OPTIMIZACION
PARA ESTUDIANTES DE INGENIERIA[1]

Jorge Amaya A.

9 de mayo de 2003

Departamento de Ingeniería Matemática y
Centro de Modelamiento Matemático

Universidad de Chile

Mayo de 2003

[1]Texto preliminar, destinado exclusivamente al uso de los estudiantes del curso **MA37A Optimización,** de la Escuela de Ingeniería de la Universidad de Chile, semestre Otoño 2003.
Los comentarios son bienvenidos: *jamaya@dim.uchile.cl*

Capítulo 1

Introducción

Un problema de optimización matemático, en términos generales, se escribe de la forma:

$$\begin{array}{c}\text{(P) minimizar} \quad \text{(o maximizar)} \quad f(x) \\ x \in S\end{array}$$

donde x es el **vector de variables de decisión**, $f : S \longrightarrow I\!R$ es la **función objetivo** y $S \subseteq I\!R^n$ es el **conjunto factible**. A menudo

$$S = \{x \in I\!R^n / g_i(x) \leq 0 \quad i = 1, ...m, \ x \in X \subseteq I\!R^n\}$$

y se dice que las expresiones $g_i(x) \leq 0 \quad i = 1, ..., m$ representan el **conjunto de restricciones** del problema (P). Si $S = I\!R^n$, el problema se dirá **irrestricto.**

Un vector $x \in I\!R^n$ que pertenezca al conjunto S se llamará **solución factible** de (P). Si además satisface que

$$f(x) \leq f(y) \quad \forall y \in S$$

(cuando se trata de un problema de minimización), se dirá que x es **solución óptima**.

Nos parece importante recordar aquí el teorema de Weierstrass, que da condiciones para la existencia de solución para un problema de minimización con características bien particulares.

Teorema 1.0.1 *Si f es una función real, continua sobre un conjunto compacto (cerrado y acotado) $K \subseteq I\!R^n$, entonces el problema*

$$\begin{array}{c} min \ \ (o \ \ max) \ f(x) \\ x \in K \end{array}$$

tiene una solución óptima $\overline{x} \in K$.

Dependiendo de las características particulares del problema, éste recibe nombres y tratamientos especiales para su resolución. Dos casos de interés, son el de la **programación lineal** (f y g_i son funciones lineales afines $\forall i$) y la **programación lineal entera** (en que además las variables sólo toman valores enteros). También trataremos la teoría y técnicas de solución de un problema con funciones no lineales.

Un concepto esencial para entender cómo plantear y resolver un problema de optimización es el de **convexidad**. Mostrar algo de la teoría básica del análisis convexo y su vinculación con la teoría de optimización son los objetivos del siguiente capítulo.

Capítulo 2

Convexidad

2.1 Conjuntos convexos

Definición 2.1.1 *Sea $S \subseteq \mathbb{R}^n$, $S \neq \phi$. Se dice que S es* **convexo** [1] *si y sólo si*

$$\lambda x + (1-\lambda) y \in S \quad \forall x, y \in S,\ \forall \lambda \in [0,1]$$

Geométricamente, esta definición se puede interpretar como sigue: *un conjunto no vacío es convexo si dados dos puntos del conjunto, el segmento de recta que los une está contenido en dicho conjunto* (ver figura 2.1).

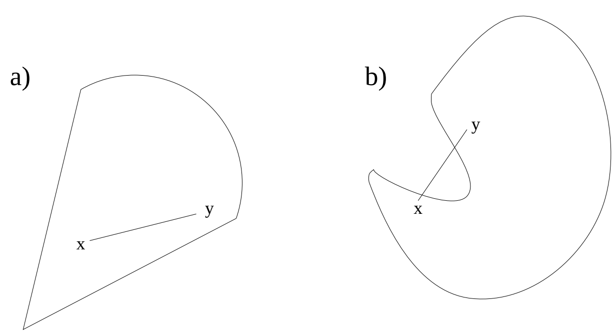

Figura 2.1: El conjunto de la figura a) es convexo. El conjunto de la figura b) no es convexo, pues existe un segmento de recta, uniendo dos puntos del conjunto, que no está incluido en el conjunto.

[1] Por convención, el conjunto vacío será considerado convexo

Ejemplo 2.1.1 *Un espacio vectorial es un conjunto convexo.*

Demostración. Directo pues, por definición, un espacio vectorial es cerrado para la suma y la ponderación por escalar. $\blacksquare$

En particular $\mathbb{R}^n$ es un convexo.

Ejemplo 2.1.2 $S = \{x \in \mathbb{R}^3 \ / \ x_1 + 2x_2 - x_3 = 2\}$ *es un conjunto convexo.*

Demostración. Sean x e $y \in S$, $\lambda \in [0,1]$. Por definición del conjunto S, esto significa que $x_1 + 2x_2 - x_3 = 2$ e $y_1 + 2y_2 - y_3 = 2$.

Veamos que $\lambda x + (1-\lambda)y = \begin{bmatrix} \lambda x_1 + (1-\lambda)y_1 \\ \lambda x_2 + (1-\lambda)y_2 \\ \lambda x_3 + (1-\lambda)y_3 \end{bmatrix}$ pertenece a S,pues

$$\begin{gathered}\lambda x_1 + (1-\lambda)y_1 + 2\{\lambda x_2 + (1-\lambda)y_2\} - \{\lambda x_3 + (1-\lambda)y_3\} = \\ \lambda(x_1 + 2x_2 - x_3) + (1-\lambda)(y_1 + 2y_2 - y_3) = 2\lambda + 2(1-\lambda) = 2 \quad \blacksquare\end{gathered}$$

Definición 2.1.2 *Sean $a \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ fijos. Se llama* **hiperplano** *al conjunto*

$$H = \{x \in \mathbb{R}^n / a^t x = \alpha\}$$

Un hiperplano H define dos **semiespacios***:*

$$H^- = \{x \in \mathbb{R}^n / a^t x \leq \alpha\}$$

$$H^+ = \{x \in \mathbb{R}^n / a^t x \geq \alpha\}$$

Por ejemplo, en el caso $H = \{x \in \mathbb{R}^3 / x_1 + 2x_2 - x_3 = 2\}$ se tiene $a^t = (1, 2, -1)$ y $\alpha = 2$. Los dos semiespacios asociados son:

$$H^- = \{x \in \mathbb{R}^3 / x_1 + 2x_2 - x_3 \leq 2\}$$

$$H^+ = \{x \in \mathbb{R}^3 / x_1 + 2x_2 - x_3 \geq 2\}.$$

Ejemplo 2.1.3 *Un semiespacio S en $\mathbb{R}^n$ es un conjunto convexo.*

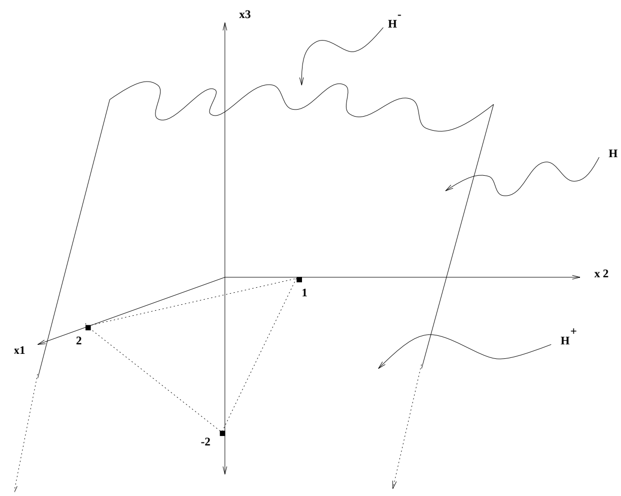

Figura 2.2: Semiespacios generados por el hiperplano H

Demostración. Consideremos $a \in I\!R^n$ y $\alpha \in I\!R$, definiendo el semiespacio $S = \{x \in I\!R^n / a^t x \leq \alpha\}$.

Sean $x, y \in S$, $\lambda \in [0,1]$, entonces

$$a^t(\lambda x + (1-\lambda)y) = \lambda(a^t x) + (1-\lambda)(a^t y) \leq \lambda\alpha + (1-\lambda)\alpha = \alpha$$

Luego $\lambda x + (1-\lambda)y \in S$ y, por lo tanto, S es convexo. ■

Proposición 2.1.1 *Sean S_1 y S_2 dos conjuntos convexos. Entonces $S_1 \cap S_2$ es un conjunto convexo.*

Demostración. Sean $x, y \in S_1 \cap S_2$, $\lambda \in [0,1]$

$$x, y \in S_1 \Rightarrow \lambda x + (1-\lambda)y \in S_1 \text{ , ya que } S_1 \text{ es convexo.}$$
$$x, y \in S_2 \Rightarrow \lambda x + (1-\lambda)y \in S_2 \text{ , ya que } S_2 \text{ es convexo.}$$

luego $\lambda x + (1-\lambda)y \in S_1 \cap S_2$, es decir, $S_1 \cap S_2$ es convexo. ■

Observación 2.1.1 *Observemos que:*

i) Esta propiedad se puede generalizar fácilmente a una intersección cualquiera de convexos. Esto es, si Γ es un conjunto arbitrario, incluso no numerable, y $\{S_\gamma\}_{\gamma\in\Gamma}$ es una clase de conjuntos convexos, entonces $\cap_{\gamma\in\Gamma}S_\gamma$ es un conjunto convexo.

ii) Aunque del ejemplo (2.1.3) puede concluirse fácilmente que un hiperplano es un conjunto convexo (reemplazando las desigualdades por igualdades), podemos usar esta proposición para probar que un hiperplano es un conjunto convexo, dado que es intersección de dos semiespacios (convexos).

Ejemplo 2.1.4 *Sistema de desigualdades lineales:*

$$\begin{gathered} a_{11}x_1 + ... + a_{1n}x_n \leq b_1 \\ \vdots \\ a_{m1}x_1 + ... + a_{mn}x_n \leq b_m \end{gathered}$$

con $a_{ij}, b_i, x_j \in I\!R$, $\forall i \in \{1, ..., m\}$, $j \in \{1, ..., n\}$

El sistema se anota $Ax \leq b$; con $A = (a_{ij})_{i=1...m;j=1...n}$, $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

El conjunto $S = \{x \in I\!R^n / Ax \leq b\}$ es la intersección de n semiespacios de la forma $S_i = \{x \in I\!R^n / A_{i\bullet}x \leq b_i\}$ (donde $A_{i\bullet}$ denota la fila i-ésima de la matriz A) los cuales, según vimos en el ejemplo (2.1.3), son conjuntos convexos. Luego, por la proposición (2.1.1), S es convexo.

Definición 2.1.3 *Sean* $x_1, ..., x_k \in I\!R^n$ *;* $\lambda_1, ..., \lambda_k \in I\!R$ *tales que* $\lambda_i \geq 0$ $\forall i$, $\sum\limits_{i=1}^{k} \lambda_i = 1$. *El vector* $x = \sum\limits_{i=1}^{k} \lambda_i x_i$ *se dice* **combinación convexa** *de los* k *vectores.*

Definición 2.1.4 *Sea* $S \subseteq I\!R^n$. *Se define el conjunto* $co(S)$, **envoltura convexa** *de* S, *de la manera siguiente:*

$$co(S) = \{x = \sum_{i=1}^{k} \lambda_i x_i \ / \ k \in I\!N, \ \{x_i\}_{i=1}^{k} \subseteq S \ , \ \{\lambda_i\}_{i=1}^{k} \subseteq [0,1] \ , \ \sum_{i=1}^{k} \lambda_i = 1\}.$$

Es decir, todo punto en $co(S)$ *es combinación convexa de puntos de* S.

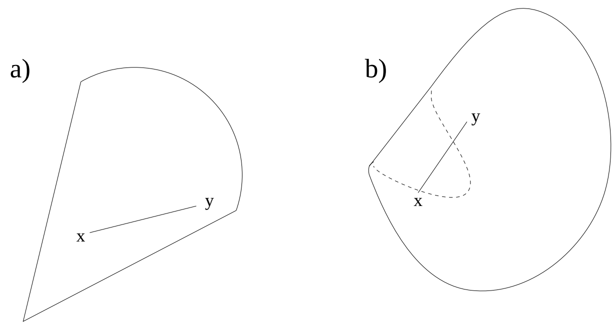

Figura 2.3: La envoltura convexa del conjunto de la figura a) coincide con él, por ser convexo. Para el conjunto de la figura b), la línea sólida corresponde a su envoltura convexa.

Observación 2.1.2 *Si S es convexo, entonces $co(S) = S$ (ver figura 2.3).*

Ejemplo 2.1.5 *La envoltura convexa de los números racionales es $I\!R$.*

Ejemplo 2.1.6 *Sean* $v_1 = \begin{bmatrix} 5 \\ 2 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 6 \\ 7 \\ 3 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, v_4 = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}$ *vectores en $I\!R^3$. Su envoltura convexa queda determinada por el poliedro de la figura (2.4), cuyos vértices están dados por el conjunto de vectores $\{v_1, v_2, v_3, v_4\}$.*

Proposición 2.1.2 *$co(S)$ es un conjunto convexo.*

Demostración. Sean $x, y \in co(S)$, es decir, $x = \sum\limits_{i=1}^{k} \nu_i x_i$, $y = \sum\limits_{i=1}^{m} \mu_i y_i$ donde $\{x_i\}_{i=1}^k \subseteq S$, $\{y_i\}_{i=1}^m \subseteq S$ y $\{\nu_i\}_{i=1}^k$, $\{\mu_i\}_{i=1}^m$ son ponderadores de la combinación convexa.
Sea $\lambda \in [0, 1]$

$$\lambda x + (1-\lambda) y = \lambda \sum_{i=1}^{k} \nu_i x_i + (1-\lambda) \sum_{i=1}^{m} \mu_i y_i.$$

Llamando

$$\begin{array}{lll} \widetilde{x}_i = x_i, & \widetilde{\lambda}_i = \lambda \nu_i & i \in \{1, ..., k\} \\ \widetilde{x}_{k+i} = y_i, & \widetilde{\lambda}_{k+i} = (1-\lambda)\mu_i & i \in \{1, ..., m\}, \end{array}$$

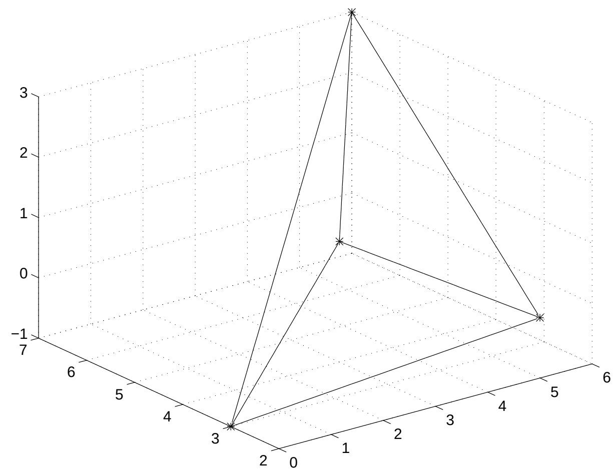

Figura 2.4: La envoltura convexa del conjunto de puntos señalados queda determinada por un poliedro cuyos vértices están dados por el conjunto de vectores.

se tiene que

$$\lambda x + (1-\lambda)y = \sum_{i=1}^{k+m} \widetilde{\lambda}_i \widetilde{x}_i$$

con

$$\{\widetilde{x}_i\}_{i=1}^{k+m} \subseteq S,\ \widetilde{\lambda}_i \in [0,1] \quad \forall i = 1, ..., k+m \quad \text{y} \quad \sum_{i=1}^{k+m} \widetilde{\lambda}_i = 1.$$

Luego por definición se tiene que, $\lambda x + (1-\lambda)y \in co(S)$, por lo tanto $co(S)$ es convexo. ■

Proposición 2.1.3 *$co(S)$es el convexo más pequeño (en el sentido de la inclusión) que contiene a S.*

Demostración. Basta probar que una caracterización equivalente es la siguiente:

$$co(S) = \bigcap \{C/C \ \ convexo,\ \ S \subseteq C\}$$

Sea $\mathcal{B} = \bigcap \{C/C \ \ convexo \ \ S \subseteq C\}$

i) Sea $x \in \mathcal{B}$, entonces para todo C convexo tal que $S \subseteq C$ se tiene que $x \in C$. Como $co(S)$ es convexo y $S \subseteq co(S)$ entonces $x \in co(S)$. Por lo tanto $\mathcal{B} \subseteq co(S)$

ii) Sea ahora $x \in co(S)$, demostremos que $x \in \mathcal{B}$. Tomemos C, un conjunto convexo cualquiera que contenga a S.

Como $x \in co(S)$, entonces $x = \sum_{i=1}^{k} \lambda_i x_i$, con $x_i \in S$, $\lambda_i \geq 0, \forall i = 1, ..., k$, y $\sum_{i=1}^{k} \lambda_i = 1$

En particular $x_i \in C$, para todo i. Así, x también pertenece a C, cualquiera sea éste. Así $co(S) \subseteq \mathcal{B}$ ■

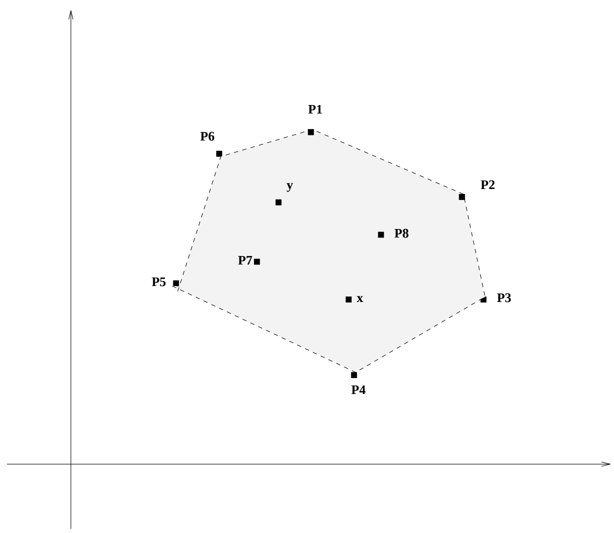

Figura 2.5: Teorema de Carathéodory.

Mostraremos a continuación una interesante propiedad de la envoltura convexa, conocida como Teorema de Carathéodory. Tomemos el ejemplo de la figura (2.5) en que

$$S = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\}$$

La envoltura convexa de S es el conjunto delimitado por las líneas punteadas. Se observa que el punto x, que pertenece a $co(S)$, pertenece también al triángulo definido por los puntos p_3, p_4 y p_5. El punto y, que pertenece a $co(S)$, está también incluído en el triángulo determinado por lo puntos p_1, p_2 y p_5 o bien en el triángulo definido por los puntos p_1, p_6 y p_4. En realidad, cualquier punto de $co(S)$ está contenido en algún triángulo definido por puntos de S. Esto se expresa de la forma siguiente.

Teorema 2.1.1 (Carathéodory) *Sean* $S \subseteq I\!R^n$ *y* $x \in co(S)$. *Entonces existen* $x_1, ..., x_{n+1} \in S$ *tales que* $x \in co\{x_1, ..., x_{n+1}\}$.

Demostración. Sea $x \in co(S)$, luego existen $\{x_i\}_{i=1}^k \subseteq S$, $\{\lambda_i\}_{i=1}^k \subseteq [0,1]$ satisfaciendo $\sum\limits_{i=1}^k \lambda_i = 1$, tales que $x = \sum\limits_{i=1}^k \lambda_i x_i$.

Sin pérdida de generalidad, podemos suponer $\lambda_i \neq 0 \ \forall i$ (si $\lambda_j = 0$ para algún j, x es combinación convexa de $k-1$ vectores)

Supongamos que $k > n+1$.

Los $k-1$ vectores $x_2 - x_1, ..., x_k - x_1$ son linealmente dependientes (pues $k-1 > n$, la dimensión del espacio). Es decir, existen $\mu_2, ..., \mu_k \in I\!R$, no todos nulos, tales que

$\sum\limits_{i=2}^k \mu_i (x_i - x_1) = 0.$

Sea $\mu_1 = -\sum\limits_{i=2}^k \mu_i$ (*)

$\sum\limits_{i=2}^k \mu_i (x_i - x_1) = \sum\limits_{i=2}^k \mu_i x_i - \sum\limits_{i=2}^k \mu_i x_1 = \sum\limits_{i=2}^k \mu_i x_i + \mu_1 x_1 = \sum\limits_{i=1}^k \mu_i x_i = 0$, con $\sum\limits_{i=1}^k \mu_i = 0.$

Entonces, $\forall \alpha \in I\!R \quad x = \sum\limits_{i=1}^k \lambda_i x_i - \alpha 0 = \sum\limits_{i=1}^k \lambda_i x_i - \alpha \sum\limits_{i=1}^k \mu_i x_i = \sum\limits_{i=1}^k (\lambda_i - \alpha\mu_i) x_i$

De (*) se deduce que $\mu_i > 0$ para al menos un i, luego podemos escoger

$$\alpha = \min_i \{ \tfrac{\lambda i}{\mu_i} \ / \ \mu_i > 0 \} = \tfrac{\lambda_I}{\mu_I}.$$

Notar que $\alpha > 0$ y que $\lambda_i - \alpha\mu_i \geq 0 \ \forall i$. (pues si $\mu_i \leq 0 \Rightarrow \lambda_i - \alpha\mu_i > 0$ y si $\mu_i > 0 \Rightarrow \lambda_i - \alpha\mu_i = \mu_i(\frac{\lambda_i}{\mu_i} - \alpha) \geq 0$)

En particular, $\lambda_I - \alpha\mu_I = 0$.

Así, $x = \sum\limits_{i=1}^k (\lambda_i - \alpha\mu_i) x_i$, donde $\sum\limits_{i=1}^k (\lambda_i - \alpha\mu_i) = 1$ y $\lambda_i - \alpha\mu_i \geq 0 \ \forall i$.

Es decir,x es combinación convexa de a lo más $k-1$ vectores (pues el coeficiente $I-$*ésimo* es nulo).

Repitiendo el argumento $k-(n+1)$ veces, se obtiene que x es combinación convexa de $n+1$ puntos. ■

Ejercicio 2.1.1 *Sean S_1 y S_2 convexos, $\alpha \in I\!R$. Se define la suma y ponderación de conjuntos como sigue:*

- $S_1 + S_2 = \{x + y \ / \ x \in S_1, y \in S_2\}$

- $\alpha S_1 = \{\alpha x \ / \ x \in S_1\}$

Pruebe que $S_1 + S_2$ y αS_1 son convexos.

2.1.1 Poliedros

Notación: Notaremos por $\mathcal{M}_{m\times n}(I\!R)$ al conjunto de las matrices de m filas y n columnas, a coeficientes reales.

Definición 2.1.5 *Se llama* **poliedro** *a un conjunto de la forma $\mathcal{P} = \{x \in I\!R^n / Ax \leq b\}$ con $A \in \mathcal{M}_{m\times n}(I\!R)$ y $b \in I\!R^m$, es decir, un poliedro es una intersección finita de semiespacios.*

Proposición 2.1.4 *$\mathcal{P}' = \{x \in I\!R^n / Ax = b \ , \ x \geq 0\}$ es un poliedro.*[2]

Demostración. Claramente, el conjunto $\mathcal{P}'$ queda representado por el siguiente sistema de ecuaciones lineales :

$$\begin{bmatrix} A \\ -A \\ -I \end{bmatrix} x \leq \begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$$

donde I la matriz identidad en dimensión n.

Llamando A'=$\begin{bmatrix} A \\ -A \\ -I \end{bmatrix}$, b'= $\begin{pmatrix} b \\ -b \\ 0 \end{pmatrix}$, se obtiene un sistema de la forma $\mathcal{P} = \{x \in I\!R^n / A'x \leq b'\}$, que es igual a $\mathcal{P}'$. Luego, $\mathcal{P}'$ es un poliedro. ■

Observación 2.1.3 *Es obvio que $\mathcal{P}'$=$\{x \in I\!R^n / Ax \geq b\}$ es un poliedro. En efecto: como x$\in I\!R^n$ es irrestricto, basta multiplicar el sistema de desigualdades por -1, y definir A'=-A, b'=-b.*

[2] $x \geq 0$ si y solamente si $x_i \geq 0 \ \forall i = 1, ..., n$

Proposición 2.1.5 *Todo poliedro es un conjunto convexo.*

Demostración. Ver ejemplo (2.1.4) $\blacksquare$

Se dirá que un poliedro $\mathcal{P} = \{x \in I\!R^n / Ax = b, x \geq 0\}$ está escrito en <u>forma canónica</u>. En lo sucesivo trabajaremos con esta representación.

Proposición 2.1.6 *Un poliedro es un conjunto cerrado.*

Demostración. Sea $\mathcal{P}$ el poliedro $\{x \in I\!R^n / Ax = b,\ x \geq 0\}$ y consideremos $\bar{x} \in \bar{\mathcal{P}}$ (adherencia o cerradura de $\mathcal{P}$). Mostraremos que $\bar{x} \in \bar{\mathcal{P}}$.
Como $\bar{x} \in \bar{\mathcal{P}}$, existe una sucesión $\{x_k\}$ en $\mathcal{P}$ tal que $\lim_{k\to\infty} x_k = \bar{x}$.
Además, $\forall k \geq 0$, el punto x_k verifica

$$\begin{array}{rcl} Ax_k & = & b \\ x_k & \geq & 0 \end{array}$$

Tomando límite (y por continuidad de la función lineal $x \longmapsto Ax$) se tiene:

$$\begin{array}{rcl} A\bar{x} & = & b \\ \bar{x} & \geq & 0 \end{array}$$

Luego $\bar{x} \in \mathcal{P}$ y por lo tanto $\bar{\mathcal{P}} \subseteq \mathcal{P}$. Dado que se cumple siempre que $\mathcal{P} \subseteq \bar{\mathcal{P}}$, se obtiene $\bar{\mathcal{P}} = \mathcal{P}$, luego $\mathcal{P}$ es cerrado. $\blacksquare$

Demostración alternativa. Sea $g : I\!R^n_+ \to I\!R^n$; $g(x) = Ax - b$. A es una forma lineal, luego la función g es lineal afín y, por lo tanto, continua en $I\!R^n_+$. El conjunto $\{x \in I\!R^n / Ax = b\}$ es pre-imagen continua de un cerrado ($\{0\}$), luego es cerrado y como $\mathcal{P}$ es igual a la intersección de este conjunto con el semiespacio cerrado $\{x \geq 0\}$, se concluye que es cerrado. $\blacksquare$

Ejemplo 2.1.7 $C = \{x \in I\!R^2 / -x_1 + x_2 \leq 2; x_1 + x_2 \geq 4; x_2 \leq 4; x_1 \geq 0; x_2 \geq 0\}$.

Matricialmente esto puede escribirse de la siguiente manera:

$$\begin{bmatrix} -1 & 1 \\ -1 & -1 \\ 0 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \leq \begin{pmatrix} 2 \\ -4 \\ 4 \\ 0 \\ 0 \end{pmatrix} .$$

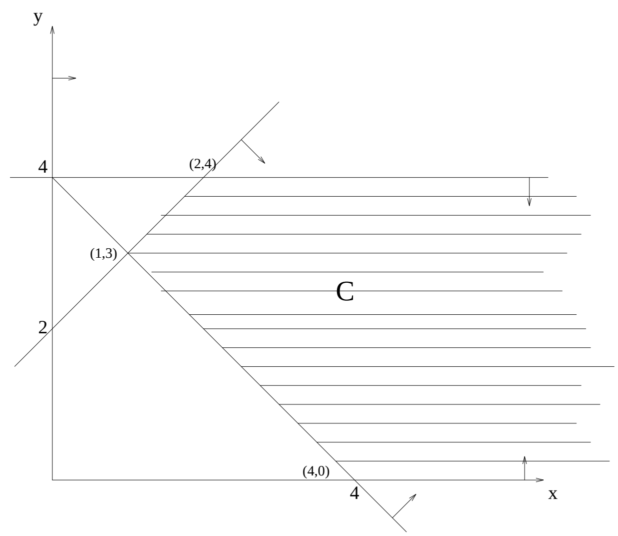

Figura 2.6: El conjunto C es un poliedro, convexo y cerrado, pero no acotado.

El conjunto C es un poliedro, convexo y cerrado, pero no acotado.

Definición 2.1.6 *Sea $S \subseteq I\!R^n$ un conjunto convexo, $S \neq \phi$. Un vector $x \in S$ se llama* **punto extremo de S** *si y sólo si no puede ser representado como combinación convexa de otros dos puntos distintos del convexo. Es decir, si $x = \lambda x_1 + (1-\lambda)x_2$, con $x_1, x_2 \in S$ y $\lambda \in]0,1[$, entonces $x = x_1 = x_2$.*

Ejemplo 2.1.8 .

a) Sea S=B(0,1), la bola unitaria en $I\!R^n$. El conjunto de puntos extremos queda representado por $\{x \in I\!R^n / \|x\| = 1\}$, que es la frontera de S

b) El conjunto de puntos extremos del poliedro del ejemplo (2.1.6) es

$$E = \left\{ \begin{bmatrix} 5 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 6 \\ 7 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} \right\}$$

c) El conjunto de puntos extremos de un semiespacio cerrado es vacío.

Ejemplo 2.1.9 *Sean* $U = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ *y* $S = co\{U\}$

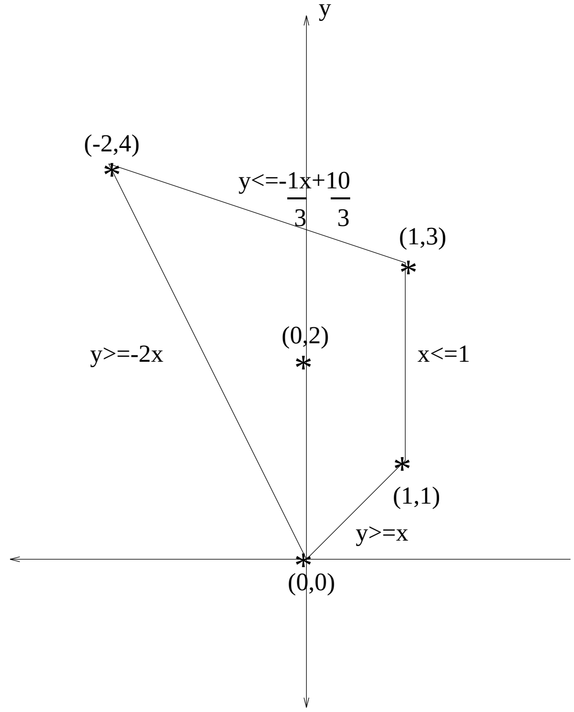

Figura 2.7: S es la envoltura convexa del conjunto U

Naturalmente, el conjunto de puntos extremos de S es $\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} -2 \\ 4 \end{pmatrix} \right\}$.

El sistema que representa a S es

$$S = \left\{ (x_1, x_2) \in I\!R^2 : \begin{bmatrix} 1 & -1 \\ -2 & -1 \\ 1 & 0 \\ \frac{1}{3} & 1 \end{bmatrix} \leq \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{10}{3} \end{bmatrix} \right\}$$

En general, fácilmente se puede ver que x es punto extremo de un convexo S si y solamente si $S \backslash \{x\}$ es un conjunto convexo, de donde se sigue que si S^* es tal que $co(S^*) = S$, entonces necesariamente S^* debe incluir al conjunto de puntos extremos de S.

La definición de punto extremo es de suma importancia en la teoría de optimización pues, como veremos más adelante, está en relación directa con el conjunto de soluciones para un problema de programación lineal, donde el conjunto de restricciones determina precisamente un poliedro. De aquí se desprende que es necesario tener una caracterización simple para estos puntos.

