Matemáticas para la economía y la empresa

M. J. Canós Darós, C. Ivorra Castillo, V. Liern Carrión

Departamento de Economía Financiera y Matemática

Índice General

Cálculo diferencial e integral

Apéndices

Prólogo

Este manual recoge los contenidos de las asignaturas Matemáticas Empresariales y Matemáticas Económico-empresariales (plan 2000) que se imparten en la Universitat de València. En cada tema, se exponen los resultados teóricos necesarios acompañados de ejercicios resueltos a modo de ejemplo y destacando los hechos más relevantes que el alumno debe recordar a la hora de resolver problemas. Se incluye también la demostración de algunos teoremas. Más concretamente, hemos seleccionado aquellas que consideramos que —sin exceder el nivel exigible a los alumnos— pueden ayudarles a familiarizarse con los conceptos que va a manejar. Así mismo, cada tema termina con una colección de ejercicios propuestos.

En todos los temas hemos intentado mostrar la conexión de las técnicas expuestas con la teoría económica y sus aplicaciones a la empresa. Para ello hemos incluido numerosos ejemplos con enunciado económico, los cuales no han de entenderse como aplicaciones realistas de la teoría, sino como una forma de que el alumno entienda el uso que se dará en otras asignaturas de su carrera a los conceptos estudiados.

Queremos agradecer a nuestros compañeros el apoyo y la ayuda que nos han prestado, especialmente a Manuel Mocholí, que nos animó a emprender el trabajo.

Valencia, octubre de 2001,

LOS AUTORES

1. Álgebra matricial

La matematización de la economía se realiza a través del concepto de número real, que nos permite asignar un valor numérico —cuantificar— cualquier magnitud económica. Una realidad económica puede tratarse matemáticamente a partir del momento en que encontramos un medio de describirla mediante magnitudes numéricas cuyo comportamiento y relaciones mutuas podemos estudiar (precios, salarios, réditos, probabilidades, tasas de inflación, de desempleo, beneficios, costes, etc.). Sin embargo, es muy raro que un problema venga determinado por un único dato numérico. Lo usual es que sea necesario trabajar simultáneamente con muchos datos. En este tema veremos los conceptos básicos para trabajar sistemáticamente con "bloques" de números.

1.1 Definición de matriz y operaciones

Matrices Si m, $n \geq 1$ son números naturales, una *matriz* $m \times n$ de números reales es una tabla A de mn números reales ordenados en m filas y n columnas. Al número que ocupa la fila i y la columna j se representa por a_{ij}, por lo que una matriz A se representa también por $A = (a_{ij})$. Así pues, una matriz $m \times n$ es de la forma

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

Por ejemplo, la matriz A es 3×3, mientras que B es 2×4:

$$A = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & 4 \\ 0 & \sqrt{2} & -8 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & 0 & -1 & -1 \\ 2 & 1 & 1/2 & 9 \end{pmatrix}.$$

Suma Si $A = (a_{ij})$ y $B = (b_{ij})$ son matrices $m \times n$, entonces $A + B = (a_{ij} + b_{ij})$. Por ejemplo,

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 9 \end{pmatrix} + \begin{pmatrix} -1 & 1 & -2 \\ 4 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 4 & -4 \\ 6 & 1 & 9 \end{pmatrix}.$$

Producto por un escalar Si $\alpha \in \mathbb{R}$ y $A = (a_{ij})$ es una matriz $m \times n$, entonces $\alpha A = (\alpha a_{ij})$. Por ejemplo,

$$-3 \begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 9 \end{pmatrix} = \begin{pmatrix} -3 & -9 & 6 \\ -6 & -3 & -27 \end{pmatrix}.$$

Producto de matrices Si $A = (a_{ij})$ es $m \times n$ y $B = (b_{ij})$ es $n \times r$, entonces AB es la matriz $m \times r$ que en la posición (i, j) tiene el número $a_{i1}b_{1j} + \cdots + a_{in}b_{nj}$. Por ejemplo,

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 9 \end{pmatrix} \begin{pmatrix} 2 & -1 & 0 \\ 1 & -1 & 3 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 2+3-2 & -1-3-2 & 0+9+0 \\ 4+1+9 & -2-1+9 & 0+3+0 \end{pmatrix} =$$

$$= \begin{pmatrix} 3 & -6 & 9 \\ 14 & 6 & 3 \end{pmatrix}.$$

Trasposición Si A es una matriz $m \times n$, se llama *matriz traspuesta* de A a la matriz $n \times m$ representada por A^t dada por $a^t_{ij} = a_{ji}$, es decir, A^t es la matriz que resulta de cambiar filas por columnas. Por ejemplo,

$$A = \begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 9 \end{pmatrix}, \qquad A^t = \begin{pmatrix} 1 & 2 \\ 3 & 1 \\ -2 & 9 \end{pmatrix}.$$

1.2 Tipos de matrices

Matrices cuadradas Las matrices con el mismo número de filas que de columnas se llaman *cuadradas,* mientras que las que no son cuadradas se llaman *rectangulares.* Normalmente, en lugar de decir que una matriz cuadrada es $n \times n$ se dice que es de orden n.

Matriz nula La *matriz nula* $m \times n$ es la matriz cuyos coeficientes son todos 0.

Matrices fila y columna Una *matriz fila* (o *vector fila*) es una matriz $1 \times n$. Una *matriz columna* (o *vector columna*) es una matriz $n \times 1$. Por ejemplo, la matriz A es una matriz fila y la matriz B es una matriz columna

$$A = (2, -2, 5), \qquad B = \begin{pmatrix} 3 \\ 0 \end{pmatrix}.$$

Matrices diagonales Decimos que una matriz cuadrada $A = (a_{ij})$ es diagonal si $a_{ij} = 0$ para $i \neq j$. Por ejemplo,

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

Matriz identidad La matriz identidad $m \times m$ es la matriz I_m que tiene unos en la diagonal y el resto ceros. Por ejemplo,

$$I_1 = (1), \qquad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \ldots$$

Matrices triangulares Una matriz cuadrada $A = (a_{ij})$ es *triangular superior* si todos los elementos que están por debajo de la diagonal son cero, es decir, si $a_{ij} = 0$ cuando $i > j$. Diremos que $A = (a_{ij})$ es *triangular inferior* si son cero los elementos que están arriba de la diagonal, es decir, $a_{ij} = 0$ cuando $i < j$. Por ejemplo, A es triangular superior, B triangular inferior y C triangular superior e inferior:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & 4 \\ 0 & 0 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -2 & 0 \\ 1 & 4 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Nótese que las matrices diagonales son triangulares superiores e inferiores a la vez.

Matrices ortogonales Una matriz cuadrada A es *ortogonal* si al multiplicarla por su traspuesta se obtiene la identidad, es decir, $A \cdot A^t = A^t \cdot A = I_m$. Por ejemplo, la matriz O es ortogonal:

$$O = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad O \cdot O^t = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Matrices simétricas Una matriz cuadrada A es *simétrica* si coincide con su traspuesta $A = A^t$, es decir, si $a_{ij} = a_{ji}$. Por ejemplo, la matriz S es simétrica, pero T no lo es:

$$S = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & 4 \\ 3 & 4 & 0 \end{pmatrix}, \qquad T = \begin{pmatrix} 1 & 2 & 3 \\ 2 & -2 & 4 \\ 1 & 4 & 0 \end{pmatrix}.$$

Matrices antisimétricas Una matriz cuadrada A es *antisimétrica* si coincide con su traspuesta cambiada de signo, $A = -A^t$, es decir, si $a_{ij} = -a_{ji}$. Por ejemplo, la matriz S es antisimétrica:

$$S = \begin{pmatrix} 0 & 2 & 3 \\ -2 & 0 & -4 \\ -3 & 4 & 0 \end{pmatrix}.$$

Matrices equivalentes Dos matrices A y B son *equivalentes*, y lo representaremos por $A \sim B$, si podemos obtener B a partir de una cantidad finita de operaciones elementales con las filas o columnas de A. Por operaciones elementales entendemos las siguientes:

a) cambiar el orden de las filas o columnas,

b) multiplicar alguna fila o columna por un escalar distinto de cero,

c) sumarle a una fila (o columna) una combinación de otras obtenida sumando filas (o columnas) multiplicadas por algún escalar.

