vienen dadas por $\vec{u}_1, \vec{u}_2, \vec{u}_3$ donde

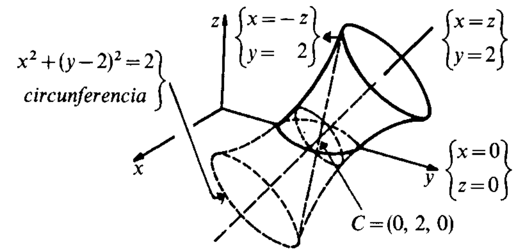
$$\lambda_1 = 1 \;;\; \begin{pmatrix} 0 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = 0 \\ z = 0 \end{Bmatrix} \;;\; \vec{u}_1 = (0, 1, 0)$$

$$\lambda_2 = -1 \;;\; \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = z \\ y = 0 \end{Bmatrix} \;;\; \vec{u}_2 = (1, 0, 1)$$

$$\lambda_3 = 3 \;;\; \begin{pmatrix} -2 & 0 & -2 \\ 0 & -2 & 0 \\ -2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = -z \\ y = 0 \end{Bmatrix} \;;\; \vec{u}_3 = (1, 0, -1)$$

Como los ejes pasan por el punto $C = (0, 2, 0)$ sus ecuaciones son

$$\lambda_1\!:\begin{Bmatrix} x = 0 \\ z = 0 \end{Bmatrix} \;;\; \lambda_2\!:\begin{Bmatrix} y = 2 \\ x = z \end{Bmatrix}. \;;\; \lambda_3\!:\begin{Bmatrix} y = 2 \\ x = -z \end{Bmatrix}$$




EJEMPLO B. *Determinar el centro y los ejes del hiperboloide de dos hojas del ejemplo B de la sección 13.2:* $y^2 + 4xz + 1 = 0$.

Su centro satisface $4z = 0$, $2y = 0$, $4x = 0$ y, por tanto, es el punto $C = (0, 0, 0)$. Los autovalores de su parte cuadrática son $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = -2$; las direcciones de sus ejes vienen dadas por $\vec{u}_1, \vec{u}_2, \vec{u}_3$ donde

$$\lambda_1 = 1 \;;\; \begin{pmatrix} -1 & 0 & 2 \\ 0 & 0 & 0 \\ 2 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = 2z \\ z = 2x \end{Bmatrix} \;;\; \vec{u}_1 = (0, 1,0)$$

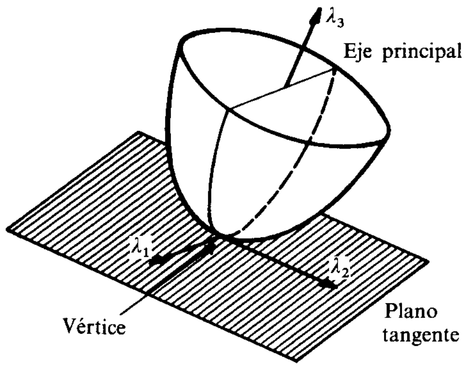
$$\lambda_2 = 2 \;;\; \begin{pmatrix} -2 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = z \\ y = 0 \end{Bmatrix} \;;\; \vec{u}_2 = (1, 0, 1)$$

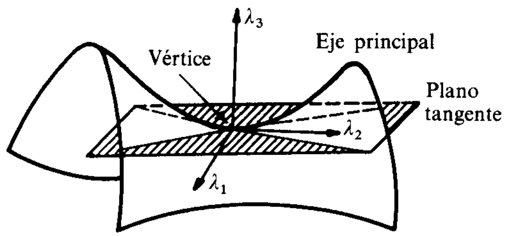
$$\lambda_2 = -2 \;;\; \begin{pmatrix} 2 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x = -z \\ y = 0 \end{Bmatrix} \;;\; \vec{u}_3 = (1, 0, -1)$$

La ecuación de su eje principal (el eje $z$ en la forma canónica) es $\{x=-z,\ y=0\}$ y los otros ejes son $\{x=0,\ z=0\}$, $\{x=z,\ y=0\}$.

* * *

Supongamos ahora que la superficie dada en (1) es un *paraboloide* ($\Delta \neq 0$, $\delta = 0$); el plano tangente al *vértice* del paraboloide tiene las direcciones de los autovectores correspondientes a los autovalores no nulos $(\lambda_1, \lambda_2)$ y el eje perpendicular a este plano tangente (que denominaremos eje principal) tiene la dirección de un autovector correspondiente a $\lambda_3=0$. Sea $(\omega_1, \omega_2, \omega_3)$ un autovector correspondiente a $\lambda_3=0$. Si $(\alpha, \beta, \gamma)$ es el vértice del paraboloide un vector perpendicular a $f$ en este punto viene dado por




$$\vec{N}=\left(\frac{\partial f}{\partial x}(\alpha, \beta, \gamma), \frac{\partial f}{\partial y}(\alpha, \beta, \gamma), \frac{\partial f}{\partial z}(\alpha, \beta, \gamma)\right)$$

Por tanto existe $\lambda \in \mathbb{R}$ tal que $\vec{N}=\lambda(\omega_1, \omega_2, \omega_3)$; si dejamos variable $\alpha$, $\beta$ y $\gamma$ en $\vec{N}$ se obtienen las ecuaciones lineales

$$\left.\begin{aligned} \omega_2\frac{\partial f}{\partial x} &= \omega_1\frac{\partial f}{\partial y} \\ \omega_3\frac{\partial f}{\partial x} &= \omega_1\frac{\partial f}{\partial z} \\ \omega_3\frac{\partial f}{\partial y} &= \omega_2\frac{\partial f}{\partial z} \end{aligned}\right\} \tag{2}$$

que determinan planos que pasan por el vértice y tienen $(\omega_1, \omega_2, \omega_3)$ como dirección común; el sistema (2) representa, por tanto, la ecuación del eje principal.

El *vértice* del paraboloide se determina hallando la intersección del eje principal (2) con la superficie.

EJEMPLO C. *Determinar el vértice, el eje principal y el plano tangente en el vértice del paraboloide* $x^2+2y^2+4xy+4z+3=0$, *así como su forma canónica.*



Tenemos

$$\Delta=\begin{vmatrix}1&2&0&0\\2&2&0&0\\0&0&0&2\\0&0&2&3\end{vmatrix}=-2\begin{vmatrix}1&2&0\\2&2&0\\0&0&2\end{vmatrix}=-4\begin{vmatrix}1&2\\2&2\end{vmatrix}=8\neq 0 \quad ; \quad \delta=|A|=\begin{vmatrix}1&2&0\\2&2&0\\0&0&0\end{vmatrix}=0$$

$$0=\begin{vmatrix}1-\lambda&2&0\\2&2-\lambda&0\\0&&0-\lambda\end{vmatrix}=-\lambda[(2-\lambda)(1-\lambda)-4]=-\lambda(\lambda^2-3\lambda-2)=0\Rightarrow$$

$$\lambda_1=\frac{3+\sqrt{17}}{2} \quad , \quad \lambda_2=\frac{3-\sqrt{17}}{2} \quad , \quad \lambda_3=0.$$

Puesto que $\sqrt{17}>3$, $\lambda_1>0$ y $\lambda_2<0$ con lo cual se trata de un *paraboloide hiperbólico.* Puesto que

$$s_2=\begin{vmatrix}1&2\\2&2\end{vmatrix}+\begin{vmatrix}1&0\\0&0\end{vmatrix}+\begin{vmatrix}2&0\\0&0\end{vmatrix}=-2$$

su forma canónica es

$$\frac{3+\sqrt{17}}{2}x^2+\frac{3-\sqrt{17}}{2}y^2\pm 4z=0.$$

El eje principal tiene como dirección:

$$\begin{pmatrix}1&2&0\\2&2&0\\0&0&0\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}0\\0\\0\end{pmatrix}\Leftrightarrow\left\{\begin{aligned}x+2y=0\\2x+2y=0\end{aligned}\right\} \quad ; \quad \vec{w}=(0, 0, 1).$$

Por tanto

$$\left(\frac{\partial f}{\partial x}, \frac{\partial f}{oy}, \frac{\partial f}{\partial z}\right)=\lambda(0, 0, 1)\Leftrightarrow(2x+4y,\ 4y+4x,\ 4)=\lambda(0, 0, 1).$$

Por tanto $\lambda=4$ y $2x+4y=0$, $4y+4x=0$; así pues $x=0$, $y=0$ y el eje principal es el eje $z$.

Sustituyendo $x=0$, $y=0$ en la superficie dada se obtiene el vértice:

$$z=-\frac{3}{4}.$$

Por tanto,

$$V=\left(0,\ 0,\ -\frac{3}{4}\right).$$

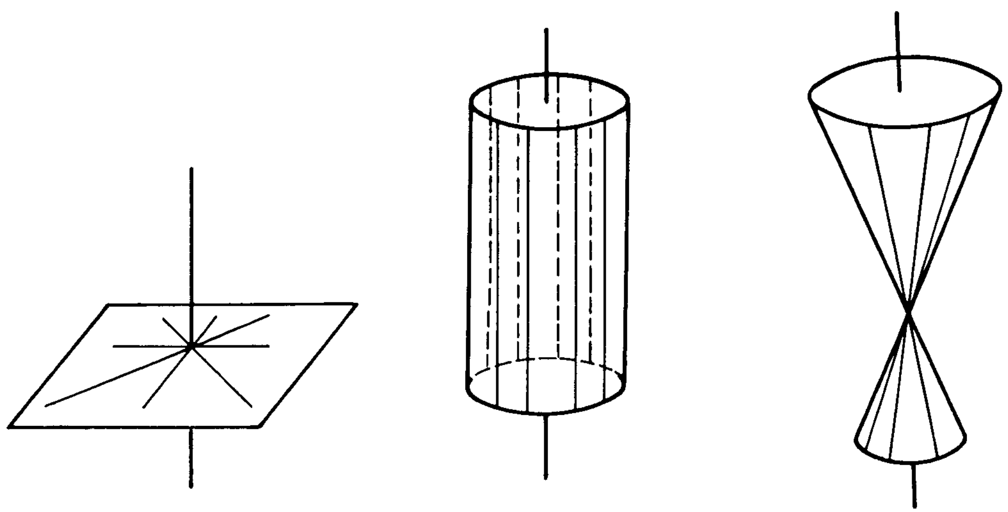
El plano tangente al paraboloide hiperbólico en $V$ tiene por ecuación

$$z=-\frac{3}{4}.$$

## 13.4. NOTAS ADICIONALES

### 1. El hiperboloide de una hoja como superficie reglada

Una superficie se denomina *superficie reglada* si por cada uno de sus puntos pasa al menos una recta totalmente contenida en la superficie. Ejemplos sencillos de superficies regladas son los «cilindros» estudiados en la sección 13.1. Otros ejemplos se obtienen girando una recta, llamada *generatriz*, alrededor de otra fija, llamada *eje*. Si la generatriz es perpendicular al eje se obtiene un *plano*, si es paralela se obtiene un *cilindro* y si le corta se obtiene un *doble cono recto* con vértice en el punto de intersección (ver las figuras adjuntas).



Podríamos ahora preguntaros cuál es la superficie que se obtiene cuando la generatriz no es paralela ni corta al eje; para tratar de identificar esta superficie encontremos sus ecuaciones en un caso sencillo. Supongamos que el eje coincide con el



eje $Oz$ y que la generatriz pasa por el punto $(a, 0, 0)$ y tiene como vector director $(0, \beta, 1)$ con $\beta \neq 0$.

Las ecuaciones paramétricas de la generatriz son

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ \beta \\ 1 \end{pmatrix}$$

y las ecuaciones del giro alrededor del eje $0z$ vienen dadas por

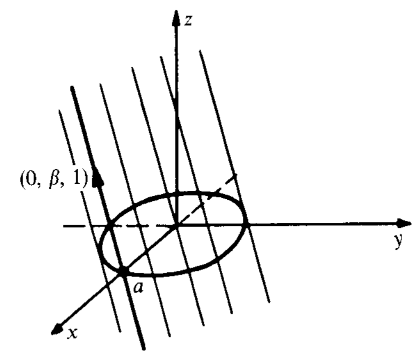
$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Por tanto,

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi & 0 \\ \operatorname{sen}\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \left[ \begin{pmatrix} a \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} 0 \\ \beta \\ 1 \end{pmatrix} \right] = \begin{pmatrix} a\cos\varphi \\ a\operatorname{sen}\varphi \\ 0 \end{pmatrix} + t \begin{pmatrix} -\beta\operatorname{sen}\varphi \\ \beta\cos\varphi \\ 1 \end{pmatrix}$$

como $z'=t$ deducimos que

$$(x')^2+(y')^2=(a\cos\varphi-z'\beta\operatorname{sen}\varphi)^2+(a\operatorname{sen}\varphi+z'\beta\cos\varphi)^2=a^2+\beta^2(z')^2$$




Por tanto, su ecuación es $(x')^2+(y')^2-\beta^2(z')^2=a^2$, que representa un *hiperboloide de una hoja* de semiejes $a$, $a$ y $\frac{a}{\beta}$ y de centro el origen de coordenadas.

Este resultado no es debido a una casualidad, sino que es cierto para todo hiperboloide reglado; en efecto, para demostrarlo observemos que es suficiente probarlo para un hiperboloide de ecuación

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}-\frac{z^2}{c^2}=1 \qquad (1)$$

ya que cualquier otro se reduce a éste mediante un movimiento y una traslación. Para el hiperboloide dado en (1) basta tomar la recta $(x, y, z)=(a, 0, 0)+t(0, \frac{b}{c}, 1)$, como puede comprobarse. (*Sugerencia*: hacer un giro elíptico.)

Observar que la recta $(x, y, z)=(a, 0, 0)+t(0, -\frac{b}{c}, 1)$ es también una generatriz del hiperboloide dado en (1) y que no es paralela a la anterior. Por tanto, un hiperboloide de una hoja tiene dos generatrices no paralelas.

* * *

## 2. CLASIFICACIÓN DE LAS CUÁDRICAS CUANDO $\Delta=0$ y $\delta=0$

En la sección 2 se han utilizado algunos invariantes de las superficies de segundo grado para obtener una fácil clasificación de ellas. El lector recordará que tal clasificación no era completa ya que era imposible encontrar la forma canónica cuando $\Delta=0$ y $\delta=0$ a partir de los invariantes encontrados en aquella sección.

El problema se resuelve demostrando que cuando $\Delta=\delta=0$ existe un nuevo invariante, a saber

$$s_3=\begin{vmatrix} a_{22} & \frac{a_{23}}{2} & \frac{a_2}{2} \\ \frac{a_{23}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_2}{2} & \frac{a_3}{2} & a \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{13}}{2} & \frac{a_1}{2} \\ \frac{a_{13}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_1}{2} & \frac{a_3}{2} & a \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a \end{vmatrix}.$$

Si estamos en el caso en que un autovalor es cero, $\lambda_3=0$, por ejemplo, teníamos $\lambda_1 x^2+\lambda_2 y^2+c=0$ como forma canónica y, se obtenían cilindros elípticos o hiperbólicos $(c\neq 0)$ o dos planos que se cortan o una recta $(c=0)$. En estos casos

$$s_2=\begin{vmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{vmatrix}+\begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}+\begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}=\lambda_1\lambda_2\neq 0$$

y

$$s_3=0+0+\begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & c \end{vmatrix}=\lambda_1\lambda_2 c=s_2 c.$$

Por tanto $c=\frac{s_3}{s_2}$ y si $s_3\neq 0$ se obtienen *cilindros elípticos* o *hiperbólicos* y si $s_3=0$ se obtienen *dos planos que se cortan* o *una recta*.

* * *



Si estamos en el caso en que dos autovalores son cero, $\lambda_2=\lambda_3=0$, por ejemplo, tenemos $\Delta=\delta=s_2=0$; si la forma canónica es

$$\lambda_1 x^2-By=0$$

se tiene que

$$s_3=0+0+\begin{vmatrix} \lambda_1 & 0 & 0 \\ 0 & 0 & -B/2 \\ 0 & -B/2 & 0 \end{vmatrix}=-\lambda_1\frac{B^2}{4}\neq 0$$

y entonces tenemos un *cilindro parabólico*; por el contrario, si la forma canónica es

$$\lambda_1 x^2+C=0$$

se tiene que $s_3=0$ y entonces si tienen dos *planos paralelos* o un *solo plano*.

