Capítulo 12

El espacio euclídeo

12.1. El espacio euclídeo n-dimensional

Definición 101 ***El espacio euclídeo*** n***-dimensional***, *denotado* $\mathbb{R}^n$, *consta de todas las n-uplas ordenadas de números reales. Simbólicamente,*

$$\mathbb{R}^n = \{(x_1, ..., x_n) \mid x_1, ..., x_n \in \mathbb{R}\}.$$

Por lo tanto, $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ *(n veces) es el producto cartesiano de* $\mathbb{R}$ *consigo mismo n veces.*

Los elementos de $\mathbb{R}^n$*, en general, se denotan con una sola letra que representa una n-upla, como* $x = (x_1, x_2, ..., x_n)$*, y decimos que* x *es un* ***punto*** *de* $\mathbb{R}^n$.

La suma y el producto por un escalar de n-uplas se definen mediante

$$(x_1, ..., x_n) + (y_1, ..., y_n) = (x_1 + y_1, ..., x_n + y_n)$$

y

$$a \cdot (x_1, ..., x_n) = (a \cdot x_1, ..., a \cdot x_n) \quad \text{para } a \in \mathbb{R}.$$

Los espacios $\mathbb{R}^2$ *y* $\mathbb{R}^3$ *y, en general,* $\mathbb{R}^n$*,son ejemplos de una estructura del álgebra lineal llamada* ***espacio vectorial****. Básicamente, un espacio vectorial es una colección de objetos, llamados* ***vectores****, que se pueden sumar, restar y multiplicar por números.*

Definición 102 *Un* ***espacio vectorial real*** $\mathcal{V}$ *es un conjunto de elementos llamados* ***vectores****, dotado de las operaciones de* ***suma de vectores*** $+ : \mathcal{V} \times \mathcal{V} \to \mathcal{V}$ *y de* ***multiplicación por un escalar*** $\cdot : \mathbb{R} \times \mathcal{V} \to \mathcal{V}$ *que cumplen las siguientes propiedades:*

i. ***Conmutatividad:*** $v + w = w + v$ *para todo* $v, w \in \mathcal{V}$.

ii. ***Asociatividad:*** $(v + u) + w = v + (u + w)$.

iii. ***Vector cero:*** *Existe un vector cero,* 0*, tal que* $v + 0 = v$ *para todo* $v \in \mathcal{V}$.

iv. ***Vector opuesto:*** *Para cada* $v \in \mathcal{V}$*, existe un vector* $-v$ *tal que* $v+(-v) = 0$.

v. *Para cada* $\lambda \in \mathbb{R}$ *y* $v,w \in \mathcal{V}$*,* $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$.

vi. *Para* $\lambda, m \in \mathbb{R}$ *y* $v \in \mathcal{V}$, $\lambda \cdot (m \cdot v) = (\lambda m) \cdot v$.

vii. *Para* $\lambda, m \in \mathbb{R}$ *y* $v \in \mathcal{V}$, $(\lambda + m) \cdot v = \lambda \cdot v + m \cdot v$.

viii. $1 \cdot v = v$, *para todo* $v \in \mathcal{V}$.

Un espacio vectorial sobre un cuerpo general $\mathbb{F}$ *se define de la misma manera, pero sustituyendo* $\mathbb{F}$ *por* $\mathbb{R}$ *en todas partes. En análisis nos ocupamos principalmente de espacios vectoriales sobre* $\mathbb{R}$, *que se denominan* ***espacios vectoriales reales****, y sobre los complejos,* $\mathbb{C}$, *a los que llamaremos* ***espacios vectoriales complejos****.*

Teorema 217 *El espacio euclídeo* n*-dimensional, con las operaciones suma y multiplicación por un escalar anteriormente definidas, es un espacio vectorial de dimensión* n.

Definición 103 *Un* ***espacio métrico*** (M, d) *es un conjunto* M *y una función* $d : M \times M \to \mathbb{R}$ *que cumplen las siguientes propiedades:*

i. $d(x, y) \geq 0$ *para todo* $x, y \in M$

ii. $d(x, y) = 0$ *si y sólo si* $x = y$

iii. $d(x, y) = d(y, x)$ *para todo* $x, y \in M$

iv. $d(x, y) \leq d(x, z) + d(z, y)$ *para todo* $x, y, z \in M$

Ejemplo 273 *La recta real,* $\mathbb{R}$, *es un espacio métrico con la métrica definida por* $d(x, y) = |x - y|$.

Ejemplo 274 $\mathbb{R}^n$ *es un espacio métrico con la* ***métrica euclídea****:*

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Ejemplo 275 ***(Métrica discreta)*** *Sea* M *cualquier conjunto y sea*

$$d(x, y) = 0 \quad si \;\; x = y$$

$$d(x, y) = 1 \quad si \;\; x \neq y.$$

Entonces d *es una métrica sobre* M.

Ejemplo 276 ***(Métrica lexicográfica)*** *Este ejemplo es útil en combinatoria y en informática. Sea* M *el conjunto de todas las "palabras", cada una de las cuales consiste en un octeto ordenado de ceros y unos:*

$$w = (w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8) \quad \text{donde cada "bit" cumple } w_k = 0 \text{ ó } 1.$$

Sea $d(v, w)$ *el número de lugares en los que* v *y* w *son diferentes. Por ejemplo, si* $v = 01100011$ *y* $w = 00110101$, *entonces* $d(v, w) = 4$, *porque* v *y* w *difieren en los bits segundo, cuarto, sexto y séptimo. Entonces* d *es una métrica en* M. *(Para probar sus propiedades puede ser útil verificar que* $d(v, w) = \sum_{k=1}^{8} |v_k - w_k|$*).*

Ejemplo 277 ***(Métrica acotada)*** *Si d es una métrica sobre un conjunto M y $\rho(x, y)$ se define mediante*

$$\rho(x, y) = d(x, y)/(1 + d(x, y)),$$

entonces ρ también es una métrica sobre M. Esta nueva métrica tiene la propiedad de que $\rho(x, y) < 1$ para todo $x, y \in M$; es decir, está acotada por 1.

Aunque en ocasiones los espacios métricos son espacios vectoriales, esto no siempre ocurre, como lo prueba el espacio discreto. Vamos a centrarnos ahora en conceptos que sólo tienen sentido en espacios vectoriales.

Definición 104 *Un **espacio normado** (o **espacio vectorial normado**) $(\mathcal{V}, \|\cdot\|)$ es un espacio vectorial $\mathcal{V}$ y una función $\|\cdot\| : \mathcal{V} \to \mathbb{R}$, llamada **norma**, tal que*

i. $\|v\| \geq 0$ *para todo* $v \in \mathcal{V}$

ii. $\|v\| = 0$ *si y sólo si* $v = 0$

iii. $\|\lambda v\| = |\lambda| \cdot \|v\|$ *para todo* $v \in \mathcal{V}, \lambda \in \mathbb{R}$

iv. ***Desigualdad triangular**:* $\|v + w\| \leq \|v\| + \|w\|$

Ejemplo 278 $\mathbb{R}$ *es un espacio normado con* $\|x\| = |x|$.

Ejemplo 279 $\mathbb{R}^n$ *es un espacio normado con la **norma euclídea**:*

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Ejemplo 280 ***(Norma del taxi)*** *Consideremos el espacio $\mathbb{R}^2$, pero en lugar de la norma usual, definamos $\|(x, y)\|_1 = |x| + |y|$; $\|\cdot\|_1$ es una norma en $\mathbb{R}^2$. El nombre proviene de la distancia asociada: si $P = (x, y)$ y $Q = (a, b)$, entonces*

$$d_1(P, Q) = \|P - Q\|_1 = |x - a| + |y - b|.$$

o sea, la suma de las separaciones vertical y horizontal. Ésta es la distancia que se debe recorrer para llegar de P a Q si siempre se viaja paralelamente a los ejes (que es lo que hace un taxi al recorrer las calles).

Ejemplo 281 ***(Norma del supremo)*** *Sea M el conjunto de todas las funciones reales definidas en el intervalo $[0, 1]$ que están acotadas; es decir,*

$$M = \{f : [0, 1] \to \mathbb{R} \mid \textit{existe } B \in \mathbb{R} \textit{ tal que } |f(x)| \leq B \textit{ para todo } x \in [0, 1]\}$$

Para cada $f \in M$, $f([0, 1])$ es un subconjunto acotado de $\mathbb{R}$ y por lo tanto $\{|f(x)| \mid x \in [0, 1]\}$ también es acotado; en consecuencia, tiene un supremo finito y

$$\|f\|_\infty = \sup\{|f(x)| \mid x \in [0, 1]\}$$

define, por lo tanto, una función $\|\cdot\|_\infty : M \to \mathbb{R}$. El conjunto M es un espacio vectorial y $\|\cdot\|_\infty$ es una norma sobre él.

Como hicimos notar anteriormente, al hablar sobre $\mathbb{R}^n$, las normas siempre engendran métricas.

Teorema 218 *Si* $(\mathcal{V}, \|\cdot\|)$ *es un espacio vectorial normado y* $d(v,w)$ *está definida por*

$$d(v,w) = \|v-w\|,$$

entonces d *es una métrica sobre* $\mathcal{V}$.
Demostración:

i. $d(v,w) \geq 0$, *porque* $d(v,w) = \|v-w\| \geq 0$.

ii. *Para probar* $d(v,w) = 0 \Leftrightarrow v = w$, *obsérvese que*

$$d(v,w) = 0 \Leftrightarrow \|v-w\| = 0 \Leftrightarrow v-w = 0 \Leftrightarrow v = w$$

iii. $d(v,w) = d(w,v)$, *porque*

$$\begin{aligned} d(v,w) &= \|v-w\| = \|-(w-v)\| \\ &= |-1| \cdot \|w-v\| = \|w-v\| = d(w-v) \end{aligned}$$

iv. *Por la desigualdad triangular,*

$$\begin{aligned} d(x,y) &= \|x-y\| = \|(x-z)+(z-y)\| \\ &\leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y) \end{aligned}$$

■

En algunas ocasiones, las métricas engendradas por normas proporcionan una medida de la diferencia entre dos vectores. Ya vimos un ejemplo en la métrica del taxi. En el ejemplo del espacio de las funciones acotadas, la métrica es

$$d(f,g) = \|f-g\|_\infty = \sup\{|f(x)-g(x)| \mid 0 \leq x \leq 1\}.$$

Por tanto, la métrica dada por la norma del supremo es la mayor separación vertical entre sus gráficas.

Observación 122 *Aunque muchas métricas interesantes provienen de normas, no todas lo hacen, como por ejemplo la métrica discreta y la métrica acotada, salvo que el espacio vectorial consista sólo en el vector cero. Para demostrarlo elijamos un vector* v *diferente de cero y observemos que* $\|\lambda v\| = |\lambda| \cdot \|v\| \to \infty$ *cuando* $|\lambda| \to \infty$, *por lo que* $d(\lambda v, 0)$ *puede ser arbitrariamente grande. En consecuencia, una norma no puede producir las métricas discreta y acotada, puesto que ambas son acotadas.*

Definición 105 *Un espacio vectorial real* $\mathcal{V}$ *con una función* $\langle \cdot,\cdot \rangle : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$ *se llama* ***espacio con producto interno*** *si*

i. $\langle v,v \rangle \geq 0$ *para todo* $v \in \mathcal{V}$

ii. $\langle v,v \rangle = 0$ *si y sólo si* $v = 0$

iii. $\langle \lambda v, w \rangle = \lambda \langle v,w \rangle$ *para todo* $v,w \in \mathcal{V}$, $\lambda \in \mathbb{R}$

iv. $\langle v, w+u \rangle = \langle v,w \rangle + \langle v,u \rangle$ *para todo* $v,w,u \in \mathcal{V}$

v. $\langle v, w\rangle = \langle w, v\rangle$ *para todo* $v, w \in \mathcal{V}$

Ejemplo 282 *En* $\mathbb{R}^n$ *podemos definir el siguiente producto interno:*

$$\langle v, w\rangle = v_1 w_1 + \cdots + v_n w_n.$$

Ejemplo 283 ***(Funciones continuas)*** *Sea* $\mathcal{V}$ *el espacio de todas las funciones reales continuas sobre el intervalo* $[0,1]$. *La suma de funciones continuas es continua; por lo tanto,* $\mathcal{V}$ *es un espacio vectorial, que habitualmente se denota por* $\mathcal{C}([0,1])$, *y que tiene un gran interés en análisis. Cualquier función continua en* $[0,1]$ *es acotada, por lo que la norma del supremo tiene sentido en* $\mathcal{C}([0,1])$. *Esta norma no se puede asociar a ningún producto interno. Existe, sin embargo, un producto interno muy interesante en* $\mathcal{C}([0,1])$ *no asociado con la norma del supremo. Dado que el producto de funciones continuas es una función continua y las funciones continuas son integrables (verificaremos estas afirmaciones más adelante, pero por ahora las tomaremos como ciertas), podemos definir un producto interno en* $\mathcal{C}([0,1])$ *mediante*

$$\langle f, g\rangle = \int_0^1 f(x)g(x)dx.$$

Es inmediato (dadas las propiedades básicas de las integrales) verificar que éste es un producto interno.

Teorema 219 *Si* $\langle\cdot,\cdot\rangle$ *es un producto interno sobre un espacio vectorial real* $\mathcal{V}$, *entonces*

i. $\langle \lambda v + \mu w, u\rangle = \lambda\langle v, u\rangle + \mu\langle w, u\rangle$

ii. $\langle u, \lambda v + \mu w\rangle = \lambda\langle u, v\rangle + \mu\langle u, w\rangle$

iii. $\langle v, \lambda w\rangle = \lambda\langle v, w\rangle$

iv. $\langle 0, w\rangle = \langle w, 0\rangle = 0.$

La demostración es inmediata y se deja al lector.

Teorema 220 ***Desigualdad de Cauchy-Schwarz.*** *Si* $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ *es un espacio con producto interno, entonces*

$$|\langle x, y\rangle| \leq \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle} \qquad \text{para todo } \ x, y \in \mathcal{V}.$$

Demostración: *Si* x *o* y *son* 0, *entonces* $\langle x, y\rangle = 0$, *y, por lo tanto, la desigualdad se cumple. Así pues, podemos suponer que* $x \neq 0$ *e* $y \neq 0$. *Entonces* $\langle x, x\rangle > 0$ *y* $\langle y, y\rangle > 0$. *Para todo* $\alpha \in \mathbb{R}$, *usando las propiedades básicas del producto interno, tenemos*

$$0 \leq \langle \alpha x + y, \alpha x + y\rangle = \alpha^2\langle x, x\rangle + 2\langle x, y\rangle\alpha + \langle y, y\rangle$$

Si definimos $\langle x, x\rangle = a, 2\langle x, y\rangle = b$ *y* $\langle y, y\rangle = c$, *esta desigualdad se convierte en*

$$a\alpha^2 + b\alpha + c \geq 0 \quad \text{para todo } \alpha \in \mathbb{R}.$$

Sabemos (ya que $a > 0$) que esta expresión cuadrática tiene un mínimo para $\alpha = -b/2a$. Si sustituimos por este valor, obtenemos

$$a(\frac{b^2}{4a^2}) + b(-\frac{b}{2a}) + c \geq 0; \qquad \textit{es decir, } c \geq \frac{b^2}{4a}.$$

Como $a > 0$, esto significa que

$$b^2 \leq 4ac; \ \textit{es decir, } \ |\langle x, y\rangle|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle .$$

Tomando la raíz cuadrada obtenemos lo que queríamos. ■

Teorema 221 *Si $(\mathcal{V}, \langle\cdot,\cdot\rangle)$ es un espacio con producto interno y $\|\cdot\|$ está definida para cada $v \in \mathcal{V}$ por*

$$\|v\| = \sqrt{\langle v, v\rangle}$$

entonces $\|\cdot\|$ es una norma sobre $\mathcal{V}$.
Demostración: *La única propiedad no trivial es la desigualdad triangular, que se puede obtener como una consecuencia sencilla de la desigualdad de Cauchy-Schwarz:*

$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w\rangle = \|v\|^2 + 2\langle v, w\rangle + \|w\|^2 \leq \\ &\leq \|v\|^2 + 2|\langle v, w\rangle| + \|w\|^2 \leq \\ &\leq \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2, \end{aligned}$$

y, por lo tanto, $\|v+w\| \leq \|v\| + \|w\|$. ■

Ejemplo 284 *Probar que, para los números reales,*

a. $x \leq |x|, -|x| \leq x$.

b. $|x| \leq a$ *si y sólo si* $-a \leq x \leq a$, *donde* $a \geq 0$.

c. $|x+y| \leq |x| + |y|$.

Solución:

a. *Si $x \geq 0$, entonces $|x| = x$; si $x < 0$, entonces $|x| \geq x$, porque $x \geq 0$. En cualquier caso $x \leq |x|$. La otra afirmación es similar.*

b. *Si $x \geq 0$, entonces debemos demostrar que $0 \leq x \leq a$ si y sólo si $-a \leq x \leq a$, lo cual es evidente. De manera similar, si $x < 0$ la afirmación toma la forma $0 \leq -x \leq a$ si y sólo si $-a \leq x \leq a$, lo que también es evidente. En este caso hemos utilizado el hecho de que si $c \leq 0$, entonces $0 \leq x \leq y$ si y sólo si $0 \geq cx \geq cy$.*

c. *Por **a**, $-|x| \leq x \leq |x|$ y $-|y| \leq y \leq |y|$. Sumamos y obtenemos*

$$-(|x| + |y|) \leq x + y \leq |x| + |y| .$$

*Entonces, por **b**, $|x+y| \leq |x| + |y|$. Esto también se puede probar por casos, como se indica en el texto. Además, obsérvese que éste es un caso especial de la desigualdad triangular en $\mathbb{R}^n$.*

■

Ejemplo 285 *Para los números* $x_1, x_2, ..., x_n$, $y_1, y_2, ..., y_n$ *y* $z_1, z_2, ..., z_n$, *probar que*

$$(\sum_{i=1}^{n} x_i\, y_i\, z_i)^4 \leq (\sum_{i=1}^{n} x_i^4)(\sum_{i=1}^{n} y_i^2)^2(\sum_{i=1}^{n} z_i^4)$$

Solución: *La desigualdad de Cauchy-Schwarz dice que*

$$(\sum w_i y_i)^2 \leq (\sum w_i^2)(\sum y_i^2).$$

Aplicando esta desigualdad a los numeros $w_i = x_i\, z_i$ *e* y_i *se obtiene*

$$(\sum x_i\, y_i\, z_i)^2 \leq (\sum (x_i\, z_i)^2)(\sum y_i^2).$$

Aplicándolo de nuevo a x_i^2 *y* z_i^2 *se sigue que*

$$(\sum x_i^2\, z_i^2)^2 \leq (\sum x_i^4)(\sum z_i^4) \quad \text{es decir,} \; (\sum (x_i z_i)^2) \leq (\sum x_i^4)^{1/2}(\sum z_i^4)^{1/2}$$

Por tanto,

$$(\sum x_i\, y_i\, z_i)^2 \leq (\sum x_i^4)^{1/2}(\sum z_i^4)^{1/2}(\sum y_i^2).$$

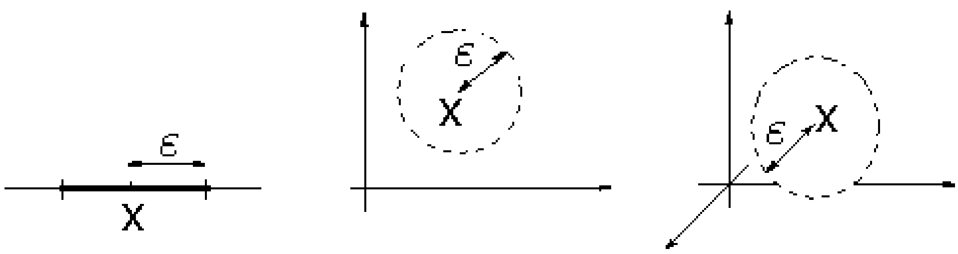
Elevando al cuadrado ambos miembros obtenemos el resultado. (Hemos utilizado el hecho de que si $a, b \geq 0$, *entonces* $a \leq b$ *si y sólo si* $a^2 \leq b^2$*).* ■

12.2. La topología del espacio euclídeo

Definición 106 *Sea* (M, d) *un espacio métrico. Para cada* $x \in M$ *y* $\varepsilon > 0$, *el conjunto*

$$D(x, \varepsilon) = \{y \in M \mid d(x, y) < \varepsilon\}$$

se denomina ***disco de radio*** ε *centrado en* x *(o también* ***bola de radio*** ε *centrada en* x*).*

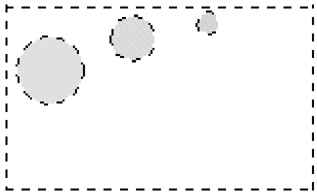


Los discos de radio ε *en* $\mathbb{R}$, $\mathbb{R}^2$, *y* $\mathbb{R}^3$ *con la distancia euclídea*

Definición 107 *Un conjunto* $A \subset M$ *es* ***abierto*** *si para cada* $x \in A$ *existe* $\varepsilon > 0$ *tal que* $D(x, \varepsilon) \subset A$. *Una* ***entorno*** *de un punto en* M *es un conjunto abierto que contiene dicho punto.*

Observación 123 *Obsérvese que el conjunto vacío* $\emptyset$ *y el conjunto total* M *son abiertos.*

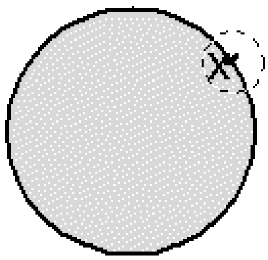
Observación 124 *Es importante observar que el valor ε de la definición de un conjunto abierto puede depender de x. Por ejemplo, el cuadrado unidad en $\mathbb{R}^2$ sin incluir su "frontera" es abierto, pero los entornos de radio ε son cada vez más pequeños al aproximarnos a la frontera; sin embargo, ε no puede anularse para ningún x.*



Un conjunto abierto

Consideremos un intervalo abierto en $\mathbb{R} = \mathbb{R}^1$, como el intervalo $(0,1)$. Éste es, de hecho, un conjunto abierto; sin embargo, si vemos este conjunto dentro de $\mathbb{R}^2$ (como un subconjunto del eje x), ya no es abierto. Así, para que un conjunto sea abierto, es esencial especificar el espacio $\mathbb{R}^n$ o, en general, el espacio métrico del que va a considerarse como subconjunto.

Hay muchos ejemplos de conjuntos que no son abiertos. El disco unidad $\left\{x \in \mathbb{R}^2 \mid \|x\| \leq 1\right\}$ es uno de estos ejemplos. Este conjunto no es abierto, ya que si un punto está en la "frontera" (es decir, un punto x tal que $\|x\| = 1$), todo disco de radio ε contiene puntos que no están en el conjunto:



Un conjunto no abierto

Teorema 222 *En un espacio métrico, todo disco de radio $\varepsilon, D(x,\varepsilon)$, es abierto.*
Demostración: *Sea $y \in D(x,\varepsilon)$. Debemos encontrar ε' tal que $D(y,\varepsilon') \subset D(x,\varepsilon)$. Sea $\varepsilon' = \varepsilon - d(x,y)$, que es estrictamente positivo, ya que $d(x,y) < \varepsilon$. Con esta elección, que depende de y, mostraremos que $D(y,\varepsilon') \subset D(x,\varepsilon)$. Sea $z \in D(y,\varepsilon')$, entonces $d(z,y) < \varepsilon'$. Es preciso demostrar que $d(z,x) < \varepsilon$. Pero, por la desigualdad triangular, $d(z,x) \leq d(z,y) + d(y,x) < \varepsilon' + d(y,x)$, y por la elección de ε', $\varepsilon' + d(y,x) = \varepsilon$.* ∎

A continuación damos algunas propiedades básicas de los conjuntos abiertos:

Teorema 223 *En un espacio métrico (M,d),*

i. *La intersección de un número finito de subconjuntos abiertos de M es abierta.*

ii. *La unión de una colección arbitraria de subconjuntos abiertos de M es abierta.*

iii. *El conjunto vacío $\emptyset$ y el espacio total M son abiertos.*

Demostración:

i. *Basta con demostrar que la intersección de dos conjuntos abiertos es abierta, puesto que podemos demostrar el resultado general por inducción, escribiendo $A_1 \cap \cdots \cap A_n = (A_1 \cap \cdots \cap A_{n-1}) \cap A_n$.*

Sean A y B conjuntos abiertos, y sea $C = A \cap B$; si $C = \emptyset$, C es abierto como un caso degenerado de la definición. Por tanto, podemos suponer que $C \neq \emptyset$, y que por tanto, existe x tal que $x \in C$. Como A y B son abiertos, existen $\varepsilon, \varepsilon' > 0$ tales que

$$D(x, \varepsilon) \subset A \qquad y \qquad D(x, \varepsilon') \subset B.$$

Sea ε'' el mínimo de ε y ε'. Entonces $D(x, \varepsilon'') \subset D(x, \varepsilon)$, y, por tanto, $D(x, \varepsilon'') \subset A$; de manera similar, $D(x, \varepsilon'') \subset B$, de modo que $D(x, \varepsilon'') \subset (A \cap B) = C$, como se pretendía.

ii. *La demostración para el caso de las uniones es más fácil. Sea $U, V, \ldots$ una colección de conjuntos abiertos cuya unión es A. Si $x \in A$, $x \in U$, para algún U en la colección. Por lo tanto, como U es abierto, $D(x, \varepsilon) \subset U \subset A$ para algún $\varepsilon > 0$, lo que demuestra que A es abierto.*

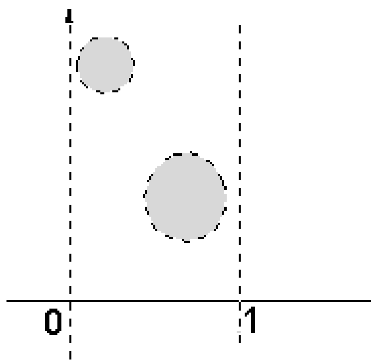
iii. *El conjunto vacío $\emptyset$ y el espacio total M son abiertos directamente por la definición.*

■

Observación 125 *Para apreciar la diferencia entre las afirmación **i** y **ii**, obsérvese que la intersección de una familia arbitraria de conjuntos abiertos no tiene por qué ser abierta. Por ejemplo, en $\mathbb{R}^1$, un punto (que no es un conjunto abierto) es la intersección de todos los conjuntos abiertos que lo contienen.*

Ejemplo 286 *Sea $S = \left\{(x, y) \in \mathbb{R}^2 \mid 0 < x < 1\right\}$. Demostrar que S es abierto.*

Solución: *En la figura vemos que podemos trazar el disco de radio $r = \min\{x, 1-x\}$ centrado en cada punto $(x, y) \in S$ y que éste está contenido completamente en S; puede verse esto en la figura y demostrarse mediante la desigualdad triangular. Por la definición, esto implica que S es abierto.* ■



Ejemplo 287 *Sea* $S = \left\{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1\right\}$. *¿Es* S *un conjunto abierto?*

Solución: *No, pues todo disco con centro en* $(1,0) \in S$ *contiene puntos* $(x,0)$ *tales que* $x > 1$ *y que por tanto, no pertenecen a* S. ■

Ejemplo 288 *Sean* $A \subset \mathbb{R}^n$ *un conjunto abierto y* $B \subset \mathbb{R}^n$. *Definamos*

$$A + B = \{x + y \in \mathbb{R}^n \mid x \in A \quad e\ y \in B\}.$$

Demostrar que $A + B$ *es abierto.*

Solución: *Sea* $w \in A + B$. *Entonces existen puntos* $x \in A$ *e* $y \in B$ *tales que* $w = x + y$. *Como* A *es abierto, existe* $\varepsilon > 0$ *tal que* $D(x,\varepsilon) \subset A$. *Veremos que* $D(w,\varepsilon) \subset A + B$. *Supongamos que* $z \in D(w,\varepsilon)$. *Entonces* $\|z - w\| = \|z - (x + y)\| < \varepsilon$. *Pero* $\|z - (x + y)\| = \|(z - y) - x\|$, *por lo que* $z - y \in D(x,\varepsilon) \subset A$. *Como* $y \in B$, *entonces* $(z - y) + y = z$ *está en* $A + B$. *Así,* $D(w,\varepsilon) \subset A + B$ *y entonces* $A + B$ *es un conjunto abierto.* ■

Observación 126 *Obsérvese que hemos demostrado en realidad que si* A *es abierto, entonces* $A + \{y\}$ *es abierto. Esto implica que* $A + B = \cup_{y \in B}(A + \{y\})$ *también es abierto.*

Ejemplo 289 *Sea* M *un conjunto arbitrario y definamos la* ***métrica discreta*** d_0 *en* M *mediante*

$$d_0(x,y) = \begin{cases} 0 & si \quad x = y \\ 1 & si \quad x \neq y \end{cases}.$$

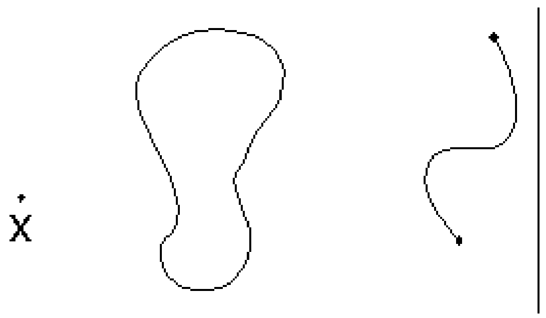
Demostrar que todo conjunto $A \subset M$ *es abierto.*

Solución: *Dado* $x \in A$, $D(x,1/2) = \{x\} \subset A$; *por definición, esto implica que* A *es abierto.* ■

Definición 108 *Un conjunto* B *en un espacio métrico* M *es* ***cerrado*** *si su complementario (es decir, el conjunto* $M \backslash B$*) es abierto.*

Por ejemplo, un punto en $\mathbb{R}^n$ es un conjunto cerrado. El conjunto en $\mathbb{R}^2$ formado por el disco unidad y su círculo frontera es cerrado. En pocas palabras, podemos decir que un conjunto es cerrado si contiene a sus "puntos frontera".

La intuición es difícil de utilizar para algunos conjuntos complicados, por lo que en esos casos se echa mano de la definición.



Ejemplos de conjuntos cerrados

Observación 127 *Es posible tener un conjunto que no sea abierto ni cerrado. Por ejemplo, en* $\mathbb{R}^1$*, un intervalo semiabierto* $(0,1]$ *no es abierto ni cerrado. Así, si sabemos que un conjunto* A *no es abierto, no podemos concluir que sea necesariamente cerrado.*

Teorema 224 *En un espacio métrico* (M,d)*,*

i. *La unión de un número finito de subconjuntos cerrados es cerrada.*

ii. *La intersección de una familia arbitraria de subconjuntos cerrados es cerrada.*

iii. *El espacio total* M *y el conjunto vacío* $\emptyset$ *son cerrados.*

Demostración:

i. *Basta con demostrar que la unión de dos conjuntos cerrados es cerrada, puesto que podemos demostrar el resultado general por inducción, escribiendo* $A_1 \cup \cdots \cup A_n = (A_1 \cup \cdots \cup A_{n-1}) \cup A_n$.

Sean A *y* B *conjuntos cerrados, y sea* $C = A \cup B$*. Veamos que* C *es cerrado, es decir, que su complementario es abierto:*

$$M \backslash C = M \backslash (A \cup B) = (M \backslash A) \cap (M \backslash B)$$

Sabemos que $M \backslash A$ *y* $M \backslash B$ *son abiertos, ya que* A *y* B *son cerrados por hipótesis. Ya vimos que la intersección finita de conjuntos abiertos es abierta, y por tanto,* C *es cerrado.*

ii. *Sea* $U, V, ...$ *una colección de conjuntos cerrados cuya intersección es* A*. Veamos que* A *es cerrado, es decir, que su complementario es abierto:*

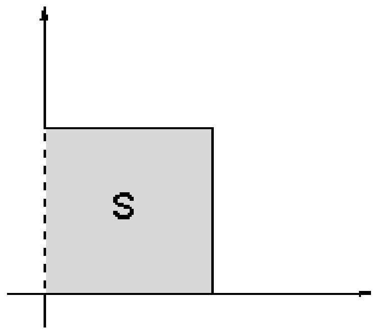
$$M \backslash A = M \backslash (U \cap V \cap \cdots) = (M \backslash U) \cup (M \backslash V) \cup \cdots$$

Sabemos que $M \backslash U$ *y* $M \backslash V$ *son abiertos (porque* $U, V, \ldots$ *son cerrados por hipótesis), y que la unión arbitraria de abiertos es abierta. Por tanto,* A *es cerrado.*

iii. *El conjunto vacío $\emptyset$ y el espacio total M son cerrados, porque son el complementario de M y $\emptyset$, respectivamente, y sabemos que M y $\emptyset$ son abiertos.*

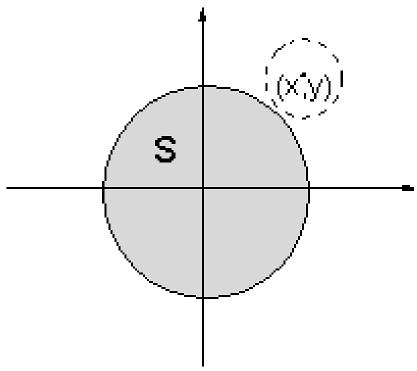
■

Ejemplo 290 *Sea $S = \{(x,y) \in \mathbb{R}^2 \mid 0 < x \leq 1, 0 \leq y \leq 1\}$. ¿Es S cerrado?*
Solución: *Véase la figura:*



Intuitivamente, S no es cerrado, pues la parte de su frontera que está en el eje y no está en S. Además, su complementario no es abierto, pues cualquier disco de radio ε con centro en un punto del eje y, como el punto $(0, 1/2)$, intersecará a S (y por lo tanto no estará contenido en $\mathbb{R}^2 \backslash S$). ■

Ejemplo 291 *Sea $S = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$. ¿Es S cerrado?*
Solución: *Sí, S es el disco unidad, con su frontera:*



El complementario es un conjunto abierto, pues para todo $(x,y) \in \mathbb{R}^2 \backslash S$, el disco de radio $\varepsilon = \sqrt{x^2+y^2} - 1$ está totalmente contenido en $\mathbb{R}^2 \backslash S$. ■

Ejemplo 292 *Demostrar que cualquier conjunto finito en $\mathbb{R}^n$ es cerrado.*
Solución: *Los puntos (individuales) son cerrados, por lo que la afirmación es una consecuencia de la teorema anterior.* ■

Ejemplo 293 *Sean (M,d) un espacio métrico y $A \subset M$ un conjunto finito. Sea*

$$B = \{x \in M \mid d(x,y) \leq 1 \ \ \text{para algún } y \in A\}.$$

Demostrar que B es cerrado.

Solución: *Veremos que $M\backslash B$ es abierto. Sea $z \in M\backslash B$, de modo que $d(z, y) > 1$ para todo $y \in A$. Sean $y_1, \ldots, y_N$ los puntos de A; entonces, $d(z, y_i) > 1$ para $i = 1, \ldots, N$. Sea*

$$\varepsilon = \text{mín}\{d(z, y_1) - 1, \ldots, d(z, y_N) - 1\}$$

Sabemos que $\varepsilon > 0$. Si $d(x, z) < \varepsilon/2$,la desigualdad triangular muestra que

$$d(x, y_i) + d(x, z) \geq d(y_i, z) \quad \Rightarrow \quad d(x, y_i) \geq d(y_i, z) - \varepsilon/2,$$

lo cual, por la construcción de ε, es estrictamente mayor que 1. Así, $D(z, \varepsilon) \subset M\backslash B$, luego $M\backslash B$ es abierto y B es cerrado. ■

Definición 109 *Sean M un espacio métrico y $A \subset M$. Un punto $x \in A$ es un* ***punto interior*** *de A si existe un conjunto abierto U tal que $x \in U \subset A$. El* ***interior de*** *A es la colección de todos los puntos interiores de A y se denota por $\mathit{int}\,(A)$. Este conjunto puede ser vacío.*

Observación 128 *La condición sobre x es equivalente a la siguiente: existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset A$. Por ejemplo, el interior de un único punto en $\mathbb{R}^n$ es vacío. El interior del disco unidad en $\mathbb{R}^2$ con su frontera, es el disco unidad sin su frontera.*

Podemos dar una descripción un poco diferente del interior de un conjunto como sigue. El interior de A es la unión de todos los subconjuntos abiertos de A. Por tanto, $\mathit{int}\,(A)$ es abierto. Por lo tanto, $\mathit{int}\,(A)$ es el mayor subconjunto abierto de A. En consecuencia, si A no tiene subconjuntos abiertos, entonces $\mathit{int}\,(A) = \emptyset$. Además, es evidente que A es abierto si y sólo si $\mathit{int}\,(A) = A$.

Ejemplo 294 *Sea $S = \left\{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq 1\right\}$. Determinar $\mathit{int}\,(S)$.*

Solución: *Para determinar los puntos interiores, buscamos los puntos para los cuales podemos trazar un disco de radio ε con centro en ellos y que esté totalmente contenido en S. Éstos son los puntos (x, y) tales que $0 < x < 1$. Así, $\mathit{int}\,(S) = \{(x, y) \mid 0 < x < 1\}$.* ■

Ejemplo 295 *¿Es cierto que $\mathit{int}\,(A) \cup \mathit{int}\,(B) = \mathit{int}\,(A \cup B)$?*

Solución: *No. Veamos un contraejemplo: consideremos los siguientes conjuntos en la recta real, sean $A = [0, 1]$ y $B = [1, 2]$. Entonces $\mathit{int}\,(A) = (0, 1)$ e $\mathit{int}\,(B) = (1, 2)$, de modo que $\mathit{int}\,(A) \cup \mathit{int}\,(B) = (0, 1) \cup (1, 2) = (0, 2)\backslash\{1\}$, mientras que $\mathit{int}\,(A \cup B) = \mathit{int}\,[0, 2] = (0, 2)$.* ■

Ejemplo 296 *¿Es cierto en un espacio métrico general (M, d) que*

$$\mathit{int}\,\{y \in M \mid d(y, x_0) \leq r\} = \{y \in M \mid d(y, x_0) < r\}$$

dados $x_0 \in M$ y $r > 0$?

Solución: *No. Por ejemplo, consideremos un conjunto M con la métrica discreta, $x_0 \in M$ y $r = 1$. Entonces $\{y \in M \mid d(y, x_0) \leq 1\} = M$, por lo que su interior es todo M. Por otro lado, $\{y \mid d(y, x_0) < 1\} = \{x_0\}$, que no es M si éste tiene más de un punto.* ■

Definición 110 *Un punto x en un espacio métrico M es un **punto de acumulación** de un conjunto $A \subset M$ si todo conjunto abierto U que contiene a x contiene también algún punto de A, distinto de x.*

En otras palabras, un punto de acumulación de un conjunto A es un punto para el cual hay puntos de A arbitrariamente cercanos a él. A los puntos de acumulación se les denomina a veces **puntos límite**. Podemos afirmar que x es un punto de acumulación de A si y sólo si para cada $\varepsilon > 0$, $D(x, \varepsilon)$ contiene algún punto $y \in A$ tal que $y \neq x$. Por ejemplo, en $\mathbb{R}^1$, un conjunto formado por un único punto no tiene puntos de acumulación, mientras que para el intervalo abierto $(0, 1)$ todos los puntos de $[0, 1]$ son puntos de acumulación. Obsérvese que un punto de acumulación de un conjunto no tiene por qué pertenecer a él.

Teorema 225 *Un conjunto $A \subset M$ es cerrado si y sólo si todos los puntos de acumulación de A pertenecen a A.*

Demostración: $\Rightarrow$) *En primer lugar, supongamos que A es cerrado. Entonces, $M \backslash A$ es abierto. Así, si $x \in M \backslash A$, existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \subset M \backslash A$; es decir, $D(x, \varepsilon) \cap A = \emptyset$. Así, x no es un punto de acumulación y entonces A contiene a todos sus puntos de acumulación.*

$\Leftarrow$) *Recíprocamente, supongamos que A contiene a todos sus puntos de acumulación. Sea $x \in M \backslash A$. Como x no es punto de acumulación de A y $x \notin A$, existe $\varepsilon > 0$ tal que $D(x, \varepsilon) \cap A = \emptyset$; es decir, $D(x, \varepsilon) \subset M \backslash A$. Por lo tanto, $M \backslash A$ es abierto y entonces A es cerrado.* ■

Observación 129 *Un conjunto no tiene por qué tener puntos de acumulación (un único punto y el conjunto de enteros en $\mathbb{R}^1$ son ejemplos de ello); pero entonces el teorema anterior se aplica trivialmente y de este modo podemos concluir que un conjunto de este tipo es cerrado.*

El teorema anterior es intuitivamente claro, pues la propiedad de ser cerrado significa, grosso modo, que un conjunto contiene todos los puntos de su "frontera", y tales puntos son puntos de acumulación. Este tipo de argumento es impreciso y tiene ciertas dificultades, por lo que debemos tener cuidado: algunos conjuntos son lo bastante complejos como para que falle nuestra intuición. Por ejemplo, considérese $A = \{1/n \in \mathbb{R} \mid n = 1, 2, 3, ...\} \cup \{0\}$. Éste es un conjunto cerrado y su único punto de acumulación es $\{0\}$, que está en A. Pero nuestra intuición de "frontera" no es clara para este conjunto; por lo tanto, necesitamos argumentos más cuidadosos.

Ejemplo 297 *Sea $S = \{x \in \mathbb{R} \mid x \in [0, 1] \text{ y } x \text{ es racional}\}$. Encontrar los puntos de acumulación de S.*

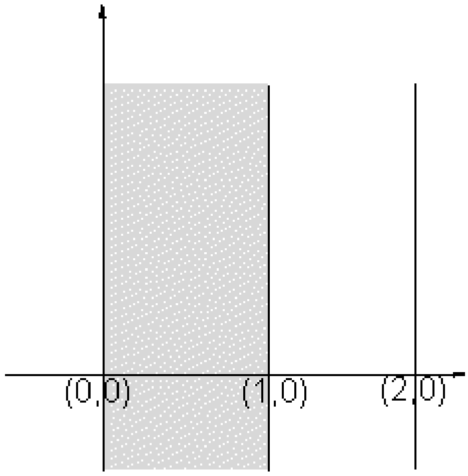
Solución: *El conjunto de puntos de acumulación está formado por todos los puntos de $[0, 1]$. De hecho, sea $y \in [0, 1]$ y sea $D(y, \varepsilon) = (y - \varepsilon, y + \varepsilon)$ un entorno de y.*

Podemos encontrar puntos racionales en $[0, 1]$ arbitrariamente cercanos a y (distintos de y), y en particular en $D(y, \varepsilon)$. Por lo tanto, y es un punto de acumulación. Ningún punto $y \notin [0, 1]$ es un punto de acumulación, pues existe un disco de radio ε centrado en y que lo contiene pero que no interseca a $[0, 1]$.
■

Ejemplo 298 *Verificar el teorema anterior para el conjunto*

$$A = \left\{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \ \text{ ó } \ x = 2\right\}$$

Solución: *El conjunto A aparece en la figura:*



Es evidente tanto que A es cerrado, como que el conjunto de puntos de acumulación de A es igual a A, de modo que cada punto de acumulación de A es un elemento de A. Obsérvese que en $\mathbb{R}$, el conjunto de puntos de acumulación de $[0,1] \cup \{2\}$ es $[0,1]$, sin el punto $\{2\}$. ■

Ejemplo 299 *Sea $S = \left\{(x, y) \in \mathbb{R}^2 \mid y < x^2 + 1\right\}$. Determínese los puntos de acumulación de S.*

Solución: *Los puntos de acumulación forman el conjunto $\left\{(x, y) \mid y \leq x^2 + 1\right\}$.* ■

Ejemplo 300 *En un espacio métrico general M, sea $B(x, r) = \{y \in M \mid d(x, y) \leq r\}$. ¿Es cierto que los puntos de $B(x, r)$ son puntos de acumulación de $D(x, r)$?*

Solución: *Esto no es cierto en general, aunque sí lo es en $\mathbb{R}^n$. Por ejemplo, si M es un conjunto con la métrica discreta, entonces $D(x, 1) = \{x\}$, mientras que $B(x, 1) = M$. El conjunto $\{x\}$ no tiene puntos de acumulación, pues está formado por un único punto, de modo que si M tiene más de un punto, $B(x, 1)$ contendrá puntos que no son puntos de acumulación de $D(x, 1)$.* ■

Ejemplo 301 *Demostrar que una sucesión acotada x_n de puntos distintos en $\mathbb{R}$ tiene un punto de acumulación; es decir, el conjunto $\{x_1, x_2, ...\}$ tiene un punto de acumulación.*

Solución: *Sabemos que para toda sucesión x_n, existe una subsucesión (que denotamos x_{n_k}) de x_n que converge: $x_{n_k} \to x$. Entonces, x es un punto de acumulación, puesto que para cada $\varepsilon > 0$, $D(x, \varepsilon)$ contiene a x_{n_k} con n_k suficiente grande (por definición de convergencia) y al menos uno de los términos x_{n_k} es distinto de x, ya que los x_{n_k} son todos distintos.* ■

Definición 111 *Sean (M, d) un espacio métrico y $A \subset M$. La* **clausura** *de A, denotada $cl(A)$, es la intersección de todos los conjuntos cerrados que contienen a A.*

Observación 130 *Puesto que la intersección de una familia arbitraria de conjuntos cerrados es cerrada,* $cl(A)$ *es cerrado; también es inmediato que* $cl(A) \supset A$ *y que* A *es cerrado si y sólo si* $cl(A) = A$*. Por ejemplo, en* $\mathbb{R}^1$, $cl((0,1)) = [0,1]$*.*

Teorema 226 *Si* $A \subset M$, $cl(A)$ *es el conjunto formado por* A *junto con los puntos de acumulación de* A*; es decir,* $cl(A) = A \cup \{\text{puntos de acumulación de } A\}$.
Demostración: *Sea* $B = A \cup \{x \mid x \text{ es un punto de acumulación de } A\}$*. Como un conjunto es cerrado si y sólo si contiene sus puntos de acumulación, cualquier conjunto cerrado que contenga a* A *también debe contener a* B*. Si* B *es cerrado, entonces será el menor conjunto cerrado que contiene a* A*, con lo que* $B = cl(A)$*. Para demostrar que* B *es cerrado, veremos que contiene todos sus puntos de acumulación: sea* y *un punto de acumulación de* B*. Si* $\varepsilon > 0$*, entonces* $D(y,\varepsilon)$ *contiene otros puntos de* B*. Si* z *es uno de tales puntos, entonces* $z \in A$ *o* z *es un punto de acumulación de* A*. En este último caso,* $D(z, \varepsilon - d(z,y))$ *es un conjunto abierto que contiene a* z *y, por definición de punto de acumulación, otros puntos de* A*, necesariamente distintos de* y*. Así,* y *es un punto de acumulación de* A *y entonces* $y \in B$*, por lo que* B *es cerrado.* ■

En otras palabras, para determinar la clausura de un conjunto A, añadimos a A todos los puntos de acumulación que no se encuentren ya en A.