Por ejemplo, las matrices A y B son equivalentes:

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ -1 & 1 & 4 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & -1 \\ 0 & -5 & 5 \\ 0 & 3 & 3 \end{pmatrix} = B$$

Basta comprobar que
1^a fila de $B = 1^a$ fila de A
2^a fila de $B = 2^a$ fila de $A + (-2)\times$ 1^a fila de A
3^a fila de $B = 3^a$ fila de $A + 1^a$ fila de A

1.3 Determinantes

Cada matriz cuadrada A tiene asociado un número real llamado *determinante* de A, que representaremos por $|A|$ o $\det A$. No vamos a dar una definición explícita de determinante, sino que en su lugar daremos criterios para calcularlos en la práctica.

Matrices 1×1 Simplemente, $|a| = a$. Por ejemplo, $|-5| = -5$.

Matrices 2×2 La fórmula es

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Por ejemplo,

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

Matrices 3×3 La fórmula para calcular determinantes 3×3 se conoce como *regla de Sarrus:*

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} = aei + bfg + cdh - ceg - afh - bdi.$$

Por ejemplo,

$$\begin{vmatrix} 1 & -1 & 3 \\ 1 & 2 & 1 \\ -3 & 2 & 0 \end{vmatrix} = 0 + 3 + 6 + 18 - 2 - 0 = 25.$$

Para dimensiones superiores conviene manipular los determinantes para simplificarlos y reducirlos a otros de dimensión menor. Para ello debemos conocer las siguientes propiedades de los determinantes (válidas para determinantes de cualquier dimensión):

1. Si una fila o columna contiene sólo ceros, el determinante es nulo.

2. Si intercambiamos dos filas (o columnas) el determinante cambia de signo.

3. Un escalar que multiplique a toda una fila (o columna) puede extraerse del determinante.

4. Si a una fila (o columna) le sumamos otra multiplicada por un número, el determinante no varía.

5. Si la matriz es triangular o diagonal, el determinante es el producto de los elementos de la diagonal principal.

6. El determinante de una matriz coincide con el de su traspuesta.

7. El determinante del producto de dos matrices es el producto de los dos determinantes.

Aplicando estas propiedades siempre podemos conseguir que una fila (o columna) de un determinante tenga nulos todos sus coeficientes salvo a lo sumo uno de ellos. Por ejemplo,

$$\begin{vmatrix} 2 & -2 & 1 & 1 \\ 3 & 0 & 2 & -2 \\ 1 & 3 & 2 & -1 \\ 1 & -3 & -1 & 2 \end{vmatrix} = \begin{vmatrix} 8/3 & 0 & 7/3 & 1/3 \\ 3 & 0 & 2 & -2 \\ 1 & 3 & 2 & -1 \\ 2 & 0 & 1 & 1 \end{vmatrix}$$

Aquí, a la primera fila le hemos sumado la tercera multiplicada por $2/3$ y a la cuarta le hemos sumado la tercera. Una vez el determinante tiene una fila o columna de ceros salvo un coeficiente a_{ij} (en nuestro ejemplo $a_{32} = 3$) el determinante es igual a $(-1)^{i+j}a_{ij}$ multiplicado por el determinante que resulta de eliminar la fila i y la columna j. En nuestro caso

$$= (-1)^{3+2}\, 3 \begin{vmatrix} 8/3 & 7/3 & 1/3 \\ 3 & 2 & -2 \\ 2 & 1 & 1 \end{vmatrix} = -3\,\frac{1}{3} \begin{vmatrix} 8 & 7 & 1 \\ 3 & 2 & -2 \\ 2 & 1 & 1 \end{vmatrix} =$$

$$= -(16 - 28 + 3 - 4 + 16 - 21) = 18.$$

1.4 Rango de matrices

Rango Dada una matriz A llamaremos *rango* de A, y lo denotaremos $\operatorname{rang} A$, al orden de la mayor submatriz cuadrada contenida en A que tenga determinante no nulo. Es decir, diremos que el rango de A es r si contiene al menos una submatriz cuadrada de orden r con determinante distinto de cero y cualquier submatriz de A de orden $r+1$ tiene determinante nulo.

Ejemplo *Calcula el rango de la matriz*

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

SOLUCIÓN: Sabemos que $\operatorname{rang} A$ es a lo sumo 3, puesto que es el orden de la mayor submatriz cuadrada de A. Ahora bien, como

$$\det A = 0 \quad \Longrightarrow \quad \operatorname{rang} A < 3.$$

Consideramos la submatriz

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

Como

$$\begin{vmatrix} 1 & 2 \\ 3 & -2 \end{vmatrix} = -8 \neq 0 \quad \Longrightarrow \quad \operatorname{rang} A = 2$$

■

Para facilitar el cálculo del rango, conviene introducir el siguiente concepto:

Matrices orladas Dada una submatriz B de A, cuando añadimos a B una fila y una columna respetando la ordenación original de la matriz A decimos que hemos *orlado* la matriz B.

Es fácil probar que el rango de una matriz A verifica las siguientes propiedades:

a) Supongamos que existe una submatriz A_r de orden r tal que $\det A_r \neq 0$. Si las matrices de orden $r+1$ obtenidas orlando la matriz A_r tienen determinante nulo, entonces todas las submatrices de A de orden $r+1$ tienen determinante nulo.

b) Las únicas matrices con rango 0 son las nulas.

c) Dos matrices equivalentes tienen el mismo rango.

Como veremos en el ejemplo siguiente, la propiedad (a) en ocasiones permite ahorrar muchos cálculos.

Ejemplo *Calcula el rango de la matriz*

$$A = \begin{pmatrix} 1 & 0 & 2 & -1 \\ 2 & 1 & 0 & 3 \\ 5 & 1 & 6 & 0 \\ 4 & 1 & 4 & 1 \\ 3 & 0 & 6 & -3 \end{pmatrix}$$

SOLUCIÓN: Sabemos que el rango está entre 1 y 4.

Orden 1: $|1| = 1 \neq 0 \quad \Longrightarrow \quad \operatorname{rang} A \geq 1.$

Orden 2: $\begin{vmatrix} 1 & 0 \\ 2 & 1 \end{vmatrix} = 1 \neq 0 \quad \Longrightarrow \quad \operatorname{rang} A \geq 2$

Orden 3: $\begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 5 & 1 & 6 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 5 & 1 & 0 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 4 & 1 & 4 \end{vmatrix} = 0,$

$\begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 4 & 1 & 1 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 3 & 0 & 6 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & -1 \\ 2 & 1 & 3 \\ 3 & 0 & -3 \end{vmatrix} = 0.$

Tenemos una submatriz de orden 2 con determinante no nulo y al orlarla para obtener submatrices de orden 3 todas tienen determinante nulo. Por tanto, el rango de A es 2. ∎

1.5 Cálculo de matrices inversas

Antes de construir la matriz inversa necesitamos introducir algunos conceptos:

Adjunto de un elemento Consideramos una matriz cuadrada A de orden n. Dado un elemento a_{ij} de A, si suprimimos la fila i-ésima y la columna j-ésima se obtiene una submatriz cuadrada de orden $n-1$. Denotamos por α_{ij} el determinante de ésta submatriz. Entonces, se define el adjunto del elemento a_{ij} como

$$A_{ij} = (-1)^{i+j}\alpha_{ij}.$$

Ejemplo *Dada la matriz*

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

calcula el adjunto del elemento a_{31}.

SOLUCIÓN: Si en A eleminamos la 3^a fila y la 1^a columna, nos queda la matriz

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

cuyo determinante vale 4. Entonces $A_{31} = (-1)^{3+1}4 = 4.$ ∎

Matriz adjunta Dada la matriz cuadrada A, se llama matriz adjunta de A, y se representa $\mathrm{ad}(A)$, a la matriz que se obtiene al sustituir cada elemento a_{ij} por su adjunto.