* * *

EJEMPLO A. *Estudiar (o clasificar) la superficie de segundo grado*

$$x^2-y^2-z^2-2yz-x+3y+3z-2=0$$

Encontramos sus invariantes:

$$\delta=\begin{vmatrix} 1 & 0 & 0 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \end{vmatrix}=0 \quad ; \quad \Delta=\begin{vmatrix} 1 & 0 & 0 & -1/2 \\ 0 & -1 & -1 & 3/2 \\ 0 & -1 & -1 & 3/2 \\ -1/2 & 3/2 & 3/2 & -2 \end{vmatrix}=0 \quad ;$$

$$s_2=\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}+\begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}+\begin{vmatrix} -1 & -1 \\ -1 & -1 \end{vmatrix}+-1-1+1-1=-2\neq 0 \quad ;$$

$$s_3=\begin{vmatrix} 1 & 0 & -1/2 \\ 0 & -1 & 3/2 \\ -1/2 & 3/2 & -2 \end{vmatrix}+\begin{vmatrix} 1 & 0 & -1/2 \\ 0 & -1 & 3/2 \\ -1/2 & 3/2 & -2 \end{vmatrix}+\begin{vmatrix} -1 & -1 & 3/2 \\ -1 & -1 & 3/2 \\ 1/2 & 3/2 & -2 \end{vmatrix}=0+0+0=0.$$

Se trata entonces de *dos planos que se cortan* o *una recta*; como sus autovalores son

$$\lambda_1=1,\ \lambda_2=-2,\ \lambda_3=0$$

ya que

$$\begin{vmatrix} 1-\lambda & 0 & 0 \\ 0 & -1-\lambda & -1 \\ 0 & -1 & -1-\lambda \end{vmatrix}=(1-\lambda)[(-1-\lambda)^2-1]=(1-\lambda)(2\lambda+\lambda^2)=(1-\lambda)\lambda(2+\lambda),$$

la forma canónica de esta cuádrica es $x^2-2y^2=0$ y, por tanto, se trata de *dos planos que se cortan.*

Para encontrar las ecuaciones de estos planos despejamos $x$ en la superficie dada considerando $y$, $z$ como fijas:

$$x=\frac{1\pm\sqrt{1-4[-y^2-z^2-2yz+3y+3z-2]}}{2}=\frac{1}{2}\pm\left(y+z-\frac{3}{2}\right).$$

Por tanto las ecuaciones de los planos son

$$x=y+z-1 \quad ; \quad x=-y-z+2$$

y la superficie dada puede escribirse de la forma

$$(x+y+z-2)(x-y-z+1)$$

* * *

## CLASIFICACIÓN DE LAS CUADRICAS

$$a_{11}x^2+a_{22}y^2+a_{33}z^2+a_{12}xy+a_{13}xz+a_{23}yz+a_1x+a_2y+a_3z+a=0$$

$$A=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} \end{pmatrix} \quad ; \quad \bar{A}=\begin{pmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_{13}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_{23}}{2} & \frac{a_2}{2} \\ \frac{a_{13}}{2} & \frac{a_{23}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & \frac{a_3}{2} & a \end{pmatrix}$$

INVARIANTES: $\Delta=|\bar{A}|$, $\delta=|A|$, $s_1=a_{11}+a_{22}+a_{33}$

$$s_2=\begin{vmatrix} a_{11} & \frac{a_{12}}{2} \\ \frac{a_{12}}{2} & a_{22} \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{13}}{2} \\ \frac{a_{13}}{2} & a_{33} \end{vmatrix}+\begin{vmatrix} a_{22} & \frac{a_{23}}{2} \\ \frac{a_{23}}{2} & a_{33} \end{vmatrix}$$

$$s_3=\begin{vmatrix} a_{22} & \frac{a_{23}}{2} & \frac{a_2}{2} \\ \frac{a_{23}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_2}{2} & \frac{a_3}{2} & a \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{13}}{2} & \frac{a_1}{2} \\ \frac{a_{13}}{2} & a_{33} & \frac{a_3}{2} \\ \frac{a_1}{2} & \frac{a_3}{2} & a \end{vmatrix}+\begin{vmatrix} a_{11} & \frac{a_{12}}{2} & \frac{a_1}{2} \\ \frac{a_{12}}{2} & a_{22} & \frac{a_2}{2} \\ \frac{a_1}{2} & \frac{a_2}{2} & a \end{vmatrix}$$



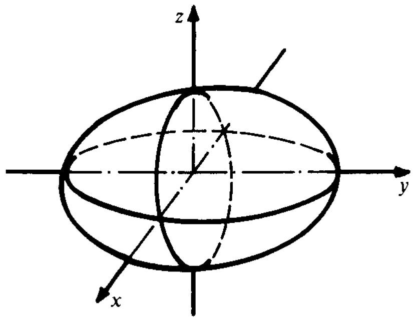
## RESUMEN

*Foma canónica*

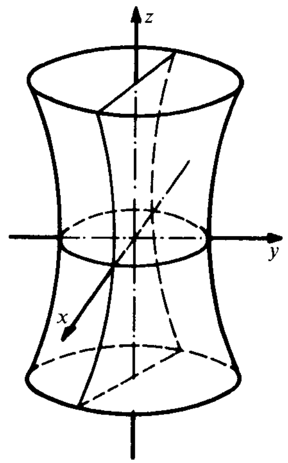
- $\Delta\neq 0$
  - $\delta\neq 0$: Elipsoide; Hiperboloide de una hoja; Hiperboloide de dos hojas — $\lambda_1X^2+\lambda_2Y^2+\lambda_3Z^2+\frac{\Delta}{\delta}=0$
  - $\delta=0$: Paraboloide elíptico; Paraboloide hiperbólico — $\lambda_1X^2+\lambda_2Y^2\pm\sqrt{-\frac{4\Delta}{\lambda_1\lambda_2}}\,Z=0$
- $\Delta=0$
  - $\delta\neq 0$: Un punto; Un cono — $\lambda_1X^2+\lambda_2Y^2+\lambda_3Z^2=0$
  - $\delta=0$
    - $S_2=0$
      - $s_3\neq 0$: cilindro elíptico o hiperbólico; $\lambda_1X^2+\lambda_2Y^2+\frac{s_3}{s_2}=0$
      - $s_3=0$: Dos planos que se cortan o una recta; $\lambda_1X^2+\lambda_2Y^2=0$
    - $S_2\neq 0$
      - $s_3\neq 0$: Cilindro parabólico; $\lambda_1X^2-BY=0$
      - $s_3=0$: Dos planos paralelos o un plano; $\lambda_1X^2+C=0$

* * *

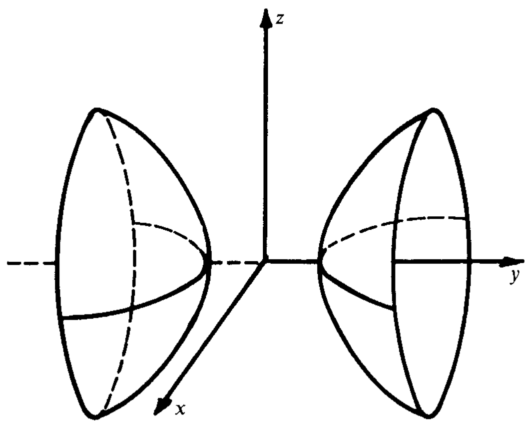
$\lambda_1$, $\lambda_2$, $\lambda_3$ son los autovalores de la matriz $A$.



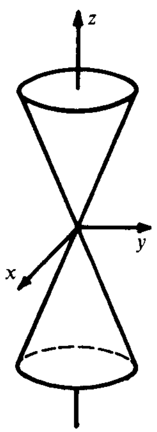

ELIPSOIDE



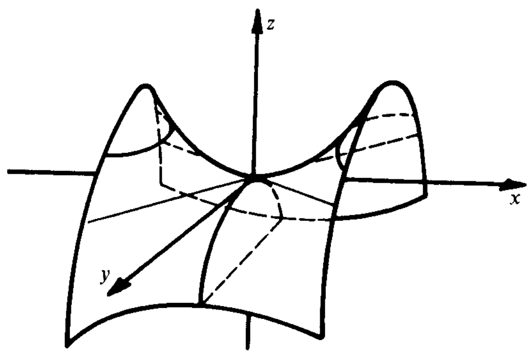

HIPERBOLOIDE REGLADO
O DE UNA HOJA




HIPERBOLOIDE DE DOS HOJAS

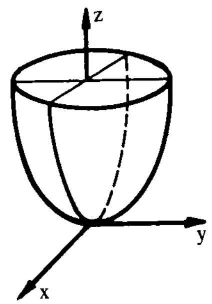



CONO ELIPTICO

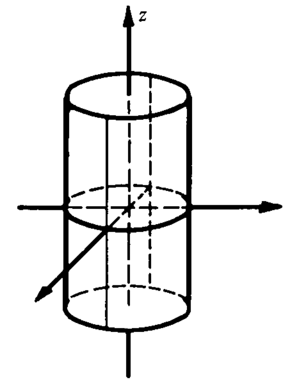



PARABOLOIDE HIPERBÓLICO

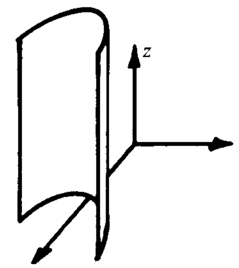




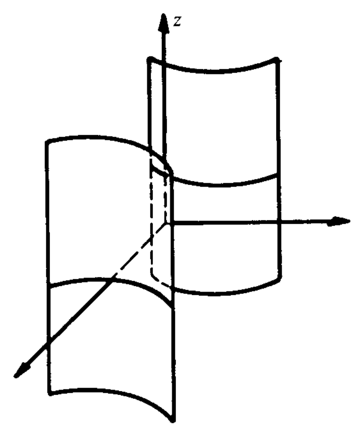

PARABOLOIDE ELIPTICO



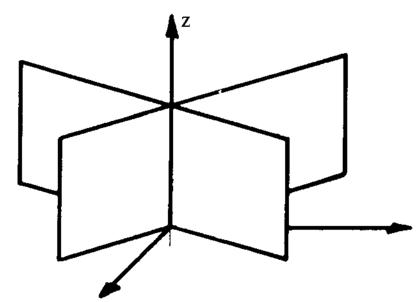

CILINDRO ELIPTICO




CILINDRO PARABÓLICO




CILINDRO HIPERBÓLICO




DOS PLANOS QUE SE CORTAN

EJERCICIOS (CAPITULO 13)

**1.** Hacer un estudio lo más detallado posible de las superficies de segundo grado que se indican (encontrar tipo, forma canónica, ejes, centro, ...):

*a)* $2x^2+2y^2+2z^2+2xz-2xy-2yz=1$

*b)* $2x^2-6y^2-2z^2-2xz+10x-6y-1=0$

*c)* $-2y^2+xz-4y+6z+5=0$

*d)* $2x^2+2y^2-4z^2-5xy-2xz-2yz-2x-2y+z=0$

**2.** Clasificar la cuádrica $3x^2+2y^2+2z^2+2yz+4x+2y-5z+7=0$ encontrando sus ejes, si los tiene.

**3.** Considerar las cuádricas que admiten por ecuaciones

$$x^2-2y^2+\alpha z^2-2xz+2yz+2x+1=0$$

para $\alpha$ real. Estudiar para que valores de $\alpha$ la cuádrica anterior es un paraboloide (elíptico o hiperbólico) y encontrar la ecuación del eje principal.

**4.** Indicar el tipo y la forma canónica de las cuádricas dadas por las siguientes matrices:

$$a)\ \begin{pmatrix} 2 & 1 & -2 & -1 \\ 1 & 2 & -2 & -1 \\ -2 & -2 & 5 & 3 \\ -1 & -1 & 3 & 2 \end{pmatrix} \qquad b)\ \begin{pmatrix} -2 & -1 & 2 & 1 \\ -1 & 0 & 2 & 1 \\ 2 & 2 & -3 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix}$$

$$c)\ \begin{pmatrix} -2 & -1 & 2 & 1 \\ -1 & -2 & 2 & 1 \\ 2 & 2 & 3 & -1 \\ 1 & 1 & -1 & 0 \end{pmatrix} \qquad d)\ \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & -1 \\ -1 & -1 & 4 & 3 \\ -1 & -1 & 3 & 2 \end{pmatrix}$$

$$e)\ \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & -1 & -1 & -1 \\ -1 & -1 & 4 & 3 \\ -1 & -1 & 3 & 2 \end{pmatrix} \qquad f)\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$g)\ \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \qquad h)\ \begin{pmatrix} 0 & -1 & 2 & 1 \\ -1 & -2 & 2 & 1 \\ 2 & 2 & -4 & -2 \\ 1 & 1 & -2 & -1 \end{pmatrix}$$

# EJERCICIOS DE REPASO: CAPITULOS 8 A 13

A continuación presentamos varios ejercicios, que pueden servir para repasar los conceptos introducidos en los capítulos 8 a 13. Los ejercicios del 1 al 24 están ordenados de acuerdo con el orden de los capítulos de este libro. Del 25 al 36 el grado de dificultad es mayor que en los anteriores. El resto son problemas que se han propuesto en varias convocatorias a los alumnos de primer curso de las licenciaturas de Matemáticas y Físicas de la Universidad Autónoma de Madrid.

**1.** Diagonalizar las matrices dadas mediante un cambio de base ortonormal e indicar el cambio de base.

$$a)\ A=\begin{pmatrix} 7 & 2 & 0 \\ 2 & 6 & 2 \\ 0 & 2 & 5 \end{pmatrix} \qquad b)\ B=\begin{pmatrix} 8 & -3 & -3 \\ -3 & 8 & -3 \\ -3 & -3 & 8 \end{pmatrix}$$

**2.** Demostrar que todos los autovalores de una aplicación antisimétrica no singular son imaginarios puros. (*Nota*: $\mathscr{A}: E\mapsto E$ es antisimétrica si $(\mathscr{A}\vec{x}, \vec{y})=-(\vec{x}, \mathscr{A}\vec{y})$.)

**3.** Aplicar el proceso de ortogonalización en $\mathbb{R}^4$ con el producto escalar usual a los vectores

$$\vec{u}_1=(1, 2, 2, -1) \quad , \quad \vec{u}_2=(1, 1, -5, 3) \quad , \quad \vec{u}_3=(3, 2, 8, -7)$$

**4.** Encontrar el coseno del ángulo que forman los planos

$$\pi_1 = \text{plano generado por } (1, 0, 0) \text{ y } (0, 1, 1)$$
$$\pi_2 = \text{plano generado por } (1, 1, 0) \text{ y } (1, -1, 1)$$

**5.** Demostrar que la suma de los cuadrados de las diagonales de un paralelogramo coincide con la suma de los cuadrados de todos sus lados.

**6.** Si $\vec{x}=(x_1, y_1, z_1)$, $\vec{y}=(x_2, y_2, z_2)\in\mathbb{R}^3$ demostrar que

$$(\vec{x}, \vec{y})=2x_1x_2+x_2y_1+x_1y_2+2y_1y_2+z_1z_2$$

es un producto escalar en $\mathbb{R}^3$; calcular la norma del vector $\vec{v}=(1, -1, -1)$ con este producto escalar y el subespacio ortogonal el subespacio generado por $\vec{v}$.

**7.** Sea $f$: $\mathbb{R}^3\mapsto\mathbb{R}^3$ la transformación ortogonal que respecto de la base canónica viene dada por la matriz.

$$A=\frac{1}{3}\begin{pmatrix} -1 & -2 & 2 \\ -2 & -1 & -2 \\ 2 & -2 & -1 \end{pmatrix}.$$

Decir qué tipo de transformación ortogonal es.

**8.** Encontrar dos subespacio invariantes $W_1$ y $W_2$ de la aplicación $T$: $\mathbb{R}^3\mapsto\mathbb{R}^3$ con matriz

$$T=\frac{1}{3}\begin{pmatrix} -1 & 2 & .5 \\ 0 & 6 & 0 \\ -2 & 7 & 1 \end{pmatrix}$$

y tal que $\dim W_1=1$, $\dim W_2=2$.

**9.** Sea la aplicación $T$: $\mathbb{R}^2\mapsto\mathbb{R}^2$ que tiene como matriz

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \text{ con respecto a la base } \vec{u}_1=\begin{pmatrix} 1 \\ 1 \end{pmatrix} \ , \ \vec{u}_2=\begin{pmatrix} 2 \\ -1 \end{pmatrix}$$

Encontrar la matriz de la aplicación adjunta , $T^*$, de $T$ con respecto a la base $\{\vec{u}_1, \vec{u}_2\}$ cuando en $\mathbb{R}^3$ se considera el producto escalar usual.

**10.** Sea $(\vec{x}, \vec{y})=(x_1+x_2)(y_1+y_2)+x_2y_2+x_3y_3$ un producto escalar $\mathbb{R}^3$, donde $\vec{x}=x_1\vec{e}_1+x_2\vec{e}_2+x_3\vec{e}_3$, $\vec{y}=y_1\vec{e}_1+y_2\vec{e}_2+y_3\vec{e}_3$ y $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ es la base canónica en $\mathbb{R}^3$. Encontrar la proyección del vector $\vec{v}=\vec{e}_1+\vec{e}_2+\vec{e}_3$ sobre el plano $\pi$: $x_2+x_3=0$ con el producto escalar dado.

**11.** Sea $\vec{u}_1=\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ , $\vec{u}_2=\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ una base de $\mathbb{R}^2$; decir si la aplicación lineal $T$: $\mathbb{R}^2\mapsto\mathbb{R}^2$ tal que $T$ tiene como matriz

$$T=\begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$$

con respecto a la base $\{u_1, u_2\}$ es autoadjunta. ¿Es $T$ una matriz ortogonal?



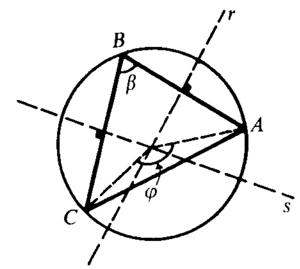
**12.** Clasificar el movimiento

$$M(P)=\begin{pmatrix} 1 \\ 1 \end{pmatrix}+\begin{pmatrix} \dfrac{2}{\sqrt{5}} & \dfrac{1}{\sqrt{5}} \\ -\dfrac{1}{\sqrt{5}} & \dfrac{2}{\sqrt{5}} \end{pmatrix}(P)$$

y encontrar sus elementos geométricos.