Veamos el siguiente ejemplo de motivación

$$\mathcal{P} = \{x \in I\!R^2 / x_1 + x_2 \leq 2,\ \ 8x_1 + 3x_2 \leq 8,\ \ x_1, x_2 \geq 0\}$$

El gráfico se muestra en la figura (2.8)

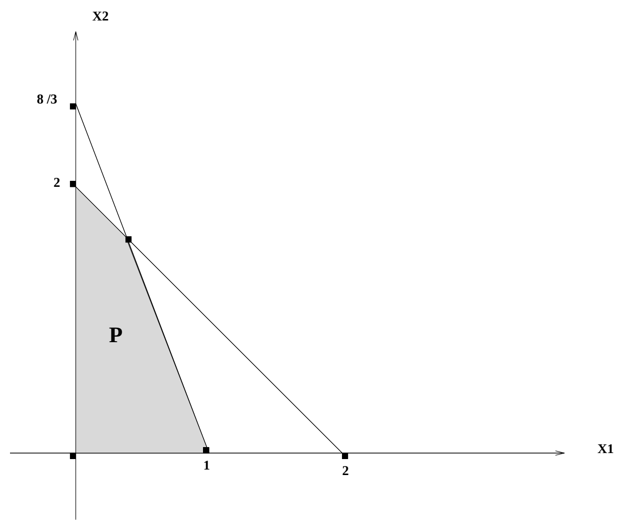

Figura 2.8: Ejemplo de motivación.

Los puntos extremos son $\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 2/5 \\ 8/5 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$. Trabajaremos con el poliedro (en $I\!R^4$)

$$\mathcal{P}' = \{x \in I\!R^4 / x_1 + x_2 + x_3 = 2, 8x_1 + 3x_2 + x_4 = 8, x_1, x_2, x_3, x_4 \geq 0\}$$

que es equivalente a $\mathcal{P}$ en el sentido siguiente: $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathcal{P} \iff \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \in \mathcal{P}'$ con $x_3, x_4 \geq 0$ Examinemos entonces el sistema

$$\begin{array}{rcl} x_1 + x_2 + x_3 \qquad & = & 2 \\ 8x_1 + 3x_2 \qquad + x_4 & = & 8 \\ x_1, x_2, x_3, x_4 & \geq & 0 \end{array}$$

Asignando valor nulo a dos variables cualesquiera podemos entonces resolver el sistema de dos ecuaciones cada vez. Esto da las soluciones

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ 2 \\ 8 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 8/3 \\ -2/3 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 0 \\ 8 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2/5 \\ 8/5 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Se observa que dos de ellas (la tercera y la cuarta) no satisfacen la condición de positividad, luego no pertenecen a $\mathcal{P}'$. Sin embargo las cuatro soluciones restantes determinan en sus dos primeras coordenadas, los puntos extremos de $\mathcal{P}$, a saber

$$\left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2/5 \\ 8/5 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

Esto se expresa en forma general en el siguiente teorema

Teorema 2.1.2 *Sea un poliedro* $\mathcal{P} = \{x \in I\!R^n / Ax = b, x \geq 0\}$, *donde* $A \in \mathcal{M}_{m\times n}(I\!R)$ *es de rango* m *y* $b \in I\!R^m$. *Un punto* x *es extremo de* $\mathcal{P}$ *si y sólo si la matriz* A *se puede descomponer, eventualmente reordenando sus columnas, en la forma* $A = [B, N]$, *donde* $B \in \mathcal{M}_{m\times m}(I\!R)$ *es invertible,* $N \in \mathcal{M}_{m\times(n-m)}(I\!R)$ *corresponde a las columnas restantes y* $x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$, *con* $B^{-1}b \geq 0$.

Demostración.($\Leftarrow$) Sea $x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} \geq 0$. Se tiene que $x \in \mathcal{P}$, pues

$Ax = [B, N] \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = BB^{-1}b + N0 = b.$

Sean $u, v \in \mathcal{P}$ tales que $x = \lambda u + (1-\lambda)v$, $\lambda \in [0, 1]$, es decir

$$\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = \lambda \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + (1-\lambda) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

De donde:

(1) $\lambda u_1 + (1-\lambda)v_1 = B^{-1}b$
(2) $\lambda u_2 + (1-\lambda)v_2 = 0$

Como $u, v \in \mathcal{P}$ necesariamente $u \geq 0, v \geq 0$. Luego de (2) se tiene que $u_2 = v_2 = 0$.

Como $u \in \mathcal{P}$ satisface $Au = b$, esto es $[B, N] \begin{pmatrix} u_1 \\ 0 \end{pmatrix} = Bu_1 = b \Rightarrow u_1 = B^{-1}b$. Por lo tanto, $u = x$.

De la misma manera se prueba que $v = x$, con lo que se concluye que x es punto extremo.

$(\Rightarrow)$ Supongamos que $x \in \mathcal{P}$ es un punto extremo, x puede escribirse $x = \begin{pmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, eventualmente reordenando las columnas del sistema, con $x_i > 0$, $i = 1, ..., k$, $k \in \{1, ..., m\}$

Notemos por $A_{\bullet k}$ el k-ésimo vector columna de A. Luego $A = [A_{\bullet 1}, ..., A_{\bullet k}, A_{\bullet k+1}, ..., A_{\bullet n}]$

$$Ax = b \Longleftrightarrow \sum_{i=1}^{k} x_i A_{\bullet i} = b$$

Probaremos que $A_{\bullet 1}, ..., A_{\bullet k}$ son linealmente independientes. Supongamos que son linealmente dependientes, es decir, que existen $\mu_1, ..., \mu_k$ no todos nulos, tales que $\sum_{i=1}^{k} \mu_i A_{\bullet i} = 0$.

Definamos $\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ y, para $\alpha > 0$, construyamos los siguientes vectores

$$\begin{aligned} y &= x + \alpha\mu \\ z &= x - \alpha\mu \end{aligned}$$

De allí: $x = \frac{1}{2}y + \frac{1}{2}z$. Es claro que $y, z \in \mathcal{P}$, para α suficientemente pequeño y además $x \neq y$, $y \neq z$, $z \neq x$. Por lo tanto x es combinación convexa de dos puntos distintos en $\mathcal{P}$, luego no es extremo (contradicción).

Así, $A_{\bullet 1}, ..., A_{\bullet k}$ son linealmente independientes, lo que implica, en particular, que $k \leq m$ Podemos agregar $A_{\bullet k+1}, ..., A_{\bullet m}$ (eventualmente reordenando columnas) para obtener un conjunto maximal (A es de rango m) y definir $B = [A_{\bullet 1}, ..., A_{\bullet m}]$, que es una matriz invertible, y $N = [A_{\bullet m+1}, ..., A_{\bullet n}]$. Con esto, A tiene la forma $A = [B, N]$.

Luego, $Ax = b \Leftrightarrow \sum_{i=1}^{n} x_i A_{\bullet i} = b \Leftrightarrow \sum_{i=1}^{m} x_i A_{\bullet i} + \sum_{i=m+1}^{n} x_i A_{\bullet i} = b$

Notando $x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, con $x_B > 0, x_N = 0$, la ecuación anterior se escribe:

$$Bx_B + Nx_N = Bx_B = b$$
$$\text{de donde } x_B = B^{-1}b \quad \blacksquare$$

Corolario 2.1.1 *El número de puntos extremos de un poliedro en la forma canónica es finito.*

Demostración. Hay a lo sumo $\begin{pmatrix} n \\ m \end{pmatrix}$ formas de elegir las m columnas independientes de A, y cada matriz B está asociada a lo más a un punto extremo. $\blacksquare$

Ejemplo 2.1.10 *Consideremos un poliedro en la forma canónica dado por las matrices* $A = \begin{bmatrix} 2 & 1 & 0 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}$ *y* $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$
Calculemos sus puntos extremos.

De acuerdo al corolario anterior, existen a lo sumo 6 puntos extremos dado que hay 6 formas posibles de elegir la matriz B.

(1) $B = \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$ no es invertible.

(2) $B = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$ es invertible, pero $B^{-1}b = \begin{bmatrix} \frac{1}{2} \\ -1 \end{bmatrix}$ no es un vector positivo.

(3) $B = \begin{bmatrix} 2 & -1 \\ 0 & 2 \end{bmatrix}$ es invertible y el vector $B^{-1}b = \begin{bmatrix} \frac{3}{4} \\ \frac{1}{2} \end{bmatrix}$ tiene todas sus coodenadas positivas.

(4) $B = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ es invertible, pero $B^{-1}b = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ no es un vector positivo.

(5) $B = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$ es invertible y el vector $B^{-1}b = \begin{bmatrix} \frac{3}{2} \\ \frac{1}{2} \end{bmatrix}$ tiene todas sus coodenadas positivas.

(6) $B = \begin{bmatrix} 0 & -1 \\ -1 & 2 \end{bmatrix}$ es invertible, pero $B^{-1}b = \begin{bmatrix} -3 \\ -1 \end{bmatrix}$ no es un vector positivo.

Los casos (3) y (5) nos entregan puntos extremos para el poliedro en estudio, sólo falta ubicar los valores resultantes en las posiciones correctas:

- La matriz del caso (3) toma las columnas primera y cuarta de la matriz A, luego el vector punto extremo correspondiente a este caso será $\begin{bmatrix} \frac{3}{4} \\ 0 \\ 0 \\ \frac{1}{2} \end{bmatrix}$
- La matriz del caso (5) toma las columnas segunda y cuarta de la matriz A, luego el vector punto extremo correspondiente a este caso será $\begin{bmatrix} 0 \\ \frac{3}{2} \\ 0 \\ \frac{1}{2} \end{bmatrix}$

Definición 2.1.7 *Se llama* **polítopo** *a la envoltura convexa de un conjunto finito de puntos.*

Definición 2.1.8 *Sean $x_1, ..., x_k \in I\!R^n$. El polítopo asociado a este conjunto de puntos es un* **simplex** *si y sólo si el conjunto $\{x_2 - x_1, ..., x_k - x_1\}$ es linealmente independiente.*

De acuerdo con estas definiciones, el conjunto S del ejemplo (2.1.9) es un polítopo y puede concluirse fácilmente que todo polítopo es envoltura convexa de sus puntos extremos.

Es obvio, además, que todo polítopo es un poliedro. Luego, parece natural preguntarse si todo poliedro puede escribirse como combinación convexa de sus puntos extremos. La respuesta es negativa, cuando el poliedro es no acotado. En el ejemplo (2.1.7) observamos que cualquier punto que no esté en la superficie del triángulo de vértices $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$, $\begin{pmatrix} 2 \\ 4 \end{pmatrix}$ y $\begin{pmatrix} 4 \\ 0 \end{pmatrix}$, no puede ser expresado como combinación convexa de esos tres puntos extremos.

Ejemplo 2.1.11 $S = \{x \in I\!R^2 / x_2 \geq |x_1|\}$

$x_2 \geq |x_1| \Longleftrightarrow x_2 \geq x_1 \geq -x_2$; $x_2 \geq 0$, por lo tanto

$$S = \{x \in I\!R^2 / x_1 - x_2 \leq 0; -x_1 - x_2 \leq 0; x_2 \geq 0\}$$

o, en forma matricial

$$\begin{bmatrix} 1 & -1 \\ -1 & -1 \\ 0 & 1 \end{bmatrix} x \leq 0$$

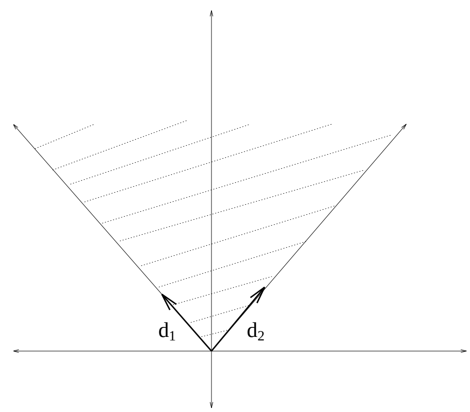

Figura 2.9: El punto (0,0) es el único punto extremo del poliedro de la figura. d_1 y d_2 son sus únicas direcciones extremas.

Como es posible ver en la figura (2.9), $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ es el <u>único</u> punto extremo y ningún punto del poliedro S puede expresarse como combinación convexa de puntos extremos. Luego, para poliedros no acotados se hace necesario definir un nuevo concepto:

Definición 2.1.9 *Un vector* $d \in I\!R^n$, $d \neq 0$, *se dice* **dirección** *de* S *si y sólo si* $\forall x \in S$ *se tiene que* $x + \lambda d \in S \quad \forall \lambda \geq 0$.

Consideremos el poliedro $\mathcal{P} = \{x \in I\!R^n / Ax = b, x \geq 0\}$. Una dirección de $\mathcal{P}$ debe satisfacer que $\forall x \in \mathcal{P}$:

$$\begin{array}{rl} A(x + \lambda d) = b & \forall \lambda \geq 0 \\ x + \lambda d \geq 0 & \forall \lambda \geq 0 \end{array}$$

Luego, d es dirección de $\mathcal{P}$ si y solamente si satisface el sistema $Ad = 0, \ d \geq 0$.

Definición 2.1.10 *Dos direcciones d_1 y d_2 se dirán* **iguales** *si y sólo si $d_1 = \alpha d_2$ para algún $\alpha > 0$.*

Observación 2.1.4 *Se escribirá $d_1 = d_2$, si no hay posible confusión.*

Definición 2.1.11 *Sea S un convexo cerrado y $d \in I\!R^n$ una dirección de S. Se dice que d es* **dirección extrema** *si dadas d_1 y d_2, direcciones de S, tales que $d = \alpha d_1 + \beta d_2$ para algún $\alpha, \beta > 0$, necesariamente se tiene que $d = d_1 = d_2$.*

Es decir, d no puede expresarse como combinación lineal positiva de otras dos direcciones distintas.

Ejemplo 2.1.12 *En la figura (2.9), d_1 y d_2 son direcciones extremas y toda otra dirección se escribe como combinación lineal positiva de ellas.*

Con lo que hemos hecho hasta aquí, una pregunta interesante es: ¿Existirá alguna caracterización de las direcciones extremas, equivalente a la obtenida para puntos extremos?

Escribamos la matriz A que representa el poliedro escrito en la forma canónica tal como en el caso de puntos extremos, es decir, $A = [B, N]$ y consideremos $d = \begin{pmatrix} -B^{-1}a_j \\ e_j \end{pmatrix}$ con $B^{-1}a_j \leq 0$, donde a_j es columna de N. Verifiquemos que d es dirección: en efecto, $d \geq 0$ y $\quad Ad = [B, N] \begin{pmatrix} -B^{-1}a_j \\ e_j \end{pmatrix} = -BB^{-1}a_j + Ne_j = -a_j + a_j = 0.$

Supongamos que no es extrema, es decir, que existen d_1 *y* d_2 direcciones de $\mathcal{P}$ distintas, tales que d es combinación lineal positiva de ellas: $d = \lambda_1 d_1 + \lambda_2 d_2$, para $\lambda_1, \lambda_2 > 0$. Entonces d_1 y d_2 tendrán la forma:

$$d_1 = \begin{bmatrix} d_{11} \\ \eta_1 e_j \end{bmatrix}, \quad d_2 = \begin{bmatrix} d_{21} \\ \eta_2 e_j \end{bmatrix}$$

para algún $\eta_1, \eta_2 \geq 0$. Como d_1 *y* d_2 son direcciones de $\mathcal{P}$ entonces $Ad_1 = Ad_2 = 0$. Luego

$$[B, N] \begin{bmatrix} d_{11} \\ e_j \end{bmatrix} = Bd_{11} + \eta_1 Ne_j = Bd_{11} + \eta_1 a_j = 0 \Rightarrow d_{11} = -\eta_1 B^{-1}a_j$$

$$[B, N] \begin{bmatrix} d_{21} \\ e_j \end{bmatrix} = Bd_{21} + \eta_2 Ne_j = Bd_{21} + \eta_2 a_j = 0 \Rightarrow d_{21} = -\eta_2 B^{-1}a_j$$

por lo tanto $d_1 = d_2 = d$ (en el sentido de la igualdad de direcciones), lo que muestra que d es dirección extrema.

Lo explicado anteriormente nos permite formular el siguiente teorema de caracterización de direcciones extremas.

Teorema 2.1.3 *Sea un poliedro $\mathcal{P} = \{x \in I\!R^n / Ax = b; x \geq 0\}$, donde $A \in \mathcal{M}_{m\times n}(I\!R)$ es de rango m y $b \in I\!R^m$. Una dirección $\overline{d} \in I\!R^n$ es dirección extrema de $\mathcal{P}$ si y sólo si la matriz A se puede descomponer, eventualmente reordenando sus columnas, en la forma $A = [B, N]$, donde $B \in \mathcal{M}_{m\times m}(I\!R)$ es invertible y $\overline{d}$ es un múltiplo positivo de $d = \begin{pmatrix} -B^{-1}a_j \\ e_j \end{pmatrix}$ con $B^{-1}a_j \leq 0$, donde $a_j \in N$ (vector columna de N) y e_j es el j-ésimo vector de la base canónica de $I\!R^{n-m}$.*

Corolario 2.1.2 *El número de direcciones extremas de un poliedro en la forma canónica es finito.*

Demostración. Hay a lo más $\begin{pmatrix} n \\ m \end{pmatrix}$ formas de elegir B^{-1} y como hay $n-m$ columnas en N, entonces $(n-m)\begin{pmatrix} n \\ m \end{pmatrix}$ es el número máximo de direcciones extremas. ■

Ejemplo 2.1.13 *Volvamos al ejemplo (2.1.10).*

De acuerdo al corolario anterior, existen 12 posibles direcciones extremas, por lo tanto, no desarrrollaremos el cálculo completo, sólo consideraremos el siguiente caso:

Tomemos la matriz B formada por la segunda y cuarta columnas de A.

$$B = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \text{ luego } N = \begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}$$

$$B^{-1}N = \begin{bmatrix} 2 & -\frac{1}{2} \\ 0 & -\frac{1}{2} \end{bmatrix}$$

El producto de B^{-1} con la primera columna de N no es negativo, por lo tanto, no nos permite calcular una dirección extrema. Sin embargo, el producto con la segunda columna de N es negativo. Tal como en el caso de puntos extremos, sólo basta ordenar la información para decir que $d = \begin{bmatrix} 0 \\ -\frac{1}{2} \\ 1 \\ -\frac{1}{2} \end{bmatrix}$ es dirección extrema del poliedro. $\blacksquare$

Para concluir esta sección, enunciaremos, sin demostrar, un teorema de caracterización que liga todo lo que hemos desarrollado hasta ahora.

Teorema 2.1.4 *Sea $\mathcal{P} = \{x \in I\!R^n / Ax = b,\, x \geq 0\}$, donde $A \in \mathcal{M}_{m\times n}(I\!R)$ es de rango m y $b \in I\!R^n$. Sean $x_1, ..., x_k$ los puntos extremos y $d_1, ..., d_l$ las direcciones extremas de $\mathcal{P}$. Entonces, $x \in \mathcal{P}$ si y sólo si puede ser escrito como la suma de una combinación convexa de los puntos extremos y una combinación lineal positiva de las direcciones extremas, es decir,*

$$x = \sum_{i=1}^{k} \lambda_i x_i + \sum_{j=1}^{l} \mu_j d_j$$

donde $\lambda_i \in [0,1] \quad \forall i = 1, ..., k,\ \sum\limits_{i=1}^{k} \lambda_i = 1; \quad \mu_j \geq 0, \forall j = 1, ..., l.$

Teorema 2.1.5 *$\mathcal{P} = \{x \in I\!R^n / Ax = b; x \geq 0\} \neq \phi$ tiene al menos una dirección extrema si y sólo si $\mathcal{P}$ es no acotado.*

Demostración.($\Rightarrow$) Si $\mathcal{P}$ tiene una dirección extrema, claramente es no acotado, pues $x + \lambda d \in \mathcal{P} \quad \forall x \in \mathcal{P}, \quad \forall \lambda \geq 0 \quad y \quad \lim\limits_{\lambda\to\infty} \|x + \lambda d\| = \infty$.

($\Leftarrow$) Supongamos que $\mathcal{P}$ es no acotado y que no posee direcciones extremas. Luego, por el teorema anterior, todo punto $x \in \mathcal{P}$ puede escribirse de la forma $x = \sum\limits_{i=1}^{k} \lambda_i x_i$, con $\lambda_i \in [0,1], \forall i = 1, ..., k, \quad \sum\limits_{i=1}^{k} \lambda_i = 1$.

Por la desigualdad de Cauchy-Schwartz

$$\|x\| = \left\| \sum_{i=1}^{k} \lambda_i x_i \right\| \leq \sum_{i=1}^{k} \lambda_i \|x_i\| \leq \sum_{i=1}^{k} \|x_i\| < \infty \quad \forall x \in \mathcal{P}$$

lo que contradice el supuesto de que $\mathcal{P}$ es no acotado. $\blacksquare$

Ejercicio 2.1.2 *Sea S un convexo. Demuestre que $x \in S$ es punto extremo si y sólo si $S \backslash \{x\}$ es convexo.*

Ejercicio 2.1.3 *Probar que todo polítopo es un poliedro.*

2.1.2 Teoremas de Proyección

Teorema 2.1.6 *Sea S un conjunto convexo, cerrado, no vacío en $I\!R^n$, $y \in I\!R^n$, $y \notin S$. Entonces, existe un único $\overline{x} \in S$ tal que minimiza la función*

$$\begin{array}{rcl} \varphi_y : S & \longrightarrow & I\!R \\ x & \longmapsto & \varphi_y(x) = \|y - x\| \end{array}$$

Demostración.

<u>Existencia</u>: Sea $\gamma = \inf\{\varphi_y(x) \ / \ x \in S\}$. Existe una sucesión minimizante $\{x_n\}_{n\in\aleph} \subseteq S$ tal que $\varphi_y(x_n) \to \gamma$, $n \to \infty$.

Usemos la propiedad conocida como ley del paralelógramo ($\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2$)

$$\begin{aligned} \|x_n - x_m\|^2 &= \|x_n - y + y - x_m\|^2 = 2\|x_n - y\|^2 + 2\|y - x_m\|^2 - \|x_n + x_m - 2y\|^2 \\ &= 2\|x_n - y\|^2 + 2\|y - x_m\|^2 - 4\left\|\tfrac{x_n + x_m}{2} - y\right\|^2 \end{aligned}$$

Notar que $\frac{x_n + x_m}{2} = \frac{1}{2}x_n + \frac{1}{2}x_m \in S$ (convexo), luego $\left\|\frac{x_n+x_m}{2} - y\right\|^2 \geq \gamma^2$, por lo tanto,

$$\|x_n - x_m\|^2 \leq 2\|x_n - y\|^2 + 2\|y - x_m\|^2 - 4\gamma^2 \ (*)$$

Si $n \to \infty, m \to \infty$, se tiene que $\|x_n - y\| \to \gamma^2$ y $\|x_m - y\| \to \gamma^2$, luego $\|x_n - x_m\|^2 \to 0$, es decir, $\{x_n\}_{n\in\aleph}$ es una sucesión de Cauchy en $I\!R^n$ y, por lo tanto, converge a $\overline{x} = \lim x_n \in S$ (cerrado).

Por continuidad de la norma, $\varphi_y(\overline{x}) = \gamma$.

<u>Unicidad</u>: Sea $\overline{\overline{x}} \in S$, $\overline{\overline{x}} \neq \overline{x}$ tal que $\varphi_y(\overline{\overline{x}}) = \gamma$.

Por (*) se tiene que $\left\|\overline{x} - \overline{\overline{x}}\right\|^2 \leq 2\|\overline{x} - y\|^2 + 2\left\|y - \overline{\overline{x}}\right\|^2 - 4\gamma^2 = 0$, luego $\overline{x} = \overline{\overline{x}}$. ■

Definición 2.1.12 *Sea S un convexo cerrado no vacío.*

i) Para $y \in I\!R^n$, se define la **distancia de** y **a** S*, por $d(y, S) = \min\{\varphi_y(x) \ / \ x \in S\}$.*

ii) Dado $y \in I\!R^n$, se define la **proyección** *de y sobre S, por $P_S(y) = \arg\min\{\varphi_y(x)/x \in S\}$.*

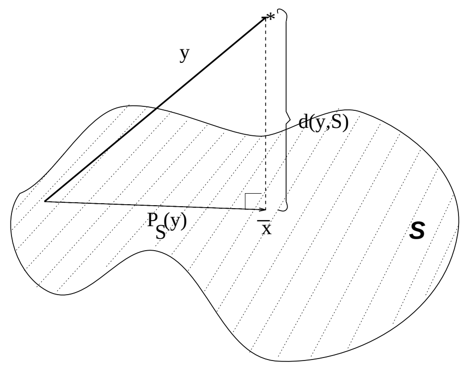

Figura 2.10: Distancia y proyección del punto y al conjunto S

Observación 2.1.5 *Claramente, si $y \in S$, $P_S(y) = y$.*

Teorema 2.1.7 *Sea S un convexo cerrado no vacío, $y \notin S$. Se tiene que*

$$\langle y - \overline{x}, x - \overline{x} \rangle \leq 0, \ \ \forall x \in S$$

si y solamente si $\overline{x}$ minimiza $\varphi_y(x)$.

Demostración. Supongamos primero que para cierto $\overline{x} \in I\!R^n$ se tiene $\langle y - \overline{x}, x - \overline{x} \rangle \leq 0 \quad \forall x \in S$. Calculemos:

$$\|y - x\|^2 = \|y - \overline{x} - (x - \overline{x})\|^2 = \|y - \overline{x}\|^2 + \|x - \overline{x}\|^2 - 2\langle y - \overline{x}, x - \overline{x} \rangle \geq$$
$$\|y - \overline{x}\|^2 + \|x - \overline{x}\|^2 \geq \|y - \overline{x}\|^2$$

Lo que implica que:

$$\|y - \overline{x}\| \leq \|y - x\| \qquad \forall x \in S$$

Es decir $\varphi_y(\overline{x}) \leq \varphi_y(x) \ \forall x \in S$.
Inversamente, tenemos que, si x minimiza φ_y en S, entonces $\forall x \in S$:

$$\|y - \overline{x}\|^2 = \|y - \overline{x} + \overline{x} - x\|^2 = \|y - \overline{x}\|^2 + \|\overline{x} - x\|^2 + 2\rangle y - \overline{x}, \overline{x} - x\langle$$

De donde,

$$\langle y - \overline{x}, x - \overline{x} \rangle \leq \tfrac{1}{2}\|x - \overline{x}\|^2 \qquad \forall x \in S$$

Como S es un conjunto convexo y $\overline{x} \in S$ podemos cambiar x por $\lambda x + (1-\lambda)\overline{x}$, con lo que queda:

$$\langle \lambda x + (1-\lambda)\overline{x} - \overline{x}, y - \overline{x} \rangle \leq \tfrac{1}{2} \| \lambda x + (1-\lambda)\overline{x} - \overline{x} \|$$

Que es lo mismo que

$$\langle x - \overline{x}, y - \overline{x} \rangle \leq \tfrac{\lambda}{2} \| x - \overline{x} \|^2$$

Tomando $\lambda \to 0^+$, se tiene el resultado ■

Geométricamente, el teorema anterior quiere decir que la proyección de y sobre S se alcanza en un punto $\overline{x}$ tal que el trazo $y - \overline{x}$ es ortogonal al conjunto.

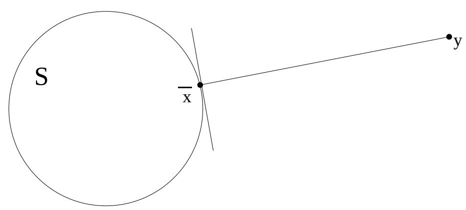

Figura 2.11: La proyección de y sobre S se alcanza en un punto $\overline{x}$ tal que el trazo $y - \overline{x}$ es ortogonal al conjunto.

Teorema 2.1.8 *Sea S un convexo cerrado no vacío, entonces* $\|P_S(x) - P_S(y)\| \leq \|x - y\|$ $\forall x, y$.

Observación 2.1.6 *Esto es equivalente a decir que si S un convexo cerrado no vacío, la función de proyección $P_S(x)$ es Lipschitz continua.*

Ejercicio 2.1.4 *Demuestre el teorema (2.1.8).*

2.1.3 Teoremas de Separación

Teorema 2.1.9 (Hahn-Banach) *Sea S un convexo cerrado no vacío, $y \notin S$. Existe $p \neq 0$, $\alpha \in I\!R$ tal que $p^t y > \alpha$ y $p^t x \leq \alpha$ $\forall x \in S$.*

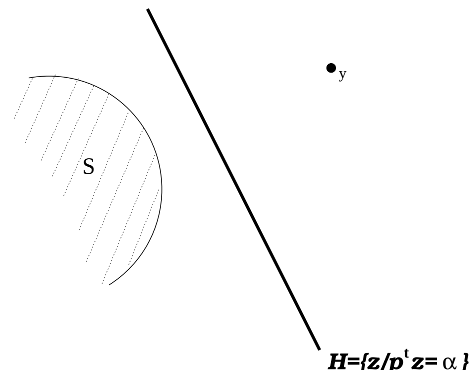

Figura 2.12: $\mathcal{H}$ es el hiperplano separador entre S e y.

Este p define lo que se conoce como **hiperplano separador**, $\mathcal{H} = \{z \in I\!R^n / p^t z = \alpha\}$ (ver figura 2.12)

Demostración. De acuerdo a lo desarrollado en la subsección anterior, existe un único $\overline{x} \in S$ tal que $\langle y - \overline{x}, x - \overline{x}\rangle \leq 0$, $\forall x \in S$.

Sea $p = y - \overline{x} \neq 0$, $\langle p, x - \overline{x}\rangle \leq 0$, $\forall x \in S \Rightarrow \langle p, x\rangle \leq \langle p, \overline{x}\rangle$, $\forall x \in S$ (*)

$\langle p, x - \overline{x}\rangle = \langle p, x - y + y - \overline{x}\rangle = \langle p, x - y\rangle + \langle p, y - \overline{x}\rangle = \langle p, x - y\rangle + \|p\|^2 \leq 0$, $\forall x \in S$

$\Rightarrow \langle p, x\rangle + \|p\|^2 \leq \langle p, y\rangle, \forall x \in S$. Como $\|p\|^2 \neq 0$, se tiene que $\langle p, x\rangle < \langle p, y\rangle$, $\forall x \in S$ (**)

Sea $\alpha = \langle p, \overline{x}\rangle$. Por (*) $\langle p, x\rangle \leq \alpha$ $\forall x \in S$ y por (**) $\alpha < \langle p, y\rangle$ lo que concluye la demostración. ■

Definición 2.1.13 *Sea S un convexo cerrado no vacío. Un* **hiperplano soportante** *de S es un hiperpalno H tal que $H \cap S \neq \phi$ y $\{S \subseteq H^+ \vee S \subseteq H^-\}$ (ver figura 2.13).*

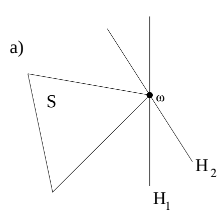

Figura 2.13: Para la figura a), H_1 y H_2 son hiperplanos soportantes en el punto señalado.

Cuando definimos los poliedros, los caracterizamos como una intersección finita de semiespacios. El siguiente teorema nos permitirá deducir una caracterización similar para un conjunto convexo no vacío cualquiera.

Teorema 2.1.10 *Sea $S \subseteq \mathbb{R}^n$ un conjunto convexo y sea $\overline{x}$ un punto en la frontera de S. Entonces S tiene un hiperplano soportante en $\overline{x}$.*

Corolario 2.1.3 *Sea S un convexo cerrado no vacío. Entonces*

$$S = \cap\{W \ \textit{semiespacio}/S \subseteq W\}$$

Observación 2.1.7 *Note que la intersección anterior no es necesariamente finita.*

Demostración. Basta tomar los semiespacios generados por todos los hiperplanos soportantes del convexo, que contengan a S. ■

Teorema 2.1.11 (Farkas) *Sea $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, $c \in \mathbb{R}^n$. Uno y sólo uno de los siguientes sistemas tiene solución:*

(1) $Ax \leq 0$, $c^t x > 0$, algún $x \in \mathbb{R}^n$.