Ejemplo *Calcula la matriz adjunta de*

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

Solución: Calculamos los adjuntos de cada elemento:

$$\mathrm{ad}(A) = \begin{pmatrix} -2 & -2 & 2 \\ 0 & -2 & 0 \\ 2 & 4 & 2 \end{pmatrix}.$$

■

Matriz inversa Dada la matriz cuadrada A de orden n, se llama matriz inversa de A a una matriz A^{-1} que cumpla

$$A \cdot A^{-1} = A^{-1} \cdot A = I_n.$$

Si existe A^{-1}, la matriz A se le llama *matriz regular* y si no existe se llama *matriz singular*. Para que exista la matriz inversa es condición necesaria y suficiente que $|A| \neq 0$ y la forma de calcularla es la siguiente:

a) Calculamos $|A|$. Si vale 0 no existe A^{-1}, y si $|A| \neq 0$ continuamos.

b) Calculamos $\mathrm{ad}(A)$, la matriz adjunta de A.

c) Calculamos la traspuesta de la adjunta, es decir $\mathrm{ad}(A)^t$.

d) La matriz inversa es
$$A^{-1} = \frac{1}{|A|}\,\mathrm{ad}(A)^t.$$

Ejemplo *Calcula la matriz inversa de*

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

Solución: Como $|A| = 4$, existe A^{-1}. La matriz adjunta y su traspuesta son

$$\mathrm{ad}(A) = \begin{pmatrix} -2 & -2 & 2 \\ 0 & -2 & 0 \\ 2 & 4 & 2 \end{pmatrix} \qquad \mathrm{ad}(A)^t = \begin{pmatrix} -2 & 0 & 2 \\ -2 & -2 & 4 \\ 2 & 0 & 2 \end{pmatrix}.$$

Por tanto,

$$A^{-1} = \frac{1}{4}\begin{pmatrix} -2 & 0 & 2 \\ -2 & -2 & 4 \\ 2 & 0 & 2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} & 0 & \frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}.$$

∎

1.6 Ejercicios

1. Consideramos las matrices

$$A = \begin{pmatrix} 1 & 3 & -1 \\ 1 & 2 & 8 \\ 2 & 2 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 3 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & 10 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 3 & -1 \\ 2 & -2 & 1 \\ 2 & 3 & 4 \end{pmatrix}.$$

Calcula:

(a) $A+B+C$.
(b) $A-2C+3B$.
(c) $2[A-3B]-2C$.
(d) A^t-2B^t.
(e) $A\cdot B$.
(f) $B\cdot A$.
(g) $(A+B)\cdot C$.

2. Comprueba:

$$\begin{vmatrix} 1 & 3 & -1 \\ 1 & 2 & 8 \\ 2 & 2 & 3 \end{vmatrix} = 31, \qquad \begin{vmatrix} 2 & 3 & 0 \\ 2 & -1 & 5 \\ 1 & 1 & 10 \end{vmatrix} = -75, \qquad \begin{vmatrix} 1 & 3 & -1 \\ 2 & -2 & 1 \\ 2 & 3 & 4 \end{vmatrix} = -39,$$

$$\begin{vmatrix} 1 & 3 & -1 & 0 \\ 1 & 2 & 8 & 0 \\ 2 & 2 & 3 & 0 \\ 2 & 5 & 1 & 2 \end{vmatrix} = 62, \qquad \begin{vmatrix} 1 & 3 & 1 & 2 \\ 0 & 1 & -1 & 1 \\ 5 & 2 & 1 & 1 \\ 1 & 3 & 2 & 4 \end{vmatrix} = -38,$$

$$\begin{vmatrix} 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 3 \\ 5 & 5 & 5 & 5 \\ -2 & 0 & -2 & -2 \end{vmatrix} = 60, \qquad \begin{vmatrix} 1 & 2 & -1 \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{vmatrix} = 6, \qquad \begin{vmatrix} a & b & c \\ 1 & 1 & 2 \\ 2 & 1 & 1 \end{vmatrix} = -a+3b-c,$$

$$\begin{vmatrix} a & 2 & -1 \\ 1 & b & 2 \\ 2 & 1 & c \end{vmatrix} = abc-2a+2b-2c+7, \qquad \begin{vmatrix} a & 3 & 1 & a \\ 0 & a & a & 1 \\ 5 & 2 & 16 & 1 \\ 1 & 3 & 2 & 4 \end{vmatrix} = -7a^2+39a-14.$$

3. Calcula el rango de las matrices siguientes:

$$A = \begin{pmatrix} 1 & 3 \\ 1 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 1 & 2 & 8 & 0 \\ 2 & 1 & 3 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 3 & -1 \\ 2 & -2 & 1 \\ 2 & 3 & 4 \end{pmatrix},$$

$$D = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \qquad E = \begin{pmatrix} 2 & 1 & 2 & 0 & -1 \\ 0 & 0 & 3 & 3 & 2 \\ 5 & 1 & 5 & 5 & 1 \\ -2 & 0 & -2 & -2 & 0 \end{pmatrix}.$$

4. Dada una matriz cuadrada A, demuestra las siguientes afirmaciones:
   (a) La matriz $A + A^t$ es simétrica.
   (b) La matriz $A - A^t$ es antisimétrica.
   (c) La matriz A se descompone como suma de una matriz simétrica y una antisimétrica de la forma siguiente: $A = \dfrac{1}{2}(A + A^t) + \dfrac{1}{2}(A - A^t)$.

5. Calcula la inversa de las matrices siguientes:

$$A = \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 3 & -1 \\ 1 & 2 & 8 \\ 2 & 2 & 3 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 3 & -1 \\ 2 & -2 & 1 \\ 2 & 3 & 4 \end{pmatrix},$$

$$D = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad E = \begin{pmatrix} 2 & 2 & 2 & 0 \\ 0 & 3 & 3 & 3 \\ 5 & 5 & 5 & 5 \\ -2 & 0 & -2 & -2 \end{pmatrix}.$$

6. Tres agentes comerciales a comisión, V_1, V_2 y V_3, venden tres productos P_1, P_2, P_3. Las matrices E, F, M y A reflejan los ingresos del primer cuatrimestre del año 2001 expresados en €:

$$E = \begin{array}{c} \begin{array}{ccc} P_1 & P_2 & P_3 \end{array} \\ \begin{pmatrix} 1150 & 1095 & 905 \\ 1230 & 1130 & 871 \\ 1050 & 1350 & 970 \end{pmatrix} \end{array} \begin{array}{c} \\ V_1 \\ V_2 \\ V_3 \end{array} \qquad F = \begin{array}{c} \begin{array}{ccc} P_1 & P_2 & P_3 \end{array} \\ \begin{pmatrix} 1202 & 1150 & 875 \\ 1135 & 1232 & 781 \\ 993 & 1250 & 863 \end{pmatrix} \end{array} \begin{array}{c} \\ V_1 \\ V_2 \\ V_3 \end{array}$$

$$M = \begin{array}{c} \begin{array}{ccc} P_1 & P_2 & P_3 \end{array} \\ \begin{pmatrix} 1090 & 1201 & 883 \\ 1140 & 1345 & 872 \\ 1090 & 1254 & 867 \end{pmatrix} \end{array} \begin{array}{c} \\ V_1 \\ V_2 \\ V_3 \end{array} \qquad A = \begin{array}{c} \begin{array}{ccc} P_1 & P_2 & P_3 \end{array} \\ \begin{pmatrix} 1223 & 1098 & 902 \\ 1142 & 1224 & 901 \\ 1100 & 1250 & 893 \end{pmatrix} \end{array} \begin{array}{c} \\ V_1 \\ V_2 \\ V_3 \end{array}$$

   (a) Calcula los ingresos totales del cuatrimestre.
   (b) Calcula el incremento de ingresos entre el mes de enero y el de febrero.
   (c) Si los vendedores reciben un 8% de los ingresos por ventas en concepto de comisión, ¿cuánto ganó cada uno en este cuatrimestre?

7. Tres empresas E_1, E_2, E_3, necesitan cuatro materias primas P_1, P_2, P_3, P_4. El consumo mensual medio de estas empresas se puede expresar mediante la matriz siguiente:

$$A = \begin{array}{c} \\ \\ \\ \\ \end{array}\begin{array}{cc} \begin{array}{cccc} P_1 & P_2 & P_3 & P_4 \end{array} & \\ \begin{pmatrix} 273 & 133 & 1375 & 62 \\ 330 & 232 & 975 & 160 \\ 257 & 161 & 770 & 76 \end{pmatrix} & \begin{array}{c} E_1 \\ E_2 \\ E_3 \end{array} \end{array}$$

donde las cifras están dadas en Tm.
En el primer trimestre del año 2001, los precios de estas materias primas, expresados en € por Tm., han sido

$$P = \begin{array}{cc} \begin{array}{ccc} \text{E} & \text{F} & \text{M} \end{array} & \\ \begin{pmatrix} 123 & 127 & 131 \\ 330 & 326 & 315 \\ 99 & 103 & 126 \\ 213 & 230 & 254 \end{pmatrix} & \begin{array}{c} P_1 \\ P_2 \\ P_3 \\ P_4 \end{array} \end{array}$$

donde las columnas E, F, M representan los meses de enero, febrero y marzo respectivamente. Expresa mediante una matriz el gasto total de cada empresa cada mes.