**13.** Demostrar que el ángulo $\beta$ es la mitad del arco que abarcan sus lados, es decir $\beta=\dfrac{\varphi}{2}$. (Ver figura adjunta.)

[*Sugerencia.* La composición $S_s$ o $S_r$ es un giro de ángulo doble del que forman $r$ y $s$, donde $r$ y $s$ son las mediatrices de $\overline{AB}$ y $\overline{BC}$, respectivamente.]




**14.** Hallar las ecuaciones de la simetría en el plano con respecto a la recta $x+y=3$.

**15.** Encontrar las ecuaciones de la simetría con respecto al plano $x+y+z=1$.

**16.** *a*) Encontrar las ecuaciones del movimiento $M$ que se obtiene como composición de la simetría con respecto al plano $2x-y-z=3$ y la traslación de vector $\vec{v}=(-2, 1, 3)$.

*b*) Decir qué tipo de movimiento es $M$ y encontrar sus elementos geométricos.

**17.** Determinar los valores de $\lambda$ para los que las formas cuadráticas dadas son definidas positivas:

*a*) $5x_1^2+x_2^2+\lambda x_3^2+4x_1x_2-2x_1x_3-2x_2x_3$

*b*) $x_1^2+4x_2^2+x_3^2+2\lambda x_1x_2+10x_1x_3+6x_2x_3$

**18.** Encontrar una base ortonormal en la cual la forma cuadrática $3x_2^2+3x_3^2+4x_1x_2+4x_1x_3-2x_2x_3$ se reduce a su forma canónica y encontrar esta forma canónica.

**19.** Clasificar la cónica $16x^2-24xy+9y^2-30x+40y=5$ y determinar su forma canónica.

**20.** Demostrar que la cónica $11x^2-24xy+4y^2+6x+8y=-15$ es una hipérbola y determinar su centro y su eje principal.

**21.** Clasificar la cuádrica $-3z^2-2xy+4xz+4yz+2x+2y-2z=0$ indicando su forma canónica.

**22.** Clasificar la cónica $x^2-2xy+2y-2=0$ indicando su forma canónica.

**23.** Reducir a una suma de cuadrados la forma cuadrática $2xy+2z^2$ mediante una transformación ortogonal e indicar la transformación.

**24.** Determinar los valores de $a$ y $b$ para los cuales la aplicación bilineal siguiente define un producto escalar en $\mathbb{R}^2$ que hace de este espacio un espacio euclídeo:

$$F\colon \mathbb{R}^2\times\mathbb{R}\mapsto\mathbb{R}\ /\ F\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\begin{pmatrix}y_1\\y_2\end{pmatrix}\right)=x_1y_1+ax_1y_1+(b-2)x_2y_1+(a^2+b)x_2y_2$$

* * *

Los ejercicios siguientes son de mayor dificultad que los anteriores.

**25.** Demostrar algebraicamente que si $M$ es un movimiento directo en el espacio con tres puntos fijos distintos y no alineados, $M$ es la identidad.

**26.** Sea $A=(a_{ij})_{i,j=1}^n$ una matriz real tal que $a_{ij}>0$ y $a_{i1}+a_{i2}+\cdots+a_{in}=1$ para todo $i,j=1,\ldots,n$. Demostrar que todos los autovalores de $A$ son en módulo menor o igual que 1.

**27.** Utilizar el problema 2 para demostrar que si $S$ es una matriz antisimétrica $(I+S)$ es invertible; demostrar, además, que $(I+S)^{-1}(I-S)$ es una matriz ortogonal con determinante $+1$.
Utilizar esta construcción para obtener una matriz ortogonal a partir de la matriz

$$s=\begin{pmatrix}0 & \operatorname{tg}\frac{1}{2}\alpha\\ -\operatorname{tg}\frac{1}{2}\alpha & 0\end{pmatrix}$$

**28.** *a*) Hallar la ecuación de la superficie que se obtiene al girar alrededor del eje $z$ la recta que pasa por el punto $(a, 0, 0)$ y que tiene como vector director $(\alpha, \beta, 1)$.
*b*) Demostrar que si $\beta\neq 0$ la superficie anterior es un hiperboloide de una hoja; encontrar su centro y sus ejes.
*Nota.* Una superficie con la propiedad de que por cada uno de sus puntos pasa al menos una recta contenida en la superficie se denomina una *superficie reglada.* ¿Qué superficies regladas conoces?

**29.** Demostrar que todos los autovalores de una matriz real simétrica $A$ están en el intervalo $[a, b]$ si y sólo si la forma cuadrática con matriz $A-\lambda_0 I$ es definida positiva para todo $\lambda_0<a$ y definida negativa para todo $\lambda_0>b$.



**30.** *a*) Sea $l$ el lugar geométrico de los puntos medios de todas las cuerdas paralelas al eje $OX$ de la cónica de ecuación

$$a_{11}x^2+a_{12}xy+a_{22}y^2+a_1x+a_2y+a=0 \qquad (1)$$

donde $a_{11}\neq 0$. Demostrar que $l$ está contenido en una recta y encontrar la ecuación de esta recta (*Sol.*: $2a_{11}x+a_{12}y+a_1=0$).
*b*) Demostrar que si (1) es una cónica *con centro*, la recta encontrada en *a*) pasa por el centro de la cónica. (Esta recta se llama *recta diametral.*)
*c*) encontrar la recta diametral de
1) $3x^2-2xy+3y^2+2x-4y+1=0$
2) $x^2-2xy+y^2+4x-6y+1=0$

**31.** Hallar el ángulo que forma la diagonal de un cubo con una de sus aristas. Demostrar que la proyección ortogonal de una arista del cubo sobre una diagonal coincide, en valor absoluto, con $\frac{1}{3}$ de la longitud de la diagonal.

**32.** En $\mathcal{M}_2(\mathbb{R})$, el conjunto de las matrices cuadradas de orden 2 con elementos reales, demostrar que $(A, B)=$ traza $(AB^t)$ es un producto escalar.
Encontrar una base ortonormal para el subespacio $M=\{A\in\mathcal{M}_2(\mathbb{R})/\text{traza}\,(A)=0\}$.

**33.** Determinar todos los movimientos planos que transforman $A=(2,0)$ en $A'(1, 2)$ y $B=(0, 0)$ en $B'=(1, 0)$, precisando tipo, subespacios invariantes y elementos geométricos.

**34.** Sea $V$ un espacio vectorial euclídeo de dimensión finita y $V$ y $W$ dos subespacios suplementarios de $V$. Probar que la proyección sobre $W$ paralelamente a $U$ es autoadjunta si y sólo si $U$ y $V$ son ortogonales.

**35.** Comprobar que todo movimiento del plano puede obtenerse como composición de tres simetrías, y que 2 son suficientes para los movimientos directos.

**36.** En un espacio vectorial euclídeo $E$ cuyo producto escalar tiene como matriz

$$\begin{pmatrix}1 & 0 & 0\\ 0 & 2 & -1\\ 0 & -1 & 2\end{pmatrix}$$

en la base $\{\vec{u}_1, \vec{u}_2, \vec{u}_3\}$, hallar la proyección ortogonal $\vec{u}$ de $\vec{u}_3$ sobre $L\{\vec{u}_1, \vec{u}_2\}$.

**37.** Determinar de las siguientes formas cuadráticas de $\mathbb{R}^3$ si son definidas positivas, definidas negativas, semidefinidas positivas, semidefinidas negativas o indefinidas:

*a)* $x^2+y^2+z^2+2xy+4y^2$

*b)* $x^2+6y+z^2+2xy+4y^2$

*c)* $x^2+y^2+z^2+2xy$

**38.** Clasificar el movimiento $M$ de ecuación

$$M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

y escribirlo en forma canónica si es posible.

**39.** Clasificar la cónica $11x^2-16xy-y^2+2y=0$; hallar su forma canónica y su eje focal.

**40.** Se consideran las formas cuadráticas de $\mathbb{R}^2$, $P$ y $Q$, dadas en el sistema canónico de coordenadas $(x, y)$ por $5x^2+2y^2+2xy$ e $y^2-2x^2-8xy$, respectivamente. Determinar un sistema de coordenadas $(x', y')$ de la forma $x'=ax+by$, $y'=cx+dy$, $a, b, c, d \in \mathbb{R}$, en el que se tenga $P=x'^2+y'^2$, $Q=\lambda x'^2+\mu y'^2$ para algunos $\lambda, \mu \in \mathbb{R}$ que también se determinarán (la solución estará completa cuando se determinen $a, b, c, d, \lambda, \mu \in \mathbb{R}$ para los que se cumpla lo pedido).

**41.** ¿Existe una transformación ortogonal o una autoadjunta de $\mathbb{R}^3$ en $\mathbb{R}^3$ que transforme $(1, 1, 1)$ en $(\sqrt{3}, 0, 0)$ y al mismo tiempo $(2, -1, -1)$ en $(\sqrt{5}, 1, 0)$?

**42.** Dada la forma cuadrática $Q(x, y, z)=35x^2+55y^2+4z^2-42xy+12xz-4yz$ encontrar una base tal que $Q$ tenga la expresión

$$Q(x', y', z')=\lambda x'^2+\mu y'^2+\nu z'^2$$

y decir si es definida.

**43.** Sea $M$ el movimiento de $\mathbb{R}^3$ dado por

$$M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}+\begin{pmatrix} \frac{3}{5} & 0 & \frac{4}{5} \\ \frac{4}{5} & 0 & -\frac{3}{5} \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

*a)* Determinar el tipo de movimiento.

*b)* Hallar los subespacios invariantes y determinar los elementos geométricos.



**44.** Clasificar la cónica de ecuación

$$x^2+2xy+y^2-6\sqrt{2}\,x+6\sqrt{2}\,y=0$$

y encontrar su forma canónica, ejes, focos, vértices, ...

**45.** Dada la cónica de ecuación

$$\alpha x^2+2\alpha xy+y^2-2x-2y+\alpha=0$$

determinar los valores del parámetro $\alpha$ para los cuales representa un par de rectas, y hallar dichas rectas.

# SOLUCIONES

En las páginas siguientes se encuentran los resultados finales de muchos de los problemas propuestos en este libro. Esperamos que sirvan al lector como comprobación de su trabajo diario. Es necesario observar que aquellos problemas en los que se pide la demostración de algún resultado no están, por su naturaleza, incluidos en estas soluciones.

Como ya se dijo en el prólogo, la matemática *no* es un deporte para espectadores. El intento repetido de solucionar un problema produce mayores beneficios que la memorización de la teoría. Esperamos que los numerosos ejercicios acumulados al final de casi todas las secciones de este libro sirvan para que el lector no sea un mero espectador.

***CONSEJO: Mirar las páginas siguientes sólo después de intentar solucionar los problemas propuestos.***

## CAPITULO 1

### 1.1

1. *a*) $x_1=0,\ x_2=1.$ *b*) $x_1=-1,\ x_2=2.$

2. *a*) $\begin{pmatrix} 1 & 0 & 32/23 & 0 \\ 0 & 1 & 43/23 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ *b*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

   *c*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ *d*) $\begin{pmatrix} 1 & 0 & -1 & 0 & -4 \\ 0 & 1 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 1/2 \end{pmatrix}$

3. *a*) $x_1=1,\ x_2=1,\ x_3=0.$ *b*) $x_3=c,\ x_2=-(1+c),\ x_1=3-2c.$
   *c*) Incompatible. *d*) $x_1=21,\ x_2=24,\ x_3=4,\ x_4=3.$
   *e*) $x_1=(15+3c)/8,\ x_2=(11-25c)/8,\ x_3=0,\ x_4=c.$
   *f*) $x_1=x_2=x_3=1.$ *g*) $x_1=-1,\ x_2=0,\ x_3=1.$
   *h*) $x_1=\frac{2}{5}c,\ x_2=\frac{4}{5}c,\ x_3=c.$ *i*) Incompatible.
   *j*) $x_1=3,\ x_2=2,\ x_3=1.$
4. *a*) Compatible indeterminado. *b*) Compatible indeterminado.
   *c*) Incompatible.

### 1.2

2. *a*) L.D. $(2, 4)=2(1, 2)$. *b*) L.I. *c*) L.D. $(1, 2, 3)=(1, 3, 2)+(0, -1, 1)$.
   *d*) L.D. $(2, 2, 4, 2)=(2, 1, 3, 1)+(0, 1, 1, 1)$.
3. *a*) 2. *b*) 1. *c*) 2. *d*) 3. *e*) 3.
4. *a*) 3. *b*) 4.
5. *a*) Compatible, $x_1=x_2=x_3=1$. *b*) Compatible indeterminado, $x_1=5-d_1-d_2$, $x_2=2-d_1-d_2,\ x_3=d_1,\ x_4=d_2$.

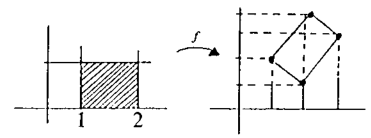
### 1.3

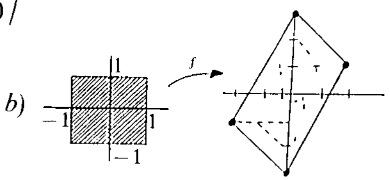
1. *a*) Nada. *b*) Inyectiva. *c*) Biyectiva.
2. Hay seis.
4. *a*) sen $[3x^2+16]$. *b*) $3[\operatorname{sen}(3x-5)]-5$. *c*) $3[\operatorname{sen}(x^2+7)]-5$.

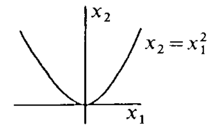




5. *a*) $\begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$. *b*) $\begin{pmatrix} 1 & -1 \\ 1 & 2 \\ 2 & -3 \end{pmatrix}$. *c*) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}$. *d*) $(3, -1, 1)$.

6. *a*) 

*b*) 


7. 


8. $\begin{pmatrix} \cos\varphi & -\operatorname{sen}\varphi \\ \operatorname{sen}\varphi & \cos\varphi \end{pmatrix}$. 9. $\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

10. *c*) $\begin{pmatrix} 172 \\ 21 \end{pmatrix}$. 11. *a*) $(4x_1,\ -3x_1+x_2,\ 5x_1+3x_2)$.
    *b*) $(7x_1-3x_2,\ -6x_1+2x_2,\ -11x_1-9x_2)$.

14. *a*) $\begin{pmatrix} 1 & 1 & 1 \\ 1 & -3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 7 \\ 3 \end{pmatrix}$. *b*) $\begin{pmatrix} 1 & 0 & 2 \\ 1 & 3 & 2 \\ 0 & -7 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix}$.

    *c*) $\begin{pmatrix} a & -2 & 3 \\ 1 & -a & 2 \\ a & 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} a^2 \\ 3 \\ a \end{pmatrix}$.

### 1.4

2. $g\circ f^{-1}(x)=\cos\left(\dfrac{x+2}{3}\right)$ ; $f^{-1}\circ f^{-1}\circ g(x)=\dfrac{1}{9}(\cos x)+\dfrac{2}{3}$.
3. *a*) 2. *b*) 2. *c*) 2. *d*) 1. *e*) 4.
4. *a*) $\begin{pmatrix} 1/3 & -2/3 \\ 1/3 & 1/3 \end{pmatrix}$. *b*) No. *c*) $\begin{pmatrix} 1 & 0 & 0 \\ -11/5 & 4/5 & 1/5 \\ -1/5 & -1/5 & 1/5 \end{pmatrix}$.
5. *a*) $\begin{pmatrix} 1 & 0 \\ -a & 1 \end{pmatrix}$. *b*) $\begin{pmatrix} 1/2 & 1/2a \\ a/2 & -1/2 \end{pmatrix}$ *c*) $\begin{pmatrix} 1 & 0 & -a \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. *d*) $\begin{pmatrix} 0 & 0 & 1/a \\ 0 & 1/a & 0 \\ 1/a & 0 & 0 \end{pmatrix}$.
6. *a*) $f^{-1}(x_1, x_2, x_3)=((x_1+x_2-x_3)/2,\ (3x_1-x_2-x_3)/2,\ -x_1+x_2)$.
   *b*) $f^{-1}(x_1, x_2)=((x_1+x_2)/2,\ (x_1-x_2)/2c)$.
7. *a*) $A^{-1}=\begin{pmatrix} 1 & -a \\ 0 & 1 \end{pmatrix}$, $\bar{x}=\begin{pmatrix} 1-2a \\ 2 \end{pmatrix}$.
   *b*) $A^{-1}=\begin{pmatrix} 1/4 & 1/4 & -1/4 \\ -37/4 & -13/4 & 17/4 \\ 11 & 4 & -5 \end{pmatrix}$, $\bar{x}=\begin{pmatrix} 0 \\ 40 \\ -48 \end{pmatrix}$.

$$c)\quad A^{-1}=\begin{pmatrix}1/2 & -1/6 & 7/3\\ 0 & 1/3 & 1/3\\ 0 & 0 & -1\end{pmatrix},\quad \vec{x}=\begin{pmatrix}5/36\\ -2/18\\ 0\end{pmatrix}.$$

## CAPÍTULO 2

### 2.1

1. *a)* 3. *b)* 12. *c)* $a(b-1)$. *d)* 1. *e)* $-1$.