(2) $A^t y = c$, $y \geq 0$, algún $y \in \mathbb{R}^n$.

Demostración. Supongamos que (2) tiene solución, es decir, que existe $y \geq 0$ tal que $A^t y = c$. Si (1) tuviese solución, existiría $x \in \mathbb{R}^n$ tal que $Ax \leq 0$, $c^t x > 0$. Premultiplicando la primera desigualdad por $y \geq 0$, se tiene que $y^t Ax = (A^t y)^t x = c^t x \leq 0$, lo cual es una contradicción. Por lo tanto (1) no tiene solución.

Supongamos ahora que (2) no tiene solución. Sea $S = \{\omega \in \mathbb{R}^n/\omega = A^t y,\ y \geq 0\}$, que es un convexo cerrado, no vacío.

Como (2) no tiene solución, $c \notin S$. Luego existe $p \neq 0$, $\alpha \in \mathbb{R}$ tal que

$$\langle p, c\rangle > \alpha \quad \text{y} \quad \langle p, \varpi\rangle \leq \alpha \ , \forall \varpi \in S.$$

Como $\omega = 0 \in S$, $\alpha \geq 0$. Así $\langle p, c\rangle > 0$. (*)

De $\langle p, \omega\rangle \leq \alpha$, $\forall \omega \in S$, se tiene que $\langle p, A^t y\rangle = \langle Ap, y\rangle \leq \alpha$, $\forall y \geq 0$.

Supongamos que Ap tiene una coordenada estrictamente positiva, digamos $(Ap)_1$, y consideremos $y = \lambda \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\lambda > 0 \Rightarrow \lambda(Ap)_1 \leq \alpha \ \forall \lambda > 0$, lo que es una contradicción, pues se puede elegir λ suficientemente grande de modo de violar la desigualdad, dado que $(Ap)_1 > 0$.

Luego, Ap no tiene coordenadas positivas, es decir, $Ap \leq 0$ (**)

Por (*) y (**), (1) tiene solución para $x = p$. $\blacksquare$

Ejemplo 2.1.14 *Sea $A \in \mathcal{M}_{m\times n}(I\!R)$, $c \in I\!R^n$. Uno y sólo uno de los siguientes sistemas tiene solución:*

(1) $Ax \leq 0$, $x \geq 0$, $c^t x > 0$, algún $x \in I\!R^n$.

(2) $A^t y = c$, $y \geq 0$, algún $y \in I\!R^n$.

Basta considerar la matriz $\widetilde{A} = \begin{bmatrix} A \\ -I \end{bmatrix}$ y aplicar Farkas. $\blacksquare$

Teorema 2.1.12 *Sean S_1 y S_2, conjuntos convexos no vacíos en $I\!R^n$, tales que $S_1 \cap S_2 = \phi$. Existe un hiperplano que separa S_1 y S_2, es decir, existe $p \in I\!R^n$ no nulo tal que*

$$p^t x_1 \geq p^t x_2 \ \forall x_1 \in S_1, x_2 \in S_2. \textit{ (Ver figura 2.14)}$$

Demostración. Consideremos el conjunto

$$S = \{x / x = x_2 - x_1, \ \ x_1 \in S_1, \ \ x_2 \in S_2\} = S_2 - S_1.$$

Se deduce del ejercicio (2.1.1) que S es un convexo. Además, es claro que $0 \notin S$ (en efecto, $0 \in S$ implicaría que $S_1 \cap S_2 \neq \phi$). Luego, usando el teorema (2.1.9), existe $p \neq 0$ tal que

$$\forall x \in S: \quad p^t x \leq 0$$

de donde se concluye que

$$\forall x_1 \in S_1, \ \forall x_2 \in S_2: \quad p^t(x_2 - x_1) \leq 0$$

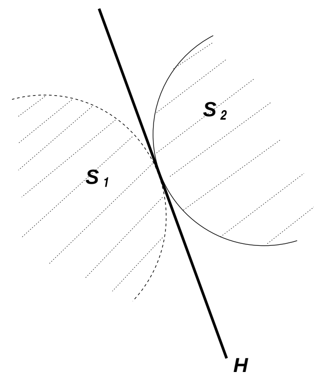

Figura 2.14: En el caso de la figura, el hiperplano separador de los dos conjuntos convexos es soportante para la clausura de ambos.

lo que prueba el teorema. ■

Teorema 2.1.13 (Gordan) *Sea $A \in \mathcal{M}_{m\times n}(I\!R)$. Entonces uno y sólo uno de los siguientes sistemas tiene solución:*

(1) $Ax < 0$, algún $x \in I\!R^n$.

(2) $A^t p = 0$, $p \geq 0$, $p \neq 0$, algún $p \in I\!R^n$.

Demostración. Supongamos que (1) tiene solución, es decir, que $Ax < 0$ para algún $x \in I\!R^n$. Si (2) tuviese solución, existiría $p \in I\!R^n, p \geq 0$, $p \neq 0$ tal que $A^t p = 0$.

Entonces, premultiplicando (1) por p^t se tiene que $p^t A x < 0 \Rightarrow (A^t p)^t x = 0 < 0$, lo que es una contradicción.

Supongamos ahora que (2) no tiene solución. Definamos

$$S_1 = \{z \in I\!R^m / z = Ax, x \in I\!R^n\} \ \ y \ \ S_2 = \{z \in I\!R^m / z < 0\}$$

Si (1) no tuviese solución, entonces

$$\begin{gathered} S_1 \cap S_2 = \phi \\ S_1 \neq \phi, S_2 \neq \phi \\ S_1 \ \ y \ \ S_2 \ \ convexos \end{gathered}$$

Luego, por el teorema anterior, existe un hiperplano separador, es decir

$$\exists p \neq 0 \text{ tal que } p^t z_1 \geq p^t z_2 \quad \forall z_1 \in S_1 \quad z_2 \in S_2$$

Luego, $p^t A x_1 \geq p^t z_2 \quad \forall x_1 \in I\!R^n \quad z_2 \in S_2$.

Como $z_2 < 0$ puede elegirse arbitrariamente negativo. Probaremos que $p \geq 0$.
Si p tuviese alguna coordenada negativa se tendría que $p^t A x_1 \geq \sup\limits_{z_2<0}\{p^t z_2\} = +\infty \Rightarrow p^t A x_1 = +\infty \quad \forall x_1 \in I\!R^n$, lo que es una contradicción. Así, $p \geq 0$.

Luego, tomando límite cuando $z_2 \to 0^-$, se tiene que $p^t A x_1 \geq 0 \quad \forall x_1 \in I\!R^n$. Tomando $x_1 = -A^t p$, se tiene que $\|A^t p\| = 0 \Rightarrow A^t p = 0$ (contradicción con el hecho de que (2) no tiene solución).

Por lo tanto, (1) tiene solución. ■

Ejercicio 2.1.5 *Demuestre, usando el teorema de Farkas, que si para todo $y \geq 0$ tal que $A^t y \geq 0$ se tiene que $b^t y \geq 0$, entonces existe $x \geq 0$ tal que $Ax \leq b$.*
Indicación: Note que el sistema $Ax \leq b$ puede reemplazarse por $Ax + s = b$, con $s \in I\!R^m_+$.

Ejercicio 2.1.6 *Sean $A \in \mathcal{M}_{p\times n}(I\!R)$ y $B \in \mathcal{M}_{q\times n}(I\!R)$. Demuestre que uno y sólo uno de los sistemas siguientes tiene solución:*

(I) $Ax < 0 \quad Bx = 0$

(II) $A^t u + B^t v = 0 \quad u \neq 0, u \geq 0$

2.2 Funciones convexas

Definición 2.2.1 *Sea $f : S \subseteq I\!R^n \longrightarrow I\!R$, con $S = Dom(f)$ convexo.*

$$x \in S \longmapsto f(x)$$

Se dice que f es **convexa** *si y sólo si*

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \quad \forall x, y \in S, \quad 0 \leq \lambda \leq 1.$$

Esta definición se puede interpretar geométricamente, diciendo que la imagen por f del segmento [x,y] queda por debajo de la recta que une (x,f(x)), (y,f(y)). (ver figura 2.15)

Por inducción es posible probar un resultado equivalente para la combinación convexa de k puntos en $I\!R^n$.

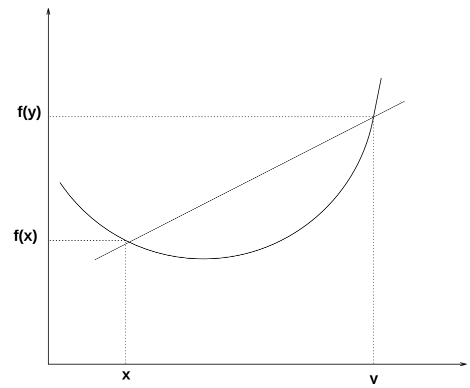

Figura 2.15: La imagen por f del segmento [x,y] queda por debajo de la recta que une (x,f(x)), (y,f(y)).

Teorema 2.2.1 (Desigualdad de Jensen) *Sea $f : S \subseteq I\!\!R^n \longrightarrow I\!\!R$, $S = Dom(f)$ convexo. Entonces, f es convexa si y sólo si $\forall \{x_i\}_{i=1}^k \subseteq S$ y $\lambda_i \geq 0, i = 1, ..., k$ tal que $\sum_{i=1}^{k} \lambda_i = 1$, se tiene*

$$f(\lambda_1 x_1 + ... + \lambda_k x_k) \leq \lambda_1 f(x_1) + ... + \lambda_k f(x_k).$$

Definición 2.2.2 *Una función f, definida como antes, se dice* **estrictamente convexa** *si y sólo si para todo $x \neq y$, $0 < \lambda < 1$, se tiene*

$$f(\lambda x + (1 - \lambda)y) < \lambda f(x) + (1 - \lambda)f(y)$$

Definición 2.2.3 *Sea $f : S \subseteq I\!\!R^n \longrightarrow I\!\!R$, con $S = Dom(f)$ convexo. Se dice que f es* **cóncava** *si y sólo si $-f$ es convexa o, equivalentemente, si*

$$f(\lambda x + (1 - \lambda)y) \geq \lambda f(x) + (1 - \lambda)f(y), \quad \forall x, y \in S, \quad 0 \leq \lambda \leq 1 \ .$$

Del mismo modo, f es **estrictamente cóncava** *si y sólo si $-f$ es estrictamente convexa.*

Ejemplo 2.2.1 .

i) Una función $f : I\!\!R^n \to I\!\!R$ tal que $f(x) = \alpha^t x + \beta$ (lineal afín) es cóncava y convexa.

ii) *La función* f *definida por* $f(x) = \begin{cases} x & x \in (-\infty, 0] \\ 0 & x \in (0,1) \\ x-1 & x \in [1, \infty) \end{cases}$ *(lineal por pedazos) es convexa, pero no estrictamente convexa.*

iii) *La función* $f(x) = -x^2$ *es estrictamente cóncava.*

iv) *La función* $f(x) = \begin{cases} x^2 & x \in (-\infty, 2] \\ 4 & x \in (2, \infty) \end{cases}$ *no es cóncava ni convexa.*

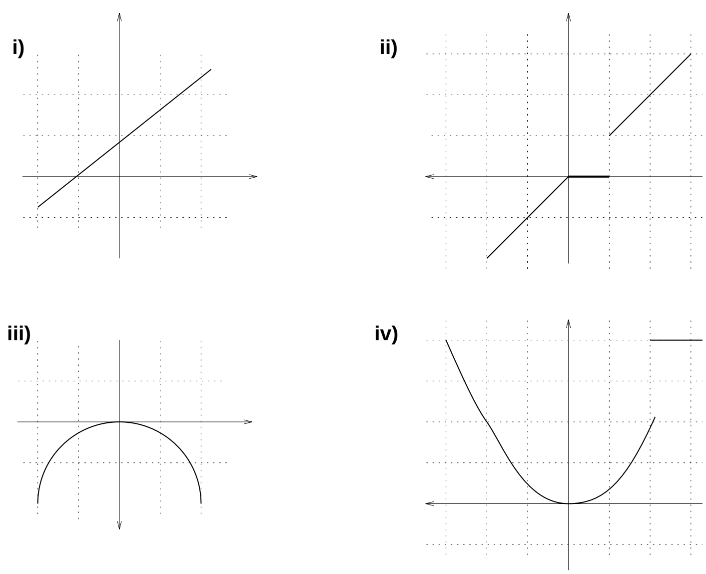

Figura 2.16: Gráfico de las funciones del ejemplo 2.2.1

Definición 2.2.4 *Sea* $f : S \subseteq I\!R^n \longrightarrow I\!R$*, con* $S = Dom(f) \neq \phi$ *un conjunto cualquiera. Para* $\alpha \in I\!R$ *se definen los siguientes conjuntos (ver figura 2.17)*

- $N_\alpha(f) = \{x \in I\!R^n \ / \ f(x) \leq \alpha\}$*, el* **conjunto de nivel** α.
- $C_\alpha(f) = \{x \in I\!R^n \ / \ f(x) = \alpha\}$, **curva de nivel** α.
- $epi(f) = \{(x, \alpha) \in S \times I\!R \ / \ f(x) \leq \alpha\}$*, el* **epígrafo de** f.

Teorema 2.2.2 *Sea una función* $f : S \subseteq I\!R^n \longrightarrow I\!R$*, con* $S = Dom(f)$ *convexo. Se tiene que*

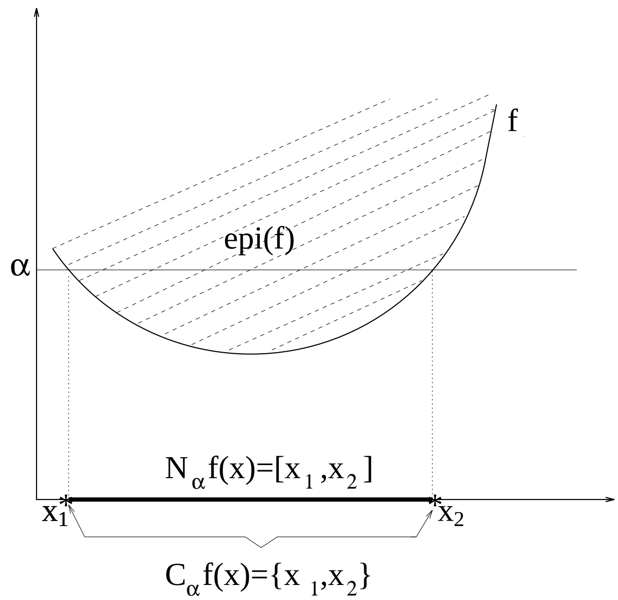

Figura 2.17: Conjunto de nivel, curva de nivel y epígrafo de una función real f.

(i) f es convexa si y sólo si $epi(f)$ es un conjunto convexo.

(ii) si f es convexa, entonces $N_\alpha(f)$ es convexo.

Demostración.

(i) ($\Rightarrow$)Sean $(x,\alpha),(y,\beta)\in epi(f),\quad \lambda\in[0,1]$.

$\lambda(x,\alpha)+(1-\lambda)(y,\beta)=(\lambda x+(1-\lambda)y,\lambda\alpha+(1-\lambda)\beta)\in S\times I\!R$ (convexo)

Como f es convexa, $f(\lambda x+(1-\lambda)y)\leq\lambda f(x)+(1-\lambda)y\leq\lambda\alpha+(1-\lambda)\beta$ (pues $f(x)\leq \alpha, f(y)\leq\beta$)

Luego, $\lambda(x,\alpha)+(1-\lambda)(y,\beta)\in epi(f)$.

($\Leftarrow$) Como $epi(f)$ es convexo, $\forall(x,\alpha),(y,\beta)\in epi(f),\quad \lambda\in[0,1]$, se tiene que $\lambda(x,\alpha)+(1-\lambda)(y,\beta)\in epi(f)$, es decir, $f(\lambda x+(1-\lambda)y)\leq\lambda\alpha+(1-\lambda)\beta$.

Claramente, $(x,f(x)),(y,f(y))\in epi(f)$. Reemplazando α por $f(x)$ y β por $f(y)$ en la expresión anterior, se concluye que f es convexa.

(ii) Directo de (i). (Ver la primera implicancia) $\blacksquare$

Veamos que la implicancia inversa en (ii) no es cierta. Para la función (iv) del ejemplo (2.2.1), $N_\alpha(f)$ es convexo $\forall\alpha\in I\!R$, sin embargo, la función no es convexa.

Teorema 2.2.3 *Sea $f : S \longrightarrow I\!R$ convexa. Entonces, f es continua en $int(S)$.*

Demostración. Sea $\overline{x} \in int(S)$. Para probar la continuidad de f en $\overline{x}$ necesitamos mostrar que, dado $\varepsilon > 0$, existe $\delta > 0$ tal que $\|x - \overline{x}\| \leq \delta \Rightarrow |f(x) - f(\overline{x})| \leq \varepsilon$.

Sea $\varepsilon > 0$.

Dado que $\overline{x} \in int(S)$ existe $\eta > 0$ tal que $B(\overline{x}, \eta) \subseteq S$. Claramente $\overline{x} \pm \eta e_i \in S$, con e_i vector de la base canónica, luego

$$
\begin{aligned}
\overline{x} &= \tfrac{1}{2}(\overline{x} + \eta e_i) + \tfrac{1}{2}(\overline{x} - \eta e_i) \\
\Rightarrow (f \text{ es convexa})\ f(\overline{x}) &\leq \tfrac{1}{2} f(\overline{x} + \eta e_i) + \tfrac{1}{2} f(\overline{x} - \eta e_i)\ \ \forall i = 1, ..., n \\
\Rightarrow 0 &\leq \tfrac{1}{2}\{f(\overline{x} + \eta e_i) - f(\overline{x})\} + \tfrac{1}{2}\{f(\overline{x} - \eta e_i) - f(\overline{x})\}.
\end{aligned}
$$

De aquí se desprende que $\forall i,\ f(\overline{x} + \eta e_i) - f(\overline{x})$ y $f(\overline{x} - \eta e_i) - f(\overline{x})$ no pueden ser simultáneamente negativos.

Sea $K = \max\{f(\overline{x} \pm \eta e_i) - f(\overline{x}),\ \forall i = 1, ..., n\}$, $\ 0 \leq K < \infty$, y definamos $\delta = \min\{\frac{\eta}{n}, \frac{\varepsilon\eta}{nK}\}$.

Sean $\alpha_i \geq 0\ \ i = 1, ..., n$, tales que $x - \overline{x} = \sum\limits_{i=1}^{n} \alpha_i d_i$, con $d_i = \left\{ \begin{array}{r} \eta e_i \ \ si\ \ x_i - \overline{x}_i \geq 0 \\ -\eta e_i \ \ si\ \ x_i - \overline{x}_i < 0 \end{array} \right.$

Luego, $\|x - \overline{x}\|^2 = \left\| \sum\limits_{i=1}^{n} \alpha_i d_i \right\|^2 = \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} \alpha_i \alpha_j d_i d_j = \sum\limits_{i=1}^{n} \alpha_i^2 \|d_i\|^2 = \eta^2 \sum\limits_{i=1}^{n} \alpha_i^2 \leq \delta^2$. Así, $\sum\limits_{i=1}^{n} \alpha_i^2 \leq \frac{\delta^2}{\eta^2} = \min\{\frac{1}{n^2}, \frac{\varepsilon^2}{n^2 K^2}\}$, lo que implica en particular que $\alpha_i \leq \min\{\frac{1}{n}, \frac{\varepsilon}{nK}\}\ \ \forall i.(*)$

Entonces,

$f(x) = f(x - \overline{x} + \overline{x}) = f(\sum\limits_{i=1}^{n} \alpha_i d_i + \overline{x}) = f(\sum\limits_{i=1}^{n} \frac{1}{n}(n\alpha_i d_i + \overline{x})) \leq \sum\limits_{i=1}^{n} \frac{1}{n} f(n\alpha_i d_i + \overline{x})$

$= \frac{1}{n} \sum\limits_{i=1}^{n} f[(1 - n\alpha_i)\overline{x} + n\alpha_i(\overline{x} + d_i)] \leq \frac{1}{n} \sum\limits_{i=1}^{n} [(1 - n\alpha_i) f(\overline{x}) + n\alpha_i f(\overline{x} + d_i)]$

$= f(\overline{x}) + \sum\limits_{i=1}^{n} \alpha_i [f(\overline{x} + d_i) - f(\overline{x})]$

Luego,

$$
f(x) - f(\overline{x}) \leq \sum_{i=1}^{n} \alpha_i [f(\overline{x} + d_i) - f(\overline{x})] \leq K \sum_{i=1}^{n} \alpha_i < K \frac{\varepsilon}{K} = \varepsilon
$$

(de la definición de K y por $(*)$)

Para terminar, falta probar que $f(\overline{x}) - f(x) < \varepsilon$. Sea $y = 2\overline{x} - x$. Notemos que $\|y - \overline{x}\| = \|\overline{x} - x\| \leq \delta$, luego, por lo anterior, $f(y) - f(\overline{x}) \leq \varepsilon$.

Pero $f(\overline{x}) = f(\frac{1}{2}y + \frac{1}{2}x) \leq \frac{1}{2}f(y) + \frac{1}{2}f(x)$

$$\Rightarrow \tfrac{1}{2}[f(\overline{x}) - f(x)] \leq \tfrac{1}{2}[f(y) - f(\overline{x})] \leq \tfrac{1}{2}\varepsilon.$$

Luego, $|f(x) - f(\overline{x})| \leq \varepsilon$ y se tiene el resultado. ■

Una función convexa podría no ser continua en todas partes. Sin embargo, del teorema anterior se puede deducir que los puntos de discontinuidad se encuentran en la frontera del dominio, como muestra el siguiente ejemplo.

Ejemplo 2.2.2 *Sea* $S = \{x / |x| \leq 1\}$ *y* $f : S \to I\!R$, *definida por* $f(x) = \left\{ \begin{array}{ll} x^2 & |x| < 1 \\ 2 & |x| = 1 \end{array} \right.$.

La función f es convexa, continua en $int\,(S)$ y los puntos de discontinuidad son $\{-1, 1\}$ (la fontera de S.)

Definición 2.2.5 *Sea* $S \subseteq I\!R^n$, *no vacío,* $f : S \longrightarrow I\!R$, $\overline{x} \in S$, *y* $d \neq 0$ *tal que*

$$\overline{x} + \lambda d \in S \ \ \forall \lambda \in [0, \eta[, \ \textit{algún} \ \eta > 0.$$

Se define la **derivada direccional** *de* f *en el punto* $\overline{x}$ *y en la dirección* d, *por el siguiente límite (cuando existe)*

$$f'(\overline{x}, d) = \lim_{\lambda \to 0^+} \tfrac{f(\overline{x} + \lambda d) - f(\overline{x})}{\lambda} \in \overline{I\!R}$$

donde $\overline{I\!R} = I\!R \cup \{-\infty, +\infty\}$ *y las operaciones se extienden como sigue:*

$a + \infty = \infty = \infty + a$ *para* $-\infty < a \leq \infty$

$a - \infty = -\infty = -\infty + a$ *para* $-\infty \leq a < \infty$

$a \cdot \infty = \infty = \infty \cdot a, \quad a \cdot (-\infty) = -\infty = (-\infty) \cdot a$ *para* $0 < a \leq \infty$

$a \cdot \infty = -\infty = \infty \cdot a, \quad a \cdot (-\infty) = \infty = (-\infty) \cdot a$ *para* $-\infty < a \leq 0$

$0 \cdot \infty = \infty \cdot 0 = 0 = 0 \cdot (-\infty) = (-\infty) \cdot 0$

$$-(-\infty) = \infty$$

y la suma de ∞ con $-\infty$ no está definida.

Definición 2.2.6 *Sea $S \subseteq I\!R^n$, no vacío. Una función $f : S \to I\!R$ se dice* **diferenciable** *en $\overline{x} \in int\,(S)$ si y sólo si existe $\nabla f(\overline{x}) \in I\!R^n$ tal que*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t\,(x - \overline{x}) + o\,(x - \overline{x}) \quad \forall x \in S$$

donde $o(x - \overline{x})$ es tal que $\lim\limits_{x \to \overline{x}} \frac{o(x-\overline{x})}{\|x-\overline{x}\|^2} = 0$.

Teorema 2.2.4 *Si f es diferenciable en $int\,(S)$, entonces $f'(\overline{x}, d) = \bigtriangledown f(\overline{x})^t d$.*

Demostración. Sea $\overline{x} \in int\,(S)$, como f es diferenciable se tiene que

$$f\,(x) = f\,(\overline{x}) + \bigtriangledown f(\overline{x})^t(x - \overline{x}) + o(x - \overline{x}) \quad \forall x \in S.$$

Sea $x = \overline{x} + \lambda d \in S$ (para $\lambda > 0$ suficientemente pequeño), luego

$$f\,(\overline{x} + \lambda d) = f\,(\overline{x}) + \bigtriangledown f(\overline{x})^t(\lambda d) + o(\lambda d) \Rightarrow \frac{f(\overline{x}+\lambda d)-f(\overline{x})}{\lambda} = \bigtriangledown f(\overline{x})^t d + \frac{o(\lambda d)\|d\|}{\lambda\|d\|} \cdot \frac{\lambda\|d\|}{\lambda\|d\|}$$

Tomando límite cuando $\lambda \to 0^+$, se obtiene

$$f'(\overline{x}, d) = \bigtriangledown f(\overline{x})^t d. \qquad \blacksquare$$

Proposición 2.2.1 *Sea $f : S \longrightarrow I\!R$, convexa. Sea $\overline{x} \in S$, y $d \neq 0$ tal que $\overline{x} + \lambda d \in S$, para todo $\lambda \in [0, \eta[$, para algún $\eta > 0$. Entonces $f'(\overline{x}, d)$ existe.*

Demostración.
Sean $0 < \lambda_1 < \lambda_2 < \overline{\lambda}$

$$f(\overline{x} + \lambda_1 d) = f(\tfrac{\lambda_1}{\lambda_2}(\overline{x} + \lambda_2 d) + (1 - \tfrac{\lambda_1}{\lambda_2})\overline{x}) \leq \tfrac{\lambda_1}{\lambda_2} f(\overline{x} + \lambda_2 d) + (1 - \tfrac{\lambda_1}{\lambda_2}) f(\overline{x})$$
$$\Rightarrow \frac{f(\overline{x}+\lambda_1 d)-f(\overline{x})}{\lambda_1} \leq \frac{f(\overline{x}+\lambda_2 d)-f(\overline{x})}{\lambda_2}$$

Así, $\varphi(\lambda) = \frac{f(\overline{x}+\lambda d)-f(\overline{x})}{\lambda}$ es una función no decreciente de λ. Luego $f'(\overline{x}, d) = \lim\limits_{\lambda \to 0^+} \varphi(\lambda) = \inf\limits_{\lambda > 0} \varphi(\lambda)$ existe. ■

Teorema 2.2.5 *Sea $f : S \longrightarrow I\!R$ una función diferenciable en $S \subseteq I\!R^n$, convexo. Entonces f es convexa si y sólo si $f(x) \geq f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x})$, $\forall x, \overline{x} \in S$.*

Demostración.

($\Rightarrow$) Sea f convexa. Dados $x, \overline{x} \in S$ se tiene que

$$f(\lambda x + (1-\lambda)\overline{x}) \leq \lambda f(x) + (1-\lambda) f(\overline{x}) \quad \forall \lambda \in [0,1]$$

Reordenando, $f(\overline{x} + \lambda(x - \overline{x})) - f(\overline{x}) \leq \lambda f(x) - \lambda f(\overline{x}) \Rightarrow \frac{f(\overline{x}+\lambda(x-\overline{x}))-f(\overline{x})}{\lambda} \leq f(x) - f(\overline{x})$.
Tomando $\lim\limits_{\lambda \to 0^+}$:

$$\begin{gathered} f'(\overline{x}, d) = \bigtriangledown f(\overline{x})^t(x - \overline{x}) \leq f(x) - f(\overline{x}) \\ \Rightarrow f(x) \geq f(\overline{x}) + \bigtriangledown f(\overline{x})^t(x - \overline{x}), \forall \overline{x}, x \in S. \end{gathered}$$

($\Leftarrow$) Sean $x, \overline{x} \in S$,

$$\begin{gathered} f(x) \geq f(\lambda x + (1-\lambda)\overline{x}) + \langle \nabla f(\lambda x + (1-\lambda)\overline{x}), (1-\lambda)(x - \overline{x}) \rangle \quad \forall \lambda \in [0,1] \\ f(\overline{x}) \geq f(\lambda x + (1-\lambda)\overline{x}) + \langle \nabla f(\lambda x + (1-\lambda)\overline{x}), \lambda(\overline{x} - x) \rangle \quad \forall \lambda \in [0,1] \end{gathered}$$

Multilplicando la primera desigualdad por λ, la segunda por $(1-\lambda)$ y sumando, se tiene

$$\lambda f(x) + (1-\lambda) f(\overline{x}) \geq f(\lambda x + (1-\lambda)\overline{x}) \quad \forall \lambda \in [0,1]$$

Es decir, f es convexa. ■

Observación 2.2.1 *Es directo probar que f, satisfaciendo las hipótesis del teorema anterior, es estrictamente convexa si y solo si $f(x) > f(\overline{x}) + \nabla f(\overline{x})^t(x - \overline{x})$, $\quad \forall x, \overline{x} \in S$.*

Corolario 2.2.1 *Sea $f : S \longrightarrow I\!R$ una función diferenciable en $S \subseteq I\!R^n$, convexo. Entonces f es convexa si y sólo si $\langle \nabla f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0, \quad \forall x_1, x_2 \in S$.*

Demostración. De acuerdo al teorema anterior,

$$\begin{array}{ll} f(x_1) \geq f(x_2) + \langle \nabla f(x_2), x_1 - x_2 \rangle & \forall x_1, x_2 \in S \\ f(x_2) \geq f(x_1) + \langle \nabla f(x_1), x_2 - x_1 \rangle & \forall x_1, x_2 \in S \end{array}$$

Sumando las desigualdades anteriores, se tiene que

$$0 \geq \langle \nabla f(x_2), x_1 - x_2 \rangle + \langle \nabla f(x_1), x_2 - x_1 \rangle = \langle -\nabla f(x_2) + \nabla f(x_1), x_2 - x_1 \rangle$$

es decir, $\langle \nabla f(x_2) - \nabla f(x_1), x_2 - x_1 \rangle \geq 0, \quad \forall x_1, x_2 \in S.$ ■

Hasta aquí hemos desarrollado caracterizaciones de convexidad que nos serán de mucha utilidad, pero sólo para funciones diferenciables. Existe una forma sencilla de extender estas caracterizaciones a funciones no diferenciables, mediante el concepto de *subgradiente*, que se define a continuación.

Definición 2.2.7 *Sea $f : S \longrightarrow I\!R$, convexa. Un vector $\xi \in I\!R^n$ se llama* **subgradiente** *de f en $\overline{x}$ si y solo si*

$$f(x) \geq f(\overline{x}) + \xi^t (x - \overline{x}) \quad \forall x \in S$$

El conjunto de subgradientes de f en $\overline{x}$ se denota por $\partial f(\overline{x})$ y se llama **subdiferencial** *de f en $\overline{x}$.*

Proposición 2.2.2 *Si f es convexa y diferenciable en $\overline{x} \in int(S)$, entonces*

$$\partial f(\overline{x}) = \{\nabla f(\overline{x})\}$$

Demostración. Notemos primero que $\nabla f(\overline{x}) \in \partial f(\overline{x})$, pues f es convexa (Teorema 2.2.5).