Ejemplo 302 *Determinar la clausura de* $A = [0,1) \cup \{2\}$ *en* $\mathbb{R}$*.*

Solución: *Los puntos de acumulación son* $[0,1]$*, por lo que la clausura es* $[0,1] \cup \{2\}$*, que claramente es el menor conjunto cerrado que contiene a* A *que podríamos encontrar.* ■

Ejemplo 303 *Para cualquier* $A \subset \mathbb{R}^n$*, demostrar que* $\mathbb{R}^n \backslash cl(A)$ *es abierto.*

Solución: $cl(A)$ *es un conjunto cerrado; por definición de conjunto cerrado, su complementario es abierto.* ■

Ejemplo 304 *¿Es cierto que* $cl(A \cap B) = cl(A) \cap cl(B)$*?*

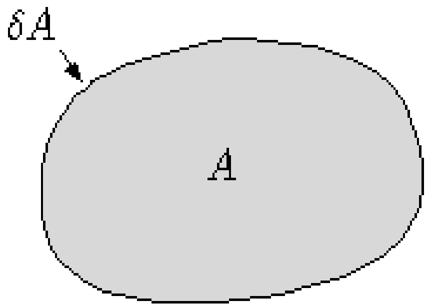
Solución: *No. Tomemos, por ejemplo,* $A = [0,1]$ *y* $B = (1,2]$*. Entonces* $A \cap B = \emptyset$ *y, por otro lado,* $cl(A) \cap cl(B) = \{1\}$. ■

Ejemplo 305 *Dado un subconjunto* A *de un espacio métrico, demostrar que* $x \in cl(A)$ *si y sólo si* $\inf \{d(x,y) \mid y \in A\} = 0$.

Solución: *En primer lugar, sean* $x \in cl(A)$ *y* $\alpha = \inf \{d(x,y) \mid y \in A\}$*. Si* $x \in A$*, entonces tomamos* $x = y$ *y obtenemos* $\alpha = 0$*. Si* x *es un punto de acumulación de* A*, entonces para cada* $\varepsilon > 0$ *existe* $y \in A$ *tal que* $d(x,y) < \varepsilon$*. Así, de nuevo,* $\alpha = 0$*. Recíprocamente, si* $\alpha = 0$ *y* $x \notin A$*, entonces para cada* $\varepsilon > 0$ *existe* $y \in A$ *tal que* $d(x,y) < \varepsilon$ *(por las propiedades del ínfimo), de modo que* x *es un punto de acumulación de* A *y entonces* $x \in cl(A)$*.* ■

Definición 112 *Para un conjunto dado* A *en un espacio métrico* (M,d)*, la* ***frontera*** *se define como el conjunto*

$$\partial(A) = cl(A) \cap cl(M \backslash A).$$



Observación 131 *Puesto que la intersección de dos conjuntos cerrados es cerrada, ∂A es un conjunto cerrado. Obsérvese también que $\partial A = \partial(M\backslash A)$.*

Teorema 227 *Sea $A \subset M$. Entonces $x \in \partial A$ si y sólo si para todo $\varepsilon > 0, D(x,\varepsilon)$ contiene puntos de A y de $M\backslash A$. (Estos puntos pueden incluir al propio x).*

Demostración: *Sea $x \in \partial A = cl(A) \cap cl(M\backslash A)$. Entonces, $x \in cl(A)$ y $x \in cl(M\backslash A)$. Por otro lado, o bien $x \in A$, o bien $x \in M\backslash A$. Si $x \in A$, por la caracterización de la clausura, la única opción para que $x \in cl(M\backslash A)$ es que x sea un punto de acumulación de $M\backslash A$ y por tanto, se cumple el resultado. El caso $x \in M\backslash A$ y el recíproco son similares.* ■

Observación 132 *La definición original establece que ∂A es la "línea divisoria" entre A y $M\backslash A$. Esto es también lo que afirma la teorema anterior, por lo que debe ser intuitivamente clara.*

Ejemplo 306 *Sea $A = \{x \in \mathbb{R} \mid x \in [0,1] \quad y\ x\ es\ racional\}$. Determinar ∂A.*

Solución: *$\partial A = [0,1]$, pues para todo $\varepsilon > 0$ y $x \in [0,1]$, $D(x,\varepsilon) = (x-\varepsilon, x+\varepsilon)$ contiene puntos racionales e irracionales.* ■

Observación 133 *El anterior ejemplo muestra que si $A \subset B$, esto no necesariamente implica que $\partial A \subset \partial B$. Tómese A como en este ejemplo y $B = [0,1]$ en $\mathbb{R}$.*

Ejemplo 307 *Si $x \in \partial A$, ¿debe x ser necesariamente un punto de acumulación?*

Solución: *No. Por ejemplo, sea $A = \{0\} \subset \mathbb{R}$. Entonces A no tiene puntos de acumulación, pero $\partial A = \{0\}$.* ■

Ejemplo 308 *Sea $S = \left\{(x,y) \in \mathbb{R}^2 \mid x^2 - y^2 > 1\right\}$. Determinar ∂S.*

Solución: *Claramente ∂S está formada por la hipérbola $x^2 - y^2 = 1$.* ■

Ejemplo 309 *Si $x = \sup(S)$ para $S \subset \mathbb{R}$, demostrar que $x \in cl(S)$.*

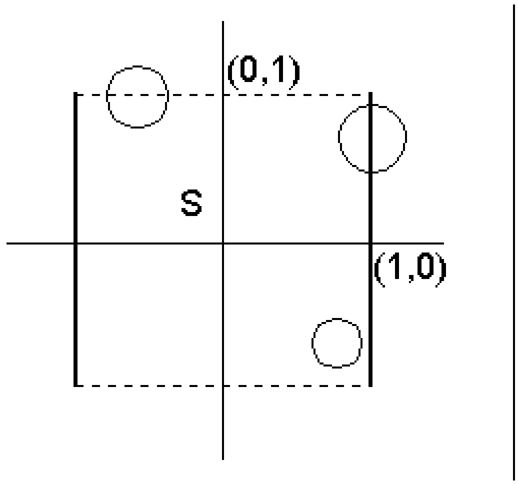
Solución: *Por la caracterización de la clausura, basta mostrar que, o bien $x \in S$, o bien x es un punto de acumulación de S. Como $x = \sup(S)$, para cualquier $\varepsilon > 0$ existe $y \in S$ tal que $d(x,y) < \varepsilon$. Esto significa que si $x \notin S$, entonces x es un punto de acumulación de S.* ■

Ejemplo 310 *Sea* $S = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \leq 1, |x_2| < 1\right\}$. *¿Es* S *abierto, cerrado o ninguna de las dos cosas? ¿Cuál es el interior de* S*?*

Solución: *S no es abierto, pues no existe ningún entorno con centro en los puntos de S con $x_1 = 1$ que esté totalmente contenido en S. Por otro lado, S no es cerrado, pues*

$$\mathbb{R}^2 \backslash S = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| > 1 \quad \acute{o} \quad |x_2| \geq 1\right\}$$

y no existe ningún entorno centrado en un punto de $\mathbb{R}^2\backslash S$ con $x_2 = 1$ que esté contenido en $\mathbb{R}^2\backslash S$.



También podemos ver que S no es cerrado si observamos que la sucesión $(0, 1 - 1/n)$, contenida en S, converge al punto $(0, 1)$, que no está en S.

Afirmamos que $int(S) = \left\{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| < 1, |x_2| < 1\right\}$, lo que verificamos mostrando que los elementos de este conjunto son los puntos interiores de S : Si $|x_1| < 1$ y $|x_2| < 1$, entonces el disco con centro en (x_1, x_2) y radio $r = \min\{1 - |x_1|, 1 - |x_2|\}$ está en S. Como ya hemos visto, los demás puntos de S no son puntos interiores. ∎

Ejemplo 311 *Demostrar que si x es un punto de acumulación de un conjunto $S \subset \mathbb{R}^n$, entonces todo conjunto abierto que contenga a x contiene a su vez infinitos puntos de S.*

Solución: *Usaremos el método de reducción al absurdo. Supongamos que existe un conjunto abierto U que contiene a x pero solamente un número finito de puntos de S. Sean $x_1, x_2, ...x_m$ los puntos de S contenidos en U distintos de x. Sea ε el mínimo de los números $d(x, x_1), d(x, x_2), ..., d(x, x_m)$, de modo que $\varepsilon > 0$. Entonces $D(x, \varepsilon)$ no contiene puntos de S distintos de x, lo que contradice el hecho de que x es un punto de acumulación de S.* ∎

12.3. Sucesiones

La definición de convergencia de una sucesión en $\mathbb{R}^n$ es similar a la de convergencia de una sucesión de números reales. De hecho, la definición tiene perfecto sentido en un espacio métrico.

Definición 113 *Sea (M,d) un espacio métrico y x_k una sucesión de puntos en M. Decimos que x_k* ***converge*** *a un punto $x \in M$ y escribimos*

$$\lim_{k\to\infty} x_k = x \qquad \text{ó} \qquad x_k \to x \ \text{ cuando } k \to \infty,$$

si para todo conjunto abierto U que contiene a x existe un entero N tal que $x_k \in U$ si $k \geq N$.

Esta definición coincide con la definición usual $\varepsilon - N$, como lo muestra el siguiente teorema.

Teorema 228 *Una sucesión x_k en M converge a $x \in M$ si y sólo si para todo $\varepsilon > 0$ existe N tal que si $k \geq N$ entonces $d(x, x_k) < \varepsilon$.*
Demostración: $\Rightarrow$) *Sea $\varepsilon > 0$, y supongamos que $x_k \to x$. Como $D(x,\varepsilon)$ es abierto, existe N tal que $k \geq N$ implica $x_k \in D(x,\varepsilon)$ o $d(x,x_k) < \varepsilon$, como se pedía.*

$\Leftarrow$) *Recíprocamente, supongamos que la condición es válida y U es un entorno de x. Sea $\varepsilon > 0$ tal que $D(x,\varepsilon) \subset U$; entonces, existe N tal que $k \geq N$ implica $d(x,x_k) < \varepsilon$; es decir, $x_k \in D(x,\varepsilon) \subset U$, por lo que $x_k \to x$, por definición.* ■

Particularicemos nuestro estudio de la convergencia al caso de $\mathbb{R}^n$, donde la definición de convergencia se lee: una sucesión de puntos v_k en $\mathbb{R}^n$ converge a v si para todo $\varepsilon > 0$ existe N tal que $\|v_k - v\| < \varepsilon$ si $k \geq N$. De hecho, esto se debe a que $d(v, v_k) = \|v - v_k\|$. Obsérvese además que para $n = 1$ recuperamos la definición de convergencia en $\mathbb{R}$. La definición de convergencia en cualquier otro espacio normado es esencialmente la misma: si $\mathcal{V}$ es un espacio normado y v_k es una sucesión en $\mathcal{V}$, entonces v_k **converge** a $v \in \mathcal{V}$ si $\|v_k - v\| \to 0$ cuando $k \to \infty$. Gran parte de lo que ya hemos hecho en el caso de dimensión 1 sigue siendo válido en $\mathbb{R}^n$, pero hay otra parte que no. Lo más destacable es que no existe un orden natural impuesto sobre $\mathbb{R}^n$, por lo que no podemos aplicar el estudio de las sucesiones monótonas a este caso general. La mayor parte del resto se obtiene reemplazando los valores absolutos por normas o distancias en todas partes. Por ejemplo, puesto que los espacios normados son espacios vectoriales, sabemos sumar vectores y multiplicarlos por números, por lo que podemos analizar la aritmética de las sucesiones.

Teorema 229 *Sean v_k y w_k sucesiones de vectores en un espacio normado (como por ejemplo $\mathbb{R}^n$), λ_k una sucesión de números en $\mathbb{R}$, $\lambda \in \mathbb{R}$ una constante y u un vector constante. Si $v_k \to v$, $w_k \to w$ y $\lambda_k \to \lambda$, entonces*

i. $v_k + w_k \to v + w$.

ii. $\lambda v_k \to \lambda v$.

iii. $\lambda_k u \to \lambda u$.

iv. $\lambda_k v_k \to \lambda v$.

v. *Si $\lambda_k \neq 0$ y $\lambda \neq 0$, entonces* $\frac{1}{\lambda_k} v_k \to \frac{1}{\lambda} v$.

Demostración:

i. *Veamos que* $v_k + w_k \to v + w$: *Sea* $\varepsilon > 0$, *y sean* N_1 *y* N_2 *tales que* $||v_k - v|| < \varepsilon/2$ *si* $k \geq N_1$, *y* $||w_k - w|| < \varepsilon/2$ *si* $k \geq N_2$. *Sea* $N = \max(N_1, N_2)$. *Entonces, para todo* $n \geq N$, *tenemos:*

$$\begin{aligned} ||(v_k + w_k) - (v + w)|| &= ||(v_k - v) - (w_k - w)|| \leq ||v_k - v|| + ||w_k - w|| < \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Por lo tanto, $v_k + w_k \to v + w$.

ii. *Éste es un caso especial de* ***iv*** *en el que* λ_k *es una sucesión constante.*

iii. *Éste es un caso especial de* ***iv*** *en el que* v_k *es una sucesión constante.*

iv. *Primero escribimos*

$$\begin{aligned} ||\lambda_k v_k - \lambda v|| &= ||\lambda_k v_k - \lambda_k v + \lambda_k v - \lambda v|| \leq ||\lambda_k v_k - \lambda_k v|| + ||\lambda_k v - \lambda v|| = \\ &= |\lambda_k| \cdot ||v_k - v|| + |\lambda_k - \lambda| \cdot ||v|| \end{aligned}$$

Como la sucesión λ_k *es una sucesión numérica convergente, está acotada, es decir, existe una constante* M *tal que* $|\lambda_k| < M$ *para todo* n. *Sean* N_1 *y* N_2 *tales que*

$$\begin{aligned} |\lambda_k - \lambda| &< \frac{\varepsilon}{2(||v|| + 1)} \quad \textit{cuando } n \geq N_1 \\ ||v_k - v|| &< \frac{\varepsilon}{2(M + 1)} \quad \textit{cuando } n \geq N_2 \end{aligned}$$

Sea $N = \max(N_1, N_2)$. *Si* $n \geq N$, *entonces*

$$\begin{aligned} ||\lambda_k v_k - \lambda v|| &\leq |\lambda_k| \cdot ||v_k - v|| + |\lambda_k - \lambda| \cdot ||v|| \leq \\ &\leq \frac{M\varepsilon}{2(M + 1)} + \frac{||v||\varepsilon}{2(||v|| + 1)} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

Por lo tanto, $\lambda_k v_k \to \lambda v$.

v. *Si* $\lambda \neq 0$, *entonces* $|\lambda| > 0$, *con lo cual existe* N_1 *tal que* $|\lambda_k - \lambda| < |\lambda|/2$ *cuando* $n \geq N_1$, *así que* $|\lambda_k| > |\lambda|/2$. *Para* $n \geq N_1$, *tenemos:*

$$\begin{aligned} \left\|\frac{v_k}{\lambda_k} - \frac{v}{\lambda}\right\| &= \left\|\frac{v_k\lambda - v\lambda_k}{\lambda_k\lambda}\right\| = \left\|\frac{(v_k - v)\lambda + v(\lambda - \lambda_k)}{\lambda_k\lambda}\right\| \leq \\ &\leq \frac{|\lambda| \cdot ||v_k - v||}{|\lambda_k\lambda|} + \frac{||v|| \cdot |\lambda - \lambda_k|}{|\lambda_k\lambda|} \leq \\ &\leq \frac{2}{|\lambda|} \cdot ||v_k - v|| + \frac{2||v||}{|\lambda|^2} \cdot |\lambda_k - \lambda|. \end{aligned}$$

Ahora elijamos N_2 *y* N_3 *de tal modo que* $||v_k - v|| < \varepsilon|\lambda|/4$ *siempre que* $n \geq N_2$, *y que* $|\lambda_k - \lambda| < \varepsilon|\lambda|^2/4||v||$ *siempre que* $n \geq N_3$. *Si* $N = \max(N_1, N_2, N_3)$ *y* $n \geq N$, *entonces*

$$\left|\frac{v_k}{\lambda_k} - \frac{v}{\lambda}\right| < \frac{2}{|\lambda|} \cdot \frac{\varepsilon|\lambda|}{4} + \frac{4||v||}{|\lambda|^2} \cdot \frac{\varepsilon|\lambda|^2}{4||v||} = \varepsilon$$

Por tanto,

$$\frac{v_k}{\lambda_k} \to \frac{v}{\lambda}.$$

$\blacksquare$

Para algunos cálculos emplearemos la representación específica de los vectores en $\mathbb{R}^n$ en términos de un número finito de coordenadas. Si v y v_k están en $\mathbb{R}^n$, entonces escribimos $v = (v^1, ..., v^n)$ y $v_k = (v_k^1, v_k^2, ..., v_k^n)$, donde $v^i, v_k^j \in \mathbb{R}$.

Teorema 230 *$v_k \to v$ en $\mathbb{R}^n$ si y sólo si cada sucesión de coordenadas converge a la coordenada correspondiente de v como una sucesión en $\mathbb{R}$. Es decir,*

$$\lim_{k\to\infty} v_k = v \quad \text{en } \mathbb{R}^n \iff \lim_{k\to\infty} v_k^i = v^i \quad \text{en } \mathbb{R} \quad \text{para todo} \quad i = 1, 2, \ldots, n.$$

Demostración: $\Rightarrow$) *Si definimos* $\delta(v, v_k) = \max\left\{\left|v^1 - v_k^1\right|, \ldots, |v^n - v_k^n|\right\}$ *entonces sabemos que*

$$\delta(v, v_k) \leq \|v - v_k\| \leq \sqrt{n} \cdot \delta(v, v_k).$$

Si $v_k \to v$ *en* $\mathbb{R}^n$, $1 \leq j \leq n$ *y* $\varepsilon > 0$, *entonces existe un entero* K *tal que* $\|v - v_k\| < \varepsilon$ *siempre que* $k \geq K$. *En consecuencia, para ese* k, *tenemos que*

$$\left|v^j - v_k^j\right| \leq \delta(v, v_k) \leq \|v - v_k\| < \varepsilon,$$

por lo que $v_k^j \to v^j$ *en* $\mathbb{R}$.

$\Leftarrow$) *La demostración de la recíproca muestra la importancia de que* n *sea finito. Supongamos que para cada* $j = 1, 2, \ldots, n$, *se cumple* $\lim_{k\to\infty} v_k^j = v^j$, *y sea* $\varepsilon > 0$. *Para todo* $\varepsilon_0 > 0$ *existen enteros* $K_1, K_2, K_3, \ldots, K_n$ *tales que*

$$\begin{array}{rcll} \left|v^1 - v_k^1\right| & < & \varepsilon_0 & \text{si} \;\; k \geq K_1, \\ \left|v^2 - v_k^2\right| & < & \varepsilon_0 & \text{si} \;\; k \geq K_2, \\ & & \ldots\ldots & \ldots\ldots \\ |v^n - v_k^n| & < & \varepsilon_0 & \text{si} \;\; k \geq K_n, \end{array}$$

Sólo hay un número finito de K_i, *luego uno de ellos es el máximo. Sea* $K = \max(K_1, K_2, \ldots, K_n)$. *Si* $k \geq K$, *entonces*

$$\|v - v_k\| \leq \sqrt{n} \cdot \delta(v, v_k) \leq \sqrt{n} \cdot \varepsilon_0.$$

Si hacemos esto con $\varepsilon_0 = \varepsilon/\sqrt{n}$, *obtenemos*

$$\|v - v_k\| < \varepsilon \quad \text{cuando} \quad k \geq K,$$

así que $v_k \to v$ *en* $\mathbb{R}^n$. $\blacksquare$

Observación 134 *Esto puede escribirse de una forma más compacta como*

$$\lim_{k\to\infty} (v_k^1, \ldots, v_k^n) = (\lim_{k\to\infty} v_k^1, \ldots, \lim_{k\to\infty} v_k^n).$$

Obsérvese que esta teorema no tiene sentido para un espacio métrico general.

Ejemplo 312 *Demostrar que la sucesión de vectores* $v_k = (1/k, 1/k^2)$ *converge a* $(0, 0)$ *en* $\mathbb{R}^2$ *cuando* $k \to \infty$.

Solución: *Cada una de las sucesiones componentes,* $1/k$ *y* $1/k^2$, *converge a* 0. *Por la teorema anterior, los vectores* $(1/k, 1/k^2)$ *convergen a* $(0, 0)$ *en* $\mathbb{R}^2$. $\blacksquare$

Las sucesiones pueden emplearse para determinar si un conjunto es cerrado. El método es el siguiente:

Teorema 231

i. *Un conjunto $A \subset B$ es cerrado si y sólo si para cada sucesión $x_k \in A$ que converge en M, el límite es un elemento de A.*

ii. *Para un conjunto $B \subset M$, $x \in cl(B)$ si y sólo si existe una sucesión $x_k \in B$ tal que $x_k \to x$.*

*Debemos hacer notar que las sucesiones en **i** e **ii** pueden ser triviales; es decir, se permite que $x_k = x$ para todo k.*

Demostración:

i. $\Rightarrow)$ *En primer lugar, supongamos que A es cerrado, y que $x_k \to x$. Entonces x es un punto de acumulación de A, ya que cualquier entorno de x contiene a $x_k \in A$ para k suficientemente grande. Como A es cerrado, entonces $x \in A$.*

$\Leftarrow)$ *Recíprocamente, para demostrar que A es cerrado veremos que todo punto de acumulación de A pertenece a A. Sean x un punto de acumulación de A y elijamos $x_k \in D(x, 1/k) \cap A$. Entonces $x_k \to x$, ya que para cada $\varepsilon > 0$ podemos elegir $N \geq 1/\varepsilon$, con lo que $k \geq N$ implica $x_k \in D(x, \varepsilon)$. Por hipótesis, $x \in A$, luego A es cerrado.*

ii. *El argumento en este caso es similar.*

■

Ejemplo 313 *Sea $x_n \in \mathbb{R}^m$ una sucesión convergente tal que $\|x_n\| \leq 1$ para todo n. Demostrar que el límite x también satisface $\|x\| \leq 1$. Si $\|x_n\| < 1$, ¿debe cumplirse entonces que $\|x\| < 1$?*

Solución: *La bola unidad $B = \{y \in \mathbb{R}^m \mid \|y\| \leq 1\}$ es cerrada. Por lo tanto, por la teorema anterior, $x_n \in B$ y $x_n \to x$ implican que $x \in B$. Esto no es cierto si reemplazamos $\leq$ por $<$; por ejemplo, considérese la sucesión $x_n = 1 - (1/n)$ en $\mathbb{R}$.* ■

Ejemplo 314 *Determinar la clausura de $A = \{1/n \in \mathbb{R} \mid n = 1, 2, \ldots\}$.*

Solución: *Podemos usar, por ejemplo, la teorema anterior. La sucesión $1/n \to 0$, por lo que $0 \in cl(A)$. Tomando otras sucesiones de A no se obtiene ningún punto nuevo, por lo que $cl(A) = A \cup \{0\}$.* ■

Ejemplo 315 *Demostrar que una sucesión de un espacio métrico puede converger, como mucho, a un punto (es decir, el límite de una sucesión es único).*

Solución: *Supongamos que $x_k \to x$ y $x_k \to y$. Dado $\varepsilon > 0$, sea N tal que $k \geq N$ implique que $d(x_k, y) < \varepsilon/2$ y sea M tal que $k \geq M$ implique que $d(x_k, y) < \varepsilon/2$. Si $k \geq N$ y $k \geq M$, por la desigualdad triangular tenemos:*

$$d(x, y) \leq d(x, x_k) + d(x_k, y) < \varepsilon$$

Como $0 \leq d(x, y) < \varepsilon$ se cumple para cada $\varepsilon > 0$, concluimos que $d(x, y) = 0$, por lo que $x = y$. ■

12.4. Completitud

En un espacio métrico general, o incluso en $\mathbb{R}^n$, los conceptos de sucesión monótona y supremo no tienen sentido. Sin embargo, la caracterización de la completitud por medio de las sucesiones de Cauchy sí lo tiene.

Definición 114 *Sea (M, d) un espacio métrico. Una* ***sucesión de Cauchy*** *es una sucesión $x_k \in M$ tal que para todo $\varepsilon > 0$, existe N tal que si $k, l \geq N$, entonces $d(x_k, x_l) < \varepsilon$.*

Observación 135 *En un espacio normado, como $\mathbb{R}^n$, una sucesión v_k es una sucesión de Cauchy si para todo $\varepsilon > 0$ existe N tal que $\|v_k - v_j\| < \varepsilon$ siempre que $k, j \geq N$.*

Definición 115 *Un espacio M es* **completo** *si y sólo si toda sucesión de Cauchy en M converge a un punto de M.*

Definición 116 *Una sucesión x_k en un espacio normado está* ***acotada*** *si existe un número C tal que $\|x_k\| \leq C$ para todo k. En un espacio métrico se requiere que exista un punto x_0 tal que $d(x_k, x_0) \leq C$ para todo k.*

Teorema 232 *Una sucesión convergente en un espacio normado o métrico está acotada.*
Demostración: *Si $x_k \to x$, existe N tal que $d(x_n, x) < 1$ cuando $n \geq N$. Así, $x_n \in D(x, 1)$ si $n \geq N$. Sea*

$$R = \max \{1, d(x_1, x), \ldots, d(x_{N-1}, x)\}.$$

Entonces $d(x, x_n) \leq R$ para todo n, con lo que $x_n \in D(x, R)$ para todo n. ■

Las siguentes son algunas propiedades generales de las sucesiones de Cauchy.

Teorema 233

i. *Toda sucesión convergente en un espacio métrico es una sucesión de Cauchy.*

ii. *Una sucesión de Cauchy en un espacio métrico debe estar acotada.*

iii. *Si una subsucesión de una sucesión de Cauchy converge a x, entonces la sucesión converge a x.*

Demostración:

i. *Si $x_n \to x$, para todo $\varepsilon > 0$ existe N tal que $d(x_n, x) < \varepsilon/2$ si $n \geq N$. Po lo tanto, si $n, m \geq N$,*

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Por tanto, (x_n) es de Cauchy.

ii. *Supongamos que* (x_n) *es una sucesión de Cauchy. Entonces existe* N *tal que*

$$d(x_n, x_k) < 1 \quad si \ \ n \geq N \ \ y \ k \geq K.$$

Por tanto,

$$||x_n|| \leq ||x_N|| + 1 \quad para \ \ n \geq N.$$

Si definimos $M = \text{máx}\{||x_1||, ||x_2||, \ldots, ||x_N||\} + 1$, *entonces* $||x_n|| \leq M$ *para todo* n.

iii. *Sea* (x_n) *una sucesión de Cauchy que tiene una subsucesión* (x_{n_k}) *que converge a* x. *Dado* $\varepsilon > 0$, *entonces existe* N *tal que* $d(x_n, x_m) < \varepsilon/2$ *cuando* $n \geq N$ *y* $m \geq N$. *Como* x_{n_k} *es una subsucesión convergente a* x, *existe* $n_0 > N$ *tal que* $d(x_{n_0}, x) < \varepsilon/2$. *Si* $m > N$, *entonces*

$$d(x_m, x) \leq d(x_m, x_{n_0}) + d(x_{n_0}, x) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Por tanto, x_n *converge a* x.

■

Teorema 234 ***(Completitud de*** $\mathbb{R}^n$***)*** *Una sucesión* x_k *en* $\mathbb{R}^n$ *converge a un punto en* $\mathbb{R}^n$ *si y sólo si es una sucesión de Cauchy.*
Demostración: $\Rightarrow$) *Si* x_k *converge a* x, *entonces para cada* $\varepsilon > 0$, *elegimos* N *tal que* $k \geq N$ *implique* $\|x_k - x\| < \varepsilon/2$. *En consecuencia, si* $k, l \geq N$, *entonces*

$$\|x_k - x_l\| = \|(x_k - x) + (x - x_l)\| \leq \|x_k - x\| + \|x - x_l\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Por tanto, x_k *es una sucesión de Cauchy.*

$\Leftarrow$) *Recíprocamente, sea* x_k *una sucesión de Cauchy. Como* $\left|x_k^i - x_l^i\right| \leq \|x_k - x_l\|$, *las componentes también son sucesiones de Cauchy en* $\mathbb{R}$. *Como toda sucesión de Cauchy en* $\mathbb{R}$ *es convergente, entonces cada sucesión* x_k^i *converge, digamos, a* x^i. *Así pues,* x_k *converge a* $x = (x^1, \ldots, x^n)$. ■

Teorema 235 ***(Teorema de los intervalo encajados)*** *Sea* (F_k) *una sucesión de conjuntos no vacíos en* $\mathbb{R}^p$ *encajados, es decir* $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$. *Supongamos que todos los* F_n, $n \geq 1$, *son cerrados y que* F_1 *es además acotado. Si*

$$diam\,(F_n) = \sup\{d(x, y) : x, y \in F_n\} \longrightarrow 0 \qquad cuando \ \ n \to \infty,$$

entonces $\bigcap_{n\geq 1} F_n$ *contiene un único punto.*
Demostración: *Para cada* $n \geq 1$, *escojamos* $x_n \in F_n$. *Si* $m \geq n$, *entonces*

$$d(x_n, x_m) \leq diam\,(F_n) \longrightarrow 0 \qquad cuando \ \ n \to \infty.$$

Por tanto, (x_n) *es una sucesión de Cauchy, y como* $\mathbb{R}^p$ *es completo, sabemos que* (x_n) *converge a algún* $x \in \mathbb{R}^p$. *Como* $(x_n) \subset F_1$ *para todo* $n \geq 1$, *entonces* $x \in F_1$, *ya que* F_1 *es cerrado. Análogamente,* $x \in F_m$, *ya que* $(x_n) \subset F_m$ *para todo* $n \geq m$ *y* F_m *es cerrado. Por tanto,*

$$x \in \bigcap_{n\geq 1} F_n \quad \Rightarrow \quad \bigcap_{n\geq 1} F_n \neq \emptyset.$$

Como diam $(\bigcap_{n\geq 1} F_n) = 0$, *entonces* $\bigcap_{n\geq 1} F_n$ *sólo contiene un único elemento, tal como queríamos ver.* ■

Teorema 236 ***(Teorema de Bolzano-Weierstrass)***

a. *Si* $A \subset \mathbb{R}^p$*, A acotado e infinito, entonces A contiene algún punto de acumulación.*

b. *Si* $(x_n) \subset \mathbb{R}^p$ *es una sucesión acotada, entonces tiene una subsucesión convergente.*

Demostración:

a. *Si* $A \subset \mathbb{R}^p$ *es acotado, entonces existe* r *tal que* A *está incluido en el rectángulo* $[-r, r] \times [-r, r] \times \cdots \times [-r, r]$*. Sea* $F_1 = [-r, r] \times [-r, r] \times \cdots \times [-r, r]$*. Subdividimos* F_1 *en* 2^p *rectángulos bisecando cada uno de sus lados. Como* F_1 *contiene una infinidad de puntos de* A*, al menos una de las partes obtenidas en esta subdivisión también contendrá una infinidad de puntos de* A*. Sea* F_2 *una de estas partes de la subdivisión de* F_1 *que contiene una infinidad de elementos de* A*. Continuamos este proceso: dado* F_n*, subdividimos* F_n *en* 2^p *rectángulos iguales bisecando sus lados, y escogemos como* F_{n+1} *una de las partes de la subdivisión que contenga una infinidad de puntos de* A*. Así obtenemos una sucesión de rectángulos cerrados encajados:*

$$F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

Además,

$$\begin{aligned} diam\,(F_n) &= \frac{1}{2^p} \cdot diam\,(F_{n-1}) = (\frac{1}{2^p})^2 \cdot diam\,(F_{n-2}) \\ &= \cdots = (\frac{1}{2^p})^{n-1} \cdot diam\,(F_1) \longrightarrow 0 \end{aligned}$$

Por el teorema de los intervalos encajados, existe un único punto x *tal que* $x \in \bigcap_{n\geq 1} F_n$*. Demostraremos que* x *es un punto de acumulación de* A*, viendo que cualquier bola centrada en* x *contiene algún punto de* A *diferente de* x*: para todo* $r > 0$*, existe* k *tal que* $diam\,(F_k) < r$*. Por tanto,*

$$F_k \subset D(x, r) \quad \Rightarrow \quad ((F_k \backslash \{x\}) \cap A) \subset ((D(x, r) \backslash \{x\}) \cap A)$$

Como $(F_k \backslash \{x\}) \cap A \neq \emptyset$*, entonces* $(D(x, r) \backslash \{x\}) \cap A \neq \emptyset$*.*

b. *Si* (x_n) *es finita, algún elemento ha de aparecer infinitas veces en* (x_n)*. La subsucesión constante igual a este elemento es convergente.*

Si (x_n) *es infinita, como es acotada por hipótesis, por el apartado* **a** *sabemos que existe un punto de acumulación de* (x_n)*. Para cada* k *escogemos un índice* n_k *tal que* $x_{n_k} \in D(x, \frac{1}{k})$ *y* $n_k > n_{k-1}$*. Entonces la subsucesión* $(x_{n_k})_{k\in\mathbb{N}}$ *converge a* x*.*

■

Definición 117 *Un punto* x *en un espacio métrico es un* ***punto límite*** *de la sucesión* x_k *si para todo* $\varepsilon > 0$ *hay infinitos valores de* k *tales que* $d(x_k, x) < \varepsilon$*.*

Teorema 237 *Si x_k es una sucesión en un espacio métrico M y $x \in M$, entonces*

i. *x es un punto límite si y sólo si para cada $\varepsilon > 0$ y cada entero N, existe $k > N$ tal que $\mathrm{d}(x_k, x) < \varepsilon$.*

ii. *x es un punto límite si y sólo si existe una subsucesión convergente a x.*

iii. *$x_k \to x$ si y sólo si toda subsucesión converge a x.*

iv. *$x_k \to x$ si y sólo si toda subsucesión de x_k tiene a su vez una subsucesión que converge a x.*

Demostración:

i. $\Rightarrow$) *Si x es un punto límite de x_n y $\varepsilon > 0$, entonces hay infinitos índices para los cuales $\mathrm{d}(x_n, x) < \varepsilon$. Sólo un número finito de ellos es menor que N, y por tanto al menos uno de ellos es mayor que N, para cualquier N.*

$\Leftarrow$) *Podemos escoger n_1 tal que $\mathrm{d}(x_{n_1}, x) < \varepsilon$, $n_2 > n_1$ tal que $\mathrm{d}(x_{n_2}, x) < \varepsilon$, $n_3 > n_2$ tal que $\mathrm{d}(x_{n_3}, x) < \varepsilon$, y así sucesivamente. Obtenemos $n_1 < n_2 < n_3 < \cdots$ tales que $\mathrm{d}(x_{n_k}, x) < \varepsilon$, y por lo tanto, x es un punto límite.*

ii. $\Rightarrow$) *Si x es un punto límite, seleccionamos el índice n_1 de modo que $\mathrm{d}(x_{n_1}, x) < 1$. Usando el apartado **i**, elegimos $n_2 > n_1$ tal que $\mathrm{d}(x_{n_2}, x) < \frac{1}{2}$. Usando **i** reiteradamente, podemos ir eligiendo los índices $n_1 < n_2 < n_3 < \cdots$ de modo que $\mathrm{d}(x_{n_k}, x) < 1/k$. Esto produce una subsucesión que converge a x, ya que si $\varepsilon > 0$, sabemos que existe un entero K tal que $1/K < \varepsilon$; entonces, $k \geq K$ implica*

$$d(x_{n_k}, x) < \frac{1}{k} \leq \frac{1}{K} < \varepsilon.$$

$\Leftarrow$) *Si hay una subsucesión convergente a x y $\varepsilon > 0$, entonces hay un índice tal que los índices posteriores a éste en la subsucesión distan menos que ε de x. Por lo tanto, x es un punto límite.*

iii. $\Rightarrow$) *Si $x_n \to x$ y x_{n_k} es una subsucesión, sea $\varepsilon > 0$ y elijamos N de tal modo que $\mathrm{d}(x_n, x) < \varepsilon$ para $n \geq N$. Como $n_1 < n_2 < n_3 < \cdots$, por inducción se demuestra que $n_k \geq k$. Por lo tanto, $k \geq N$ implica $n_k \geq N$, y, en consecuencia, $\mathrm{d}(x_{n_k}, x) < \varepsilon$, con lo cual la subsucesión converge a x.*

$\Leftarrow$) *Demostraremos este apartado por el contrarecíproco: si x_n fuera una sucesión que no convergiera a x, entonces por el apartado **i** existiría $\varepsilon > 0$ tal que para cada N habría un índice mayor que N tal que $\mathrm{d}(x_n, x) > \varepsilon$. Usando este hecho repetidamente, podemos seleccionar $n_1 < n_2 < n_3 < \cdots$ tales que $\mathrm{d}(x_{n_k}, x) > \varepsilon$. Esto nos da una subsucesión que no converge a x.*

iv. $\Rightarrow$) *Si $x_k \to x$, por el apartado **iii**, toda subsucesión converge a x, por lo cual no necesitamos pasar a una sub-subsucesión.*

$\Leftarrow$) *Si toda subsucesión tiene una sub-subsucesión convergente a x, entonces x es un punto límite de toda subsucesión y por lo tanto de x_n. Si existiera otro punto límite $y \neq x$, entonces, por **ii**, habría una subsucesión convergente a él. Sería el único punto límite de esa subsucesión y, en consecuencia, no podría tener una sub-subsucesión convergente a x. Así pues, x debe ser el punto límite de x_n, con lo cual $x_n \to x$.*

■

Ejemplo 316 *Sea (M, d) un espacio métrico completo y $N \subset M$ un subconjunto cerrado. Demostrar que N también es completo.*

Solución: *Se entiende que la métrica utilizada en N es la misma que la de M. Sea x_n una sucesión de Cauchy en N, la cual, considerada como una sucesión en M, sigue siendo una sucesión de Cauchy. Como M es completo, esta sucesión converge en M a un punto x. Como $x_n \in N$, $x_n \to x$ y N es cerrado, $x \in N$. Así, x_n converge en N y, por tanto, N es completo.* ■

Ejemplo 317 *Demostrar que el conjunto de puntos límite de una sucesión x_k es cerrado.*

Solución: *Sea y_n una sucesión de puntos límite tales que $y_n \to y$. Para cada $\varepsilon > 0$, existe N tal que $d(y_n, y) < \varepsilon/2$ si $n \geq N$. Hay infinitos puntos x_k tales que $d(x_k, y_N) < \varepsilon/2$, pues y_N es un punto límite. La desigualdad triangular implica que*

$$d(x_k, y) \leq d(x_k, y_N) + d(y_N, y) < \varepsilon/2 + \varepsilon/2 = \varepsilon,$$

por lo que y es también un punto límite de x_k. ■

12.5. Conjuntos compactos

Definición 118 *Sea (M, d) un espacio métrico. Se dice que $\mathcal{U} = \{U_i, i \in I\}$ es un **recubrimiento abierto** de $A \subset M$ si U_i es abierto para todo $i \in I$ y $A \subset \bigcup_{i \in I} U_i$. Una subcolección finita de $\mathcal{U}$ que sea un recubrimiento de A se llama **subrecubrimiento finito**.*

Definición 119 *Un subconjunto $K \subset M$ es compacto de si de todo recubrimiento abierto de K se puede extraer un subrecubrimiento finito.*

Ejemplo 318 *Una colección finita de puntos es un conjunto compacto. Todo recubirmiento abierto $\mathcal{U}$ tiene un subrecubrimiento finito, que se puede obtener asignando a cada punto un elemento de $\mathcal{U}$ en el que se encuentre.*

Ejemplo 319 *$\mathbb{R}^p$ no es compacto, ya que el siguiente recubrimiento abierto*

$$\mathcal{U} = \{D(0, n), \quad n \in \mathbb{N}\}$$

no admite un subrecubrimiento finito.

Ejemplo 320 *$(a, b) \subset \mathbb{R}$ no es compacto, ya que el siguiente recubrimiento abierto*

$$\mathcal{U} = \{(a + \frac{1}{n}, b - \frac{1}{n}), \quad n \in \mathbb{N}\}$$

no admite un subrecubrimiento finito.

Ejemplo 321 *El conjunto formado por la sucesión* $(\frac{1}{n})$ *y 0 es compacto. Cualquiera que sea el recubrimiento abierto* $\mathcal{U}$ *del conjunto, si* $U \in \mathcal{U}$ *contiene a 0, teniendo en cuenta la definición de límite de una sucesión, fuera de* U *sólo existe un número finito de elementos de la sucesión, los cuales son recubiertos por un nombre finito de elementos de* $\mathcal{U}$*. Estos últimos junto con* U *forman un subrecubrimiento finito del conjunto.*

Teorema 238 *Todo conjunto* $K \subset M$ *compacto es cerrado y acotado.*
Demostración: *Sea* $K \subset M$ *compacto. Veamos en primer lugar que* K *es cerrado o, equivalentemente, que* $M \backslash K$ *es abierto. Sea* $x \in M \backslash K$*. Encontraremos un entorno* V *de* x *contenido en* $M \backslash K$*. Por las propiedades de los entornos, para cada punto* $y \in K$*, existe un entorno abierto* V_y *de* x *y un entorno abierto* W_y *de* y*, tales que* $V_y \cap W_y = \emptyset$*. La familia* $\{W_y : y \in K\}$ *es un recubrimiento abierto de* K*. Por ser* K *compacto, existe un subrecubrimiento finito* $W_1, W_2, \ldots, W_p$*. Si* $V_1, V_2, \ldots, V_p$ *son los entornos de* x *correspondientes, entonces* $V = \bigcap_{j=1}^{p} V_j$ *está contenido en* $M \backslash K$*; es decir,* $V \cap K = \emptyset$*, ya que*

$$V \cap K \subset V \cap (\bigcup_{j=1}^{p} W_j) = \bigcup_{j=1}^{p} (V \cap W_j) = \emptyset.$$

Veamos ahora que K *es acotado. En efecto, la familia de discos abiertos*

$$\{D(a,n) : n \in \mathbb{N}\} \quad \textit{de centro } a \in M$$

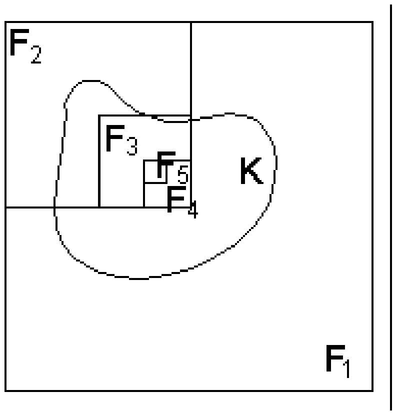
es un recubrimiento abierto de K*. Como* K *es compacto, existe un número finito de discos que recubre* K*. Por tanto,* K *está contenido en el disco de mayor radio y por consiguiente,* K *es acotado.* ∎

Teorema 239 ***(Teorema de Heine-Borel)*** *En* $\mathbb{R}^p$*, un conjunto* K *es compacto si y sólo si es cerrado y acotado.*
Demostración: $\Rightarrow$) *Este resultado ha sido demostrado en la teorema anterior.*

$\Leftarrow$) *Lo demostraremos por reducción al absurdo: sea* K *un conjunto cerrado y acotado pero no compacto. Por lo tanto, existe un recubrimiento abierto* $\mathcal{U}$ *de* K *que no admite un subrecubrimiento finito. Como* K *es acotado, existe un número* r *tal que*

$$K \subset [-r, r] \times [-r, r] \times \cdots \times [-r, r]$$

Sea $F_1 = [-r, r] \times [-r, r] \times \cdots \times [-r, r]$*. Si bisecamos los lados de* F_1*, obtenemos* 2^p *rectángulos. Sea* F_2 *cualquiera de los rectángulos de esta subdivisión tal que el conjunto* $K \cap F_2$ *no esté contenido en la unión de cualquier número finito de conjuntos en* $\mathcal{U}$ *(sabemos que* F_2 *existe, porque si cada una de las* 2^p *partes de* K *estuvieran contenidas en un número finito de los conjuntos en* $\mathcal{U}$*, entonces* K *estaría contenido en la unión de un número finito de conjuntos de* $\mathcal{U}$*, lo que contradiría la hipótesis). Se continúa este proceso bisecando los lados de* F_2 *para obtener* 2^p *rectángulos y se toma* F_3 *como uno de los rectángulos tal que el conjunto no vacío* $K \cap F_3$ *no esté contenido en la unión de un número finito de conjuntos de* $\mathcal{U}$*, y así sucesivamente.*



De esta manera se obtiene una sucesión de rectángulos cerrados y encajados:

$$F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

Además,

$$diam\,(F_n) = \frac{1}{2^p} \cdot diam\,(F_{n-1}) = \cdots = (\frac{1}{2^p})^{n-1} \cdot diam\,(F_1) \longrightarrow 0 \quad cuando\ n \to \infty.$$

Por el teorema de los intervalos encajados, existe un punto x *tal que* $\bigcap_{n\geq 1} F_n = \{x\}$, *y* x *es un punto de acumulación de* K. *Como* K *es cerrado,* $x \in \bar{K}$, *luego existe algún conjunto abierto* $V \in \mathcal{U}$ *tal que* $x \in V$. *Por tanto, existe* $r > 0$ *tal que* $D(x, r) \subset V$. *Existe* n *tal que* $diam\,(F_n) < r$. *Por consiguiente,*

$$F_n \subset D(x, r) \subset V,$$

lo que contradice la elección de F_n. *Por tanto, no puede existir ningún recubrimiento* $\mathcal{U}$ *de* K *por abiertos que no admita un subrecubrimiento finito.* ■

12.6. Conjuntos conexos

Definición 120 *Se dice que un subconjunto* $D \subset \mathbb{R}^p$ *es* ***conexo*** *si no existen dos conjuntos abiertos* U *i* V *tales que* $V \cap D$ *y* $U \cap D$ *sean disjuntos, no vacíos y* $D \subset (U \cup V)$. *En caso de que existan* U *y* V *con las mencionadas condiciones, se dice que forman una* ***separación*** *de* D.