8. Una empresa de importación de vehículos recibe pedidos de tres concesionarios A, B y C. El primer concesionario ha solicitado 50 coches del modelo T_1, 15 del modelo T_2, 10 coches del modelo T_3 y 2 del modelo T_4, el concesionario B ha solicitado 17 coches del modelo T_1, 12 del modelo T_2, 7 del modelo T_3 y 3 del modelo T_4; y el concesionario C ha pedido $11, 7, 5$ y 4 coches de los modelos T_1, T_2, T_3 y T_4 respectivamente. Los concesionarios aportan una parte del capital al efectuar la compra y aplazan a 90 días el resto. El concesionario A paga el 50 por cien del total y aplaza el resto, B aplaza un tercio y C aplaza un cuarto del pago. Calcula la cantidad de coches de los tipos T_1, T_2, T_3 y T_4 que la empresa vende al contado y cuántos con pago aplazado.

9. Una empresa produce cuatro bienes diferentes P_1, P_2, P_3 y P_4, para los que utiliza cuatro materias primas m_1, m_2, m_3 y m_4. El consumo en kg. para obtener 1 unidad de cada producto es el siguiente:

$$A = \begin{array}{cc} \begin{array}{cccc} m_1 & m_2 & m_3 & m_4 \end{array} & \\ \begin{pmatrix} 56 & 32 & 21 & 43 \\ 62 & 23 & 15 & 54 \\ 57 & 17 & 21 & 61 \\ 75 & 28 & 35 & 42 \end{pmatrix} & \begin{array}{c} P_1 \\ P_2 \\ P_3 \\ P_4 \end{array} \end{array}$$

y los costes, en € por kg., de cada una de las materias es:

$$B = \begin{pmatrix} 2.7 \\ 3.3 \\ 2.5 \\ 1.3 \end{pmatrix} \begin{array}{c} m_1 \\ m_2 \\ m_3 \\ m_4 \end{array}$$

Dos distribuidores, D_1 y D_2, adquieren las siguientes unidades:

$$C = \begin{array}{c} \\ \\ \end{array}\begin{array}{cccc} P_1 & P_2 & P_3 & P_4 \end{array}$$

$$C = \left(\begin{array}{rrrr} 270 & 130 & 1370 & 60 \\ 230 & 175 & 972 & 121 \end{array} \right) \begin{array}{c} D_1 \\ D_2 \end{array}$$

(a) Calcula e interpreta el significado de los productos AB y CAB.

(b) ¿Cuántos kg. se consumen de cada materia prima para satisfacer las demandas de D_1 y D_2?

2. Sistemas de ecuaciones lineales

2.1 Conceptos básicos

La relación más simple que puede darse entre varias magnitudes es que satisfagan una o varias ecuaciones. En esta sección consideraremos las ecuaciones del tipo más sencillo posible: aquellas en las que las variables aparecen únicamente multiplicadas por escalares y sumadas. Por ejemplo, el sistema

$$\left.\begin{array}{rcl} x+2y-\ \ z & = & 3 \\ 2x-\ \ y+3z & = & 6 \\ -x+\ \ y+4z & = & 3 \end{array}\right\}$$

es un sistema de tres ecuaciones lineales con tres incógnitas. Observemos que admite la expresión matricial

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ -1 & 1 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 3 \\ 6 \\ 3 \end{pmatrix}$$

En general, un sistema de m ecuaciones lineales con n incógnitas,

$$\left.\begin{array}{rl} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n & =b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n & =b_2 \\ \cdots & \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n & =b_m \end{array}\right\}$$

se puede expresar matricialmente como $A\bar{x}^t=\bar{b}^t$, donde A es una matriz $m\times n$ llamada *matriz de coeficientes* del sistema, $\bar{b}\in\mathbb{R}^m$ es el *vector de términos independientes* y $\bar{x}\in\mathbb{R}^n$ es el *vector de incógnitas.*

Cuando todos los términos del vector $\bar{b}^t$ son ceros, se tiene un *sistema lineal homogéneo*, $A\bar{x}^t=\bar{0}^t$. Estos sistemas siempre tienen al menos una solución, $(x_1,x_2,\ldots,x_n)=(0,0,\ldots,0)$, que recibe el nombre de trivial.

Según el número de soluciones, podemos clasificar los sistemas de la forma siguiente:

Sistemas de ecuaciones lineales	Compatibles	Determinados: *Solución única*	
		Indeterminados: *Infinitas soluciones*	
	Incompatibles:	*No tienen solución*	

Sistemas equivalentes Dos sistemas S_1 y S_2 son *equivalentes* si tienen las mismas soluciones.

Para encontrar la solución de un sistema puede convenir encontrar otro sistema equivalente que sea más sencillo de resolver. La forma de obtener estos sistemas equivalentes es utilizar alguna de las operaciones siguientes:

a) cambiar el orden de las ecuaciones,
b) multiplicar alguna ecuación por un escalar distinto de cero,
c) sumarle a una ecuación otra multiplicada por un escalar.

2.2 Resolución de sistemas

2.2.1 Método de Gauss

El método consiste en partir de un sistema S y llegar a uno equivalente que sea triangular.

Sistemas determinados Se trata de sistemas con solución única. Veamos un ejemplo:

$$\left.\begin{array}{r} x+2y-\ z=3 \\ 2x-\ y+3z=6 \\ -x+\ y+4z=3 \end{array}\right\} \Rightarrow \left[\begin{array}{r} \\ -2\times 1^a \text{ ec.} + 2^a \text{ ec.} \\ 1^a \text{ ec.} + 3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} x+2y-\ z=3 \\ -5y+5z=0 \\ 3y+3z=6 \end{array}\right\}$$

$$\Rightarrow \left[\begin{array}{l} \\ 1/5\times 2^a \text{ ec.} \\ 1/3\times 3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} x+2y-z=3 \\ -y+z=0 \\ y+z=2 \end{array}\right\} \Rightarrow \left[\begin{array}{l} \\ \\ 2^a \text{ ec.} + 3^a \text{ ec.} \end{array}\right] \Rightarrow$$

$$\Rightarrow \left.\begin{array}{r} x+2y-z=3 \\ -y+z=0 \\ 2z=2 \end{array}\right\}$$

Con esto hemos *triangulado* el sistema, es decir, hemos dejado la x sólo en la primera ecuación, la y sólo en las dos primeras ecuaciones y la z (sólo) en las tres primeras ecuaciones. Resolver un sistema triangulado es inmediato:

$$z=\frac{2}{2}=1, \qquad y=z=1, \qquad x=3-2y+z=3-2+1=2.$$

La solución es, pues, $(x,y,z)=(2,1,1)$. ∎

Sistemas indeterminados En general un sistema de ecuaciones lineales no tiene por qué tener una única solución. El método de Gauss es aplicable también aunque haya más de una. Veamos un ejemplo:

$$\left.\begin{array}{r} x+2y+z=\ 4\\ 2x+\ y-z=\ 2\\ 7x+8y+z=16 \end{array}\right\} \Rightarrow \left[\begin{array}{c} \\ -2\times 1^a \text{ ec. } + 2^a \text{ ec.}\\ -7\times 1^a \text{ ec. } +3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} x+2y+\ z=\ 4\\ -3y-3z=\ -6\\ -6y-6z=-12 \end{array}\right\}$$

$$\Rightarrow \left[\begin{array}{c} \\ -1/3\times 2^a \text{ ec.}\\ -1/6\times 3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} x+2y+z=4\\ y+z=2\\ y+z=2 \end{array}\right\} \Rightarrow \left[\begin{array}{c} \\ \\ -1\times 2^a \text{ ec. } + 3^a \text{ ec.} \end{array}\right] \Rightarrow$$

$$\Rightarrow \left.\begin{array}{r} x+2y+z=4\\ y+z=2 \end{array}\right\}$$

Ahora el sistema ha quedado triangulado, pero hay menos ecuaciones que incógnitas. En tal caso asignamos valores arbitrarios a todas las variables de la última ecuación excepto a una. Por ejemplo, hacemos $z=\lambda$, donde $\lambda\in\mathbb{R}$ es un número real arbitrario. Al despejar queda:

$$z=\lambda, \qquad y=2-\lambda, \qquad x=4-2y-z=4-2(2-\lambda)-\lambda=\lambda.$$

Las soluciones del sistema son $(x,y,z)=(\lambda,2-\lambda,\lambda)$, para todo $\lambda\in\mathbb{R}$. El hecho de que λ pueda tomar cualquier valor se expresa diciendo que λ es un *parámetro*. Como λ puede tomar infinitos valores, el sistema tiene infinitas soluciones.