Sea $\xi \in \partial f(\overline{x})$. Por definición $f(x) \geq f(\overline{x}) + \xi^t (x - \overline{x}) \quad \forall x \in S.$

Sea $x = \overline{x} + \lambda d \in S$ (para $\lambda > 0$ suficientemente pequeño), se tiene que

$$f(\overline{x} + \lambda d) \geq f(\overline{x}) + \lambda \xi^t d \Rightarrow \tfrac{f(\overline{x}+\lambda d) - f(\overline{x})}{\lambda} \geq \xi^t d$$

Tomando límite cuando $\lambda \to 0^+$, $\nabla f(\overline{x})^t d \geq \xi^t d$, luego $(\xi - \nabla f(\overline{x}))^t d \leq 0$. Escogiendo $d = \xi - \nabla f(\overline{x})$ y reemplazando en la ecuación anterior, obtenemos $\|\xi - \nabla f(\overline{x})\|^2 \leq 0$, lo que implica que $\xi = \nabla f(\overline{x})$. ■

Proposición 2.2.3 *Sea* $S \subseteq I\!R^n$ *un convexo no vacío y* $f : S \to I\!R$. *Si* $\forall \overline{x} \in int(S)$ $\exists \xi \in \partial f(\overline{x})$, *entonces* f *es convexa en* $int(S)$.

Demostración. Sean $x_1, x_2 \in int(S)$, $\lambda \in [0,1]$

S es convexo $\Rightarrow int(S)$ es convexo, luego $\overline{x} = \lambda x_1 + (1-\lambda) x_2 \in int(S) \Rightarrow \exists \xi \in \partial f(\overline{x})$, es decir, $f(x) \geq f(\overline{x}) + \xi^t (x - \overline{x}) \quad \forall x \in S$.

En particular,

$$\begin{array}{l} f(x_1) \geq f(\overline{x}) + \xi^t (x_1 - \overline{x}) \\ f(x_2) \geq f(\overline{x}) + \xi^t (x_2 - \overline{x}) \end{array}$$

Pero $x_1 - \overline{x} = (1-\lambda)(x_1 - x_2)$ y $x_2 - \overline{x} = -\lambda (x_1 - x_2)$. Luego, multiplicando la primera desigualdad por λ, la segunda por $(1-\lambda)$ y sumando, se tiene que

$$\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2) \quad \forall \lambda \in [0,1],$$

es decir, f es convexa. $\blacksquare$

Definición 2.2.8 *Sea* $S \subseteq I\!R^n$, *no vacío. Una función* $f : S \to I\!R$ *se dice* **dos veces diferenciable** *en* $\overline{x}$ *si y sólo si existe* $\nabla f(\overline{x}) \in I\!R^n$ *y* $H(\overline{x}) \in I\!R^{n \times n}$ *tal que*

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t (x - \overline{x}) + \tfrac{1}{2} (x - \overline{x})^t H(\overline{x}) (x - \overline{x}) + o(x - \overline{x}) \quad \forall x \in S$$

$H(\overline{x})$ se llama **matriz hessiana** de f en $\overline{x}$,

$$H(\overline{x}) = \begin{bmatrix} \frac{\partial^2 f(\overline{x})}{\partial x_1^2} & \frac{\partial^2 f(\overline{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\overline{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \frac{\partial^2 f(\overline{x})}{\partial x_i \partial x_j} & \vdots \\ \frac{\partial^2 f(\overline{x})}{\partial x_n \partial x_1} & \cdots & \cdots & \frac{\partial^2 f(\overline{x})}{\partial x_n^2} \end{bmatrix}$$

Teorema 2.2.6 *Sea* $S \subseteq I\!R^n$ *un abierto, convexo, no vacío, y sea* $f : S \to I\!R$ *dos veces diferenciable en* S. *Entonces* f *es convexa si y solo si* $H(x)$ *es semi-definida positiva* $\forall x \in S$.

Demostración.

($\Rightarrow$) Sea $\overline{x} \in S$. Queremos probar que $\forall x \in I\!R^n, \quad x^t H(\overline{x})\, x \geq 0$.

Como S es abierto, $\forall x \in I\!R^n, \quad \overline{x} + \lambda x \in S$, para λ suficientemente pequeño. Del teorema (2.2.5) se tiene que

$$f(\overline{x} + \lambda x) \geq f(\overline{x}) + \lambda \nabla f(\overline{x})^t x \quad \forall x \in I\!R^n$$

Además,

$$f(\overline{x} + \lambda x) = f(\overline{x}) + \lambda \nabla f(\overline{x})^t x + \tfrac{\lambda^2}{2} x^t H(\overline{x})\, x + o(\lambda x) \quad \forall x \in I\!R^n$$

Restando las dos ecuaciones, se tiene que

$$\begin{array}{l} 0 \geq -\frac{\lambda^2}{2} x^t H(\overline{x})\, x - o(\lambda x) \quad \forall x \in I\!R^n \\ \Rightarrow x^t H(\overline{x})\, x + \frac{2}{\lambda^2} o(\lambda x) \geq 0 \quad \forall x \in I\!R^n \end{array}$$

Para $x \neq 0$ (el caso $x = 0$ es directo), dividamos por $\|x\|^2$ y tomemos límite cuando $\lambda \rightarrow^+ 0$ para obtener que $x^t H(\overline{x})\, x \geq 0 \quad \forall x \in I\!R^n$, es decir, que $H(\overline{x})$ es semi-definida positiva.

($\Leftarrow$) Sean $x, \overline{x} \in S$. Por teorema del valor medio

$$f(x) = f(\overline{x}) + \nabla f(\overline{x})^t (x - \overline{x}) + \tfrac{1}{2} (x - \overline{x})^t H(\widehat{x}) (x - \overline{x})$$

con $\widehat{x} = \lambda \overline{x} + (1 - \lambda) x \in S$, para algún $\lambda \in (0, 1)$.

Como $H(\widehat{x})$ es semi-definida positiva, $(x - \overline{x})^t H(\widehat{x}) (x - \overline{x}) \geq 0$, luego

$$f(x) \geq f(\overline{x}) + \nabla f(\overline{x})^t (x - \overline{x}) \quad \forall x, \overline{x} \in S$$

Por el teorema (2.2.5), f es convexa.

Ejemplo 2.2.3 *Sea $f(x_1, x_2) = -x_1^2 - 5x_2^2 + 2x_1x_2 + 10x_1 - 10x_2$. Deseamos verificar si f es convexa, cóncava o ninguna de ellas.*

Podemos escribir f de una manera más conveniente como sigue:

$$f(x_1, x_2) = \begin{pmatrix} 10 & -10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{bmatrix} -1 & 1 \\ 1 & -5 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \begin{pmatrix} 10 & -10 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{bmatrix} -2 & 2 \\ 2 & -10 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Luego, $H(x) = \begin{bmatrix} -2 & 2 \\ 2 & -10 \end{bmatrix}$ (constante para todo x).

Calculemos sus valores propios:

$$\det\left(\begin{bmatrix} -2-\lambda & 2 \\ 2 & -10-\lambda \end{bmatrix}\right) = (2+\lambda)\,(10+\lambda) - 4 = \lambda^2 + 12\lambda + 20 = 0$$

$\Rightarrow \lambda_1 = -1.5279 \;\; y \;\; \lambda_2 = -10.4721$

Como ambos valores son negativos, $H(x)$ es definida negativa. Luego, por el teorema anterior, f es cóncava. Más aún, como lo demuestra el siguiente resultado, f es estrictamente cóncava.

Corolario 2.2.2 *Sea $S \subseteq I\!R^n$ un abierto, convexo, no vacío, y sea $f : S \to I\!R$ dos veces diferenciable en S. Se tiene que,*

(i) *si $H(x)$ es definida positiva en cada punto de S, entonces f es estrictamente convexa.*

(ii) *si f es estrictamente convexa, entonces $H(x)$ es semi-definida positiva en todo punto de S.*

Demostración.

(i) Directo de la segunda implicancia del teorema anterior. Basta ver $H(x)$ definida positiva implica que $f(x) > f(\overline{x}) + \nabla f(\overline{x})^t\,(x-\overline{x}) \;\; \forall x, \overline{x} \in S$, lo que por la observación (2.2.1) es equivalente a decir que f es estrictamente convexa.

(ii) Notar que $\lim\limits_{\lambda\to 0} \left\{ \frac{x^t}{\|x\|} H\left(\overline{x}\right) \frac{x}{\|x\|} + \frac{2}{\lambda^2\|x\|^2} o\left(\lambda x\right) \right\} > 0 \Rightarrow x^t H\left(\overline{x}\right) x \geq 0 \;\; \forall x \in I\!R^n$, es decir, $H\left(\overline{x}\right)$ es semi-definida positiva. ■

Ejemplo 2.2.4 .

*(i) Consideremos la función $f(x) = -\ln(x)$.
La matriz hessiana, $H\left(x\right) = f^{(2)}(x) = \frac{1}{x^2} > 0 \;\; \forall x \in S = \{x \in I\!R^n : x > 0\}$, es definida positiva y por el corolario anterior, f es estrictamente convexa.*

(ii) La función $f(x) = x^4$ es estrictamente convexa en todo $I\!R^n$. Sin embargo,

$$H(x) = f^{(2)}(x) = 12x^2 \geq 0 \;\; \forall x \in I\!R^n$$

es semi-definida positiva (notar que $H(0) = 0$).

2.3 Definición del problema de optimización

Sea una función $f : S \to I\!R$ definida sobre un conjunto cerrado S, y consideremos el problema de encontrar $\overline{x} \in S$ tal que $f(\overline{x}) \leq f(x) \quad \forall x \in S$.

Este es un problema de optimización y se escribe de la siguiente manera:

$$(P) \qquad \min_{x \in S} f(x)$$

Definición 2.3.1 *Un elemento $\overline{x} \in S$ se llama* **solución factible** *de (P).*
Si $\overline{x}$ resuelve (P) se dice que es **mínimo, solución óptima** *o* **solución global** *del problema.*
Si existe $\varepsilon > 0$ tal que $f(\overline{x}) \leq f(x) \quad \forall x \in V_\varepsilon(\overline{x})$, donde $V_\varepsilon(\overline{x}) = \{x \in S / \|x - \overline{x}\| \leq \varepsilon\}$, se dice que $\overline{x}$ es **solución local** *o* **mínimo local** *del problema.*

Teorema 2.3.1 *Sea $f : S \to I\!R$, con S convexo no vacío, y sea $\overline{x}$ solución local del problema (P). Entonces,*

(i) *si f es convexa, $\overline{x}$ es mínimo global.*

(ii) *si f es estrictamente convexa, $\overline{x}$ es el único mínimo global.*

Demostración.

(i) Sea $\varepsilon > 0$, $f(x) \geq f(\overline{x}) \quad \forall x \in V_\varepsilon(\overline{x})$.

Supongamos que $\overline{x}$ no es óptimo global, es decir, $\exists y \in S$ tal que $f(y) < f(\overline{x})$. Luego, $f(\lambda y + (1-\lambda)\overline{x}) \leq \lambda f(y) + (1-\lambda) f(\overline{x}) < \lambda f(\overline{x}) + (1-\lambda) f(\overline{x}) = f(\overline{x})$

Pero para $\overline{\lambda}$ suficientemente pequeño, $\overline{\lambda} y + (1-\overline{\lambda})\overline{x} \in V_\varepsilon(\overline{x})$, lo cual es una contradicción pues $\overline{x}$ es mínimo local.

(ii) f estrictamente convexa $\Rightarrow f$ convexa. Luego, por (i), $\overline{x}$ es mínimo global.

Supongamos que no es único, esto es, que existe $y \in S \;\; (y \neq \overline{x})$, tal que $f(y) = f(\overline{x})$.

$f(\frac{1}{2}y + \frac{1}{2}\overline{x}) < \frac{1}{2}f(y) + \frac{1}{2}f(\overline{x}) = f(\overline{x}) \Rightarrow \exists z = \frac{1}{2}y + \frac{1}{2}\overline{x} \neq \overline{x}$ tal que $f(z) < f(\overline{x})$, lo que contradice el hecho de $\overline{x}$ es mínimo global. ■

Cuando la función f es lineal, es decir, $f(x)$ es de la forma $c^t x$, y el conjunto de restricciones S es un poliedro cerrado, $S = \{Ax = b, x \geq 0\}$, el problema (P) se conoce como *problema de programación lineal*, y su estudio es el objetivo del próximo capítulo.

Ejercicio 2.3.1 *Considere las funciones $f_i : I\!R^n_+ \Rightarrow I\!R,\ i = 1, ..., k,$ convexas y diferenciables y sea $f(x) = \max\{f_1(x), ..., f_k(x)\}$.*

Considere además el problema

$$(P) \min \quad \begin{array}{l} f(x) \\ x \geq 0 \end{array}$$

Demuestre que todo mínimo local de (P) es un mínimo global.

2.4 Ejercicios Resueltos

Ejercicio 2.4.1 *Una función f se dice "homogénea de primer grado" si satisface la siguiente igualdad:*

$$f(\lambda x) = \lambda f(x) \quad \forall x \in I\!R^n, \quad \lambda \geq 0.$$

Además, una función homogenea se dice subaditiva si satisface la siguiente desigualdad:

$$f(x + y) \leq f(x) + f(y) \quad \forall x, y \in I\!R^n.$$

Pruebe que una una función homogenea es convexa si y sólo si es subaditiva.

Ejercicio 2.4.2 *Sean $f_i : I\!R^n \to I\!R$ m funciones convexas $i = 1, ..., m$. Sean además λ_i $i = 1, ..., m$ escalares no negativos. Pruebe que:*

$$g(x) = \sum_{i=1}^{k} \lambda_i f_i(x),$$

es convexa.

Solución. Sean $x, y \in I\!R^n$. Sea además $\lambda \in (0,1)$. De la definición de g se tiene que

$$g(\lambda x + (1-\lambda)y) = \sum_{i=1}^{m} \lambda_i f_i(\lambda x + (1-\lambda)y)$$

como cada f_i es convexa, obtenemos además que

$$f_i(\lambda x + (1-\lambda)y) \leq \lambda f_i(x) + (1-\lambda) f_i(y) \quad \forall i$$

por lo tanto:

$$\begin{aligned} g(\lambda x + (1-\lambda)y) &= \sum_{i=1}^{m} \lambda_i f_i(\lambda x + (1-\lambda)y) \leq \sum_{i=1}^{m} \lambda_i [\lambda f_i(x) + (1-\lambda) f_i(y)] \\ &= \lambda \sum_{i=1}^{m} \lambda_i f_i(x) + (1-\lambda) \sum_{i=1}^{m} \lambda_i f_i(y) = \lambda g(x) + (1-\lambda) g(y) \end{aligned}$$

De la definición, g es convexa. ■

Ejercicio 2.4.3 *Considere $f_i : I\!R^n \to (-\infty, \infty]$ con $i \in I$. I un conjunto arbitrario. Considere además $g : I\!R^n \to (-\infty, \infty]$, dada por:*

$$g(x) = \sup_{i \in I} f_i(x)$$

Si f_i es convexa $\forall i \in I$. Muestre que g es convexa.

Solución Usaremos el hecho de que g es convexa ssi $epi(g)$ es convexo. Así, un par $(x, \alpha) \in epi(g)$ ssi $g(x) \leq \alpha$. Lo que es equivalente a $f_i(x) \leq \alpha \;\; \forall i \in I$. Por lo tanto,

$$(x, \alpha) \in \bigcap_{i \in I} epi(f_i)$$

i.e, $epi(g) = \bigcap\limits_{i \in I} epi(f_i)$. Así, como f_i es convexa $\forall i \in I$, y además la intersección arbitraria de convexos es convexa (Observación 2.1.1), entonces $epi(g)$ es convexo $\Leftrightarrow$ g es convexa ■

Ejercicio 2.4.4 *Muestre que la función $f : I\!R^n \to I\!R$ es lineal afín ssi es cóncava y convexa.*

Solución. Una función lineal afín es aquella que se escribe como:

$$f(x) = c^t x + \alpha.$$

($\Rightarrow$) Evidentemente f es cóncava y convexa.

($\Leftarrow$) Sea f cóncava y convexa, *i.e*

$$f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y)$$

Sea $L(x) = f(x) - f(0)$. Demostremos que L es lineal. Sea $0 \leq \lambda \leq 1$
$f(\lambda x) = f(\lambda x + (1-\lambda)0) = \lambda f(x) + (1-\lambda)f(0)$
De donde $L(\lambda x) = f(\lambda x) - f(0) = \lambda f(x) + (1-\lambda)f(0) - f(0) = \lambda f(x) - \lambda f(0) = \lambda L(x)$
Por otro lado,
$L(x+y) = f(x+y) - f(0) = f(\frac{1}{2}2x + \frac{1}{2}2y) - f(0) = \frac{1}{2}f(2x) + \frac{1}{2}f(2y) - f(0)$
$= f(\frac{1}{2}2x) - \frac{1}{2}f(0) + f(\frac{1}{2}2y) - \frac{1}{2}f(0) - f(0) = L(x) + L(y)$
Ahora, $L(0) = f(0) - f(0) = 0$. Con lo cual, $0 = L(x-x) = L(x) + L(-x)$. Asi obtenemos,

$$L(-x) = -L(x)$$

Como todo $\alpha \in I\!R$ se puede escribir como $sgn(\alpha)(n+\lambda)$,con $n \in I\!N$ y $\lambda \in (0,1)$, deducimos que L es lineal, y $L + f(0)$ es lineal afín, con lo cual f es lineal afín. ■

Ejercicio 2.4.5 *Sea S un conjunto convexo cerrado no vacío en $I\!R^n$ y sea $f : I\!R^n \rightarrow I\!R$ definida por:*

$$f(y) = \inf_{x \in S} \|x - y\|$$

Demuestre que f es convexa.

Solución. Supongamos que no es convexa, *i.e* $\exists z_0, x_0$y λ_0 tales que:

$$\inf_{x \in S} \|\lambda_0 z_0 + (1-\lambda_0)y_0 - x\| > \lambda_0 \inf_{x \in S} \|x - z_0\| + (1-\lambda_0) \inf_{x \in S} \|x - y_0\|$$

Sea $x_0 \in S$ tal que $\|x_0 - z_0\| = \inf\limits_{x \in S} \|x - z_0\|$

De esta menera:

$$\begin{gathered} \lambda_0 f(z_0) + (1-\lambda_0) f(y_0) = \inf_{x \in S} \|\lambda_0 (x_0 - z_0)\| + \|(1-\lambda_0)(x - y_0)\| \\ \geq \inf_{x \in S} \|\lambda_0 (z_0 - x_0) + (1-\lambda_0)(y_0 - x)\| \geq \inf_{x \in S} \|\lambda_0 z_0 + (1-\lambda_o) y_0 - [\lambda_0 x_0 + (1-\lambda_0) x]\| \end{gathered}$$

De lo cual obtenemos que $\inf\limits_{x \in S} \|\lambda_0 z_0 + (1-\lambda_0) y_0 - x\| > \inf\limits_{x \in S} \|\lambda_0 z_0 + (1-\lambda_0) y_0 - [\lambda_0 x_0 + (1-\lambda_0 x]\|$

Para $\lambda_0 \neq 1$ tenemos que $\lambda_0 x_0 + (1-\lambda_0) x \in S$, por ser convexo. Así definamos $L = \{l \in S / l = \lambda_0 x_0 + (1-\lambda_0) x\}$, es claro que $L \subseteq S$. Con esto podemos reescribir la última desigualdad como:

$$\inf_{x \in S} \|\lambda_0 z_0 + (1-\lambda_0) y_0 - x\| > \inf_{l \in L} \|\lambda_0 z_0 + (1-\lambda_0) y_0 - l\|$$

o de otra forma,

$$\inf_{x \in S} \|\eta - x\| > \inf_{x \in L} \|\eta - x\|$$

Lo cual es una contradicción ■

Ejercicio 2.4.6 *Sea $A \in \mathcal{M}_{m \times n}(I\!R)$, $B \in \mathcal{M}_{l \times n}(I\!R)$, $c \in I\!R^n$. Uno y sólo uno de los siguientes sistemas tiene solución:*

(1) $Ax \leq 0$, $Bx = 0$, $c^t x > 0$, algún $x \in I\!R^n$.

(2) $A^t y + B^t z = c$, $y \geq 0$, algún $y \in I\!R^n$, $z \in I\!R^n$.

Solución Consideremos la matriz $\widetilde{A} = \begin{bmatrix} A \\ B \\ -B \end{bmatrix}$. Notemos que (1) es equivalente a $\widetilde{A} x \leq 0$, $c^t x > 0$, algún $x \in I\!R^n$.

Sea $u = \begin{pmatrix} y \\ z_1 \\ z_2 \end{pmatrix}$, con $z_1, z_2 \geq 0$ tales que $z = z_1 + z_2$. De este modo, $u \geq 0$.

$\widetilde{A}^t u = \begin{bmatrix} A^t & B^t & -B^t \end{bmatrix} \begin{pmatrix} y \\ z_1 \\ z_2 \end{pmatrix} = A^t y + B^t z_1 - B^t z_2 = A^t y + B^t z = c$, define un sistema equivalente a (2).

Se concluye usando Farkas. $\blacksquare$

Ejercicio 2.4.7 *Dados $x, b, c \in \mathbb{R}^n$, con $x \geq 0$ y $\mathcal{M}_{m\times n}(\mathbb{R})$ se definen:*

$$Z(b) = \max_{Ax\leq b} c^t x$$
$$V(c) = \max_{Ax\leq b} c^t x$$

Demuestre que Z es convexa y V es cóncava; asumiendo que b y c están en dominios convexos en que los dos problemas son factibles y acotados. Comente que sucede si levantamos estas suposiciones.

Solución Sea $\lambda \in [0,1]$. Debemos probar que $\forall b_1, b_2$ lados derechos, se tiene que:

$$Z(\lambda b_1 + (1-\lambda)b_2) \leq \lambda Z(b_1) + (1-\lambda)Z(b_2)$$

y que $\forall c_1, c_2$ vectores de costo, se tiene que:

$$V(\lambda c_1 + (1-\lambda)c_2) \geq \lambda V(c_1) + (1-\lambda)V(c_2).$$

- Probemos que $Z(b)$ es convexa:

Sean $Z(b_1) = c^t x_1$ con $x_1 \geq 0 \quad Ax_1 = b_1$

$Z(b_2) = c^t x_2$ con $x_2 \geq 0 \quad Ax_2 = b_2$

(x_1, x_2 son soluciones al problema (P) con lado derecho b_1, b_2, respectivamente)

Tomando $x = \lambda x_1 + (1-\lambda)x_2$, se tiene que $Ax = A(\lambda x_1 + (1-\lambda)x_2) = \lambda b_1 + (1-\lambda)b_2$. (1)
Claramente $x \geq 0$. (2)

De (1) y (2) se tiene que x es solución del problema (P) con lado derecho $\lambda b_1 + (1-\lambda)b_2$. El valor óptimo de este problema es $Z(\lambda b_1 + (1-\lambda)b_2) \leq c^t(\lambda x_1 + (1-\lambda)x_2) = \lambda Z(b_1) + (1-\lambda)Z(b_2)$, luego $Z(b)$ es convexa.

- Probemos ahora que $V(c)$ es cóncava:

Sea $\overline{x}$ solución óptima del problema

$$(P_{\overline{c}}) \quad \max \quad \overline{c}^t x$$
$$Ax = b$$
$$x \geq 0$$

con $\overline{c} = \lambda c_1 + (1-\lambda)c_2$.
Claramente $\overline{x}$ es factible para (Pc_1) y (Pc_2), luego $V(c_1) \geq c_1^t\overline{x}$ y $V(c_2) \geq c_2^t\overline{x}$.
Así,

$$\lambda V(c_1) + (1-\lambda)V(c_2) \geq \lambda c_1^t\overline{x} + (1-\lambda)c_2^t\overline{x} = \overline{c}^t\overline{x} = V(\lambda c_1 + (1-\lambda)c_2)$$

Luego, $V(c)$ es cóncava. $\blacksquare$

Capítulo 3

Programación Lineal

3.1 Introducción a la programación lineal

Un problema de programación lineal se escribe de manera explícita como:

$$
\begin{array}{lrrrrrrcl}
(PL) & \min z = & c_1x_1+ & c_2x_2+ & \ldots & + & c_{n-1}x_{n-1}+ & c_nx_n & & \\
 & & a_{1,1}x_1+ & a_{1,2}x_2+ & \ldots & + & a_{1,n-1}x_{n-1}+ & a_{1,n}x_n & = & b_1 \\
 & & a_{2,1}x_1+ & a_{2,2}x_2+ & \ldots & + & a_{2,n-1}x_{n-1}+ & a_{2,n}x_n & = & b_2 \\
 & & \vdots & \vdots & & & \vdots & & & \vdots \\
 & & a_{m,1}x_1+ & a_{m,2}x_2+ & \ldots & + & a_{m,n-1}x_{n-1}+ & a_{m,n}x_n & = & b_m \\
 & & & & & & & x_i & \geq & 0 \quad \forall i
\end{array}
$$

O en forma compacta como:

$$
\begin{array}{rrl}
(PL) & \min & z = c^tx \\
 & sa. & Ax = b \\
 & & x \geq 0
\end{array}
$$

con $x, c \in I\!R^n$, $b \in I\!R^m$, $A \in \mathcal{M}_{m\times n}(I\!R)$, con $m \leq n$.

En la **función objetivo** o **criterio** c^tx, la variable x se conoce como **variable de decisión** o **nivel de actividad** y c como **vector de costos**.

El conjunto de restricciones $S = \{Ax = b, x \geq 0\}$ es un poliedro cerrado y se llama **conjunto factible**. La matriz A se conoce como la **matriz de coeficientes tecnológicos** y b como **vector de recursos** o,simplemente, **lado derecho**.

Otras definiciones preliminares:

- Si S es acotado, existe solución, pues se minimiza una función lineal continua sobre un conjunto compacto.
- Si S es no acotado, puede ocurrir que $c^t x \to -\infty$, con $x \in S$.
- Se dirá que el problema (PL) es **acotado** si y solamente si $\exists \overline{x} \in S \quad c^t\overline{x} \leq c^t x \ \ \forall x \in S$.
- Se dirá que el problema (PL) es **no acotado** si y solamente si $\exists d \in I\!R^n, x_o \in S$ tal que $c^t(x_o + \lambda d) \to -\infty$ si $\lambda \to \infty$ con $x_o + \lambda d \in S \ \ \forall \lambda > 0$ Es evidente que si (PL) es no acotado, entonces S es no acotado. La otra implicancia no siempre es cierta, como veremos cuando estudiemos cómo resolver un programa lineal.
- Se dirá que el problema (PL) es infactible si y sólo si $S = \phi$.

En lo que sigue, revisaremos algunos ejemplos clásicos de la programación lineal.

Ejemplo 3.1.1 *Problema de transporte*

Consideremos una industria que tiene dos fábricas, una en la ciudad O1 y otra en la ciudad O2. Ella produce un bien que es consumido en D1, D2 y D3. La oferta en O1 es de 10.000 unidades diarias, mientras que la oferta en O2 es de 5.000 unidades. La demanda diaria en todos los centros de consumo es de 5.000 unidades.

Los costos de transporte (peajes, impuestos, pago de camiones) están dados por el vector c, donde c_{ij} es el costo de transportar una unidad de producción desde la fábrica i a la ciudad j, con $c_{11} = 2, c_{12} = 1, c_{13} = 3, c_{21} = 1, c_{22} = 2, c_{23} = 1$.

El objetivo del industrial es determinar la cantidad de producto que debe enviar desde cada fábrica a cada centro de demanda, minimizando los costos de transporte.

Si llamamos x_{ij} a la cantidad de producto enviada desde la fábrica i a la ciudad j, el problema y sus restricciones pueden plantearse como sigue:

$$\begin{array}{lllllllllll}
\text{(P)} & \min & 2x_{11} & +x_{12} & +3x_{13} & +x_{21} & +2x_{22} & +x_{23} & & \\
 & & & & & & & & & \\
\text{oferta} & \text{O1} & x_{11} & +x_{12} & +x_{13} & & & & = & 10.000 \\
 & \text{O2} & & & & x_{21} & +x_{22} & +x_{23} & = & 5.000 \\
\text{demanda} & \text{D1} & x_{11} & & & +x_{21} & & & = & 5.000 \\
 & \text{D2} & & x_{12} & & & +x_{22} & & = & 5.000 \\
 & \text{D3} & & & x_{13} & & & +x_{23} & = & 5.000 \\
 & & x_{ij} & \geq 0 & \forall i,j & & & & &
\end{array}$$

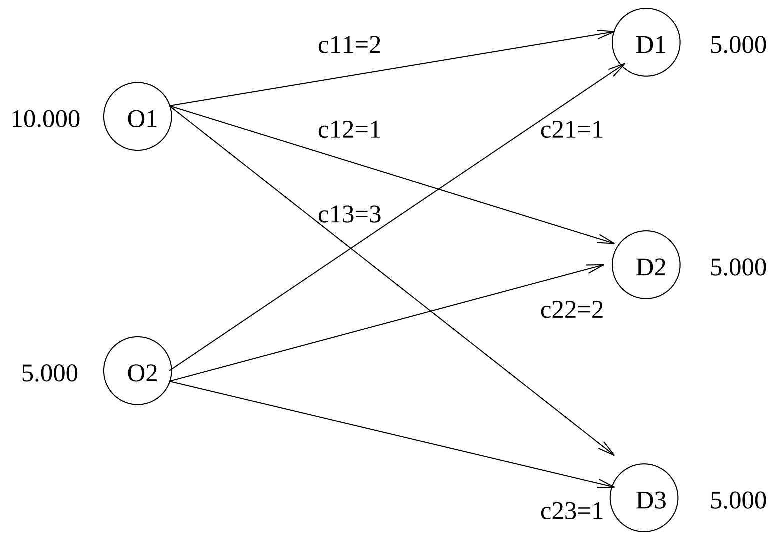

Figura 3.1: Problema de transporte: grafo de la red oferta-demanda

O equivalentemente,

$$(\mathrm{P}) \ \min_x \sum_{i=1}^{N} \sum_{j=1}^{M} c_{ij} x_{ij}$$

$$\text{sa} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \\ x_{21} \\ x_{22} \\ x_{23} \end{bmatrix} = \begin{bmatrix} 10.000 \\ 5.000 \\ 5.000 \\ 5.000 \\ 5.000 \end{bmatrix}$$

$$x_{ij} \geq 0 \ \forall i, j$$

es decir, como un problema lineal, donde S (el poliedro) está escrito en la forma canónica $Ax = B \ , x \geq 0$.

En términos más generales, si a_i denota la oferta del nodo $i = 1, ..., N$ y b_j la demanda en el nodo $j = 1, ..., M$, el problema de transporte puede escribirse como sigue:

$$
\begin{array}{ll}
\min & \sum_{i=1}^{N}\sum_{j=1}^{M} c_{ij}x_{ij} \\
\\
\text{oferta} & \sum_{j=1}^{M} x_{ij} = a_i \quad \forall i = 1,...,N \\
\text{demanda} & \sum_{i=1}^{N} x_{ij} = b_j \quad \forall j = 1,...,M \\
 & x_{ij} \geq 0 \qquad \forall i,j
\end{array}
$$

En general, supondremos que $\sum_{i=1}^{N} a_i = \sum_{j=1}^{M} b_j$, esto garantiza la factibilidad del problema.

Ejemplo 3.1.2 *Planificación de la producción*

Se necesita planificar la producción para k meses, con demandas conocidas al fin de cada mes $d_1, ..., d_k$.