Ejemplo 322 *El conjunto* $\mathbb{N} \subset \mathbb{R}$ *no es conexo, ya que se puede tomar*

$$\begin{aligned} U &= \{x \in \mathbb{R} : x < 3/2\} \\ V &= \{x \in \mathbb{R} : x > 3/2\}. \end{aligned}$$

Ejemplo 323 *El conjunto* $\mathbb{Q} \subset \mathbb{R}$ *no es conexo, ya que se puede tomar*

$$\begin{aligned} U &= \{x \in \mathbb{R} : x < \sqrt{2}\} \\ V &= \{x \in \mathbb{R} : x > \sqrt{2}\}. \end{aligned}$$

Teorema 240 *El intervalo unitario cerrado* $\mathbf{I} = [0, 1]$ *es un suconjunto conexo de* $\mathbb{R}$.
Demostración: *Lo demostraremos por reducción al absurdo, suponiendo que existen dos conjuntos abiertos* U *y* V *que forman una separación de* $\mathbf{I}$. *Por tanto,* $U \cap \mathbf{I}$ *y* $V \cap \mathbf{I}$ *son conjuntos disjuntos no vacíos acotados tales que* $\mathbf{I} \subset (U \cup V)$. *Dado que los conjuntos* U *y* V *son abiertos, los conjuntos* $U \cap \mathbf{I}$ *y* $V \cap \mathbf{I}$ *no pueden constar de un solo punto. Para ser más precisos, supongamos que existen* $a \in U, b \in V$ *tales que* $0 < a < b < 1$. *Consideremos el conjunto* $\{x \in U : x < b\}$. *Es un conjunto no vacío, porque al menos a pertenece a él. Como es un conjunto acotado, tiene supremo. Sea*

$$c = \sup\{x \in U : x < b\}.$$

Por tanto, $0 < c < 1$, *por lo que* $c \in U \cup V$. *Si* $c \in U$, *entonces* $c \neq b$, *y como* U *es abierto, existe un punto* $a_1 \in U$, $c < a_1$, *tal que el intervalo* $[c, a_1]$ *está contenido en* $\{x \in U : x < b\}$, *contrario a la definición de c. Análogamente, si* $c \in V$, *entonces, puesto que* V *es abierto, hay un punto* $b_1 \in V$, $b_1 < c$, *tal que el intervalo* $[b_1, c]$ *está contenido en* $V \cap \mathbf{I}$, *contrario a la definición de c. Por tanto, la hipótesis de que* $\mathbf{I}$ *no es conexo lleva a contradicción.* ■

Observación 136 *El intervalo abierto* $(0, 1)$ *es conexo en* $\mathbb{R}$, *y se puede probar con la misma demostración que el teorema anterior.*

Teorema 241 *El espacio* $\mathbb{R}^p$ *es conexo.*
Demostración: *Supongamos que* $\mathbb{R}^p$ *no es conexo, es decir, supongamos que existen dos conjuntos* U *y* V *abiertos, disjuntos, y no vacíos tales que* $\mathbb{R}^p = U \cup V$. *Sean* $x \in U$ *e* $y \in V$. *Consideremos el segmento de línea* S *que une* x *con* y, *es decir,*

$$S = \{x + t \cdot (y - x) : t \in \mathbf{I}\}.$$

Sean

$$\begin{aligned} U_1 &= \{t \in \mathbb{R} : x + t \cdot (y - x) \in U\} \\ V_1 &= \{t \in \mathbb{R} : x + t \cdot (y - x) \in V\} \end{aligned}$$

U_1 *y* V_1 *son subconjuntos abiertos disjuntos no vacíos de* $\mathbb{R}$, *y constituyen una separación de* $\mathbf{I}$, *lo que contradice el teorema anterior.* ■

Teorema 242 *Los únicos subconjuntos de* $\mathbb{R}^p$ *que son tanto abiertos como cerrados son* $\emptyset$ *y* $\mathbb{R}^p$.
Demostración: *Supongamos que existe* $U \subset \mathbb{R}^p$ *que es tanto abierto como cerrado en* $\mathbb{R}^p$. *Entonces* $V = \mathbb{R}^p \backslash U$ *también lo es. Si* U *es no vacío y no es todo* $\mathbb{R}^p$, *entonces el par* U, V *constituye una separación de* $\mathbb{R}^p$, *lo que contradice el teorema anterior. Por tanto, los únicos subconjuntos de* $\mathbb{R}^p$ *abiertos y cerrados a la vez sólo pueden ser* $\emptyset$ *y* $\mathbb{R}^p$. ■

Capítulo 13

Funciones continuas

13.1. Propiedades locales de funciones continuas

Se habrá de considerar a f como una función con dominio $D(f)$ contenido en $\mathbb{R}^p$ y con imagen $\mathrm{Im}(f)$ contenida en $\mathbb{R}^q$. En general, no se requerirá que $D(f) = \mathbb{R}^p$ o que $p = q$. Primero se definirá continuidad en términos de entornos y después se mencionarán algunas condiciones equivalentes.

Definición 121 *Si $a \in D(f)$, se dice que f es **continua en** a si para todo entorno V de $f(a)$ existe un entorno U (dependiendo de V) de a tal que si x es cualquier elemento de $V \cap D(f)$, entonces $f(x)$ es un elemento de V. Si $A \subseteq D(f)$, se dice que f es **continua en** A siempre que sea continua en todo punto de A.*

A continuación se dan dos afirmaciones equivalentes que se podían haber utilizado como definición.

Teorema 243 *Sea a un punto en el dominio $D(f)$ de al función f. Las siguientes afirmaciones son equivalentes:*

a. *f es continua en a.*

b. *Si $\varepsilon > 0$, existe un número $\delta(\varepsilon) > 0$ tal que si $x \in D(f)$ y $\|x - a\| < \delta(\varepsilon)$, entonces $\|f(x) - f(a)\| < \varepsilon$.*

c. *Si (x_n) es cualquier sucesión de elementos de $D(f)$ que converge a a, entonces la sucesión $(f(x_n))$ converge a $f(a)$.*

Demostración: *$a \Rightarrow b$)   Supongamos que (a) es cierto y que $\varepsilon > 0$, entonces la bola $V_\varepsilon = \{y \in \mathbf{R}^q : \|y - f(a)\| < \varepsilon\}$ es un entorno del punto $f(a)$. Por la definición de continuidad, hay un entorno U de a tal que si $x \in U \cap D(f)$, entonces $f(x) \in V_\varepsilon$. Dado que U es un entorno de a, existe un número real positivo $\delta(\varepsilon)$ tal que la bola abierta con radio $\delta(\varepsilon)$ y centro a está contenida en U. Por lo tanto, la condición (a) implica (b).*

$b \Rightarrow c$)   Supongamos que (b) es cierto y sea (x_n) una sucesión de elementos en $D(f)$ que converge a a. Dado $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ con la propiedad establecida en (b). Como (x_n) converge a a, existe un número natural $N(\delta(\varepsilon))$

tal que si $n \geq N(\delta(\varepsilon))$, entonces $\|x_n - a\| < \delta(\varepsilon)$. Dado que $x_n \in D(f)$ para todo n, de (b) se deduce que $\|f(x_n) - f(a)\| < \varepsilon$. Por lo tanto, se cumple (c).

$c \Rightarrow a$) Demostraremos que si la condición (a) no es válida, entonces la condición (c) tampoco lo es. Si (a) no se cumple, existe un entorno V_0 de $f(a)$ tal que para cualquier entorno U de a existe un elemento x_U que pertenece a $D(f) \cap U$ pero tal que $f(x_U)$ no pertenece a V_0. Para cada número natual n, consideremos el entorno U_n de a definido por $U_n = \{x \in \mathbf{R}^p : \|x - a\| < 1/n\}$; por lo establecido en la frase anterior, para cada n en $\mathbb{N}$ existe un elemento x_n que pertenece a $D(f) \cap U_n$ pero tal que $f(x_n)$ no pertenece a V_0. La sucesión (x_n) recién construida pertenece a $D(f)$ y converge a a; sin embargo, ninguno de los elementos de la sucesión $(f(x_n))$ pertenece al entorno V_0 de $f(a)$. Por lo tanto, se ha construido una sucesión para la cual no es válida la condición (c). Esto demuestra que la parte (c) implica (a). ■

El siguiente criterio sobre discontinuidad es una consecuencia de lo que se acaba de hacer.

Teorema 244 ***(Criterio de discontinuidad)*** *La función f no es continua en un punto $a \in D(f)$ si y sólo si existe una sucesión (x_n) de elementos de $D(f)$ que converge a a pero tal que la sucesión $(f(x_n))$ de imágenes no converge a $f(a)$.*

El siguiente resultado es una simple reformulación de la definición de continuidad. Recordamos que la **imagen inversa** $f^{-1}(H)$ de un subconjunto H de $\mathbb{R}^q$ bajo f está definida por

$$f^{-1}(H) = \{x \in D(f) : f(x) \in H\}.$$

Teorema 245 *La función f es continua en un punto a en $D(f)$ si y sólo si para todo entorno V de $f(a)$ existe un entorno V_1 de a tal que*

$$V_1 \cap D(f) = f^{-1}(V).$$

Demostración: *Si V_1 es un entorno de a que satisface esta ecuación, entonces se puede tomar $U = V_1$. Inversamente, si la definición de continuidad se satisface, entonces se puede tomar $V_1 = U \cup f^{-1}(V)$ para obtener la ecuación del enunciado.* ■

Antes de seguir adelante con la teoría se hará una pausa para dar algunos ejemplos. Por simplicidad, la mayoría de los ejemplos son para el caso en que $\mathbb{R}^p = \mathbb{R}^q = \mathbb{R}$.

Ejemplo 324 *Sea $D(f) = \mathbb{R}$ y sea f la función "constante" definida como la función que es igual al número real c para todos los números reales x. Entonces, f es continua en todo punto de $\mathbb{R}$; de hecho, se puede tomar el entorno U de la definición de continuidad igual a $\mathbb{R}$ para cualquier $a \in D(f)$. Análogamente, la función g definida por*

$$g(x) = \begin{cases} 1 & si \quad 0 \leq x \leq 1 \\ 2 & si \quad 2 \leq x \leq 3 \end{cases}$$

es continua en cada punto de su dominio.

Ejemplo 325 *Sea $D(f) = \mathbb{R}$ y sea f la función identidad definida como $f(x) = x$, $x \in \mathbb{R}$. Si a es un número real dado, sea $\varepsilon > 0$ y sea $\delta(\varepsilon) = \varepsilon$. Si $|x-a| < \delta(\varepsilon)$, se tiene $|f(x) - f(a)| = |x-a| < \varepsilon$.*

Ejemplo 326 *Sea $D(f) = \mathbb{R}$ y sea f la función cuadrática definida como $f(x) = x^2$, $x \in \mathbb{R}$. Sean a un elemento de $\mathbb{R}$ y $\varepsilon > 0$; entonces,*

$$|f(x) - f(a)| = \left|x^2 - a^2\right| = |x-a| \cdot |x+a| .$$

Se quiere hacer esta última expresión menor que ε haciendo a $|x-a|$ lo suficientemente pequeña. Si $a = 0$, se elige $\delta(\varepsilon) = \sqrt{\varepsilon}$. Si $a \neq 0$, se desea obtener una cota para $|x+a|$ en un entorno de a. Por ejemplo, si $|x-a| < |a|$, entonces $0 < |x| < 2|a|$ y $|x+a| \leq |x| + |a| < 3|a|$. De donde

$$|f(x) - f(a)| \leq 3|a|\,|x-a| ,$$

siempre que $|x-a| < |a|$. Por tanto, si se define $\delta(\varepsilon) = \inf\left\{|a|, \frac{\varepsilon}{3|a|}\right\}$, entonces cuando $|x-a| < \delta(\varepsilon)$, se cumple:

$$|f(x) - f(a)| \leq 3|a|\,|x-a| < \varepsilon.$$

Ejemplo 327 *Tomaremos ahora la misma función que en (c) pero usaremos una técnica un poco distinta: en vez de factorizar $x^2 - a^2$, lo escribiremos como un polinomio en $x - a$.*

$$x^2 - a^2 = (x^2 - 2ax + a^2) + (2ax - 2a^2) = (x-a)^2 + 2a(x-a).$$

Usando la desigualdad triangular se obtiene

$$|f(x) - f(a)| \leq |x-a|^2 + 2|a| \cdot |x-a| .$$

Si $\delta \leq 1$ y $|x-a| < \delta$, entonces $|x-a|^2 < \delta^2 \leq \delta$ y el término de la derecha está acotado superiormente por $\delta + 2|a|\delta = \delta(1 + 2|a|)$. Por lo que se escoge

$$\delta(\varepsilon) = \inf\left\{1, \frac{\varepsilon}{1 + 2|a|}\right\}.$$

Ejemplo 328 *Consideremos $D(f) = \{x \in \mathbb{R} : x \neq 0\}$ y definamos f como $f(x) = 1/x$, $x \in D(f)$. Si $x \in D(f)$, entonces*

$$|f(x) - f(a)| = \left|\frac{1}{x} - \frac{1}{a}\right| = \frac{|x-a|}{|ax|}.$$

De nuevo se desea encontrar una cota para el coeficiente $|x-a|$ que sea válida en un entorno de $a \neq 0$. Obsérvese que si $|x-a| < \frac{1}{2}|a|$, entonces $\frac{1}{2}|a| < |x|$, y se tiene

$$|f(x) - f(a)| \leq \frac{2}{|a|^2}\,|x-a| .$$

Por tanto, basta tomar $\delta(\varepsilon) = \inf\left\{\frac{1}{2}|a|, \frac{1}{2}\varepsilon|a|^2\right\}$.

Ejemplo 329 *Definamos f de la forma*

$$f(x) = \begin{cases} 0 & si \quad x \leq 0 \\ 1 & si \quad x > 0 \end{cases}$$

Se puede ver que f es continua en todos los puntos distintos de 0. Se habrá de probar que f no es continua en 0 empleando el criterio de discontinuidad. De hecho, si $x_n = 1/n$, la sucesión $(f(1/n)) = (1)$ no converge a $f(0)$.

Ejemplo 330 *Sea $D(f) = \mathbb{R}$ y sea f la función discontinua de Dirichlet definida como*

$$f(x) = \begin{cases} 1 & si \quad x \text{ es racional} \\ 0 & si \quad x \text{ es irracional.} \end{cases}$$

Si a es un número racional, podemos tomar $X = (x_n)$ como una sucesión de números irracionales que convergen a a. Dado que $f(x_n) = 0$ para toda $n \in \mathbb{N}$, la sucesión $(f(x_n))$ no converge a $f(a) = 1$ y f no es continua en el número racional a. Por otro lado, si b es un número irracional, existe una sucesión $Y = (y_n)$ de números racionales que converge a b. La sucesión $(f(y_n))$ no converge a $f(b)$ por lo que f no es continua en b. Por lo tanto, la función de Dirichlet no es continua en ningún punto.

Ejemplo 331 *Sea $D(f) = \{x \in \mathbb{R} : x > 0\}$. Para cualquier número irracional $x > 0$, se define $f(x) = 0$; para un número racional de la forma m/n, donde los números m y n no tienen ningún factor común excepto 1, se define $f(m/n) = 1/n$. Probaremos que f es continua en todo número irracional de $D(f)$ y discontinua en todo número racional de $D(f)$. La última afirmación resulta de tomar una sucesión de números irracionales que converjan al número racional dado y de usar el criterio de discontinuidad. Por otro lado, dado un número irracional a y $\varepsilon > 0$, entonces existe un número natural n tal que $1/n < \varepsilon$. Si δ se elige tan pequeña que el intervalo $(a-\delta, a+\delta)$ no contenga a ningún número racional con denominador menor que n, entonces se deduce que para toda x en este intervalo se tiene:*

$$|f(x) - f(a)| = |f(x)| \leq 1/n < \varepsilon.$$

Por tanto, f es continua en el número irracional a. Así pues, esta función es continua precisamente en los puntos irracionales de su dominio.

Ejemplo 332 *Sea $D(f) = \mathbb{R}^2$ y sea f la función en $\mathbb{R}^2$ con valores en $\mathbb{R}^2$ definida como*

$$f(x, y) = (2x + y, x - 3y).$$

Sea (a, b) un punto fijo en $\mathbb{R}^2$; veremos que f es continua en este punto. Para ello, debemos probar que la expresión

$$\|f(x, y) - f(a, b)\| = \left\{(2x + y - 2a - b)^2 + (x - 3y - a + 3b)^2\right\}^{1/2}$$

se puede hacer arbitrariamente pequeña escogiendo a (x, y) lo suficientemente cerca de (a, b), y utilizaremos la desigualdad $\left\{p^2 + q^2\right\}^{1/2} \leq \sqrt{2} \sup\{|p|, |q|\}$. Resulta bastante evidente probar que los términos

$$|2x + y - 2a - b|, \qquad |\, x - 3y - a + 3b|,$$

se pueden hacer arbitrariamente pequeños eligiendo (x, y) *lo suficientemente cerca de* (a, b) *en* $\mathbb{R}^2$*. De hecho, por la desigualdad triangular,*

$$|2x + y - 2a - b| = |2(x - a) + (y - b)| \leq 2\,|x - a| + |y - b|\,.$$

Ahora, $|x - a| \leq \left\{(x - a)^2 + (y - b)^2\right\}^{1/2} = \|(x, y) - (a, b)\|$*, y análogamente para* $|y - b|$*; entonces se tiene*

$$|2x + y - 2a - b| \leq 3\,\|(x, y) - (a, b)\|\,.$$

De modo análogo,

$$|\; x - 3y - a + 3b| \leq |x - a| + 3\,|y - b| \leq 4\,\|(x, y) - (a, b)\|\,.$$

Por lo tanto, si $\varepsilon > 0$*, se puede tomar* $\delta(\varepsilon) = \varepsilon/(4\sqrt{2})$ *con la seguridad de que si* $\|(x, y) - (a, b)\| < \delta(\varepsilon)$*, entonces* $\|f(x, y) - f(a, b)\| < \varepsilon$*, aunque se puede lograr un valor más grande de* δ *a través de un análisis más refinado (por ejemplo, usando la desigualdad de Schwarz).*

Ejemplo 333 *Sea* $D(f) = \mathbb{R}^2$ *y definimos* f *por medio de*

$$f(x, y) = (x^2 + y^2, 2xy)$$

Si (a, b) *es un punto fijo en* $\mathbb{R}^2$*,*

$$\|f(x, y) - f(a, b)\| = \left\{(x^2 + y^2 - a^2 - b^2)^2 + (2xy - 2ab)^2\right\}^{1/2}.$$

Igual que en el ejemplo anterior, examinaremos por separado los dos términos de la derecha. Por la desigualdad triangular, se tiene

$$\left|x^2 + y^2 - a^2 - b^2\right| \leq \left|x^2 - a^2\right| + \left|y^2 - b^2\right|.$$

Si el punto (x, y) *está dentro de una distancia de 1 de* (a, b)*, entonces,* $|x| \leq |a|+1$ *por lo que* $|x + a| \leq 2\,|a|+1$ *y* $|y| \leq |b|+1$*, de manera que* $|y + b| \leq 2\,|b|+1$*. Por lo tanto se tiene*

$$\begin{aligned} \left|x^2 + y^2 - a^2 - b^2\right| &\leq |x - a|\,(2\,|a| + 1) + |y - b|\,(2\,|b| + 1) \\ &\leq 2(|a| + |b| + 1)\,\|(x, y) - (a, b)\|\,. \end{aligned}$$

De manera análoga se tiene

$$\begin{aligned} |2xy - 2ab| &= 2\,|xy - xb + xb - ab| \leq 2\,|x|\,|y - b| + 2\,|b|\,|x - a| \\ &\leq 2(|a| + |b| + 1)\,\|(x, y) - (a, b)\|\,. \end{aligned}$$

Por lo tanto, basta establecer

$$\delta(\varepsilon) = \inf\left\{1, \frac{\varepsilon}{2\sqrt{2}(|a| + |b| + 1)}\right\};$$

Si $\|(x, y) - (a, b)\| < \delta(\varepsilon)$*, se tiene* $\|f(x, y) - f(a, b)\| < \varepsilon$*, con lo cual tenemos que* f *es continua en el punto* (a, b)*.*

13.1.1. Combinaciones de funciones

Recordamos que si f y g son funciones con dominios $D(f)$ y $D(g)$ en $\mathbb{R}^p$ e imágenes en $\mathbb{R}^q$, entonces se definen la **suma** $f+g$, la **diferencia** $f-g$ y el **producto interno** $f \cdot g$ para cada x en $D(f) \cap D(g)$ por medio de las fórmulas

$$f(x)+g(x), \quad f(x)-g(x), \quad f(x) \cdot g(x).$$

Análogamente, si c es un número real y si φ es una función con dominio $D(\varphi)$ en $\mathbb{R}^p$ e imagen en $\mathbb{R}$, se definen los productos $c \cdot f$ para x en $D(f)$ y $\varphi \cdot f$ para x en $D(\varphi) \cap D(f)$ por medio de las fórmulas

$$c \cdot f(x), \quad \varphi(x) \cdot f(x).$$

En particular, si $\varphi(x) \neq 0$ para $x \in D_0$, se puede definir el **cociente** f/φ para x en $D(f) \cap D_0$ como

$$f(x)/\varphi(x).$$

Teorema 246 *Sean* $f : D(f) \subset \mathbb{R}^p \to \mathbb{R}^q$, $g : D(g) \subset \mathbb{R}^p \to \mathbb{R}^q$, $\varphi : D(\varphi) \subset \mathbb{R}^p \to \mathbb{R}$, $c \in \mathbb{R}$. *Si las funciones* f, g, φ *son continuas en un punto, entonces las funciones*

$$f+g, \quad f-g, \quad f \cdot g, \quad c \cdot f, \quad \varphi \cdot f \quad y \quad f/\varphi$$

también son continuas en ese punto.
Demostración: *En la unidad El espacio euclídeo n-dimensional, vimos que se cumple:*

a. *Si* X *e* Y *son dos sucesiones en* $\mathbb{R}^p$ *que convergen a* x *e* y*, respectivamente, entonces las sucesiones* $X+Y$, $X-Y$ *e* $X \cdot Y$ *convergen a* $x+y$, $x-y$ *y* $x \cdot y$*, respectivamente.*

b. *Si* $X=(x_n)$ *es una sucesión en* $\mathbb{R}^p$ *que converge a* x *y* $A=(a_n)$ *es una sucesión en* $\mathbb{R}$ *que converge a* a*, entonces la sucesión* $(a_n \cdot x_n) \subset \mathbb{R}^p$ *converge a* ax*.*

c. *Si* $X=(x_n)$ *es una sucesión en* $\mathbb{R}^p$ *que converge a* x *y* $B=(b_n)$ *es una sucesión de números reales distintos de cero que converge a* b*, entonces la sucesión* $(\frac{x_n}{b_n}) \subset \mathbb{R}^p$ *converge a* $\frac{x}{b}$*.*

Utilizamos, además, el tercer apartado del primer teorema: para demostrar que una función f *es continua, basta ver que para cualquier sucesión de elementos de su dominio que converjan a* x_0*, la sucesión de las imágenes converge a* $f(x_0)$*.*

De estos dos resultados se deduce trivialmente el enunciado. Veamos, por ejemplo, la demostración de que $f+g$ *es continua: sea* (x_n) *una sucesión en* $\mathbb{R}^p$ *convergente a* x*. Debemos ver que* $(f+g)(x_n)$ *converge a* $(f+g)(x)$*:*

$$(f+g)(x_n) = f(x_n)+g(x_n) \to f(x)+g(x) = (f+g)(x) \quad \textit{porque } f \textit{ y } g \textit{ son continuas}$$

Análogamente para las demás combinaciones de funciones. ■

Existe otra combinación algebraica que a menudo resulta útil. Si f está definida para $D(f)$ de $\mathbb{R}^p$ a $\mathbb{R}$, se define el valor absoluto $|f|$ de f como la función con imagen en $\mathbb{R}$ tal que $|f|(x) = |f(x)|$ para toda $x \in D(f)$.

Teorema 247 *Si f es continua en un punto, entonces $|f|$ también es continua en ese punto.*
Demostración: *Por la desigualdad triangular, se tiene*

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)|,$$

de donde el resultado es inmediato. ∎

Sea f con dominio $D(f)$ en $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$ y sea g con dominio $D(g)$ en $\mathbb{R}^q$ e imagen en $\mathbb{R}^{'}$. Se define la composición $h = g \circ f$ como la función que tiene dominio $D(h) = \{x \in D(f) \text{ tales que } f(x) \in D(g)\}$ y en la que para $x \in D(h)$ se tiene $h(x) = g[f(x)]$. Por tanto, $h = g \circ f$ es una función que transforma a $D(h)$, que es un subconjunto de $D(f) \subseteq \mathbb{R}^p$, en un subconjunto de $\mathbb{R}^{'}$.

Teorema 248 *Si f es continua en a y g es continua en $b = f(a)$, entonces la composición $g \circ f$ es continua en a.*
Demostración: *Sea W un entorno del punto $c = g(b)$. Dado que g es continua en b, existe un entorno V de b tal que si y pertenece a $V \cap D(g)$, entonces $g(y) \in W$. Como f es continua en a, existe un entorno U de a tal que si x pertenece a $U \cap D(f)$, entonces $f(x)$ está en V. Por lo tanto, si x pertenece a $U \cap D(g \circ f)$, entonces $f(x)$ está en $V \cap D(g)$ y $g[f(x)]$ pertenece a W. Esto prueba que $h = g \circ f$ es continua en a.* ∎

13.2. Funciones lineales

Definición 122 *Se dice que una función $f : \mathbb{R}^p \to \mathbb{R}^q$ es lineal cuando*

$$f(ax + by) = a \cdot f(x) + b \cdot f(y)$$

para toda $a, b \in \mathbb{R}$ y $x, y \in \mathbb{R}^p$.

Por inducción se deduce de que si $a, b, ..., c$ son n números reales y $x, y, ..., z$ son n elementos de $\mathbb{R}^p$, entonces

$$f(ax + by + \cdots + cz) = af(x) + bf(y) + \cdots + cf(z).$$

Observación 137 *Si f y g son funciones lineales de $\mathbb{R}^p$ a $\mathbb{R}^q$, entonces $f + g$ es una función lineal de $\mathbb{R}^p$ a $\mathbb{R}^q$. Análogamente, si $c \in \mathbb{R}$, entonces cf es una función lineal. La colección $\mathcal{L}(\mathbb{R}^p, \mathbb{R}^q)$ de todas las funciones lineales de $\mathbb{R}^p$ a $\mathbb{R}^q$ es un espacio vectorial bajo estas operaciones vectoriales.*

Teorema 249 *Si f es una función lineal con dominio $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$, entonces existen $p \cdot q$ números reales (c_{ij}), $1 \leq i \leq q$, $1 \leq j \leq p$, tales que si $x = (x_1, x_2, ..., x_p)$ es cualquier punto en $\mathbb{R}^p$, y si $y = (y_1, y_2, ..., y_q) = f(x)$, entonces*

$$\begin{aligned} y_1 &= c_{11}x_1 + c_{12}x_2 + \cdots + c_{1p}x_p, \\ y_2 &= c_{21}x_1 + c_{22}x_2 + \cdots + c_{2p}x_p, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y_q &= c_q x_1 + c_q x_2 + \cdots + c_{qp}x_p. \end{aligned}$$

De manera inversa, si (c_{ij}) *es una colección de* $p \cdot q$ *números reales, entonces la función que asigna a* $x \in \mathbb{R}^p$ *el elemento* $y \in \mathbb{R}^q$ *según las ecuaciones anteriores, es una función lineal con dominio* $\mathbb{R}^p$ *e imagen en* $\mathbb{R}^q$.
Demostración: *Sean* $e_1, e_2, ..., e_p$ *los elementos de* $\mathbb{R}^p$ *dados por* $e_1 = (1, 0, ..., 0)$, $e_2 = (0, 1, ..., 0), ...,\ e_p = (0, 0, ..., 1)$. *Supongamos que*

$$\begin{array}{rcl} f(e_1) & = & (c_{11}, c_{21}, ..., c_{q1}), \\ f(e_2) & = & (c_{12}, c_{22}, ..., c_{q2}), \\ & & \cdots\cdots\cdots\cdots\cdots\cdots \\ f(e_p) & = & (c_{1p}, c_{2p}, ..., c_{qp}). \end{array}$$

Así pues, el número real c_{ij} *es la i-ésima coordenada del punto* $f(e_j)$.

Un elemento arbitrario $x = (x_1, x_2, ..., x_p)$ *de* $\mathbb{R}^p$ *se puede expresar en términos de los vectores* $e_1, e_2, ..., e_p$*:*

$$x = x_1e_1 + x_2e_2 + \cdots + x_pe_p.$$

Dado que f *es lineal, se cumple que*

$$f(x) = x_1f(e_1) + x_2f(e_2) + \cdots + x_pf(e_p).$$

Usando las expresiones de $f(e_i)$, *se tiene*

$$\begin{array}{rcl} f(x) & = & x_1(c_{11}, c_{21}, ..., c_{q1}) + x_2(c_{12}, c_{22}, ..., c_{q2}) \\ & & + \cdots + x_p(c_{1p}, c_{2p}, ..., c_{qp}) = \\ & = & (c_{11}x_1, c_{21}x_1, ..., c_{q1}x_1) + (c_{12}x_2, c_{22}x_2, ..., c_{q2}x_2) \\ & & + \cdots + (c_{1p}x_p, c_{2p}x_p, ..., c_{qp}x_p) = \\ & = & (c_{11}x_1 + c_{12}x_2 + \cdots + c_{1p}x_p, c_{21}x_1 + c_{22}x_2 + \cdots + c_{2p}x_p, \\ & & \cdots, c_{q1}x_1 + c_{q2}x_2 + \cdots + c_{qp}x_p). \end{array}$$

Esto prueba que las coordenadas de $f(x)$ *están dadas por las relaciones que el enunciado aseguraba.*

A la inversa, es fácil probar, por cálculos directos, que si las relaciones del enunciado se usan para obtener las coordenadas y_i *de* y *de las coordenadas* x_j *de* x, *entonces la función que resulta satisface la definición de linealidad y por lo tanto es lineal.* ■

Definición 123 *Se debe mencionar que al arreglo rectangular de números*

$$\left[\begin{array}{cccc} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ c_{q1} & c_{q2} & \cdots & c_{qp} \end{array}\right],$$

que consta de q *filas y* p *columnas a menudo se le llama la* ***matriz*** *correspondiente a la función lineal* f.

Hay una correspondencia uno a uno entre las funciones lineales de $\mathbb{R}^p$ a $\mathbb{R}^q$ y las matrices de $q \times p$ con números reales. Como se ha visto, la acción de f está completamente descrita en términos de su matriz. Se hará referencia a la matriz anterior como una síntesis de una descripción más elaborada de la función lineal f.

Teorema 250 *Si f es una función lineal con dominio $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$, entonces existe una constante positiva A tal que si u,v son dos vectores cualesquiera en $\mathbb{R}^p$, entonces*

$$\|f(u) - f(v)\| \leq A\,\|u - v\|\,.$$

Por lo tanto, una función lineal de $\mathbb{R}^p$ a $\mathbb{R}^q$ es continua en todo punto.
Demostración: *Escribimos la desigualdad de Schwarz en la forma*

$$|a_1b_1 + a_2b_2 + \cdots + a_pb_p|^2 \leq \left\{a_1^2 + a_2^2 + \cdots + a_p^2\right\}\left\{b_1^2 + b_2^2 + \cdots + b_p^2\right\}.$$

Se aplica esta desigualdad a cada expresión del enunciado del teorema anterior y obtenemos:

$$|y_i|^2 \leq (|c_{i1}|^2 + |c_{i2}|^2 + \cdots + |c_{ip}|^2)\,\|x\|^2 = \sum_{j=1}^{p} |c_{ij}|^2\,\|x\|^2\,.$$

Sumando estas desigualdades se tiene

$$\|y\|^2 \leq \left\{\sum_{i=1}^{q}\sum_{j=1}^{p} |c_{ij}|^2\right\}\|x\|^2\,,$$

de donde se concluye que

$$\|y\| = \|f(x)\| \leq \left\{\sum_{i=1}^{q}\sum_{j=1}^{p} |c_{ij}|^2\right\}^{1/2}\|x\|\,.$$

Por tanto, existe una constante A tal que si x es cualquier elemento de $\mathbb{R}^p$, entonces $\|f(x)\| \leq A\,\|x\|$. Ahora, sea $x = u - v$. Usando la linealidad de f se obtiene

$$f(x) = f(u - v) = f(u) - f(v),$$

tal como queríamos ver. Es claro que esta relación implica la continuidad de f, ya que se puede hacer $\|f(u) - f(v)\| < \varepsilon$ tomando $\|u - v\| < \varepsilon/A$ si $A > 0$. ■

13.3. Propiedades globales de funciones continuas

En la primera sección se consideró la continuidad "local"; es decir, se estaba tratando la continuidad en un punto. En esta sección se establecerán algunas propiedades más profundas acerca de funciones continuas. Aquí se estará tratando la continuidad "global" en el sentido de que se supondrá que las funciones son continuas en todo punto de sus dominios.

A menos que se especifique lo contrario, f designará a una función con dominio $D(f)$ contenido en $\mathbb{R}^p$ y con imagen $\mathrm{Im}(f)$ en $\mathbb{R}^q$.

Definición 124 *Si B es subconjunto del espacio de rango $\mathbb{R}^q$, la* ***imagen inversa*** *de B bajo f es el conjunto*

$$f^{-1}(B) = \{x \in D(f) : f(x) \in B\}\,.$$

Observe que $f^{-1}(B)$ de manera automática es un subconjunto de $D(f)$ aun cuando B no necesariamente sea un subconjunto de $\mathrm{Im}(f)$.

En cursos de topología, en que se trata más continuidad global que continuidad local, el siguiente resultado a menudo se toma como la definición de continuidad (global). Su importancia muy pronto será evidente.

Teorema 251 ***(Teorema de continuidad global)*** *Sea $f : D(f) \subset \mathbb{R}^p \to \mathbb{R}^q$. Las siguientes afirmaciones son equivalentes:*

a. *f es continua en su dominio $D(f)$.*

b. *Si G es cualquier conjunto abierto en $\mathbb{R}^q$, entonces existe un conjunto abierto G_1 en $\mathbb{R}^p$ tal que $G_1 \cap D(f) = f^{-1}(G)$.*

c. *Si H es cualquier conjunto cerrado en $\mathbb{R}^q$, entonces existe un conjunto cerrado H_1 en $\mathbb{R}^p$ tal que $H_1 \cap D(f) = f^{-1}(H)$.*

Demostración: *$a \Rightarrow b)$ Supongamos que f es continua en $D(f)$, y sea G un subconjunto abierto de $\mathbb{R}^q$. Si a pertenece a $f^{-1}(G)$, como G es un entorno de $f(a)$ y f es continua en a, se infiere que hay un conjunto abierto $U(a)$ tal que si $x \in D(f) \cap U(a)$, entonces $f(x) \in G$. Podemos elegir $U(a)$ para cada a en $f^{-1}(G)$, y sea G_1 la unión de los conjuntos $U(a)$. Por las propiedades de los conjuntos abiertos, G_1 es abierto y está claro que $G_1 \cap D(f) = f^{-1}(G)$. Por lo tanto, (a) implica (b).*

$b \Rightarrow a)$ Si a es un punto arbitrario de $D(f)$ y G es un entorno abierto de $f(a)$, la condición (b) implica que existe un conjunto abierto G_1 en $\mathbb{R}^p$ tal que $G_1 \cap D(f) = f^{-1}(G)$. Dado que $f(a) \in G$, se deduce que $a \in G_1$, por lo que G_1 es un entorno de a. Si $x \in G_1 \cap D(f)$, $f(x) \in G$ por lo que f es continua en a. Esto prueba que la condición (b) implica (a).

A continuación demostraremos la equivalencia de las condiciones (b) y (c). Primero, se observa que si B es cualquier subconjunto de $\mathbb{R}^q$ y si $C = \mathbb{R}^q \backslash B$, se tiene $f^{-1}(B) \cap f^{-1}(C) = \emptyset$ y

$$D(f) = f^{-1}(B) \cup f^{-1}(C).$$

Si B_1 es un subconjunto de $\mathbb{R}^p$ tal que $B_1 \cap D(f) = f^{-1}(B)$ y $C_1 = \mathbb{R}^p \backslash B_1$, entonces $C_1 \cap f^{-1}(B) = \emptyset$ y

$$D(f) = (B_1 \cap D(f)) \cup (C_1 \cap D(f)) = f^{-1}(B) \cup (C_1 \cap D(f)).$$

Estas fórmulas son dos representaciones de $D(f)$ como la unión de $f^{-1}(B)$ con otro conjunto con el que no tiene puntos en común. Por lo tanto, se tiene $C_1 \cap D(f) = f^{-1}(C)$.

$b \Rightarrow c)$ Supongamos que (b) es válido y que H es cerrado en $\mathbb{R}^q$. Apliquemos el argumento anterior en el caso en que $B = \mathbb{R}^q \backslash H$ y $C = H$. Entonces, B y B_1 son conjuntos abiertos en $\mathbb{R}^q$ y $\mathbb{R}^p$, respectivamente, de manera que $C_1 = \mathbb{R}^q \backslash B_1$ es cerrado en $\mathbb{R}^p$. Esto prueba que (b) implica (c).

$c \Rightarrow b)$ Para ver que (c) implica (b), se usa el argumento anterior con $B = \mathbb{R}^p \backslash G$, en donde G es un conjunto abierto en $\mathbb{R}^q$. ∎

En el caso en que $D(f) = \mathbb{R}^p$, el resultado anterior se simplifica hasta cierto punto:

Corolario 34 *Sea f una función definida en todo $\mathbb{R}^p$ y con imagen en $\mathbb{R}^q$. Las siguientes afirmaciones son equivalentes:*

a. *f es continua en $\mathbb{R}^p$.*

b. *Si G es abierto en $\mathbb{R}^q$, entonces $f^{-1}(G)$ es abierto en $\mathbb{R}^p$.*

c. *Si H es cerrado en $\mathbb{R}^q$, entonces $f^{-1}(H)$ es cerrado en $\mathbb{R}^p$.*

Observación 138 *El teorema de continuidad global no dice que si f es continua y G un conjunto abierto en $\mathbb{R}^p$, entonces la imagen directa $f(G) = \{f(x) : x \in G\}$ sea abierta en $\mathbb{R}^q$. En general, una función continua no necesariamente manda conjuntos abiertos a conjuntos abiertos o conjuntos cerrados a conjuntos cerrados. Por ejemplo, la función f de $\mathbb{R}$ a $\mathbb{R}$ definida como*

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

es continua en $\mathbb{R}$, puesto que es una combinación de funciones continuas, y el denominador nunca se anula. Si G es el conjunto abierto $G = (-1, 1)$, entonces $f(G) = (\frac{1}{2}, 1]$, que no es abierto en $\mathbb{R}$. Análogamente, si H es el conjunto cerrado $H = \{x \in \mathbb{R} : x \geq 1\}$, entonces $f(H) = (0, \frac{1}{2}]$, que no es cerrado en $\mathbb{R}$. De manera similar, la función f aplica el conjunto $\mathbb{R}$, que es abierto y cerrado en $\mathbb{R}$, al conjunto $f(\mathbb{R}) = (0, 1]$ que no es ni abierto ni cerrado en $\mathbb{R}$.

Lo que nos dicen estas afirmaciones es que la propiedad de un conjunto de ser cerrado o abierto no necesariamente se preserva con la aplicación de una función continua. Sin embargo, existen propiedades importantes de conjunto que se preservan con aplicaciones continuas. Por ejemplo, se va a demostrar que las propiedades de conexidad y compacidad de conjuntos tienen esta característica.

13.3.1. Conservación de conexidad

Recordamos que un conjunto $H \subset \mathbb{R}^p$ es inconexo si existen conjuntos abiertos A, B en $\mathbb{R}^p$ tales que $A \cap H$ y $B \cap H$ son conjuntos disjuntos no vacíos cuya unión es H. Un conjunto es conexo si no es inconexo.

Teorema 252 ***(Conservación de conexidad)*** *Si $H \subseteq D(f)$ es conexo en $\mathbb{R}^p$ y f es continua en H, entonces $f(H)$ es conexo en $\mathbb{R}^q$.*
Demostración: *Sea h la restricción de f al conjunto H de tal manera que $D(h) = H$ y $h(x) = f(x)$ para toda $x \in H$. Observese que $f(H) = h(H)$ y que h es continua en H.*

Si $f(H) = h(H)$ es inconexo en $\mathbb{R}^q$, entonces existen dos conjuntos abiertos $A, B \subset \mathbb{R}^q$ tales que $A \cap h(H)$ y $B \cap h(H)$ son conjuntos disjuntos no vacíos cuya unión es $h(H)$. Por el teorema de continuidad global, existen conjuntos abiertos A_1, B_1 en $\mathbb{R}^p$ tales que

$$A_1 \cap H = h^{-1}(A), \quad B_1 \cap H = h^{-1}(B).$$

Estas intersecciones son no vacías y son disjuntas porque los conjuntos $A \cap h(H)$ y $B \cap h(H)$ son disjuntos. El supuesto de que la unión de $A \cap h(H)$ y $B \cap h(H)$ sea $h(H)$ implica que la unión de $A_1 \cap H$ y $B_1 \cap H$ es H. Por lo tanto, la inconexion de $f(H) = h(H)$ implica la inconexion de H. ■

La misma palabra "continuo" indica que ho hay "interrupciones" bruscas en la gráfica de la función, por lo que el siguiente resultado no es de ninguna manera inesperado.

Teorema 253 ***(Teorema de valor intermedio de Bolzano)*** *Sea $H \subseteq D(f)$ un subconjunto conexo de $\mathbb{R}^p$ y sea f continua en H y con valores en $\mathbb{R}$. Si k es cualquier número real que satisface*

$$\inf \{f(x) : x \in H\} < k < \sup \{f(x) : x \in H\},$$

entonces hay al menos un punto de H en donde f toma el valor k.
Demostración: *Si $k \notin f(H)$, entonces los conjuntos $A = \{t \in \mathbb{R} : t < k\}$ y $B = \{t \in \mathbb{R} : t > k\}$ forman una separación de $f(H)$, lo que contradice el teorema anterior.* ∎

13.3.2. Conservación de compacidad

En seguida demostraremos que la importante propiedad de compacidad se conserva con la aplicación continua. Recordamos que por el teorema de Heine-Borel un subconjunto $K \subset \mathbb{R}^p$ es compacto si y sólo si es cerrado y acotado en $\mathbb{R}^p$. De modo que el siguiente resultado se puede expresar diciendo que si K es cerrado y acotado en $\mathbb{R}^p$ y si f es continua en K y con imagen en $\mathbb{R}^q$, entonces $f(K)$ es cerrado y acotado en $\mathbb{R}^q$.