A veces podemos necesitar una solución particular del sistema. Para encontrarla basta elegir valores concretos para los parámetros de los que dependa la solución general. Por ejemplo, si hacemos $\lambda=3$ obtenemos la solución particular $(x,y,z)=(3,-1,3)$. ■

Sistemas incompatibles También puede suceder que un sistema de ecuaciones lineales no tenga solución. El método de Gauss nos permite reconocer si se da el caso:

$$\left.\begin{array}{r} 2x-y+3z=2\\ x+y-\ z=1\\ 4x+y+\ z=6 \end{array}\right\} \Rightarrow \left[\begin{array}{c} \\ 1^a \text{ ec. } + -2\times 2^a \text{ ec.}\\ -2\times 1^a \text{ ec. } + 3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} 2x-y+3z=2\\ -3y+5z=0\\ 3y-5z=2 \end{array}\right\}$$

$$\Rightarrow \left[\begin{array}{c} \\ \\ 2^a \text{ ec. } + 3^a \text{ ec.} \end{array}\right] \Rightarrow \left.\begin{array}{r} 2x-y+3z=2\\ -3y+5z=0\\ 0=2 \end{array}\right\}$$

Como la última ecuación es imposible, concluimos que el sistema no tiene solución. (En realidad esto se ve ya al comparar las dos últimas ecuaciones del sistema del centro.) ■

Ejemplo Una empresa de productos alimenticios tiene un stock de 114 kilos de chocolate y 111 litros de leche, con los que puede elaborar tres productos distintos A, B y C. El producto A requiere un 40% de chocolate y un 10% de leche, el producto B requiere un 25% de chocolate y un 25% de leche, mientras que C requiere un 20% de chocolate y un 30% de leche. Del resto de ingredientes (azúcar, etc.) la empresa dispone de reservas abundantes. Determina las posibilidades que tiene la empresa para consumir su stock con los productos A, B y C. ¿Cuál de todas le proporcionará más beneficios si la empresa obtiene 10 € por cada kilo de A, 8 € por cada kilo de B y 6 € por cada kilo de C?

Solución: Llamemos x, y, z a las cantidades respectivas de los productos A, B y C que puede producir la empresa. En total se requieren $0.4x + 0.25y + 0.2z$ kilos de chocolate y $0.1x + 0.25y + 0.3z$ kilos de leche. Por consiguiente hemos de resolver el sistema

$$\left.\begin{array}{r} 0.4x + 0.25y + 0.2z = 114 \\ 0.1x + 0.25y + 0.3z = 111 \end{array}\right\} \Rightarrow \left.\begin{array}{r} 40x + 25y + 20z = 11400 \\ 10x + 25y + 30z = 11100 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \quad \left.\begin{array}{r} 40x + 25y + \ \ 20z = 11400 \\ 75y + 100z = 33000 \end{array}\right\}$$

Hacemos $z = \lambda$, con lo que la solución es

$$(x, y, z) = (10 + \frac{1}{3}\lambda, 440 - \frac{4}{3}\lambda, \lambda), \qquad \text{para todo } \lambda \in \mathbb{R}.$$

Ésta es la solución general del sistema, pero no es cierto que cualquier valor de λ nos de una producción aceptable para la empresa. Hemos de exigir que x, y, z sean mayores o iguales que 0, es decir,

$$10 + \frac{\lambda}{3} \geq 0, \quad 440 - \frac{4\lambda}{3} \geq 0, \quad \lambda \geq 0 \Leftrightarrow \lambda \geq -30, \quad \lambda \leq 330, \quad \lambda \geq 0.$$

En definitiva, los valores aceptables para el parámetro son los que cumplen $0 \leq \lambda \leq 330$. De este modo tenemos un único parámetro λ que determina cada una de las posibilidades de la empresa. Si expresamos el beneficio correspondiente en función de λ estaremos en condiciones de determinar qué opción es la más ventajosa:

$$B = 10\Big(10 + \frac{\lambda}{3}\Big) + 8\Big(440 - \frac{4\lambda}{3}\Big) + 6\lambda = 3620 - \frac{4\lambda}{3}.$$

Ahora es claro que el beneficio será mayor cuanto menor sea λ, luego será máximo para $\lambda = 0$. La solución más conveniente para la empresa es, pues, $(x, y, z) = (10, 440, 0)$. ■

2.2.2 La regla de Cramer

La regla de Cramer es otro método para resolver sistemas de ecuaciones lineales mediante determinantes. El caso principal es el de un sistema S, $A\bar{x}^t = \bar{b}^t$, con el mismo número de ecuaciones que incógnitas cuya matriz de coeficientes tiene determinante no nulo. El término x_i de la la solución $(x_1, x_2, \cdots, x_n)$ se calcula como un cociente de determinantes. El numerador es el determinante de la matriz A sustituyendo la columna i-ésima por el vector de términos independientes, y el denominador es $|A|$, es decir

$$x_i = \frac{\begin{array}{c} \begin{array}{ccccc} & & (i) & & \end{array} \\ \begin{vmatrix} a_{11} & \cdots & b_1 & \cdots & a_{1n} \\ a_{21} & \cdots & b_2 & \cdots & a_{2n} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & b_n & \cdots & a_{nn} \end{vmatrix} \end{array}}{|A|}, \quad i = 1, 2, \cdots, n.$$

Sistemas determinados Se tiene el mismo número de ecuaciones que de incógnitas y $|A| \neq 0$. Veámoslo en un ejemplo:

$$\left.\begin{array}{r} x + 2y - z = 3 \\ 2x - y + 3z = 6 \\ -x + y + 4z = 3 \end{array}\right\}$$

Según la regla de Cramer,

$$x = \frac{\begin{vmatrix} 3 & 2 & -1 \\ 6 & -1 & 3 \\ 3 & 1 & 4 \end{vmatrix}}{\begin{vmatrix} 1 & 2 & -1 \\ 2 & -1 & 3 \\ -1 & 1 & 4 \end{vmatrix}} = \frac{-60}{-30} = 2.$$

El denominador es el determinante de la matriz de coeficientes del sistema, mientras que el numerador resulta de sustituir en esta matriz la columna de los coeficientes de x por la columna de los términos independientes. Si sustituimos los coeficientes de y obtenemos el valor de y:

$$y = \frac{\begin{vmatrix} 1 & 3 & -1 \\ 2 & 6 & 3 \\ -1 & 3 & 4 \end{vmatrix}}{-30} = \frac{-30}{-30} = 1.$$

Igualmente:

$$z = \frac{\begin{vmatrix} 1 & 2 & 3 \\ 2 & -1 & 6 \\ -1 & 1 & 3 \end{vmatrix}}{-30} = \frac{-30}{-30} = 1.$$

La solución es, por tanto, $(x, y, z) = (2, 1, 1)$.