El stock inicial es de $s_o \leq d_1$. Los costos de producción en cada período son $c_1, ..., c_k$. Se puede guardar producto de un mes a otro, con costos unitarios $q_1, ..., q_k$.

Al final del horizonte de producción el stock debe ser nulo.

Se desea encontrar el plan de producción que minimize los costos totales de la empresa.

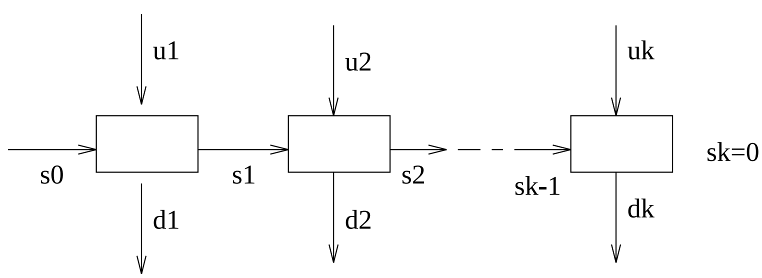

Figura 3.2: Problema de planificación de producción

Escribamos primero las restricciones del problema:

Llamando u_i a la producción del período $i = 1, ...k$, del diagrama de la figura (3.1) es claro que

$$
\begin{array}{lcl}
variables & & datos \\
u_1 - s_1 & = & d_1 - s_o \\
s_1 + u_2 - s_2 & = & d_2 \\
\quad\vdots & & \quad\vdots \\
s_{k-1} + u_k & = & d_k \\
u_i \geq 0, s_i \geq 0 \quad \forall i = 1, ..., k & &
\end{array}
$$

Luego, el problema puede escribirse como:

$$
\min \sum_{i=1}^{k} c_i u_i + \sum_{i=1}^{k-1} q_i s_i
$$

$$
\begin{bmatrix}
1 & 0 & \dots & \dots & 0 & \text{-1} & 0 & \dots & \dots & 0 \\
0 & 1 & 0 & \dots & 0 & 1 & \text{-1} & & & \vdots \\
\vdots & & 1 & 0 & & 0 & 1 & \text{-1} & & \vdots \\
\vdots & & & \ddots & & & & \ddots & \ddots & 0 \\
\vdots & & & & \ddots & & & & 1 & \text{-1} \\
0 & \dots & \dots & \dots & 0 & 1 & 0 & \dots & 0 & 1
\end{bmatrix}
\begin{bmatrix}
u_1 \\ \vdots \\ u_k \\ s_1 \\ \vdots \\ s_{k-1}
\end{bmatrix}
=
\begin{bmatrix}
d_1 - s_o \\ d_2 \\ \vdots \\ \vdots \\ d_k
\end{bmatrix}
$$

$$
u_i \geq 0, s_i \geq 0
$$

Ejemplo 3.1.3 *Problema de la dieta*

Un hacendado está criando cerdos y necesita definir la cantidad de alimentos que hay que dar a cada uno diariamente, para satisfacer los requerimientos nutricionales mínimos, de modo de minimizar el costo por alimentación.

El Servicio Nacional de Alimentación de cerdos entrega a los empresarios del rubro la siguiente carta de nutrientes por Kg. de los alimentos más comunes:

Alimentos [Kg]	**Nutrientes**		
	carbohidratos	**proteínas**	**vitaminas**
maíz	90	30	10
cebada	20	80	20
alfalfa	40	60	60
Req. mínimo diario	200	180	150

El precio del maíz, la cebada y la alfalfa es de \$42, \$36 y \$30 por Kg., respectivamente.

Planteemos el problema de optimización asociado: llamemos $m \geq 0$ a la cantidad de maíz, $c \geq 0$ a la de cebada y $a \geq 0$ a la de alfalfa, todas medidas en Kg.

El objetivo del hacendado es minimizar sus costos por alimentación, esto es

$$\min 42m + 36c + 30a$$

Las restricciones nutricionales, de acuerdo a la tabla, están dadas por

$$\begin{array}{ll} 90m + 20c + 40a \geq 200 & \text{carbohidratos} \\ 30m + 80c + 60a \geq 180 & \text{proteínas} \\ 10m + 20c + 60a \geq 150 & \text{vitaminas} \end{array}$$

Si llamamos $x = \begin{pmatrix} m \\ c \\ a \end{pmatrix}$ a la variable de decisión, $c = \begin{pmatrix} 42 \\ 36 \\ 30 \end{pmatrix}$ al vector de costos,

$A = \begin{pmatrix} 90 & 20 & 40 \\ 30 & 80 & 60 \\ 10 & 20 & 60 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ a la matriz de coeficientes tecnológicos y

$b = \begin{pmatrix} 200 \\ 180 \\ 150 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ el vector de recursos,

el problema de la dieta puede escribirse como un problema de programación lineal:

$$\begin{array}{ll} \min & c^t x \\ & x \in S \end{array}$$

donde $S = \{x / Ax \geq b\}$[1] es un poliedro cerrado.

[1]Recordemos del capítulo anterior que $\{Ax = b, x \geq 0\}$ y $Ax \geq b$ son formas equivalentes de describir un poliedro cerrado.

El problema de la dieta es un problema típico en programación lineal. Una variación interesante es la siguiente:

Supongamos que el precio del maíz es estable ($42 / Kg) y que el precio de la cebada y la alfalfa toman, cualquiera de los siguientes valores:

Alimentos	**Precio**	**por Kg.**
cebada	36	31
alfalfa	31	32

¿Cuál será ahora la función objetivo?

Al hacendado le interesará minimizar sus costos para la peor condición de precios (pues esto implica que si los precios son favorables los costos serán bajos también), esto es

$$\min \max\{42m + 36c + 31a, 42m + 31c + 32a\}$$

O de otra forma,

$$\begin{array}{c} \min \varphi(m, c, a) \\ (m, c, a)^t \in S \end{array}$$

Con $\varphi(m, c, a) = \max\{42m + 36c + 31a, 42m + 31c + 32a\}$, que es una función no lineal de $(m, c, a)^t$.

Un problema equivalente al problema de la dieta es:

$$\begin{array}{lrcl} (P') & & \min \lambda & \\ & 42m + 36c + 31a & \leq & \lambda \\ & 42m + 31c + 32a & \leq & \lambda \\ & x & \in & S \end{array}$$

Observación 3.1.1 *Cuando se crean variables artificiales para transformar un problema, sólo se consideran las variables primitivas en el concepto de problema equivalente. En el caso de nuestro ejemplo, la variable artificial λ no es una variable de (P).*

Proposición 3.1.1 .

$$\begin{array}{llll} \textit{(P1)} & \min\{\max\{f_i(x)\}_{i=1,\ldots,n}\} & \textit{es equivalente a} \quad \textit{(P2)} & \min \lambda \\ & \textit{sa. } x \in S & & \textit{sa. } f_i(x) \leq \lambda \quad \forall i = 1, ..., n \\ & & & (x, \lambda) \in S \times I\!R^n \end{array}$$

En el sentido de que $(\overline{x})$ es solución de $(P1) \Leftrightarrow (\overline{x}, \overline{\lambda})$ es solución de $(P2)$.

Formas canónicas de un programa lineal

Un general un PL se encuentra en una de las dos siguientes formas:

1) Forma de desigualdad:

$$\begin{array}{cc} \min\limits_{x} & z = c^t x \\ \text{sa} & Ax \geq b \\ & x \geq 0 \end{array}$$

2) Forma estándar:

$$\begin{array}{cc} \min\limits_{x} & z = c^t x \\ \text{sa} & Ax = b \\ & x \geq 0 \end{array}$$

Estas dos formas son equivalentes pues, como vimos en el capítulo anterior, los poliedros $\{x/Ax = b, x \geq 0\}$ y $\{x/Ax \leq b\}$ son equivalentes, en el sentido que se puede pasar de una forma a otra.

En general, cualquier problema de programación lineal pude escribirse en una de estas dos formas, usando transformaciones sencillas, como por ejemplo:

- Pasar de $\geq$ a $\leq$ multiplicando la restricción por -1.
- Pasar de $\leq$ a $=$ usando una variable de holgura:

  $$\sum_{j=1}^{n} a_{ij}x_j \leq b_j \text{ puede escribirse de la forma } \sum_{j=1}^{n} a_{ij}x_j + x_{n+1} = b_j,$$

  con $x_{n+1} \geq 0$.[2]
- Una variable irrestricta x puede ser reemplazada por dos variables no negativas $x_1 \geq 0$ y $x_2 \geq 0$, escribiendo $x = x_1 - x_2$
- Maximizar $c^t x$ es eqivalente a minimizar $-c^t x$

[2]Notar que en este caso el problema aumenta en una variable por cada restricción.

Ejemplo 3.1.4 *El problema*

$$\begin{array}{rrrcr} \max & c_1x_1+ & c_2x_2 & & \\ & x_1- & 3x_2 & \geq & 8 \\ & x_1 & & \geq & 0 \\ & & x_2 & \in & I\!R \end{array}$$

es equivalente al problema

$$\begin{array}{rrrrcr} -\min & -c_1x_1- & c_2x_3+ & c_2x_4 & & \\ & -x_1+ & 3x_3- & 3x_4 & \leq & -8 \\ & x_1 & & & \geq & 0 \\ & & x_3 & & \geq & 0 \\ & & & x_4 & \geq & 0 \end{array}$$

que a su vez, es equivalente al problema:

$$\begin{array}{rrrrrcr} -\min & -c_1x_1 & -c_2x_3 & +c_2x_4 & & & \\ & -x_1 & +3x_3 & -3x_4 & +x_5 & = & -8 \\ & x_1 & & & & \geq & 0 \\ & & x_3 & & & \geq & 0 \\ & & & x_4 & & \geq & 0 \\ & & & & x_5 & \geq & 0 \end{array}$$

Plantear un problema de optimización requiere comprender la estructura del mismo y ser ordenado y creativo a la hora de darle forma. Veamos algunos ejemplos más de planteamiento de programas lineales.

Ejemplo 3.1.5 *Una constructora de viviendas de madera acaba de ganarse una propuesta para edificar un conjunto de casas. Los ingenieros de la constructora están preocupados de minimizar los costos tanto como sea posible. Ellos han estimado que requerirán madera aserrada de* 4×4 *de diferentes longitudes: de 80, 60 y 50 cm. En el mercado existe este producto en dos dimensiones: 1.20 y 2.10 m. con un costo de 300 y 350 pesos, respectivamente. La cantidad de piezas de cada largo, que se emplearán en la construcción, se indican en la siguiente tabla:*

Longitud (cm)	*Cantidad mínima requerida*
80	*1500*
60	*800*
50	*2400*

Para satisfacer sus necesidades, la empresa tiene que trozar los productos que compre para obtener las piezas deseadas. ¿Cuántos produactos de 1.20 m. y de 2.10 m.debe comprar la empresa para satisfacer sus requerimientos y minimizar los costos? Formule un modelo de programación lineal para resolver el problema.

Solución:

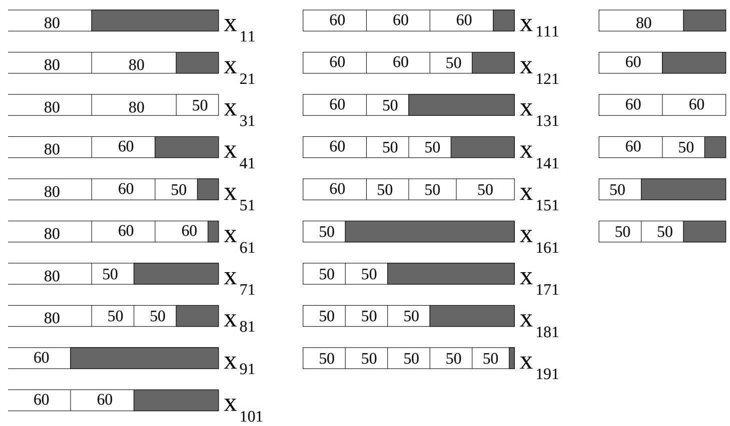

Figura 3.3: Posibles configuraciones de corte

El mercado ofrece madera aserrada en dos dimensiones

j	*dimensión*
1	1.20
2	2.10

con costos $c_1 = 300$ y $c_2 = 350$, y la empresa requiere 3 longitudes distintas: 80, 60 y 30 cm, en las cantidades señaladas en la tabla del enunciado.

Una placa 2.10 m. puede cortarse en 19 formas distintas (configuraciones), para satisfacer los requerimientos de la empresa. Por su parte, una placa de 1.20 m. puede cortarse en 6 configuraciones, tal como se indica en la figura (3.3).

Luego, el conjunto de variables de decisión natural será

x_{ij} = cantidad de madera de dimensión j, cortada de acuerdo a la configuaración i.

Con esto, la función objetivo se escribirá:

$$\min 300 \sum_{i=1}^{19} x_{i1} + 350 \sum_{i=1}^{6} x_{i2}$$

Las restricciones serán:

$$\begin{array}{rcl}
x_{11}+2x_{21}+2x_{31}+x_{41}+x_{51}+x_{61}+x_{71}+x_{81}+x_{12} & \geq & 1500 \\
x_{41}+x_{51}+2x_{61}+x_{91}+2x_{101}+3x_{111}+2x_{121}+x_{131}+x_{141}+x_{151}+x_{22}+2x_{32}+x_{42} & \geq & 800 \\
x_{31}+x_{51}+x_{71}+2x_{81}+x_{121}+x_{131}+2x_{141}+3x_{151}+x_{161}+2x_{171}+3x_{181}+4x_{191}+x_{42}+x_{52}+2x_{62} & \geq & 2400 \\
x_{ij} & \geq & 0 \ \ \forall i,j
\end{array}$$

Analicemos la primera restricción. En ella se impone que la cantidad de piezas de 80 cms. debe ser al menos de 1500. En efecto, la configuracion x_{11} tiene exactamente una pieza de 80 cms, la configuración x_{21} contiene 2 piezas de 80cms, etc.

Ejemplo 3.1.6 *Un pequeño banco asigna un máximo de $20.000 para préstamos personales y para automóvil durante el mes siguiente. El banco cobra una tasa de interés anual del* 14% *a préstamos personales y del* 12% *a préstamos para automóvil. Ambos tipos de préstamos se saldan en períodos de 3 años. El monto de los préstamos para automóvil debe ser, cuando menos, dos veces mayor que el de los préstamos personales. La experiencia pasada ha demostrado que los adeudos no cubiertos constituyen el* 1% *de todos los préstamos personales. ¿Cómo deben asignarse los fondos?*

Solución:

Variables de decisión:

x_1 : dinero asignado a préstamos personales
x_2 : dinero asignado a préstamos para automóvil

Planteamiento:

$$\begin{array}{rrrcr}
\max z = & (0,14 \cdot 0,99 - 0,01)x_1 & +0,12x_2 & & \\
 & x_1 & +x_2 & \leq & 20.000 \\
 & 2x_1 & -x_2 & \leq & 0 \\
 & x_1 & & \geq & 0 \\
 & & x_2 & \geq & 0
\end{array}$$

Donde maximizamos la cantidad total de intereses a recibir, menos la fracción de créditos personales que no se pagan. La primera restricción corresponde a la cantidad de dinero a repartir en créditos y la segunda dice que lo destinado a préstamos para automóviles debe ser al menos el doble que lo que se destina a préstamos personales. Esto podría obedecer a alguna política del banco para recuperar el dinero en caso de no pago, por ejemplo, embargando los vehículos.

Ejemplo 3.1.7 *Una empresa dedicada a la elaboración de trajes de seguridad para obreros forestales ha desarroallado dos nuevos tipos de trajes, que vende a tiendas en todo el país.*

Aunque la demanda por estos trajes excede a su capacidad de producción, la empresa sigue trabajando a un ritmo constante, limitando su trabajo en estos nuevos artículos a 50 horas/semana. El traje tipo I se produce en 3.5 horas y genera una ganacia de US\$28, mientras que el traje tipo II toma 4 horas para su producción y da una ganancia de US\$31. ¿Cuántos trajes de cada tipo deberá producir semanalmente la empresa, si su objetivo es maximizar su ganacia total?

Solución:

Variables de decisión:

x_1 : número de trajes tipo I
x_2 : número de trajes tipo II

Planteamiento:

$$\begin{array}{rrrcl} (P)\max z = & 28x_1 & +31x_2 & & \\ & 3.5x_1 & +4x_2 & \leq & 50 \\ & x_1 & & \geq & 0 \\ & & x_2 & \geq & 0 \end{array}$$

Ejemplo 3.1.8 *Suponga que una persona acaba de heredar \$6000 y desea invertirlos. Al oir esta noticia, dos amigos distintos le ofrecen la oportunidad de participar como socio en sus negocios. En ambos casos, la inversión significa dedicar un poco de tiempo el verano siguiente, al igual que invertir efectivo. Con el primer amigo, al convertirse en socio completo, tendría que invertir \$5000 y 4000 horas, con una ganacia estimada (ignorando el valor del tiempo) de \$4500. Las cifras correspondientes a la proposición del segundo amigo son \$4000 y 500 horas, con una ganacia estimada de \$4500. Sin embargo, ambos am,igos son flexibles y le permitirán entrar al negocio con cualquier fracción de la sociedad. La participación de las utilidades sería proporcional a esa fracción. Como de todas maneras esta persona está buscandoun trabajo interesante para el verano (600 horas a lo sumo), ha decidido participar en una o ambas sociedades, con la combinación que maximize la ganacia total estimada. Formule un modelo de programación lineal para este problema.*

Solución:

Variables de decisión:

x_1 : dinero invertido en el primer negocio
x_2 : dinero invertido en el primer negocio

Planteamiento:

$$
\begin{array}{rrrcr}
(P)\max z = & \frac{9}{10}x_1+ & \frac{9}{8}x_2 & & \\
 & x_1 & +x_2 & \leq & 6000 \\
 & \frac{2}{25}x_1 & +\frac{1}{8}x_2 & \leq & 600 \\
 & x_1 & & \leq & 5000 \\
 & & x_2 & \leq & 4000 \\
 & & x_1, x_2 & \geq & 0
\end{array}
$$

Ejercicio 3.1.1 *Demuestre, usando el teorema de Farkas, que si el problema*

$$
\begin{array}{rl}
\min & c^t x \\
Ax & = b \\
x & \geq 0
\end{array}
$$

es no acotado, entonces no existe $y \in I\!R^m$ *tal que* $A^t y \leq c$*, donde* $A \in \mathcal{M}_{m\times n}(I\!R^n)$.

Plantee los siguientes problemas de programación lineal.

Ejercicio 3.1.2 *La National Free Transportation Agency (NAFTA), debe decidir un programa de formación y contratación de nuevas azafatas para los próximos seis meses.*

Las exigencias a respetar son expresadas en horas de vuelo de azafatas: 8.000 en enero, 9.000 en febrero, 8.000 en marzo, 10.000 en abril, 9.000 en mayo y 12.000 en junio.

La formación de una nueva azafata dura un mes. Esta formación comprende 100 horas de vuelo en líneas de la compañía. Estas 100 horas se pueden deducir de exigencias que las azafatas deben cumplir, es decir, sirven para satisfacer las exigencias de horas de vuelo de azafatas de la compañía.

Cada azafata <u>*experimentada*</u> *puede entregar hasta 150 horas de vuelo por mes. La compañía dispone de 60 azafatas experimentadas al 1 de enero.*

Cada azafata experimentada recibe un sueldo de US\$800 por mes, independientemente del número de horas que preste servicio. Cada mes, el 10% de las azafatas experimentadas deja su trabajo por diversas razones.

Al cabo de un mes de formación, que cuesta US\$400 a la compañía, una azafata <u>*aprendiz*</u> *se convierte en azfata experimentada.*

Ejercicio 3.1.3 *Un proyecto de construcción municipal requiere fondos de 2 millones, 4 millones, 8 millones y 5 millones, durante los próximos 4 años, respectivemente. Suponga que todo el dinero para un año se requiere al principio del año. El municipio intenta vender el número exacto de bonos a largo plazo, suficiente para cubrir los fondos requeridos para el* *proyecto, y todos estos bonos, independientemenet de cuándo sean vendidos, serán pagados (se vencerán) en la misma fecha de algún año futuro distante. Se ha estimado que los procentajes de interés en el mercado (es decir, los costos de vender los bonos) de bonos a largo plazo en los próximos 4 años serán del 7%, 6%, 6.5% y 7.5%, respectivemente. El pago de intereses de los bonos empezará un año después de haber completado el proyecto y continuará durante 20 años, después de lo cual los bonos serán pagados. Por otra parte, se ha estimado que durante el mismo período, los porcentajes de interés a corto plazo sobre los depósitos a tiempo fijo (es decir, lo que la ciudad puede ganar en depósitos) serán del 6%, 5.5% y 4.5%, respectivamente (es claro que el municipio no invertirá dinero en depósitos a corto plazo durante el cuarto año). Cuál es la estrategia óptima que debe seguir el gobierno municipal en la venta de bonos y en el depósito de fondos en cuentas a tiempo fijo para poder completar el proyecto de construcción?*

Ejercicio 3.1.4 *Un granjero posee 100 hectáreas (ha.) que pueden ser utilizadas para el cultivo de trigo y maíz. El rendimiento por ha. es de 60 quintales anuales de trigo y de 95 quintales de maíz.*

Cualquier fracción de las 100 ha. puede ser destinada al cultivo de trigo o maíz. El trabajo necesario es de 4 hrs. por ha. anuales, más 0.15 hr. por quintal de trigo y 0.70 hr. por quintal de maíz. El costo de las semillas y avono es de $20 por quintal de trigo y $12 por quintal de maíz.

El granjero puede vender su trigo a $175 el quintal y su maíz a $95 el quintal. A la compra, le costarían respectivemente $250 y $150. Puede también criar cerdos y pollos. Los vende cuando han alcanzado la edad de 12 meses. Un cerdo se vende a $4.000. Un ave se vende en términos de cerdo-equivalente (el número de pollos necesarios para obtener $4.000 al memento de la venta).

Un cerdo requiere 25 quintales de trigo o 20 quintales de maíz, aí como 25 hrs. de trabajo y 25 m^2 *de terreno. Un cerdo-equivalente de pollos requiere 25 quintales de maíz o 10 quintales de trigo, así como 40 hrs. de trabajo y 15* m^2 *de terreno.*

El granjero dispone de 10.000 m^2 *de terreno para la crianza. Dispone también de 2.000 hrs. de trabajo anuales y puede poner a su familia a trabajar, disponiendo así de 2.000 hrs. suplementarias. Puede también contratar horas suplementarias de obreros agrícolas al costo de $150 la hora.*

Cada hora de obrero agrícola demanda 0.15 hr. de trabajo de supervisión de parte del granjero.

Determine las superficies a destinar al cultivo de trigo y/o maíz y las cantidades de cerdos y/o pollos a producir, de manera de maximizar el beneficio.

Explicite los supuestos usados en la modelación.

3.2 Solución de un LP

Definición 3.2.1 *Sea $f : I\!R^n \to I\!R$ una función continuamente diferenciable. Una dirección $d \in I\!R^n$ se dirá* ***dirección de descenso*** *de f en el punto $\overline{x}$ factible, si y solamente si $\forall \lambda \in [0, \overline{\lambda}]$ se tiene que $f(\overline{x} + \lambda d) \leq f(\overline{x})$.*

Proposición 3.2.1 *Se cumple que:*

- *Si $\nabla f(\overline{x}) < 0$ entonces d es dirección de descenso.*
- *Si existe $\overline{\lambda} > 0$ tal que $f(\overline{x} + \lambda d) \leq f(\overline{x})$ $\forall \lambda \in [0, \overline{\lambda}]$, entonces $\nabla f(\overline{x})^t d \leq 0$*

Demostración. Si $\nabla f(\overline{x})^t d < 0$ entonces, usando la definición de derivada direccional:

$$\lim_{\lambda \to 0^+} \frac{f(\overline{x}+\lambda d) - f(\overline{x})}{\lambda} < 0$$

Entonces,$f(\overline{x} + \lambda d) - f(\overline{x}) < 0$. Con lo que se tiene que $\exists \overline{\lambda} > 0$ tal que

$$f(\overline{x} + \lambda d) < f(\overline{x}) \;\; \forall \lambda \in [0, \overline{\lambda}]$$

y d es dirección de descenso.

Por otra parte, si $\exists \overline{\lambda} > 0$ tal que:

$$f(\overline{x} + \lambda d) - f(\overline{x}) < 0 \;\; \forall \lambda \in [0, \overline{\lambda}]$$

tenemos que $\lim\limits_{\lambda \to 0^+} \frac{f(\overline{x}+\lambda d) - f(\overline{x})}{\lambda} < 0$ y por lo tanto, $\nabla f(\overline{x})^t d \leq 0$ ■

Proposición 3.2.2 *El vector $d = -\frac{\nabla f(\overline{x})}{\|\nabla f(\overline{x})\|}$ es la dirección unitaria de máximo descenso de la función f en el punto $\overline{x}$.*

Demostración. Notemos primero que $-\bigtriangledown f(\overline{x})$ efectivamente es dirección de descenso, pues $\bigtriangledown f(\overline{x})^t(-\bigtriangledown f(\overline{x})) = -\left\|\bigtriangledown f(\overline{x})\right\|^2 \leq 0$.

El problema que se plantea es el siguiente: $\min\limits_{\|d\|\leq 1} \nabla f(\overline{x})^t d$. De la desigualdad de Cauchy-Schwartz se tiene que

$$|\nabla f(\overline{x})^t d| \leq \|\nabla f(\overline{x})\| \, \|d\| \leq \|\nabla f(\overline{x})\|$$

Esto implica que $-\|\nabla f(\overline{x})\| \leq \nabla f(\overline{x})^t d$, que se alcanza con igualdad para $d = -\frac{\bigtriangledown f(\overline{x})}{\|\bigtriangledown f(\overline{x})\|}$. ■

Teorema 3.2.1 *Sea S un poliedro no vacío en la forma canónica y sea el problema*

$$\begin{array}{ll} (P) \min & c^t x \\ & x \in S \end{array}$$

Entonces una de las siguientes afirmaciones es verdadera:

(i) (P) es no acotado.
(ii) (P) tiene un vértice (punto extremo) como solución.

Demostración. Sea $x \in S$. Del teorema (2.1.4) sabemos que x puede escribirse como la suma de una combinación convexa de puntos extremos y una combinación lineal positiva de direcciones extremas de S, es decir, $x = \sum\limits_{i=1}^{k} \lambda_i x_i + \sum\limits_{j=1}^{l} \mu_j d_j$, donde $\lambda_i \in [0,1]$ $\forall i = 1,...,k$, $\sum\limits_{i=1}^{k} \lambda_i = 1$; $\mu_j \geq 0$ $\forall j \in \{1,...,l\}$.

Si existe $\widetilde{j}$ tal que $c^t d_{\widetilde{j}} < 0$, entonces $c^t x = \sum\limits_{i=1}^{K} \lambda_i c^t x_i + \sum\limits_{j=1}^{l} \mu_j c^t d_j \to -\infty$ cuando $\mu_{\widetilde{j}} \to -\infty$. Es decir, el problema es no acotado.

Si no existe una dirección satisfaciendo que $c^t d_j < 0$, entonces $c^t d_j \geq 0 \quad \forall j = 1,...,l$, luego

$$c^t x = \sum_{i=1}^{k} \lambda_i c^t x_i + \underbrace{\sum_{j=1}^{l} \mu_j c^t d_j}_{\geq 0} \geq \sum_{i=1}^{K} \lambda_i c^t x_i$$

Sea x_v tal que $c^t x_v = \min\limits_{i=1,..,k} \{c^t x_i\}$ (que existe, pues el conjunto es finito). Luego $c^t x \geq c^t x_v \sum\limits_{i=1}^{K} \lambda_i = c^t x_v$, es decir, x_v es óptimo. ■

Motivación: Solución gráfica en $\mathbf{R}^2$

Consideremos el siguiente problema lineal

$$
(\mathrm{P}) \qquad \begin{array}{lrrcl}
\min & -2x_1 & +3x_2 & & \\
 & 2x_1 & -x_2 & \leq & 3 \\
 & -x_1 & +2x_2 & \leq & 2 \\
 & & x_2 & \geq & 1 \\
 & x_1 & ,x_2 & \geq & 0
\end{array}
$$

Gráficamente, cuando se minimizan los costos lo que se hace es desplazar la curva de costos en la dirección $-c\ (= -\bigtriangledown(cx))$ que, como ya probamos, es la dirección de máximo descenso. Con esto tenemos que el valor óptimo será aquel último punto factible que alcance la función objetivo en su descenso.

(insertar gráfico)

En la figura podemos observar que el óptimo para este problema es el punto $(2, 1)$, que es un punto extremo!

Observación 3.2.1 *La observación anterior es cierta cuando las variables de decisión de un programa lineal son continuas (pertenecen a un dominio acotado de $\mathbb{R}^n$). En el caso de que las variables de*

decisión sean discretas, puede ocurrir que el óptimo no corresponda a un vértice. Este caso será tratado más adelante, en el Capítulo de Programación Entera.

Desgraciadamente, cuando aumenta la dimensión del problema, ya no es posible resolverlo de manera gráfica.

3.2.1 Algoritmo simplex

Vimos que para poder resolver un problema de programación lineal necesitamos solo considerar los vértices del poliedro $S = \{x \in \mathbb{R}^n / Ax = b\}$ como candidatos a solución óptima del problema. Para un número grande de variables (n) y restricciones (m) vimos también que el número de vértices puede ser enorme, $\begin{pmatrix} n \\ m \end{pmatrix}$, por lo que una metodología más sistemática se hace necesaria.

El método simplex, desarrollado por Dantzig (1949), tiene una idea geométrica muy simple: primero encuentra una base factible (un vértice de S). Luego el método se mueve de vértice en vértice, a través de las aristas de S que sean direcciones de descenso para la función objetivo, generando una sucesión de vértices cuyos valores por f son estrictamente decrecientes, con lo que se asegura que un mismo vértice no es visitado dos veces. Así, como el número de vértices es finito,el algoritmo converge en tiempo finito; esto significa que encuentra una solución óptima, o una arista a lo largo de la cual la función objetivo es no acotada.

A la búsqueda de una base factible se la llama Fase I del algoritmo simplex. El resto del procedimiento se conoce como Fase II.

Fase II del Algoritmo Simplex

Consideremos el problema

$$\text{(PL)} \min_{x\in S} c^t x$$

con $S = \{x \in I\!R^n / Ax = b\}$ poliedro cerrado, convexo.

Supongamos que $A \in \mathcal{M}_{mn}$ es de rango m, entonces, por lo visto en el Capítulo 1, A puede escribirse de la forma $A = [B, N]$, con $B \in \mathcal{M}_{mm}$ invertible.

Notemos $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}, \quad c = \begin{bmatrix} c_B \\ c_N \end{bmatrix}$

Entonces, $Ax = Bx_B + Nx_N = b$ con lo que finalmente $x_B = B^{-1}b - B^{-1}Nx_N$

El problema (PL) es equivalente a

$$\begin{array}{lll} \min\limits_{s.a} & c_B^t B^{-1}b + (c_N^t - c_B^t B^{-1}N)x_N & \\ & x_B + B^{-1}Nx_N & = B^{-1}b \\ & x_N & \geq 0 \\ & x_B & \geq 0 \end{array}$$

Consideremos un punto x extremo (es factible), $\begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \geq 0$

Con esto, $c^t x = c_B^t x_B + c_N^t x_N = c_B^t B^{-1}b$, por lo tanto, si $c_N^t - c_B^t B^{-1}N \geq 0$ no es aconsejable dar valor positivo a las variables en x_N.