Teorema 254 ***(Conservación de compacidad)*** *Si $K \subseteq D(f)$ es compacto y f es continua en K, entonces $f(K)$ es compacto.*
Demostración: *Demostraremos este teorema de dos formas:*

Primera demostración: *Supongamos que K es cerrado y acotado en $\mathbb{R}^p$; debemos probar que $f(K)$ es cerrado y acotado en $\mathbb{R}^q$. Si $f(K)$ no fuera acotado, para cada $n \in \mathbb{N}$ existiría un punto $x_n \in K$ con $\|f(x_n)\| \geq n$. Dado que K es acotado, la sucesión $X = (x_n)$ es acotada, de donde, por el teorema de Bolzano-Weierstrass, se deduce que hay una subsucesión de X que converge a un elemento x. Dado que $x_n \in K$ para toda $n \in \mathbb{N}$ y K es cerrado, x pertenece a K. Por lo tanto, f es continua en x y f está acotada por $\|f(x)\| + 1$ en un entorno de x. Dado que esto contradice al supuesto de que $\|f(x_n)\| \geq n$, el conjunto $f(K)$ es acotado.*

Para probar que $f(K)$ es cerrado demostraremos que cualquier punto de acumulación $y \in f(K)$ debe estar contenido en $f(K)$. De hecho, si n es un número natural, existe un punto $z_n \in K$ tal que $\|f(z_n) - y\| < 1/n$. Por el teorema de Bolzano-Weierstrass, la sucesión $Z = (z_n)$ tiene una subsucesión $Z' = (z_{n_k})$ que converge a un elemento z. Dado que K es cerrado, $z \in K$ y f es continua en z. Por lo tanto,

$$f(z) = \lim_k (f(z_{n_k})) = y,$$

lo cual demuestra que y pertenece a $f(K)$. Por lo tanto, $f(K)$ es cerrado.

Segunda demostración: *Restringiendo f a K se puede suponer que $D(f) = K$. Supongamos ahora que $\mathcal{G} = \{G_\alpha\}$ es una familia de conjuntos abiertos en $\mathbb{R}^q$ cuya unión contiene a $f(K)$. Por el teorema de continuidad global,* *para cada conjunto G_α en $\mathcal{G}$ hay un subconjunto abierto C_α de $\mathbb{R}^p$ tal que $C_\alpha \cap D = f^{-1}(G_\alpha)$. La familia $\mathcal{C} = \{C_\alpha\}$ consta de subconjuntos abiertos de $\mathbb{R}^p$; la unión de estos conjuntos contiene a K, ya que si $x \in K$, entonces $f(x)$ está contenido en $f(K)$. Por lo tanto, $f(x)$ pertenece a algun conjunto G_α y, por construcción, x pertenece al conjunto correspondiente C_α. Dado que K es compacto, está contenido en la unión de un número finito de conjuntos de c y su imagen $f(K)$ está contenida en la unión del número finito de conjuntos correspondientes en $\mathcal{G}$. Dado que esto es válido para una familia arbitraria $\mathcal{G}$ de conjuntos abiertos que cubren a $f(K)$, el conjunto $f(K)$ es compacto en $\mathbb{R}^q$.*

■

Cuando la imagen de la función es $\mathbb{R}$, el siguiente teorema se puede expresar de otra manera diciendo que una función continua de valor real en un conjunto compacto alcanza sus valores máximos y mínimo.

Teorema 255 ***(Teorema del valor máximo y mínimo)*** *Sea $K \subseteq \mathrm{Im}(f)$ compacto en $\mathbb{R}^p$ y sea f una función continua de valor real. Entonces existen puntos x^* en K tales que*

$$f(x^*) = \sup\{f(x) : x \in K\}, \qquad f(x_*) = \inf\{f(x) : x \in K\}.$$

Demostración: *Demostraremos este teorema de dos formas:*

Primera demostración: *Dado que K es compacto en $\mathbb{R}^p$, del teorema de conservación de compacidad se deduce que $f(K)$ está acotado en $\mathbb{R}$. Sean $M = \sup f(K)$ y (x_n) una sucesión en K tal que*

$$f(x_n) \geq M - 1/n, \quad n \in \mathbb{N}.$$

Por el teorema de Bolzano-Weierstrass, alguna subsucesión (x_{n_k}) converge a un límite $x^* \in K$. Dado que f es continua en x^*, se debe tener $f(x^*) = \lim(f(x_{n_k})) = M$. La demostración de la existencia de x_* es análoga.

Segunda demostración: *Restringiendo f a K, se puede suponer que $D(f) = K$. Se fija $M = \sup f(K)$. Entonces, para cada $n \in \mathbb{N}$, sea*

$$G_n = \{u \in \mathbb{R} : u < M - 1/n\}$$

Dado que G_n es abierto, del teorema de continuidad global se deduce que existe un conjunto abierto C_n en $\mathbb{R}^p$ tal que

$$C_n \cap K = \{x \in K : f(x) < M - 1/n\}.$$

Si no se alcanzara el valor M entonces la unión de la familia $\mathcal{C} = \{C_n\}$ de conjuntos abiertos contendría a todo K. Dado que K es compacto y la familia $\{C_n \cap K\}$ es creciente, existe $r \in \mathbb{N}$ tal que $K \subseteq C_r$ pero entonces se tendría $f(x) < M - 1/r$ para toda $x \in K$, lo que contradice al hecho de que $M = \sup f(K)$.

■

Si $\mathrm{Im}(f) \subset \mathbb{R}^q$, con $q > 1$, el siguiente corolario algunas veces resulta útil.

Corolario 35 *Sea $f : D(f) \subseteq \mathbb{R}^p \to \mathbb{R}^q$ y sea $K \subseteq D(f)$ compacto. Si f es continua en K, existen puntos x^* en K tales que*

$$\|f(x^*)\| = \sup\{\|f(x)\| : x \in K\}, \qquad \|f(x_*)\| = \inf\{\|f(x)\| : x \in K\}.$$

Sabemos que si $f : \mathbb{R}^p \to \mathbb{R}^q$ es lineal, entonces existe una constante $M > 0$ tal que $\|f(x)\| \leq M\|x\|$ para toda $x \in \mathbb{R}^p$. Sin embargo, no siempre ocurre que exista una constante $m > 0$ tal que $\|f(x)\| \geq m\|x\|$ para toda $x \in \mathbb{R}^p$. A continuación demostraremos que este es el caso si y sólo si f es una función lineal inyectiva.

Corolario 36 *Sea $f : \mathbb{R}^p \to \mathbb{R}^q$ una función lineal. Entonces f es inyectiva si y sólo si existe $m > 0$ tal que $\|f(x)\| \geq m\|x\|$ para toda $x \in \mathbb{R}^p$.*
Demostración: $\Rightarrow$) *Supongamos que f es inyectiva y sea $S = \{x \in \mathbb{R}^p : \|x\| = 1\}$ la esfera compacta unitaria en $\mathbb{R}^p$. Por el corolario anterior, existe $x_* \in S$ tal que $\|f(x_*)\| = m = \inf\{\|f(x)\| : x \in S\}$. Dado que f es inyectiva, $m = \|f(x_*)\| > 0$. Por lo tanto, $\|f(x)\| \geq m > 0$ para toda $x \in S$. Ahora, si $u \in \mathbb{R}^p$, $u \neq 0$, entonces $u/\|u\|$ pertenece a S y por la linealidad de f se tiene*

$$\frac{1}{\|u\|}\|f(u)\| = \left\|f(\frac{u}{\|u\|})\right\| \geq m,$$

por lo que se deduce que $\|f(u)\| \geq m\|u\|$ para toda $u \in \mathbb{R}^p$ (ya que el resultado es trivial para $u = 0$).

$\Leftarrow$) *De manera inversa supongamos que $\|f(x)\| \geq m\|x\|$ para toda $x \in \mathbb{R}^p$. Si $f(x_1) = f(x_2)$, se tiene*

$$0 = \|f(x_1) - f(x_2)\| = \|f(x_1 - x_2)\| \geq m\|x_1 - x_2\|,$$

que implica $x_1 = x_2$. Por lo tanto, f es inyectiva. ■

Una de las consecuencias más notables del teorema de conservación de compacidad es que si f es continua e inyectiva en un dominio compacto entonces la función inversa f^{-1} automáticamente es continua.

Teorema 256 ***(Continuidad de la función inversa)*** *Sea K un subconjunto compacto de $\mathbb{R}^p$ y sea f una función continua inyectiva con dominio K e imagen en $\mathbb{R}^q$. Entonces la función inversa es continua con dominio $f(K)$ e imagen K.*
Demostración: *Como K es compacto, el teorema de conservación de compacidad implica que $f(K)$ es compacto y por lo tanto cerrado. Dado que por hipótesis f es inyectiva, la función inversa $g = f^{-1}$ está definida. Sea H cualquier conjunto cerrado en $\mathbb{R}^p$ y consideremos a $H \cap K$; dado que este conjunto es cerrado y acotado, el teorema de Heine-Borel asegura que $H \cap K$ es un subconjunto compacto de $\mathbb{R}^p$. Por el teorema de conservación de compacidad se concluye que $H_1 = f(H \cap K)$ es compacto y por lo tanto cerrado en $\mathbb{R}^q$. Si $g = f^{-1}$, entonces*

$$H_1 = f(H \cap K) = g^{-1}(H).$$

Dado que H_1 es un subconjunto de $f(K) = D(g)$, esta última ecuación se puede escribir de la siguiente manera

$$H_1 \cap D(g) = g^{-1}(H).$$

Del teorema de continuidad global se deduce que $g = f^{-1}$ es continua. ■

Definición 125 *Si $D \subseteq \mathbb{R}^p$, entonces la colección de todas las funciones continuas de D a $\mathbb{R}^q$ se denota por $\mathcal{C}_{pq}(D)$. La colección de todas las* ***funciones continuas acotadas*** *de D a $\mathbb{R}^q$ se designa por medio de $\mathcal{BC}_{pq}(D)$. Cuando p y q se sobreentienden, estas colecciones se designan simplemente como $\mathcal{C}(D)$ y $\mathcal{BC}(D)$.*

Teorema 257

a. *Los espacios $\mathcal{C}_{pq}(D)$ y $\mathcal{BC}_{pq}(D)$ son espacios vectoriales bajo las operaciones vectoriales*

$$(f+g)(x) = f(x) + g(x), \qquad (cf)(x) = cf(x) \quad \text{para } x \in D.$$

b. *El espacio $\mathcal{BC}_{pq}(D)$ es un espacio normado bajo la norma*

$$\|f\|_D = \sup\{\|f(x)\| : x \in D\}.$$

Por supuesto, en el caso especial en que D sea un subconjunto compacto de $\mathbb{R}^p$, $\mathcal{C}_{pq}(D) = \mathcal{BC}_{pq}(D)$.

13.4. Continuidad uniforme y puntos fijos

Defínase f en un subconjunto $D(f)$ de $\mathbb{R}^p$ a $\mathbb{R}^q$. Es fácil ver que los siguientes resultados son equivalentes:

i. f es continua en todo punto en $D(f)$.

ii. Dada $\varepsilon > 0$ y $u \in D(f)$, existe una $\delta(\varepsilon, u) > 0$ tal que si x pertenece a $D(f)$ y $\|x - u\| \leq \delta$, entonces $\|f(x) - f(u)\| \leq \varepsilon$.

En general, δ depende tanto de ε como de u. El que δ dependa de u es consecuencia del hecho de que la función f puede cambiar sus valores con rapidez cerca de ciertos puntos y lentamente cerca de otros.

Puede suceder que la función sea tal que el número δ se pueda escoger independiente del punto u en $D(f)$ y dependiente únicamente de ε. Por ejemplo, si $f(x) = 2x$, entonces

$$|f(x) - f(u)| = 2|x - u|$$

de manera que se puede elegir $\delta(\varepsilon, u) = \varepsilon/2$ para todos los valores de u.

Por otro lado, si $g(x) = 1/x$ para $x > 0$, entonces

$$g(x) - g(u) = \frac{u - x}{ux}.$$

Si $0 < \delta < u$ y $|x - u| \leq \delta$, entonces

$$|g(x) - g(u)| \leq \frac{\delta}{u(u - \delta)},$$

y esta desigualdad no se puede mejorar, ya que, de hecho, la igualdad es válida para $x = u - \delta$. Si se quiere hacer $|g(x) - g(u)| \leq \varepsilon$, entonces el valor más grande para δ que se puede elegir es

$$\delta(\varepsilon, u) = \frac{\varepsilon u^2}{1 + \varepsilon u}.$$

De modo que si $u > 0$, g es continua en u ya que se puede elegir $\delta(\varepsilon, u) = \varepsilon u^2/(1+\varepsilon u)$, y éste es el valor más grande que se puede escoger. Dado que

$$\inf \left\{ \frac{\varepsilon u^2}{1+\varepsilon u} : u > 0 \right\} = 0,$$

no se puede obtener una $\delta(\varepsilon, u) > 0$ que sea independiente de la elección de u para todos los puntos $u > 0$.

Ahora restringimos g a un dominio menor. De hecho, sea $a > 0$ y definamos $h(x) = 1/x$ para $x \geq a$. El análisis recién hecho prueba que se puede usar el mismo valor de $\delta(\varepsilon, u)$. Sin embargo, esta vez el dominio es más pequeño y

$$\inf \left\{ \frac{\varepsilon u^2}{1+\varepsilon u} : u \geq a \right\} = \frac{\varepsilon a^2}{1+\varepsilon a} > 0.$$

De modo que si se define $\delta(\varepsilon) = \varepsilon a^2/(1+\varepsilon a)$, se puede usar este número para todos los puntos $u \geq a$.

Con esta introducción se establece ahora la definición formal.

Definición 126 *Sea f la función con dominio $D(f)$ en $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$. Se dice que f es* ***uniformemente continua*** *en un conjunto $A \subseteq D(f)$ si para cada $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si x y u pertenecen a A y $\|x - u\| \leq \delta(\varepsilon)$, entonces $\|f(x) - f(u)\| \leq \varepsilon$.*

Observación 139 *Es claro que si f es uniformemente continua en A entonces es continua en todo punto de A. Sin embargo, en general lo inverso no es válido.*

Es conveniente tener en mente qué se entiende al decir que una función no es uniformemente continua:

Lema 11 *Una condición necesaria y suficiente para que la función f no sea uniformemente continua en $A \subseteq D(f)$ es que existan $\varepsilon_0 > 0$ y dos sucesiones $X = (x_n)$, $Y = (y_n)$ en A tales que si $n \in \mathbb{N}$, entonces $\|x_n - y_n\| \leq \frac{1}{n}$ y $\|f(x_n) - f(y_n)\| > \varepsilon_0$.*
Demostración: $\Rightarrow$) *Por la definición de continuidad uniforme, sabemos que una función no es uniformemente continua si existe $\varepsilon > 0$ tal que para toda $\delta > 0$, existen dos puntos, llamémosles x_δ, y_δ tales que $\|x_\delta - y_\delta\| < \delta$ y $\|f(x_\delta) - f(y_\delta)\| > \varepsilon_0$. Consideremos $\delta = \frac{1}{n}$: sabemos que existen x_n, y_n en A tales que $\|x_n - y_n\| < \frac{1}{n}$ y $\|f(x_n) - f(y_n)\| > \varepsilon_0$. Así pues, las sucesiones $X = (x_n)$ e $Y = (y_n)$ cumplen: $\|x_n - y_n\| < \frac{1}{n}$ y $\|f(x_n) - f(y_n)\| > \varepsilon_0$ para toda $n \in \mathbb{N}$.*

$\Leftarrow$) *Tomemos el ε_0 del enunciado, y sea $\delta > 0$ cualquiera. Sabemos que existen dos sucesiones (x_n) e (y_n) en A tales que $\|x_n - y_n\| \leq \frac{1}{n}$ y $\|f(x_n) - f(y_n)\| > \varepsilon_0$ para toda $n \in \mathbb{N}$. Por tanto, existe $K(\delta)$ tal que $\|x_n - y_n\| < \delta$ para toda $n > K(\delta)$, y $\|f(x_n) - f(y_n)\| \geq \varepsilon_0$. Así pues, basta tomar $x_\delta = x_{K(\delta)+1}$, e $y_\delta = y_{K(\delta)+1}$.* ■

En seguida se ofrece un resultado muy útil que asegura que una función continua de manera automática es uniformemente continua en cualquier conjunto compacto en su dominio.

Teorema 258 ***(Teorema de continuidad uniforme)*** *Sea f una función continua con dominio $D(f) \subset \mathbb{R}^p$ e $\mathrm{Im}(f) \subset \mathbb{R}^q$. Si $K \subseteq D(f)$ es compacto, entonces f es uniformemente continua en K.*
Demostración:

Primera demostración: *Suponga que f no es uniformemente continua en K. Por el lema anterior, existen $\varepsilon_0 > 0$ y dos sucesiones (x_n) e (y_n) en K tales que si $n \in \mathbb{N}$, entonces*

$$\|x_n - y_n\| \leq \frac{1}{n}, \qquad \|f(x_n) - f(y_n)\| > \varepsilon_0.$$

Como K es compacto en $\mathbb{R}^p$, la sucesión X es acotada; por el teorema de Bolzano-Weierstrass, hay una subsucesión (x_{n_k}) de (x_n) que converge a un elemento z. Dado que K es cerrado, el límite z pertenece a K y f es continua en z. Está claro que la subsucesión correspondiente (y_{n_k}) de Y también converge a z.

Como f es continua, ambas sucesiones $(f(x_{n_k}))$ y $(f(y_{n_k}))$ convergen a $f(z)$. Por lo tanto, cuando k es suficientemente grande, se tiene

$$\|f(x_{n_k}) - f(y_{n_k})\| < \varepsilon_0$$

Pero esto contradice que $\|f(x_n) - f(y_n)\| > \varepsilon_0$.

Segunda demostración: *Supongamos que f es continua en todo punto del conjunto compacto K. Entonces, dada $\varepsilon > 0$ y $u \in K$ existe un número $\delta(\frac{1}{2}\varepsilon, u) > 0$ tal que si $x \in K$ y*

$$\|x - u\| < \delta(\frac{1}{2}\varepsilon, u)$$

entonces

$$\|f(x) - f(u)\| < \frac{1}{2}\varepsilon$$

Para cada $u \in K$, formamos la bola abierta

$$G(u) = \left\{ x \in \mathbb{R}^p : \|x - u\| < \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u) \right\}$$

entonces el conjunto K ciertamente está contenido en la unión de la familia $\mathcal{G} = \{G(u) : u \in K\}$, ya que para cada punto u en K hay una bola abierta $G(u)$ que lo contiene. Dado que K es compacto, está contenido en la unión de un número finito de conjuntos en la familia $\mathcal{G}$, digamos $G(u_1), ..., G(u_N)$. Definimos ahora

$$\delta(\varepsilon) = \frac{1}{2} \inf \left\{ \delta(\frac{1}{2}\varepsilon, u_1), ..., \delta(\frac{1}{2}\varepsilon, u_N) \right\},$$

y se habrá de probar que $\delta(\varepsilon)$ tiene la propiedad deseada. Supongamos que x, u pertenecen a K y que $\|x - u\| \leq \delta(\varepsilon)$. Entonces existe un número natural k con $1 \leq k \leq \mathbb{N}$ tal que x pertenece al conjunto $G(u_k)$; es decir,

$$\|x - u_k\| < \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u_k)$$

Dado que $\delta(\varepsilon) \leq \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u_k)$, se deduce que

$$\|u - u_k\| \leq \|u - x\| + \|x - u_k\| < \delta(\frac{1}{2}\varepsilon, u_k).$$

Por lo tanto, se tienen las relaciones

$$\|f(x) - f(u_k)\| < \frac{1}{2}\varepsilon, \qquad \|f(u) - f(u_k)\| < \frac{1}{2}\varepsilon,$$

de donde $\|f(x) - f(u)\| < \varepsilon$*. Se ha demostrado que si* x, u *son dos puntos arbitrarios de* K *para los cuales* $\|x - u\| \leq \delta(\varepsilon)$*, entonces* $\|f(x) - f(u)\| < \varepsilon$*.*

■

En secciones posteriores se hará uso en muchos casos del concepto de continuidad uniforme, de modo que en este momento no se darán aplicaciones. Sin embargo se introducirá otra propiedad que a menudo es asequible y suficiente para garantizar la continuidad uniforme.

Definición 127 *Si* f *tiene dominio* $D(f) \subset \mathbb{R}^p$ *e* $\mathrm{Im}(f) \subset \mathbb{R}^q$*, se dice que* f *satisface una* ***condición de Lipschitz*** *si existe una constante* $A > 0$ *tal que*

$$\|f(x) - f(u)\| \leq A \|x - u\|$$

para todos los puntos x, u *en* $D(f)$*. En el caso en que la desigualdad anterior sea válida para una constante* $A < 1$*, la función recibe el nombre de* ***contracción****.*

Observación 140 *Si* f *satisface una condición de Lipschitz, entonces* f *es uniformemente continua en* $D(f)$ *(basta fijar* $\delta(\varepsilon) = \varepsilon/A$*). Sin embargo, lo inverso no se cumple, como se puede ver al considerar la función* $f(x) = \sqrt{x}$ *definida en* $D(f) = [0, 1]$*. Si* f *cumpliera una condición de Lipschitz, entonces, fijando* $u = 0$ *se debería tener* $|f(x)| \leq A|x|$ *para alguna constante* A*, pero fácilmente se ve que esta última desigualdad no se puede satisfacer.*

Observación 141 *Una función lineal con dominio* $\mathbb{R}^p$ *e imagen en* $\mathbb{R}^q$ *satisface una condición de Lipschitz.*

13.4.1. Teoremas del punto fijo

Definición 128 *Si* f *es una función con dominio e imagen en el mismo espacio* $\mathbb{R}^p$*, entonces se dice que un punto* $u \in D(f)$ *es un* ***punto fijo*** *de* f *si* $f(u) = u$*.*

Teorema 259 ***(Teorema del punto fijo para contracciones)*** *Sea* f *una contracción con dominio* $\mathbb{R}^p$ *e imagen contenida en* $\mathbb{R}^p$*. Entonces,* f *tiene un punto fijo único.*
Demostración: *Como* f *es una contracción, existe una constante* C *con* $0 < C < 1$ *tal que* $\|f(x) - f(y)\| \leq C \|x - y\|$ *para toda* $x, y \in \mathbb{R}^p$*. Sea* x_1 *un punto arbitrario en* $\mathbb{R}^p$ *y definimos* $x_2 = f(x_1)$*; inductivamente, fijamos*

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}.$$

Probaremos que la sucesión (x_n) *converge a un punto fijo único de* f *y calcularemos la rapidez de convergencia.*

Para hacer esto, observemos que

$$\|x_3 - x_2\| = \|f(x_2) - f(x_1)\| \leq C \|x_2 - x_1\|,$$

e, inductivamente,

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq C\,\|x_n - x_{n-1}\| \leq C^{n-1}\,\|x_2 - x_1\|\,.$$

Si $m \geq n$, con el uso repetido de esta desigualdad obtenemos

$$\begin{aligned} \|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \cdots + \|x_{n+1} - x_n\| \leq \\ &\leq \left\{C^{m-2} + C^{m-3} + \cdots + C^{n-1}\right\} \|x_2 - x_1\|\,. \end{aligned}$$

Por lo que se deduce que, para $m \geq n$,

$$\|x_m - x_n\| \leq \frac{C^{n-1}}{1-C}\,\|x_2 - x_1\|\,.$$

Puesto que $0 < C < 1$, la sucesión (C^{n-1}) converge a cero. Por lo tanto, (x_n) es una sucesión de Cauchy. Si $u = \lim(x_n)$, como $x_{n+1} = f(x_n)$, está claro que u es un punto fijo de f. Además, se cumple

$$\|u - x_n\| \leq \frac{C^{n-1}}{1-C}\,\|x_2 - x_1\|\,,$$

lo que nos da la velocidad de convergencia.

Por último, demostraremos que sólo hay un punto fijo para f. De hecho, si u, v son dos puntos fijos distintos, de f, entonces

$$\|u - v\| = \|f(u) - f(v)\| \leq C\,\|u - v\|\,.$$

Puesto que $u \neq v$, entonces $\|u - v\| \neq 0$, por lo que esta relación implica que $1 \leq C$, contrario a la hipótesis de que $C < 1$. ∎

Notamos que lo que realmente se ha hecho es demostrar el siguiente resultado:

Corolario 37 *Si f es una contracción con $C < 1$, si x_1 es un punto arbitrario en $\mathbb{R}^p$, y si la sucesión $X = (x_n)$ está definida por la ecuación $x_{n+1} = f(x_n)$, entonces X converge al punto fijo único u de f con la rapidez dada por $\|u - x_n\| \leq \frac{C^{n-1}}{1-C}\,\|x_2 - x_1\|$.*

Para el caso en que la función f no esté definida en todo $\mathbb{R}^p$, se deberá tener más cuidado para asegurar que la definición iterativa $x_{n+1} = f(x_n)$ de la sucesión se puede llevar a cabo y que los puntos permanezcan en el dominio de f. Aun cuando son posibles otras formulaciones, habrá que conformarse con la siguiente.

Teorema 260 *Supongamos que f es una contracción con C constante definida para $D(f) = \{x \in \mathbb{R}^p : \|x\| \leq B\}$ y que $\|f(0)\| \leq B(1-C)$. Entonces la sucesión*

$$x_1 = 0, x_2 = f(x_1), ..., x_{n+1} = f(x_n), ...$$

converge al punto fijo único de f que se encuentra en el conjunto $D(f)$.
Demostración: *En efecto, si $x \in D = D(f)$, entonces $\|f(x) - f(0)\| \leq C\,\|x - 0\| \leq CB$, por lo que se deduce que*

$$\|f(x)\| \leq \|f(0)\| + CB \leq (1-C)B + CB = B.$$

Por lo tanto, $f(D) \subseteq D$. De modo que la sucesión (x_n) se puede definir y permanece en D y la demostración del teorema del punto fijo para contracciones es aplicable. ∎

El teorema de contracción que se acaba de establecer tiene ciertas ventajas: es constructivo, el error de aproximación se puede calcular y garantiza un punto fijo único. Sin embargo, tiene la desventaja de que el requisito de que f sea una contracción es una restricción muy severa. Un hecho importante y profundo demostrado por primera vez en 1910 por L.E.J. Brouwer, es que cualquier función continua con dominio $D = \{x \in \mathbb{R}^p : \|x\| \leq B\}$ e imagen en D debe tener al menos un punto fijo.

Teorema 261 ***(Teorema del punto fijo de Brouwer)*** *Sea $B > 0$ y sea $D = \{x \in \mathbb{R}^p : \|x\| \leq B\}$. Entonces cualquier función continua con dominio D e imagen contenida en D tiene al menos un punto fijo.*

No demostraremos este teorema, ya que la demostración se alejaría mucho del tema.

13.5. Sucesiones de funciones continuas

En muchas ocasiones es necesario considerar una sucesión de funciones continuas. En esta sección se dan varios teoremas interesantes e importantes acerca de dichas sucesiones.

En esta sección se deberá aclarar la importancia de la convergencia uniforme. Recordamos que una sucesión (f_n) de funciones en un subconjunto $D \subset \mathbb{R}^p$ a $\mathbb{R}^q$ se dice que converge uniformemente en D a f si para toda $\varepsilon > 0$ existe una $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$ y $x \in D$, entonces $\|f_n(x) - f(x)\| < \varepsilon$. Esto es válido si y sólo si $\|f_n - f\|_D \to 0$, cuando (f_n) es una sucesión acotada.

13.5.1. Intercambio de límite y continuidad

Observamos que el límite de una sucesión de funciones continuas puede no ser continuo. Consideramos, por ejemplo, las funciones $f_n(x) = x^n$ para $n \in \mathbb{N}$, y $x \in [0, 1]$. Esta sucesión converge en $[0, 1]$ a la función f definida por

$$f(x) = \begin{cases} 0 & \text{si} \quad 0 \leq x < 1 \\ 1 & \text{si} \quad x = 1 \end{cases}$$

Por lo que, a pesar de la índole tan sencilla de las funciones continuas f_n, la función límite no es continua en el punto $x = 1$.

Aún cuando la extensión de la discontinuidad de la función límite en el ejemplo recién dado no es muy grande, es evidente que se pueden construir ejemplos más complicados en los que se obtenga una discontinuidad más extensa. Sería interesante investigar qué tan discontinuo puede ser el límite de una sucesión de funciones continuas; sin embargo, esta investigación se saldría mucho del campo. Más aún, para la mayoría de las aplicaciones es más importante encontrar condiciones adicionales que garanticen que la función límite es continua.

En seguida probaremos el hecho de que la convergencia uniforme de una sucesión de funciones continuas es suficiente para garantizar la continuidad de la función límite.

Teorema 262 *Sea $F = (f_n)$ una sucesión de funciones continuas con dominio $D \subset \mathbb{R}^p$ e imagen en $\mathbb{R}^q$ tal que converge uniformemente en D a la función f. Entonces f es continua en D.*

Demostración: *Como (f_n) converge uniformemente en D a f, dada $\varepsilon > 0$ existe un número natural $N = N(\varepsilon/3)$ tal que $\|f_N(x) - f(x)\| < \varepsilon/3$ para toda x en D. Para probar que f es continua en un punto a en D, observamos que*

$$\begin{aligned} \|f(x) - f(a)\| &= \|f(x) - f_N(x) + f_N(x) - f_N(a) + f_N(a) - f(a)\| \leq \\ &\leq \|f(x) - f_N(x)\| + \|f_N(x) - f_N(a)\| + \|f_N(a) - f(a)\| \leq \\ &\leq \frac{\varepsilon}{3} + \|f_N(x) - f_N(a)\| + \frac{\varepsilon}{3}. \end{aligned}$$

Dado que f_N es continua, existe un número $\delta = \delta(\varepsilon/3, a, f_N) > 0$ tal que si $\|x - a\| < \delta$ y $x \in D$, entonces $\|f_N(x) - f_N(a)\| < \varepsilon/3$. Por lo tanto, para tal x se tiene $\|f(x) - f(a)\| < \varepsilon$. Esto demuestra la continuidad de la función f en un punto arbitrario $a \in D$. ∎

Observación 142 *Aun cuando la convergencia uniforme de la sucesión de funciones continuas es suficiente para la continuidad de la función límite, no es necesaria. De modo que si (f_n) es una sucesión de funciones continuas que converge a una función continua f, no podemos deducir que la convergencia sea uniforme.*

Teorema 263 *Si (f_n) es una sucesión de funciones de $BC_{pq}(D)$ tal que*

$$\|f_n - f\|_D \to 0$$

entonces $f \in BC_{pq}(D)$.

13.5.2. Teoremas de aproximación

En muchas aplicaciones es conveniente "aproximar" funciones continuas por medio de funciones de naturaleza elemental. A pesar de que existen varias definiciones razonables que se pueden emplear para hacer más precisa la palabra "aproximar", una de las más naturales, así como importante, es la que exige que en todo punto del dominio dado la función de aproximación no debe diferir de la función dada en más del error prefijado. A este juicio algunas veces se le conoce como "aproximación uniforme" y está íntimamente ligado a la convergencia uniforme.

Definición 129 *Suponiendo que f es una función dada con dominio $D = D(f)$, contenido en $\mathbb{R}^p$ y con imagen en $\mathbb{R}^q$, se dice que una función g* ***aproxima a*** *f* ***uniformemente en*** *D* ***en*** *$\varepsilon > 0$, si*

$$\|g(x) - f(x)\| \leq \varepsilon \quad \text{para toda } x \in D;$$

o de modo equivalente si

$$\|g - f\|_D = \sup\{\|g(x) - f(x)\| : x \in D\} \leq \varepsilon.$$

Se dice que la función f se puede ***aproximar uniformemente en*** *D por funciones de una clase $\mathcal{G}$ si para cada número $\varepsilon > 0$ hay una función g_ε en $\mathcal{G}$ tal que $\|g_\varepsilon - f\|_D < \varepsilon$; o, lo que es lo mismo, si existe una sucesión de funciones en $\mathcal{G}$ que converge uniformemente en D a f.*

Definición 130 *A una función g con dominio $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$ se le llama una* ***función escalonada*** *si toma únicamente un número finito de valores distintos en $\mathbb{R}^q$, tomándose cada valor distinto de cero en un intervalo en $\mathbb{R}^p$.*

Por ejemplo, si $p = q = 1$, entonces la función g definida explícitamente por

$$g(x) = \begin{cases} 0 & si \quad x \leq -2 \\ 1 & si \quad -2 < x \leq 0 \\ 3 & si \quad 0 < x < 1 \\ -5 & si \quad 1 \leq x \leq 3 \\ 0 & si \quad x > 3 \end{cases}$$

es una función escalonada.

En seguida se demuestra que una función continua cuyo dominio es una celda compacta se puede aproximar uniformemente por medio de funciones escalonadas.

Teorema 264 *Sea f una función continua cuyo dominio D es una celda compacta en $\mathbb{R}^p$ y cuyos valores pertenecen a $\mathbb{R}^q$. Entonces f se puede aproximar uniformemente en D por medio de funciones escalonadas.*

Demostración: *Sea $\varepsilon > 0$ dado; como f es uniformemente continua, existe un número $\delta(\varepsilon) > 0$ tal que si x, y cumplen $\|x - y\| < \delta(\varepsilon)$, entonces $\|f(x) - f(y)\| < \varepsilon$. Dividimos el dominio D de f en celdas disjuntas $I_1, ..., I_n$ tales que si x, y pertenecen a I_k, entonces $\|x - y\| < \delta(\varepsilon)$. Sea x_k cualquier punto perteneciente a la celda I_k, $k = 1, ..., n$ y definimos $g_\varepsilon(x) = f(x_k)$ para $x \in I_k$ y $g_\varepsilon(x) = 0$ para $x \notin D$. Está claro entonces que $\|g_\varepsilon(x) - f(x)\| < \varepsilon$ para $x \in D$ de tal manera que g_ε aproxima a f uniformemente en D en no más de ε.* ■

Es natural suponer que una función continua se puede aproximar uniformemente por funciones simples que también sean continuas (las funciones escalonadas no lo son). Para hacerlo más sencillo, el siguiente resultado se dará sólo para el caso en que $p = q = 1$ aunque es evidente que hay una generalización para dimensiones más altas.

Definición 131 *Se dice que una función g definida en un intervalo compacto $J = [a, b]$ de $\mathbb{R}$ con valores en $\mathbb{R}$ es* ***lineal por partes*** *si hay un número finito de puntos c_k con $a = c_0 < c_1 < c_2 < \cdots < c_n = b$ y números reales correspondientes A_k, B_k, $k = 0, 1, ..., n$, tales que cuando x cumple $c_{k-1} < x < c_k$, la función g es de la forma*

$$g(x) = A_k x + B_k, \quad k = 0, 1, ..., n.$$

Si g es continua en J, las constantes A_k, B_k deben satisfacer, desde luego, ciertas relaciones.

Teorema 265 *Sea f una función continua cuyo dominio es un intervalo compacto J en $\mathbb{R}$. Entonces f se puede aproximar uniformemente en J por medio de funciones continuas lineales por partes.*

Demostración: *Como en el caso anterior, f es uniformemente continua en el conjunto compacto J. Por lo tanto, dada $\varepsilon > 0$, se divide $J = [a, b]$ en subintervalos agregando puntos intermedios c_k, $k = 0, 1, ..., n$, con $a = c_0 < c_1 < c_2 < \cdots < c_n = b$ de tal manera que $c_k - c_{k-1} < \delta(\varepsilon)$. Unimos los puntos $(c_k, f(c_k))$ mediante segmentos de línea y definimos la función continua lineal por partes g que resulta. Está claro que g_ε aproxima a f uniformemente en J en no más de ε.* ■

13.5.3. Aproximación por polinomios

Demostraremos ahora un resultado más profundo, útil e interesante que se refiere a la aproximación por polinomios. Primero probaremos el teorema de aproximación de Weierstrass para $p = q = 1$, usando los polinomios de Bernstein.

Definición 132 *Sea f una función con dominio $\mathbf{I} = [0,1]$ e imagen en $\mathbb{R}$. El enésimo polinomio de Bernstein para f se define como*

$$B_n(x) = B_n(x; f) = \sum_{k=0}^{n} f(\frac{k}{n}) \binom{n}{k} x^k (1-x)^{n-k}.$$

Estos polinomios de Bernstein no son tan complicados como parecen a primera vista. El lector que tenga cierta experiencia en probabilidad podrá ver que detrás de esto está el binomio de distribución. Aun sin este tipo de experiencia, el lector deberá observar que el valor $B_n(x; f)$ del polinomio en el punto x se calcula a partir de los valores $f(0), f(1/n), f(2/n), ..., f(1)$, con ciertos pesos no negativos $\varphi_k(x) = \binom{n}{k} x^k (1-x)^{n-k}$ que se podrán ver muy pequeños para aquellos valores de k para los cuales k/n está lejos de x. De hecho, la función φ_k es no negativa en $\mathbf{I}$ y toma su valor máximo en el punto k/n. Más aún, como se verá en seguida, la suma de todas las $\varphi_k(x)$, $k = 0, 1, ..., n$ es 1 para cada $x \in \mathbf{I}$.

Recordamos que el teorema del binomio dice que

$$(s+t)^n = \sum_{k=0}^{n} \binom{n}{k} s^k t^{n-k}, \qquad (1)$$

donde $\binom{n}{k}$ denota al coeficiente binomial

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Directamente se puede observar que

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{k}{n}\binom{n}{k}, \qquad (2)$$

$$\binom{n-2}{k-2} = \frac{(n-2)!}{(k-2)!(n-k)!} = \frac{k(k-1)}{n(n-1)}\binom{n}{k}. \qquad (3)$$

Ahora, si $s = x$ y $t = 1 - x$, sustituyendo en (1) se obtiene

$$1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}. \qquad (4)$$

Escribiendo (4) reemplazando n por $n-1$ y k por j, se tiene

$$1 = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j}.$$

Multiplicando esta última relación por x y aplicar la igualdad (2), se obtiene

$$x = \sum_{j=0}^{n-1} \frac{j+1}{n} \binom{n}{j+1} x^{j+1} (1-x)^{n-(j+1)}.$$

Ahora si $k = j + 1$, entonces

$$x = \sum_{k=1}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k}.$$

Observamos también que el término correspondiente a $k = 0$ se puede incluir, ya que es cero. Por lo tanto, se tiene

$$x = \sum_{k=0}^{n} \frac{k}{n} \binom{n}{k} x^k (1-x)^{n-k}. \qquad (5)$$

Un cálculo análogo basado en (4) con n reemplazada por $n-2$ y la igualdad (3) prueba que

$$(n^2 - n)x^2 = \sum_{k=0}^{n} (k^2 - k) \binom{n}{k} x^k (1-x)^{n-k}.$$

Por lo tanto, se concluye que

$$(1 - \frac{1}{n})x^2 + \frac{1}{n}x = \sum_{k=0}^{n} (\frac{k}{n})^2 \binom{n}{k} x^k (1-x)^{n-k}. \qquad (6)$$

Multiplicando (4) por x^2, (5) por $-2x$, y sumándolas a (6), se obtiene

$$\frac{1}{n}\, x\,(1-x) = \sum_{k=0}^{n} (x - \frac{k}{n})^2 \binom{n}{k} x^k (1-x)^{n-k}, \qquad (7)$$

que es cálculo que se necesitará más adelante.

Examinando la definición del polinomio de Bernstein, la fórmula (4) dice que el enésimo polinomio de Bernstein para la función constante $f_0(x) = 1$ coincide con f_0. La fórmula (5) dice lo mismo para la función $f_1(x) = x$. La fórmula (6) asegura que el enésimo polinomio de Berstein para la función $f_2(x) = x^2$ es

$$B_n(x; f_2) = (1 - \frac{1}{n})x^2 + \frac{1}{n}\, x,$$

que converge uniformemente en $\mathbf{I}$ a f_2. A continuación demostraremos que si f es cualquier función continua de $\mathbf{I}$ a $\mathbb{R}$ entonces la sucesión de polinomios de Bernstein tiene la propiedad de que converge uniformemente en $\mathbf{I}$ a f. Esto nos dará una demostración constructiva del teorema de aproximación de Weierstrass. En el desarrollo de esta demostración será necesaria la fórmula (7).

Teorema 266 ***(Teorema de aproximación de Bernstein)*** *Sea f continua en $\mathbf{I}$ con valores en $\mathbb{R}$. Entonces la sucesión de polinomios de Bernstein para f converge uniformemente en $\mathbf{I}$ a f.*
Demostración: *Al multiplicar la fórmula (4) por $f(x)$ se obtiene*

$$f(x) = \sum_{k=0}^{n} f(x) \binom{n}{k} x^k (1-x)^{n-k}.$$

Por lo tanto, se obtiene la relación

$$f(x) - B_n(x) = \sum_{k=0}^{n} \left[f(x) - f(\frac{k}{n}) \right] \binom{n}{k} x^k (1-x)^{n-k}.$$

de donde se deduce que

$$|f(x) - B_n(x)| \leq \sum_{k=0}^{n} \left| f(x) - f(\frac{k}{n}) \right| \binom{n}{k} x^k (1-x)^{n-k}. \qquad (8)$$

f está acotada, digamos que por M, y también es uniformemente continua. Observe que si k es tal que k/n está cerca de x entonces el término correspondiente en la suma (8) es pequeño por la continuidad de f en x; por otro lado, si k/n está lejos de x el factor en relación con f sólo se puede decir que es menor que $2M$ y cualquier pequeñez debe surgir de los otros factores. Por lo tanto, el camino a seguir es dividir (8) en dos partes: aquellos valores de k en donde $x - k/n$ es pequeño y aquellos en los que $x - k/n$ es grande.

Sea $\varepsilon > 0$ y sea $\delta(\varepsilon)$ como en la definición de continuidad uniforme de f. Resulta conveniente elegir n de tal magnitud que

$$n \geq \sup \left\{ (\delta(\varepsilon))^{-4}, M^2/\varepsilon^2 \right\}, \qquad (9)$$

y partir (8) en dos sumas. La suma tomada sobre aquellas k para las cuales $|x - k/n| < n^{-1/4} \leq \delta(\varepsilon)$ da

$$\sum_k |f(x) - f(\frac{k}{n})| \binom{n}{k} x^k (1-x)^{n-k} \leq \varepsilon \sum_{k=1}^{n} \binom{n}{k} x^k (1-x)^{n-k} = \varepsilon.$$

La suma tomada sobre aquellas k para las cuales $|x - k/n| \geq n^{-1/4}$, es decir $(x - k/n)^2 \geq n^{-1/2}$, se puede calcular usando la fórmula (7). Para esta parte de la suma en (8) se obtiene la cota superior

$$\begin{aligned} \sum_k 2M \binom{n}{k} x^k (1-x)^{n-k} &= 2M \sum_k \frac{(x-k/n)^2}{(x-k/n)^2} \binom{n}{k} x^k (1-x)^{n-k} \leq \\ &\leq 2M\sqrt{n} \sum_{k=1}^{n} (x-k/n)^2 \binom{n}{k} x^k (1-x)^{n-k} \leq \\ &\leq 2M\sqrt{n} \left\{ \frac{1}{n} x(1-x) \right\} \leq \frac{M}{2\sqrt{n}}, \end{aligned}$$

ya que $x(1-x) \leq \frac{1}{4}$ en el intervalo $\mathbf{I}$. Recordando la resolución (9) para n, se concluye que cada una de estas dos partes de (8) está acotada por arriba por ε. Por lo tanto, para alguna n elegida en (9) se tiene

$$|f(x) - B_n(x)| < 2\varepsilon,$$

independientemente del valor de x. Esto prueba que la sucesión (B_n) converge uniformemente en $\mathbf{I}$ a f. ∎

Como corolario directo del teorema de Bernstein se tiene el siguiente resultado importante.

Teorema 267 ***(Teorema de aproximación de Weierstrass)*** *Sea f una función continua en un intervalo compacto de $\mathbb{R}$ y con valores en $\mathbb{R}$. Entonces f se puede aproximar uniformemente por polinomios.*
Demostración: *Si f está definida en $[a,b]$, entonces la función g definida en $\mathbf{I} = [0,1]$ por medio de*

$$g(t) = f((b-a)t + a), \qquad t \in \mathbf{I},$$

es continua. Por lo tanto, g se puede aproximar uniformemente por polinomios de Bernstein y un simple cambio de variable da una aproximación polinomial de f. ■

Se han mostrado los detalles de la demostración del teorema de Bernstein porque proporciona un método constructivo para encontrar una sucesión de polinomios que converja uniformemente en $\mathbf{I}$ a la función continua dada. Además, para demostrar algunos resultados más generales de aproximación más adelante será necesario saber que la función valor absoluto se puede aproximar uniformemente en un intervalo compacto por polinomios, y la demostración directa de este resultado no es tan sencilla.

13.6. Límites de funciones

En esta sección nos ocuparemos del límite de una función en un punto y de algunas ligeras variaciones del mismo tema. Por lo general, este tema se estudia antes de continuidad; de hecho, la misma definición de una función continua en ocasiones se expresa en términos de este límite en vez de usar la definición que vimos. Una de las razones por las que se eligió estudiar por separado continuidad de límite es que se habrán de introducir dos definiciones ligeramente distintas del límite de una función en un punto. Debido a que las dos definiciones se usan bastante, se ofrecerán ambas tratando de relacionarlas entre sí.

A menos que se especifique lo contrario, f será una función con dominio D, contenido en $\mathbb{R}^p$ y con valores en $\mathbb{R}^q$ y se habrá de considerar al límite de f en un punto de acumulación c de D. Por lo tanto, todo entorno de c contiene una infinidad de puntos de D.