Sistemas indeterminados Si tenemos menos ecuaciones que incógnitas también podemos aplicar la regla de Cramer del modo siguiente:

$$\left.\begin{array}{r}x+\ \ y-2z+\ \ w=4\\2x+2y+\ \ z+2w=3\end{array}\right\}$$

Buscamos una submatriz 2×2 de la matriz de coeficientes con determinante no nulo. Vemos que el formado por las dos primeras columnas no sirve, pero el formado por la segunda y la tercera sí:

$$\begin{vmatrix}1 & 1\\2 & 2\end{vmatrix}=0, \qquad \begin{vmatrix}1 & -2\\2 & 1\end{vmatrix}=5.$$

Entonces dejamos la segunda y la tercera columna a la izquierda y pasamos las restantes a la derecha:

$$\left.\begin{array}{r}y-2z=4-\ \ x-\ \ w\\2y+\ \ z=3-2x-2w\end{array}\right\}$$

Las variables de la derecha las convertimos en parámetros: $x=\lambda$, $w=\mu$, y las de la izquierda las calculamos por la regla de Cramer:

$$y=\frac{\begin{vmatrix}4-\ \ \lambda-\ \ \mu & -2\\3-2\lambda-2\mu & 1\end{vmatrix}}{\begin{vmatrix}1 & -2\\2 & 1\end{vmatrix}}=\frac{1}{5}(4-\lambda-\mu+6-4\lambda-4\mu)=2-\lambda-\mu,$$

$$z=\frac{\begin{vmatrix}1 & 4-\ \ \lambda-\ \ \mu\\2 & 3-2\lambda-2\mu\end{vmatrix}}{5}=\frac{1}{5}(3-2\lambda-2\mu-8+2\lambda+2\mu)=-1.$$

La solución es $(x,y,z,w)=(\lambda,2-\lambda-\mu,-1,\mu)$, para todo λ, $\mu\in\mathbb{R}$. ■

Un sistema de ecuaciones lineales puede clasificarse sin necesidad de resolverlo. Dado el sistema $A\bar{x}^t=\bar{b}^t$ tiene asociadas dos matrices: la *matriz de coeficientes,* A, y la *matriz ampliada,* A^*, que se obtiene añadiendo a A la columna de términos independientes:

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

A partir de los rangos de estas matrices podemos saber el tipo de solución del sistema de la forma siguiente:

$$\textbf{Sistema lineal de m ecuaciones n incógnitas}\left\{\begin{array}{l}\begin{array}{c}\text{Compatible}\\rg(A)=rg(A^*)=k\end{array}\left\{\begin{array}{c}\text{Determinado:}\\k=n\\\text{Indeterminado:}\\k<n\end{array}\right.\\[2ex]\text{Incompatible:}\quad rg(A)\neq rg(A^*)\end{array}\right.$$

2.3 Aplicaciones

Aplicación 1: Precios de equilibrio Sea $\bar{p} = (p_1, p_2, p_3)$ el vector de precios de un mercado en el que se venden tres productos A, B y C. Se estima que la oferta y la demanda de cada uno de ellos viene dada por

$$\begin{array}{ll} S_A = 15p_1 + p_2 + 3p_3 - 13 & D_A = 70 - 8p_1 - p_2 - p_3 \\ S_B = p_1 + 20p_2 + 10p_3 - 10 & D_B = 93 - 2p_1 - 4p_2 - p_3 \\ S_C = 10p_1 + 15p_2 + 30p_3 - 50 & D_C = 107 - p_1 - 3p_2 - 5p_3 \end{array}$$

Vamos a calcular los precios de equilibrio, es decir, los precios para los cuales la oferta coincide con la demanda. Esto nos lleva a resolver el sistema:

$$\left.\begin{array}{r} 15p_1 + p_2 + 3p_3 - 13 = 70 - 8p_1 - p_2 - p_3 \\ p_1 + 20p_2 + 10p_3 - 10 = 93 - 2p_1 - 4p_2 - p_3 \\ 10p_1 + 15p_2 + 30p_3 - 50 = 107 - p_1 - 3p_2 - 5p_3 \end{array}\right\}$$

En primer lugar lo ordenamos y luego aplicamos el método de Gauss:

$$\left.\begin{array}{r} 23p_1 + 2p_2 + 4p_3 = 83 \\ 3p_1 + 24p_2 + 11p_3 = 103 \\ 11p_1 + 18p_2 + 35p_3 = 157 \end{array}\right\} \Rightarrow \left.\begin{array}{r} 23p_1 + 2p_2 + 4p_3 = 83 \\ 546p_2 + 241p_3 = 2120 \\ 392p_2 + 761p_3 = 2698 \end{array}\right\} \Rightarrow$$

$$\Rightarrow \left.\begin{array}{r} 23p_1 + 2p_2 + 4p_3 = 83 \\ 546p_2 + 241p_3 = 2120 \\ 321034p_3 = 642068 \end{array}\right\}$$

De aquí se sigue fácilmente que el vector de precios de equilibrio es $\bar{p} = (3, 3, 2)$.

■

Aplicación 2. Modelo input-output de Leontief Consideramos una economía formada por n industrias interrelacionadas $I_1, I_2, \cdots, I_n$ de modo que cada una produce un único bien $b_1, b_2, \cdots, b_n$ respectivamente. Cada industria debe atender las demandas de *inputs* de las n industrias (incluida ella misma) y las demandas externas (*demanda final*). Se trata de calcular el nivel de producción de cada industria para que se satisfagan estos requisitos.

Si x_1 es el nivel de producción de b_1, debe verificarse la ecuación siguiente:

$$x_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + d_1$$

donde $a_{1j}x_j$ representa la demanda de b_1 desde la industria I_j y d_1 es la demanda exterior del producto b_1.

Si repetimos el proceso con la producción de las n industrias se tiene el sistema

$$\left.\begin{array}{c} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + d_1 = x_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + d_2 = x_2 \\ \cdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + d_n = x_n \end{array}\right\}$$

y agrupando las variables,

$$\left.\begin{array}{rl}(1-a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n & = d_1 \\ -a_{21}x_1 + (1-a_{22})x_2 + \cdots - a_{2n}x_n & = d_2 \\ \cdots & \\ -a_{n1}x_1 - a_{n2}x_2 - \cdots + (1-a_{nn})x_n & = d_n\end{array}\right\}$$

En notación matricial se expresaría

$$(I_n - A)\bar{x}^t = \bar{d}^t,$$

siendo $A = (a_{ij})$. A los coeficientes de A se les denomina *coeficientes input*, a la matriz A se le llama *matriz input-output* y $\bar{d}$ es el *vector de demanda final*.

Se pueden presentar dos situaciones diferentes: un *modelo cerrado*, en el que todo se produce y consume internamente, o un *modelo abierto* en el que parte de la producción se destina al consumo exterior. En un modelo cerrado todas las demandas exteriores serán cero, es decir $d_i = 0$, $i = 1, \cdots, n$, y se tendrá por tanto un sistema lineal homogéneo. En el modelo abierto el sistema será completo.

Ejemplo *Supongamos tres industrias interrelacionadas I_1, I_2, I_3 que producen un único bien cada una y cuya producción se obtiene de la forma siguiente: Cada unidad de I_1 requiere 0.3 unidades de I_1, 0.2 unidades de I_2 y 0.3 unidades de I_3. Cada unidad producida en I_2 necesita 0.1 unidades de I_1, 0.2 de I_2 y 0.3 de I_3, y cada unidad de I_3 precisa $0,1$, 0.5 y 0.1 unidades producidas en I_1, I_2 e I_3 respectivamente. Si las demandas exteriores son 45, 50 y 51 unidades de I_1, I_2 e I_3, determina cuáles son los niveles de producción que permiten el equilibrio de esta economía.*

Solución: Llamemos x_1, x_2, x_3 a las unidades producidas por las industrias I_1, I_2, I_3 respectivamente. Se tiene

$$\left.\begin{array}{rcl}0.3x_1 + 0.1x_2 + 0.2x_3 + 45 & = & x_1 \\ 0.2x_1 + 0.2x_2 + 0.5x_3 + 50 & = & x_2 \\ 0.3x_1 + 0.3x_2 + 0.1x_3 + 51 & = & x_3\end{array}\right\}$$

Por tanto hemos de resolver el sistema

$$\left.\begin{array}{rcl}0.7x_1 - 0.1x_2 - 0.2x_3 & = & 45 \\ -0.2x_1 + 0.8x_2 - 0.5x_3 & = & 50 \\ -0.3x_1 - 0.3x_2 + 0.9x_3 & = & 51\end{array}\right\} \Rightarrow \left.\begin{array}{rcl}7x_1 - x_2 - 2x_3 & = & 450 \\ -2x_1 + 8x_2 - 5x_3 & = & 500 \\ -3x_1 - 3x_2 + 9x_3 & = & 510\end{array}\right\}$$