Definición 3.2.2 *La cantidad $\pi^t = c_B^t B^{-1}$ se conoce como* ***vector de multiplicadores del simplex****. Esta terminología proviene de la interpretación de las componenetes de π como multiplicadores de Lagrange y como precio de equilibrio en el óptimo, como veremos más adelante.*

Ejemplo 3.2.1

$$\begin{array}{lrrcl} \min & 3x_3 & +x_4 & & \\ & -3x_3 & +3x_4 & \leq & 6 \\ & -8x_3 & +4x_4 & \leq & 4 \\ & x_3 & ,x_4 & \geq & 0 \end{array}$$

Escribamos el problema en la forma canónica:

$$\begin{array}{lccccl} \min & 0x_1 & +0x_2 & +3x_3 & +x_4 & \\ & x_1 & & -3x_3 & +3x_4 & = & 6 \\ & & x_2 & -8x_3 & +4x_4 & = & 4 \\ & & & & x_i & \geq & 0 \quad \forall i \in \{1,2,3,4\} \end{array}$$

Elijamos $B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $N = \begin{pmatrix} -3 & 3 \\ -8 & 4 \end{pmatrix}$, $\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$, $\begin{pmatrix} c_B \\ c_N \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 3 \\ 1 \end{pmatrix}$

Donde: x_B : variables básicas o en la base

x_N : variables no-básicas o fuera de la base

x_1, x_2 : variables de holgura

x_3, x_4 : variables estructurales

Se puede despejar x_1, x_2 en función de x_3, x_4 y reemplazar en la función objetivo. Notar que todo queda igual, pues $B = I$ y $\overline{c}_B = 0$.
Como $c_N^t - \pi^t N = (3, 1) \geq 0$, la solución es óptima.

$$\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} 6 \\ 4 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

- Criterio de Optimalidad

En un problema de minimización escrito en la forma canónica, si las variables no básicas tienen asociado un coeficiente $\overline{c}_N^t = c_N^t - \pi^t N \geq 0$, entonces la solución $\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ es óptima.

Habitualmente, los datos se ordenan en un cuadro:

$$\begin{array}{cc|c} 0 & \overline{c}_N^t & -\pi^t b \\ \hline I & B^{-1}N & B^{-1}b \end{array} \; (*)$$

$\overline{c}_N^t = c_N^t - \pi^t N$ se llama vector de costos reducidos.

Consideremos el siguiente ejemplo:

Ejemplo 3.2.2

$$\begin{array}{rrrcl} \min & -3x_3 & +x_4 & & \\ & -3x_3 & +3x_4 & \leq & 6 \\ & -8x_3 & +4x_4 & \leq & 4 \\ & x_3 & ,x_4 & \geq & 0 \end{array}$$

Siguiendo el procedimiento anterior, escribamos la siguiente tabla:

x_1	x_2	$\downarrow$ x_3	x_4	$-z$
0	0	-3	1	$0 = -\pi^t b$
1	0	-3	3	6
0	1	-8	4	4

Tal como en el ejemplo anterior, $x = \begin{pmatrix} 6 & 4 & 0 & 0 \end{pmatrix}^t$ es una solución factible; pero no es óptima, pues $\bar{c}_3 < 0$

¿Conviene hacer crecer una variable no básica a un valor positivo? Sí.

$$x_1 = -3x_4 + 6 + 3x_3$$
$$x_2 = -4x_4 + 4 + 8x_3$$

Tomando $x_4 = 0$ (fuera de la base), x_3 puede crecer indefinidamente, disminuyendo el costo total. Más aún, este problema es no acotado.

- Criterio de no acotamiento

Si un costo reducido es negativo y los elementos en la columna correspondiente son negativos o nulos, en al menos una de las filas, entonces el problema es no acotado.

Observación 3.2.2 *Si la columna es enteramente nula, la variable es irrelevante.*

Ejemplo 3.2.3

$$\begin{array}{rrrcl} \min & 3x_3 & -x_4 & & \\ & -3x_3 & +3x_4 & \leq & 6 \\ & -8x_3 & +4x_4 & \leq & 4 \\ & x_3 & ,x_4 & \geq & 0 \end{array}$$

Escribamos la tabla:

			$\downarrow$	
x_1	x_2	x_3	x_4	$-z$
0	0	3	-1	0
1	0	-3	3	6
0	1	-8	4	4

x_4 es una variable no básica cuyo costo reducido es negativo[3], luego conviene hacerla entrar a la base.

¿Cuánto puede crecer? Los datos en la columna asociada son todos positivos, luego el problema es acotado.

$$\begin{aligned} x_1 &= -3x_4 + 6 \\ x_2 &= -4x_4 + 4 \\ x_3 &= 0 \qquad \text{(fuera de la base)} \end{aligned}$$

Se necesita que $x_1 \geq 0, x_2 \geq 0$, lo que implica que $x_4 \leq 1$.

Más aún, $x_4 = \min\{\frac{6}{3}, \frac{4}{4}\} = \min\limits_{\overline{a}_{ij}>0}\{\frac{\overline{b}_j}{\overline{a}_{ij}}\}$

			$\downarrow r$		
x_1	x_2	x_3	x_4	z	
0	0	3	-1	0	
1	0	-3	3	6	
0	1	-8	$\boxed{4}$	4	$\leftarrow s$

- Criterio de pivoteo

Se hace entrar a la base aquella variable cuyo costo reducido sea negativo. Sea x_r la elegida para entrar a la base, ¿cuál sale?

Se busca s tal que $\dfrac{\overline{b}_s}{\overline{a}_{sr}} = \min\limits_i\{\dfrac{\overline{b}_i}{\overline{a}_{ir}} / \overline{a}_{ir} > 0\}$

Se pivotea sobre (s, r)

Volvamos a la tabla:

[3]Si bien no hay resultados respecto a la elección entre las variables con costos negativos, conviene hacer ciertas convenciones. En este caso eligiremos siempre la variable de costo reducido *más* negativo

$$\begin{array}{ccc|c} \overline{c}_1 & \cdots & \overline{c}_n & -\overline{z} \\ \hline \overline{a}_{11} & & \overline{a}_{1n} & \overline{b}_1 \\ \vdots & & \vdots & \\ \overline{a}_{m1} & \cdots & \overline{a}_{mn} & \overline{b}_m \end{array}$$

Donde suponemos $\overline{b}_i \geq 0$

(1) Si $\overline{c}_j \geq 0 \quad \forall j = 1, ..., n$, entonces la solución en curso en óptima. Las variables básicas son iguales a $\overline{b}_i$ y las no básicas son nulas.

(2) Si $\overline{c}_j < 0$ para algún j, la elegimos para entrar a la base. Usaremos el criterio descrito anteriormente de elegir la variable cuyo costo reducido es menor. Supongamos que dicha variable es la s.

(3) Si $\overline{a}_{is} \leq 0 \quad \forall i = 1, ..., m$, el problema es no acotado.

(4) Si $\overline{a}_{is} > 0$ para algún(os) i, se determina r tal que $\frac{\overline{b}_r}{\overline{a}_{rs}} = \min\{\frac{\overline{b}_i}{\overline{a}_{is}} / \overline{a}_{is} > 0\}$ y se pivotea en $\overline{a}_{rs}$:

$$\overline{a}_{ij} \leftarrow \overline{a}_{ij} - \frac{\overline{a}_{is}\overline{a}_{rj}}{\overline{a}_{rs}}$$

$$\overline{b}_i \leftarrow \overline{b}_i - \frac{\overline{a}_{is}\overline{b}_r}{\overline{a}_{rs}}$$

$$\overline{c}_j \leftarrow \overline{c}_j - \frac{\overline{c}_s\overline{a}_{rj}}{\overline{a}_{rs}}$$

$$\overline{z} \leftarrow \overline{z} - \frac{\overline{c}_s\overline{b}_r}{\overline{a}_{rs}}$$

Observación 3.2.3 *Notar que esto corresponde precisamente al pivoteo de Gauss para la inversión de matrices, lo que es consistente con el esquema presentado en la tabla (*)*

Ejemplo 3.2.4

$$\begin{array}{lrrcl} \min & -x_1 & -3x_2 & & \\ & x_1 & +x_2 & \leq & 3 \\ & -3x_1 & +x_2 & \leq & 2 \\ & x_1 & x_2 & \geq & 0 \end{array}$$

El problema, escrito en la forma canónica, queda:

$$\begin{array}{lrrrrcll}
\min & -x_1 & -3x_2 & +0x_3 & +0x_4 & & & \\
 & x_1 & +x_2 & +x_3 & & = & 3 & \\
 & -3x_1 & +x_2 & & +x_4 & = & 2 & \\
 & & & & x_i & \geq & 0 & \forall i \in \{1,2,3,4\}
\end{array}$$

Escribamos la tabla:

$$\begin{array}{cccc|cc}
 & \downarrow r & & & & \\
x_1 & x_2 & x_3 & x_4 & -z & \\
\text{-1} & \text{-3} & 0 & 0 & 0 & \\
\hline
1 & 1 & 1 & 0 & 3 & \\
\text{-3} & \boxed{1} & 0 & 1 & 2 & \overrightarrow{s}
\end{array}
\rightsquigarrow (\textit{pivoteando})
\begin{array}{cccc|cc}
\downarrow r & & & & & \\
x_1 & x_2 & x_3 & x_4 & -z & \\
\text{-10} & 0 & 0 & 3 & 6 & \\
\hline
\boxed{4} & 0 & 1 & \text{-1} & 1 & \overrightarrow{s} \\
\text{-3} & 1 & 0 & 1 & 2 &
\end{array}$$

$$\rightsquigarrow
\begin{array}{cccc|c}
x_1 & x_2 & x_3 & x_4 & -z \\
0 & 0 & 5/2 & 1/2 & 17/2 \\
\hline
1 & 0 & 1/4 & \text{-}1/4 & 1/4 \\
0 & 1 & 3/4 & 1/4 & 11/4
\end{array}$$

Identificamos $x^* = \begin{pmatrix} \frac{1}{4} & \frac{11}{4} & 0 & 0 \end{pmatrix}^t$ como la solución óptima, luego el valor óptimo es $z = -17/2$

Definición 3.2.3 *Si una o más variables básicas se anulan, entonces la solución de un PL se dice **degenerada**. En caso contrario, la llamaremos **no-degenerada**.*

Observación 3.2.4 *Cuando la solución x de un PL es degenerada, existe más de una base asociada a ese vértice. En efecto, si $x \in I\!R^n$ tiene $p < m$ componentes positivas, donde m es el número de restricciones del problema, entonces podrían haber $\begin{pmatrix} n-p \\ m-p \end{pmatrix}$ soluciones básicas factibles diferentes correspondientes a x.*
El punto x es el mismo, pero los conjuntos de variables etiquetadas básicas y no básicas son diferentes.

Fase I del Algoritmo Simplex
Obtención de una solución básica factible

Hasta ahora, sabemos resolver un PL dada una solución básica inicial. Esta generalmente viene dada por la base de las variables de holgura. La Fase I del Simplex consiste en encontrar una base factible cuando no se tiene directamente desde el problema.

Sabemos que el problema

$$\text{(P)}\quad \begin{array}{lll} \min & c^t x & \\ & Ax & = b \\ & x & \geq 0 \end{array}$$

tiene por solución (si el problema es acotado) a

$$x = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \geq 0$$

que se llama solución básica factible.

El problema es conocer una solución básica factible para comenzar el algoritmo.

El problema de factibilidad de (P) puede plantearse como el siguiente problema auxiliar:

$$(P_a)\quad \begin{array}{lll} \min & \sum\limits_i x_{ai} & \\ & Ax + x_a & = b \\ & x, x_a & \geq 0 \end{array}$$

Notemos que $x \in I\!R^n$ $x_a \in I\!R^m$. Bajo el supuesto razonable de $b \geq 0$, una solución básica factible evidente para este problema es $x = 0$ y $x_a = b$, luego podemos usar la Fase II del algoritmo simplex para resolverlo.

(P_a) se resuelve considerando $B = I$ (x_a en la base), $N = A, c_B = 1^t$ y $c_N = 0^t$.
Así, para este problema el cuadro es:

$$\begin{array}{cc|c} \overline{c}_N^t & 0 & -c_B^t b \\ \hline A & I & b \end{array} \quad \Leftrightarrow \quad \begin{array}{cc|c} c_N^t - c_B^t A & 0 & -c_B^t b \\ \hline A & I & b \end{array} \quad \Leftrightarrow \quad \begin{array}{cc|c} -1^t A & 0 & -1^t b \\ \hline A & I & b \end{array}$$

Las variables x_{ai} son llamadas *variables artificiales* y el propósito del problema (P_a) es llevarlas a tomar valores nulos. Esto es posible, siempre que el problema original tenga una solución factible. En tal caso,el método simplex terminará con una solución básica factible, donde $x_{ai} = 0 \ \ \forall i$.

- Si en el óptimo de (P_a) se tiene algún $x_{ai} = 0 \ \ \forall i = 1, ..., m$, entonces la solución en curso es solución factible de (P).

La solución de (P_a) satisface

$$\begin{pmatrix} A & I \end{pmatrix} \begin{pmatrix} x \\ x_a \end{pmatrix} = b,$$

con $\begin{pmatrix} x \\ x_a \end{pmatrix} \geq 0$, luego $\begin{pmatrix} x \\ x_a \end{pmatrix}$ es solución de (P_a) si y sólo si x es solución de (P)

- Si en el óptimo de (P_a) se tiene algún $x_{ai} > 0$, entonces el poliedro es vacío, es decir, (P) es infactible.

Observación 3.2.5 *Otra forma de abordar el problema de obtener una solución básica factible es considerar el siguiente problema de optimización:*

$$\min_{(x,x_a)} \sum_{i=1}^{n} c_i x_i + M \sum_{j=1}^{m} x_{ai}$$

donde M se escoge suficientemente grande, de modo que eventualmente todas las variables artificiales tomarán el valor cero, cuando el problema original tanga una solución básica factible.

Ejemplo 3.2.5
$$\begin{array}{rrrcll} \max & -x_1 & +2x_2 & & & \\ & x_1 & +x_2 & \geq & 1 & \\ & x_1 & -x_2 & = & 0 & \\ & x_1 & & \leq & 4 & \\ & & x_i & \geq & 0 & i \in \{1,2\} \end{array}$$

Escribamos el problema en su forma canónica:

$$
(P)\quad \begin{array}{rrrrrrl}
-\min & -x_1 & -2x_2 & & & & \\
 & x_1 & +x_2 & -x_3 & & = & 1 \\
 & x_1 & & & +x_4 & = & 4 \\
 & x_1 & -x_2 & & & = & 0 \\
 & & & & x_i & \geq & 0 \quad \forall i \in \{1,2,3,4\}
\end{array}
$$

Luego, agregando las variables artificiales x_5, x_6, x_7 (3, dado el número de restricciones del problema $(P1)$) queda:

$$
\begin{array}{rrrrrrrrrl}
\min & & & & & x_5 & +x_6 & +x_7 & & \\
 & x_1 & +x_2 & -x_3 & & +x_5 & & & = & 1 \\
 & x_1 & & & +x_4 & & +x_6 & & = & 4 \\
 & x_1 & -x_2 & & & & & +x_7 & = & 0 \\
 & & & & & & & x_i & \geq & 0 \quad \forall i \in \{1,2,3,4,5,6,7\}
\end{array}
$$

Ya tenemos planteado el problema que necesitamos resolver. Escribamos la tabla y apliquemos el método simplex [4]:

$$
\begin{array}{ccccccc|cc}
\downarrow & & & & & & & & \\
-3 & 0 & 1 & -1 & 0 & 0 & 0 & -5 & \\
\hline
1 & 1 & -1 & 0 & 1 & 0 & 0 & 1 & \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 4 & \\
1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & \leftarrow
\end{array}
\qquad x^1 = (\ 0\ \ 0\ \ 0\ \ 0\ \ \mathbf{1}\ \ \mathbf{4}\ \ \mathbf{0}\ )^t
$$

Primera iteración:

$$
\begin{array}{ccccccc|cc}
 & \downarrow & & & & & & & \\
0 & -3 & 1 & -1 & 0 & 0 & 3 & -5 & \\
\hline
0 & 2 & -1 & 0 & 1 & 0 & -1 & 1 & \leftarrow \\
0 & 1 & 0 & 1 & 0 & 1 & -1 & 4 & \\
1 & -1 & 0 & 0 & 0 & 0 & 1 & 0 &
\end{array}
\qquad x^2 = (\ \mathbf{0}\ \ 0\ \ 0\ \ 0\ \ \mathbf{1}\ \ \mathbf{4}\ \ 0\ )^t
$$

Segunda iteración:

$$
\begin{array}{ccccccc|cc}
 & & & \downarrow & & & & & \\
0 & 0 & -1/2 & -1 & 3/2 & 0 & 3/2 & -7/2 & \\
\hline
0 & 1 & -1/2 & 0 & 1/2 & 0 & -1/2 & 1/2 & \\
0 & 0 & 1/2 & 1 & -1/2 & 1 & -1/2 & 7/2 & \leftarrow \\
1 & 0 & -1/2 & 0 & 1/2 & 0 & 1/2 & 1/2 &
\end{array}
\qquad x^3 = (\ \mathbf{\tfrac{1}{2}}\ \ \mathbf{\tfrac{1}{2}}\ \ 0\ \ 0\ \ 0\ \ \mathbf{\tfrac{1}{2}}\ \ 0\ )^t
$$

[4] las negrillas señalan las variables en la base para cada iteración.

Tercera iteración:

$$\begin{array}{ccccccc|c} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ \hline 0 & 1 & -1/2 & 0 & 1/2 & 0 & -1/2 & 1/2 \\ 0 & 0 & 1/2 & 1 & -1/2 & 1 & -1/2 & 7/2 \\ 1 & 0 & -1/2 & 0 & 1/2 & 0 & 1/2 & 1/2 \end{array} \qquad x^4 = \begin{pmatrix} \mathbf{\frac{1}{2}} & \mathbf{\frac{1}{2}} & 0 & \mathbf{\frac{7}{2}} & 0 & 0 & 0 \end{pmatrix}^t$$

En la última iteración todas las variables artificiales salen de la base, los costos reducidos asociados toman todos el valor 1 y $z = 0$.

Ya tenemos una solución básica factible. Ahora eliminamos las variables artificiales y recalculamos los costos reducidos y el valor objetivo para esta solución:

$$\begin{array}{cccc|c} 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & -1/2 & 0 & 1/2 \\ 0 & 0 & 1/2 & 1 & 7/2 \\ 1 & 0 & -1/2 & 0 & 1/2 \end{array}$$

El vector de costos está dado por $c^t = \begin{pmatrix}1 & -2 & 0 & 0\end{pmatrix}^t$, por lo tanto:

$$c_B^t = \begin{pmatrix}c_2 & c_4 & c_1\end{pmatrix}^t = \begin{pmatrix}-2 & 0 & 1\end{pmatrix}^t \quad , c_N^t = \begin{pmatrix}c_3\end{pmatrix} = \begin{pmatrix}0\end{pmatrix}$$

El orden en que se escribe el costo para las variables básicas, depende del vector en la base canónica que las mismas tienen asociado.

Reconozcamos las matrices B y N en el problema original (P):

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}, N = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

Los costos reducidos serán:

$\overline{c_B}^t = \begin{pmatrix}0 & 0 & 0\end{pmatrix}$, como siempre

$$\overline{c_N}^t = c_N^t - \begin{pmatrix}-2 & 0 & 1\end{pmatrix} \underbrace{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}}_{\begin{pmatrix}-\frac{1}{2} & \frac{1}{2} & -\frac{1}{2}\end{pmatrix}^t \text{ columna no básica del cuadro}} = -\frac{1}{2}$$

En tanto que:

$$-\pi^t b = \begin{pmatrix}-2 & 0 & 1\end{pmatrix}\begin{pmatrix}1/2\\7/2\\1/2\end{pmatrix} = \tfrac{1}{2}$$

Con esto el cuadro queda:

$$\begin{array}{ccccc|l}
 & & \downarrow & & & \\
0 & 0 & -1/2 & 0 & 1/2 & \\
\hline
0 & 1 & -1/2 & 0 & 1/2 & \\
0 & 0 & 1/2 & 1 & 7/2 & \leftarrow \\
1 & 0 & -1/2 & 0 & 1/2 &
\end{array}$$

Aplicando la segunda fase del algoritmo simplex, obtenemos:

$$\begin{array}{cccc|c}
0 & 0 & 0 & 1 & 4 \\
\hline
0 & 1 & 0 & 1 & 4 \\
0 & 0 & 1 & 2 & 7 \\
1 & 0 & 0 & 1 & 4
\end{array}$$

con lo cual $x^t = \begin{pmatrix}4 & 4 & 7 & 0\end{pmatrix}$ es solución óptima para el problema, y el valor objetivo de la función es $-(-\pi^t b) = 4$ (recordemos que hicimos el cambio $\max z = -\min -z$).

3.3 Introducción a la dualidad en programación lineal

Comenzaremos el estudio de la dualidad en programación lineal con un ejemplo.

Ejemplo 3.3.1 *Una fábrica produce tres artículos en cantidades x_1, x_2, x_3,los cuales utilizan dos materias primas en su elaboración, digamos a y b.*

El proceso de producción debe satisfacer lo siguiente:

1) Para producir una unidad del artículo 1 se necesitan 2 unidades del recurso a y 5 del recurso b.

   Para producir una unidad del artículo 2 se necesitan 3 unidades del recurso a y 2 del recurso b.

   Para producir una unidad del artículo 3 se necesitan 1 unidades del recurso a y 1 del recurso b.

2) El recurso a está disponible hasta 10 unidades y el recurso b hasta 20 unidades, sin costo para el fabricante.

El precio de venta del producto 1 es \$4, el del producto 2 es \$1 y el del producto 3 es \$5.

El problema del fabricante será el de maximizar sus utilidades (sus ingresos por venta, en este ejemplo) sujeto a sus restricciones en la producción, es decir:

$$
(P) \quad \begin{array}{lrrrcll}
\max & 4x_1 & +x_2 & +5x_3 & & & \\
 & 2x_1 & +3x_2 & +x_3 & \leq & 10 & \\
 & 5x_1 & +2x_2 & +x_3 & \leq & 20 & \\
 & & & x_i & \geq & 0 & \forall i \in \{1,2,3\}
\end{array}
$$

Tomemos una combinación lineal positiva de las restricciones, con multiplicadores y_1, y_2:

$$y_1(2x_1 + 3x_2 + x_3) + y_2(5x_1 + 2x_2 + x_3) \leq 10y_1 + 20y_2$$

Esto que puede reescribirse de la forma:

$$x_1(2y_1 + 5y_2) + x_2(3y_1 + 2y_2) + x_3(y_1 + y_2) \leq 10y_1 + 20y_2$$

Si imponemos que

$$\begin{array}{rrcl} 2y_1 & +5y_2 & \geq & 4 \\ 3y_1 & +2y_2 & \geq & 1 \quad (1) \\ y_1 & +y_2 & \geq & 5 \end{array}$$

entonces, $z = 4x_1 + x_2 + 5x_3 \leq 10y_1 + 20y_2 = \omega$, es decir, ω acota superiormente a la función objetivo de (P) cuando se satisface (1). Luego, es razonable plantear el siguiente problema asociado a los multiplicadores y_i:

$$\begin{array}{llrrcl} (D) & \min & 10y_1 & +20y_2 & & \\ & & 2y_1 & +5y_2 & \geq & 4 \\ & & 3y_1 & +2y_2 & \geq & 1 \\ & & y_1 & +y_2 & \geq & 5 \\ & & y_1 & ,y_2 & \geq & 0 \end{array}$$

es decir, minimizar el gasto de recursos sujeto a restricciones en los precios.

Notemos que partimos de un problema de la forma

$$(\mathrm{P}) \qquad \max \begin{pmatrix} 4 & 1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

$$\begin{bmatrix} 2 & 3 & 1 \\ 5 & 2 & 1 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq \begin{pmatrix} 10 \\ 20 \end{pmatrix}$$

$$x_i \geq 0 \qquad \forall i = 1, 2, 3$$

y llegamos a otro de la forma

$$(\mathrm{D}) \qquad \min \begin{pmatrix} 10 & 20 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\begin{bmatrix} 2 & 3 & 1 \\ 5 & 2 & 1 \end{bmatrix}^t \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \geq \begin{pmatrix} 4 \\ 1 \\ 5 \end{pmatrix}$$

$$y_i \geq 0, \, \forall i = 1, 2$$

Esto motiva la siguiente definición.

Definición 3.3.1 *Sea:*

$$
\begin{array}{llll}
(P) & \max & c^t x & \\
 & & Ax & \leq b \\
 & & x & \geq 0
\end{array}
$$

un problema que llamaremos ***Problema Primal.*** *El problema:*

$$
\begin{array}{llll}
(D) & \min & b^t y & \\
 & & A^t y & \geq c \\
 & & y & \geq 0
\end{array}
$$

se llamará ***Problema Dual de (P).***

Teorema 3.3.1 *El dual de (D) es (P)*

Demostración. El problema:

$$
\begin{array}{llll}
(D) & \min & b^t y & \\
 & & A^t y & \geq c \\
 & & y & \geq 0
\end{array}
$$

es equivalente a:

$$
\begin{array}{llll}
(\widetilde{D}) & -\max \& (-b)^t y & & \\
 & & (-A)^t y & \leq -c \\
 & & y & \geq 0
\end{array}
$$

El problema dual asociado a $(\widetilde{D})$, según la definición anterior, es:

$$
\begin{array}{lll}
-\min & (-c)^t x & \\
 & (-A)x & \geq -b \\
 & x & \geq 0
\end{array}
$$

que a su vez es equivalente a:

$$
\begin{array}{llll}
 & \max & c^t x & \\
\text{(P)} & & Ax & \leq b \quad \blacksquare \\
 & & x & \geq 0
\end{array}
$$

Teorema 3.3.2 (Teorema de Dualidad Débil) *Sea (P) un problema primal de maximización y (D) su dual, y consideremos x e y, puntos factibles de (P) y (D), respectivamente. Entonces $c^t x \leq b^t y$, es decir, la función objetivo del problema dual acota superiormente a la del primal.*

Demostración. Si multiplicamos por $x^t \quad (\geq 0)$ la inecuación $A^t y \geq c$, se obtiene que $x^t A^t y \geq x^t c$, de donde $c^t x \leq (Ax)^t y \leq b^t y$, pues $y \geq 0$ ■

Corolario 3.3.1 *Sean $\widetilde{x}$ e $\widetilde{y}$ puntos factibles para (P) y (D). Si $c^t\widetilde{x} = b^t\widetilde{y}$, entonces $\widetilde{x}$ e $\widetilde{y}$ son óptimos respectivos.*

Demostración. Es consecuencia directa del teorema de Dualidad Débil:

$b^t\widetilde{y} = c^t\widetilde{x} \leq b^t y \quad \forall y$ punto factible de (D), es decir, $\widetilde{y}$ es óptimo de (D).

$c^t\widetilde{x} = b^t\widetilde{y} \geq c^t x \quad \forall x$ punto factible de (P), es decir, $\widetilde{x}$ es óptimo de (P). ■

Consideremos ahora un PL en la forma estándar

$$
\begin{array}{lll}
(P) \quad \min & c^t x & \\
 & Ax & = b \\
 & x & \geq 0
\end{array}
$$

y calculemos un dual para (P).

Notemos que (P) es equivalente a

$$
\begin{array}{lrcr}
(\widetilde{P}) \quad \min & (-c)^t x & & \\
 & Ax & \geq & b \\
 & (-A)x & \geq & -b \\
 & x & \geq & 0
\end{array}
$$

Cuyo dual es, de acuerdo a la definición:

$$
\begin{array}{lrlcc}
(\widetilde{D}) \quad \max & \begin{pmatrix} b^t & -b^t \end{pmatrix} & \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} & & \\
 & \begin{pmatrix} A^t & -A^t \end{pmatrix} & \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} & \leq & c \\
 & & \begin{pmatrix} y_1 & y_2 \end{pmatrix}^t & \geq & 0
\end{array}
$$

Es decir,

$$\begin{array}{rrcl} \max & b^t & (y_1 - y_2) & \\ & A^t & (y_1 - y_2) & \leq \; c \\ & & \begin{pmatrix} y_1 & y_2 \end{pmatrix}^t & \geq \; 0 \end{array}$$

que, tomando $y = y_1 - y_2 \in I\!R^m$ (irrestricto), es equivalente a:

$$\text{(D)} \qquad \begin{array}{rrccl} \max & b^t & y & & \\ & A^t & y & \leq & c \\ & & y & \in & I\!R^m \end{array}$$

Con esto, concluimos el siguiente teorema:

Teorema 3.3.3 *El problema dual de (P)* $\begin{array}{rrcl} \min & c^t x & & \\ & Ax & = & b \\ & x & \geq & 0 \end{array}$ *es (D)* $\begin{array}{rrcl} \max & b^t y & & \\ & A^t y & \leq & c \\ & y & \in & I\!R^m \end{array}$

La inversa también es cierta.

Ejercicio 3.3.1 *Escriba el dual del siguiente problema:*

$$\begin{array}{rrcl} \max & b^t y & & \\ & A^t y & \leq & b \end{array}$$

Teorema 3.3.4 *Teorema de Dualidad Fuerte*

Sean (P) $\begin{array}{rrcl} \min & c^t x & & \\ & Ax & = & b \\ & x & \geq & 0 \end{array}$ $\quad y \quad$ (D) $\begin{array}{rrcl} \max & b^t y & & \\ & A^t y & \leq & c \\ & y & \in & I\!R^m \end{array}$

Entonces:

a) *Si $\widetilde{z}$ (valor óptimo de (P)) es finito, entonces $\widetilde{\omega}$ (el óptimo de (D)) también lo es y se cumple $\widetilde{z} = \widetilde{\omega}$*

b) *Si $\widetilde{\omega}$ es finito, entonces $\widetilde{z}$ también lo es y $\widetilde{z} = \widetilde{\omega}$*

c) *Si (P) es no acotado, entonces (D) es infactible*

d) *Si (D) es no acotado, entonces (P) es infactible*

Demostración.

a) Dado que $\widetilde{z}$ es finito, existe un $\widetilde{x}$ solución óptima básica factible de (P). Entonces existe también B, submatriz de $A = [B, N]$, tal que $\widetilde{x} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = \begin{pmatrix} \widetilde{x}_B \\ \widetilde{x}_N \end{pmatrix}$

Además, los costos reducidos de las variables no básicas son positivos, es decir para $\pi = B^{-t}c_B$:

$$c_N^t - \pi^t N \geq 0 \qquad \text{lo que es equivalente a} \qquad N^t\,\pi \leq c_N$$

Probaremos que π es solución básica factible óptima de (D), con lo cual, $\widetilde{\omega} = b^t\pi$ será finito.

En efecto, π es factible para (D), pues

$A^t\pi = \begin{pmatrix} B^t \\ N^t \end{pmatrix} B^{-t}c_B = \begin{pmatrix} c_B \\ N^t\pi \end{pmatrix} \leq \begin{pmatrix} c_B \\ c_N \end{pmatrix} = c$

y π es óptimo para (D), pues

$\widetilde{\omega} = b^t\pi = b^tB^{-t}c_B = \pi^t b = \begin{pmatrix} c_B^t & c_N^t \end{pmatrix} \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = c^t\widetilde{x} = \widetilde{z}$

y por el teorema de Dualidad Débil, π es óptimo.

b) Análogo.

c) Supongamos que existe $\overline{y}$ tal que $A^t\overline{y} \leq c$ (factibilidad del dual). Por el teorema de Dualidad Débil, $b^t\overline{y} \leq c^tx \;\; \forall x$ punto primal factible. Pero (P) es no acotado, luego $b^t\overline{y} \leq -\infty$, lo que contradice que (P) es no acotado. Por lo tanto, (D) es infactible.

d) Análogo. ■

Resumamos los resultados anteriores en el siguiente cuadro:

Primal

Dual

	$\widetilde{z}$ finito	(P) no acotado	(P) infactible
$\widetilde{\omega}$ finito	Si	No	No
(D) no acotado	No	No	Si
(D) infactible	No	Si	Si

Teorema 3.3.5 (Holgura Complementaria) *Sean*

$$(P) \quad \begin{array}{rrcl} \min & c^t x & & \\ & Ax & = & b \\ & x & \geq & 0 \end{array} \qquad y \ (D) \quad \begin{array}{rrrcl} \max & b^t y & & & \\ & A^t y & +s & = & c \\ & & s & \geq & 0 \end{array}$$

Si x^* e y^* (con $s^* = c - A^t y^*$) son óptimos respectivos de (P) y (D), entonces $x^{*t} s^* = 0$.