Definición 133 *Se dice que un elemento $b \in \mathbb{R}^q$ es el* ***límite de f en c*** *si para todo entorno V de b hay un entorno U de c tal que si x pertenece a $U \cap D$ y $x \neq c$, entonces $f(x)$ pertenece a V. En este caso se escribe*

$$b = \lim_c f \qquad \text{ó} \qquad b = \lim_{x \to c} f(x).$$

La unicidad del límite, cuando existe, se plantea con facilidad. Será suficiente la siguiente afirmación.

Lema 12 *Si $\lim_c f$ existe, entonces está determinado de manera única.*

Se establecerán ahora algunas condiciones necesarias y suficientes para la existencia del límite. Se deberá observar que en la parte (c) el límite se refiere al límite de una sucesión.

Teorema 268 *Las siguientes afirmaciones son equivalentes.*

a. *El límite* $b = \lim_c f$ *existe.*

b. *Si* $\varepsilon > 0$, *existe* $\delta > 0$ *tal que si* $x \in D$ *y* $0 < \|x - c\| < \delta$, *entonces* $\|f(x) - b\| < \varepsilon$.

c. *Si* (x_n) *es cualquier sucesión en* D *tal que* $x_n \neq c$ *y* $c = \lim(x_n)$, *entonces* $b = \lim(f(x_n))$.

Demostración: *Veamos en primer lugar:* **i** $\Longleftrightarrow$ **ii**

$\Rightarrow$) *Supongamos que* $b = \lim_c f$ *existe. Entonces, para todo* $\varepsilon > 0$, *consideramos el entorno* $V = D(b;\varepsilon)$ *de* b. *Por hipótesis, existe un entorno* U *de* c *tal que para toda* $x \in U \cap D, x \neq c$, *el valor* $f(x)$ *pertenece a* V. *Puesto que* U *es un entorno, existe* $\delta > 0$ *tal que* $D(x;\delta) \subset U$. *Sin embargo,* $x \neq c$ *está en* $D(x,\delta)$ *si y sólo si* $0 < \|x - c\| < \delta$. *Asimismo,* $f(x)$ *pertenece a* V *si y sólo si* $\|f(x) - f(c)\| < \varepsilon$. *Por tanto, si* $x \in D$ *satisface* $0 < \|x - c\| < \delta$, *entonces* $f(x)$ *satisface* $\|f(x) - f(c)\| < \varepsilon$.

$\Leftarrow$) *Supongamos que se cumple la condición enunciada en* **ii**. *Dado* $\varepsilon > 0$ *cualquiera, consideramos los entornos* $U = D(c;\delta)$ *y* $V = D(b;\varepsilon)$. *Entonces la condición* **ii** *indica que si* x *está en* $D \cap U, x \neq c$, *entonces* $f(x)$ *pertenece a* V. *Por lo tanto, el límite* $b = \lim_c f$ *existe.*

Veamos ahora: **i** $\Longleftrightarrow$ **iii**

$\Rightarrow$) *Supongamos que* b *es el límite de* f *en* c. *Sea* $(x_n), x_n \neq c$, *una sucesión en* D *con* $\lim(x_n) = c$. *Se debe probar que la sucesión* $(f(x_n))$ *converge a* b. *Sea* $\varepsilon > 0$ *dada. Entonces, por* **ii** *(podemos usarlo porque ya hemos demostrado que* **i** $\Longleftrightarrow$ **ii***), existe* $\delta > 0$ *tal que si* x *satisface* $0 < \|x - c\| < \delta$, *donde* $x \in D$, *entonces* $f(x)$ *satisface* $\|f(x) - b\| < \varepsilon$. *Se aplica ahora la definición de sucesión convergente de la* δ *dada para obtener un número natural* $K(\delta)$ *tal que si* $n > K(\delta)$ *entonces* $0 < \|x_n - c\| < \delta$. *Pero para cada una de estas* x_n *se tiene* $\|f(x_n) - b\| < \varepsilon$. *Por tanto, si* $n > K(\delta)$, *entonces* $\|f(x_n) - b\| < \varepsilon$. *Por lo tanto, la sucesión* $(f(x_n))$ *converge a* b.

$\Leftarrow$) *[La demostración es un razonamiento contrarrecíproco.] Si la igualdad de* **i** *no se cumple, entonces existe un entorno* V *de* b *tal que, independientemente del entorno* U *de* c *que se elija, habrá al menos un número* x_0 *en* $D \cap U$, $x_0 \neq c$, *tal que* $f(x_0) \notin V$. *Por tanto, para toda* $n \in \mathbb{N}$, *el entorno* $D(c; \frac{1}{n})$ *de* c *contiene un número* $x_n \neq c$ *tal que*

$$0 < \|x_n - c\| < \frac{1}{n} \quad \text{y} \quad x_n \in D,$$

pero tal que

$$\|f(x_n) - b\| \geq \varepsilon_0 \quad \text{para toda} \quad n \in \mathbb{N}$$

Se concluye que la sucesión (x_n) *en* D *converge a* c, *pero la sucesión* $(f(x_n))$ *no converge a* $f(c)$. *Por lo tanto, se ha demostrado que si* **i** *no es verdadera, entonces* **iii** *tampoco lo es. Se concluye que* **iii** *implica* **i**. ■

Teorema 269 *Si* c *es un punto de acumulación que pertenece al dominio* D *de* f, *entonces las siguientes afirmaciones son equivalentes.*

a. *La función* f *es continua en* c.

b. *El límite* $\lim_c f$ *existe y es igual a* $f(c)$.

Demostración: *Si (a) es válido y V es un entorno de $f(c)$, entonces existe un entorno U de c tal que si x pertenece a $U \cap D$, entonces $f(x)$ pertenece a V. Como $f(x)$ pertenece a V para toda $x \neq c$ para la cual $x \in U \cap D$, entonces $\lim f$ existe y es igual a $f(c)$. A la inversa, es fácil ver que la afirmación (b) implica (a).* ■

Si f y g son dos funciones que tienen límite en un punto de acumulación c de $D(f+g) = D(f) \cap D(g)$, entonces la suma $f+g$ tiene límite en c y

$$\lim_c (f+g) = \lim_c f + \lim_c g.$$

Resultados análogos son válidos para otras combinaciones algebraicas de funciones, como se puede ver fácilmente. El siguiente resultado, referente a la composición de dos funciones, es ligeramente más profundo.

Teorema 270 *Supongamos que f tiene dominio $D(f)$ en $\mathbb{R}^p$ e imagen en $\mathbb{R}^q$ y que g tiene dominio $D(g)$ en $\mathbb{R}^q$ e imagen en $\mathbb{R}^r$. Sea $g \circ f$ la composición de g y f y sea c un punto de acumulación de $D(g \circ f)$. Entonces se cumple:*

Si ambos límites $b = \lim_c f$ y $a = \lim_b g$ existen y si g es continua en b o bien $f(x) \neq b$ para x en un entorno de c, entonces el límite de $g \circ f$ existe en c y $a = \lim_c g \circ f$.

Demostración: *Sea W un entorno de a en $\mathbb{R}^r$; dado que $a = \lim_b g$, hay un entorno V de b tal que si $y \in V \cap D(g)$ e $y \neq b$, entonces $g(y) \in W$. Como $b = \lim_c f$, hay un entorno U de c tal que si $x \in U \cap D(f)$ y $x \neq c$, entonces $f(x) \in V$. Por lo tanto, si x pertenece al conjunto, posiblemente más pequeño, $U \cap D(g \circ f)$, y $x \neq c$, entonces $f(x) \in V \cap D(g)$. Si $f(x) \neq b$ en algún entorno U_1 de c, se deduce que para $x \neq c$ en $(U_1 \cap U) \cap D(g \circ f)$, $(g \circ f)(x) \in W$, de tal manera que a es el límite de $g \circ f$ en c. Si g es continua en b, entonces $(g \circ f)(x) \in W$ para $U \cap D(g \circ f)$ y $x \neq c$.* ■

El teorema anterior puede no ser válido si se omite la condición de que g sea continua en b o que $f(x) \neq b$ en un entorno de c. Para concretar esta observación, sea f la siguiente función de $\mathbb{R}$ a $\mathbb{R}$:

$$f(x) = \begin{cases} 0 & \text{si} \quad x \neq 0 \\ 1 & \text{si} \quad x = 0 \end{cases}$$

y sean $g = f$ y $c = 0$. Entonces, $g \circ f$ está dada por

$$(g \circ f)(x) = \begin{cases} 1 & \text{si} \quad x \neq 0 \\ 0 & \text{si} \quad x = 0 \end{cases}$$

Más aún, se tiene $\lim_{x \to 0} f(x) = 0$, y $\lim_{y \to 0} g(y) = 0$, mientras que está claro que $\lim_{x \to 0} (g \circ f)(x) = 1$.

13.6.1. Límites superiores en un punto

En lo que queda de esta sección se habrá de considerar el caso en que $q = 1$. Así pues, f es una función con dominio D en $\mathbb{R}^p$ y valores en $\mathbb{R}$ y el punto $c \in \mathbb{R}^p$ es un punto de acumulación de D. Definiremos el límite por arriba o el límite superior de f en c. También se puede definir el límite inferior de manera análoga. Una cosa que se debe observar aquí es que, aun cuando la existencia del límite en $\mathbb{R}$ es un asunto relativamente delicado, los límites superiores que se habrán de definir tienen la propiedad de que si f está acotada, entonces su existencia está garantizada.

Definición 134 *Supongamos que f esta acotada en un entorno del punto c. Si $r > 0$, definimos $\varphi(r)$ de la forma*

$$\varphi(r) = \sup\{f(x) : 0 < \|x - c\| < r,\, x \in D\}$$

y establecemos

$$\operatorname*{lím\,sup}_{x \to c} f = \inf\{\varphi(r) : r > 0\},$$

$\operatorname{lím\,sup} f_{x \to c}$ *se llama **límite superior** de f en c.*

Debido a que esta cantidad está definida como el ínfimo de la imagen bajo f de entornos siempre decrecientes de c, probablemente no sea claro que merezca la terminología de "límite superior". En el siguiente lema se da una justificación de esta terminología.

Lema 13 *Si φ se define como antes, entonces*

$$\operatorname{lím\,sup} f_{x\to c} = \lim_{r\to 0} \varphi(r),$$

Demostración: *Observemos que si $0 < r < s$, entonces*

$$\operatorname*{lím\,sup}_{x\to c} f \le \varphi(r) \le \varphi(s).$$

Más aún, por la definición de límite superior, si $\varepsilon > 0$ existe una $r_\varepsilon > 0$ tal que

$$\varphi(r_\varepsilon) < \operatorname*{lím\,sup}_{x\to c} f + \varepsilon.$$

Por lo tanto, si r satisface $0 < r < r_\varepsilon$, se tiene $|\varphi(r) - \operatorname{lím\,sup}_{x\to c} f| < \varepsilon$, que demuestra

$$\operatorname{lím\,sup} f_{x\to c} = \lim_{r\to 0} \varphi(r).$$

■

Lema 14 *Si $M > \operatorname{lím\,sup}_{x\to c} f$, entonces existe un entorno U de c tal que*

$$f(x) < M \quad \text{para} \quad x \in D \cap U,\, x \ne c.$$

Demostración: *Por la definición de límite superior, se tiene $\inf\{\varphi(r) : r > 0\} < M$. Por tanto, existe un número real $r_1 > 0$ tal que $\varphi(r_1) < M$ y se puede tomar*

$$U = \{x \in \mathbb{R}^p : \|x - c\| < r_1\}.$$

■

Lema 15 *Sean f y g funciones acotadas en un entorno de c y supongamos que c es un punto de acumulación de $D(f+g)$. Entonces*

$$\operatorname*{lím\,sup}_{x\to c}(f+g) \le \operatorname*{lím\,sup}_{x\to c} f + \operatorname*{lím\,sup}_{x\to c} g$$

Demostración: *Por la relación*

$$\sup\{f(x)+g(x) : x \in A\} \leq \sup\{f(x) : x \in A\} + \sup\{g(x) : x \in A\},$$

está claro que usando la notación de la definición de límite superior se tiene

$$\varphi_{f+g}(r) \leq \varphi_f(r) + \varphi_g(r).$$

Haciendo $r \to 0$ se obtiene

$$\limsup_{x\to c}(f+g) \leq \limsup_{x\to c} f + \limsup_{x\to c} g.$$

■

13.7. Otros resultados

En esta sección plantearemos algunos teoremas que no se aplicarán posteriormente pero que a menudo son útiles en topología y análisis. Los primeros resultados son extensiones (de largo alcance) del teorema de aproximación de Weierstrass, después hay un teorema que ofrece condiciones en que una funciónn continua tiene una extensión continua y el último resultado es análogo al de Bolzano-Weierstrass en el espacio $C_{pq}(K)$ de funciones continuas en un conjunto compacto K.

13.7.1. El teorema de Stone-Weierstrass

Para facilitar el análisis se introduce la siguiente terminología.

Definición 135 *Si f y g son funciones con dominio D en $\mathbb{R}^p$ y con valores en $\mathbb{R}$ entonces las funciones h y k, definidas para x en D por medio de*

$$h(x) = \sup\{f(x), g(x)\}, \qquad k(x) = \inf\{f(x), g(x)\},$$

reciben el nombre de ***supremo*** *e* ***ínfimo****, respectivamente, de las funciones f y g.*

Observación 143 *Si f y g son continuas en D, entonces h y k también son continuas. Esto se deduce de la sección Propiedades locales de las funciones continuas y la observación de que si a, b son números reales entonces*

$$\begin{aligned} \sup\{a,b\} &= \frac{1}{2}\{a+b+|a-b|\}, \\ \inf\{a,b\} &= \frac{1}{2}\{a+b-|a-b|\}. \end{aligned}$$

A continuación se demuestra un modelo de la generalización de Stone al teorema de aproximación de Weierstrass.

Teorema 271 ***(Teorema de aproximación de Stone)*** *Sea K un subconjunto compacto de $\mathbb{R}^p$ y sea $\mathcal{L}$ una colección de funciones continuas de K a $\mathbb{R}$ con las siguientes propiedades:*

a. *Si f y g pertenecen a $\mathcal{L}$, entonces* $\sup\{f,g\}$ *e* $\inf\{f,g\}$ *pertenecen a $\mathcal{L}$.*

b. *Si* $b \in \mathbb{R}$ *y* $x \neq y \in K$, *con* x, *entonces existe una función* f *en* $\mathcal{L}$ *tal que* $f(x) = a$, $f(y) = b$.

Entonces cualquier función continua de K *a* $\mathbb{R}$ *se puede aproximar uniformemente en* K *por funciones de* $\mathcal{L}$.

Demostración: *Sea* F *una función continua de* K *a* $\mathbb{R}$. *Si* x, y *pertenecen a* K, *sea* $g_{xy} \in \mathcal{L}$ *tal que* $g_{xy}(x) = F(x)$ *y* $g_{xy}(y) = F(y)$. *Dado que las funciones* F, g_{xy} *son continuas y tienen el mismo valor en* y, *dada* $\varepsilon > 0$, *hay un entorno abierto* $U(y)$ *de* y *tal que si* z *pertenece a* $K \cap U(y)$, *entonces*

$$g_{xy}(z) > F(z) - \varepsilon.$$

Fijando x, *para cada* $y \in K$ *elegimos un entorno abierto* $U(y)$ *con esta propiedad. Por la compacidad de* K, *se deduce que* K *está contenido en la unión de un número finito de entornos:* $U(y_1), ..., U(y_n)$. *Si* $h_x = \sup\{g_{xy_1}, ..., g_{xy_n}\}$, *entonces se deduce que*

$$h_x(z) > F(z) - \varepsilon \quad \text{para} \quad z \in K.$$

Dado que $g_{xy_1}(x) = F(x)$, *se puede ver que* $h_x(x) = F(x)$ *y, por lo tanto, hay un entorno abierto* $V(x)$ *de* x *tal que si* z *pertenece a* $K \cap V(x)$, *entonces*

$$h_x(z) < F(z) + \varepsilon$$

Usando una vez más la compacidad de K *se obtiene un número finito de entornos en cuya unión está contenido* $K : V(x_1), ..., V(x_m)$. *Sea* $h = \inf\{h_{x_1}, ..., h_{x_m}\}$. *Entonces* h *pertenece a* $\mathcal{L}$ *y se deduce que*

$$h(z) > F(z) - \varepsilon \quad \text{para} \quad z \in K.$$

y que

$$h(z) < F(z) + \varepsilon \quad \text{para} \quad z \in K.$$

Combinando estos resultados se tiene

$$|h(z) - F(z)| < \varepsilon \quad \text{para} \quad z \in K,$$

lo que da la aproximación deseada. ■

El lector puede observar que en este último resultado no se ha utilizado el teorema de aproximación de Weierstrass. En el siguiente resultado se sustituye la condición (a) por tres condiciones algebraicas en el conjunto de funciones. En esta ocasión se utiliza el teorema de Weierstrass para el caso especial de la función valor absoluto φ definida para $t \in \mathbb{R}$ como $\varphi(t) = |t|$, para concluir que φ se puede aproximar por polinomios en todo conjunto compacto de números reales.

Teorema 272 ***(Teorema de Stone-Weierstrass)*** *Sea* K *un subconjunto compacto de* $\mathbb{R}^p$ *y sea* $\mathcal{A}$ *una colección de funciones continuas de* K *a* $\mathbb{R}$ *con las siguientes propiedades:*

a. *La función constante* $e(x) \to 1$, $x \in K$, *pertenece a* $\mathcal{A}$.

b. *Si* f, g *pertenecen a* $\mathcal{A}$, *entonces* $\alpha f + \beta g$ *pertenecen a* $\mathcal{A}$ *para toda* $\alpha, \beta \in \mathbb{R}$.

c. *Si f, g pertenecen a $\mathcal{A}$, entonces fg pertenece a $\mathcal{A}$.*

d. *Si $x \neq y$ son dos puntos de K, existe una función f en $\mathcal{A}$ tal que $f(x) \neq f(y)$.*

Entonces, cualquier función continua de K a $\mathbb{R}$ se puede aproximar uniformemente en K por funciones en $\mathcal{A}$.

Demostración: *Sean $a, b \in \mathbb{R}$ y $x \neq y$ tales que pertenezcan a K. De acuerdo con (d), hay una función f en $\mathcal{A}$ tal que $f(x) \neq f(y)$. Dado que $e(x) = 1 = e(y)$, se deduce que hay números reales α, β tales que*

$$\alpha f(x) + \beta e(x) = a, \qquad \alpha f(y) + \beta e(y) = b.$$

Por lo tanto, por (b) existe una función $g \in a$ tal que $g(x) = a$ y $g(y) = b$.

Ahora, sea $\mathcal{L}$ el conjunto de todas las funciones continuas en K que se pueden aproximar uniformemente por funciones en $\mathcal{A}$. Es obvio que $\mathcal{A} \subseteq \mathcal{L}$, de tal manera que $\mathcal{L}$ tiene la propiedad (b) del teorema de aproximación de Stone. Demostraremos ahora que si $h \in \mathcal{L}$, entonces $|h| \in \mathcal{L}$. Dado que

$$\sup\{f, g\} = \frac{1}{2}(f + g + |f - g|),$$

$$\inf\{f, g\} = \frac{1}{2}(f + g - |f - g|),$$

esto implicará que $\mathcal{L}$ tiene la propiedad (a) del teorema de aproximación de Stone y por lo tanto que toda función continua de K a $\mathbb{R}$ pertenece a $\mathcal{L}$.

Como h es continua y K es compacto, se deduce que existe una $M > 0$ tal que $\|h\|_K \leq M$. Dado que $h \in \mathcal{L}$, hay una sucesión (h_n) de funciones en $\mathcal{A}$ que convergen uniformemente a h en K y se puede suponer que $\|h_n\|_K \leq M + 1$ para toda $n \in \mathbb{N}$. Si $\varepsilon > 0$ está dada, se aplica el teorema de aproximación de Weierstrass a la función valor absoluto en el intervalo $[-(M+1), M+1]$ para obtener un polinomio p_ε tal que

$$||t| - p_\varepsilon(t)| \leq \frac{1}{2}\varepsilon \qquad \textit{para} \quad |t| \leq M + 1.$$

Se deduce por tanto que

$$||h_n(x)| - p_\varepsilon(h_n(x))| \leq \frac{1}{2}\varepsilon \qquad \textit{para} \quad x \in \mathbf{K}.$$

Pero $p_\varepsilon \circ f_n$ pertenece a $\mathcal{A}$ debido a las hipótesis (a), (b) y (c). Como

$$||h(x)| - |(h_n(x))|| \leq \|h - h_n\|_{K,} .$$

se deduce que si n es suficientemente grande entonces se tiene

$$||h(x)| - p_\varepsilon \circ h(x)| \leq \varepsilon \qquad \textit{para} \quad x \in \mathbf{K}.$$

Dado que $\varepsilon > 0$ es arbitrario, se deduce que $|h| \in \mathcal{L}$ y ahora el resultado se deduce del teorema de aproximación de Stone. ■

Como caso especial del teorema de Stone-Weierstrass obtendremos una forma más fuerte del teorema de aproximación de Weierstrass. Este resultado es más fuerte que el anterior por dos razones: (i) permite que el dominio sea un subconjunto compacto arbitrario de $\mathbb{R}^p$ y no únicamente una intervalo compacto en $\mathbb{R}$ y (ii) permite que la imagen esté en cualquier espacio $\mathbb{R}^q$, y no sólo en $\mathbb{R}$.

Definición 136 *Una función f con dominio $D \subset \mathbb{R}^p$ e imagen en $\mathbb{R}^q$ se puede considerar como q funciones de D a $\mathbb{R}$ con la representación de coordenadas:*

$$f(x) = (f_1(x), ..., f_q(x)) \qquad \text{para } \ x \in D.$$

Si cada función coordenada f_j es un polinomio en las p coordenadas $(x_1, ..., x_p)$, entonces se dice que f es una ***función polinomial****.*

Teorema 273 ***(Teorema de aproximación polinomial)*** *Sea f una función continua cuyo dominio K es un subconjunto compacto de $\mathbb{R}^p$ y cuya imagen pertenece a $\mathbb{R}^q$ y sea $\varepsilon > 0$. Entonces existe una función polinomial p de $\mathbb{R}^p$ a $\mathbb{R}^q$ tal que $\|f(x) - p(x)\| < \varepsilon$ para $x \in K$.*
Demostración: *Representar a f por medio de sus q funciones coordenadas:*

$$f(x) = (f_1(x), ..., f_q(x)) \qquad \text{para } \ x \in D.$$

Dado que f es continua en K, cada una de las funciones coordenadas f_i es continua de K a $\mathbb{R}$. Es evidente que las funciones polinomiales definidas de $\mathbb{R}^p$ a $\mathbb{R}$ satisfacen las propiedades del teorema de Stone-Weierstrass. Por lo tanto, la función coordenada f_j se puede aproximar uniformemente en K en $\varepsilon/\sqrt{q}$ por una función polinomial p_j. Si definimos p por

$$p(x) = (p_1(x), ..., p_q(x)),$$

obtenemos una función polinomial de $\mathbb{R}^p$ a $\mathbb{R}^q$ que da la aproximación deseada en K a la función dada f. ■

13.7.2. Extensión de funciones continuas

Algunas veces es deseable extender el dominio de una función continua a un conjunto más grande sin cambiar los valores del dominio original. Esto siempre se puede hacer de una manera trivial estableciendo que la función sea 0 fuera del dominio original, pero, en general, este método de extensión no da una función continua. Después de reflexionar un poco, el lector podrá ver que no siempre es posible obtener una extensión continua. Por ejemplo, si $D = \{x \in \mathbb{R}: x \neq 0\}$ y si f está definida para $x \in D$ como $f(x) = 1/x$, entonces no es posible extender a f de tal manera que se obtenga una función continua en todo $\mathbb{R}$. Sin embargo, es importante saber que una extensión siempre es posible cuando el dominio es un conjunto cerrado. Además, no es necesario incrementar la cota de la función (si es que es acotada).

Observación 144 *Si A y B son dos subconjuntos cerrados disjuntos de $\mathbb{R}^p$, entonces existe una función continua φ definida en $\mathbb{R}^p$ con valores de $\mathbb{R}$ tal que*

$$\begin{aligned} \varphi(x) &= 0 \quad \text{si } x \in A \\ \varphi(x) &= 1 \quad \text{si } x \in B \\ 0 \leq \varphi(x) &\leq 1 \quad \text{para todo } x \in \mathbb{R}^p. \end{aligned}$$

De hecho, si $d(x, A) = \inf\{\|x - y\| : y \in A\}$ y $d(x, B) = \inf\{\|x - y\| : y \in B\}$, entonces se puede definir φ para $x \in \mathbb{R}^p$ por medio de la ecuación

$$\varphi(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}.$$

Teorema 274 ***(Teorema de extensión de Tietze)*** *Sea f una función continua acotada definida en un subconjunto cerrado D de $\mathbb{R}^p$ y con valores en $\mathbb{R}$. Entonces existe una función continua g de $\mathbb{R}^p$ a $\mathbb{R}$ tal que $g(x) = f(x)$ para $x \in D$ y tal que* $\sup\{\|g(x)\| : x \in \mathbb{R}^p\} = \sup\{\|f(x)\| : x \in D\}$.
Demostración: *Sea* $M = \sup\{|f(x)| : x \in D\}$ *y definimos*

$$A_1 = \left\{x \in D : f(x) \leq -\frac{M}{3}\right\}$$

$$B_1 = \left\{x \in D : f(x) \geq \frac{M}{3}\right\}.$$

Por la continuidad de f y el hecho de que D sea cerrado, del teorema de continuidad global se deduce que A_1 y B_1 son subconjuntos cerrados de $\mathbb{R}^p$. De acuerdo con la observación que precede al planteamiento del teorema, hay una función continua φ_1 de $\mathbb{R}^p$ a $\mathbb{R}$ tal que

$$\begin{aligned}
\varphi_1(x) &= -\frac{1}{3}M \quad \text{si } x \in A_1 \\
\varphi_1(x) &= \frac{1}{3}M \quad \text{si } x \in B_1
\end{aligned}$$

$$-\frac{1}{3}M \leq \varphi_1(x) \leq \frac{1}{3}M \quad \text{para todo } x \in \mathbb{R}^p.$$

Fijamos $f_2 = f - \varphi_1$ y observamos que f_2 es continua en D y que

$$\sup\{|f_2(x)| : x \in D\} \leq \frac{2}{3}M.$$

Definimos

$$\begin{aligned}
A_2 &= \left\{x \in D : f_2(x) \leq -\frac{1}{3} \cdot \frac{2}{3}M\right\} \\
B_2 &= \left\{x \in D : f_2(x) \geq \frac{1}{3} \cdot \frac{2}{3}M\right\}
\end{aligned}$$

y se obtiene una función continua φ_2 de $\mathbb{R}^p$ a $\mathbb{R}$ tal que

$$\begin{aligned}
\varphi_2(x) &= -\frac{1}{3} \cdot \frac{2}{3}M \quad \text{si } x \in A_2 \\
\varphi_2(x) &= -\frac{1}{3} \cdot \frac{2}{3}M \quad \text{si } x \in B_2
\end{aligned}$$

$$-\frac{1}{3} \cdot \frac{2}{3}M \leq \varphi_2(x) \leq \frac{1}{3} \cdot \frac{2}{3}M, \quad \text{para todo } x \in \mathbb{R}^p.$$

A continuación, fijamos $f_3 = f_2 - \varphi_2$ y observamos que $f_3 = f - \varphi_1 - \varphi_2$ es continua en D y que $\sup\{|f_3(x)| : x \in D\} \leq (\frac{2}{3})^2 M$.

Siguiendo de esta manera se obtiene una sucesión (φ_n) de funciones definidas de $\mathbb{R}^p$ a $\mathbb{R}$ tales que, para cada n,

$$|f(x) - [\varphi_1(x) + \varphi_2(x) + \cdots + \varphi_n(x)]| \leq (\frac{2}{3})^n M,$$

para toda x en D y tal que

$$|\varphi_n(x)| \le (\frac{1}{3})(\frac{2}{3})^{n-1} M \quad \text{para toda } x \in \mathbb{R}^p.$$

Definimos g_n de $\mathbb{R}^p$ a $\mathbb{R}$ de la forma $g_n = \varphi_1 + \varphi_2 + \cdots + \varphi_n$, por lo que tenemos que g_n es continua. De la desigualdad anterior para $|\varphi_n(x)|$ se deduce que si $m \ge n$ y $x \in \mathbb{R}^p$, entonces

$$\begin{aligned} |g_m(x) - g_n(x)| &= \left|\varphi_{n+1}(x) + \cdots + \varphi_m(x)\right| \le \\ &\le (\frac{1}{3})(\frac{2}{3})^n M \left[1 + \frac{2}{3} + (\frac{2}{3})^2 + \cdots\right] \le \\ &\le (\frac{2}{3})^n M, \end{aligned}$$

lo cual prueba que la sucesión (g_n) converge uniformemente en $\mathbb{R}^p$ a una función que denotaremos por g. Dado que cada g_n es continua en $\mathbb{R}^p$, g también es continua en todo punto de $\mathbb{R}^p$. Además, se cumple

$$|f(x) - g_n(x)| \le (\frac{2}{3})^n M \quad \text{para } x \in D.$$

Por lo tanto, se concluye que $f(x) = g(x)$ para x en D. Además, para cualquier $x \in \mathbb{R}^p$ se tiene

$$|g_n(x)| \le \frac{1}{3} M \left[1 + \frac{2}{3} + \cdots + (\frac{2}{3})^{n-1}\right] \le M,$$

que prueba la última afirmación del teorema ■

Corolario 38 *Sea f una función continua acotada definida en un subconjunto cerrado D de $\mathbb{R}^p$ y con valores en $\mathbb{R}^q$. Entonces existe una función continua g de $\mathbb{R}^p$ a $\mathbb{R}^q$ con $g(x) = f(x)$ para x en D y tal que*

$$\sup\{\|g(x)\| : x \in \mathbb{R}^p\} \le \sqrt{q} \sup\{\|f(x)\| : x \in D\}.$$

Demostración: *Este resultado se acaba de probar para $q = 1$. Para el caso general, observamos que f define q funciones coordenadas continuas de valor real en D, digamos*

$$f(x) = (f_1(x), f_2(x), ..., f_q(x)).$$

Dado que cada una de las $f_j, 1 \le j \le q$ tiene una extensión continua g_j de $\mathbb{R}^p$ a $\mathbb{R}$, se define g de $\mathbb{R}^p$ a $\mathbb{R}^q$ por medio de $g(x) = (g_1(x), g_2(x), ..., g_q(x))$. Se puede ver que la función g tiene las propiedades requeridas. ■

13.7.3. Equicontinuidad

A continuación presentamos un teorema enteramente análogo al teorema de Bolzano-Weierstrass pero para conjuntos de funciones continuas y no para conjuntos de puntos. Para que sea más sencillo y breve, ofreceremos sólo una forma secuencial del teorema.

En lo sucesivo, K será un subconjunto compacto fijo de $\mathbb{R}^p$, y se estará tratando con funciones continuas en K con imagen en $\mathbb{R}^q$. Por el teorema de conservación de compacidad, cada una de estas funciones es acotada y por lo tanto $C_{pq}(K) = BC_{pq}(K)$.

Definición 137 *Se dice que un conjunto $\mathcal{F} \subset C_{pq}(K)$ es* ***acotado*** *(o* ***uniformemente acotado****) en K si existe una constante M tal que $\|f\|_K \leq M$, para toda $f \in \mathcal{F}$.*

Observación 145 *Está claro que cualquier conjunto finito $\mathcal{F}$ de dichas funciones es acotado, ya que si $F = \{f_1, f_2, ..., f_n\}$, entonces se puede fijar*

$$M = \sup \{\|f_1\|_K, \|f_2\|_K, ..., \|f_n\|_K\}.$$

En general, un conjunto infinito de funciones continuas de K a $\mathbb{R}^q$ no es acotado. Sin embargo, una sucesión uniformemente convergente de funciones continuas es acotada.

Si f es una función continua en un conjunto compacto $K \subset \mathbb{R}^p$, entonces el teorema de continuidad uniforme implica que es uniformemente continua. Por lo tanto, si $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$, tal que si x, y pertenecen a K y $\|x - y\| < \delta(\varepsilon)$, entonces $\|f(x) - f(y)\| < \varepsilon$. Desde luego, el valor de δ puede depender de la función f, así como de ε, por lo que a menudo se escribe $\delta(\varepsilon, f)$. (Cuando se está tratando con más de una función, es conveniente indicar explícitamente esta dependencia). Observemos que si $F = \{f_1, ..., f_n\}$ es un conjunto finito en $C_{pq}(K)$, entonces, fijando

$$\delta(\varepsilon, F) = \inf \{\delta(\varepsilon, f_1), ..., \delta(\varepsilon, f_n)\},$$

se obtiene una δ que "sirve" para todas las funciones de este conjunto finito.

Definición 138 *Se dice que un conjunto de funciones de K a $\mathbb{R}^q$ es uniformemente equicontinuo en K si para cada número real $\varepsilon > 0$ hay un número $\delta(\varepsilon) > 0$ tal que si x, y pertenecen a K y $\|x - y\| < \delta(\varepsilon)$ y f es una función de F, entonces $\|f(x) - f(y)\| < \varepsilon$.*

Ya hemos visto que un conjunto finito de funciones continuas en K es equicontinuo. También es cierto que una sucesión de funciones continuas que converge uniformemente en K también es equicontinua.

Para que una sucesión en $C_{pq}(K)$ sea uniformemente convergente en K es necesario que la sucesión sea acotada y uniformemente equicontinua en K. Se habrá de probar ahora que estas dos propiedades son necesarias y suficientes para que un conjunto $\mathcal{F}$ en $C_{pq}(K)$ tengo la propiedad de que toda sucesión de funciones de $\mathcal{F}$ tiene una subsucesión que converge uniformemente en K. Esto se puede considerar como una generalización del teorema de Bolzano-Weierstrass para conjuntos de funciones continuas y desempeña un papel importante en la teoría de ecuaciones diferenciales e integrales.

Teorema 275 ***(Teorema de Arzela-Ascoli)*** *Sea K un subconjunto compacto de $\mathbb{R}^p$ y sea $\mathcal{F}$ una colección de funciones continuas en K y con valores en $\mathbb{R}^q$. Las siguientes propiedades son equivalentes:*

a. *La familia $\mathcal{F}$ es acotada y uniformemente equicontinua en K.*

b. *Toda sucesión de $\mathcal{F}$ tiene una subsucesión que es uniformemente convergente en K.*

Demostración:

$b \Rightarrow a$) *Primero probaremos que si la condición (a) es falsa, entonces también lo es la condición (b) (razonamiento contrarecíproco). Si $\mathcal{F}$ no es acotada, entonces existe una sucesión (f_n) en $\mathcal{F}$ tal que $\|f_n\|_K \geq n$ para $n \in \mathbb{N}$. Pero entonces ninguna subsucesión de (f_n) puede ser uniformemente convergente. Además, si el conjunto F no es uniformemente equicontinuo, entonces para alguna $\varepsilon_0 > 0$ existe una sucesión (f_n) en F y sucesiones (x_n) y (y_n) en K con $\|x_n - y_n\| < 1/n$ pero tales que $\|f_n(x_n) - f_n(y_n)\| > \varepsilon_0$. Pero entonces ninguna subsucesión de (f_n) puede ser uniformemente convergente en K.*

$a \Rightarrow b$) *Demostraremos ahora que si el conjunto F satisface (a) entonces dada cualquier sucesión (f_n) en F hay una subsucesión que converge uniformemente en K. Existe un conjunto numerable C en K tal que si $y \in K$ y $\varepsilon > 0$, entonces existe un elemento x en C tal que $\|x - y\| < \varepsilon$. Si $C = \{x_1, x_2, ...\}$, entonces la sucesión $(f_n(x_1))$ es acotada en $\mathbb{R}^q$, y del teorema de Bolzano-Weierstrass se deduce que hay una subsucesión*

$$(f_1^1(x_1), f_2^1(x_1), ..., f_n^1(x_1), ...)$$

de $(f_n(x_1))$ que es convergente. Observamos ahora que la sucesión $(f_k^1(x_2))_{k\in\mathbb{N}}$ es acotada en $\mathbb{R}^q$; por tanto, tiene una subsucesión

$$(f_1^2(x_2), f_2^2(x_2), ..., f_n^2(x_2), ...)$$

que es convergente. Nuevamente, la sucesión $(f_n^2(x_3))_{n\in\mathbb{N}}$ es acotada en $\mathbb{R}^q$, por lo que alguna subsucesión

$$(f_1^3(x_3), f_2^3(x_3), ..., f_n^3(x_3), ...)$$

es convergente. Se prosigue de la misma manera y después se fija $g_n = f_n^n$ de tal manera que g_n sea la enésima función en la enésima subsucesión. Por la construcción, es claro que la sucesión (g_n) converge en cada punto de C.

Probaremos ahora que la sucesión (g_n) converge en cada punto de K y que la convergencia es uniforme. Para hacer esto, sea $\varepsilon > 0$ y sea $\delta(\varepsilon)$ como en la definición de equicontinuidad. Sea $C_1 = \{y_1, ..., y_k\}$ un subconjunto finito de C tal que todo punto en K dista menos de $\delta(\varepsilon)$ de algún punto en C_1. Dado que las sucesiones

$$(g_n(y_1)),\ (g_n(y_2)), ...,\ (g_n(y_k))$$

convergen, existe un número natural M tal que si $m, n \geq M$, entonces

$$\|g_m(y_i) - g_n(y_i)\| < \varepsilon \quad \text{para} \quad i = 1, 2, ..., k.$$

Dada $x \in K$, existe una $y_j \in C_1$ tal que $\|x - y_j\| < \delta(\varepsilon)$ de donde, por la equicontinuidad uniforme, se tiene $\|g_n(x) - g_n(y_i)\| < \varepsilon$ para toda $n \in \mathbb{N}$; específicamente, esta desigualdad es válida para $n \geq M$. Por lo tanto, se tiene

$$\begin{array}{rcl} \|g_n(x) - g_m(x)\| & \leq & \|g_n(x) - g_n(y_i)\| + \|g_n(y_i) - g_m(y_i)\| + \\ \|g_m(y_i) - g_m(x)\| & < & \varepsilon + \varepsilon + \varepsilon = 3\varepsilon, \end{array}$$

siempre que $m, n \geq M$. *Esto prueba que*

$$\|g_n - g_m\|_K \leq 3\varepsilon \quad \textit{para} \;\; m, n \geq M,$$

por lo que la convergencia uniforme de la sucesión (g_n) *en* K *se sigue del criterio de Cauchy para convergencia uniforme.*

■

En la demostración de este resultado se ha construido una sucesión de subsucesiones de funciones y después se ha seleccionado la sucesión "diagonal" (g_n), en donde $g_n = f_n^n$. A dicha construcción se la acostumbra a llamar "proceso diagonal" o "método diagonal de Cantor" y con frecuencia es útil.

Capítulo 14

Funciones diferenciables

14.1. Introducción

En este capítulo se estudia la teoría de funciones diferenciables en $\mathbb{R}^n$ en donde $n > 1$. La teoría es parecida a la que se vio en la asignatura de Cálculo I, pero surgen algunas complicaciones y aspectos nuevos. Varias de estas complicaciones se deben sólo a la invevitable complejidad de la notacion, pero otras surgen porque es posible llegar a un punto $c \in \mathbb{R}^n$ desde "muchas direcciones", de tal manera que pueden ocurrir fenómenos nuevos.

En la asignatura de Cálculo I, definimos la derivada de una función $f : \mathbb{R} \to \mathbb{R}$ como el número $L \in \mathbb{R}$ tal que

$$L = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

cuando este límite existe. De manera equivalente se pudo haber definido esta derivada como el número L tal que

$$\lim_{x \to c} \frac{|f(x) - f(c) - L \cdot (x - c)|}{|x - c|} = 0$$

Se puede considerar que dicha relación de límite hace preciso el sentido en el que se aproximan los valores $f(x)$, para x suficientemente cerca de c, por medio de la aplicación

$$x \mapsto f(c) + L(x - c)$$

cuya gráfica da la recta tangente a la gráfica de f en el punto $(c, f(c))$.

Es este acceso a la derivada el que se usará para funciones de $\mathbb{R}^n$ a $\mathbb{R}^m$. De manera que la derivada de una función f definida en un entorno de un punto $c \in \mathbb{R}^n$ con valores en $\mathbb{R}^m$ será una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ tal que

$$\lim_{x \to c} \frac{\|f(x) - f(c) - L(x - c)\|}{\|x - c\|} = 0$$

Se está aproximando $f(x)$, para x suficientemente cerca de c, por medio de la aplicación

$$x \mapsto f(c) + L(x - c)$$

de $\mathbb{R}^n$ hacia $\mathbb{R}^m$. [Observamos que si $n = 1$, entonces la notación $L(x - c)$ representa el producto de los números reales L y $x - c$; sin embargo, si $n > 1$, entonces $L(x - c)$ denota el valor de la aplicación lineal L en el vector $x - c$].

14.2. Funciones diferenciables

Definición 139 *Sea A un abierto de $\mathbb{R}^n$, f una aplicación de A en $\mathbb{R}^m$, y c un punto interior de A. Dado un vector u cualquiera de $\mathbb{R}^n$, se define la* ***derivada de*** f ***según el vector*** *u en el punto c, que designaremos por $D_u f(c)$, como el vector de $\mathbb{R}^m$ dado por*

$$D_u f(c) = \lim_{t \to 0} \frac{f(c+tu) - f(c)}{t}$$

Si f es de valor real ($m = 1$) y u es el i-ésimo vector de la base canónica de $\mathbb{R}^n$, $e_i = (0, 0, \ldots, 1, \ldots, 0)$, entonces la derivada de f según e_i se llama ***derivada parcial de*** f ***respecto*** e_i ***en*** *c, y se denota por medio de*

$$D_i f(c) \qquad ó \qquad \frac{\partial f(c)}{\partial x_i}$$

Si u es un vector unitario en $\mathbb{R}^n$, entonces la derivada parcial $D_u f(c)$ se llama la ***derivada direccional de*** f ***en*** c ***en la dirección de*** u.

Ejemplo 334 *La función $f : \mathbb{R}^3 \to \mathbb{R}$ definida por*

$$f(x,y,z) = \begin{cases} \frac{xyz}{z^2+x^2y^2} & si \quad xy \neq 0 \quad ó \quad z \neq 0 \\ 0 & si \quad xy = 0 \quad y \quad z = 0 \end{cases}$$

posee derivada según cualquier vector $v = (a, b, c)$ en el punto $(0, 0, 0)$.

Solución: *Si $c \neq 0$, entonces*

$$D_v f(0,0,0) = \lim_{t \to 0} \frac{f(ta,tb,tc) - f(0,0,0)}{t} = \lim_{t \to 0} \frac{t^3abc}{t^3c^2 + t^5a^2b^2} = \frac{ab}{c}$$

y si $c = 0$, es claro que $D_v f(0,0,0) = 0$. ■

Definición 140 *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}$ posee derivadas parciales en todos los puntos de A, se define el* ***gradiente*** *de f en x como*

$$\nabla f(x) = (D_1 f(x), D_2 f(x), \ldots, D_n f(x))$$

Definición 141 *Si dada una función $g : A \subset \mathbb{R}^n \to \mathbb{R}$ existe una función $f : A \subset \mathbb{R}^n \to \mathbb{R}$ tal que $\nabla f = g$, f se llama* ***función potencial*** *de g.*

Definición 142 *Dada una función $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$, se define la* ***matriz jacobiana*** *de f en a como la matriz $m \times n$ de las derivadas parciales:*

$$f'(a) = \begin{pmatrix} D_1 f_1(a) & D_2 f_1(a) & \cdots & D_n f_1(a) \\ D_1 f_2(a) & D_2 f_2(a) & \cdots & D_n f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m(a) & D_2 f_m(a) & \cdots & D_n f_m(a) \end{pmatrix}$$

Cuando $m = n$, el determinante de la matriz jacobiana se llama el ***jacobiano*** *de f en c, y frecuentemente se denota por $J_f(c)$.*

Observación 146 *Las filas de $f'(a)$ son los gradientes de cada f_i en a.*

Ejemplo 335

a. *El gradiente de la función* $f(x,y,z) = xe^z + ye^x + ze^y$ *en un punto genérico* (x,y,z) *es*

$$(e^z + ye^x, e^x + ze^y, e^y + xe^z)$$

b. *La función potencial de* $g(x,y) = (3x^2y^2, 2x^3y)$ *es*

$$f(x,y) = x^3y^2$$

c. *La matriz jacobiana de* $f(x,y,z) = (xyz, -3xz^2)$ *en un punto genérico* (x,y,z) *es*

$$f'(x,y,z) = \begin{pmatrix} yz & xz & xy \\ -3z^2 & 0 & -6xz \end{pmatrix}$$

El principal inconveniente de la derivada parcial de una función f en un punto c con respecto a un vector u es que sólo da una visión del comportamiento de f cerca de c en el conjunto unidimensional $\{c + tu : t \in \mathbb{R}\}$. Para poder obtener una información más completa acerca de f en un entorno de $c \in \mathbb{R}^n$, se introduce el concepto de la diferencial de f en c, que es una aplicación lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$.

Definición 143 *Una función* f *de un abierto* $A \subset \mathbb{R}^n$ *en* $\mathbb{R}^m$ *se dice que es* ***diferenciable*** *en* $c \in A$ *si existe una aplicación lineal* $L : \mathbb{R}^n \to \mathbb{R}^m$ *tal que*

$$\lim_{\|u\| \to 0} \frac{\|f(c+u) - f(c) - L(u)\|}{\|u\|} = 0$$

L *recibe el nombre de* ***diferencial de*** f *en* c*, y se designa por* df_c.