2. $$\begin{pmatrix}\dfrac{b}{ab-a} & -\dfrac{1}{ab-a}\\ \dfrac{-1}{b-1} & \dfrac{1}{b-1}\end{pmatrix}\ \text{y}\ \begin{pmatrix}\cos\alpha & \operatorname{sen}\alpha\\ -\operatorname{sen}\alpha & \cos\alpha\end{pmatrix}$$

3. *a)* $x_1=-1,\ x_2=1$. *b)* $x_1=\pi/2,\ x_2=0$. *c)* $x_1=x_2=0$.

4. *a)* 11. *b)* 0. *c)* $(x-y)(x-z)(z-y)$. *d)* 1.

7. $\cos\alpha$.

### 2.2

1. *a)* 8. *b)* 2. *c)* $-56$.
2. *a)* $-16$. *b)* 0. *c)* $-14$. *d)* $-90$.
3. *a)* 0. *b)* 4. *c)* 1. *d)* $-60$. *e)* $-162$.
   *f)* $-\lambda^3+5\lambda^2-3\lambda-1$. *g)* $(x-1)(y-1)$.
4. *a)* $F_1=-F_2$. *b)* $C_3=2C_1+3C_2$. *c)* $F_2=F_1+F_3$.
   *d)* $F_1=2F_2-F_3$. *e)* $F_4-F_1=3(F_3-F_2)$.

8. $$a)\ \begin{pmatrix}0 & a_{12} & a_{13} & a_{14}\\ -a_{12} & 0 & a_{23} & a_{24}\\ -a_{13} & -a_{23} & 0 & a_{34}\\ -a_{14} & -a_{24} & -a_{34} & 0\end{pmatrix}.\quad b)\ |A|=|A^t|=(-1)^n|A|.\quad c)\ \text{No.}$$

13. $$\begin{pmatrix}1 & n\\ 0 & 1\end{pmatrix}.$$

### 2.3

1. $$a)\ \begin{pmatrix}1 & 0\\ -2 & 1\end{pmatrix}A=\begin{pmatrix}1 & 0 & 2\\ 0 & 1 & -3\end{pmatrix}.$$

$$b)\ \begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 5/7\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1/8\end{pmatrix}.$$



$$\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & -1/7 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}A=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

$$c)\ \begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1/3 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}A=\begin{pmatrix}1 & 0\\ 0 & 1\\ 0 & 0\end{pmatrix}.$$

$$d)\ \begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -1\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -1/2\end{pmatrix}.$$

$$\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix}A=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

$$e)\ \begin{pmatrix}1 & -2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & -1/3 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -2 & 0 & 1\end{pmatrix}.$$

$$\cdot\begin{pmatrix}1 & 0 & 0\\ -3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}A=\begin{pmatrix}1 & 0 & -1/3\\ 0 & 1 & 2/3\\ 0 & 0 & 0\end{pmatrix}.$$

2. $$b)\ A=\begin{pmatrix}1 & 0 & 0\\ 3 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ -1 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & -7 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 8\end{pmatrix}.$$

$$\cdot\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & -5/7\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

3. $AA^{-1}=I\Rightarrow 1=|AA^{-1}|=|A||A^{-1}|$.

5. $AA^t=(a^2+b^2+c^2+d^2)I\Rightarrow |A|=(a^2+b^2+c^2+d^2)^2$.

6. $$A=\begin{pmatrix}1 & x_1 & 0\\ 1 & x_2 & 0\\ 1 & x_3 & 0\end{pmatrix}\begin{pmatrix}1 & 1 & 1\\ y_1 & y_2 & y_3\\ 0 & 0 & 0\end{pmatrix}\Rightarrow |A|=0.$$

7. *a)* $2(-x)^{n-1}$. *b)* $(x-1)(x-2)\cdots(x-n+1)$. *c)* $(n-1)^{n-1}(2n-1)$.

9. $D_n=D_{n-1}-D_{n-2}$.

11. $(x+2)(x-2)^3$.

12. $(-1)^{n-1}\,2^{n-2}(n-1)$.

13. $-2(x-1)^3$.

**2.4**

1. *a*) $\operatorname{cof}(A)=\begin{pmatrix} 5 & -11 & -1 \\ 3 & -5 & -1 \\ -3 & 7 & 1 \end{pmatrix}$ ; $A^{-1}=\frac{1}{2}[\operatorname{cof}(A)]^t$.

   *b*) $\operatorname{cof}(A)=\begin{pmatrix} 2 & -3 \\ -1 & 2 \end{pmatrix}$ ; $A^{-1}=[\operatorname{cof}(A)]^t=\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$.

   *c*) $\operatorname{cof}(A)=\begin{pmatrix} 15 & -1 & -6 \\ -3 & 5 & 6 \\ -3 & -3 & 6 \end{pmatrix}$ ; $A^{-1}=\frac{1}{24}[\operatorname{cof}(A)]^t$

   *d*) $\operatorname{cof}(A)=\begin{pmatrix} -1 & 0 & 0 & -2 \\ -3 & 0 & 0 & 1 \\ 0 & -5 & 0 & 19 \\ -15 & 20 & -5 & -19 \end{pmatrix}$ ; $A^{-1}=-\frac{1}{7}[\operatorname{cof}(A)]^t$.

4. $A^{-1}=\frac{1}{|A|}C^t \Rightarrow |C|=|C^t|=|A|^n|A^{-1}|=|A|^{n-1}$.

5. *a*) y *b*) son inversas entre sí. *c*) $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{a}{a^2+b^2} & \frac{b}{a^2+b^2} \\ 0 & 0 & \frac{-b}{a^2+b^2} & \frac{a}{a^2+b^2} \end{pmatrix}$

6. *a*) $\begin{pmatrix} 1 & -2 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & -2 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}$

   *b*) $\begin{pmatrix} \frac{x+(n-1)}{x(n+x)} & -\frac{1}{x(n+x)} & \cdots & \frac{1}{x(n+x)} \\ \frac{1}{x(n+x)} & \frac{x+(n-1)}{x(n+x)} & \cdots & -\frac{1}{x(n+x)} \\ \vdots & \vdots & & \vdots \\ -\frac{1}{x(n+x)} & -\frac{1}{x(n+x)} & \cdots & \frac{x+(n+1)}{x(n+x)} \end{pmatrix}$

   *c*) $a \neq 0,\ b \neq 0,\ a \neq 1,\ b \neq 1,\ a \neq b$.

7. *a*) $a \neq -1,\ \frac{-1 \pm \sqrt{17}}{2}$. *b*) $a \neq 1 \pm \sqrt{3}$.



8. $T^{-1}(x, y, z)=(-2x-2y+3z,\ -y+z,\ 3x+4y-5z)$.

10. *a*) $x_1=5/3,\ x_2=8/3,\ x_3=0$.

    *b*) $x_1=-1/5,\ x_2=14/5,\ x_3=6/5$.

    *c*) $x_1=\frac{332}{145},\ x_2=\frac{-239}{145},\ x_3=-\frac{55}{58},\ x_4=-\frac{1}{58}$.

11. $m=-5$.

12. *a*) $\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$. *b*) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \operatorname{sen}\alpha \\ 0 & -\operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$. *c*) $\begin{pmatrix} \cos\alpha & \operatorname{sen}\alpha \\ \operatorname{sen}\alpha & -\cos\alpha \end{pmatrix}$.

13. $f(x)=2x^2-x+1$.

16. $|A| = a(a-1)^{n-1}$ ; $A$ es invertible si $a \neq 0$ y $a \neq 1$.

**2.5**

1. *a*) $r=2$. *b*) $r=3$. *c*) $r=2$. *d*) $r=2$.

2. $r=3$ para toda $a \neq 3$; $r=2$ si $a=3$.

4. *b*) $x_1=8/9,\ x_2=-5/9,\ x_3=x_4=0$.

   *c*) $x_1=3\lambda,\ x_2=\lambda,\ x_3=1-4\lambda$.

5. *a*) Incompatible. *b*) $x_1=11/4,\ x_2=3/2,\ x_3=-11/2$.

   *c*) $x=\frac{4+5t}{3},\ y=\frac{-2-10t}{3},\ z=\frac{4-10t}{9},\ t=t$ (indeterminado).

   *d*) $x=\frac{23+66t}{62},\ y=\frac{27+2t}{62},\ z=\frac{7-19t}{31},\ t=t$ (indeterminado).

6. Todas las soluciones son $(33\lambda+17\mu,\ 7\lambda+4\mu,\ \lambda,\ 0,\ \mu)$.

## CAPITULO 3

**3.1**

1. *a*) $\begin{Bmatrix} x=1+2t \\ y=-2t \end{Bmatrix}$, $x+y-1=0$. *b*) $\begin{Bmatrix} x=1+4t \\ y=-3t \end{Bmatrix}$, $3x+4y-3=0$.

   *c*) $\begin{Bmatrix} x=-5t-1 \\ y=3t+2 \end{Bmatrix}$, $5y+3x-7=0$.

3. *a*) Coinciden. *b*) Se cortan en $(1, 0)$. *c*) Coinciden.
   *d*) Se cortan en $(3/7, -6/7)$.

4. *a*) $\begin{Bmatrix} x=1+2t \\ y=-2+t \end{Bmatrix}$. *b*) $\begin{Bmatrix} x=2+t \\ y=1+2t \end{Bmatrix}$. *c*) $\begin{Bmatrix} x=3 \\ y=t \end{Bmatrix}$. *d*) $\begin{Bmatrix} x=t \\ y=0 \end{Bmatrix}$.

5. $y+x+t=0,\ t \in \mathbb{R}$.

6. *a*) $\begin{Bmatrix} x=1+2t \\ y=t \end{Bmatrix}$, $x-2y-1=0$. *b*) $\begin{Bmatrix} x=2+t \\ y=2t \end{Bmatrix}$, $2x-y-4=0$.

   *c*) $\begin{Bmatrix} x=-2+t \\ y=-3-t \end{Bmatrix}$, $x+y-5=0$.

9. *a)* $x+y-2=0$. *b)* $t(4x-y)+s(x+y-5)=0,\ t,\ s\in\mathbb{R}$.
*c)* $tx+sy=0,\ t,\ s\in\mathbb{R}$.

**3.2**

1. *a)* $\begin{cases} x=1-2t \\ y=1+3t \\ z=-1+t \end{cases}$. *b)* $\begin{cases} x=1+t \\ y=1 \\ z=-1-3t \end{cases}$. *c)* $\begin{cases} x=t \\ y=1-t \\ z=2-t \end{cases}$.

2. *a)* Se cruzan. *b)* Se cortan en (1, 4, 3). *c)* Se cruzan.

3. $\begin{cases} x=1+t/2 \\ y=1+5t/2 \\ z=1+t \end{cases}$.

4. $\left.\begin{array}{r} 3x+6y+z-11=0 \\ -4x+6y-3z-15=0 \end{array}\right\}$.

5. *a)* $\left\{\begin{array}{l} x=t-s \\ y=1+t-3s \\ z=-1+4t-2s \end{array}\right\}$, $5x-y+z=0$.

*b)* $\left\{\begin{array}{l} x=1-t+5s \\ y=2-t \\ z=3-4t+3s \end{array}\right\}$, $3x+17y-5z-22=0$.

*c)* $\left\{\begin{array}{l} x=10t+5s \\ y=1+3t+3s \\ z=-1+3t \end{array}\right\}$, $3x-5y-5z=0$.

6. *a)* $\begin{cases} x=2/3+t/2 \\ y=-1/3+3t/2 \\ z=t \end{cases}$. *b)* $\begin{cases} x=t \\ y=1-3t \\ z=-2+7t \end{cases}$. *c)* $\begin{cases} x=0 \\ y=t \\ z=-1/3+2t \end{cases}$

9. *a)* $\left\{\begin{array}{l} 3x-z+1=0 \\ 2x+y-2z=0 \end{array}\right\}$. *b)* $\begin{cases} x=t+s \\ y=2+2t+s \\ z=1+t-s \end{cases}$.

*c)* $\left.\begin{array}{l} x=1+t\alpha \\ y=-2t\alpha \\ z=1-7td \end{array}\right\}$, $\alpha\in\mathbb{R}$ ; $\left.\begin{array}{l} x=1+t \\ y=-2t \\ z=1-7t \end{array}\right\}$.

*d)* $x+y-3=0$. *e)* $\begin{cases} x=2-t \\ y=1-t \\ z=1 \end{cases}$.

**3.3**

1. *a)* $\sqrt{5}$. *b)* $-\dfrac{1}{\sqrt{10}}$. *c)* $\dfrac{1}{5\sqrt{2}}$. *d)* $x-3y-1=0$. *e)* $-2x-y+5=0$.



*f)* $\left(-\frac{1}{2}, \frac{3}{2}\right)+\left(-\frac{3}{2}, -\frac{1}{2}\right)$. *g)* $\left(-\frac{3}{5}, \frac{1}{5}\right)$. *h)* $\dfrac{6}{\sqrt{10}}, \dfrac{1}{\sqrt{10}}$.

2. *a)* 1, $\sqrt{19}$. *b)* 0. *c)* $z+2=0$. *d)* (1, 0, 0).
*e)* $\left(\frac{5}{3}, \frac{10}{3}, -\frac{5}{3}\right)+\left(-\frac{2}{3}, -\frac{1}{3}, -\frac{4}{3}\right)$, *f)* $\left\{\begin{array}{l} x+3y-11=0 \\ y+z-1=0 \end{array}\right.$
*g)* $-x+4y+7z-6=0$. *h)* $\dfrac{2}{\sqrt{6}}$, 0. *i)* (1, 0, 0).

3. *b)* $2x+z-1=0$. *c)* $\left\{\begin{array}{l} x=2t \\ y=0 \\ z=-2+t \end{array}\right\}$ ó $\left\{\begin{array}{l} y=0 \\ z=-2+x/2 \end{array}\right\}$.

*d)* $\left\{\begin{array}{l} x=t\left(\dfrac{2}{\sqrt{5}}-\dfrac{1}{\sqrt{11}}\right) \\ y=t/\sqrt{11} \\ z=2+t\left(-\dfrac{1}{\sqrt{5}}-\dfrac{3}{\sqrt{11}}\right) \end{array}\right\}$ y $\left\{\begin{array}{l} x=t\left(\dfrac{2}{\sqrt{5}}+\dfrac{1}{\sqrt{11}}\right) \\ y=t\left(-\dfrac{1}{\sqrt{11}}\right) \\ z=2+t\left(\dfrac{1}{\sqrt{5}}+\dfrac{3}{\sqrt{11}}\right) \end{array}\right\}$.

*e)* $\left(\frac{1}{5}, 1, -\frac{2}{5}\right)$. *f)* $\dfrac{\sqrt{104}}{5}$. *g)* $\left(-\frac{2}{5}, 0, -\frac{1}{5}\right)+\left(\frac{2}{5}, 3, -\frac{4}{5}\right)$.
*h)* $x-5y-2z-4=0$. *i)* $x-5y-2z+4=0$. *j)* $8/\sqrt{30}$.

6. Punto de corte (1, 1, 1) $\left\{\begin{array}{l} x=1+t\left(-\dfrac{1}{\sqrt{14}}-\dfrac{3}{\sqrt{11}}\right) \\ y=1+t\left(\dfrac{2}{\sqrt{14}}-\dfrac{1}{\sqrt{11}}\right) \\ z=1+t\left(\dfrac{3}{\sqrt{14}}+\dfrac{1}{\sqrt{11}}\right) \end{array}\right\}$ y

$\left\{\begin{array}{l} x=1+t\left(-\dfrac{1}{\sqrt{14}}+\dfrac{3}{\sqrt{11}}\right) \\ y=1+t\left(\dfrac{2}{\sqrt{14}}+\dfrac{1}{\sqrt{11}}\right) \\ z=1+t\left(\dfrac{3}{\sqrt{14}}-\dfrac{1}{\sqrt{11}}\right) \end{array}\right\}$.

7. $x-2y+z-6=0$. 8. $3x+2y+6z-6=0$.
9. $2x-3y-6z\pm35=0$. 10. $2x+4y-3z+18=0$.

**3.4**

1. $\left(\frac{2}{3}, 1, -\frac{1}{3}\right), \left(\frac{7}{3}, 2, \frac{7}{3}\right)$.

3. *a)* $\left(\frac{3+\sqrt{11}}{2}, \frac{1+\sqrt{11}}{2}\right)$ y $\left(\frac{3-\sqrt{11}}{2}, \frac{1-\sqrt{11}}{2}\right)$. *b)* $(x-2)^2+y^2=6$.
*c)* $\left\{\begin{matrix} 2x-y>0 \\ x+2y-5>0 \end{matrix}\right\}$.

10. *a)* $\left(-\frac{1}{3}, \frac{4}{3}\right)$. *b)* $\left(-\frac{1}{2}, \frac{1}{2}\right)$, $\left(x+\frac{1}{2}\right)^2+\left(y-\frac{1}{2}\right)^2=\frac{\sqrt{26}}{2}$.
*d)* $(0, 3)$.

11. $\left\{\begin{matrix} x=1-t \\ y=1+t \\ z=1-4t \end{matrix}\right\}$.