Aplicando la regla de Cramer,

$$x_1 = \frac{\begin{vmatrix} 450 & -1 & -2 \\ 500 & 8 & -5 \\ 510 & -3 & 9 \end{vmatrix}}{\begin{vmatrix} 7 & -1 & -2 \\ -2 & 8 & -5 \\ -3 & -3 & 9 \end{vmatrix}} = \frac{430}{3}\text{ u.}, \quad x_2 = \frac{\begin{vmatrix} 7 & 450 & -2 \\ -2 & 500 & -5 \\ -3 & 510 & 9 \end{vmatrix}}{\begin{vmatrix} 7 & -1 & -2 \\ -2 & 8 & -5 \\ -3 & -3 & 9 \end{vmatrix}} = \frac{620}{3}\text{ u.}$$

$$x_3 = \frac{\begin{vmatrix} 7 & -1 & 450 \\ -2 & 8 & 500 \\ -3 & -3 & 510 \end{vmatrix}}{\begin{vmatrix} 7 & -1 & -2 \\ -2 & 8 & -5 \\ -3 & -3 & 9 \end{vmatrix}} = \frac{520}{3} \text{ u.}$$

2.4 Ejercicios

1. Resuelve:

$$\left.\begin{array}{r} x+2y+\ \ z=3 \\ 2x+y-\ \ z=0 \\ -x+y+3z=5 \end{array}\right\} \quad \left.\begin{array}{r} z+w=5 \\ x+y+z\qquad=4 \\ -y\qquad+w=1 \\ 2z+w=7 \end{array}\right\} \quad \left.\begin{array}{r} 2x+\ \ y=\ \ 4 \\ x-2y=-3 \\ 3x+\ \ y=\ \ 5 \end{array}\right\}$$

$$\left.\begin{array}{r} x+\ \ y+z=3 \\ 2x-\ \ y-z=1 \\ 4x+3y+z=7 \\ 3x+2y\qquad=5 \end{array}\right\} \quad \left.\begin{array}{r} 2x+3y+2z+t=4 \\ x+y-z+t=3 \end{array}\right\} \quad \left.\begin{array}{r} 2x+\ \ y=4 \\ x-2y=0 \\ 3x+\ \ y=5 \\ 3x+\ \ y=5 \end{array}\right\}$$

2. Plantea un sistema de ecuaciones lineales compatible determinado cuya solución sea $(x, y, z, w) = (1, 2-1, 1)$ y resuélvelo por el método de Gauss y por la regla de Cramer. (Cada ecuación deberá contener al menos dos de las incógnitas).

3. Calcula la solución general de los sistemas siguientes así como tres soluciones particulares de cada uno de ellos:

$$\left.\begin{array}{r} -x+2y+3z=3 \\ x+3y-\ \ z=1 \\ x+8y+\ \ z=5 \end{array}\right\} \quad \left.\begin{array}{r} -x+2y+3z+u-5v=2 \\ x+3y-\ \ z-u+3v=0 \end{array}\right\}$$

4. Resuelve la ecuación $x + y - 2z + w = 3$.

5. El comportamiento en el mercado de tres productos, A, B y C vienen expresados por las siguientes curvas de oferta y demanda:

$$\begin{array}{rclcrcl} D_A & = & 13-2x+\ \ y+z & \qquad & S_A & = & 12+x \\ D_B & = & 5+\ \ x-\ \ y+z & & S_B & = & 5+y \\ D_C & = & 10+\ \ x+3y-z & & S_C & = & 2+z \end{array}$$

donde x, y, z son los precios unitarios de los productos A, B, C, respectivamente. Calcula las cantidades que se deben ofrecer de cada producto para alcanzar el equilibrio entre ofertas y demandas.

6. Una empresa fabrica tres productos A, B y C. El vector de precios es $\bar{p}_0 = (4, 4, 5)$, pero la empresa advierte que tiene un exceso de demanda, por lo que decide aumentar los precios a la vez que aumenta su producción buscando una situación de equilibro. Un análisis de la empresa muestra que su capacidad de producción para un vector de precios dado $\bar{p}$ viene dada por

$$\begin{aligned} S_A &= 2p_1 + 3p_2 + 7p_3 - 30 \\ S_B &= 4p_1 + 2p_2 + 3p_3 - 10 \\ S_C &= p_1 + 3p_2 + 3p_3 - 20 \end{aligned}$$

Por otra parte, un estudio de mercado indica que la demanda prevista para un vector de precios $\bar{p}$ es

$$\begin{aligned} D_A &= 140 - 8p_1 - 5p_2 - 2p_3 \\ D_B &= 107 - p_1 - 8p_2 - p_3 \\ D_C &= 78 - p_1 - p_2 - 5p_3 \end{aligned}$$

Calcula el vector de incrementos de precios $\Delta\bar{p}$ necesario para alcanzar los precios de equilibrio, el incremento de producción $\Delta S = (\Delta S_A, \Delta S_B, \Delta S_C)$ que tendrá que efectuar la empresa para satisfacer toda la demanda y el incremento de demanda ΔD que producirá el aumento de los precios. Interpreta lo que se obtiene al calcular la oferta y la demanda correspondientes a los vectores de precios $\bar{p} = (1, 1, 1)$ y $\bar{p} = (100, 100, 100)$.

7. Una empresa fabrica tres artículos A, B y C en cantidades x, y, z y en la producción intervienen dos grupos de trabajadores especializados I y II. Para elaborar una unidad de producto se requieren 4 minutos (o sea, 1/15 de hora) de mano de obra de tipo I y 1/2 hora, 3/4 de hora o 1 hora de mano de obra de tipo II según si el producto es A, B o C. La plantilla de la empresa proporciona 20 horas diarias de tipo I y 160 horas diarias de tipo II. El coste de producción unitario es de 3 u.m. para el producto A, 5 u.m para el producto B y 2 u.m. para C. Expresa las posibilidades de producción de la empresa en términos de un único parámetro. Indica los valores admisibles para éste. ¿Cuál es la producción que minimiza el coste? Calcula el coste mínimo. ¿Y si la empresa tiene comprometida la producción de al menos 5 unidades de cada producto?

8. Una empresa exporta su producto a seis países A, B, C, D, E y F en cantidades anuales determinadas por el vector

$$(x, y, z, u, v, w) = (20, 20, 30, 70, 10, 20).$$

La empresa desea modificar su política de exportaciones de modo que la cantidad u exportada al país D sea lo menor posible. Ahora bien, la capacidad de producción de la empresa está limitada a 170 unidades de producto; los costes de producción son distintos según el país de destino (a causa del transporte), y son, respectivamente, de 2, 1, 3, 2, 1 y 1 unidad monetaria por unidad producida, el presupuesto de la empresa es de 320 unidades monetarias; por último, los países A, B y C forman parte de una asociación económica que impone a la empresa una cuota de importación (máxima) de 70 unidades de producto anuales, y lo mismo sucede con los países A, B, E, y F, con la misma cuota. Determina las posibilidades de la empresa indicando los valores admisibles para los parámetros. (Suponemos que la empresa quiere exportar todo lo que las cuotas impuestas le permiten.) Calcula después la solución que más le conviene a la empresa. ¿Qué valores de los parámetros corresponden a la exportación actual?

9. Tres productores A, B, C interrelacionados distribuyen sus outputs de la forma siguiente:

   a) La producción de A se destina en partes iguales a A, B y C.

   b) La mitad de la producción de B se destina a C, la cuarta parte a A y el resto lo consume internamente.

   c) La mitad de la producción de C se destina a A y el resto lo consumen By C en partes iguales.

   ¿Cuál debe ser la relación entre las producciones para que ninguno deba pagar al otro?

10. Consideramos una economía formada por tres industrias interrelacionadas A, B, C. Sabemos que la matriz input-output es

$$M = \begin{pmatrix} 0.2 & 0.1 & 0 \\ 0.2 & 0.2 & 0.2 \\ 0 & 0.3 & 0.3 \end{pmatrix}$$

   y que el vector de demanda externa es $\bar{d} = (210, 111, 37)$

   (a) Interpreta el significado económico de la suma de los elementos de la tercera columna de M.

   (b) Interpreta el significado económico del elemento $a_{21} = 0.1$ de la matriz M.

   (c) Calcula los niveles de producción que permiten el equilibrio del modelo.

   Obtén la relación de precios que permite que nadie pague nada.

11. En una economía formada por cuatro productores, A, B, C, D, se sabe que la matriz input-output es

$$M = \begin{pmatrix} 0.2 & 0.1 & 0 & 0.3 \\ 0.2 & 0.2 & 0.4 & 0.2 \\ 0 & 0.3 & 0.3 & 0.5 \\ 0.2 & 0.4 & 0.3 & 0 \end{pmatrix}$$

   Razona si puede tratarse de un modelo cerrado de Leontief. ¿Y de un modelo abierto?