Si f *es diferenciable en cada uno de los puntos del abierto* A*, se dice que* f *es diferenciable en* A.

Teorema 276 *Si existe, la diferencial de* f *en* c *es única.*
Demostración: *Supongamos que* L_1, L_2 *son diferenciales de* f *en* c*. Para abreviar la notación, definimos* $d(u) = f(c+u) - f(c)$*.Entonces, se tiene*

$$\begin{aligned} \frac{\|L_1(u) - L_2(u)\|}{\|u\|} &= \frac{\|L_1(u) - d(u) + d(u) - L_2(u)\|}{\|u\|} \leq \\ &\leq \frac{\|d(u) - L_1(u)\|}{\|u\|} + \frac{\|d(u) - L_2(u)\|}{\|u\|} \end{aligned}$$

Tomando límites cuando $\|u\| \to 0$*, y teniendo en cuenta que, por definición de diferencial, se cumple*

$$\begin{aligned} \lim_{\|u\| \to 0} \frac{\|d(u) - L_1(u)\|}{\|u\|} &= 0 \\ \lim_{\|u\| \to 0} \frac{\|d(u) - L_2(u)\|}{\|u\|} &= 0 \end{aligned}$$

entonces deducimos que

$$\lim_{\|u\| \to 0} \frac{\|L_1(u) - L_2(u)\|}{\|u\|} = 0$$

Por lo tanto, para cada $x \in \mathbb{R}^n$, $x \neq 0$, *si hacemos* $u = tx$, *obtenemos:*

$$\begin{aligned} 0 &= \lim_{\|u\| \to 0} \frac{\|L_1(u) - L_2(u)\|}{\|u\|} = \lim_{t \to 0} \frac{\|L_1(tx) - L_2(tx)\|}{\|tx\|} = \\ &= \lim_{\|t\| \to 0} \frac{\|t\| \cdot \|L_1(x) - L_2(x)\|}{\|t\| \cdot \|x\|} = \frac{\|L_1(x) - L_2(x)\|}{\|x\|} \end{aligned}$$

Así pues, $\|L_1(x) - L_2(x)\| = 0$ *para todo* $x \in \mathbb{R}^n$, *de donde se deduce que* $L_1 = L_2$. ■

Ejemplo 336 *Toda función constante es diferenciable y su diferencial en cualquier punto es la aplicación lineal cero.*

Solución: *Sea* $A \subseteq \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$, *y supongamos que* $f : A \to \mathbb{R}^m$ *es la función constante definida por*

$$f(x) = y_0$$

para $x \in A$. *Si* c *es un punto interior de* A *y* $x \in A$, *entonces* $f(x) - f(c) = 0$. *Se deduce que* f *es diferenciable en* c *y que*

$$df_c = 0$$

es decir, df_c *es la función lineal que aplica todo elemento de* $\mathbb{R}^n$ *al elemento cero de* $\mathbb{R}^m$. *Por lo tanto, la derivada en cualquier punto de una función constante es la función lineal cero.* ■

Ejemplo 337 *Toda aplicación lineal es diferenciable y su diferencial en cualquier punto es ella misma.*

Solución: *Sea* $A = \mathbb{R}^n$ *y sea* $f : A \to \mathbb{R}^m$ *una función lineal. Si* $c \in A$ *y* $x \in A$, *entonces*

$$f(c+u) - f(c) - f(u) = 0$$

De aquí que se deduce que f *es diferenciable en* c *y que* $df(c) = f$. *Por lo tanto, la diferencial en cualquier punto de una función lineal es la función lineal misma.*
■

Teorema 277 *Si* $f : A \to \mathbb{R}^m$ *es diferenciable en* $c \in A$, *entonces es continua en* c.

Demostración: *Sea* $L = df_c$. *Definimos*

$$R(x-c) = \frac{f(x) - f(c) - L(x-c)}{\|x-c\|}$$

De aquí se deduce la fórmula

$$f(x) = f(c) + L(x-c) + \|x-c\| \cdot R(x-c)$$

Como f *es diferenciable en* c, $\lim_{x \to c} R(x-c) = 0$. *Tomando límites cuando* $x \to c$ *en la igualdad anterior, se tiene*

$$\lim_{x \to c} f(x) = f(c)$$

ya que el límite de $L(x-c)$ *es cero porque* L *es continua.* ■

Teorema 278 *Sea A un abierto de $\mathbb{R}^n$ y sea c un punto de A.*

a. *Si f y g están definidas de A en $\mathbb{R}^m$ y son diferenciables en c, y si $\alpha, \beta \in \mathbb{R}$, entonces la función $\alpha f + \beta g$ es diferenciable en c y se cumple:*

$$d(\alpha f + \beta g)_c = \alpha df_c + \beta dg_c$$

b. *Si $\varphi : A \to \mathbb{R}$ y $f : A \to \mathbb{R}^m$ son diferenciables en c, entonces la función $\varphi \cdot f$ es diferenciable en c y se cumple:*

$$d(\varphi \cdot f)_c = d\varphi_c \cdot f(c) + \varphi(c) \cdot df_c$$

Demostración:

a. *Puesto que f y g son diferenciables en c, existen dos aplicaciones lineales, df_c y dg_c, tales que*

$$\lim_{\|u\|\to 0} \frac{\|f(c+u) - f(c) - df_c(u)\|}{\|u\|} = 0$$

$$\lim_{\|u\|\to 0} \frac{\|g(c+u) - g(c) - dg_c(u)\|}{\|u\|} = 0$$

Por tanto,

$$\lim_{\|u\|\to 0} \frac{\|(\alpha f + \beta g)(c+u) - (\alpha f + \beta g)(c) - (\alpha df_c + \beta dg_c)(u)\|}{\|u\|}$$

$$\begin{aligned} \leq \quad & \lim_{\|u\|\to 0} \frac{\|(\alpha f)(c+u) - (\alpha f)(c) - (\alpha df_c)(u)\|}{\|u\|} \\ & + \lim_{\|u\|\to 0} \frac{\|(\beta g)(c+u) - (\beta g)(c) - (\beta dg_c)(u)\|}{\|u\|} \end{aligned}$$

$$\begin{aligned} = \quad & \alpha \cdot \lim_{\|u\|\to 0} \frac{\|f(c+u) - f(c) - df_c(u)\|}{\|u\|} \\ & + \beta \cdot \lim_{\|u\|\to 0} \frac{\|g(c+u) - g(c) - dg_c(u)\|}{\|u\|} \\ & = \alpha \cdot 0 + \beta \cdot 0 = 0 \end{aligned}$$

Como $\alpha df_c + \beta dg_c$ es una función lineal de $\mathbb{R}^n$ a $\mathbb{R}^m$, se deduce que es diferenciable en c y que $d(\alpha f + \beta g)_c = \alpha df_c + \beta dg_c$.

b. *Puesto que φ y f son diferenciables en c, existen dos aplicaciones lineales, $d\varphi_c$ y df_c, tales que*

$$\lim_{\|u\|\to 0} \frac{\|\varphi(c+u) - \varphi(c) - d\varphi_c(u)\|}{\|u\|} = 0$$

$$\lim_{\|u\|\to 0} \frac{\|f(c+u) - f(c) - df_c(u)\|}{\|u\|} = 0$$

Se cumple:

$$\begin{aligned}
& (\varphi \cdot f)(c+u) - (\varphi \cdot f)(c) - \{d\varphi_c(u)f(c) + \varphi(c)df_c(u)\} \\
= \; & \varphi(c+u)f(c+u) - \varphi(c)f(c) - d\varphi_c(u)f(c) - \varphi(c)df_c(u) + \\
& +\varphi(c)f(c+u) - \varphi(c)f(c+u) + d\varphi_c(u)f(c+u) - d\varphi_c(u)f(c+u) \\
= \; & \{\varphi(c+u) - \varphi(c) - d\varphi_c(u)\}f(c+u) + d\varphi_c(u)\{f(c+u) - f(c)\} + \\
& +\varphi(c)\{f(c+u) - f(c) - df_c(u)\}
\end{aligned}$$

Dado que df_c existe, f es continua en c. Por lo tanto, existe una constante M tal que si $\|u\| \leq \delta$, entonces $\|f(c+u)\| \leq M$. De aquí se puede ver que:

$$\begin{aligned}
& \lim_{\|u\|\to 0} \frac{\|(\varphi \cdot f)(c+u) - (\varphi \cdot f)(c) - \{d\varphi_c(u)f(c) + \varphi(c)df_c(u)\}\|}{\|u\|} \\
\leq \; & \lim_{\|u\|\to 0} \frac{\|\{\varphi(c+u) - \varphi(c) - d\varphi_c(u)\}f(c+u)\|}{\|u\|} + \\
& + \lim_{\|u\|\to 0} \frac{\|d\varphi_c(u)\{f(c+u) - f(c)\}\|}{\|u\|} \\
& + \lim_{\|u\|\to 0} \frac{\|\varphi(c)\{f(c+u) - f(c) - df_c(u)\}\|}{\|u\|} \\
= \; & 0
\end{aligned}$$

■

Teorema 279 ***(Regla de la cadena)*** *Sean A un abierto de $\mathbb{R}^n, B$ un abierto de $\mathbb{R}^m$, f una aplicación de A en B y g una aplicación de B en $\mathbb{R}^p$. Si f es diferenciable en a y g lo es en $b = f(a)$, entonces la aplicación composición $h = g \circ f$ es diferenciable en a y*

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

Demostración: *Sean $\lambda = df_a$ y $\mu = dg_b$. Definimos*

$$\begin{aligned}
\Psi(x) &= f(x) - f(a) - \lambda(x-a) && (1)\\
\Phi(y) &= g(y) - g(b) - \mu(y-b) && (2)\\
\rho(x) &= (g \circ f)(x) - (g \circ f)(a) - (\mu \circ \lambda)(x-a) && (3)
\end{aligned}$$

Por hipótesis, tenemos (4)

$$\lim_{x\to a} \frac{\|\Psi(x)\|}{\|x-a\|} = 0$$

y (5)

$$\lim_{y\to b} \frac{\|\Phi(y)\|}{\|y-b\|} = 0$$

Tenemos que demostrar que

$$\lim_{x\to a} \frac{\|\rho(x)\|}{\|x-a\|} = 0$$

Si sustituimos en (3) el valor de $\lambda(x-a)$ obtenido de (1)

$$\rho(x) = g(f(x)) - g(f(a)) - \mu(f(x) - f(a) - \Psi(x)) = \Phi(f(x)) + \mu(\Psi(x)).$$

Dividiendo por $\|x-a\|, x \neq a$, se tiene

$$\frac{\|\rho(x)\|}{\|x-a\|} \leq \frac{\|\Phi(f(x))\|}{\|x-a\|} + \frac{\|\mu(\Psi(x))\|}{\|x-a\|}$$

Por tanto, es suficiente probar que el límite cuando $x \to a$ de cada uno de dos sumandos del segundo miembro es 0.

En efecto, μ es continua por ser una aplicación lineal entre espacios de dimensión finita. Como consecuencia existe una constante $M > 0$ tal que

$$\|\mu(\Psi(x))\| \leq M\|\Psi(x)\|$$

de aquí

$$\frac{\|\mu(\Psi(x))\|}{\|x-a\|} \leq M\frac{\|\Psi(x)\|}{\|x-a\|}$$

y por la hipótesis (4) el límite es cero.

Veamos ahora que el límite del primer sumando es también 0. Sea $\varepsilon > 0$. Por ser λ continua existe $k > 0$ tal que $\|\lambda(x)\| \leq k\|x\|$. Por la hipótesis (5), si $\varepsilon' = \frac{\varepsilon}{1+k}$, existe $\alpha > 0$ tal que si $0 < \|y-b\| < \alpha$, $y \in B$, se tiene $\|\Phi(y)\| < \varepsilon'\|y-b\|$. Por ser f continua en a, dado $\alpha > 0$ existe $\delta_1 > 0$ tal que si $0 < \|x-a\| < \delta_1, x \in A$, se tiene $\|f(x) - b\| < \alpha$. Como consecuencia, para todos los elementos $x \in A$ tales que $\|x-a\| < \delta_1$, se cumple

$$\|\Phi(f(x))\| < \varepsilon'\|f(x) - b\|$$

sustituyendo el valor de $f(x) - b$ obtenido de (1)

$$\|\Phi(f(x))\| < \varepsilon'\|\Psi(x) + \lambda(x-a)\| \leq \varepsilon'\|\Psi(x)\| + \varepsilon' k\|x-a\|.$$

Además, por la hipótesis (4) existe $\delta_2 > 0$ tal que si $0 < \|x-a\| < \delta_2, x \in A$, se tiene $\|\Psi(x)\|/\|x-a\| < 1$. Por consiguiente,

$$\|\Phi(f(x))\| < \varepsilon'\|x-a\| + \varepsilon' k\|x-a\| = \varepsilon'(1+k)\|x-a\|$$

para todo $x \in A, 0 < \|x-a\| < \delta$, en donde $\delta = \min\{\delta_1, \delta_2\}$. Dividiendo por $\|x-a\|$ y teniendo en cuenta que $\varepsilon' = \varepsilon/(1+k)$, obtenemos

$$\frac{\|\Phi(f(x))\|}{\|x-a\|} < \varepsilon.$$

■

Ejemplo 338 *Sean las funciones $f : \mathbb{R}^3 \to \mathbb{R}^2$ y $g : \mathbb{R}^2 \to \mathbb{R}^4$ tales que la diferencial de f en el punto (a,b,c) es la aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^2$ determinada por la matriz*

$$f'(a,b,c) = \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix}$$

y la diferencial de g en el punto $f(a,b,c)$ es la aplicación de $\mathbb{R}^2$ en $\mathbb{R}^4$ determinada por la matriz

$$g'(f(a,b,c)) = \begin{pmatrix} 2 & -1 \\ 0 & 1 \\ 1 & -1 \\ -1 & 0 \end{pmatrix}.$$

Encontrar la diferencial de la función compuesta $h = g \circ f$.

Solución: *La diferencial de la función compuesta $h = g \circ f$ en (a,b,c) es la aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^4$ dada por la matriz*

$$h'(a,b,c) = \begin{pmatrix} 2 & -1 \\ 0 & 1 \\ 1 & -1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 3 \end{pmatrix} = \begin{pmatrix} -2 & 5 & -1 \\ 0 & -1 & 3 \\ -1 & 3 & -2 \\ 1 & -2 & -1 \end{pmatrix}.$$

■

Teorema 280 *Una aplicación $f = (f_1, f_2, \ldots, f_m)$ de $A \subset \mathbb{R}^n$ en $\mathbb{R}^m$ es diferenciable en un punto $c \in A$ si y sólo si cada f_i es diferenciable en c. La diferencial de f en c es la aplicación lineal df_c que tiene por componentes las diferenciales $d(f_i)_c$ de las componentes de f, es decir,*

$$df_c = (d(f_1)_c, d(f_2)_c, \ldots, d(f_m)_c).$$

Demostración: $\Rightarrow$) *Supongamos que f es diferenciable en c. Cada f_i es la composición de la proyección $i-$ésima π_i con f. Como π_i es lineal, entonces es diferenciable en todos los puntos de $\mathbb{R}^n$ y su diferencial es ella misma. Por la regla de la cadena, $f_i = \pi_i \circ f$ es diferenciable en c y $d(f_i)_c = \pi_i \circ df_c$.*

$\Leftarrow$) *Supongamos que cada f_i es diferenciable en c. Entonces se tiene*

$$\lim_{\|u\| \to 0} \frac{|f_i(c+u) - f_i(c) - d(f_i)_c(u)|}{\|u\|} = 0.$$

Ahora probaremos que la aplicación lineal $\lambda = (d(f_1)_c, d(f_2)_c, \ldots, d(f_m)_c)$ de $\mathbb{R}^n$ en $\mathbb{R}^m$ es la aplicación df_c. En efecto:

$$\lim_{\|u\| \to 0} \frac{\|f(c+u) - f(c) - \lambda(u)\|}{\|u\|} \leq \sum_{i=1}^{m} \lim_{\|u\| \to 0} \frac{|f_i(c+u) - f_i(c) - d(f_i)_c(u)|}{\|u\|} = 0.$$

Por tanto, f es diferenciable en c y su diferencial es la aplicación λ. ■

14.3. Condiciones de diferenciabilidad

Teorema 281 *Si $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$ es diferenciable en un punto $c \in A$, entonces existe $D_u f(c)$ para todo $u \in \mathbb{R}^n, u \neq 0$, y se cumple*

$$D_u f\,(c) = df_c\,(u)\,.$$

Demostración: *Puesto que f es diferenciable en c, se cumple*

$$\lim_{u \to 0} \frac{\|f(c+u) - f(c) - df_c(u)\|}{\|u\|} = 0$$

Considerando este límite a través de la recta $u = tv$, *se tiene*

$$0 = \lim_{t \to 0} \frac{\|f(c+tv) - f(c) - df_c(tv)\|}{\|tv\|} = \lim_{t \to 0} \frac{1}{\|v\|} \left\| \frac{f(c+tv) - f(c)}{t} - df_c(v) \right\|.$$

Como $1/\|v\|$ *es constante, se deduce que*

$$df_c(v) = \lim_{t \to 0} \frac{f(c+tv) - f(c)}{t} = D_v f(c).$$

■

Corolario 39 *Sea* $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *y sea* $c \in A$. *Si la diferencial* df_c *existe, entonces cada una de las derivadas parciales* $D_1 f(c), \ldots, D_n f(c)$ *existen en* $\mathbb{R}$ *y si* $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$, *entonces*

$$df_c(u) = u_1 D_1 f(c) + \cdots + u_n D_n f(c).$$

Demostración: *El teorema anterior implica que para cada uno de los vectores* $e_1, \ldots, e_n$ *las derivadas parciales* $D_1 f(c), \ldots, D_n f(c)$ *existen y son iguales a* $df_c(e_1), \ldots, df_c(e_n)$. *Sin embargo, dado que* df_c *es lineal y* $u = u_1 e_1 + \cdots + u_n e_n$, *se deduce que*

$$df_c(u) = \sum_{j=1}^{n} u_j df_c(e_j) = \sum_{j=1}^{n} u_j D_j f(c).$$

■

Observación 147 *El inverso del corolario anterior no siempre es válido, ya que las derivadas parciales pueden existir sin que exista la derivada. Por ejemplo, definimos la función* $f : \mathbb{R}^2 \to \mathbb{R}$ *como*

$$f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases}.$$

La derivada parcial de f *con respecto al vector* (a,b) *en* $(0,0)$ *está dada por*

$$D_{(a,b)} f(0,0) = \frac{ab^2}{a^2+b^2}, \qquad si \ (a,b) \neq (0,0).$$

En particular, $D_1 f(0,0) = 0$ *y* $D_2 f(0,0) = 0$. *Si la diferencial* df *existiese en* $(0,0)$, *el corolario anterior implicaría que*

$$D_{(a,b)} f(0,0) = df_{(0,0)}(a,b) = a \cdot 0 + b \cdot 0 = 0,$$

lo que es una contradicción.

Teorema 282 *Si* $f = (f_1, f_2, \ldots, f_m)$ *de* $A \subset \mathbb{R}^n$ *en* $\mathbb{R}^m$ *es diferenciable en un punto* c *del abeirto* A, *entonces* df_c *es la aplicación lineal definida por la matriz jacobiana* $f'(c) = (D_i f_j(c))$.

Demostración: *Consideremos en primer lugar* $m = 1$. *Sea* e_j *el vector* $j-$*ésimo de la base canónica de* $\mathbb{R}^m$. *Si* $(d_1, d_2, \ldots, d_n)$ *es la matriz que define la aplicación lineal* df_c, *se tiene:*

$$df_c(e_j) = \begin{pmatrix} d_1 & d_2 & \cdots & d_j & \cdots & d_n \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = d_j.$$

Ahora bien, puesto que f *es diferenciable en* c, *se cumple* $df_c(e_j) = D_j f(c)$, *con lo cual se concluye que* $D_j f(c) = d_j$.

Supongamos ahora que $m > 1$. *Como* $df_c = (d(f_1)_c, d(f_2)_c, \ldots, d(f_m)_c)$, *la matriz jacobiana*

$$\begin{pmatrix} D_1 f_1(c) & D_2 f_1(c) & \cdots & D_n f_1(c) \\ D_1 f_2(c) & D_2 f_2(c) & \cdots & D_n f_2(c) \\ \vdots & \vdots & \ddots & \vdots \\ D_1 f_m(c) & D_2 f_m(c) & \cdots & D_n f_m(c) \end{pmatrix},$$

cuyas filas son las matrices que definen $d(f_i)(c)$, *determina* df_c. ■

Teorema 283 *Si* $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *es diferenciable en un punto* c *del abierto* A,*entonces la derivada direccional de* f *en* c *es máxima en la dirección del vector gradiente y su valor es el módulo del gradiente.*

Demostración: *Sea* v *un vector de* $\mathbb{R}^n$ *tal que* $\|v\| = 1$. *Como*

$$D_v f(c) = df_c(v) = \nabla f(c) \cdot v$$

y por la desigualdad de Cauchy-Schwarz,

$$|\nabla f(c) \cdot v| \leq \|\nabla f(c)\| \cdot \|v\| = \|\nabla f(c)\|,$$

la derivada direccional $D_v f(c)$ *está acotada por* $\|\nabla f(c)\|$. *Ahora bien, para el vector* $v = \frac{\nabla f(c)}{\|\nabla f(c)\|}$ *se tiene :*

$$D_v f(c) = \nabla f(c) \cdot \frac{\nabla f(c)}{\|\nabla f(c)\|} = \|\nabla f(c)\|.$$

Por tanto, $D_v f(c)$ *es máxima en la dirección del vector gradiente, y su valor es* $\|\nabla f(c)\|$. ■

Ejemplo 339 *Sea* $f(x, y, z) = xz^2 - \sin yz^2$. *Calcular su derivada siguiendo el vector* $(1, 2, -1)$ *y su derivada direccional correspondiente en el punto* $c = (-3, \pi, 1)$. *¿Cuál es el valor de la derivada direccional máxima?*

Solución: *Como* f *es diferenciable en* c, *se cumple:*

$$\begin{aligned} D_v f(c) &= df_c(v) = \nabla f(c) \cdot v \\ \nabla f(c) &= (z^2, -z^2 \cos yz^2, 2xz - 2yz \cos yz^2)_c = (1, 1, -6 + 2\pi) \\ D_v f(c) &= (1, 1, -6 + 2\pi) \cdot (1, 2, -1) = 9 - 2\pi. \end{aligned}$$

La derivada direccional respecto a $\frac{v}{\|v\|}$ *es*

$$D_{v/\|v\|}f(c) = (1,1,-6+2\pi)\cdot\left(\frac{1}{\sqrt{6}},\frac{2}{\sqrt{6}},-\frac{1}{\sqrt{6}}\right) = \frac{9-2\pi}{\sqrt{6}}.$$

Observamos que si f *es diferenciable siempre se cumple*

$$D_{v/\|v\|}f(c) = \frac{1}{\|v\|}\,D_v f(c).$$

La función alcanza su derivada direccional máxima en la direccion del gradiente en el punto, $\nabla f(c) = (1,1,-6+2\pi)$, *y su valor es*

$$\|\nabla f(c)\| = \sqrt{2+(2\pi-6)^2}.$$

■

Teorema 284 ***(Condición suficiente de diferenciabilidad)*** *Sea* $f : A \subset \mathbb{R}^n \to \mathbb{R}^m$. *Si existen todas las derivadas parciales de* f *en un entorno de* c *y son continuas en* c, *entonces* f *es diferenciable en* c.

Demostración: *Como* f *es diferenciable en* c *si y sólo si lo es cada una de sus componentes, podemos suponer que* $m = 1$. *Consideramos la identidad*

$$\begin{aligned} f(c+u)-f(c) &= f(c_1+u_1,c_2,\ldots,c_n)-f(c_1,c_2,\ldots,c_n)\\ &+f(c_1+u_1,c_2+u_2,\ldots,c_n)-f(c_1+u_1,c_2,\ldots,c_n)\\ &+f(c_1+u_1,c_2+u_2,c_3+u_3,\ldots,c_n)-f(c_1+u_1,c_2+u_2,\ldots,c_n)\\ &+\cdots+f(c_1+u_1,c_2+u_2,\ldots,c_n+u_n)-f(c_1+u_1,c_2+u_2,\ldots,c_n)\end{aligned}$$

Para cada $j = 1,2,\ldots,n$, *definimos la función* $g_j : \mathbb{R} \to \mathbb{R}$ *dada por* $g_j(t) = f(c_1+u_1,\ldots,c_{j-1}+u_{j-1},t,c_{j+1},\ldots,c_n)$, *que es claramente continua en el intervalo* $[c_j, c_j+u_j]$ *y derivable en su interior. Aplicando el teorema del valor medio a* g_j *en este intervalo, se cumple:*

$$g_j(c_j+u_j)-g_j(c_j) = u_j\cdot g_j'(b_j),$$

donde $b_j \in (c_j, c_j+u_j)$. *Teniendo en cuenta la definición de* g_j, *si designamos por* a_j *el punto* $(c_1+u_1,\ldots,c_{j-1}+u_{j-1},b_j,c_{j+1},\ldots,c_n)$ *y sustituimos en la identidad*

$$f(c+u)-f(c) = \sum_{j=1}^{n} g_j(c_j+u_j)-g_j(c_j) = \sum_{j=1}^{n} u_j\cdot D_j f(a_j).$$

Probemos que f *es diferenciable en* c, *demostrando que*

$$\lim_{\|u\|\to 0}\frac{|f(c+u)-f(c)-\sum_{j=1}^{n}u_jD_jf(c)|}{\|u\|} = 0$$

En efecto,

$$\begin{aligned} \frac{1}{\|u\|}\left|f(c+u)-f(c)-\sum_{j=1}^{n}u_jD_jf(c)\right| &= \frac{1}{\|u\|}\left|\sum_{j=1}^{n}u_jD_jf(a_j)-\sum_{j=1}^{n}u_jD_jf(c)\right| =\\ &= \left|\sum_{j=1}^{n}(D_jf(a_j)-D_jf(c))\,\frac{u_j}{\|u\|}\right| \leq\\ &\leq \sum_{j=1}^{n}|D_jf(a_j)-D_jf(c)|\,,\end{aligned}$$

ya que $|u_j| \leq \|u\|$. *Ahora bien, como* $a_j \to c$ *si* $\|u\| \to 0$, *(ya que* $\|a_j - c\| \leq \|u\|$*), y las derivadas parciales* $D_j f(x)$ *son continuas en* c, *se tiene:*

$$\lim_{\|u\| \to 0} D_j f(a_j) = D_j f(c).$$

Tomando límites en la desigualdad anterior cuando $\|u\|$ *tiende a 0, se obtiene el resultado deseado.* ■

Ejemplo 340 *Demostrar que la función*

$$f(x, y, z) = z^2 e^{x^2} - \cos xyz + y^2 x$$

es diferenciable en cualquier punto de $\mathbb{R}^3$.

Solución: *Las derivadas parciales de la función* f *existen y son continuas en cualquier punto de* $\mathbb{R}^3$*:*

$$\begin{aligned} D_1 f(x, y, z) &= 2xz^2 e^{x^2} + yz \sin xyz + y^2 \\ D_2 f(x, y, z) &= xz \sin xyz + 2yx \\ D_3 f(x, y, z) &= 2ze^{x^2} + xy \sin xyz. \end{aligned}$$

Por tanto, f *es diferenciable en cualquier punto de* $\mathbb{R}^3$ *y su diferencial es la aplicación lineal de* $\mathbb{R}^3$ *en* $\mathbb{R}$ *determinada por la matriz jacobiana de* f *en el punto.* ■

Teorema 285 *Sea* $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *diferenciable en el abierto* A. *Si* a *es un punto de* A *y* h *un vector de* $\mathbb{R}^n$ *tales que* $a + th \in A$ *para todo* $t \in [0, 1]$, *entonces existe un número real* θ, $0 < \theta < 1$, *tal que* $f(a + h) - f(a) = df_z(h)$, *donde* $z = a + \theta h$.

Demostración: *Sea* Ψ *la función de* $[0, 1]$ *en* A *definida por*

$$\Psi(t) = a + th = (a_1 + th_1, a_2 + th_2, \ldots, a_n + th_n)$$

y $g : [0, 1] \to \mathbb{R}$ *la composición de* f *y* Ψ. *Por ser* f *y* Ψ *continuas y diferenciables en sus dominios,* $g = f \circ \Psi$ *es continua en* $[0, 1]$ *y derivable en* $(0, 1)$. *Aplicando a* g *el teorema del valor medio en* $[0, 1]$, *existe* $\theta \in (0, 1)$ *tal que* $g(1) - g(0) = g'(\theta)$. *Sustituyendo los valores de* $g(1) = f(a + h), g(0) = f(a)$ *y* $g'(\theta) = df_{a+\theta h} \circ d\Psi_\theta = df_{a+\theta h}(h)$, *si definimos* $z = a + \theta h$, *se tiene*

$$f(a + h) - f(a) = df_z(h).$$

■

Corolario 40 *Sea* A *un subconjunto abierto y convexo de* $\mathbb{R}^n$, f *una función de* A *en* $\mathbb{R}$ *diferenciable y con derivadas parciales acotadas en todo* A. *Entonces* f *es una función de Lipschitz en* A.

Demostración: *Supongamos que* $|D_j f(x)| \leq k$ *para todo* $x \in A, j = 1, 2, \ldots, n$. *Dados dos puntos* $x, y \in A$, *aplicando el teorema del valor medio a* f *en el segmento de extremos* x, y, *sabemos que existe* $z \in (x, y)$ *tal que*

$$|f(y) - f(x)| = |df_z(y - x)| = \left| \sum_{j=1}^{n} D_j f(z)(y_j - x_j) \right| \leq k \sum_{j=1}^{n} |y_j - x_j| \leq kn\|y - x\|.$$

■

Observación 148 *El resultado anterior es válido para f con valores en $\mathbb{R}^m$, ya que para cada componente f_i se tiene: $|f_i(y) - f_i(x)| \leq M\|y - x\|$, donde $M = kn$, y k es una cota común para las derivadas de todas la componentes. Por consiguiente,*

$$\|f(y) - f(x)\| \leq \sqrt{m}\,\text{máx}\{|f_i(y) - f_i(x)| : i = 1, 2, \ldots, n\} \leq \sqrt{m}M\|x - y\|.$$

El teorema del valor medio no es cierto para funciones con valores vectoriales.

Ejemplo 341 *La función $f(t) = (\sin t, \cos t)$ es continua en el intervalo $[0, 2\pi]$ y diferenciable en $(0, 2\pi)$. Sin embargo, para todo $\theta \in (0, 2\pi)$, se tiene:*

$$f(2\pi) - f(0) \neq 2\pi f'(\theta),$$

ya que $f(2\pi) - f(0) = (0, 0)$ y $2\pi f'(\theta) = (2\pi \cos\theta, -2\pi \sin\theta)$ y no existe ningún ángulo θ que anule simultáneamente al seno y al coseno.

14.4. Derivadas de orden superior

Definición 144 *Consideremos una función f de un abierto $A \subset \mathbb{R}^n$ en $\mathbb{R}$, y supongamos que existe la derivada parcial $D_i f(x)$ para todos los puntos $x \in A$. Entonces podemos definir una nueva función $g_i(x) = D_i f(x) : A \to \mathbb{R}$. Si en un punto $a \in A$ existe $D_j g_i(a)$, diremos que es la* ***derivada segunda*** *de f en a respecto a las variables i, j y la denotaremos por $D_{ij} f(a)$ o bien por*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Repitiendo el proceso anterior se pueden definir las derivadas parciales de cualquier orden. La notación para la derivada parcial $p-$ésima de f en a, primero respecto a la variable x_{i_1}, después respecto a la variable x_{i_2}, etc., es $D_{i_1, i_2, \ldots, i_n}(a)$ o bien

$$\frac{\partial^p f}{\partial x_{i_1} x_{i_2} \cdots x_{i_n}}(a).$$

Definición 145 *Si existen todas la derivadas parciales de f hasta las de orden q y son continuas en A, se dice que f es una* ***función de clase*** *q. Las funciones de clase q en A se representan por $C^q(A)$. Si $f \in C^q(A)$ para cualquier número natural q, entonces se dice que f es de* ***clase infinito*** *y se escribe $f \in C^\infty(A)$.*

Observación 149 *Evidentemente, toda función de clase q en A es de clase $p \leq q$ en A.*

Observación 150 *$C^q(A)$ es un álgebra respecto a las operaciones usuales, o en otras palabras la suma, el producto por un escalar y el producto de dos funciones de clase q es una función de clase q. Naturalmente, una función $f = (f_1, f_2, \ldots, f_m)$ con valores en $\mathbb{R}^m$ es de clase q en A si lo es cada una de sus componentes. Además, se cumple la siguiente regla de la cadena: Sean A un abierto de $\mathbb{R}^n$ y B un abierto de $\mathbb{R}^m$. Si $f : A \to B$ es de clase q en A y $g : B \to \mathbb{R}$ es de clase p en B, entonces $g \circ f$ es de clase $r = \text{mín}\{p, q\}$ en A.*

Teorema 286 ***(Teorema de Schwarz)*** *Sea A un abierto de $\mathbb{R}^2$ y f una función de A en $\mathbb{R}$. Si $f \in C^1(A)$, existe $D_{12}f(x,y)$ para todo $(x,y) \in A$ y la función $D_{12}f$ es continua en un punto $(a,b) \in A$, entonces existe $D_{21}f(a,b)$ y coincide con $D_{12}f(a,b)$.*

Demostración: *Consideremos*

$$F(h) = \frac{D_2 f(a+h,b) - D_2 f(a,b)}{h}.$$

Probemos que existe $\lim_{h\to 0} F(h)$, es decir, existe $D_{21}f(a,b)$, y que su valor es $D_{12}f(a,b)$. De esta manera, tendremos que $D_{21}f(a,b) = D_{12}f(a,b)$.

En primer lugar, si hacemos

$$A(h,k) = \frac{f(a+h,b+k) - f(a,b+k) - f(a+h,b) + f(a,b)}{hk},$$

como

$$F(h) = \frac{1}{h}\left[\lim_{k\to 0} \frac{f(a+h,b+k) - f(a+h,b)}{k} - \lim_{k\to 0} \frac{f(a,b+k) - f(a,b)}{k}\right]$$

se tiene

$$F(h) = \lim_{k\to 0} A(h,k).$$

Sea $\varepsilon > 0$. Por ser $D_{12}f$ continua en (a,b) y A abierto, dado $\varepsilon > 0$ existe $\delta > 0$ tal que si $|h| < \delta$ y $|k| < \delta$, se tiene:

$$(a+h,b+k) \in A \quad y \quad |D_{12}f(a+h,b+k) - D_{12}f(a,b)| < \varepsilon.$$

Definamos en el intervalo $[a-\delta, a+\delta]$ la función real

$$g(x) = f(x,b+k) - f(x,b),$$

con lo que

$$A(h,k) = \frac{g(a+h) - g(a)}{hk}.$$

Por las hipótesis, g es continua en el intervalo $[a-\delta, a+\delta]$ y derivable en su interior. Aplicando el teorema del valor medio a g en el intervalo $[a, a+h]$, existe $\theta, 0 < \theta < 1$, tal que $g(a+h) - g(a) = h \cdot g'(a+\theta h)$. Entonces resulta

$$A(h,k) = \frac{1}{k} g'(a+\theta h) = \frac{1}{k}[D_1 f(a+\theta h, b+k) - D_1 f(a+\theta h, b)].$$

Por otra parte, sea $l(y) = D_1 f(a+\theta h, y)$ definida en el intervalo $[b-\delta, b+\delta]$. Por las hipótesis, l es continua en este intervalo y derivable en su interior. Aplicando a l el teorema del valor medio en el intervalo $[b, b+k]$, existe $\theta', 0 < \theta' < 1$ tal que

$$l(b+k) - l(b) = k \cdot l'(b+\theta' k),$$

luego

$$A(h,k) = l'(b+\theta k) = D_{12}f(a+\theta h, b+\theta' k).$$

Como consecuencia, al ser $|\theta h| < \delta$ y $|\theta' k| < \delta$, se tiene

$$|A(h,k) - D_{12}f(a,b)| = |D_{12}f(a+\theta h, b+\theta' k) - D_{12}f(a,b)| < \varepsilon$$

gracias a la continuidad de $D_{12}f$. *Tomando límites cuando* k *tiende a cero en la desigualdad anterior, y teniendo en cuenta que*

$$\lim_{k\to 0} A(h,k) = F(h),$$

resulta

$$|F(h) - D_{12}f(a,b)| < \varepsilon \qquad \text{para todo } |h| < \delta.$$

Así pues,

$$\lim_{h\to 0} F(h) = D_{12}f(a,b).$$

■

Corolario 41 *Sea* A *un abierto de* $\mathbb{R}^n$ *y* $f : A \to \mathbb{R}$. *Si* $f \in C^1(A)$, *existe* $D_{ij}f(x)$ *para todo* $x \in A$ *y la función* $D_{ij}f(x)$ *es continua en un punto* $a \in A$, *entonces existe* $D_{ji}f(a)$ *y coincide con* $D_{ij}f(a)$.

Demostración: *Supongamos, por ejemplo,* $i < j$, *y consideremos la función*

$$g(x,y) = f(a_1,\ldots,x,a_{i+1},\ldots,y,a_{j+1},\ldots,a_n)$$

definida en un entorno abierto del punto (a_i,a_j) *de* $\mathbb{R}^2$. *Aplicando a* g *el teorema de Schwarz, existe* $D_{21}g(a_i,a_j)$ *y coincide con* $D_{12}g(a_i,a_j)$. *Como*

$$\begin{aligned} D_{21}g(a_i,a_j) &= D_{ji}f(a) \\ D_{12}g(a_i,a_j) &= D_{ij}f(a), \end{aligned}$$

se tiene el resultado. ■

Ejemplo 342 *La función* $f(x,y) = y\arctan x$ *cumple todas las hipótesis del teorema de Schwarz en cualquier punto de* R^2 *y se tiene*

$$D_{12}f(x,y) = \frac{1}{1+x^2} = D_{21}f(x,y).$$

Ejemplo 343 *La función*

$$f(x,y) = \begin{cases} xy\,\frac{y^2-x^2}{y^2+x^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases}$$

es de clase C^1 *en todo* $\mathbb{R}^2$, *ya que existen y son continuas sus derivadas primeras*

$$D_1f(x,y) = \begin{cases} \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases}$$

$$D_2f(x,y) = \begin{cases} \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2} & si \quad (x,y) \neq (0,0) \\ 0 & si \quad (x,y) = (0,0) \end{cases}$$

existen sus derivadas cruzadas $D_{12}f$ *y* $D_{21}f$ *en el origen*

$$D_{12}f(0,0) = \lim_{t\to 0} \frac{D_1f(0,t) - D_1f(0,0)}{t} = -1$$

$$D_{21}f(0,0) = \lim_{t\to 0} \frac{D_2f(t,0) - D_2f(0,0)}{t} = 1$$

pero son diferentes. La hipótesis del teorema de Schwarz que no se cumple es la continuidad de una de las derivadas segundas en el punto, tanto $D_{12}f(x,y)$ *como* $D_{21}f(x,y)$ *y son discontinuas en* $(0,0)$.

Capítulo 15

Funciones continuas y diferenciables

15.1. Fórmula de Taylor

En la unidad de *Funciones diferenciables* vimos que dada una función $f : \mathbb{R}^n \to \mathbb{R}$, su diferencial en c es la aplicación lineal de $\mathbb{R}^n$ a $\mathbb{R}$ que más se aproxima a la diferencia $f(c+z) - f(c)$ cuando z es pequeña. Cualquier otra función lineal llegaría a una aproximación menos exacta para z pequeña. Vimos también que si $Df(c)$ existe, entonces necesariamente está dada por la fórmula

$$Df(c)(z) = D_1 f(c) z_1 + \cdots + D_n f(c) z_n,$$

donde $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$.

A pesar de que las aproximaciones lineales son en particular sencillas y son lo suficientemente exactas para muchos propósitos, algunas veces es deseable obtener un mayor grado de aproximación. En tales casos resulta natural dirigirse a funciones cuadráticas, funciones cúbicas, etcétera, para efectuar aproximaciones más precisas.

Definición 146 *Se define la* ***segunda derivada*** *de f en $c, D^2 f(c)$, como la función de $\mathbb{R}^n \times \mathbb{R}^n$ a $\mathbb{R}$ tal que si $(y,z) \in \mathbb{R}^n \times \mathbb{R}^n$ y $y = (y_1, \ldots, y_n), z = (z_1, \ldots, z_n)$, entonces*

$$D^2 f(c)(y,z) = \sum_{i,j=1}^{n} D_{ji} f(c) y_i z_j.$$

Al estudiar la segunda derivada se habrá de suponer en lo subsiguiente que las segundas derivadas parciales de f existen y son continuas en un entorno de c. De manera análoga se define la ***tercera derivada*** *$D^3 f(c)$ de f en c como la función de $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ en $\mathbb{R}$ dada por*

$$D^3 f(c)(y,z,w) = \sum_{i,j,k=1}^{n} D_{kji} f(c) y_i z_j w_k.$$

Al analizar la tercera derivada se habrá de suponer que todas las terceras derivadas parciales de f existen y son continuas en un entorno de c.

Análogamente se forman las derivadas de orden superior. (Observamos que si las derivadas parciales son continuas, entonces las derivadas parciales mixtas son independientes del orden de diferenciación).

Otra notación: se escribe

$$\begin{array}{ccc} D^2f(c)(w)^2 & \text{para} & D^2f(c)(w,w), \\ D^3f(c)(w)^3 & \text{para} & D^3f(c)(w,w,w), \\ \vdots & \vdots & \vdots \\ D^nf(c)(w)^n & \text{para} & D^nf(c)(w,w,\ldots,w). \end{array}$$

Si $n = 2$ y $w = (h, k)$, entonces $D^2f(c)(w)^2$ es igual a la expresión

$$D_{xx}f(c)h^2 + 2D_{xy}f(c)hk + D_{yy}f(c)k^2.$$

De forma análoga, $D^3f(c)(w)^3$ es igual a

$$D_{xxx}f(c)h^3 + 3D_{xxy}f(c)h^2k + 3D_{xyy}f(c)hk^2 + D_{yyy}f(c)k^3,$$

y $D^mf(c)(w)^m$ es igual a la expresión

$$D_{x\cdots x}f(c)h^m + \binom{n}{1}D_{x\cdots xy}f(c)h^{m-1}k + \binom{n}{2}D_{x\cdots xyy}f(c)h^{m-2}k^2 + \cdots + D_{y\cdots y}f(c)k^m.$$

Una vez introducida esta notación, probaremos una generalización importante del teorema de Taylor para funciones de $\mathbb{R}^n$ a $\mathbb{R}$.

Teorema 287 ***(Teorema de Taylor)*** *Sea A un abierto de $\mathbb{R}^n$, a un punto de A y $f : A \to \mathbb{R}$ una función de clase $C^m(A)$. Entonces, para cada $x \in A$ tal que el segmento $[a, x]$ está contenido en A, existe $\xi \in (a, x)$ tal que*

$$\begin{aligned} f(x) &= f(a) + \frac{1}{1!}\,Df(a)\,(x-a) + \frac{1}{2!}\,D^2f(a)\,(x-a)^2 + \\ &\quad + \cdots + \frac{1}{(m-1)!}\,D^{m-1}f(a)\,(x-a)^{m-1} + \frac{1}{m!}\,D^mf(\xi)\,(x-a)^m. \end{aligned}$$

Demostración: *Sea $h = x - a$, y sea $F : [0, 1] \to \mathbb{R}$ la función definida por*

$$F(t) = f(a + th).$$

Por el supuesto de que existen las derivadas parciales de f se sigue que

$$\begin{aligned} F'(t) &= \nabla f(a+th)\cdot h = \sum_{i=1}^{n} D_if(a+th)\cdot h_i = Df(a+th)(h), \\ F''(t) &= \sum_{i=1}^{n} h_i[\nabla D_if(a+th)\cdot h] = \sum_{i=1}^{n} h_i \sum_{j=1}^{n} D_{ij}f(a+th)\cdot h_j] = \\ &= \sum_{i,j=1}^{n} h_ih_j\,D_{ij}f(a+th) = D^2f(a+th)(h)^2 \end{aligned}$$

Por inducción, para todo $p \leq m$, se tiene

$$F^{(p)}(t) = D^pf(a+th)(h)^p.$$

Si se aplica la versión unidimensional del teorema de Taylor a la función F en $[0,1]$,se tiene que existe un número real $t_0 \in [0,1]$ tal que

$$F(1) = F(0) + \frac{1}{1!}F'(0) + \cdots + \frac{1}{(m-1)!}F^{(m-1)}(0) + \frac{1}{m!}F^{(m)}(t_0).$$

Sustituyendo los valores

$$\begin{aligned} F(0) &= f(a) \\ F(1) &= f(a+h) = f(x) \\ F^{(p)}(0) &= D^p f(a)(h)^p \end{aligned}$$

en la fórmula de Taylor para F, se obtiene la expresión del enunciado, en donde $\xi = a + t_0 h$ y $h = x - a$. ∎

Definición 147 *El polinomio*

$$P_m(x) = f(a) + \frac{1}{1!}\,Df(a)\,(x-a) + \frac{1}{2!}\,D^2 f(a)\,(x-a)^2 + \cdots + \frac{1}{m!}\,D^m f(a)\,(x-a)^m$$

se llama ***polinomio de Taylor de** f **en** a **de grado** m.*

El error cometido al aproximar $f(x)$ por el polinomio de Taylor de grado m se llama ***resto de orden*** *$m+1$, y por el teorema de Taylor sabemos que viene dado por*

$$R_{m+1}(x) = \frac{1}{(m+1)!}\,D^{m+1} f(\xi)\,(x-a)^{m+1}.$$

Si f es de clase C^∞, podemos asociarle la serie de potencias

$$S(x) = f(a) + \sum_{m=1}^{\infty} \frac{1}{m!}\,D^m f(a)\,(x-a)^m,$$

que recibe el nombre de ***serie de Taylor de** f **en** a. *Si $a = 0$, recibe el nombre de* ***serie de MacLaurin***.