Demostración: Por el teorema de dualidad fuerte $c^t x^* = b^t y^*$, luego

$c^t x^* = x^{*t} A^t y^* = x^{*t}(c - s^*) = x^{*t} c - x^{*t} s^* \Rightarrow x^{*t} s^* = 0$ ■

Observación 3.3.1 *La condición de holgura complementaria $x^{*t} s^* = 0$ se puede cambiar en forma equivalente, por $x_i^* s_i^* = 0 \quad \forall i = 1, .., n$*

3.4 Interpretación económica de la dualidad

El dual de un programa lineal surge naturalmente de las condiciones de optimalidad del problema primal.

Probaremos que si el problema primal tiene una interpretación económica, entonces también el dual y los valores óptimos de las variables duales pueden ser interpretados como precios.

Como ya vimos $x = \begin{bmatrix} x_B \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1} b \\ 0 \end{bmatrix}$ es una solución básica factible para un programa lineal en la forma estándar. Como $x_B \geq 0$, una pequeña perturbación del lado derecho Δb no provoca un cambio en la base óptima. Luego, cuando b es reemplazado por $b + \Delta b$, la nueva solución óptima se transforma en $x' = \begin{bmatrix} x'_B \\ 0 \end{bmatrix} = \begin{bmatrix} B^{-1}(b + \Delta b) \\ 0 \end{bmatrix}$

y el valor óptimo de la función objetivo se transforma en

$$\Delta z = c_B^t B^{-1} \Delta b = \pi^{*T} \Delta b$$

donde $\pi^* = B^{-T} c_B$ es el multiplicador del problema primal en el óptimo. Como probamos en el teorema ? π es la solución óptima del problema dual. Claramente, π_i puede verse como el precio marginal del i-ésimo recurso (es decir, lado derecho b_i), ya que da el cambio en el valor objetivo óptimo por unidad de incremento en ese recurso. Esta interpretación puede ser muy útil ya que indica la cantidad máxima que uno debe estar dispuesto a pagar por aumentar la cantidad del i-ésimo recurso. Note que las condiciones de holgura complementaria implican que el precio marginal para un recurso es cero si tal recurso no fue completamente utilizado en el óptimo. Otros nombres dados a este precio en el óptimo son *precio sombra* y *precio de equilibrio.*

Estos precios sombras son útiles también para determinar cuando es conveniente agregar una nueva actividad.

Veamos ahora otra interpretación económica posible. Supongamos que estamos bajo competencia perfecta, es decir los agentes de la economía actúan como tomadores de precios. Un productor resuelve:

$$\begin{array}{rrcl} \max & c^t x & & \\ & Ax & \leq & b \\ & x & \geq & 0 \end{array}$$

O sea, maximiza las utilidades dadas por el vector de precios c, sujeto a las restricciones de capacidad de su firma. En el óptimo las restricciones no necesariamente se cumplen con igualdad, es decir podrían sobrar ciertas materias primas que vende luego en el mercado en un precio dado por el vector $\lambda \geq 0$. Entonces, lo que resolvemos es:

$$\begin{array}{rc} \max & c^t x + \lambda^t (b - Ax) \\ & x \geq 0 \end{array}$$

Así, las utilidades de la firma están dadas por:

$$\varphi(\lambda) = \lambda^t b \; + \max\{(c - A^t \lambda)^t x \; , x \geq 0\}$$

Las posibles soluciones de este problema son dos:

- Si el vector $(c - A^t\lambda)$ tiene todas sus componentes negativas, dado que el vector x es positivo, se tiene que el máximo es cero, y $\varphi(\lambda) = \lambda^t b$.

- Si el vector $(c - A^t\lambda)$ tiene alguna coordenada positiva, entonces, por el mismo argumento el subproblema de maximización es no acotado, luego $\varphi(\lambda) = \infty$.

Sabemos que por arbitraje (o por el gran tamaño del mercado, o la crueldad de éste), no podemos obtener utilidades infinitas. O sea, el "mercado"me asigna la menor cantidad posible para mis utilidades, y eso puede hacerlo a través de precios, como conoce mi situación de costos e infraestructura (suponiendo información completa), resuelve:

$$\begin{array}{rrcl} \min & b^t\lambda & & \\ & A^t\lambda & \geq & c \\ & \lambda & \geq & 0 \end{array}$$

Que es el problema dual asociado al inicial. Esta idea inspiró en los años 50 muchos trabajos relacionados a la teoría del Equilibrio General de la economía. [5]

3.4.1 Algoritmo simplex-dual

Supongamos que tenemos el siguiente cuadro *dual factible*, es decir, **los costos reducidos son positivos**:

0	$\overline{c}_n^t$	$-\overline{z}$
		$\overline{b_1}$
I	$\tilde{\text{N}}$	$\vdots$
		$\overline{b_n}$

Donde,

- $\overline{c}_n^t \geq 0$ (condición de dual factibilidad).
- $\overline{c}_n^t = c_n^t - c_b^t\widetilde{N}$
- $\widetilde{N} = B^{-1}N$
- $\overline{b} = B^{-1}b$

[5] Un famoso trabajo en este ámbito es "The Coefficient of Resource Utilization"de Gerard Debreu. (Disponible en http://cowles.econ.yale.edu/P/cp/p00a/p0045.pdf)

Los siguientes pasos resumen el Algoritmo Simplex Dual:

(1) Si $\overline{b} \geq 0 \;\; i = 1, .., m$, entonces la solución en curso es óptima. Si no, ir a (2).

(2) Elegir $\overline{b}_r = \min\limits_{\overline{b}_i < 0}\{\overline{b}_i\}$. En realidad, se puede elegir cualquier $\overline{b}_i < 0$.
Si $\overline{a}_{rj} \geq 0 \;\; \forall j = 1, .., n$ entonces el problema *dual es no acotado*, es decir, el problema primal es infactible.
Si algún $\overline{a}_{rj} < 0$ pasar a (3).

(3) Elegir la columna s tal que:

$$\frac{\overline{c}_s}{\overline{a}_{rs}} = \max_{\overline{a}_{rj} < 0}\{\frac{\overline{c}_j}{\overline{a}_{rj}}\}$$

e ir a (4).

(4) Pivotear en la posición (r,s).

3.5 Introducción al análisis post-optimal -Análisis de sensibilidad-

Muchas veces, una vez resuelto el problema lineal:

$$\begin{array}{ll} \min & c^t x \\ & Ax = b \\ & x \geq 0 \end{array}$$

se desea examinar el comportamiento de la solución si se modifica alguno de sus parámetros. Estos cambios pueden ser:

- Variación en los coeficientes de la función objetivo.
- Variación en el vector de recursos.
- Introducción de una nueva variable.
- Introducción de una nueva restricción.

Puede suceder que nuestro problema sea muy complejo y no se desee resolver completamente de nuevo parta analizar estos cambios, por ejemplo por problemas de tiempo. La siguiente sección examina los mencionados casos.

3.5.1 Variación en los coeficientes de la función objetivo

Consideremos el siguiente problema:

$$\begin{array}{lll} \min & -20x_1 - 16x_2 - 12x_3 & \\ & x_1 & \leq 400 \\ & 2x_1 + x_2 + x_3 & \leq 1000 \\ & 2x_1 + 2x_2 + x_3 & \leq 1600 \\ & x_i & \geq 0 \end{array}$$

Una solución inicial es $x_0 = (0,0,0)^t$ (notar que es un punto extremo del poliedro). El cuadro simplex inicial es:

-20	-16	-12	0	0	0	0
1	0	0	1	0	0	400
2	1	1	0	1	0	1000
2	2	1	0	0	1	1600

Luego de pivotear, se llega al siguiente cuadro final: Por lo tanto, la solución es:

4	0	0	0	8	4	14400
2	0	1	0	2	-1	400
0	1	0	0	-1	1	600
1	0	0	1	0	0	400

$$x^* = \begin{pmatrix} 0 \\ 600 \\ 400 \end{pmatrix} \qquad x_{holg} = \begin{pmatrix} 400 \\ 0 \\ 0 \end{pmatrix}$$

La base está compuesta por $[x_3, x_2, x_4]$ y el valor óptimo es -14400. Qué sucede si nos informan que el coeficiente c_1 vale -30 en lugar de -20?.
Examinemos los costos reducidos (los demás elementos del cuadro, es decir $B^{-1}N, B^{-1}b, c_B^t B^{-1}b$ no sufen alteraciones, dado que c_1 no participa en ellos).

$\bar{c}_5 = 8$ no se modifica.

$\bar{c}_6 = 4$ no se modifica.

$$\overline{c}_1 = c_1 - c_B^t B^{-1} a_1 = -30 - \begin{pmatrix} -12 & -16 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = -6$$

Por lo cual el cuadro final ha cambiado, transformándose en:

-6	0	0	0	8	4	14400
2	0	1	0	2	-1	400
0	1	0	0	-1	1	600
1	0	0	1	0	0	400

que no es óptimo, por lo tanto al pivotear se llega a:

0	0	3	0	14	1	15600
1	0	$\frac{1}{2}$	0	1	$-\frac{1}{2}$	200
0	1	0	0	-1	1	600
0	0	$-\frac{1}{2}$	1	-0	$\frac{1}{2}$	200

Y la nueva solución es:

$$x^* = \begin{pmatrix} 200 \\ 600 \\ 0 \end{pmatrix} \qquad x_{holg} = \begin{pmatrix} 200 \\ 0 \\ 0 \end{pmatrix}$$

La base cambió a $[x_1, x_2, x_4]$ y el valor mínimo cayó a -15600.
Qué sucede si $c_1 = -20$ se modifica, en un contexto un poco más general a $-20 + \theta$? Retomemos el asunto:

$$\overline{c}_1 = c_1 - c_B^t B^{-1} a_1 = (-20 + \theta) - \begin{pmatrix} -12 & -16 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} = 4 + \theta$$

Que es positivo cuando $\theta \geq -4$.
Es decir, el rango para el coeficiente c_1 con el que la base óptima $[x_3, x_2, x_4]$ no cambie es $c_1 \geq -24$.
Veamos otro caso. Supoongamos ahora que el coeficiente perturbado es $c_2 = -16$ y pasa a ser $-16 + \gamma$. El vector de costos queda:

$$c = \begin{pmatrix} -20 \\ -16+\gamma \\ -12 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Recordemos que las tres primeras componentes de este vector corresponden a los costos estructurales y las últimas tres corresponden a las variables de holgura, y por lo tanto son cero. Examinemos los costos reducidos:
$\overline{c}_N^t = c_N^t - c_b^t B^{-1} N$
$= \begin{pmatrix} -20 & 0 & 0 \end{pmatrix} - \begin{pmatrix} -12 & -16+\gamma & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$
$= \begin{pmatrix} 4 & 8+\gamma & 4-\gamma \end{pmatrix}$
Estos costos reducidos de las variables no básicas son positivos, es decir preservan la optimalidad, cuando:

$$-8 \leq \gamma \leq 4$$

O sea, la base $[x_2, x_3, x_4]$ no cambia si:

$$-24 \leq c_2 \leq -12$$

Finalmente, en general si el vector c cambia a $\widetilde{c}$ se debe evaluar $\overline{\widetilde{c}}_N^t = \widetilde{c}_N^t - \widetilde{c}_B^t B^{-1} N$ y decidir:

- si $\overline{\widetilde{c}}_N^t \geq 0$, la base óptima no cambia y sólo hay que reevaluar $\widetilde{c}_B^t B^{-1} b = z^*$.
- si $\overline{\widetilde{c}}_N^t \ngeq 0$, se itera con algoritmo simplex.

3.5.2 Variación en el vector de recursos (lado derecho).

Tomemos el mismo ejemplo y supongamos que el vector b cambia a $\widetilde{b}$. La base óptima para b es $[x_3, x_2, x_4]$ entonces se tiene que:

$$B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 0 \end{pmatrix} \text{ y } B^{-1} = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Notemos que la matriz B^{-1} es parte del cuadro final. Se tiene que:

- Si $B^{-1}\widetilde{b} \geq 0$, la solución en curso *aún es óptima.*
- Si $B^{-1}\widetilde{b} \ngeq 0$, la solución no es factible (primal), pero los costos reducidos no han sufrido cambios, luego el cuadro final presenta una solución primal-infactible y dual-factible. Entonces se debe iterar con el algoritmo simplex dual.

Veamoslo con un ejemplo:
Supongamos que $\widetilde{b} = \begin{pmatrix} 100 \\ 1000 \\ 1600 \end{pmatrix}$ por lo tanto:

$$B^{-1}\widetilde{b} = \begin{pmatrix} 400 \\ 600 \\ 100 \end{pmatrix} \geq 0$$

Así, la base óptima no cambia, pero:

$$x^* = \begin{pmatrix} x_1^* \\ x_2^* \\ x_3^* \\ x_4^* \\ x_5^* \\ x_6^* \end{pmatrix} = \begin{pmatrix} 0 \\ 600 \\ 400 \\ 100 \\ 0 \\ 0 \end{pmatrix}$$

Además, $z^* = c_B^t B^{-1}\widetilde{b} = \begin{pmatrix} -12 & -16 & 0 \end{pmatrix} \begin{pmatrix} 400 \\ 600 \\ 100 \end{pmatrix} = -14400.$

Una pregunta interesante es Cuál es el rango para $\widetilde{b}$, de modo que la base óptima no se modifique?. Para ello, basta calcular:

$$B^{-1}\widetilde{b} = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \widetilde{b}_1 \\ \widetilde{b}_2 \\ \widetilde{b}_3 \end{pmatrix} = \begin{pmatrix} 2\widetilde{b}_2 - \widetilde{b}_3 \\ -\widetilde{b}_2 + \widetilde{b}_3 \\ \widetilde{b}_1 \end{pmatrix} \geq 0$$

De aquí se deduce que para que se cumpla la condición de optimalidad se debe tener:

$$\widetilde{b}_1 \geq 0 \qquad \widetilde{b}_2 \leq \widetilde{b}_3 \leq 2\widetilde{b}_2$$

(Notemos que que los datos originales satisfacen estas condiciones)
Por ejemplo, si b_1 y b_3 quedan fijos en sus valore originales (400 y 1600, respectivamente).

3.5.3 Introducción de una actividad (o variable)

Supongamos que, en el ejemplo, se introduce la variable x_4 con costo $c_4 = -10$ y coeficientes $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ en la matriz, es decir, el problema se transforma en:

$$\begin{array}{lll} \min & -20x_1 - 16x_2 - 12x_3 - 10x_4 & \\ & x_1 + x_4 & \leq 400 \\ & 2x_1 + x_2 + x_3 & \leq 1000 \\ & 2x_1 + 2x_2 + x_3 + x_4 & \leq 1600 \\ & x_i & \geq 0 \end{array}$$

Y el cuadro inicial es:

-20	-16	-12	-10	0	0	0	0
1	0	0	1	1	0	0	400
2	1	1	0	0	1	0	1000
2	2	1	1	0	0	1	1600

Si se realiza la misma secuencia de iteraciones, el cuadro final es:

4	0	0	-6	0	8	4	14400
2	0	1	-1	0	2	-1	400
0	1	0	1	0	-1	1	600
1	0	0	1	1	0	0	400

Aquí, conviene observar que: $\overline{c}_4 = c_4 - c_B^t B^{-1} a_4$, en que $c_4 = -10$ y $a_4 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. Así:

$$\overline{c}_4 = -10 - \begin{pmatrix} -12 & -16 & 0 \end{pmatrix} \begin{pmatrix} 0 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = -6$$

Además:

$$B^{-1} a_4 = \begin{pmatrix} 0 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

Por lo tanto, basta agregar la columna correspondiente a la nueva variable en el cuadro final original. Esta columna es:

$$\begin{pmatrix} c_4 - c_B^t B^{-1} a_4 \\ B^{-1} a_4 \end{pmatrix}$$

en que c_4 es el costo de la variable nueva y a_4 es el vector columna de dicha variable, en la matriz de restricciones.
En este caso, la nueva variable tiene costo reducido $-6 < 0$, y por lo tanto puede entrar a la base. Así el cuadro queda:

10	0	0	0	6	8	4	16800
3	0	1	0	1	2	-1	800
-1	1	0	0	-1	-1	1	200
1	0	0	1	1	0	0	400

La nueva variable permite disminuir el costo total desde -14400 a -16800 siendo la solución final:

$$x^* = \begin{pmatrix} 0 \\ 200 \\ 800 \\ 400 \end{pmatrix}$$

Observación 3.5.1 *Podría la nueva variable producir no acotamiento? La respuesta es sí. La condición para ello es que* $B^{-1}a_4 \leq 0$.

En el ejemplo, nos interesa calcular para qué valores del nuevo costo c_4 la variable introducida es irrelevante en el óptimo (es decir, no pertenece a la base óptima).
La condición para ello es que :

$$\overline{c}_4 = c_4 - c_B^t B^{-1} a_4 \geq 0 \qquad \Leftrightarrow \qquad c_4 \geq c_B^t B^{-1} a_4 = -4$$

3.5.4 Introducción de una nueva restricción

Estamos tratando el problema:

$$\begin{array}{ll} \min & c^t x \\ & \begin{array}{ll} Ax & = b \\ x_i & \geq 0 \end{array} \end{array}$$

Cuyo cuadro óptimo, salvo reorden, está dado por:

0	$c_N^t - c_B^t B^{-1}N$	$c_b^t B^{-1}b$
I	$B^{-1}N$	$B^{-1}b$

Supongamos que se agrega la restricción:

$$d^t x \leq d_0$$

En que $d \in I\!R^n$ y $d_0 \in I\!R$. Es decir, agregamos:

$$d^t x + x_{n+1} = d_0$$

Con x_{n+1} una variable de holgura. Así, el problema original se puede plantear como:

$$\begin{array}{rll} \min & c_B^t x_B + c_N^t x_N + 0x_{n+1} & \\ & Bx_B + Nx_N + \overrightarrow{0}\, x_{n+1} & = b \\ & d_B x_B + d_N x_N + x_{n+1} & = d_0 \\ & x_B, x_N, x_{n+1} & \geq 0 \end{array}$$

en que $d = \begin{pmatrix} d_B \\ d_N \end{pmatrix}$. O bien, el nuevo problema es:

$$\begin{array}{rll} \min & \begin{pmatrix} c^t & 0 \end{pmatrix} \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} & \\ & \begin{bmatrix} A & 0 \\ d^t & 1 \end{bmatrix} \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} & = \begin{pmatrix} b \\ d_0 \end{pmatrix} \\ & x, x_{n+1} & \geq 0 \end{array}$$

Agregemos x_{n+1} a la base, es decir propongamos:

$$\widetilde{B} = \begin{bmatrix} B & 0 \\ d_B^t & 1 \end{bmatrix}$$

Cuya inversa es:

$$\widetilde{B}^{-1} = \begin{bmatrix} B^{-1} & 0 \\ -d_B^t B^{-1} & 1 \end{bmatrix}$$

Cómo se modifica el cuadro final? Veamos término a término:

- $\overline{c}_N^t = c_N^t - \begin{pmatrix} c_B^t & 0 \end{pmatrix} \widetilde{B}^{-1} \begin{pmatrix} N \\ d_N^t \end{pmatrix} = c_N^t - c_B^t B^{-1} N$
  Así, los costos reducidos no cambian.
- $\widetilde{B}^{-1} \begin{pmatrix} b \\ d_0 \end{pmatrix} = \begin{bmatrix} B^{-1} & 0 \\ -d_B^t B^{-1} & 1 \end{bmatrix} \begin{pmatrix} b \\ d_0 \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ -d_B^t B^{-1} b + d_0 \end{pmatrix}$
- $\begin{pmatrix} c_B^t & 0 \end{pmatrix} \widetilde{B}^{-1} \begin{pmatrix} b \\ d_0 \end{pmatrix} = c_B^t B^{-1} b.$

Por lo tanto, el cuadro es:

0	$c_N^t - c_B^t B^{-1} N$	0	$c_b^t B^{-1} b$
I	$B^{-1} N$	0	$B^{-1} b$
0	$d_N^t - d_B^t B^{-1} N$	1	$d_0 - d_B^t B^{-1} b$

Luego,

- Si $d_0 - d_B^t B^{-1} b \geq 0$, la solución propuesta en los datos originales sigue siendo óptima, sólo que la holgura x_{n+1} entra a la base con valor $d_0 - d_B^t B^{-1} b$.
- Si $d_0 - d_B^t B^{-1} b < 0$, la solución propuesta *no es factible*, pues la holgura x_{n+1} toma un valor negativo. Iterar con algoritmo Simplex Dual, pivoteando sobre la fila agregada. En este caso, si $d_0 - d_B^t B^{-1} N \geq 0$, el problema (primal) es infactible, dado que el dual es no acotado.

Retomemos el problema del inicio:

$$\begin{array}{lll} \min & -20x_1 - 16x_2 - 12x_3 & \\ & x_1 + x_4 & = 400 \\ & 2x_1 + x_2 + x_3 + x_5 & = 1000 \\ & 2x_1 + 2x_2 + x_3 + x_6 & = 1600 \\ & x_i & \geq 0 \end{array}$$

El cuadro óptimo es:

4	0	0	0	8	4	14400
2	0	1	0	2	-1	400
0	1	0	0	-1	1	600
1	0	0	1	0	0	400

Si se agrega la restricción

$$x_1 + x_2 + x_3 \leq 800$$

Es decir,

- $d_0 = 800$
- $d^t = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$

Entonces:

- $d_B^t = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \qquad [x_3, x_2, x_4]$
- $d_N^t = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \qquad [x_1, x_5, x_6]$
- $d_N^t - d_B^t B^{-1} N = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{bmatrix} 2 & 2 & -1 \\ 0 & -1 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{pmatrix} -1 & -1 & 0 \end{pmatrix}$
- $d_0 - d_B^t B^{-1} b = 800 - \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} 400 \\ 600 \\ 400 \end{pmatrix} = -200$

Ahora completamos el cuadro final:

4	0	0	0	8	4	0	14400
2	0	1	0	2	-1	0	400
0	1	0	0	-1	1	0	600
1	0	0	1	0	0	0	400
-1	0	0	0	-1	0	1	-200

0	0	0	0	4	4	4	14000
0	0	1	0	0	-1	2	0
0	1	0	0	-1	1	0	600
0	0	0	1	-1	0	1	200
1	0	0	0	1	0	-1	200

Al pivotear con Simplex Dual, la última variable sale de la base y el cuadro óptimo queda:
La nueva solución óptima es:

$$x^* = \begin{pmatrix} 200 \\ 600 \\ 0 \\ 200 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y $z^* = -14000$.

Capítulo 4

Descripción de problemas típicos. Modelo de flujo en redes

Cabe destacar que esta área de la programación lineal es importante ya que existen varios problemas de estructura espacial que son parte de la programacion lineal. Estos problemas son de gran interes pues, curiosamente, muchos de ellos se formularon originalmente antes del desarrollo general de la programacion lineal y continuan presentandose en diversas aplicaciones.

En éste capítulo se estudiará el problema de flujo al costo mínimo, el cual se divide en cuatro problemas menos generales:

a) Problema de transporte

b) Problema de asignación

c) Problema de flujo máximo

d) Problema del camino más corto

Definición 4.0.1 *Un **grafo** es un par $(\mathcal{N}, \mathcal{A})$, donde $\mathcal{N}$ es un conjunto finito y $(\mathcal{A} \subseteq \mathcal{N} \times \mathcal{N})$. A los elemnetos en $\mathcal{N}$ se les llama **nodos** y a los pares ordenados en $\mathcal{A}$ se les llama **arcos**.*

En el grafo de la figura la cantidad entre paréntesis (b) representa la oferta en cada nodo (si $b \geq 0$ el nodo **ofrece** la cantidad b, si $b \leq 0$ el nodo **demanda** la cantidad b)

La notación (u, c) indica la capacidad del arco (u) y el costo unitario del arco (c).

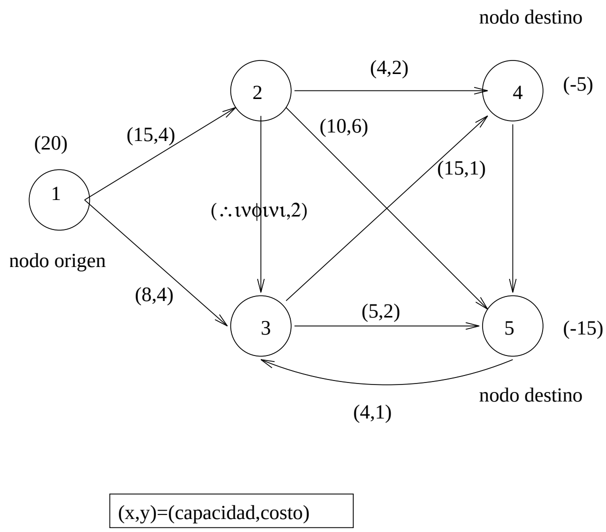

Figura 4.1: Ejemplo de un grafo

El problema es encontrar un flujo factible, de costo mínimo. Si x_{ij} es la cantidad enviada de i a j, entonces el problema es:

$$\begin{array}{lll} \min & \displaystyle\sum_{(i,j)\in A} c_{ij}x_{ij} & \\ \text{s.a.} & \displaystyle\sum_{j/(i,j)\in A} x_{ij} - \sum_{k/(k,i)\in A} x_{ki} = b_i & \forall i \in \mathcal{N} \\ & 0 \leq x_{ij} \leq u_{ij} & \forall (i,j) \in \mathcal{A} \end{array}$$

La primera restricción dice que la oferta en el nodo i es igual a lo que entrega menos lo que recibe y la segunda, que el flujo sobre un arco debe estar entre las capacidades del mismo.

Los datos de un grafo se pueden resumir en una matriz S, cuyas filas son los nodos del grafo y cuyas columnas son los arcos, de manera que $\forall n \in \mathcal{N}, m \in \mathcal{A}$.

$$S_{nm} = \begin{cases} 1 & \text{si el arco } m \text{ sale del nodo } n \\ -1 & \text{si el arco } m \text{ llega al nodo } n \\ 0 & \text{si no} \end{cases}$$

Notemos que cada arco aparece sólo en dos retricciones ya que un arco participa en solo dos nodos, uno que indica la cantidad entrante a éste, y otro que indica la cantidad saliente. Cabe destacar que la matriz resultante será de rango incompleto (n-1), esto es, existirá una ecuación redundante, lo cual se ve al sumar las columnas ya que como en cada columna sólo aparece un 1 y un -1 (los demás son sólo ceros) éstos se anulan, lo que hace que la suma por columnas sea cero. Siempre actuamos bajo el supuesto que $\sum b_i = 0$, supuesto que ayuda a que el sistema sea factible.

Así , el problema puede escribirse de la forma:

$$\begin{array}{ll} \min & \sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij} \\ \text{s.a.} & S\,x = b \\ & 0 \leq x_{ij} \leq u_{ij} \quad \forall (i,j) \in \mathcal{A} \end{array}$$

Los datos del problema del ejemplo de la figura pueden entonces resumirse en la siguiente tabla:

Tabla 4.1:

	x_{12}	x_{13}	x_{23}	x_{24}	x_{25}	x_{34}	x_{35}	x_{45}	x_{53}	
nodo/costo	4	4	2	2	6	1	3	2	1	oferta
1	1	1	0	0	0	0	0	0	0	20
2	-1	0	1	1	1	0	0	0	0	0
3	0	-1	-1	0	0	1	1	0	-1	0
4	0	0	0	-1	0	-1	0	1	0	-5
5	0	0	0	0	-1	0	-1	-1	1	-15
capacidad	15	8	∞	4	10	15	15	∞	4	

Ahora entraremos a estudiar los cuatro problemas antes mencionados, para lugo solucionar dos de ellos.

4.1 Motivación: Problema de Asignación

Supongamos que somos gerente de algún prestigioso supermercado que consta de 50 cajas, éstas, diferenciadas según el tipo de cliente (embarazadas, facturas, tercera edad, etc.). Como gerente, uno pretende maximizar la eficiencia de su supermercado para atraer clientela, uno de nuestros recursos es la disponibilidad de 50 cajeras. Nuestro problema es como asignar estas 50 cajeras en las 50 cajas de la mejor manera posible. Si hicieramos ésta asignacion probando a cada cajera en nuestras distintas cajas tardaríamos más tiempo del cuál disponemos (pues son 50! configuraciones), aquí es donde entra a jugar un papel importante la programación lineal, y en particular, el problema de asignación, que gracias a las distintas restricciones de las cajeras (experiencia, habilidad, carácter, etc.), podemos obtener una solución óptima.

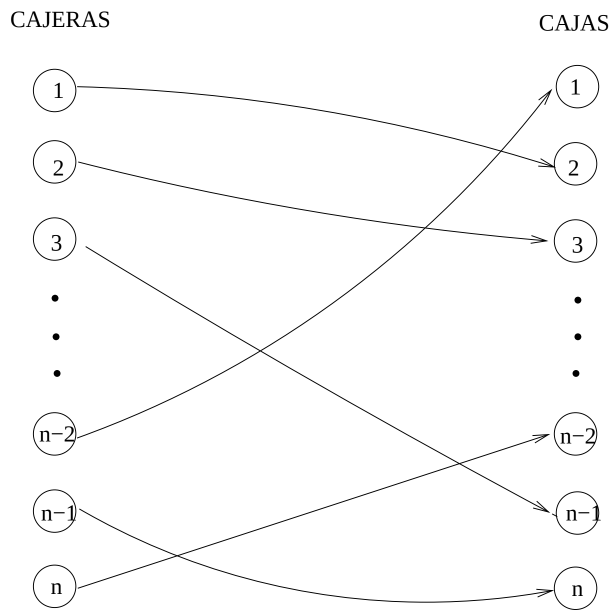

Figura 4.2: Problema de Asignación

Las variables en éste caso son:

$$x_{i,j} = \begin{cases} 1 & \text{si al nodo } i \text{ le corresponde el nodo } j \\ 0 & \text{si no} \end{cases}$$

Este problema se escribe:

$$\begin{aligned} \max \quad & \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}x_{ij} \\ \text{s.a.} \quad & \sum_{i=1}^{n} x_{ij} = 1 \quad \forall j \quad \text{(a cada caja una sola cajera)} \\ & \sum_{j=1}^{n} x_{ij} = 1 \quad \forall i \quad \text{(a cada cajera una sola caja)} \\ & x_{ij} \in \{0,1\} \end{aligned}$$

Notemos que $\sum\limits_{i=1}^{n} x_{ij} = 1$ pues a cada nodo de la izquierda le corresponde un único nodo de la derecha (análogo para le segunda restricción).

4.2 Problema de Transporte

Figura 4.3: Ejemplo de un problema de transporte

Consideremos un grafo con un conjunto de m nodos de partida, con ofertas $a_i \geq 0 \;\; i = 1, ..., m$ y n nodos de llegada con demandas $b_j \geq 0 \;\; j = 1, ..., n$.

Cada arco tiene asociado un costo unitario de transporte c_{ij}.

Supongamos, por ahora, que $\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$, aunque lo natural es que $\sum_{i=1}^{m} a_i \geq 0 \sum_{j=1}^{n} b_j$ pues la oferta siempre es mayor que la demanda. este supuesto se hace para que el problema de transporte tenga sentido.