**3.5**

1. 8. 2. *a)* $x+3y+z-8=0$. *b)* $\left.\begin{matrix} 25x-3y-144=0 \\ 8x-3z-36=0 \end{matrix}\right\}$.
*c)* $-x+2y-5=0$.

## CAPITULO 4

1. *a)* $2.016+92i$. *b)* $-1$. *c)* 25.
2. *a)* $-i$. *b)* $\frac{22-4i}{10}$. *c)* $(7-i)^2/100$.
3. $a^2-b^2$, $(a^2-b^2)/(a^2+b^2)^2$, $(a^2-1+b^2)/(a^2+b^2-2a+1)$.

**4.2**

1. Solamente escribimos módulo y argumento.
*a)* $\sqrt{2}$, $\pi/4$. *b)* 1, $-\pi/3$. *c)* 1, $7\pi/6$. *d)* $\sqrt{8}$, $5\pi/4$.
2. $-1, 1, \frac{1}{\sqrt{2}}+i\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}-i\frac{1}{\sqrt{2}}$.
4. *a)* 2, 0. *b)* 1, 0. *c)* $\sqrt{18}$, $3\pi-\frac{\pi}{4}$.

**4.3**

1. *a)* $\sqrt{2}e^{\left(\pi-\frac{\pi}{8}\right)i}$, $\sqrt{2}e^{\left(2\pi-\frac{\pi}{8}i\right)}$. *b)* $2e^{\frac{\pi}{3}i}$, $-2$, $2e^{\frac{5\pi}{3}i}$.
*c)* $i$, $e^{\frac{7\pi}{6}i}$, $e^{\frac{11\pi}{6}i}$.
*d)* $2e^{\frac{\pi}{8}i}$, $2e^{\left(\frac{\pi}{8}+\frac{\pi}{2}\right)i}$, $2e^{\left(\frac{\pi}{8}+\pi\right)i}$, $2e^{\left(\frac{\pi}{8}+\frac{3\pi}{2}i\right)}$.
*e)* $3i$, $-3i$.
2. 1, $e^{2\pi i/6}$, $e^{4\pi i/6}$, $e^{8\pi i/6}$, $e^{10\pi i/6}$.
3. $\omega_0=1$, $\omega_1=e^{2\pi i/3}$, $\omega_2=e^{\frac{4\pi}{3}i}$.



**4.4**

1. *a)* 1, 2, 3, 4. *b)* $-1, 1, -\frac{1}{2}+\frac{\sqrt{3}}{2}i, -\frac{1}{2}-\frac{\sqrt{3}}{2}i$. *c)* $-2, \frac{3\pm\sqrt{3}i}{2}$.
*d)* $\frac{3}{2}, \frac{1}{2}, -1$. *e)* $3, -3, 4i, -4i$.
2. Las soluciones son $e^{2\pi i/5}$, $e^{4\pi i/5}$, $e^{6\pi i/5}$, $e^{8\pi i/5}$.
4. No lo contradice.

## CAPITULO 5

**5.1, 5.2 y 5.3**

1. *a)* Sí. *b)* Sí.
2. *a)* Linealmente independientes. *b)* Linealmente dependientes.
*c)* Linealmente dependientes. *d)* Linealmente independientes.
*e)* Linealmente independientes.
3. $z=0, \pm\sqrt{2}$.
5. *a)* Sí. *b)* Sí. *c)* No.
6. $\{(0, 0, 1, 1), (1, 1, 0, 0), (1, 0, 0, 0), (0, 0, 0, 1)\}$.
7. No.
9. $\dim T=n^2-1$.
10. *b)* $\vec{v}=4\vec{u}_1+2\vec{u}_2-\vec{u}_3-6\vec{u}_4$.
11. *b)* $\vec{x}=-(\vec{u}_1+2\vec{u}_2+3\vec{u}_3+\cdots+(n-1)\vec{u}_{n-1}+n\vec{u}_n)$.
12. $S_1=\{x+1, x-1, x^2-1\}$.
15. $a\neq 1$ y $a\neq -2$

**5.4**

1. *a)* No. *b)* Sí, no. *c)* Sí, no.
2. *a)* No. *b)* Sí.
3. *a)* 2. *b)* 1.

4. *a*) $\operatorname{Com}\begin{pmatrix}2 & 1\\ 0 & 2\end{pmatrix}+\operatorname{Com}\begin{pmatrix}2 & 0\\ 1 & 2\end{pmatrix}=L\left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}, \begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}\right\}$

$\operatorname{Com}\begin{pmatrix}2 & 1\\ 0 & 2\end{pmatrix}\cap\operatorname{Com}\begin{pmatrix}2 & 0\\ 1 & 2\end{pmatrix}=L\left\{\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\right\}.$

*b*) $V_1+V_2=P_{\mathbb{C}}^{(3)}[x]$ y $V_1\cap V_2=L\{x^2-1,\ x(x^2-1)\}$.

*c*) $L(\operatorname{sen} t, \cos t)+L(e^{it}, e^{-it})=L(\operatorname{sen} t, e^{it}, e^{-it})$,
$L(\operatorname{sen} t, \cos t)\cap L(e^{it}, e^{-it})=L(\cos t)$.

5. *a*) Sí, $\mathscr{S}\oplus\mathscr{A}=\mathscr{M}_{2\times 2}(\mathbb{C})$.
*b*) Sí, $P\oplus I=\{f\colon [-1, 1]\mapsto\mathbb{R}\}$.
*c*) Sí, $V_1\oplus V_2=\{f\colon [-1, 1]\mapsto\mathbb{R}/f(0)=0\}$.

## 5.5

1. *a*) $(-4/3, 5/3)$. *b*) $x_1'+2x_2'+1=0$.
2. $(x_1'-\cos\alpha)^2+(x_2'+\operatorname{sen}\alpha)^2=4$.
3. $P=(3, 0, 0)$, $\vec{u}_1=(1, 1, 0)$, $\vec{u}_2=(1, -1, 1)$, $\vec{u}_3=(0, 1, 1)$.
4. $\dfrac{(x_1'+2\sqrt{2})^2}{9}+\dfrac{x_2'^2}{9/4}=1.$

# CAPITULO 6

## 6.1, 6.2 y 6.3

1. *a*) Lineal. *b*) No lineal. *c*) Lineal. *d*) Lineal.
*e*) No lineal. *f*) Lineal. *g*) No lineal.

2. $\begin{pmatrix}1 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 1\end{pmatrix}$. *a*) $L\{(1, -1, 0), (1, -1, -1)\}$, $\dim(V_1)=2$, $\dim(A(V_1))=2$.

*b*) $L\{(1, 0, 1), (1, 0, 0)\}$, $\dim(V_2)=2$, $\dim(A(V_2))=2$.
*c*) $L\{(0, 1, 2)\}$, $\dim(V_3)=1$, $\dim(A(V_3))=1$.

3. *a*) $M_B=\begin{pmatrix}1 & -1 & 0 & 0\\ 0 & 0 & 1 & -1\end{pmatrix}$, $C_B=\begin{pmatrix}0 & 0 & -1 & 0\\ 1 & 0 & 0 & -1\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\end{pmatrix}$. *b*) $\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 1\\ 1 & 2 & 4\\ 1 & 3 & 9\end{pmatrix}$.

*c*) $\begin{pmatrix}1 & 1 & 1 & 1\\ 0 & 1 & 2 & 3\\ 0 & 0 & 1 & 3\\ 0 & 0 & 0 & 1\end{pmatrix}$.



4. *a*) $\begin{pmatrix}\cos\alpha & -\operatorname{sen}\alpha & 0\\ \operatorname{sen}\alpha & \cos\alpha & 0\\ 0 & 0 & 1\end{pmatrix}$. *b*) $\begin{pmatrix}-1 & 0 & 0\\ 0 & -1 & 0\\ 0 & 0 & 1\end{pmatrix}$. *c*) $\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & -1\end{pmatrix}$.

*d*) $\begin{pmatrix}2/3 & 1/3 & -1/3\\ 1/3 & 2/3 & 1/3\\ -1/3 & 1/3 & 2/3\end{pmatrix}$. *e*) $\begin{pmatrix}-1/3 & 2/3 & 2/3\\ 2/3 & -1/3 & 2/3\\ 2/3 & 2/3 & -1/3\end{pmatrix}$.

6. $A\circ A(x_1, x_2, x_3)=(x_3, 0, 0)$ y $A^n=A\circ A\circ\cdots\circ A=0$ si $n\geqslant 3$
$B\circ A(x_1, x_2, x_3)=(x_1, x_3)$.

7. *a*) $\begin{pmatrix}2 & 1 & 3\\ 1 & 0 & 1\\ 2 & 0 & 1\end{pmatrix}$. *b*) $\begin{pmatrix}1 & 2 & 4\\ -1 & -1 & -1\\ 2 & 2 & 3\end{pmatrix}$.

8. *a*) $\left.\begin{aligned} x'&=-\frac{1}{3}x+\frac{2}{3}y+\frac{2}{3}z\\ y'&=\frac{2}{3}x-\frac{1}{3}y+\frac{2}{3}z\\ z'&=\frac{2}{3}x+\frac{2}{3}y-\frac{1}{3}z\end{aligned}\right\}.$ *b*) $\left.\begin{aligned} x'&=\frac{2}{3}x+\frac{1}{3}y-\frac{1}{3}z\\ y'&=\frac{1}{3}x+\frac{2}{3}y+\frac{1}{3}z\\ z'&=-\frac{1}{3}x+\frac{1}{3}y+\frac{2}{3}z\end{aligned}\right\}.$

*c*) $\left.\begin{aligned} x'&=\frac{1}{2}x-\frac{1}{2}y+\frac{1}{2}z\\ y'&=-\frac{1}{2}x+\frac{1}{2}y+\frac{1}{2}z\\ z'&=-x-y\end{aligned}\right\}.$

9. $T(x^2+2x+1)=1-2x+x^2+x^3$
$T((x-2)^2+x^3)=3+5x+5x^2+x^3$ $\begin{pmatrix}1 & 0 & 0 & -1\\ 0 & -1 & 0 & 1\\ 1 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{pmatrix}$

10. *a*) $B=\{(2, 1, 0), (3, 0, -1)\}$. $\begin{pmatrix}3 & 3\\ 1 & -1\end{pmatrix}$ es la matriz de $A|_{V_1}$ con respecto a $B$ y la base canónica de $\mathbb{R}^2$.
*b*) *Observación*: $\mathbb{C}^2$ y $\mathbb{C}^3$ son espacios vectoriales sobre $\mathbb{C}$; la base canónica de $\mathbb{C}^2$ es $\{(1, 0), (0, 1)\}$ y la base canónica de $\mathbb{C}^3$ es $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$.
$B=\{(i, 1)\}$ ; $\begin{pmatrix}1+i\\ i\\ -1\end{pmatrix}$

11. *a*) $\{(0, 1, 0)\}$. *b*) $\lambda=2$, $\{(0, 0, 3, 1)\}$; $\lambda=1$, $\{(1, 0, 0, 0), (0, 1, 0, 0)\}$; $\lambda=0$, $\{(0, 0, 0, 0)\}$.
*c*) $E(e^{\frac{\pi}{3}i})=\{(0, 0, 0)\}$.

12. 1.

### 6.4

1. *a*) $\ker(A)=\{t(1,0,-1)/t\in\mathbb{R}\}$, $\operatorname{img}(A)=\{t(1,0,1)+s(0,1,0)/t,s\in\mathbb{R}\}$. No inyectiva, no suprayectiva.
   *b*) $\mathbb{C}^2$ se considera como espacio vectorial sobre $\mathbb{C}$. $\ker(B)=\{(0,0)\}$, $\operatorname{img}(B)=\mathbb{C}^2$; inyectiva y suprayectiva. *c*) $\ker(C)=\{(0,0)\}$, $\operatorname{img}(C)=\{t(1,1,2,1)+s(0,1,0,2)/t,s\in\mathbb{R}\}$. Inyectiva, no suprayectiva.
   *d*) $\ker(D)=\{\alpha(2i,-2,1)/\alpha\in\mathbb{C}\}$, $\operatorname{img}(D)=\{\alpha(1,0,i)+\beta(0,1,0)/\alpha,\beta\in\mathbb{C}\}$; no inyectiva, no suprayectiva.
2. *a*) $\ker M_B=\mathscr{L}\left\{\begin{pmatrix}1&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&1\end{pmatrix}\right\}$, $\operatorname{Img} M_B=\mathscr{L}\left\{\begin{pmatrix}1\\0\end{pmatrix},\begin{pmatrix}0\\1\end{pmatrix}\right\}$; no inyectiva, suprayectiva.
   *b*) $\ker C_B=\mathscr{L}\left\{\begin{pmatrix}1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}\right\}$, $\operatorname{img} C_B=\mathscr{L}\left\{\begin{pmatrix}-1&0\\0&1\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix}\right\}$; no inyectiva, no suprayectiva.
   *c*) $\ker(A)=\{0\}$, $\operatorname{img}(A)=P_{\mathbb{R}}^{(3)}[x]$; inyectiva, suprayectiva y biyectiva.
   *d*) $\ker(D)=\mathscr{L}\{1\}$, $\operatorname{img}(D)=P_{\mathbb{R}}^{(n-1)}[x]$; no inyectiva, suprayectiva.
3. $\ker(T)=\left\{(a_{ij})_{i,j}\in\mathscr{M}_{n\times n}/\sum_{i=1}^{n}a_{ii}=0\right\}$; $\dim(\ker T)=n^2-1$;
   $\operatorname{img}(T)=\mathbb{K}$; $\dim(\operatorname{img}(T))=1$.
5. *a*) $f:\mathbb{C}^5\mapsto\mathbb{R}^{10}/f(a_1+ib_1,a_2+ib_2,a_3+ib_3,a_4+ib_4,a_5+ib_5)=$ $=(a_1,b_1,a_2,b_2,a_3,b_3,a_4,b_4,a_5,b_5)$.
   *b*) $f\colon P_{\mathbb{C}}^{(n)}[x]\mapsto\mathbb{C}^{n+1}/f\left(\sum_{i=0}^{n}a_ix_i\right)=(a_0,a_1,\dots,a_n)$.
   *c*) $f\colon\mathscr{M}_{3\times2}(\mathbb{R})\to\mathbb{R}^6/f\begin{pmatrix}a&b\\c&d\\e&f\end{pmatrix}=(a,b,c,d,e,f)$.

### 6.5

1. $E_1(x_1,x_2,x_3)=x_2-x_3$, $E_2(x_1,x_2,x_3)=x_3-x_1$, $E_3(x_1,x_2,x_3)=x_1-x_2+x_3$.
2. $E_1(a_0+a_1x+a_2x^2+a_3x^3)=a_0-a_1+2a_2+2a_3$
   $E_2(a_0+a_1x+a_2x^2+a_3x^3)=a_2$
   $E_3(a_0+a_1x+a_2x^2+a_3x^3)=a_2+a_3$
   $E_4(a_0+a_1x+a_2x^2+a_3x^3)=a_3$

## CAPITULO 7

### 7.2

1. *a*) 1, $-2$. *b*) 3. *c*) No hay. *d*) $-1$. *e*) $\pm1$. *f*) 0, $\sqrt{5}$, $-\sqrt{5}$. *g*) 1. *h*) $-1$, $-2$. *i*) 2.
2. *a*) $\begin{pmatrix}1&0\\0&-2\end{pmatrix}$. *d*) $\begin{pmatrix}-1&0\\0&-1\end{pmatrix}$. *f*) $\begin{pmatrix}0&0&0\\0&1&0\\0&0&3\end{pmatrix}$. *h*) $\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&-2\end{pmatrix}$.



4. *a*) $\lambda_1=\lambda_2=\lambda_3=-1$; $c(1,-1,-1)$ con $c\neq0$.
   *b*) $\lambda_1=1$, $\lambda_2=\lambda_3=0$; $c(1,1,1)$, $c(1,2,3)$, con $c\neq0$.
   *c*) $\lambda_1=\lambda_2=1$; $c(0,0,0,1)$, $c\neq0$; $\lambda_3=\lambda_4=0$;
   $c_1(0,1,0,0)+c_2(0,0,1,0)$, $c_1\cdot c_2\neq0$.
5. *a*) $\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}$ y $P=\begin{pmatrix}1&1&1\\1&1&0\\1&0&-3\end{pmatrix}$. *b*) $\begin{pmatrix}2&0&0\\0&3&0\\0&0&3\end{pmatrix}$ y $P=\begin{pmatrix}1&1&1\\1&1&0\\1&0&1\end{pmatrix}$.
   *c*) $\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&-1&0\\0&0&0&-1\end{pmatrix}$ y $P=\begin{pmatrix}1&0&0&-1\\0&1&-1&0\\0&1&1&0\\1&0&0&1\end{pmatrix}$.
6. $\lambda=0$ y polinomios de grado cero.
7. Si $b=0$, la matriz ya es diagonal; si $b\neq0$, la matriz es diagonalizable si y sólo si $a\neq-1$.
8. No es diagonalizable en ningún caso.
10. Autovalores $\lambda=b$, $\lambda=\pm\sqrt{a}$; $a=0$ no diagonalizable; $a>0$ diagonalizable; $a<0$: autovalores complejos y, por tanto, no diagonalizable en forma real.
11. $\lambda_1=5$, $\lambda_2=0$; $c_1(2,1)$, $c_2(-1,2)$, $c_1\neq0$, $c_2\neq0$.
12. $\lambda_1=1$, $\lambda_2=-1$; $(\beta,1-\alpha)$, $(\beta,-1-\alpha)$. Simetría: $S^2=I$.
13. $\lambda_1=1$, $\lambda_2=0$; $(\alpha,\beta)$, $(\beta,-\alpha)$. Proyección: $P^2=P$.