Observación 151 *De acuerdo con el teorema anterior, la fórmula de Taylor es válida para todos los puntos x pertenecientes al mayor disco abierto $D(a,r)$ contenido en A. Si consideramos un disco cerrado o en general, cualquier compacto K contenido en $D(a,r)$, como las derivadas de orden m de f son continuas en K, alcanzan allí su máximo. Si M es una cota superior de estas derivadas en K, el valor absoluto del resto de orden m*

$$R_m(x) = \frac{1}{m!}\,D^m f(\xi)\,(x-a)^m$$

puede ser acotado por $Mr^m n^m/m!$, ya que $|x_i - a_i| < r$ y el número de términos del desarrollo de las derivadas es n^m, lo cual permite obtener una cota del error cometido al aproximar $f(x)$ en K por el polinomio de Taylor $P_{m-1}(x)$ de f en el punto a.

Ejemplo 344 *Consideramos la función $f : \mathbb{R}^3 \to \mathbb{R}$ definida por*

$$f(x,y,z) = e^{ax+by+cz},$$

donde a, b, c son números reales fijados. Calcular su serie de Taylor en el punto $(0,0,0)$.

Solución: *f es una función de clase $C^\infty(\mathbb{R}^3)$, luego existen las derivadas parciales de cualquier orden. Éstas valen*

$$\frac{\partial^{p+q+r} f(x,y,z)}{\partial x^p y^q z^r} = a^p b^q c^r \cdot e^{ax+by+cz},$$

por lo tanto,

$$\frac{\partial^{p+q+r} f(0,0,0)}{\partial x^p y^q z^r} = a^p b^q c^r.$$

Por consiguiente, la serie de Taylor asociada a la función f en el punto $(0,0,0)$ (serie de MacLaurin) viene dada por

$$S(x) = 1 + \sum_{m=1}^{\infty} \frac{1}{m!}(ax+by+cz)^m.$$

■

Definición 148 *Una función f de clase $C^\infty(A)$ se dice que es **analítica** en A si para cada punto $a \in A$ existe un disco $D(a,\delta) \subseteq A$ tal que la serie de Taylor de f en a converge uniformemente a f en $D(a,\delta)$.*

Teorema 288 *Si f es de clase $C^\infty(A)$ y $k > 0$ es una constante tal que para cualquier derivada de orden m se tiene*

$$|D_{i_1,i_2,\ldots,i_m} f(x)| < k^m$$

para todo $x \in A$, entonces f es analítica en A.

Demostración: *Sea a un punto de A. Probaremos que la serie de Taylor de f en a converge uniformemente en $D(a,r)$. Como para todo $x \in D(a,r)$ la convergencia de la serie es equivalente a la existencia de $\lim_{m\to 0} P_m(x)$, es suficiente probar que este límite es $f(x)$. En efecto, teniendo en cuenta la observación al teorema de Taylor, para todo $x \in D(a,r)$, resulta*

$$|f(x) - P_m(x)| = |R_{m+1}(x)| \leq \frac{k^{m+1} r^{m+1} n^{m+1}}{(m+1)!} = \frac{(krn)^{m+1}}{(m+1)!}$$

y el último término de la desigualdad converge a cero cuando $m \to \infty$, ya que es el término general de la serie que define e^{krn} (recordar que $e^t = \sum_{m=0}^{\infty} t^m/m!$), por tanto

$$f(x) = \lim_{m\to 0} P_m(x)$$

uniformemente en $D(a,r)$. ■

15.2. Teorema de la función inversa

Sea A un conjunto abierto en $\mathbb{R}^n$ y sea $f : A \to \mathbb{R}^m$.Probaremos que, con algunas hipótesis, el "caracter local" de f en $c \in A$ está indicado por df_c. De manera un poco más precisa:

(i) si $n \leq m$ y df_c es inyectiva, entonces f es inyectiva en pequeños entornos de c

(ii) si $n \geq m$ y df_c es exhaustiva, entonces la imagen de un entorno pequeño de c es un entorno de $f(c)$

(iii) si $n = m$ y df_c es biyectiva, entonces f aplica un entorno U de c biyectivamente sobre un entorno V de $f(c)$. Existe también una función definida en V que es inversa a la restricción de f a U.

Recordamos que si $f : A \to \mathbb{R}^m$ es diferenciable en todo punto de $A \subseteq \mathbb{R}^n$, entonces la función $x \mapsto df_x$ es una aplicación de A hacia $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ (conjunto de todas las aplicaciones lineales de $\mathbb{R}^n$ a $\mathbb{R}^m$). El conjunto $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ es un espacio vectorial normado según la norma

$$\|L\|_{nm} = \sup\{\|L(x)\| : x \in \mathbb{R}^p, \ \|x\| \leq 1\}.$$

El siguiente lema, que que es una variante del teorema del valor medio, será necesario más adelante.

Lema 16 *Sea $A \subseteq \mathbb{R}^n$ un conjunto abierto y sea $f : A \to \mathbb{R}^m$ diferenciable en A. Supongamos que A contiene a los puntos a, b y al segmento de línea S que une a estos puntos, y sea $x_0 \in A$. Entonces se tiene:*

$$\|f(b) - f(a) - df_{x_0}(b-a)\| \leq \|b-a\| \cdot \sup_{x \in S}\{\|df_x - df_{x_0}\|_{nm}\}.$$

Demostración: *Definimos $g : A \to \mathbb{R}^m$ de la forma*

$$g(x) = f(x) - df_{x_0}(x).$$

Dado que df_{x_0} es lineal, se deduce que $dg_x = df_x - df_{x_0}$ para $x \in A$. Si se aplica el teorema del valor medio, existe un punto $c \in S$ tal que

$$\|f(b) - f(a) - df_{x_0}(b-a)\| = \|g(b) - g(a)\| \leq$$

$$\leq \|dg_c(b-a)\| = \|(df_c - df_{x_0})(b-a)\| \leq$$
$$\leq \|b-a\| \cdot \sup_{x \in S}\{\|df_x - df_{x_0}\|_{nm}\}.$$

■

El siguiente resultado es el lema clave para los teoremas sobre transformaciones.

Lema 17 ***(Lema de aproximación)*** *Sea $A \subseteq \mathbb{R}^n$ abierto y supongamos que $f : A \to \mathbb{R}^m$ es de clase $C^1(A)$. Si $x_0 \in A$ y $\varepsilon > 0$, entonces existe $\delta(\varepsilon) > 0$ tal que si $\|x_1 - x_0\| \leq \delta(\varepsilon)$ y $\|x_2 - x_0\| \leq \delta(\varepsilon)$, entonces $x_1, x_2 \in A$ y*

$$\|f(x_1) - f(x_2) - df_{x_0}(x_1 - x_2)\| \leq \varepsilon \|x_1 - x_2\|.$$

Demostración: *Dado que $x \mapsto df_x$ es una aplicación de A a $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ continua, dado $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que si $\|x - x_0\| < \delta(\varepsilon)$, entonces $x \in A$ y $\|df_x - df_{x_0}\|_{nm} \leq \varepsilon$. Ahora, supongamos que x_1, x_2 satisfacen $\|x_1 - x_0\| \leq \delta(\varepsilon)$ y $\|x_2 - x_0\| \leq \delta(\varepsilon)$, con lo cual el segmento de línea que une x_1 y x_2 queda dentro del disco cerrado con centro x_0 y radio $\delta(\varepsilon)$, y, por lo tanto, dentro de A. Aplicando ahora el lema anterior obtenemos la conclusión deseada.* ■

A continuación probaremos que si f es de clase $C^1(A)$ y si df_c es inyectiva, entonces la restricción de f a un entorno de c adecuado es también inyectiva.

Teorema 289 ***(Teorema de la transformación inyectiva)*** *Supongamos que $A \subseteq \mathbb{R}^n$ es abierto, que $f : A \to \mathbb{R}^m$ es de clase $C^1(A)$ y que $L = df_c$ es inyectiva. Entonces existe $\delta > 0$ tal que la restricción de f a $B_\delta = \{x \in \mathbb{R}^n : \|x - c\| \leq \delta\}$ es inyectiva. Más aún, la inversa de la restricción de f a B_δ es una función continua de $f(B_\delta) \subseteq \mathbb{R}^m$ a $B_\delta \subseteq \mathbb{R}^n$.*
Demostración: *Dado que la función lineal $L = df_c$ es inyectiva, tal como vimos en la unidad Funciones continuas esto implica que existe $r > 0$ tal que*

$$r \|u\| \leq \|df_c(u)\| \qquad \text{para } u \in \mathbb{R}^n \qquad (1)$$

Aplicando el lema de aproximación anterior con $\varepsilon = \frac{1}{2} r$, obtenemos un número $\delta > 0$ tal que si $\|x_1 - c\| \leq \delta$ y $\|x_2 - c\| \leq \delta$, entonces

$$\|f(x_1) - f(x_2) - L(x_1 - x_2)\| \leq \frac{1}{2} r \, \|x_1 - x_2\| .$$

Aplicando la desigualdad triangular al lado izquierdo, obtenemos

$$\|L(x_1 - x_2)\| - \|f(x_1) - f(x_2)\| \leq \frac{1}{2} r \, \|x_1 - x_2\| .$$

Tomando $u = x_1 - x_2$, y usando (1), tenemos:

$$r\|u\| - \|f(x_1) - f(x_2)\| \leq \frac{1}{2} r \, \|x_1 - x_2\|$$

Por tanto,

$$\frac{1}{2} r \, \|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\| \qquad (2)$$

para $x_1, x_2 \in B_\delta$. Esto prueba que la restricción de f a B_δ es una función inyectiva; por lo tanto, esta restricción tiene una función inversa a la que se designará por medio de g. Si $y_1, y_2 \in f(B_\delta)$, entonces existen puntos únicos $x_i = g(y_i)$ en B_δ tales que $y_i = f(x_i), i = 1, 2$. De (2) se infiere que

$$\|g(y_1) - g(y_2)\| \leq \frac{2}{r} \, \|y_1 - y_2\| ,$$

por lo que $g = (f_{|B_\delta})^{-1}$ es uniformemente continua de $f(B_\delta)$ a $\mathbb{R}^n$.

Observamos que g no está definida necesariamente en un entorno de $f(c)$, es decir $f(c)$ no necesariamente es un punto interior de $f(B_\delta)$. Por este motivo, no se puede asegurar nada acerca de la diferenciabilidad de g. ■

El siguiente resultado asegura que si f es de clase $C^1(A)$ y si para alguna $c \in A$, la aplicación lineal $df_c : \mathbb{R}^n \to \mathbb{R}^m$ es exhaustiva, entonces f aplica un entorno apropiado de c sobre un entorno de $f(c)$. Así pues, todo punto de $\mathbb{R}^m$ suficientemente cerca de $f(c)$ es la imagen por f de un punto cerca de c.

Teorema 290 ***(Teorema de la transformación exhaustiva)*** *Sea $A \subseteq \mathbb{R}^n$ abierto y supongamos que $f : A \to \mathbb{R}^m$ pertenece a la clase $C^1(A)$. Supongamos que para alguna $c \in A$, la función lineal $L = df_c : \mathbb{R}^n \to \mathbb{R}^m$ es exhaustiva.*

Entonces existen dos números $p > 0$ *y* $\alpha > 0$ *tales que si* $y \in \mathbb{R}^m$ *y* $\|y - f(c)\| \leq \alpha/2p$*, entonces existe* $x \in A$ *tal que* $\|x - c\| \leq \alpha$ *y* $f(x) = y$*.*
Demostración: *Dado que* L *es exhaustiva, cada uno de los vectores de la base canónica de* $\mathbb{R}^m$

$$e_1 = (1, 0, \ldots, 0), e_2 = (0, 1, \ldots, 0), \ldots, e_m = (0, 0, \ldots, 1)$$

es la imagen por L *de algún vector de* $\mathbb{R}^n$*, digamos* $u_1, u_2, \ldots, u_m$*. Definimos* $M : \mathbb{R}^m \to \mathbb{R}^n$ *como la función lineal que aplica* e_j *en* u_j *para* $j = 1, 2, \ldots, m$*; es decir,*

$$M\left(\sum_{i=1}^{m} a_i e_i\right) = \sum_{i=1}^{m} a_i u_i.$$

Entonces $L \circ M$ *es la aplicación identidad en* $\mathbb{R}^m$*; es decir* $(L \circ M)(y) = y$ *para toda* $y \in \mathbb{R}^m$*. Si se toma*

$$p = \left\{\sum_{i=1}^{m} \|u_j\|^2\right\}^{1/2},$$

entonces las desigualdades triangular y de Schwarz implican que si $y = \sum_{i=1}^{m} a_i e_i$*, entonces*

$$\begin{aligned} \|M(y)\| &= \left\|\sum_{i=1}^{m} a_i u_i\right\| \leq \sum_{i=1}^{m} |a_1| \, \|u_i\| \leq \\ &\leq \left\{\sum_{i=1}^{m} |a_i|^2\right\}^{1/2} \cdot \left\{\sum_{i=1}^{m} \|u_i\|^2\right\}^{1/2} = \\ &= p \, \|y\| . \end{aligned}$$

Por el lema de aproximación, existe un número $\alpha > 0$ *tal que si* $\|x_k - c\| \leq \alpha, k = 1, 2,$ *entonces* $x_k \in A$ *y*

$$\|f(x_1) - f(x_2) - L(x_1 - x_2)\| \leq \frac{1}{2p} \, \|x_1 - x_2\| \qquad (1)$$

Ahora, sea $B_\alpha = \{x \in \mathbb{R}^n : \|x - c\| \leq \alpha\}$ *y supongamos que* $y \in \mathbb{R}^m$ *cumple* $\|y - f(c)\| \leq \alpha/(2p)$*. Probaremos que existe* $x \in B_\alpha$ *tal que* $y = f(x)$*.*

Sea $x_0 = c$ *y sea* $x_1 = x_0 + M(y - f(c))$ *de tal manera que* $\|x_1 - x_0\| \leq p \, \|y - f(c)\| \leq \frac{1}{2} \alpha$*, de donde obtenemos:*

$$\|x_1 - x_0\| \leq \frac{\alpha}{2} \quad y \quad \|x_1 - c\| \leq \left(1 - \frac{1}{2}\right) \alpha.$$

Supongamos que $c = x_0, x_1, \ldots, x_q$ *se han elegido en* $\mathbb{R}^n$ *de forma que*

$$\|x_k - x_{k-1}\| \leq \frac{\alpha}{2^k}, \quad \|x_k - c\| \leq \left(1 - \frac{1}{2^k}\right) \alpha, \qquad (2)$$

para $k = 1, \ldots, q$*. Definimos ahora* x_{q+1} $(q \geq 1)$ *como*

$$x_{q+1} = x_q - M\left[f(x_q) - f(x_{q-1}) - L(x_q - x_{q-1})\right]. \qquad (3)$$

De (1) se infiere que

$$\begin{aligned} \|x_{q+1} - x_q\| &\leq p\,\|f(x_q) - f(x_{q-1}) - L(x_q - x_{q-1})\| \leq \\ &\leq \frac{1}{2}\,\|x_q - x_{q-1}\|\,, \end{aligned}$$

por lo que $\|x_{q+1} - x_q\| \leq \frac{1}{2}(\alpha/2^q) = \alpha/2^{q+1}$ *y*

$$\begin{aligned} \|x_{q+1} - c\| &\leq \|x_{q+1} - x_q\| + \|x_q - c\| \leq \\ &\leq (\alpha/2^{q+1}) + (1 - 1/2^q)\alpha = \\ &= (1 - 1/2^{q+1})\alpha. \end{aligned}$$

Por tanto, (2) también queda probado para $k = q+1$. *Así pues, se puede construir una sucesión* (x_q) *en* B_α *de esta manera. Si* $p \geq q$, *se tiene*

$$\begin{aligned} \|x_q - x_p\| &\leq \|x_q - x_{q+1}\| + \|x_{q+1} - x_{q+2}\| + \cdots + \|x_{p-1} - x_p\| \leq \\ &\leq \frac{\alpha}{2^{q+1}} + \frac{\alpha}{2^{q+2}} + \cdots + \frac{\alpha}{2^p} \leq \frac{\alpha}{2^q}. \end{aligned}$$

Se deduce que (x_q) *es una sucesión de Cauchy en* $\mathbb{R}^n$ *y por lo tanto converge a algún elemento* x. *Dado que* $\|x_q - c\| \leq (1 - 1/2^q)\alpha$, *se sigue que* $\|x - c\| \leq \alpha$, *de manera que* $x \in B_\alpha$.

Como $x_1 - x_0 = M(y - f(c))$, *se deduce que*

$$L(x_1 - x_0) = (L \circ M)(y - f(c)) = y - f(x_0).$$

Más aún, por (3) se tiene

$$\begin{aligned} L(x_{q+1} - x_q) &= -(L \circ M)\,[f(x_q) - f(x_{q-1}) - L(x_q - x_{q-1})] = \\ &= -\{f(x_q) - f(x_{q-1}) - L(x_q - x_{q-1})\} = \\ &= L(x_q - x_{q-1}) - [f(x_q) - f(x_{q-1})] \end{aligned}$$

Por inducción se obtiene

$$L(x_{q+1} - x_q) = y - f(x_q),$$

por lo que se infiere que $y = \lim f(x_q) = f(x)$. *Por lo tanto, todo punto que satisfaga* $\|y - f(c)\| \leq \alpha/(2p)$ *es la imagen bajo* f *de un punto* $x \in A$ *con* $\|x - c\| \leq \alpha$. ■

A continuación, combinaremos el teorema de la transformación inyectiva con el de la transformación exhaustiva para el caso $n = m$. En este caso supondremos que df_c es una biyección (es decir, la derivada df_c tiene inversa, lo que se da si y sólo si el jacobiano $\det[D_j f_i(c)]$ es distinto de cero).

Teorema 291 ***(Teorema de inversión)*** *Sea* $A \subseteq \mathbb{R}^n$ *abierto y supongamos que* $f : A \to \mathbb{R}^n$ *es de clase* $C^1(A)$. *Si* $c \in A$ *es tal que* df_c *es biyectiva, entonces existe un entorno abierto* U *de* c *tal que* $V = f(U)$ *es un entorno abierto de* $f(c)$ *y la restricción de* f *a* U *es una biyección sobre* V *con inversa continua. Además,* $g = (f_{|U})^{-1}$ *es de clase* $C^1(V)$ *y cumple*

$$dg_y = \left[df_{g(y)}\right]^{-1} \quad \textit{para} \;\; y \in V.$$

Demostración: *Por hipótesis $L = df_c$ es inyectiva. En la unidad de Funciones continuas vimos que esto implica que existe $r > 0$ tal que*

$$2r\,\|z\| \leq \|df_c(z)\| \quad para \;\; z \in \mathbb{R}^n.$$

Dado que f es de clase $C^1(A)$, hay un entorno de c en el que df_x es invertible y satisface

$$r\,\|z\| \leq \|df_x(z)\| \quad para \;\; z \in \mathbb{R}^n \qquad (1)$$

Nos limitaremos a un entorno U de c en que f es inyectiva y está contenida en un disco con centro c y radio α (como en el teorema de la transformación exhaustiva). Entonces, $V = f(U)$ es un entorno de $f(c)$ y de los teoremas de la tranformación inyectiva y exhaustiva anteriores se infiere que la restricción $f_{|U}$ tiene una inversa continua: $g : V \to \mathbb{R}^n$.

Queda por demostrar que g es diferenciable en un punto arbitrario $y_1 \in V$. Sea $x_1 = g(y_1) \in U$. Dado que f es diferenciable en x_1,se deduce que si $x \in U$, entonces

$$f(x) - f(x_1) - df_{x_1}(x - x_1) = \|x - x_1\| \cdot u(x),$$

en donde $\|u(x)\| \to 0$ si $x \to x_1$. Si M_1 es el inverso de la función lineal df_{x_1}, entonces

$$\begin{aligned} x - x_1 &= M_1\left[df_{x_1}(x - x_1)\right] = \\ &= M_1\left[f(x) - f(x_1) - \|x - x_1\| \cdot u(x)\right]. \end{aligned}$$

Si $x \in U$, entonces $x = g(y)$ para alguna $y = f(x) \in V$; además $y_1 = f(x_1)$, y esta ecuación se puede escribir en la forma

$$g(y) - g(y_1) - M_1(y - y_1) = -\|x - x_1\| \cdot M_1(u(x)).$$

Dado que df_{x_1} es inyectiva, como en la demostración del teorema de la transformación inyectiva se infiere que

$$\|y - y_1\| = \|f(x) - f(x_1)\| \geq \frac{1}{2}\, r\, \|x - x_1\|$$

siempre que y esté lo suficientemente cerca de y_1. Además, de (1) se deduce que $\|M_1(u)\| \leq \frac{1}{r}\|u\|$ para toda $u \in \mathbb{R}^n$. Por lo tanto, se tiene

$$\|g(y) - g(y_1) - M_1(y - y_1)\| \leq \frac{2}{r^2}\,\|u(x)\|\,\|y - y_1\|.$$

Ahora, si $y \to y_1$, entonces $x = g(y) \to g(y_1) = x_1$ y $\|u(x)\| \to 0$. Se concluye, por lo tanto, que dg_{y_1} existe y es igual a $M_1 = (df_{x_1})^{-1}$.

El hecho de que g pertenezca a la clase $C^1(V)$ se sigue de la relación $dg_y = \left[df_{g(y)}\right]^{-1}$ para $y \in V$, y de la continuidad de las aplicaciones

$$y \mapsto g(y), \quad x \mapsto df_x, \quad L \mapsto L^{-1}$$

de $V \to U$, $U \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, y $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, respectivamente. ■

Observación 152 *El teorema de inversión tiene un carácter local, en el sentido de que si f cumple las hipótesis, lo único que podemos asegurar es la existencia de f^{-1} en un entorno del punto. Se puede dar el caso de que f cumpla las hipótesis del teorema en cada uno de los puntos de un abierto A (con lo cual puede tener inversa local en un entorno de cada uno de sus puntos) y sin embargo no existir inversa global por no ser la función inyectiva en todo A.*

Ejemplo 345 *La función* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *definida por*

$$f(x, y) = (e^x \cos y, e^x \sin y)$$

tiene inversa local en todos los puntos de $\mathbb{R}^2$ *y sin embargo no tiene inversa global.*

Solución: *La función* f *es de clase* $C^\infty(\mathbb{R}^2)$ *por ser infinitamente derivable. Además,*

$$\det f'(x, y) = \begin{vmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{vmatrix} = e^{2x} > 0 \qquad \text{para todo } (x, y) \in \mathbb{R}^2.$$

Entonces f *verifica las hipótesis del teorema de inversión, con lo cual posee inversa local en un entorno de cada punto.*

Sin embargo, f *tiene inversa global porque no es inyectiva en* $\mathbb{R}^2$*, ya que* $f(x, y) = f(x, y + 2\pi)$. ■

Ejemplo 346 *Las hipótesis del teorema son una condición suficiente pero no necesaria para la existencia de inversa local. Por ejemplo, la función* $f(x) = x^3$ *no cumple las hipótesis en el punto* $x = 0$*, ya que* $\det f'(0) = 0$*, y sin embargo existe* $f^{-1}(y) = y^{1/3}$.

Teorema 292 *Si* f *es de clase* $C^1(A)$ *y* $\det f'(x) \neq 0$ *para todo* $x \in A$*, entonces la imagen de cualquier subconjunto abierto de* A *es un subconjunto abierto.*

Demostración: *Sea* U *un subconjunto abierto de* A*. Probaremos que* $f(U)$ *es entorno de cada uno de sus puntos y por lo tanto abierto. Si* $y \in f(U)$*, existe* $x \in A$ *tal que* $y = f(x)$ *(puede existir más de un punto de* A *que cumpla esto). Como* $\det f'(x) \neq 0$*, aplicando el teorema de inversión a la restricción de* f *a* U*, existe un entorno abierto* V *de* $x, V \subset U$ *tal que* $f(V) = W$ *es un entorno abierto de* y *contenido en* $f(U)$*. Por tanto,* $f(U)$ *es un entorno de* y. ■

Definición 149 *Sea* A *un abierto de* $\mathbb{R}^n$*. Una aplicación* $f : A \to \mathbb{R}^n$ *se dice que es un cambio de variable, una aplicación regular o un difeomorfismo de clase* q *si se verifica:*

a. $f \in C^q(A)$, $q \geq 1$.

b. $\det f'(x) \neq 0$ *para toda* $x \in A$.

c. f *es inyectiva en* A.

Teorema 293 *Si* f *es una aplicación regular de clase* $C^1(A)$*, entonces* $f(A)$ *es abierto y* $f^{-1} : f(A) \to A$ *es de clase* $C^1(f(A))$.

Demostración: *Por el teorema anterior,* $f(A)$ *es abierto. Por ser* f *inyectiva para todo* $y \in f(A)$*, existe un único* $x \in A$ *tal que* $y = f(x)$*. Por el teorema de inversión, existe un entorno abierto* V *de* x *y un entorno abierto* W *de* y *tales que la restricción de* f^{-1} *a* W *es de clase* $C^1(W)$*. Como esto se cumple para cada punto* $y \in f(A)$*, entonces* f^{-1} *es de clase* $C^1(f(A))$. ■

Ejemplo 347 ***(Coordenadas polares)*** *Sea la aplicación* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *definida por las ecuaciones*

$$\begin{aligned} x &= \rho \cos \theta \\ y &= \rho \sin \theta \end{aligned}$$

Evidentemente, f es de clase $C^\infty(\mathbb{R}^2)$, y su jacobiano

$$\det f'(\rho,\theta) = \begin{vmatrix} \cos\theta & -\rho\sin\theta \\ \sin\theta & \rho\cos\theta \end{vmatrix} = \rho$$

es distinto de 0 si $\rho \neq 0$. Como

$$f(\rho,\theta) = f(\rho,\theta+2\pi),$$

es claro que f no es inyectiva en todo $\mathbb{R}^2$. Por ejemplo, si para el valor fijo $\theta_0 = 0$ consideramos el abierto

$$A = \{(\rho,\theta) \in \mathbb{R}^2 : \rho > 0,\ -\pi < \theta < \pi\},$$

la restricción de f a A es inyectiva. Entonces f es un cambio de variables de clase C^∞ entre A y $f(A)$, donde $f(A)$ es todo $\mathbb{R}^2$ menos el semieje negativo OX^-. Resolviendo las ecuaciones dadas se obtiene el cambio inverso f^{-1} de $f(A)$ en A que queda determinado por

$$\begin{aligned} \rho &= \sqrt{x^2+y^2} \\ \theta &= \arctan\frac{y}{x} \end{aligned}$$

para $(x,y) \in f(A)$.

15.3. Teorema de la función implícita

Supongamos que F es una función definida en un subconjunto de $\mathbb{R}^n \times \mathbb{R}^m$ (si se hace la identificación obvia de $\mathbb{R}^n \times \mathbb{R}^m$ con $\mathbb{R}^{n+m}$, entonces no es necesario definir lo que significa que F sea continua, que sea diferenciable en un punto o que sea de clase C^1 en un conjunto). Supongamos que F aplica el punto (a,b) en el vector cero de $\mathbb{R}^m$. El problema de las funciones implícitas consiste en resolver la ecuación

$$F(x,y) = 0$$

para un elemento (digamos y) en términos del otro, en el sentido de encontrar una función φ definida en un subconjunto de $\mathbb{R}^n$ con valores en $\mathbb{R}^m$ tal que $b = \varphi(a)$ y

$$F(x,\varphi(x)) = 0$$

para toda x del dominio de φ.

Veamos ahora la interpretación del problema en términos de coordenadas. Si $x = (x_1,\ldots,x_n)$ y $y = (y_1,\ldots,y_n)$, entonces la ecuación $F(x,y) = 0$ toma la forma de m ecuaciones en los $n+m$ argumentos $x_1,\ldots,x_n,y_1,\ldots,y_m$ dadas por

$$\begin{aligned} f_1(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0, \\ &\vdots \\ f_m(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0. \end{aligned}$$

Por conveniencia, supondremos que $a = 0$ y $b = 0$, de tal manera que este sistema se satisface para $x_1 = 0,\ldots,x_n = 0, y_1 = 0,\ldots,y_m = 0$, y se desea resolver para las y_j en términos de las x_i, al menos cuando las $|x_i|$ son suficientemente pequeñas. Si las funciones f_j son lineales, entonces la condición de solubilidad es que el determinante de coeficientes de las y_j sea distinto de cero. Si las funciones f_j no son lineales, entonces la condición es que el determinante jacobiano sea distinto de cero. En ese caso, hay funciones φ_j, $j = 1, \ldots, m$, definidas y continuas cerca de $a = 0$ tales que si se sustituye

$$\begin{aligned} y_1 &= \varphi_1(x_1, \ldots, x_n), \\ &\vdots \\ y_m &= \varphi_m(x_1, \ldots, x_n). \end{aligned}$$

en el sistema anterior, se obtiene una identidad en las x_i.

Teorema 294 ***(Teorema de la función implícita)*** *Sea $A \subseteq \mathbb{R}^n \times \mathbb{R}^m$ abierto y sea $(a,b) \in A$. Supongamos que $F : A \to \mathbb{R}^m$ pertenece a la clase $C^1(A)$, que $F(a,b) = 0$, y que la aplicación lineal definida por*

$$L_2(v) = dF_{(a,b)}(0,v), \quad v \in \mathbb{R}^m,$$

es una biyección de $\mathbb{R}^m$ sobre $\mathbb{R}^m$. Entonces:

a) Existe un entorno abierto W de $a \in \mathbb{R}^n$ y una única función $\varphi : W \to \mathbb{R}^m$ de clase $C^1(W)$ tal que $b = \varphi(a)$ y

$$F(x, \varphi(x)) = 0 \quad \textit{para toda } x \in W.$$

b) Existe un entorno abierto U de $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$ tal que $(x,y) \in U$ satisface $F(x,y) = 0$ si y sólo si $y = \varphi(x)$ para $x \in W$.

Demostración: *Sin perder generalidad se puede suponer que $a = 0$ y $b = 0$. Supongamos que $H : A \to \mathbb{R}^n \times \mathbb{R}^m$ está definida como*

$$H(x,y) = (x, F(x,y)) \quad \textit{para } (x,y) \in A.$$

Evidentemente, H es de clase $C^1(A)$ y

$$dH_{(x,y)}(u,v) = (u, dF_{(x,y)}(u,v))$$

para $(x,y) \in A$ y $(u,v) \in \mathbb{R}^n \times \mathbb{R}^m$. $dH_{(0,0)}$ es invertible en $\mathbb{R}^n \times \mathbb{R}^m$. De hecho, si se toma $L_1 \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ definida por

$$L_1(u) = dF_{(0,0)}(u,0) \quad \textit{para } u \in \mathbb{R}^n,$$

entonces el hecho de que $dF_{(0,0)}(u,v) = L_1(u) + L_2(v)$ prueba que el inverso de $dH_{(0,0)}$ es la aplicación lineal K definida por

$$K(x,z) = (x, L_2^{-1}[z - L_1(x)]).$$

Por lo tanto, del teorema de inversión se deduce que existe un entorno abierto U de $(0,0) \in \mathbb{R}^n \times \mathbb{R}^m$ tal que $V = H(U)$ es un entorno abierto de $(0,0) \in \mathbb{R}^n \times \mathbb{R}^m$ y la restricción de H a U es una biyección sobre V con inversa continua $\Phi : V \to U$ que pertenece a la clase $C^1(V)$ y tal que $\Phi(0,0) = (0,0)$. Supongamos que Φ es de la forma

$$\Phi(x,z) = (\varphi_1(x,z), \varphi_2(x,z)) \quad \textit{para } (x,z) \in V,$$

donde $\varphi_1 : V \to \mathbb{R}^n$ *y* $\varphi_2 : V \to \mathbb{R}^m$*. Dado que*

$$\begin{aligned}(x,z) &= (H \circ \Phi)(x,z) = H\left[\varphi_1(x,z), \varphi_2(x,z)\right] = \\ &= \left[\varphi_1(x,z), F\left(\varphi_1(x,z), \varphi_2(x,z)\right)\right],\end{aligned}$$

se deduce que $\varphi_1(x,z) = x$ *para toda* $(x,z) \in V$*, por lo que* Φ *en realidad puede expresarse como*

$$\Phi(x,z) = (x, \varphi_2(x,z)) \quad \textit{para } (x,z) \in V.$$

Ahora definimos $P : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ *por* $P(x,z) = z$*. Entonces,* P *es lineal y continua y* $\varphi_2 = P \circ \Phi$*; por lo tanto,* φ_2 *es de clase* $C^1(V)$ *y se tiene*

$$z = F(x, \varphi_2(x,z)) \quad \textit{para } (x,z) \in V.$$

Sea ahora $W = \{x \in \mathbb{R}^n : (x,0) \in V\}$*, de manera que* W *es un entorno abierto de* 0 *en* $\mathbb{R}^n$*, y si se define* $\varphi(x) = \varphi_2(x,0)$ *para* $x \in W$*, es evidente que* $\varphi(0) = 0$*, y de la fórmula anterior se infiere que*

$$F(x, \varphi(x)) = 0 \quad \textit{para } x \in W.$$

Más aún, $d\varphi_x(u) = d\varphi_{2(x,0)}(u,0)$ *para* $x \in W, u \in \mathbb{R}^n$*, por lo que se concluye que* φ *pertenece a la clase* $C^1(W)$*. Esto prueba el apartado (a).*

Para completar la demostración del apartado (b), supongamos que $(x,y) \in U$ *satiface* $F(x,y) = 0$*. Entonces* $H(x,y) = (x, F(x,y)) = (x,0) \in V$*, por lo que se sigue que* $x \in W$*. Además,* $(x,y) = \Phi(x,0) = (x, \varphi_2(x,0)) = (x, \varphi(x))$*, de tal manera que* $y = \varphi(x)$*.* ■

Observación 153 *Por conveniencia de notación, hemos supuesto siempre que son las últimas* m *variables las que se definen como función implícita de las* n *primeras, pero siempre que se cumplan las hipótesis del teorema, el orden de las* m *variables es indiferente.*

Definición 150 *Si* $(x,y) \in A$*, la* ***derivada parcial de bloque*** $D_{(1)}F(x,y)$ *es la función lineal* $\mathbb{R}^n \to \mathbb{R}^m$ *dada por*

$$D_{(1)}F(x,y)(u) = dF_{(x,y)}(u,0) \quad \textit{para } u \in \mathbb{R}^n,$$

y la derivada parcial de bloque $D_{(2)}F(x,y)$ *es la función lineal* $\mathbb{R}^n \to \mathbb{R}^m$ *dada por*

$$D_{(2)}F(x,y)(v) = dF_{(x,y)}(0,v) \quad \textit{para } v \in \mathbb{R}^m.$$

Observación 154 *Dado que* $(u,v) = (u,0) + (0,v)$*, está claro que*

$$dF_{(x,y)}(u,v) = D_{(1)}F(x,y)(u) + D_{(2)}F(x,y)(v).$$

Observamos que las transformaciones L_1 y L_2 de la demostración anterior son $D_{(1)}F(0,0)$ y $D_{(2)}F(0,0)$, respectivamente.

Ejemplo 348 *La función* $f(x,y,z) = (x^2 + y^2 - xz^2 - 1, xz - yx^2)$ *es de clase* $C^\infty(\mathbb{R}^3)$*, cumple* $f(1,1,1) = (0,0)$ *y su matriz jacobiana*

$$f'(1,1,1) = \begin{pmatrix} 2x - z^2 & 2y & -2xz \\ z - 2xy & -x^2 & x \end{pmatrix}_{(1,1,1)} = \begin{pmatrix} 1 & 2 & -2 \\ -1 & -1 & 1 \end{pmatrix}$$

es de rango máximo para las variables x, y, también para las variables x, z pero no para las variables y, z, ya que

$$\begin{vmatrix} 1 & 2 \\ -1 & -1 \end{vmatrix} = 1, \qquad \begin{vmatrix} 1 & -2 \\ -1 & 1 \end{vmatrix} = -1, \qquad \begin{vmatrix} 2 & -2 \\ -1 & 1 \end{vmatrix} = 0.$$

Por tanto, podemos asegurar que f define en un entorno de $(1,1,1)$ a x, y como función implícita de z, y también a x, z como función implícita de y.

Corolario 42 *Con la hipótesis del teorema de la función implícita, existe $\gamma > 0$ tal que si $\|x - a\| < \gamma$, entonces la diferencial de φ en x es el elemento de $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ dado por*

$$d\varphi_x = -\left[D_{(2)}F(x, \varphi(x))\right]^{-1} \circ \left[D_{(1)}F(x, \varphi(x))\right].$$

Demostración: *Supongamos que $K : W \to \mathbb{R}^n \times \mathbb{R}^m$ está definida por*

$$K(x) = (x, \varphi(x)) \quad \text{para } x \in W.$$

Entonces, dado que $(F \circ K)(x) = F(x, \varphi(x)) = 0$, se tiene que $F \circ K : W \to \mathbb{R}^m$ es una función constante. Además, como fácilmente se puede ver que

$$dK_x(u) = (u, d\varphi_x(u)) \quad \text{para } u \in \mathbb{R}^n,$$

aplicando la regla de la cadena a la función constante $F \circ K$, deducimos que

$$0 = d(F \circ K)_x = dF_{K(x)} \circ dK_x.$$

Por definición de las derivadas de bloque, se tiene

$$dF_{(x,\varphi(x))}(u, v) = D_{(1)}F(x, \varphi(x))(u) + D_{(2)}F(x, \varphi(x))(v).$$

De aquí se deduce que si $u \in \mathbb{R}^n$, entonces

$$\begin{aligned} 0 &= dF_{(x,\varphi(x))}(u, d\varphi_x(u)) = D_{(1)}F(x, \varphi(x))(u) + D_{(2)}F(x, \varphi(x))(d\varphi_x(u)) = \\ &= D_{(1)}F(x, \varphi(x))(u) + \left[D_{(2)}F(x, \varphi(x)) \circ d\varphi_x\right](u). \end{aligned}$$

Por tanto, se tiene

$$0 = D_{(1)}F(x, \varphi(x)) + D_{(2)}F(x, \varphi(x)) \circ d\varphi_x$$

para toda $x \in W$. Por hipótesis, $L_2 = D_{(2)}F(a, b)$ es invertible. Dado que φ y F son continuas, existe $\gamma > 0$ tal que si $\|x - a\| < \gamma$, entonces $D_{(2)}F(x, \varphi(x))$ también es invertible. Por lo tanto, a partir de la ecuación anterior, deducimos el resultado deseado:

$$d\varphi_x = -\left[D_{(2)}F(x, \varphi(x))\right]^{-1} \circ \left[D_{(1)}F(x, \varphi(x))\right].$$

■

Observación 155 *Puede ser útil interpretar la fórmula anterior en términos de matrices. Supongamos que se tiene un sistema de m ecuaciones en $n + m$ argumentos dado por*

$$\begin{aligned} f_1(x_1, \ldots, x_n, y_1, \ldots, y_m) &= 0, \\ &\vdots \\ f_m(x_1, \ldots, x_n, y_1, \ldots, y_m) &= 0. \end{aligned}$$

Como ya se ha dicho antes, la hipótesis del teorema de la función implícita requiere que la matriz

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_{n+1}} & \cdots & \frac{\partial f_1}{\partial x_{n+m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_{n+1}} & \cdots & \frac{\partial f_m}{\partial x_{n+m}} \end{bmatrix}$$

sea invertible en el punto (a, b). En este caso la diferencial de la función solución φ en un punto x está dada por

$$-\begin{bmatrix} \frac{\partial f_1}{\partial x_{n+1}} & \cdots & \frac{\partial f_1}{\partial x_{n+m}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_{n+1}} & \cdots & \frac{\partial f_m}{\partial x_{n+m}} \end{bmatrix}^{-1} \cdot \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix},$$

donde se entiende que ambas matrices están calculadas en el punto $(x, \varphi(x))$ cercano a (a, b).

Ejemplo 349 *Supongamos que $x = h_1(z), y = h_2(z)$ es la función implícita definida por $f(x, y, z) = (x^2 + y^2 - xz^2 - 1, xz - yx^2)$, en un entorno del punto $(1, 1, 1)$. Calculemos las derivadas $h_1'(1), h_2'(1), h_1''(1), h_2''(1)$. Para ello derivamos respecto a z en las ecuaciones*

$$\begin{aligned} h_1^2(z) + h_2^2(z) - h_1(z)\, z^2 - 1 &= 0 \\ h_1(z)\, z - h_2(z)\, h_1^2(z) &= 0. \end{aligned}$$

Obtenemos:

$$\begin{aligned} 2h_1(z)\, h_1'(z) + 2h_2(z)\, h_2'(z) - z^2\, h_1'(z) - 2z\, h_1(z) &= 0 \\ z\, h_1'(z) + h_1(z) - h_2'(z)\, h_1^2(z) - 2h_2(z)\, h_1(z)\, h_1'(z) &= 0 \end{aligned}$$

Particularizando para $z = 1, x = h_1(1) = 1, y = h_2(1) = 1$, resulta

$$\begin{aligned} 2\, h_1'(1) + 2\, h_2'(1) - \, h_1'(1) - 2 &= 0 \\ h_1'(1) + 1 - h_2'(1) - 2\, h_1'(1) &= 0, \end{aligned}$$

de donde se obtiene que $h_1'(1) = 0$ y $h_2'(1) = 1$.

Para calcular las derivadas segundas, derivamos en las ecuaciones de las derivadas primeras:

$$2h_1'(z)^2 + 2h_1(z)\, h_1''(z) + 2h_2'(z)^2 + 2h_2(z)\, h_2''(z) - 2z\, h_1'(z) - z^2\, h_1''(z) - 2h_1(z) - 2z\, h_1'(z) = 0$$

$$h_1'(z) + zh_1''(z) + h_1'(z) - h_2''(z)\, h_1^2(z) - 2h_2'(z)\, h_1'(z)\, h_1(z) - 2h_2'(z)\, h_1(z)\, h_1'(z) -$$

$$-2h_2(z)\, h_1'(z)^2 - 2h_2(z)\, h_1(z)\, h_1''(z) = 0.$$

Sustituyendo $z = 1$ y los valores de las derivadas primeras, se obtiene:

$$h_1''(1) = h_2''(1) = 0.$$

15.4. Los teoremas de parametrización y de rango

Estudiaremos ahora otro teorema que ofrece condiciones bajo las cuales la imagen de una función que transforma un subconjunto abierto de $\mathbb{R}^n$ hacia $\mathbb{R}^m$ se puede parametrizar por medio de una función φ definida en un conjunto abierto en un espacio de menor dimensión.

Recuerde que si $L : \mathbb{R}^n \to \mathbb{R}^m$ es una aplicación lineal, entonces la **imagen** R_L de L es el subespacio de $\mathbb{R}^m$ dado por

$$R_L = \{L(x) : x \in \mathbb{R}^n\},$$

y el **núcleo** (o el **kernel**) N_L de L es el subespacio de $\mathbb{R}^n$ definido por

$$N_L = \{x \in \mathbb{R}^n : L(x) = 0\}.$$

La dimensión $r(L)$ de R_L se llama **rango** de L y la dimensión $n(L)$ de N_L se llama **nulidad** de L. (De modo que el rango de L es el número de vectores linealmente independientes en $\mathbb{R}^m$ necesarios para generar la imagen R_L, y la nulidad de L es el número de vectores linealmente independientes en $\mathbb{R}^n$ necesarios para generar el núcleo N_L). Se cumple: $n = n(L) + r(L)$; es decir, la dimensión del dominio de L es igual a la suma de la nulidad y el rango de L.

Si L se representa por una matriz $m \times n$, entonces se puede probar que el rango de L es el máximo número r tal que exista al menos una submatriz $r \times r$ con determinante distinto de cero. El teorema de parametrización asegura que si f es una aplicación C^1 de un conjunto abierto $A \subseteq \mathbb{R}^n$ hacia $\mathbb{R}^m$ tal que df_x tenga rango igual a r para toda $x \in A$ y si $f(a) = b \in \mathbb{R}^m$ para alguna $a \in A$, entonces hay un entorno V de a tal que la restricción de f a V se pueda dar como una aplicación C^1 definida en un entorno en $\mathbb{R}^r$.