Se conoce el problema de transporte como el de minimización de los costos de transporte por los arcos del grafo, de manera de satisfacer la demanda en cada nodo de llegada, sujeto a restricciones en la oferta de cada nodo de partida.

Podemos notar que el problema de asignación es un caso particular del problema de transporte, donde cada oferta y cada demanda consta de una sola unidad.

El problema de transporte se escribe:

$$\begin{aligned}
\min \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} \\
\text{s.a.} \quad & \sum_{j=1}^{n} x_{ij} = a_i \quad i = 1, ..., m \quad (oferta) \\
& \sum_{i=1}^{m} x_{ij} = b_j \quad j = 1, ..., n \quad (demanda) \\
& 0 \leq x_{ij} \qquad \forall i, j
\end{aligned}$$

Las restricciones quedan definidas de esa forma ya que

$$\sum_{j/(i,j)\in A} x_{ij} - \sum_{k/(i,j)\in A} x_{ki} = a_i$$

pues en este caso

$$\sum_{k/(i,j)\in A} x_{ki} = 0$$

y en el caso de la demanda se tiene que

$$\sum_{j/(i,j)\in A} x_{ij} - \sum_{k/(i,j)\in A} x_{ki} = -b_j$$

dado que

$$\sum_{j/(i,j)\in A} x_{ij} = 0$$

(pues "sale - entra = demanda ").

4.3 Problema de Flujo Máximo

Este problema es el de determinar el flujo maximal posible de un nodo origen o fuente (s) dado a un nodo destino o sumidero (t) con restricciones de capacidad en los arcos.

Si denotamos v al flujo correspondiente a transportar la cantidad final en t, desde t a s, notaremos que $max\ v$ es equivalente a maximizar el flujo total transportado por el resto del grafo ya que dado que todo lo que sale de s llega t, entonces en el sistema se mantiene un equilibrio que permite que los problemas sean análogos.

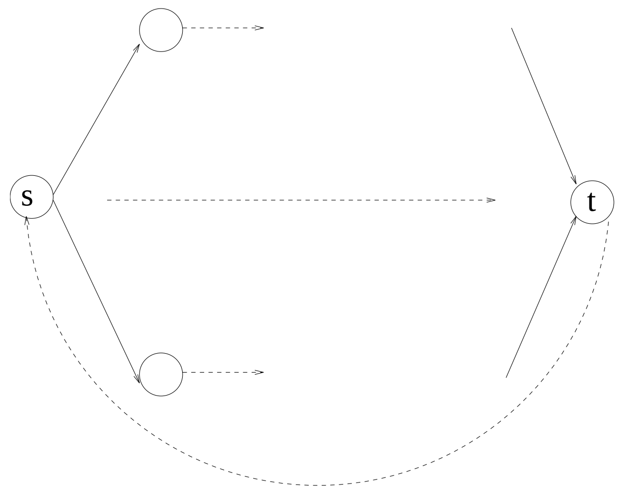

Figura 4.4: Problema de flujo máximo

Luego, el problema se escribe de la siguiente manera

$$
\begin{array}{lll}
\max & v & \\
\text{s.a.} & \sum\limits_j x_{sj} - \sum\limits_k x_{ks} - v = 0 & \text{sale de } s \\
& \sum\limits_j x_{tj} - \sum\limits_k x_{kt} + v = 0 & \text{llega a } s \\
& \sum\limits_{j \in N} x_{ij} - \sum\limits_{k \in N} x_{ki} \quad = 0 & i \neq s,\, t \text{ (balance entre nodos intermedios)} \\
& 0 \leq x_{ij} \leq u_{ij} & \forall (i,j) \in \mathcal{A} \cup \{(t,s)\}
\end{array}
$$

notemos que el problema original es de la forma $min(-v)$ donde el problema es de flujo de costo mínimo, en el cual el vector c es de la forma $c^t = (-1, 0, .., 0)$.

4.4 Problema de Flujo a Costo Mínimo

Este problema generaliza algo el problema de transporte. El problema es minimizar los costos de transporte por los arcos de una red, que tienen asociados capacidades técnicas.

El problema de flujo a costo mínimo se plantea como sigue:

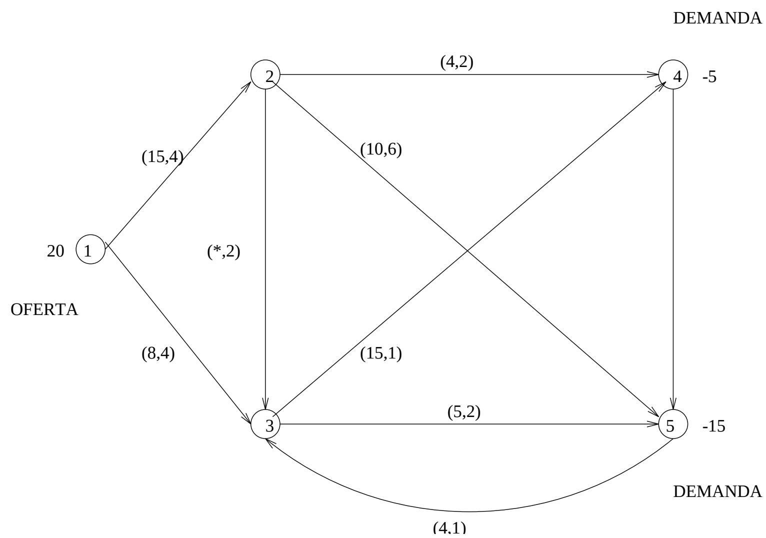

Figura 4.5: Problema de flujo a costo mínimo

$$
\begin{array}{llll}
\min & \sum_{i=1}^{m}\sum_{j=1}^{n} c_{ij}x_{ij} & & \\
\text{s.a.} & \sum_{j=1}^{n} x_{sj} = a_s & s \in S & \text{(conjunto de nodos iniciales)} \\
& \sum_{i=1}^{m} x_{jt} = b_t & t \in T & \text{(conjunto de nodos terminales)} \\
& \sum_{j\in N} x_{ij} - \sum_{k\in N} x_{ki} = 0 & i \notin S \cup T & \text{(balance entre nodos intermedios)} \\
& 0 \leq x_{ij} \leq u_{ij} & \forall i,j &
\end{array}
$$

4.5 Problema del Camino más corto

Este problema tiene como objetivo encontar el camino más corto entre el nodo s y el nodo t en un grafo dado, ésto, encontrando una secuencia de arcos dirigidos y adyacentes entre s y t. La longitud del arco puede ser expresada en términos de costo, tiempo, distancia, etc., es ésto lo que se debe minimizar para solucionar tal problema. El problema del camino mas corto se ve caracterizado por el siguiente grafico:

El problema se escribe

Figura 4.6: Ejemplo problema camino más corto

$$
\begin{array}{llll}
\min & \sum_i \sum_j c_{ij} x_{ij} & & \\
\text{s.a.} & \sum_{j/(i,j)\in A} x_{sj} = 1 & & \text{(ofrece una unidad)} \\
& -\sum_{k\in N} x_{kt} = -1 & & \text{(demanda una unidad)} \\
& \sum_{j\in N} x_{ij} - \sum_{k\in N} x_{ki} = 0 & i \neq s, t & \text{(balance entre nodos intermedios)} \\
& x_{ij} \in \{0,1\} & \forall (i,j) \in \mathcal{A} &
\end{array}
$$

A continuación resolveremos dos de los cuatro problemas que se plantearon anteriormente. El primero será el problema de Transporte (que como dijimos anterirmente generaliza al problema de Asignación), y luego será resuelto el problema de Flujo a costo Mínimo.

4.6 Solución al Problema de Transporte

Solución básica factible: Fase I El proceso de saturación es, en realidad, la búsqueda de una solucion básica factible del sistema, es decir, la determinacion de una base factible.

Definición 4.6.1 *Un* ***árbol*** *es un grafo conexo, ie, existe una cadena entre dos nodos cualesquiera (todos los nodos están conectados por una secuencia de arcos, sin considerar la orientación), que no contiene ciclos, es decir, partiendo de un nodo no se puede volver a él por una secuencia de arcos adyacentes (sin importar la direccion de los arcos).*

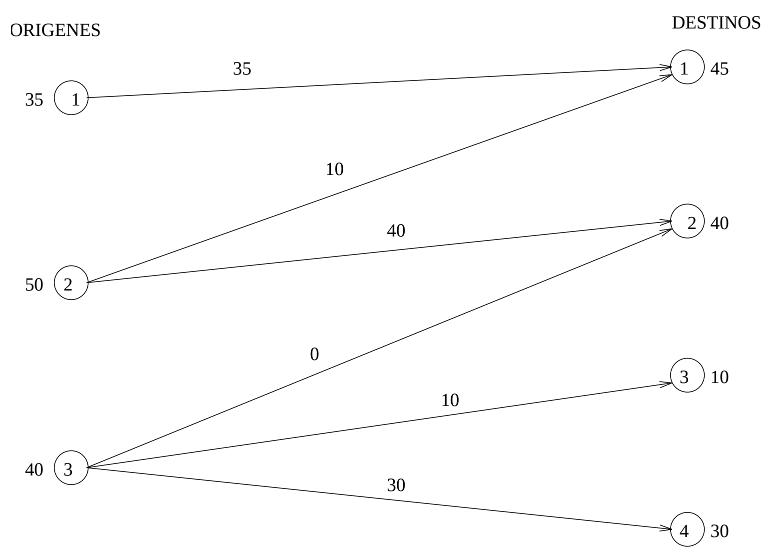

Figura 4.7: Procedimiento de saturación

Procedimiento de saturación: Sea $m+n$ el número de arcos del grafo. En la solución factible deberan haber $m+n-1$ arcos con flujos positivos en la solución (los demás están en cero, es decir, $mn-(m+n-1)$ arcos nulos).

El método de saturación empieza cuando se satura el primer arco (elejido arbitrariamente), esto es, se elije un arco al cual se le entrega el maximo flujo posible, satisfaciendo así la demanda de los nodos demandantes, luego se prosigue de la misma manera con el resto de los arcos, hasta satisfacer la demanda de todos los nodos.

El sistema de ecuaciones del problema de transporte tiene $m+n$ ecuaciones, pero recordemos que una es redundante. Luego, cuando saturo en orden arbitario, la solución propuesta es básica.

Las bases son **árboles**, si se formaran ciclos, se traicionaría la idea de base y la solución sería múltiple.

Procedimiento de saturación a Costo Mínimo Ahora saturemos guiados por el costo mínimo, es decir, comenzamos saturando desde el arco que posee el menor costo al de mayor costo. Este procedimiento produce un árbol (solución básica factible).

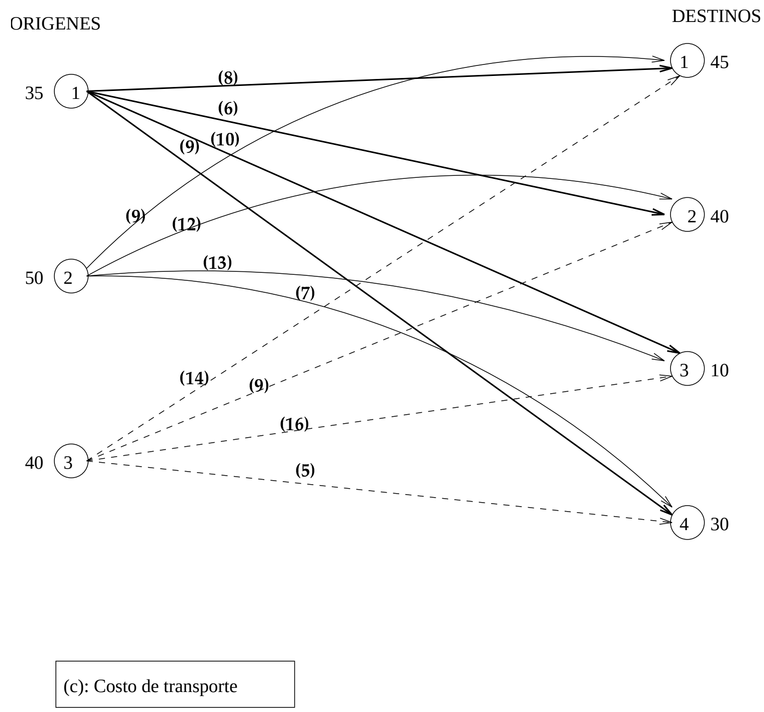

Figura 4.8: Procedimiento de saturación a costo mínimo

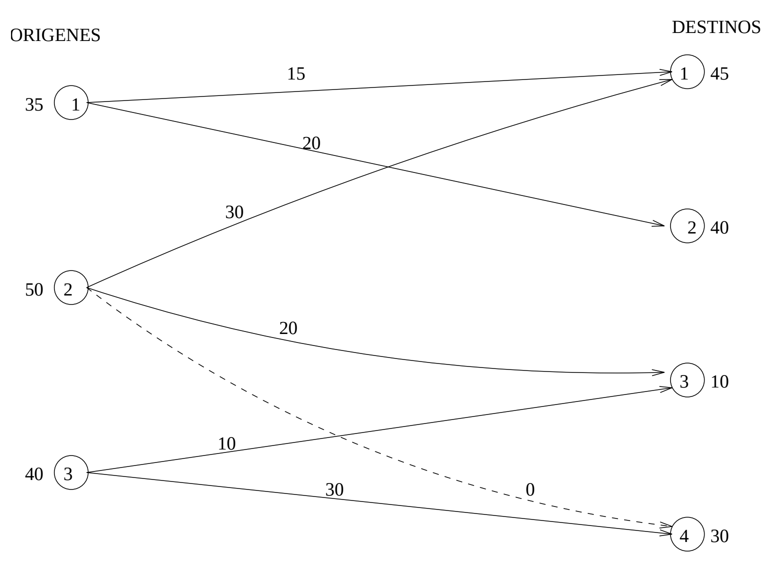

Figura 4.9: Base factible

En el caso de que se sepa cuáles son los arcos básicos, ¿cómo determinar los flujos?

Procedimiento:

(1) Elegir un nodo final del árbol (nodo al cual llega un solo arco, o del cual emerge un solo arco) y saturar el arco correspondiente.

(2) Eliminar el nodo y su arco saturado y repetir el procedimiento en el subgrafo resultante.

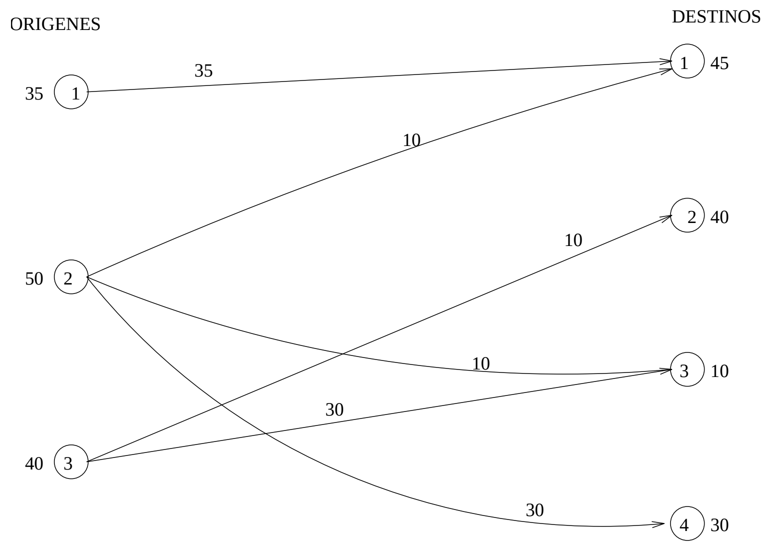

Figura 4.10: Nueva base factible

Si toda la demanda no puede ser cubiereta, la solución es infactible.

Observación 4.6.1 *Si los datos a_i, b_j son enteros, entonces los flujos x_{ij} son enteros (de acuerdo a los procedimientos que hemos descrito), pues se trata de diferencias de números enteros (dados por la oferta y demanda en cuestión). Esto muestra que todos los puntos extremos del problema de transporte con datos enteros, tienen coordenadas enteras.*

4.7 Mejora de una solución extrema: Fase II (simplex)

En esta etapa se supone ya conocida una solucion básica factible, a partir de la cual se construye el problema dual del original y se procede segun se explica a continuacion.

Recordemos que este problema de transporte (P) y su dual (D) son de la forma

$$
\begin{array}{llll}
(P) & \min & c^t\, x & \\
 & \text{s.a.} & S\,x = b & \\
 & x_{ij} \leq 0 & &
\end{array}
\qquad
\begin{array}{llll}
(D) & \max & b^t\, y & \\
 & s.a. & S^t\, y & \leq \;\; c
\end{array}
$$

Calculemos el dual del problema de transporte

$$
\begin{array}{lll}
(D) & \max & \sum_i a_i u_i + \sum_j b_j v_j \\
 & \text{s.a.} & u_i + v_j \leq c_{ij} \quad \forall i, j \\
 & & u_i, v_j \in I\!R^n
\end{array}
$$

Supongamos que tenemos una solución básica factible. Los costos reducidos para las variables básicas son

(α) $\bar{c}_{ij} = c_{ij} - u_i - v_j = 0$

Los costos reducidos para las variables no básicas son

(β) $\bar{c}_{ij} = c_{ij} - u_i - v_j$

Donde los u_i son los nodos de oferta y los v_j los de demanda.

El conjunto de ecuaciones (α) representa un sistema de $m + n - 1$ ecuaciones y $m + n$ incógnitas, de rango $m + n - 1$. Entonces podemos fijar una variable dual en un valor arbitrario y usar (α) para encontrar todas las restantes variables duales.

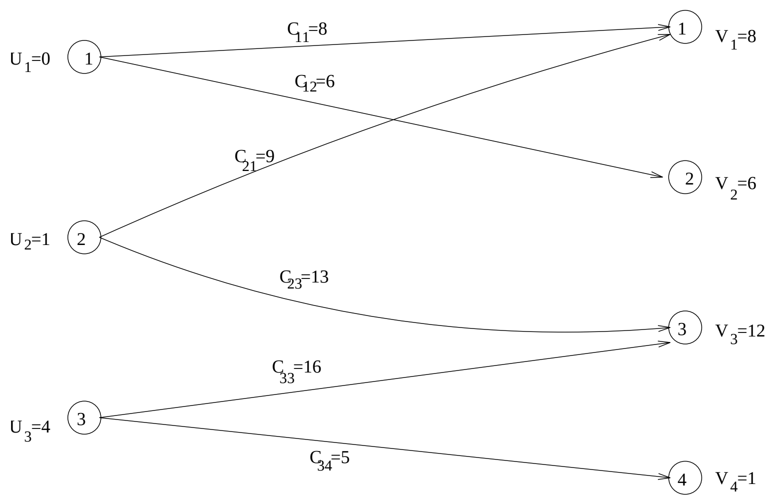

Figura 4.11:

Usemos las ecuaciones (β) para determinar los costos reducidos de los demás arcos (no básicos)

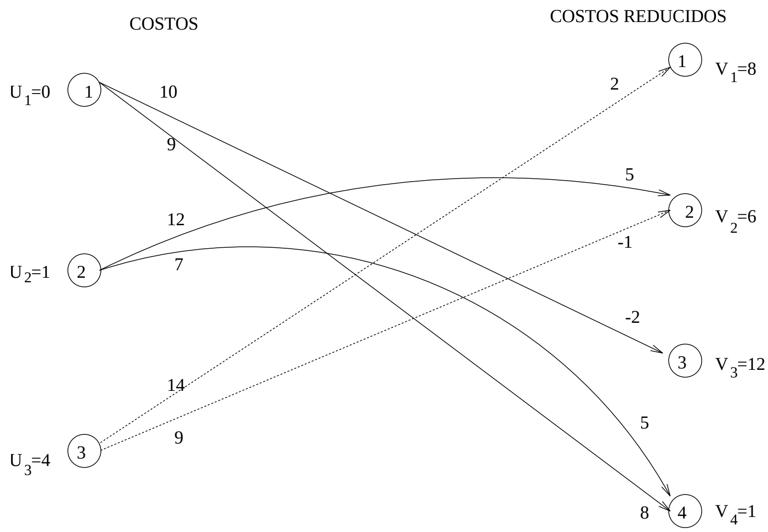

Figura 4.12:

Se ingresa a la base un arco de costo reducido negativo. Si todos los costos reducidos son positivos, llegamos al óptimo.

En el caso del ejemplo, elegimos el arco (1,3)

Si se agrega un arco (i, j) al conjunto de arcos básicos, se genera un ciclo en el grafo.

Se asigna un flujo $\lambda \geq 0$ a ese arco. Los flujos son positivos

$$\left.\begin{array}{rcl} 15-\lambda & \geq & 0 \\ \lambda & \geq & 0 \\ 30+\lambda & \geq & 0 \\ 20-\lambda & \geq & 0 \end{array}\right\} \Longrightarrow 0 \leq \lambda \leq 15$$

Se elije λ de modo de minimizar las diferencias y ese arco sale de la base.

En este caso, $\lambda = 15$ y el arco $(1, 1)$ sale de la base. Notar que las modificaciones sólo afectan al ciclo.

Reiterar hasta que todos los costos reducidos sean positivos.

Ejercicio 4.7.1 *Resolver el problema de transporte para los datos*

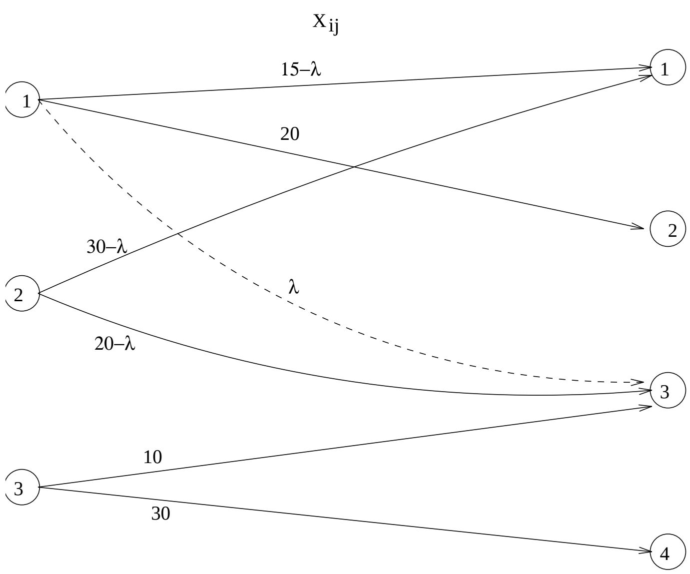

Figura 4.13:

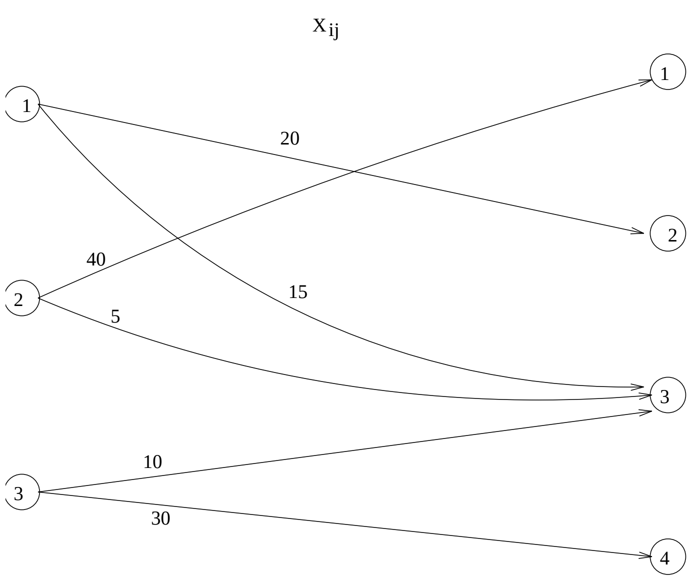

Figura 4.14:

$$a = \begin{bmatrix} 10 \\ 17 \\ 3 \end{bmatrix}, \; b = \begin{bmatrix} 6 \\ 9 \\ 3 \\ 12 \end{bmatrix}, \; c = \begin{bmatrix} 3 & 6 & 8 & 4 \\ 2 & 7 & 3 & 11 \\ 12 & 3 & 4 & 6 \end{bmatrix}$$

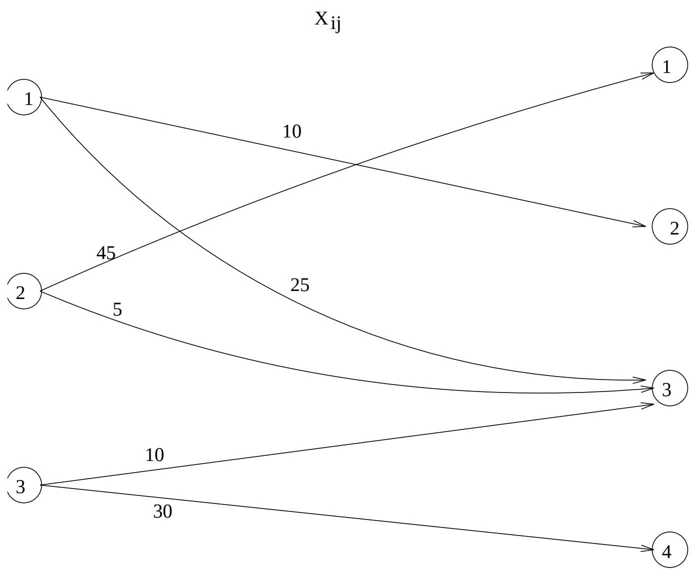

Figura 4.15:

4.7.1 Solución del problema de flujo a costo mínimo

Si n es el número de nodos de la red y m el número de arcos, la matriz de incidencia del grafo es de $n \times m$ y las bases están compuestas por $n-1$ arcos.

No hay un método fácil para encontrar soluciones básicas factibles, por lo tanto, es necesario usar Fase I de simplex.

Consideremos el siguiente ejemplo:

Si tenemos la siguiente base $\{(1,2),(2,4),(2,3),(3,4)\}$. Los arcos $(1,3)$ y $(3,5)$ no son básicas, pero tampoco son nulas (se debe extender el concepto de base).

Extensión del concepto de base Una variable no básica es fijada en alguna de sus cotas. Las variables básicas se determinan resolviendo el sistema de ecuaciones (respetando sus cotas)

Consideremos el problema escrito en la forma canónica:

$$\begin{array}{ll} \min & c^t x \\ \text{s.a.} & Ax = b \\ & l \leq x \leq u \end{array}$$

donde $l, u \in I\!R^n$ y A de rango completo (si es necesario, eliminando filas)

Tomemos la partición $A = [B, N]$ y supongamos que $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ donde $[x_N]_j = l_j \vee u_j$

Se dice que x_B es la base si y solamente si $x_B = B^{-1}(b - Nx_N)$ ($\neq 0$ en este caso).

$[x_B]_j$ debe satisfacer $l_j \leq [x_B]_j \leq u_j$

En el caso del ejemplo, $(2, 3)$ y $(\infty, 2)$ son no básicas y están en su cota inferior, cero en este caso.

Para el problema de transporte, la degenerancia se traduce en que un arco básico esté en alguna de sus cotas.

Veamos el cuadro resumen del problema planteado:

Restricciones: sale-entra=oferta

	x_{12}	x_{13}	x_{23}	x_{24}	x_{25}	x_{34}	x_{35}	x_{45}	x_{53}		$\downarrow$ *var.duales*
1	1	1	0	0	0	0	0	0	0	20	π_1
2	-1	0	1	1	1	0	0	0	0	0	π_2
3	0	-1	-1	0	0	1	1	0	-1	0	π_3
4	0	0	0	-1	0	-1	0	1	0	-5	π_4
5	0	0	0	0	-1	0	-1	-1	1	-15	π_5

La fórmula general de un costo reducido es $\overline{c}^t = c^t - \pi^t N$

$\overline{c}_{ij} = c_{ij} - \pi_i + \pi_j$

Condición de optimalidad:

$$\begin{array}{ll} \overline{c}_{ij} \geq 0 & \text{si } l_{ij} = x_{ij} \text{ (cota inferior)} \\ \overline{c}_{ij} = 0 & \text{si } l_{ij} < x_{ij} < u_{ij} \\ \overline{c}_{ij} \leq 0 & \text{si } x_{ij} = u_{ij} \text{ (cota superior)} \end{array}$$

Las ecuaciones $\overline{c}_{ij} = c_{ij} - \pi_i + \pi_j = 0$ para las variables básicas, permiten determinar los valores de las variables duales. Dado que hay n nodos y $n - 1$ arcos básicos, basta fijar arbitrariamente el valor de una variable dual.

Para el caso del ejemplo, fijemos $\pi_2 = 0$

$c_{12} = \pi_1 - \pi_2 \Rightarrow \pi_1 = 4$

Así se obtiene que $\pi = \begin{bmatrix} 4 \\ 0 \\ -1 \\ -2 \\ -6 \end{bmatrix}$

Luego, se calculan los costos reducidos de los arcos no básicos: $\overline{c}_{ij} = c_{ij} - \pi_i + \pi_j$

Para el ejemplo:

$$\begin{aligned}
\overline{c}_{13} = &\quad c_{13} - (\pi_1 - \pi_3) = 4 - 4 - 1 = & -1 \\
\overline{c}_{23} = &\quad c_{23} - (\pi_2 - \pi_3) = 2 - 0 - 1 = & 1 \\
\overline{c}_{35} = &\quad c_{35} - (\pi_3 - \pi_5) = 3 + 1 - 6 = & -2 \\
\overline{c}_{45} = &\quad c_{45} - (\pi_4 - \pi_5) = 2 + 2 - 6 = & -2 \\
\overline{c}_{53} = &\quad c_{53} - (\pi_5 - \pi_3) = 1 + 6 - 1 = & 6
\end{aligned}$$

Regla de entrada a la base Son candidatos para ingresar a la base los siguientes arcos:

a) Un arco de costo reducido negativo que está en su cota inferior.

b) Un arco de costo reducido positivo que está en su cota superior.

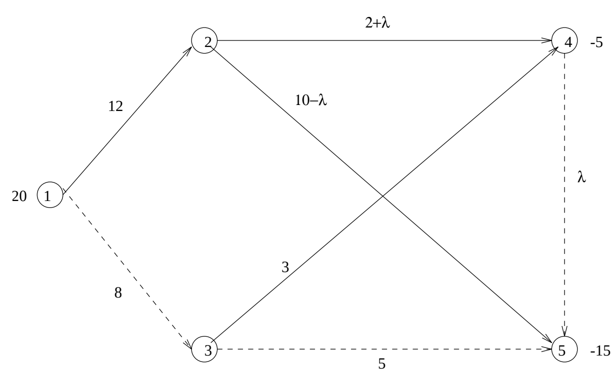

Figura 4.16: primero

Con cada nuevo grafo se tiene una nueva base. Se recalculan las variables duales y los costos reducidos, para llegar al óptimo.

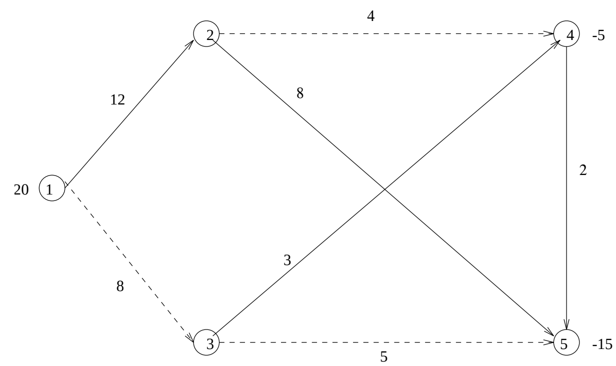

Figura 4.17: segundo

Observación 4.7.1 *Se asume que los arcos que no se dibujan están en cero.*