### 7.3

1. $J=\begin{pmatrix}\dfrac{1+\sqrt{5}}{2}&0\\0&\dfrac{1-\sqrt{5}}{2}\end{pmatrix}$; $A^5=\begin{pmatrix}8&5\\5&3\end{pmatrix}$.
2. $A^{10}=\begin{pmatrix}1&1\\-1&0\end{pmatrix}\begin{pmatrix}2^{10}&5\cdot2^{10}\\0&2^{10}\end{pmatrix}\begin{pmatrix}0&-1\\1&1\end{pmatrix}=2^{10}\begin{pmatrix}5&4\\-4&-3\end{pmatrix}$.
3. $P(A)=\begin{pmatrix}0&0\\0&0\end{pmatrix}$.

### 7.4

1. A: $J=\begin{pmatrix}1&0&0\\0&2&0\\0&0&2\end{pmatrix}$, $P=\begin{pmatrix}1&1&1\\1&1&0\\1&0&1\end{pmatrix}$; B: $J=\begin{pmatrix}0&0&0\\0&-1&0\\0&0&3\end{pmatrix}$, $P=\begin{pmatrix}1&1&1\\1&-3&1\\-3&4&0\end{pmatrix}$
   C: $J=\begin{pmatrix}1&0&0\\0&2&0\\0&0&3\end{pmatrix}$, $P=\begin{pmatrix}1&1&1\\1&0&1\\1&1&0\end{pmatrix}$; D: $J=\begin{pmatrix}3&0&0\\0&3&1\\0&0&3\end{pmatrix}$, $P=\begin{pmatrix}1&2&0\\0&1&1\\0&1&0\end{pmatrix}$

E: $J=\begin{pmatrix}2&1&0\\0&2&1\\0&0&2\end{pmatrix}$, $P=\begin{pmatrix}1&-2&1\\0&1&0\\-1&1&0\end{pmatrix}$; F: $J=\begin{pmatrix}-1&0&0\\0&-1&1\\0&0&-1\end{pmatrix}$,

$P=\begin{pmatrix}0&-1&1\\1&-1&0\\0&1&0\end{pmatrix}$; G: $J=\begin{pmatrix}2&0&0\\0&2&1\\0&0&2\end{pmatrix}$, $P=\begin{pmatrix}1&1&1\\1&4&0\\1&3&0\end{pmatrix}$;

H: $J=\begin{pmatrix}0&0&0\\0&1&0\\0&0&1\end{pmatrix}$, $P=\begin{pmatrix}1&1&0\\-1&1&1\\-1&0&1\end{pmatrix}$

2. $C^6=\begin{pmatrix}1&1&1\\1&0&1\\0&1&0\end{pmatrix}\begin{pmatrix}1&0&0\\0&2^6&0\\0&0&3^6\end{pmatrix}\begin{pmatrix}1&1&1\\1&0&1\\0&1&0\end{pmatrix}^{-1}$

3. $V_1=\mathscr{L}\{(2,2,-1)\}$; $V_2=\mathscr{L}\{(1,1,0)\}$; $V_3=\mathscr{L}\{(0,1,1)\}$
$V_4=\mathscr{L}\{(1,1,0),(0,1,1)\}$; $V_5=\mathscr{L}\{(2,2,-1),(0,1,1)\}$
$V_6=\mathscr{L}\{(2,2,-1),(1,1,0)\}$; $V_7=\{\overline{0}\}$; $V_8=\mathbb{R}^3$.

4. $\lambda_1=0$ doble; $(-\lambda,\alpha,0)$, $(-\mu,0,\alpha)$; $\lambda_2=1$; $(\mu,\nu,\gamma)$.

**7.6**

4. *c*) $\lambda=0,1$; $\lambda=\pm 1$; $\lambda=e^{2\pi ik/n}$, $k=0,1,\dots,n-1$; $\lambda=\pm i$.

**7.7**

1. $J=\begin{pmatrix}2&1&0&0\\0&2&0&0\\0&0&2&0\\0&0&0&1\end{pmatrix}$, $P=\begin{pmatrix}-1&2&1&3\\-1&1&0&6\\-1&0&0&7\\0&0&-1&1\end{pmatrix}$

2. $J=\begin{pmatrix}2&1&0&0\\0&2&0&0\\0&0&2&1\\0&0&0&2\end{pmatrix}$, $P=\begin{pmatrix}1&1&-1&0\\0&0&1&0\\1&0&0&0\\0&0&0&1\end{pmatrix}$.

3. $J=\begin{pmatrix}1+i&1&0&0\\0&1+i&0&0\\0&0&1-i&1\\0&0&0&1-i\end{pmatrix}$, $P=\begin{pmatrix}0&1+i&0&1-i\\3-i&1-5i&3+i&1+5i\\1-2i&0&1+2i&0\\0&1&0&1\end{pmatrix}$.

**7.8**

1. *a*) $J=\begin{pmatrix}1&0&0\\0&0&\sqrt{3}\\0&-\sqrt{3}&0\end{pmatrix}$, $P=\begin{pmatrix}0&1&0\\1&0&0\\1&0&-\sqrt{3}\end{pmatrix}$.



*b*) $J=\begin{pmatrix}1&\sqrt{6}\\-\sqrt{6}&1\end{pmatrix}$, $P=\begin{pmatrix}0&2\\\sqrt{6}&0\end{pmatrix}$.

*c*) $J=\begin{pmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&1\\0&0&-1&0\end{pmatrix}$, $P=\begin{pmatrix}0&1&-1&0\\0&0&1&0\\1&0&0&0\\0&0&0&-1\end{pmatrix}$.

**7.9.**

2. La matriz nula y todas aquellas equivalentes a la matriz $\begin{pmatrix}0&1\\0&0\end{pmatrix}$

3. Todas las de la forma $P\begin{pmatrix}1&0\\0&-1\end{pmatrix}P^{-1}$ con $|P|\neq 0$.

**Soluciones a algunos ejercicios de repaso de los capítulos 1 a 7**

1. $(x_1,x_2,x_3,x_4)=\left(-\frac{1}{3},\frac{4}{3},0,0\right)+c\left(\frac{2}{3},\frac{4}{3},1,0\right)$, $c\in\mathbb{R}$.

2. $r=2$. 3. $(x_1,x_2,x_3,x_4,x_5)=(7,-1,-2,-3,0)$.

4. Si $a+b-2ab\neq 0$, sistema compatible determinado. En este caso:

$$x=\frac{(1-a)(b-1)}{(a+b)-2ab},\quad y=\frac{1-a}{(a+b)-2ab},\quad z=\frac{1-b}{(a+b)-2ab}$$

Si $a+b-2ab=0$, $a\neq 1$ y $b\neq 1$ Incompatible.
Si $a+b-2ab=0$, $a=1$, $b=1$ Compatible indeterminado.
$x=0$, $y=1-t$, $z=t$, $t\in\mathbb{R}$
Si $a=0$ y $b=0$ Incompatible.

5. Si $x\neq a$ y $x\neq -a\pm 2b$, $r=4$. [*Sug.*: Utilizar transformaciones elementales.)
Si $x=a$, $r=2$; si $x=-a\pm 2b$, $r=2$.

6. Si $x\neq 0$ existe inversa y es

$$\frac{1}{x^4}\begin{pmatrix}x^3&-x^2&x&-1\\0&x^3&-x^2&x\\0&0&x^3&-x^2\\0&0&0&x^3\end{pmatrix}$$

[*Sug.*: Utilizar transformaciones elementales.)

7. $(x_1,x_2,x_3)=c(1,0,1)+d(0,1,0)$.

8. *a*) $a\neq 2$. *b*) $5x_1'+2x_2'+10x_3'=1$.

9. $a_1 a_2 \cdots a_n$.

12. *a)* $(x, y)=(3, 1)+t(a, 1)$, $t\in\mathbb{R}$, $a\in\mathbb{R}$. *b)* $3x-4y-5=0$.

13. *a)* $(2, -1, 8)$. *b)* $\dfrac{x-2}{1}=\dfrac{y+1}{7}=\dfrac{z-8}{2}$.

14. $\dfrac{x-1}{-3}=\dfrac{y-1}{5}=\dfrac{z-1}{-2}$.

15. $P_1=\left(\dfrac{2}{3}, \dfrac{1}{3}, \dfrac{-5}{3}\right)$, $P_2=\left(-\dfrac{7}{9}, \dfrac{17}{9}, \dfrac{-2}{9}\right)$. Área $=\dfrac{5}{54}\sqrt{2}$.

16. $1/\sqrt{138}$. 17. $\left(\dfrac{15}{25}, \dfrac{6}{25}, -\dfrac{8}{25}\right)+t(5, 3, -4)$.

18. $\left(\dfrac{13}{7}, -\dfrac{4}{7}, \dfrac{2}{7}\right), \left(-\dfrac{5}{7}, \dfrac{8}{7}, -\dfrac{4}{7}\right), \left(-\dfrac{1}{7}, \dfrac{10}{7}, \dfrac{9}{7}\right), \left(\dfrac{9}{7}, -\dfrac{6}{7}, -\dfrac{11}{7}\right)$

$$\mathrm{Vol}=\frac{1}{49}\begin{vmatrix} -2 & 2 & 1 \\ 13 & -11 & -12 \\ -5 & 1 & -18 \end{vmatrix}$$

19. $-4x+y+3z=-4$. 20. $(0, 2, -2)+t(0, 1, 1)$.

22. *a)* $x=1$, $x=2+i$, $x=2-i$.

23. *a)* $A=\begin{pmatrix} 1 & -1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 3 & -4 \\ 0 & 0 & 1 & -3 & 6 \\ 0 & 0 & 0 & 1 & -4 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$. *b)* $7+8(x-1)+3(x-1)^2+2(x-1)^3+(x-1)^4$.

24. Dimensión $=3$; $\{(1, 0, 0, -1), (2, 1, 1, 0), (1, 2, 3, 4)\}$ es una base; $x-y-z+t=0$.

25. Está generado por dos vectores: $\vec{u}_1$ y $\vec{u}_2$ (por ejemplo).

31. $J_A=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$ con $C=\begin{pmatrix} -1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$ ; $J_B=\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ con

$C=\begin{pmatrix} -1 & 1 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$

32. $J_A=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -1 \end{pmatrix}$ ; $A^3=\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 8 \end{pmatrix}$.

33. No. Sus autovalores son 0 y 2 con $\dim(\ker(A))=1$ y $\dim(\ker(A-2I))=1$.



34. $J_A=\begin{pmatrix} 1 & 1 & & & 0 \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ 0 & & & & 1 \end{pmatrix}=J_B$

35. *a)* $-\lambda^3+4\lambda^2-5\lambda+2$.

*b)* $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$. *c)* $-3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

36. *a)* Autovalores $\lambda_1=(n-1)$ doble, $\lambda_2=n+2$; $\dim(\ker(A_n-(n-1)I))=2$.

*b)* $n\neq 1$, $n\neq -2$ ; $A_n^{-1}=\dfrac{1}{(n-1)(n+2)}\begin{pmatrix} n+1 & -1 & -1 \\ -1 & n+1 & -1 \\ -1 & -1 & n+1 \end{pmatrix}$.

*c)* $n=1$, $\ker A_1=\{(x, y, z)/x+y+z=0\}$, $\mathrm{img}(A_1)=\{(u, \omega, t)/u=\omega=t\}$
$n=-2$, $\ker A_{-2}=\{(x, y, z)/x=y=z\}$, $\mathrm{img}(A_{-2})=\{(u, w, t)/u+w+t=0\}$.

37. $0, 1, 3^{10}$.

38. $a\neq 1$ y $a\neq 2$ Compatible determinado
$a=1$, $b\neq -1$ Incompatible
$a=1$, $b=-1$ Compatible indeterminado: $\begin{cases} x_1=10x_3-1 \\ x_2=-1+6x_3 \end{cases}$.
$a=2$, $b\neq -\dfrac{3}{5}$ Incompatible
$a=2$, $b=-\dfrac{3}{5}$ Compatible indeterminado: $\begin{cases} x_1=2/5 \\ x_2=-1/5+2x_3 \end{cases}$.

39. $\dfrac{x-1}{-1}=\dfrac{y-2}{-1}=\dfrac{z+1}{2}$.

40. *a)* $c_1=(2, 1, 3)$, $c_2=(-2, 1, -1)$. *b)* $D_1=(1, 2, 1)$, $D_2=(1, 0, 1)$.

41. $\{x^3+2x^2+x, -2x^3-4x^2+1\}$ es base del núcleo.
$\{x^2-2x, x+1\}$ es base de la imagen.

42. Sí. *b)* Recta $t(2, 5, -1)$ y todas las rectas contenidas en el plano $t(-2, 1, 0)+s(0, 0, 1)$.

43. 5, 3 y 2.

44. *a)* $\left(\dfrac{7}{3}, \dfrac{7}{6}, -\dfrac{1}{6}\right)$. *b)* $(1, 1, 0)$.

45. *a)* $D=(0, 1, 1)$. *b)* $E=\left(\dfrac{1}{2}, \dfrac{1}{2}, 4\right)$ y $E'=\left(\dfrac{1}{2}, \dfrac{1}{2}, -2\right)$.

46. *a)* No hay. *b)* $L(\{(2, 1, 1)\})$. *c)* $\mathbb{R}^3$.

47. $\begin{pmatrix} 2 & -11 & 6 \\ 1 & -7 & 4 \\ 2 & -1 & 0 \end{pmatrix}$.

48. $a=1$; $-\vec{u}_1-\vec{u}_2+\vec{u}_3=\vec{0}$.

49. $(0,\ -1,\ -1)+t(1,\ 2,\ 4)$ ; $\left.\begin{matrix} 2x-3y+z=2 \\ 2x+y-z=0 \end{matrix}\right\}$.

50. $3x-4y=10$ ó $3x-4y=-10$.

51. *a*) $\begin{pmatrix} 5-i & -4+i & i-1 \\ -1-5i & 1+4i & 1+i \end{pmatrix}$.

*b*) $\ker(A)=L\{(4-i,\ 5-i,\ 0),\ (1-i,\ 0,\ 5-i)\}$; $\operatorname{img}(A)=L\{(c,\ 1)\}$.

52. La matriz es diagonalizable para todo $a\neq -1$.

## CAPITULO 8

### 8.1 y 8.2

1. *a*) No. *b*) No. *c*) No. *d*) No. *e*) Sí.

2. $\|\vec{x}\|=\sqrt{2(x_1+x_2)^2+x_2^2}$; $\cos \measuredangle (\vec{x},\ \vec{y})=\dfrac{2x_1y_1+3x_2y_2+2x_1y_2+2x_2y_1}{\sqrt{2(x_1+x_2)^2+x_2^2}\ \sqrt{2(y_1+y_2)^2+y_2^2}}$.

3. $\pi/2$. 4. $\arccos(1/\sqrt{n})$.

### 8.3

1. La matriz del producto escalar es $\begin{pmatrix} 2 & 1 & 5 \\ 1 & 1 & 3 \\ 5 & 3 & 14 \end{pmatrix}$

2. $\vec{y}_1=(1,\ 0,\ 0)$, $\vec{y}_2=(0,\ -2,\ 0)$, $\vec{y}_3=(0,\ 0,\ 5)$.

4. $\{(1,\ 2,\ 1,\ 3),\ \left(\frac{10}{3},\ \frac{-1}{3},\ \frac{1}{3},\ -1\right),\ \left(\frac{-19}{185},\ \frac{87}{185},\ \frac{61}{185},\ \frac{-72}{185}\right)$.

5. *a*) $\{(1,\ 1,\ 1)\}$. *b*) $\{(2,\ -1,\ 0),\ (-4,\ 11,\ -6)\}$.

6. $\left\{\left(\frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}},\ 0,\ 0\right),\ \left(0,\ \frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}},\ 0\right),\ \left(\frac{1}{\sqrt{6}},\ 0,\ \frac{-1}{\sqrt{6}},\ 0,\ \frac{\sqrt{2}}{\sqrt{3}}\right)\right.$.

7. $\frac{1}{202}(187,\ 55,\ 33,\ -99)$.

8. $\vec{c}=-\frac{7}{17}\vec{a}+\frac{8}{17}\vec{b}+\frac{1}{17}(7,\ 6,\ 25,\ 19,\ 1)$.

10. *a*) $\left\{\left(\frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}}\right),\ \left(-\frac{1}{\sqrt{6}},\ \frac{2}{\sqrt{6}},\ \frac{1}{\sqrt{6}}\right)\right\}$.



*b*) $\frac{\sqrt{3}}{\sqrt{2}}x$.

*c*) $\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\right.$

$\left.\frac{\sqrt{2}}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\right\}$

### 8.4

1. *a*) $W^{\perp}=L\{(1,\ 1,\ -1)\}$. *b*) $W^{\perp}=L\{(1,\ 1,\ 1,\ 1),\ (3,\ 0,\ 1,\ 1)\}$.