Teorema 295 ***(Teorema de parametrización)*** *Sea $A \subseteq \mathbb{R}^n$ abierto y supongamos que $f : A \to \mathbb{R}^m$ pertenece a la clase $C^1(A)$. Supongamos que df_x tiene rango r para toda $x \in A$ y sea $f(a) = b \in \mathbb{R}^m$ para alguna $a \in A$. Entonces existe un entorno abierto $V \subseteq A$ de a y una función $\alpha : V \to \mathbb{R}^r$ de clase $C^1(V)$, y existe un conjunto abierto $W \subseteq \mathbb{R}^r$ y funciones $\beta : W \to \mathbb{R}^n$ y $\varphi : W \to \mathbb{R}^m$, tales que*

$$f(x) = (\varphi \circ \alpha)(x) \quad \textit{para toda } x \in V,$$

y

$$\varphi(t) = (f \circ \beta)(t) \quad \textit{para toda } t \in W.$$

Demostración: *Sin perder generalidad, se puede suponer que $a = 0 \in \mathbb{R}^n$ y $b = 0 \in \mathbb{R}^m$.*

Sea $L = df_0$, de manera que $L : \mathbb{R}^n \to \mathbb{R}^m$ tenga rango r y sea $\{x_1, \ldots, x_n\}$ una base en $\mathbb{R}^n$ tal que $\{x_{r+1}, \ldots, x_n\}$ genere el núcleo de L. Se toma X_1 como el subespacio generado por $\{x_1, \ldots, x_r\}$ y $X_2 = N_L$ como el subespacio generado por $\{x_{r+1}, \ldots, x_n\}$. Entonces, por álgebra lineal sabemos que $Y_1 = R_L$ está generado por $\{y_1 = L(x_1), \ldots, y_r = L(x_r)\}$. Elegimos $\{y_{r+1}, \ldots, y_m\}$ tal que $\{y_1, \ldots, y_m\}$ sea base de $\mathbb{R}^m$, y supongamos que Y_2 es el subespacio generado por $\{y_{r+1}, \ldots, y_m\}$.

Se sigue que todo vector $x \in \mathbb{R}^n$ *tiene una representación única de la forma* $x = c_1x_1 + \cdots + c_nx_n$. *Sean* P_1 *y* P_2 *las transformaciones lineales en* $\mathbb{R}^n$ *definidas como*

$$P_1(x) = \sum_{j=1}^{r} c_jx_j, \qquad P_2(x) = \sum_{j=r+1}^{n} c_jx_j.$$

Es claro que la imagen de P_j *es igual a* $X_j, j = 1, 2$. *De manera análoga, se toman* Q_1 *y* Q_2 *como las transformaciones lineales en* $\mathbb{R}^m$ *definidas para* $y = c_1y_1 + \cdots + c_qy_q$ *como*

$$Q_1(y) = \sum_{j=1}^{r} c_jy_j, \qquad Q_2(y) = \sum_{j=r+1}^{m} c_jy_j.$$

Claramente, la imagen de Q_j *es* Y_j, $j = 1, 2$.

Si L_1 *es la restricción de* L *a* X_1, *entonces* L_1 *es una biyección de* X_1 *sobre* Y_1. *Sea* $B : Y_1 \to X_1$ *el inverso de* L_1, *con lo cual* $(B \circ L)(x) = x$ *para toda* $x \in X_1$ *y* $(L \circ B)(y) = y$ *para toda* $y \in Y_1$. *Definimos ahora* $u : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ *como*

$$u(x) = (B \circ Q_1 \circ f)(x) + P_2(x),$$

de tal manera que $u(0) = 0$, u *aplica* $X_1 \cap A$ *en* X_1, *y*

$$du_x = (B \circ Q_1 \circ df_x) + P_2,$$

de donde se deduce que u *pertenece a la clase* $C^1(A)$. *Puesto que fácilmente se puede ver que* du_0 *es la transformación identidad en* $\mathbb{R}^n$, *del teorema de inversión se infiere que existe un entorno abierto* U *de* $a = 0$ *tal que* $U' = u(U)$ *es un entorno abierto de* 0 *y que la restricción de* u *a* U *es una biyección sobre* U' *con inverso* $w = u^{-1} : U' \to \mathbb{R}^n$ *que pertenece a la clase* $C^1(U')$. *Además, reemplazando* U *y* U' *por conjuntos más pequeños también se puede suponer que* U' *es convexo (es decir, contiene al segmento de línea que une a dos cualesquiera de sus puntos).*

Ahora, supongamos que $g : U' \to \mathbb{R}^m$ *se define como*

$$g(z) = f(w(z)), \quad z \in U' \subseteq \mathbb{R}^n$$

Está claro que g *pertenece a la clase* $C^1(U')$ *y*

$$dg_z = df_{w(z)} \circ dw_z, \quad z \in U'.$$

Dado que df_x *tiene rango* r *para toda* $x \in A$ *y* dw_z *es invertible para* $x \in U'$, *entonces se sigue de un teorema de álgebra lineal que* dg_z *tiene rango* r *para toda* $z \in U'$. *Por otro lado,*

$$\begin{aligned} g(z) &= (Q_1 + Q_2) \circ f(w(z)) = \\ &= (Q_1 \circ f)(w(z)) + (Q_2 \circ f)(w(z)). \end{aligned}$$

Dado que $w = u^{-1}$, *de la definición de* u *se deduce que*

$$z = u(w(z)) = (B \circ Q_1 \circ f)(w(z)) + P_2(w(z)), \quad z \in U'.$$

Pero como $L \circ B \circ Q_1 = Q_1$ *en* $\mathbb{R}^m$ *y* $L \circ P_2 = 0$ *en* $\mathbb{R}^n$, *se tiene*

$$L(z) = (Q_1 \circ f)(w(z)) = (Q_1 \circ g)(z),$$

de donde $L = Q_1 \circ dg_z$ *para* $z \in U'$. *Por lo tanto, si* $z \in U'$, *el operador* Q_1 *aplica la imagen de* dg_z *(que tiene dimensión* r*) sobre la imagen* L *(que también tiene dimensión* r*). De aquí se sigue que* Q_1 *es inyectiva en la imagen de* dg_z *para* $z \in U'$*; por lo tanto, si* $z \in U'$ *y* $x \in \mathbb{R}^n$ *es tal que* $L(x) = 0$, *entonces* $dg_z(x) = 0$. *Así pues, si* $z \in U'$ *y* $z_2 \in X_2 = N_L$, *se deduce que* $dg_z(z_2) = 0$.

Demostraremos ahora que $g : U' \to \mathbb{R}^m$ *depende sólo de* $z_1 \in X_1$, *en el sentido de que si* $z \in U'$ *y* $z_2 \in X_2$ *son tales que* $z + z_2 \in U'$, *entonces,* $g(z + z_2) = g(z)$. *Para verlo, se aplica el teorema del valor medio para deducir que existe un punto* z_0 *en el segmento de línea que une a* z *y* $z + z_2$ *(y por lo tanto incluido en* U'*) tal que*

$$0 \leq \|g(z + z_2) - g(z)\| \leq \|dg_{z_0}(z_2)\| = 0;$$

por lo tanto, $g(z + z_2) = g(z)$, *como se quería ver.*

Ahora ya se pueden definir las transformaciones α,β,φ. *Sea* $C : \mathbb{R}^r \to \mathbb{R}^n$ *la aplicación lineal que envía los elementos de la base canónica* $e_1, \ldots, e_r$ *de* $\mathbb{R}^r$ *a los vectores* $x_1, \ldots, x_r$ *que forman una base para* X_1. *Por lo tanto,* C *es una biyección de* $\mathbb{R}^r$ *sobre* X_1, *luego* $C^{-1} : X_1 \to \mathbb{R}^r$ *existe. Sea* $W = C^{-1}(U') = C^{-1}(U' \cap X_1)$, *de tal manera que* $W \subseteq \mathbb{R}^r$ *es un entorno abierto de* 0 *en* $\mathbb{R}^r$, *y sea* $V \subseteq U$ *un entorno abierto de* $a = 0$ *tal que* $(P_1 \circ u)(V) \subseteq U'$. *Definimos ahora* $\alpha : V \to \mathbb{R}^r$ *y* $\beta : W \to \mathbb{R}^n$ *como*

$$\alpha(x) = (C^{-1} \circ P_1 \circ u)(x), \qquad \beta(t) = (w \circ C)(t)$$

para $x \in V$ *y* $t \in W$. *Está claro que* α *es de clase* $C^1(V)$ *y* $\alpha(V) \subseteq W$, *y que* β *es de clase* $C^1(W)$ *y* $\beta(W) \subseteq U$. *Se define* $\varphi : W \to \mathbb{R}^m$ *para* $t \in W$ *como*

$$\varphi(t) = (g \circ C)(t),$$

por lo que se sigue que

$$\varphi(t) = (f \circ w \circ C)(t) = f \circ \beta(t).$$

Más aún, si $x \in V$, *entonces*

$$f(x) = f((w \circ u)(x)) = ((f \circ w) \circ u)(x) = (g \circ u)(x).$$

Sin embargo, hemos visto que $(g \circ u)(x) = (g \circ P_1 \circ u)(x)$, *de manera que*

$$\begin{aligned} f(x) &= (g \circ u)(x) = (g \circ (C \circ C^{-1}) \circ (P_1 \circ u))(x) = \\ &= ((g \circ C) \circ (C^{-1} \circ P_1 \circ u))(x) = \\ &= (\varphi \circ \alpha)(x). \end{aligned}$$

Por lo tanto, $f(x) = (\varphi \circ \alpha)(x)$ *para toda* $x \in V$. ■

Durante esta construcción, se ha establecido en realidad un poco más de información. En este corolario se hace uso de la notación que se desarrolló en la demostración del teorema.

Corolario 43 *Con las notaciones utilizadas en la demostración del teorema de parametrización, se tiene:*

(a) La aplicación $\varphi : W \to \mathbb{R}^m$ *es de la forma* $\varphi_1 + \varphi_2$*, donde* φ_1 *es la restricción a* W *de la aplicación lineal* $\mathbb{R}^r \to \mathbb{R}^m$ *que manda* $e_j \in \mathbb{R}^r$ *a* $y_i = L(x_j)$*,* $j = 1, \ldots, r$*, y donde* $\varphi_2(W) \subseteq Y_2$*.*

(b) Si $t \in W$*, entonces* $(\alpha \circ \beta)(t) = t$*.*

(c) Si $x \in U \cap X_1$*, entonces* $x \in V$ *y* $(\beta \circ \alpha)(x) = x$*.*

Demostración: *(a) Dado que* $g = Q_1 \circ g + Q_2 \circ g$*, y* $L(z) = (Q_1 \circ g)(z)$*, entonces se tiene* $g = L + Q_2 \circ g$*. Por lo tanto, a partir de la definición de* φ*, se obtiene*

$$\varphi = L \circ C + Q_2 \circ g \circ C,$$

que es de la forma establecida en (a).

(b) Si $t \in W$*, entonces* $x = \beta(t) = (w \circ C)(t) \in U$ *tiene la propiedad de que* $u(x) = (u \circ w \circ C)(t) = C(t) \in U' \cap X_1$*; por lo tanto* $(P_1 \circ u)(x) = C(t) \in U'$*, de modo que* $x \in V$ *y*

$$\alpha(x) = (C^{-1} \circ P_1 \circ u)(x) = (C^{-1} \circ C)(t) = t,$$

que prueba la afirmación (b).

(c) Si $x \in A \cap X_1$*, entonces, como* $u(x) = (B \circ Q_1 \circ f)(x) + P_2(x)$ *y* $P_2(x) = 0$*, se deduce que* $u(x) \in X_1$*. Por lo tanto, si* $x \in U \cap X_1$ *se sigue que* $(P_1 \circ u)(x) = u(x) \in U' \cap X_1$*, de tal manera que* $x \in V$*. Más aún,*

$$\begin{aligned}(\beta \circ \alpha)(x) &= [(w \circ C) \circ (C^{-1} \circ P_1 \circ u)](x) \\ &= (w \circ C \circ C^{-1} \circ u)(x) = (w \circ u)(x) = x.\end{aligned}$$

■

Ahora podemos usar el resultado del teorema de parametrización para demostrar el teorema del rango.

Teorema 296 ***(Teorema del rango)*** *Sea* $A \subseteq \mathbb{R}^n$ *abierto y supongamos que* $f : A \to \mathbb{R}^m$ *es de clase* $C^1(A)$*. Supongamos que* df_x *tiene rango* r *para toda* $x \in A$*, y sea* $f(a) = b \in \mathbb{R}^m$ *para alguna* $a \in A$*. Entonces*

(i) existen entornos abiertos V *de* a *y* V' *de* 0 *en* $\mathbb{R}^n$ *y una función* $\sigma : V \to V'$ *de clase* $C^1(V)$ *que tiene una inversa* $\sigma^{-1} : V' \to V$ *de clase* $C^1(V')$*;*

(ii) existen entornos abiertos Z *de* b *y* Z' *de* 0 *en* $\mathbb{R}^m$*, y una función* $\tau : Z' \to Z$ *de* $C^1(Z')$ *que tiene una inversa* $\tau^{-1} : Z \to Z'$ *de clase* $C^1(Z)$*;*

(iii) si $x \in V$ *entonces* $f(x) = (\tau \circ i_r \circ \sigma)(x)$*, donde* $i_r : \mathbb{R}^n \to \mathbb{R}^m$ *es la aplicación definida por*

$$i_r(c_1, \ldots, c_r, c_{r+1}, \ldots, c_n) = (c_1, \ldots, c_r, 0, \ldots, 0) \in \mathbb{R}^m.$$

Demostración: *Supondremos que* $a = 0$ *y* $b = 0$*, y usaremos la notación y los resultados establecidos en la demostración del teorema de parametrización. Sea* $D : \mathbb{R}^n \to \mathbb{R}^n$ *la función lineal que transforma a los elementos de la base canónica* $e_1, \ldots, e_n$ *de* $\mathbb{R}^n$ *en los vectores* $x_1, \ldots, x_n$*; entonces* D *es una biyección de* $\mathbb{R}^n$ *sobre* $\mathbb{R}^n$*, y* D^{-1} *existe. La aplicación* $\sigma : A \to \mathbb{R}^n$ *definida por* $\sigma(x) = (D^{-1} \circ\ u)(x)$ *pertenece a la clase* $C^1(A)$*, y puesto que la restricción de* u *a* U *tiene una inversa* $w : U' \to \mathbb{R}^n$ *que transforma* U' *sobre* U*, se sigue que la restricción de* σ *a* U *tiene una inversa* $\sigma^{-1} = w \circ D$ *que aplica* $D^{-1}(U')$ *sobre* U*.*

Sean $W \subseteq \mathbb{R}^r$ *y* $\varphi : W \to \mathbb{R}^m$ *como en el teorema de parametrización y sea* $H : \mathbb{R}^m \to \mathbb{R}^m$ *la función lineal que aplica los elementos de la base canónica* $e_1, \ldots, e_m$ *de* $\mathbb{R}^m$ *a los vectores* $y_1, \ldots, y_m$*; por lo tanto,* H *es una biyección de* $\mathbb{R}^m$ *sobre* $\mathbb{R}^m$ *y* H^{-1} *existe. Se define*

$$W' = \{(c_1, \ldots, c_m) \in \mathbb{R}^m : (c_1, \ldots, c_r) \in W\}$$

y se considera a $\tau : W' \to \mathbb{R}^m$ *definida como*

$$\tau(c_1, \ldots, c_m) = \varphi(c_1, \ldots, c_r) + H(0, \ldots, 0, c_{r+1}, \ldots, c_m).$$

Del primer apartado del corolario anterior se infiere que $d\tau_0 = H$*, por lo tanto, el teorema de inversión implica que la restricción de* τ *a algún entorno* Z' *de* 0 *es una biyección sobre algún entorno* Z *de* $\tau(0) = 0$.

Restringiendo aún más V*, si es necesario, podemos suponer que* $f(V) \subseteq Z$*. Ahora, sea* $x \in V$ *y consideremos* $\sigma(x) = (D^{-1} \circ u)(x)$*. Si* i_r *se define como antes, entonces* $(i_r \circ \sigma)(x) = ((C^{-1} \circ P_1 \circ u)(x), 0) = (\alpha(x), 0)$*. Por lo tanto,*

$$(\tau \circ i_r \circ \sigma)(x) = (\varphi \circ \sigma)(x) = f(x) \qquad \textit{para toda } x \in V.$$

■

Capítulo 16

Problemas sobre extremos

16.1. Extremos relativos

Definición 151 *Sean $A \subseteq \mathbb{R}^n$ y $f : A \to \mathbb{R}$. Un punto $c \in A$ se dice que es un* ***punto de mínimo (máximo) relativo*** *de f si existe $\delta > 0$ tal que $f(c) \leq f(x)$ $(f(c) \geq f(x))$ para toda $x \in A$ tal que $\|x - c\| < \delta$. Un punto $c \in A$ se dice que es un* ***punto de mínimo (máximo) relativo estricto*** *de f si existe $\delta > 0$ tal que $f(c) < f(x)$ $(f(c) > f(x))$ para toda $x \in A$ tal que $0 < \|x - c\| < \delta$. Más aún, si $c \in A$ es un punto de mínimo (o máximo) relativo (estricto) de f, se dice que c es un punto* ***extremo relativo (estricto)*** *de f o que f tiene un* ***extremo relativo (estricto)*** *en c.*

Con frecuencia es útil el siguiente resultado.

Teorema 297 *Sean $A \subseteq \mathbb{R}^n$ y $f : A \to \mathbb{R}$. Si un punto c interior de A es un extremo relativo de f, y si la derivada parcial $D_u f(c)$ de f con respecto a un vector $u \in \mathbb{R}^n$ existe, entonces $D_u f(c) = 0$.*
Demostración: *Por hipótesis, la restricción de f a la intersección de A con la recta $\{c + tu : t \in \mathbb{R}\}$ tiene un extremo relativo en c, y es una función de variable real. Por uno de los resultados de la unidad de derivadas, se sigue que $D_u f(c) = 0$.* ■

Teorema 298 *Sean $A \subseteq \mathbb{R}^n$ y $f : A \to \mathbb{R}$. Si un punto c interior de A es un extremo relativo de f, y si la diferencial df_c existe, entonces $df_c = 0$.*
Demostración: *Del uno de los resultados de la unidad Funciones diferenciables, se deduce que cada una de las derivadas parciales $D_j f(c), j = 1, \ldots, n$, existe y que si $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$, entonces*

$$df_c(u) = \sum_{j=1}^{n} u_j D_j f(c).$$

Por el teorema anterior, $D_j f(c) = 0$ para $j = 1, \ldots, n$, por lo que $df_c(u) = 0$ para toda $u \in \mathbb{R}^n$. ■

Observación 156 *Así pues, si $f : A \to \mathbb{R}$ tiene un extremo relativo en $c \in A$ y df_c existe, entonces*

$$D_1 f(c) = 0, \ldots, D_n f(c) = 0.$$

Definición 152 *Un punto interior c tal que* $df_c = 0$ *se llama* ***punto crítico*** *de* f.

Observación 157 *De los teoremas anteriores se deduce que si* A *es un conjunto abierto en* $\mathbb{R}^n$ *en el cual* f *es diferenciable, entonces el conjunto de puntos críticos de* f *contendrá a todos los puntos extremos relativos de* f. *Desde luego, el conjunto de puntos críticos también puede contener puntos en los que* f *no tenga un extremo relativo. Además,* f *puede tener un extremo relativo en un punto interior* c *de* A *en el que* df_c *no exista, o puede tener un extremo relativo en un punto* $c \in A$ *que no sea interior; en cualquiera de los casos, el punto* c *no será un punto crítico de* f.

Definición 153 *Se dice que* $c \in A$ *es un punto de ensilladura de* $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *si todo entorno de* c *contiene puntos en los que* f *es estrictamente mayor que* $f(c)$ *y también contiene puntos en los que* f *es estrictamente menor que* $f(c)$.

Ejemplo 350 *Sea* $f(x) = x^3$ *para* $x \in [-1, 1]$. *Entonces,* $df_0 = 0$*; sin embargo,* f *no tiene un extremo en* $x = 0$. *Por otro lado,* f *tiene extremos estrictos en los puntos* ± 1 *(que no son puntos interiores del dominio y no son puntos críticos).*

Ejemplo 351 *Sea* $f(x) = |x|$ *para* $x \in [-1, 1]$. *Entonces,* df_0 *no existe; sin embargo,* f *tiene un mínimo relativo estricto en el punto interior* 0. *Por otro lado,* f *tiene extremos relativos estrictos en los puntos* ± 1.

Ejemplo 352 *Supongamos que* $f : \mathbb{R}^2 \to \mathbb{R}$ *está definida por* $f(x, y) = xy$. *Entonces,* $df_{(0,0)} = 0$, *de manera que el origen* $(0, 0)$ *es un punto crítico de* f*; sin embargo, no es un extremo relativo de* f *ya que*

$$\begin{array}{lcll} f(0,0) & < & f(x,y) & \text{para } xy > 0, \\ f(0,0) & > & f(x,y) & \text{para } xy < 0. \end{array}$$

Por tanto, $(0, 0)$ *es un* ***punto de ensilladura*** *de* f.

En vista de los ejemplos que se acaban de dar, es conveniente tener condiciones que sean necesarias (o sean suficientes) para garantizar que un punto crítico es un extremo o es un punto de ensilladura. Los siguientes resultados dan condiciones en términos de la segunda derivada de f.

Teorema 299 *Sea* $A \subseteq \mathbb{R}^n$ *abierto y supongamos que* $f : A \to \mathbb{R}$ *tiene segundas derivadas parciales continuas en* A. *Si* $c \in A$ *es un punto de mínimo relativo (respectivamente, máximo relativo) de* f, *entonces*

$$D^2 f(c)(w)^2 = \sum_{i,j=1}^{n} D_{ij} f(c) w_i w_j \geq 0$$

(respectivamente, $D^2 f(c)(w)^2 \leq 0$*) para toda* $w \in \mathbb{R}^n$.

Demostración: *Sea* $w \in \mathbb{R}^n$, $\|w\| = 1$. *Si* c *es un punto de mínimo relativo, existe* $\delta > 0$ *tal que si* $|t| < \delta$, *entonces,* $f(c + tw) - f(c) \geq 0$. *Dado que* A *es abierto, existe* $\delta_1 > 0$ *con* $\delta_1 \leq \delta$ *tal que* $c + tw$ *pertenece a* A *para* $0 \leq t \leq \delta_1$. *Por el teorema de Taylor, existe* t_1 *con* $0 \leq t_1 \leq t \leq \delta_1$ *tal que si* $c_t = c + t_1 w$, *entonces*

$$f(c + tw) = f(c) + df_c + \frac{1}{2} D^2 f(c_t)(tw)^2.$$

Dado que c es un punto de mínimo relativo, se infiere que $df_c = 0$*; por lo tanto, se tiene*

$$\frac{1}{2}\, D^2 f(c_t)(tw)^2 \geq 0$$

para $0 \leq t \leq \delta_1$*. De aquí se deduce que* $D^2 f(c_t)(w)^2 \geq 0$*. Dado que* $\|c_t - c\| = |t_1| \leq |t|$*, entonces* $c_t \to c$ *cuando* $t \to 0$*. Como las segundas derivadas parciales de* f *son continuas, entonces,* $D^2 f(c)(w)^2 \geq 0$ *para toda* $w \in \mathbb{R}^n$ *con* $\|w\| = 1$*, de donde se obtiene el resultado.* ■

Teorema 300 *Sea* $A \subseteq \mathbb{R}^n$ *abierto, supongamos que* $f : A \to \mathbb{R}$ *tiene segundas derivadas parciales continuas en* A*, y sea* $c \in A$ *un punto crítico de* f*.*

a. *Si* $D^2 f(c)(w)^2 > 0$ *para toda* $w \in \mathbb{R}^n$*,* $w \neq 0$*, entonces* f *tiene un mínimo relativo estricto en c.*

b. *Si* $D^2 f(c)(w)^2 < 0$ *para toda* $w \in \mathbb{R}^n$*,* $w \neq 0$*, entonces* f *tiene un máximo relativo estricto en c.*

c. *Si* $D^2 f(c)(w)^2$ *toma valores estrictamente positivos así como estrictamente negativos para* $w \in \mathbb{R}^n$*, entonces* f *tiene un punto de ensilladura en c.*

Demostración:

a. *Por hipótesis,* $D^2 f(c)(w)^2 > 0$ *para* w *en el conjunto compacto*

$$\{w \in \mathbb{R}^n : \|w\| = 1\}$$

Dado que la aplicación $w \longmapsto D^2 f(c)(w)^2$ *es continua, existe* $m > 0$ *tal que*

$$D^2 f(c)(w)^2 \geq m \quad \text{para} \quad \|w\| = 1$$

Dado que las segundas derivadas parciales de f *son continuas en* A*, existe* $\delta > 0$ *tal que si* $\|x - c\| < \delta$ *entonces*

$$D^2 f(x)(w)^2 \geq \frac{1}{2} m \quad \text{para} \quad \|w\| = 1.$$

Por el teorema del Taylor, si $0 \leq t \leq 1$ *entonces hay un punto* c_t *en el segmento de línea que une a* c *con* $c + tw$ *tal que*

$$f(c + tw) = f(c) + df_c(tw) + \frac{1}{2}\, D^2 f(c_t)(tw)^2.$$

Dado que c es un punto crítico, se infiere que si $\|w\| = 1$ *y* $0 < t < \delta$*, entonces*

$$f(c + tw) - f(c) = \frac{1}{2} t^2 D^2 f(c_t)(w)^2 \geq \frac{1}{4} m t^2 > 0,$$

de modo que $f(c + u) > f(c)$ *para* $0 < \|u - c\| < \delta$*, y* f *tiene un mínimo relativo estricto en c. Por tanto, el apartado (a) está demostrado*

b. *La demostración es análoga a la del apartado (a).*

c. *Sean w_+, w_- vectores unitarios en $\mathbb{R}^n$ tales que*

$$D^2 f(c)(w_+)^2 > 0, \qquad D^2 f(c)(w_-)^2 < 0.$$

Del teorema de Taylor se deduce que para $t > 0$ suficientemente pequeña, se tiene

$$f(c + tw_+) > f(c), \qquad f(c + tw_-) < f(c).$$

Por lo tanto, c es un punto de ensilladura de f.

■

Observación 158 *Comparando los teoremas anteriores, se tiende a hacer las siguientes conjeturas:*

i. *Si $c \in A$ es un punto de mínimo relativo estricto, entonces $D^2 f(c)(w)^2 > 0$ para toda $w \in \mathbb{R}^n$, $w \neq 0$.*

ii. *Si $c \in A$ es un punto de ensilladura de f, entonces $D^2 f(c)(w)^2$ toma valores estrictamente positivos, así como estrictamente negativos.*

iii. *Si $D^2 f(c)(w)^2 \geq 0$ para toda $w \in \mathbb{R}^n$, entonces c es un punto de mínimo relativo.*

Observación 159 *Para poder utilizar el teorema anterior, es necesario saber si la forma cuadrática $w \longmapsto D^2 f(c)(w)^2$ es definida. Para cada $j = 1, 2, \ldots, n$, definimos Δ_j como el determinante de la matriz simétrica (llamada **matriz hessiana** cuando $j = n$)*

$$\begin{pmatrix} D_{11}f(c) & \cdots & D_{1j}f(c) \\ \vdots & \ddots & \vdots \\ D_{j1}f(c) & \cdots & D_{jj}f(c) \end{pmatrix}.$$

Tal como se puede ver en la unidad de formas cuadráticas de álgebra lineal,

i. *si los números $\Delta_1, \Delta_2, \ldots, \Delta_n$ son todos estrictamente positivos, entonces $D^2 f(c)(w)^2 > 0$ para toda $w \neq 0$ y f tiene un mínimo relativo estricto en c.*

ii. *si los números $\Delta_1, \Delta_2, \ldots, \Delta_n$ son en forma alternante estrictamente negativos y estrictamente positivos, entonces $D^2 f(c)(w)^2 < 0$ para toda $w \neq 0$ y f tiene un máximo relativo estricto en c.*

En otros casos puede haber puntos extremos o puntos de ensilladura.

En el caso especial e importante $n = 2$, se puede usar el siguiente resultado:

Sea $A \subseteq \mathbb{R}^n$ abierto, supongamos que $f : A \to \mathbb{R}$ tiene segundas derivadas parciales continuas en A, sea $c \in A$ un punto crítico de f y sea

$$\Delta = D_{11}f(c) \cdot D_{22}f(c) - [D_{12}f(c)]^2.$$

a. *Si $\Delta > 0$ y si $D_{11}f(c) > 0$, entonces f tiene un mínimo relativo estricto en c.*

b. *Si $\Delta > 0$ y si $D_{11}f(c) < 0$, entonces f tiene un máximo relativo estricto en c.*

c. *Si* $\Delta < 0$*, entonces tiene un punto de ensilladura en c.*

d. *Si* $\Delta = 0$*, entonces no se puede afirmar nada: c puede ser o no extremo.*

Ejemplo 353 *Encontrar los extremos de la función*

$$f(x, y, z) = x^2 + y^2 + z^2 + xyz.$$

Solución: *Para encontrar los posibles extremos debemos resolver el sistema*

$$\begin{aligned} 2x + yz &= 0 \\ 2y + xz &= 0 \\ 2z + xy &= 0 \end{aligned}$$

Las soluciones son

$$\begin{aligned} P_1 &= (0, 0, 0) \\ P_2 &= (2, 2, -2) \\ P_3 &= (2, -2, 2) \\ P_4 &= (-2, 2, 2) \\ P_5 &= (-2, -2, -2) \end{aligned}$$

Para decidir si son máximos o mínimos, calculamos la matriz de las derivadas segundas:

$$\begin{pmatrix} 2 & z & y \\ z & 2 & x \\ y & x & 2 \end{pmatrix}.$$

Para el punto $P_1 = (0, 0, 0)$*, tenemos:*

$$\Delta_1 = 2, \qquad \Delta_2 = 4, \qquad \Delta_3 = 8.$$

Por tanto, P_1 *es un mínimo relativo.*

Para $P_2 = (2, 2, -2)$*, tenemos:*

$$\Delta_1 = 2, \qquad \Delta_2 = 0, \qquad \Delta_3 = -32.$$

Por tanto, la derivada segunda no es ni definida positiva ni definida negativa, por lo que P_2 *no es un extremo relativo. Análogamente,* P_3, P_4 *y* P_5 *tampoco son extremos relativos.* ■

16.2. Problemas de extremos con restricciones

Hasta este momento se ha estado analizando el caso en que los extremos de la función $f : A \to \mathbb{R}$ pertenecen al interior de su dominio $A \subseteq \mathbb{R}^n$. Ninguna de las observaciones es aplicable a la localización de los extremos en la frontera.

Supongamos que S es una "superficie" contenida en el dominio A de la función f. A menudo es deseable encontrar los valores de f que sean máximo o mínimo entre todos los que se tienen en S. Supongamos que los puntos de S se pueden expresar como aquellos que satisfacen una relación de la forma.

$$g(x) = 0,$$

para una función $g : A \to \mathbb{R}$. En un problema de extremos con restricciones se quiere encontrar los valores extremos relativos de f para aquellos puntos $x \in A$ que satisfagan la **restricción** (o **condición lateral**) $g(x) = 0$. Si suponemos que f y g son de clase $C^1(A)$ y que $dg_c \neq 0$, entonces una condición necesaria para que c sea un punto extremo de f con respecto a los puntos x que satisfacen $g(x) = 0$ es que dg_c sea un múltiplo de df_c. En términos de derivadas parciales, esto significa que existe un número real λ tal que

$$\begin{aligned} D_1 f(c) &= \lambda D_1 g(c), \\ &\vdots \\ D_n f(c) &= \lambda D_n g(c). \end{aligned}$$

En la práctica se desea determinar las p coordenadas del punto c que satisfagan esta condición necesaria. Sin embargo, tampoco se conoce el número real λ, al que por lo general se le llama el **multiplicador de Lagrange**. Las n ecuaciones dadas antes, junto con la ecuación

$$g(c) = 0$$

se resuelven para las $n + 1$ cantidades desconocidas, de las cuales las coordenadas de c son de interés primordial.

Teorema 301 ***(Teorema de Lagrange)*** *Sea $A \subseteq \mathbb{R}^n$ abierto y supongamos que $f, g : A \to \mathbb{R}$ son de clase $C^1(A)$. Supongamos que $c \in A$ cumple $g(c) = 0$, y que existe un entorno U de c tal que*

$$f(x) \leq f(c) \qquad (\text{ó } f(x) \geq f(c))$$

para todos los puntos $x \in U$ que satisfacen $g(x) = 0$. Entonces, existen números reales μ, λ, no nulos a la vez, tales que

$$\mu\, df_c = \lambda dg_c.$$

Más aún, si $dg_c \neq 0$, entonces se puede tomar $\mu = 1$.
Demostración: *Supongamos que $F : U \to \mathbb{R}^2$ está definida por*

$$F(x) = (f(x), g(x)) \quad \text{para } x \in U,$$

de manera que F es de clase $C^1(U)$ y

$$dF_x(v) = (df_x(v), dg_x(v)), \qquad x \in U, v \in \mathbb{R}^n.$$

Además, observamos que un punto $x \in U$ satisface la restricción $g(x) = 0$ si y sólo si $F(x) = (f(x), 0)$.

Si $f(x) \leq f(c)$ para toda $x \in U$ que satisfaga $g(x) = 0$, entonces los puntosde la forma $(r, 0)$ con $f(c) < r$ no pertenecen a $F(U)$; por lo tanto, dF_c no es una aplicación exhaustiva de $\mathbb{R}^n$ sobre $\mathbb{R}^2$. Pero dado que es una aplicación lineal, la imagen de la transformación lineal dF_c está contenida en alguna recta en $\mathbb{R}^2$ que pasa por $(0, 0)$. Por lo tanto, existe un punto $(\lambda, \mu) \neq (0, 0)$ tal que la imagen de dF_c está contenida en la recta que pasa por $(0, 0)$ y (λ, μ). Así pues, se tiene

$$\mu\, df_c(v) = \lambda dg_c(v) \text{ para toda } v \in \mathbb{R}^n,$$

lo que implica que

$$\mu\, df_c = \lambda dg_c.$$

Por último, supongamos que $dg_c \neq 0$. *Si* $\mu = 0$, *entonces obtenemos*

$$0 = \lambda dg_c,$$

lo que implica que $\lambda = 0$. *Esto contradice el hecho de que* $(\mu, \lambda) \neq (0, 0)$. *Por lo tanto, en este caso se debe tener* $\mu \neq 0$ *y así podemos dividir entre* μ *y reemplazar* λ/μ *por* λ. ■

Dado que $U \subseteq \mathbb{R}^n$, la ecuación

$$\mu\, df_c(v) = \lambda dg_c(v) \text{ para toda } v \in \mathbb{R}^n,$$

tomando $v = e_1, \ldots, e_n$ origina el sistema de n ecuaciones:

$$\begin{aligned}
\mu\, D_1 f(c) &= \lambda\, D_1 g(c), \\
&\vdots \\
\mu\, D_n f(c) &= \lambda\, D_n g(c).
\end{aligned}$$

Si alguna $D_i g(c), i = 1, \ldots, n$, es no nula, entonces se puede tomar $\mu = 1$ para obtener el sistema

$$\begin{aligned}
D_1 f(c) &= \lambda\, D_1 g(c), \\
&\vdots \\
D_n f(c) &= \lambda\, D_n g(c).
\end{aligned}$$

Observación 160 *El teorema de Lagrange da sólo una condición necesaria, y los puntos que se obtienen al resolver las ecuaciones (lo que a menudo es difícil hacer) pueden ser máximos relativos, mínimos relativos o ninguno de los dos. Sin embargo, con frecuencia se pueden usar los resultados de la sección anterior para probar si son máximos o mínimos relativos. Más aún, en muchas aplicaciones, la determinación de si los puntos son realmente extremos se puede basar en consideraciones geométricas o físicas.*

Ejemplo 354 *Encontrar el punto del plano* $\{(x, y, z) : 2x + 3y - z = 5\} \subset \mathbb{R}^3$ *más cercano al origen.*

Solución: *Para resolver este problema, minimizaremos la función que da el cuadrado de la distancia al origen*

$$f(x, y, z) = x^2 + y^2 + z^2,$$

con la restricción

$$g(x, y, z) = 2x + 3y - z - 5 = 0.$$

Dado que $dg_c \neq 0$ *para toda* $c \in \mathbb{R}^3$, *aplicando el teorema de Lagrange obtenemos el sistema*

$$\begin{aligned}
2x &= 2\lambda \quad \Rightarrow \quad x = \lambda \\
2y &= 3\lambda \quad \Rightarrow \quad y = \frac{3}{2}\lambda \\
2z &= -\lambda \quad \Rightarrow \quad z = -\frac{1}{2}\lambda \\
2x + 3y - Z - 5 &= 0
\end{aligned}$$

Por lo tanto, al substituir x, y, z se obtiene

$$2\lambda + 3\left(\frac{3}{2}\lambda\right) - \left(-\frac{1}{2}\lambda\right) - 5 = 0 \quad \Rightarrow \quad 14\lambda = 4\lambda + 9\lambda + \lambda = 10 \quad \Rightarrow \quad \lambda = \frac{5}{7}.$$

Por lo tanto, el punto del plano más cercano a $(0,0,0)$ es $(\frac{5}{7}, \frac{15}{14}, -\frac{5}{14})$. ■

Ejemplo 355 *Encontrar las dimensiones de la caja rectangular, abierta por arriba, con volumen máximo y superfície dada A.*

Solución: *Sean x, y, z las dimensiones de la caja con z como altura. Entonces, se desea maximizar la función*

$$V(x, y, z) = xyz$$

sujeta a la restricción

$$g(x, y, z) = xy + 2xz + 2yz - A = 0.$$

Dado que el punto deseado tendrá coordenadas estrictamente positivas, aplicando el teorema de Lagrange obtenemos el sistema

$$\begin{aligned} yz &= \lambda(y + 2z), \\ xz &= \lambda(x + 2z), \\ xy &= \lambda(2x + 2y), \\ xy + 2xz + 2yz - A &= 0. \end{aligned}$$

Si se multiplican las tres primeras ecuaciones por $x, y,$ y z, respectivamente,se igualan y se dividen por λ (ya que $\lambda \neq 0$), se llega a

$$xy + 2xz = xy + 2yz = 2xz + 2yz.$$

La primera igualdad implica $x = y$, y la segunda implica $y = 2z$. Por tanto, $x = 2z$. De la última ecuación se deduce que

$$4z^2 + 4z^2 + 4z^2 = A \quad \Rightarrow \quad 12z^2 = A \quad \Rightarrow \quad z = \frac{1}{2}\sqrt{\frac{A}{3}}.$$

Por consiguiente, el volumen de la caja es

$$V = xyz = 2\left(\frac{1}{2}\sqrt{\frac{A}{3}}\right) \cdot 2\left(\frac{1}{2}\sqrt{\frac{A}{3}}\right) \cdot \frac{1}{2}\sqrt{\frac{A}{3}} = \frac{1}{2}\left(\sqrt{\frac{A}{3}}\right)^3.$$

■

Con frecuencia hay más de una restricción; en este caso, el siguiente resultado es útil.

Teorema 302 *Sea $A \subseteq \mathbb{R}^n$ abierto y supongamos que f y $g_1, \ldots, g_k$ son funciones reales de clase $C^1(A)$. Supongamos que $c \in A$ satisface las restricciones*

$$g_1(x) = 0, \ldots, g_k(x) = 0,$$

y que existe un entorno abierto U de a tal que $f(x) \leq f(c)$ (o $f(x) \geq f(c)$) para toda $x \in U$ que satisfaga estas restricciones. Entonces existen números reales $\mu, \lambda_1, \ldots, \lambda_k$ no todos iguales a cero tales que

$$\mu\, df_c = \lambda_1 (dg_1)_c + \cdots + \lambda_k (dg_k)_c.$$

Demostración: *Definimos la función $F : U \to \mathbb{R}^{k+1}$ de la forma*

$$F(x) = (f(x), g_1(x), \ldots, g_k(x)) \quad \textit{para} \quad x \in U,$$

y usamos el mismo argumento que en la demostración del teorema de Lagrange. ■

Corolario 44 *Además de las hipótesis del teorema anterior, supongamos que el rango de la matriz*

$$\begin{pmatrix} D_1 g_1(c) & \cdots & D_1 g_k(c) \\ \vdots & \ddots & \vdots \\ D_n g_1(c) & \cdots & D_n g_k(c) \end{pmatrix}$$

es igual a k $(\leq n)$. Entonces existen números reales $\lambda_1, \ldots, \lambda_k$ no todos iguales a cero tales que

$$\begin{aligned} D_1 f(c) &= \lambda_1 D_1 g_1(c) + \cdots + \lambda_k D_1 g_k(c), \\ &\vdots \\ D_n f(c) &= \lambda_1 D_n g_1(c) + \cdots + \lambda_k D_n g_k(c). \end{aligned}$$

Demostración: *Si aplicamos la fórmula*

$$\mu\, df_c = \lambda_1 (dg_1)_c + \cdots + \lambda_k (dg_k)_c.$$

a la base canónica $e_1, \ldots, e_n \in \mathbb{R}^n$, obtenemos el sistema de ecuaciones

$$\begin{aligned} \mu\, D_1 f(c) &= \lambda_1 D_1 g_1(c) + \cdots + \lambda_k D_1 g_k(c), \\ &\vdots \\ \mu\, D_n f(c) &= \lambda_1 D_n g_1(c) + \cdots + \lambda_k D_n g_k(c). \end{aligned}$$

Si $\mu = 0$, entonces el supuesto de que el rango es igual a k implica que $\lambda_1 = \cdots = \lambda_k = 0$, contrario a la hipótesis. Por lo tanto, $\mu \neq 0$ y podemos dividir todas la ecuaciones por μ. Llamando de nuevo λ_i a λ_i/μ, se obtiene el sistema del enunciado. ■

Ejemplo 356 *Encontrar los puntos en la intersección del cilindro*

$$\left\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 4\right\}$$

y el plano

$$\left\{(x, y, z) \in \mathbb{R}^3 : 6x + 3y + 2z = 6\right\}$$

más cercanos al origen y los más lejanos al origen.

Solución: *Buscaremos los extremos relativos de la función distancia elevada al cuadrado*

$$f(x,y,z) = x^2 + y^2 + z^2$$

sujeta a las restricciones

$$\begin{aligned} g_1(x,y,z) &= x^2 + y^2 - 4 = 0, \\ g_2(x,y,z) &= 6x + 3y + 2z - 6 = 0. \end{aligned}$$

Como la matriz

$$\begin{pmatrix} 2x & 6 \\ 2y & 3 \\ 0 & 2 \end{pmatrix}$$

tiene rango 2 *(excepto en el punto* $(x,y) = (0,0)$*, que no satisface las restricciones), podemos aplicar el corolario, y obtenemos el sistema*

$$\begin{aligned} 2x &= \lambda_1 \cdot 2x + \lambda_2\, 6, \\ 2y &= \lambda_1 \cdot 2y + \lambda_2\, 3, \\ 2z &= \lambda_2\, 2, \\ x^2 + y^2 &= 4, \\ 6x + 3y + 2z &= 6, \end{aligned}$$

de cinco ecuaciones con cinco variables. De la tercera ecuación obtenemos

$$\lambda_2 = z,$$

de tal manera que se puede sustituir λ_2 *en las dos primeras ecuaciones:*

$$\begin{aligned} 2x &= \lambda_1 \cdot 2x + 6z \\ 2y &= \lambda_1 \cdot 2y + 3z \end{aligned}$$

Para sustituir λ_1*, multiplicamos la primera ecuación por* y *y la segunda por* x*, y restándolas obtenemos*

$$0 = 6yz - 3xz \quad \Rightarrow \quad 0 = z(2y - x).$$

De aquí se deduce que $z = 0$ *o bien* $x = 2y$.

Si $z = 0$*, la quinta ecuación da* $2x + y = 2$*. Al combinarlo con la cuarta ecuación, se obtiene*

$$\begin{aligned} x^2 + (2-2x)^2 &= x^2 + 4 - 8x + 4x^2 = 4 \quad \Rightarrow \quad 5x^2 - 8x = 0 \quad \Rightarrow \\ &\Rightarrow \quad x = 0 \quad ó \quad x = \frac{8}{5} \end{aligned}$$

Esto nos lleva a los puntos $(0,2,0)$ *y* $(8/5, -6/5, 0)$*, que están ambos a distancia* 2 *del origen.*

Por otro lado, si $x = 2y$*, la cuarta ecuación da* $5y^2 = 4$*, de tal manera que* $y = 2/\sqrt{5}$ *(y* $x = 4/\sqrt{5}$*) o* $y = -2/\sqrt{5}$ *(y* $x = -4/\sqrt{5}$*). Sustituyendo en la quinta ecuación, se obtiene* $z = 3(1 - \sqrt{5})$ *y* $z = 3(1 + \sqrt{5})$*, respectivamente. Por lo tanto, este caso nos lleva a los puntos* $(4/\sqrt{5}, 2/\sqrt{5}, 3(1-\sqrt{5}))$ *y* $(-4/\sqrt{5}, -2/\sqrt{5}, 3(1+\sqrt{5}))$*. La distancia entre estos puntos y el origen vale* $\sqrt{58 - 18\sqrt{5}}$ *y* $\sqrt{58 + 18\sqrt{5}}$*, respectivamente.*

Por tanto, los puntos $(0,2,0)$ *y* $(8/5, -6/5, 0)$ *son lo que minimizan la distancia desde el origen a la intersección del enunciado, y el punto* $(-4/\sqrt{5}, -2/\sqrt{5}, 3(1+\sqrt{5}))$ *maximiza dicha distancia.* ■