2. *a*) $P_W(\vec{x})=\frac{1}{19}[(2x_1+4x_2-3x_3-3x_4)(1,\ 2,\ 0,\ -1)+(-3x_1-6x_2+14x_3$ $-5x_4)(0,\ 0,\ 1,\ -1)$.

*b*) $P_{W(3,2)}(\vec{x})=\frac{1}{19}(17,\ -4,\ 3,\ 3)+P_W(\vec{x})$.

4. $\vec{g}=(3,\ 1,\ -1,\ -2)\in L(\vec{b}_1,\ \vec{b}_2)$, $\vec{h}=(2,\ 1,\ -1,\ 4)\perp L(\vec{b}_1,\ \vec{b}_2)$.

5. [Ejercicio 10, sección 8.3: *a*) $W^{\perp}=L(1,\ 1,\ -1)$. *b*) $W^{\perp}=L\{x^3,\ x^2,\ 1\}$.

*c*) $W^{\perp}=L\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right\}$.]

*a*) $W^{\perp}=\{A\in \mathcal{M}_{3\times 3}(\mathbb{R})/a_{11}=a_{22}=a_{33}=0\}$.

*b*) $P_W(X)=\begin{pmatrix} x_{11} & 0 & 0 \\ 0 & x_{22} & 0 \\ 0 & 0 & x_{33} \end{pmatrix}$ ; $\min d(X,\ W)=\sum_{i=1}^{3}\sum_{j=1}^{3} x_{ij}^2-(x_{11}^2+x_{22}^2+x_{33}^2)$.

6. $P_W(\vec{x})=\frac{1}{11}[(5x_1+5x_2-x_3+2x_4)(1,\ 1,\ -1,\ 0)+(2x_1+2x_2+4x_3+3x_4)(0,\ 0,\ 2,\ 1)]$.

8. $\vec{u}_f=(2,\ 1,\ -1)$.

### 8.5 y 8.6

1. Todas las aplicaciones dadas son autoadjuntas.

2. *a*) *No* es autoadjunta, ya que su matriz en la base canónica de $\mathbb{R}^3$ es $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$.

*b*) *Sí* es autoadjunta, ya que su matriz en la base canónica de $\mathbb{R}^3$ es $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 1 \end{pmatrix}$.

47. $\begin{pmatrix} 2 & -11 & 6 \\ 1 & -7 & 4 \\ 2 & -1 & 0 \end{pmatrix}$.

48. $a=1$; $-\vec{u}_1-\vec{u}_2+\vec{u}_3=\vec{0}$.

49. $(0,\ -1,\ -1)+t(1,\ 2,\ 4)$ ; $\left.\begin{matrix} 2x-3y+z=2 \\ 2x+y-z=0 \end{matrix}\right\}$.

50. $3x-4y=10$ ó $3x-4y=-10$.

51. *a)* $\begin{pmatrix} 5-i & -4+i & i-1 \\ -1-5i & 1+4i & 1+i \end{pmatrix}$.

*b)* $\ker(A)=L\{(4-i,\ 5-i,\ 0),\ (1-i,\ 0,\ 5-i)\}$; $\operatorname{img}(A)=L\{(c,\ 1)\}$.

52. La matriz es diagonalizable para todo $a\neq -1$.

## CAPITULO 8

### 8.1 y 8.2

1. *a)* No. *b)* No. *c)* No. *d)* No. *e)* Sí.

2. $\|\vec{x}\|=\sqrt{2(x_1+x_2)^2+x_2^2}$; $\cos \sphericalangle(\vec{x},\ \vec{y})=\dfrac{2x_1y_1+3x_2y_2+2x_1y_2+2x_2y_1}{\sqrt{2(x_1+x_2)^2+x_2^2}\ \sqrt{2(y_1+y_2)^2+y_2^2}}$.

3. $\pi/2$. 4. $\arccos(1/\sqrt{n})$.

### 8.3

1. La matriz del producto escalar es $\begin{pmatrix} 2 & 1 & 5 \\ 1 & 1 & 3 \\ 5 & 3 & 14 \end{pmatrix}$

2. $\vec{y}_1=(1,\ 0,\ 0)$, $\vec{y}_2=(0,\ -2,\ 0)$, $\vec{y}_3=(0,\ 0,\ 5)$.

4. $\{(1,\ 2,\ 1,\ 3),\ \left(\frac{10}{3},\ \frac{-1}{3},\ \frac{1}{3},\ -1\right),\ \left(\frac{-19}{185},\ \frac{87}{185},\ \frac{61}{185},\ \frac{-72}{185}\right)$.

5. *a)* $\{(1,\ 1,\ 1)\}$. *b)* $\{(2,\ -1,\ 0),\ (-4,\ 11,\ -6)\}$.

6. $\left\{\left(\frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}},\ 0,\ 0\right),\ \left(0,\ \frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}},\ 0\right),\ \left(\frac{1}{\sqrt{6}},\ 0,\ \frac{-1}{\sqrt{6}},\ 0,\ \frac{\sqrt{2}}{\sqrt{3}}\right)\right\}$.

7. $\frac{1}{202}(187,\ 55,\ 33,\ -99)$.

8. $\vec{c}=-\frac{7}{17}\vec{a}+\frac{8}{17}\vec{b}+\frac{1}{17}(7,\ 6,\ 25,\ 19,\ 1)$.

10. *a)* $\left\{\left(\frac{1}{\sqrt{2}},\ 0,\ \frac{1}{\sqrt{2}}\right),\ \left(-\frac{1}{\sqrt{6}},\ \frac{2}{\sqrt{6}},\ \frac{1}{\sqrt{6}}\right)\right\}$.



*b)* $\frac{\sqrt{3}}{\sqrt{2}}x$.

*c)* $\left\{\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},\right.$

$\left.\frac{\sqrt{2}}{\sqrt{3}}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix},\ \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\right\}$

### 8.4

1. *a)* $W^{\perp}=L\{(1,\ 1,\ -1)\}$. *b)* $W^{\perp}=L\{(1,\ 1,\ 1,\ 1),\ (3,\ 0,\ 1,\ 1)\}$.

2. *a)* $P_W(\vec{x})=\frac{1}{19}[(2x_1+4x_2-3x_3-3x_4)(1,\ 2,\ 0,\ -1)+(-3x_1-6x_2+14x_3-5x_4)(0,\ 0,\ 1,\ -1)$.

*b)* $P_{W(3,\,2)}(\vec{x})=\frac{1}{19}(17,\ -4,\ 3,\ 3)+P_W(\vec{x})$.

4. $\vec{g}=(3,\ 1,\ -1,\ -2)\in L(\vec{b}_1,\ \vec{b}_2)$, $\vec{h}=(2,\ 1,\ -1,\ 4)\perp L(\vec{b}_1,\ \vec{b}_2)$.

5. [Ejercicio 10, sección 8.3: *a)* $W^{\perp}=L(1,\ 1,\ -1)$. *b)* $W^{\perp}=L\{x^3,\ x^2,\ 1\}$.

*c)* $W^{\perp}=L\left\{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right\}$.]

*a)* $W^{\perp}=\{A\in\mathcal{M}_{3\times 3}(\mathbb{R})/a_{11}=a_{22}=a_{33}=0\}$.

*b)* $P_W(X)=\begin{pmatrix} x_{11} & 0 & 0 \\ 0 & x_{22} & 0 \\ 0 & 0 & x_{33} \end{pmatrix}$ ; $\min d(X,\ W)=\sum_{i=1}^{3}\sum_{j=1}^{3}x_{ij}^2-(x_{11}^2+x_{22}^2+x_{33}^2)$.

6. $P_W(\vec{x})=\frac{1}{11}[(5x_1+5x_2-x_3+2x_4)(1,\ 1,\ -1,\ 0)+(2x_1+2x_2+4x_3+3x_4)(0,\ 0,\ 2,\ 1)]$.

8. $\vec{u}_f=(2,\ 1,\ -1)$.

### 8.5 y 8.6

1. Todas las aplicaciones dadas son autoadjuntas.

2. *a) No* es autoadjunta, ya que su matriz en la base canónica de $\mathbb{R}^3$ es $\begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$.

*b) Sí* es autoadjunta, ya que su matriz en la base canónica de $\mathbb{R}^3$ es $\begin{pmatrix} 1 & 2 & 1 \\ 2 & 0 & 3 \\ 1 & 3 & 1 \end{pmatrix}$.

3. *a*) $A=\begin{pmatrix}2&1\\1&2\end{pmatrix}$; $J=\begin{pmatrix}3&0\\0&1\end{pmatrix}$, $P=\dfrac{1}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$.

*b*) $B=\begin{pmatrix}0&1&1\\1&0&1\\1&1&0\end{pmatrix}$; $J=\begin{pmatrix}-1&0&0\\0&-1&0\\0&0&2\end{pmatrix}$, $P=\begin{pmatrix}1/\sqrt{2}&1/\sqrt{6}&1/\sqrt{3}\\0&-2/\sqrt{6}&1/\sqrt{3}\\-1/\sqrt{2}&1/\sqrt{6}&1/\sqrt{3}\end{pmatrix}$.

*c*) $T=\begin{pmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{pmatrix}$; $J=\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{pmatrix}$, $P=\begin{pmatrix}1&0&0&0\\0&1/\sqrt{2}&0&1/\sqrt{2}\\0&1/\sqrt{2}&0&-1/\sqrt{2}\\0&0&-1&0\end{pmatrix}$.

**8.7 y 8.8**

1. *a*) Ortogonal. *b*) No es ortogonal. *c*) No (observar que el determinante de la matriz es $-2$.)
2. Base ortonormal $=\left\{\dfrac{1}{\sqrt{3}}(1,1,-1),\ \dfrac{1}{\sqrt{6}}(1,1,2),\ \dfrac{1}{\sqrt{6}}(\sqrt{3},-\sqrt{3},0)\right\}$.

$$J=\begin{pmatrix}1&0&0\\0&-\frac{1}{2}&\frac{\sqrt{3}}{2}\\0&-\frac{\sqrt{3}}{2}&-\frac{1}{2}\end{pmatrix}.$$

3. $\alpha=\gamma=2,\ \beta=1$.
8. *b*) $x+2y=0$. *c*) $T=\begin{pmatrix}1&0\\0&-1\end{pmatrix}$.
9. *c*) $3x+y=0$. *d*) $P=\begin{pmatrix}1&0\\0&0\end{pmatrix}$.
11. *a*) Sí. *b*) No. *c*) Sí.

**8.9**

2. *a*) $S=\begin{pmatrix}\sqrt{2}&0\\0&\sqrt{2}\end{pmatrix}$, $O=\dfrac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}$.

*b*) $S=\begin{pmatrix}1&0&1\\0&1&0\\1&0&0\end{pmatrix}$, $O=\begin{pmatrix}-1&0&0\\0&1&0\\0&0&1\end{pmatrix}$.

[*Sugerencia*: trabajar con $B=\begin{pmatrix}-1&-1\\1&0\end{pmatrix}$; hallar una raíz cuadrada $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ de $B^tB=\begin{pmatrix}1&1\\1&0\end{pmatrix}$.]



3. $\dfrac{1}{\sqrt{2}}\begin{pmatrix}1&0&1\\0&1&0\\1&0&1\end{pmatrix}$. [*Sugerencia*: Calcular *a*, *b* y *c* en la igualdad $\begin{pmatrix}1&1\\1&1\end{pmatrix}=\begin{pmatrix}a&b\\b&c\end{pmatrix}^2$.]

## CAPITULO 9

1. $\{(z_1,z_2,z_3)\in\mathbb{C}^3/z_3=iz_1\}$.
2. No, sí, no.
3. *b*) $J=\begin{pmatrix}\frac{1+\sqrt{3}i}{2}&0\\0&\frac{1-\sqrt{3}i}{2}\end{pmatrix}$. *c*) $J=\begin{pmatrix}\frac{i+\sqrt{-1-12i}}{2}&0\\0&\frac{i-\sqrt{-1-12i}}{2}\end{pmatrix}$.
4. $A$ tiene en la base $\vec{u}_1=(0,0,1)$, $\vec{u}_2=(1,1,0)$, $\vec{u}_3=(1,-1,0)$ la forma diagonal

$$\begin{pmatrix}-1&0&0\\0&i&0\\0&0&-i\end{pmatrix}$$

5. $\begin{pmatrix}1+i&1\\1&1-i\end{pmatrix}=\begin{pmatrix}1+i&0\\0&1-i\end{pmatrix}\begin{pmatrix}1&\frac{1-i}{2}\\\frac{1+i}{2}&1\end{pmatrix}=$

$$=\left[\frac{1}{\sqrt{2}}\begin{pmatrix}1+i&0\\0&1-i\end{pmatrix}\right]\left[\sqrt{2}\begin{pmatrix}1&\frac{1-i}{2}\\\frac{1+i}{2}&1\end{pmatrix}\right]$$

(La primera es unitaria y la segunda simétrica-conjugada.)

*Observaciones sobre el problema* 5

$U=\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}$ con $|a|^2+|b|^2=1$ y $S=\begin{pmatrix}r&c\\\bar{c}&s\end{pmatrix}$ con $r,s\in\mathbb{R}$. Escribiendo $\begin{pmatrix}1+i&1\\1&1-i\end{pmatrix}=\begin{pmatrix}a&b\\-\bar{b}&\bar{a}\end{pmatrix}\begin{pmatrix}r&c\\\bar{c}&s\end{pmatrix}$ e igualando, se obtienen varias ecuaciones, que se van resolviendo hasta obtener el resultado.

6. *a*) $(\vec{z}+i\vec{w},\vec{u}+i\vec{v})=(\vec{z},\vec{u})+(\vec{w},\vec{v})+i[-(\vec{z},\vec{v})+(\vec{w},\vec{u})]=\overline{(\vec{u}+i\vec{v},\vec{z}+i\vec{w})}$.
   *b*) Se deduce de la propiedad distributiva del producto euclídeo.
   *c*) Se deduce de la propiedad distributiva del producto euclídeo.
   *d*) $(\vec{u}+i\vec{v},\vec{u}+i\vec{v})=(\vec{u},\vec{u})+(\vec{v},\vec{v})+i[-(\vec{u},\vec{v})+(\vec{v},\vec{u})]=$
   $=\|\vec{u}\|^2+\|\vec{v}\|^2>0$ salvo si $\vec{u}=\vec{v}=0$.

## CAPITULO 10

### 10.1

1. $T\begin{pmatrix} x \\ y \end{pmatrix}=\begin{pmatrix} 2 \\ 0 \end{pmatrix}+\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$

2. *a*) $T\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}+\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$ *b*) $T\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}+\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$

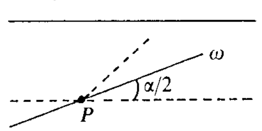
   *c*) $T \circ G\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix},$

   $(G \circ T)\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$

3. Todos los puntos del plano $2x+y+z=2$ son puntos fijos. La aplicación es una simetría con respecto a este plano.

### 10.2

1. El centro de giro es el punto de intersección de los ejes. El ángulo de giro es el doble del ángulo que forman las rectas.
2. La composición de dos simetrías de ejes paralelos es una traslación de vector $2\vec{v}$, donde $\vec{v}$ es un vector perpendicular a las rectas y $\|\vec{v}\|=d(r, l)$.
3. El eje de la simetría deslizante es paralelo a la recta $w$.




5. $S_t \circ S_s \circ S_r$ es una simetría con respecto a la recta $l$, que es bisectriz del ángulo $2 \measuredangle (s, t)$ y uno de cuyos lados es $r$.

### 10.3

1. *a*) $X'=\begin{pmatrix} 1+\frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{pmatrix}+\begin{pmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}X.$ *b*) $X'=\begin{pmatrix} 3 \\ 1 \end{pmatrix}+\begin{pmatrix} -\frac{3}{5} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{3}{5} \end{pmatrix}X.$

   *c*) $X'=\begin{pmatrix} \frac{\sqrt{3}}{2} \\ -1-\sqrt{3} \end{pmatrix}+\begin{pmatrix} \frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{1}{2} \end{pmatrix}X.$



   *d*) $X'=\begin{pmatrix} \frac{12}{5}+\frac{\sqrt{2}}{10} \\ \frac{1}{5}-\frac{7\sqrt{2}}{10} \end{pmatrix}+\begin{pmatrix} -\frac{\sqrt{2}}{10} & \frac{7\sqrt{2}}{10} \\ \frac{7\sqrt{2}}{10} & \frac{\sqrt{2}}{10} \end{pmatrix}X.$

2. *a*) Giro de ángulo $\pi/3$ y centro $P=\left(-1-\frac{\sqrt{3}}{2}, -\sqrt{3}+\frac{1}{2}\right)$.

   *b*) Simetría deslizante: Eje $(2-\sqrt{2})x_1-\sqrt{2}x_2+\frac{\sqrt{2}}{2}=0$. Vector: $\left(\frac{\sqrt{2}}{4}, -\frac{\sqrt{2}}{4}+\frac{1}{2}\right)$.
3. *a*) Giro de ángulo $\pi/2$ y centro (1, 1).
   *b*) La simetría deslizante de eje $x_2=1/2$ y vector (1, 0).
4. *a*) Traslaciones de vector $\vec{v}$ paralelo al eje de simetría.
   *b*) Giros de ángulo $\pi$ y centro perpendicular al eje.
   *c*) Simetrías respecto a rectas perpendiculares al eje.
5. *a*) Traslaciones; simetrías respecto de rectas con vector director $\vec{u}_0$; simetrías deslizantes respecto de rectas con vector director $\vec{u}_0$.
   *b*) Giros de centro $P_0$; simetrías respecto de rectas que pasan por $P_0$.

### 10.4 y 10.5

1. *a*) $S_\pi\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\frac{1}{7}\left[\begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}+\begin{pmatrix} -2 & 3 & -6 \\ 3 & 6 & 2 \\ -6 & 2 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right]$

   *b*) $M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}+\begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

   *c*) $G\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$

   *d*) $M \circ S_\pi\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\frac{1}{7}\left[\begin{pmatrix} 15 \\ -17 \\ -2 \end{pmatrix}+\begin{pmatrix} -3 & -6 & -2 \\ 2 & -3 & 6 \\ 6 & -2 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right]$

2. *a*) Simetría respecto al plano $\pi$: $x-\sqrt{2}y-z=\frac{1+\sqrt{2}}{2}$, *con vector de* deslizamiento $\frac{1}{4}(3-\sqrt{2}, \sqrt{2}-2, \sqrt{2}+1)$.

   *b*) Movimiento helicoidal: $r$: $(x, y, z)=\frac{1}{3}(1, 0, 1)+\lambda(1, 1, -1)$; giro $\alpha=120°$ en el sentido del sacacorchos para que avance en la dirección $(1, 1, -1)$; vector de deslizamiento $\frac{1}{3}(2, 2, -2)$.

*c*) Movimiento helicoidal: $r$: $\{x=1,\ z=1\}$; giro $\alpha=-90°$ (el sentido de giro positivo es el del sacacorchos para que avance en la dirección (0, 1, 0)); vector de deslizamiento (0, 1, 0).

*d*) $M_3 \circ M_2$ es un movimiento heicoidal: $r$: $\{y=1/2,\ z=3/2\}$, giro $\alpha=90°$ en el sentido del sacarcorchos para que avance en la dirección de (1, 0, 0); vector de deslizamiento: (2, 0, 0).

$M_2 \circ M_3$ es un movimiento helicoidal: $r$: $\{x=3/2,\ y=-1/2\}$, giro $\alpha=-90°$ (el sentido de giro positivo es el del sacacorchos para que avance en la dirección de (0, 0, 1); vector de deslizamiento: (0, 0, −2).

3. Todos los movimientos que afectan sólo al plano $xy$ y la composición de estos movimientos con una simetría respecto a este mismo plano $xy$.

4. $M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}+\left[\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}+\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right].$

5. *a*) $M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 3 \\ 3 \\ 0 \end{pmatrix}+\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$; vector de traslación (3, 3, 0).

*b*) Vector de traslación: $2\bar{v}$, donde $\bar{v}$ es un vector normal a los planos, que va de uno de los planos al otro.

6. *a*) $M\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 4/3 \\ 4/3 \\ 10/3 \end{pmatrix}+\begin{pmatrix} 1/3 & -2/3 & 2/3 \\ -2/3 & 1/3 & 2/3 \\ -2/3 & -2/3 & -1/3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$; $r$: $\begin{matrix} z=1 \\ x+y=3 \end{matrix}$.

Giro $\alpha=\arccos(-1/3)$ en el sentido del sacacorchos para que avance en la dirección (−1, 1, 0).

*b*) El eje es la intersección de los dos planos.

*c*) El ángulo de giro es el doble del ángulo que forman los planos.

## CAPITULO 11

### 11.2, 11.3 y 11.4

1. $\|F_1C\|=\dfrac{\varepsilon}{1-\varepsilon}\|V_1F_1\|$, $C=$centro;

$\|V_1V_2\|=\dfrac{2}{1-\varepsilon}\|V_1F_1\|$, $V_2=$vértice, $b=\sqrt{\dfrac{1+\varepsilon}{1-\varepsilon}}\|V_1F_1\|$.

2. $\|V_1C\|=\dfrac{1}{\varepsilon-1}\|V_1F_1\|$, $C=$centro; las asíntotas son las rectas de pendiente $\pm\sqrt{\varepsilon^2-1}$ que pasan por $C$.

### 11.5

2. $y^2-6x+2y+4=0$; $y^2+2y+6x-20=0$.
3. $3x^2-4xy-4=0$.



4. $x^2+9y^2+8x-36y+43=0$.
5. $20x^2-5y^2-10y-41=0$.
6. $64x^2+16xy+76y^2-393x-87y+596=0$.

### 11.7, 11.8, 11.9 y 11.10

1. Parábola; $V=\frac{1}{2}(1, 5)$; forma canónica $y^2=\dfrac{x}{\sqrt{2}}$; eje: $x+y=3$.
2. Hipérbola; $C=\frac{1}{6}(1, -2)$; forma canónica $\dfrac{12x^2}{5}-\dfrac{18y^2}{5}=1$; ejes; $y=2x-\frac{2}{3}$, $2y+x+\frac{1}{2}=0$.
3. Dos rectas paralelas; $x-y-1=0$, $x-y+5=0$.
4. Un punto $(-1, 1)$.
5. Elipse; $C=\dfrac{3}{\sqrt{5}}$; $\dfrac{5x^2}{12}+\dfrac{20y^2}{12}=1$; $\varepsilon=\dfrac{3}{2\sqrt{5}}$; $C=\frac{1}{5}(1, 2)$; ejes: $x+2y-1=0$, $y=2$.
6. Tipo parabólico: dos rectas coincidentes $(x+2y-3)^2=0$.
7. Hipérbola equilátera; $C=(-1, 3)$; $\dfrac{x^2}{5}-\dfrac{y^2}{5}=1$.
8. Tipo hiperbólico; dos rectas $x-2y+3=0$, $x+y-1=0$.
9. Hipérbola equilátera; $\dfrac{4x^2}{5}-\dfrac{4y^2}{5}=1$; asíntotas $x-y+\frac{1}{2}=0$, $x+y+\frac{1}{2}=0$
10. Elipse: $\dfrac{x^2}{2}+\dfrac{9y^2}{2}=1$; $C=\frac{1}{8}(-3, 1)$; ejes: $y=x+\frac{1}{2}$, $y=-x-\frac{1}{4}$.
11. No tiene soluciones reales.
12. Elipse; $\dfrac{x^2}{8}+\dfrac{y^2}{16}=1$; ejes: $x-y=-1$, $x+y=-7$; $C=(-4, -3)$; focos: $(-5+\sqrt{3}, -4+\sqrt{3})$ y $(-5-\sqrt{3}, -4-\sqrt{3})$.
13. Parábola; $2y^2=\sqrt{2}x$.
14. *a*) $y=\pm(1+\sqrt{2})\,x$ *b*) $y+\frac{4}{5}=\frac{11}{2}(x-\frac{3}{5})$; $y+\frac{4}{5}=\frac{1}{2}\left(x-\frac{3}{5}\right)$
15. *a*) $V=(0, 0)$, eje: $4x=3y$ *b*) $V=\left(\frac{5}{7}, \frac{5}{7}\right)$, eje: $4x=3y$.

## CAPITULO 12

### 12.1

1. *a*) $\begin{pmatrix} 2 & -3/2 & 3 \\ -3/2 & 4 & 0 \\ 3 & 0 & 0 \end{pmatrix}$. *b*) $\dfrac{1}{9}\begin{pmatrix} 6 & -3 & -3 \\ -3 & 6 & -3 \\ -3 & -3 & 6 \end{pmatrix}$. *c*) $\begin{matrix} a_{ii}=0,\ i=1, 2, ..., n \\ a_{ij}=(i-j)^2/2,\ i\neq j \end{matrix}$

### 12.2

4. *a)* $-\frac{11}{2}x_1^2+2y_1^2+z_1^2$. *b)* $x_2^2-y_2^2$. *c)* $x_1^2-4y_1^2-z_1^2$.

5. *a)* $2x_1^2-3y_1^2$. *b)* $-6x_1^2+\sqrt{5}y_1^2-\sqrt{5}z_1^2$. *c)* $x_1^2+\sqrt{3}y_1^2-\sqrt{3}z_1^2$.
   *d)* $\frac{1}{2}y_1^2-\frac{1}{2}y_2^2+\frac{1}{2}y_3^2+\frac{1}{2}y_4^2$.

6. *a)* $\left\{\left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right), \left(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right), \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right)\right\}$, $3x^2+6y_1^2+9z_1^2$.
   *b)* $\left\{\frac{1}{\sqrt{6}}(2, -1, -1), \frac{1}{\sqrt{5}}(1, 2, 0), \frac{1}{\sqrt{29}}(2, -1, 5)\right\}$, $-2x_1^2+4y_1^2+4z_1^2$.
   *c)* $\left\{\frac{1}{\sqrt{2}}(1, 1, 0), \frac{1}{\sqrt{3}}(1, -1, -1), \frac{1}{\sqrt{6}}(1, -1, 2)\right\}$, $x_1^2+2y_1^2-z_1^2$.

### 12.3

7. *a)* $x^2-y^2+z^2$; índice de inercia positivo $=2$.
   *b)* $x^2+y^2-z^2$; índice de inercia positivo $=1$.
   *c)* $x_1^2-y_1^2+z_1^2$; índice de inercia positivo $=2$.

$$\left.\begin{array}{l} x_1=2x-y+z \\ y_1=(y+z)/2 \\ z_1=(y-z)/2 \end{array}\right\}$$

### 12.4

8. Ningún valor. 9. $\beta\in(-\infty, -1)$.

10. $\{(x, y, z)\in\mathbb{R}^3/|z|<1, y\neq 0\}$.

11. *a)* $x=-1$, $y=0$ es un mínimo.
    *b)* $(2, -2)$ y $(-2, -2)$ son máximos.
    *c)* $(1, -1)$ es un mínimo.

13. *a)* $0<a<\sqrt{2}$. *b)* Nunca.

### 12.5

14. $A(\vec{x}, \vec{x})=x_1^2+x_2^2$ ; $B(\vec{x}, \vec{x})=4x_1^2-2x_2^2$ ; $C=\frac{1}{\sqrt{2}}\begin{pmatrix} 4 & 6 \\ -1 & -1 \end{pmatrix}$.



## CAPITULO 13

1. *a)* Elipsoide; $x^2+y^2+4z^2=1$; centro $=(0, 0, 0)$.
   *b)* Hiperboloide de dos hojas con eje en la dirección de $OY$; centro $=\left(-2, -\frac{1}{2}, 1\right)$.
   *c)* Hiperboloide de una hoja con eje $OY$; $-2x^2+\frac{1}{2}y^2-\frac{1}{2}z^2=-7$; centro $=(-6, -1, 0)$; ejes $\{x=-6,\ z=0\}$, $\{y=-1,\ x=z-6\}$, $\{y=-1,\ z=-x-6\}$.
   *d)* Paraboloide hiperbólico; $\frac{9}{2}x^2-\frac{9}{2}y^2\pm\frac{4}{9}\sqrt{\Delta}z=0$.

4. *a)* No tiene soluciones reales.
   *b)* Hiperboloide de dos hojas; $-x^2+(3+\sqrt{6})y^2-(3-\sqrt{6})z^2=\frac{13}{2}$.
   *c)* Hiperboloide de una hoja; $-x^2-\sqrt{17}y^2+\sqrt{17}z^2=-\frac{11}{17}$.
   *d)* Elipsoide; $x^2+\left(\frac{5}{2}+\frac{\sqrt{17}}{2}\right)y^2+\left(\frac{5}{2}-\frac{\sqrt{17}}{2}\right)z^2=\frac{1}{2}$.
   *e)* Hiperboloide de una hoja.
   *f)* El punto $(0, 0, -1)$.
   *g)* Paraboloide hiperbólico; $\frac{1+\sqrt{5}}{2}x^2+\frac{1-\sqrt{5}}{2}y^2=-2z$.

**Soluciones a los ejercicios de repaso de los capítulos 8 a 13**

1. $J_A=\begin{pmatrix} 3 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{pmatrix}$ con $C=\frac{1}{3}\begin{pmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$ ; $J_B=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 11 & 0 \\ 0 & 0 & 11 \end{pmatrix}$ con $C=\frac{1}{\sqrt{6}}\begin{pmatrix} \sqrt{2} & -\sqrt{3} & 1 \\ \sqrt{2} & 0 & -2 \\ \sqrt{2} & \sqrt{3} & 1 \end{pmatrix}$.

2. $0\leqq(\mathscr{A}\vec{x}, \mathscr{A}\vec{x})=-(\vec{x}, \mathscr{A}\mathscr{A}\vec{x})=-(\vec{x}, \mathscr{A}(\lambda\vec{x}))=-\lambda^2(\vec{x}, \vec{x}) \Rightarrow -\lambda^2\geqq 0$ (las igualdades anteriores son ciertas si $\lambda$ es un autovalor real de $\vec{x}\neq 0$).

3. $\vec{e}_1=(1, 2, 2, -1)$, $\vec{e}_2=(2, 3, -3, 2)$, $\vec{e}_3=(2, -1, -1, -2)$.

4. $-1/2\sqrt{3}$.

6. $\|\vec{v}\|=\sqrt{3}$ ; $x_1-y_1-z_1=0$.

7. Simetría con respecto a la recta $x=-y=z$.
8. $W_1=\{(x, y, z)\in\mathbb{R}^3/x=y=z\}$ ; $W_2=\{(x, y, z)\in\mathbb{R}^3/y=0\}$.
9. $T^*=\frac{1}{3}\begin{pmatrix}4 & 8\\ 1 & 2\end{pmatrix}$.
10. $P(\vec{v})=(2, 0, 0)$.
11. Sí es autoadjunta. No es ortogonal.
12. Giro de centro $\frac{1}{10-4\sqrt{5}}\begin{pmatrix}5-\sqrt{5}\\ 5-3\sqrt{5}\end{pmatrix}$ y de ángulo de giro $\varphi$ con $\cos\varphi=\frac{2}{\sqrt{5}}$, sen $\varphi$ $=-1/\sqrt{5}$.
14. $\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}3\\ 3\end{pmatrix}+\begin{pmatrix}0 & -1\\ -1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$.
15. $\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=\begin{pmatrix}2/3\\ 2/3\\ 2/3\end{pmatrix}+\begin{pmatrix}1/3 & -2/3 & -2/3\\ -2/3 & 1/3 & -2/3\\ -2/3 & -2/3 & 1/3\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}$.
16. *a*) $\begin{pmatrix}x'\\ y'\\ z'\end{pmatrix}=\begin{pmatrix}0\\ 0\\ 2\end{pmatrix}+\frac{1}{3}\begin{pmatrix}-1 & 2 & 2\\ 2 & 4 & -1\\ 2 & -1 & 4\end{pmatrix}\begin{pmatrix}x\\ y\\ z\end{pmatrix}$. *b*) Simetría deslizante.
17. *a*) Para todo $\lambda>2$. *b*) Para ningún $\lambda$.
18. Forma canónica $4y_1^2+4y_2^2-2y_3^2$ en la base $\vec{u}_1=\frac{1}{\sqrt{5}}(1, 0, 2)$, $\vec{u}_2=\frac{1}{\sqrt{30}}(2, 5, -1)$, $\vec{u}_3=\frac{1}{\sqrt{6}}(-2, 1, 1)$.
19. Parábola; forma canónica $25X^2=\frac{2}{\sqrt{5}}\sqrt{875}\,Y$.
20. Centro $\left(\frac{3}{5}, \frac{4}{5}\right)$; eje principal $4x=3y$.
21. Hiperboloide de dos hojas; $25z^2-5x^2-5y^2=1$.
23. $2X^2+Y^2-Z^2$ ; $\begin{pmatrix}x\\ y\\ z\end{pmatrix}=\begin{pmatrix}0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}}\\ 0 & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}}\\ 1 & 0 & 0\end{pmatrix}\begin{pmatrix}X\\ Y\\ Z\end{pmatrix}$.
24. $a=b-2$, $a>-2$.
28. *a*) $x^2+y^2-(\alpha^2+\beta^2)z^2-2a\alpha z=a^2$.
30. *c*) 1. $3x-y+1=0$. 2. $x-y+2=0$.



31. $\arccos\frac{1}{\sqrt{3}}$.
33. $\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix}+\begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$: giro de 90° alrededor del punto $\left(\frac{1}{2}, \frac{1}{2}\right)$.
$\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}1\\ 0\end{pmatrix}+\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$: simetría deslizante respecto a la recta $x-y=1/2$, con vector de traslación $\left(\frac{1}{2}, \frac{1}{2}\right)$.
36. $P(\vec{u}_3)=-\vec{u}_2/2$.
37. I) No es definida. II) Definida positiva. III) Semidefinida positiva.
38. Simetría deslizante; plano de simetría $x-y=1$; vector de traslación $\vec{v}=\begin{pmatrix}0\\ 0\\ 1\end{pmatrix}$.
39. Hipérbola; $\frac{x^2}{\left(\frac{11}{375}\right)}-\frac{y^2}{\left(\frac{11}{1.125}\right)}=1$; eje focal $2x-y=1/15$.
40. $\begin{pmatrix}x'\\ y'\end{pmatrix}=\begin{pmatrix}2 & 1\\ 1 & -1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$ es la matriz del cambio; esto es, $\left.\begin{array}{l}x'=2x+y\\ y'=x-y\end{array}\right\}$.

# INDICE ALFABETICO