3. Espacios vectoriales reales

Ya sabemos que, en general, la forma más habitual de tratar teóricamente varios datos numéricos es a través de un elemento de $\mathbb{R}^n$. En este tema veremos que si las componentes de los datos que estamos manipulando verifican ciertas relaciones sencillas, entonces disponemos de una potente teoría matemática para tratar con ellos. Esta teoría parte del concepto de espacio vectorial.

3.1 Espacios y subespacios vectoriales

Un *espacio vectorial* es un conjunto V (cuyos elementos se llaman *vectores*) en el que hay definidas dos operaciones: una *suma de vectores,* es decir, una operación que a cada par de vectores $\bar{x}$, $\bar{y} \in V$ les asigna una suma $\bar{x} + \bar{y}$, y un *producto escalar,* que a cada $\lambda \in \mathbb{R}$ y cada vector $\bar{x} \in V$ les asigna un vector $\lambda\bar{x} \in V$. Además, estas operaciones han de cumplir las propiedades siguientes (donde las letras griegas representan números reales y las latinas vectores):

1. $(\bar{x} + \bar{y}) + \bar{z} = \bar{x} + (\bar{y} + \bar{z})$,
2. $\bar{x} + \bar{y} = \bar{y} + \bar{x}$,
3. existe un vector $\bar{0} \in V$ tal que $\bar{x} + \bar{0} = \bar{x}$, para todo vector $\bar{x}$,
4. para cada vector $\bar{x}$, existe otro vector $-\bar{x}$ tal que $\bar{x} + (-\bar{x}) = \bar{0}$,
5. $\lambda(\bar{x} + \bar{y}) = \lambda\bar{x} + \lambda\bar{y}$,
6. $(\lambda + \mu)\bar{x} = \lambda\bar{x} + \mu\bar{x}$,
7. $\lambda(\mu\bar{x}) = (\lambda\mu)\bar{x}$,
8. $1\bar{x} = \bar{x}$.

El ejemplo más importante de espacio vectorial es $\mathbb{R}^n$ con la suma y el producto habituales, pero el interés del concepto de espacio vectorial radica en que hay otros ejemplos muy útiles. Las propiedades que aparecen en la definición de espacio vectorial garantizan que la suma y el producto en cualquier espacio vectorial se comportan de forma muy similar a la suma y el producto en $\mathbb{R}^n$. Hay algunas propiedades adicionales que se cumplen en $\mathbb{R}^n$ y que no aparecen en la definición de espacio vectorial pero que, no obstante, se deducen de las ocho propiedades, con lo que también las cumplen todos los espacios vectoriales. Por ejemplo, es cierto en general que $0\bar{x} = \bar{0}$, o que $\lambda\bar{0} = \bar{0}$.

Los ejemplos más importantes de espacios vectoriales aparte de los espacios $\mathbb{R}^n$ son los subespacios de $\mathbb{R}^n$:

Definición Si V es un espacio vectorial, un *subespacio* de V es un subconjunto W que sea también un espacio vectorial con la misma suma y el mismo producto que V. En la práctica, para que esto suceda basta con que W cumpla las propiedades siguientes:

1. $\bar{0} \in W$,
2. si $\bar{x}$, $\bar{y} \in W$, entonces $\bar{x} + \bar{y} \in W$,
3. si $\lambda \in \mathbb{R}$ y $\bar{x} \in W$, entonces $\lambda\bar{x} \in W$.

Ejemplo *Una empresa utiliza tres inputs A, B y C en cantidades x, y, z respectivamente. La elaboración de una unidad de su producto requiere* 3 *unidades de A,* 5 *unidades de B y* 2 *de C. La empresa está estudiando una modificación de su producción. Comprueba que los incrementos factibles* $(\Delta x, \Delta y, \Delta z)$ *(que no producen excedentes de inputs) determinan un subespacio vectorial de* $\mathbb{R}^3$.

Solución: Si la producción de la empresa es P, un incremento ΔP requiere un incremento de los inputs de la forma $(3\Delta P, 5\Delta P, 2\Delta P)$. Lo que nos piden es comprobar que el conjunto

$$W = \{(3\Delta P, 5\Delta P, 2\Delta P) \mid \Delta P \in \mathbb{R}\} = \{\Delta P \cdot (3, 5, 2) \mid \Delta P \in \mathbb{R}\}$$

es un subespacio vectorial de $\mathbb{R}^3$. Comprobamos las condiciones de subespacio vectorial:

1. $(0, 0, 0) \in W$. Esto se cumple tomando $\Delta P = 0$, pues entonces
$$(0, 0, 0) = (3 \cdot 0, 5 \cdot 0, 2 \cdot 0) \in W.$$
En otros términos, $(0, 0, 0)$ es un incremento de inputs factible (el correspondiente a no incrementar la producción).

2. Si $\bar{x}$, $\bar{y} \in W$, entonces $\bar{x} + \bar{y} \in W$. En efecto, del hecho de que $\bar{x}$, $\bar{y} \in W$ se sigue que $\bar{x} = \Delta P \cdot (3, 5, 2)$ e $\bar{y} = \Delta P' \cdot (3, 5, 2)$. Por consiguiente
$$\bar{x} + \bar{y} = \Delta P \cdot (3, 5, 2) + \Delta P' \cdot (3, 5, 2) = (\Delta P + \Delta P') \cdot (3, 5, 2) \in W.$$
En otros términos, si $\bar{x}$ corresponde a un incremento de producción ΔP e $\bar{y}$ corresponde a un incremento de producción $\Delta P'$, entonces $\bar{x} + \bar{y}$ es el incremento de inputs que corresponde a un incremento de producción $\Delta P + \Delta P'$.

3. Si $\bar{x} \in W$ y $\lambda \in \mathbb{R}$, entonces $\lambda\bar{x} \in W$.
En efecto, como $\bar{x} \in W$ ha de ser $\bar{x} = \Delta P \cdot (3, 5, 2)$, luego
$$\lambda\bar{x} = (\lambda\Delta P) \cdot (3, 5, 2) \in W.$$

■

Como hemos visto, la comprobación anterior ha sido un poco laboriosa. Ahora mostraremos que, sabiendo algunos resultados, comprobar que un conjunto es un subespacio vectorial es inmediato. Un caso especialmente simple es el siguiente:

Teorema *Un subconjunto de $\mathbb{R}^n$ definido por un sistema ecuaciones lineales homogéneas (o sea, igualadas a 0) es siempre un subespacio vectorial.*

DEMOSTRACIÓN: Consideremos el sistema de ecuaciones en forma matricial: $A\bar{x}^t = \bar{0}$. El conjunto será de la forma

$$W = \{\bar{x} \in \mathbb{R}^n \mid A\bar{x}^t = \bar{0}\}.$$

Comprobamos que cumple las condiciones de subespacio vectorial:

1. $\bar{0} \in W$, pues $A \cdot \bar{0} = \bar{0}$.

2. Dados $\bar{x}$, $\bar{y} \in W$, hemos de probar que $\bar{x} + \bar{y} \in W$, lo cual equivale a que $A(\bar{x} + \bar{y})^t = \bar{0}$. En efecto, teniendo en cuenta que $A\bar{x}^t = A\bar{y}^t = \bar{0}$ (porque $\bar{x}$, $\bar{y} \in W$),

$$A(\bar{x} + \bar{y})^t = A(\bar{x}^t + \bar{y}^t) = A\bar{x}^t + A\bar{y}^t = \bar{0} + \bar{0} = \bar{0}.$$

3. Dados $\bar{x} \in W$ y $\lambda \in \mathbb{R}$, hemos de probar que $\lambda\bar{x} \in W$, lo cual equivale a que $A(\lambda\bar{x})^t = \bar{0}$. Usando, como antes que $A\bar{x}^t = \bar{0}$, vemos que

$$A(\lambda\bar{x})^t = A\lambda\bar{x}^t = \lambda A\bar{x}^t = \lambda\bar{0} = \bar{0}.$$

■

Ejemplo *Una empresa distribuye su producto en tres mercados A, B y C en cantidades x, y, z respectivamente. La empresa está estudiando cambiar esta distribución sin alterar la producción. Prueba que el conjunto de los incrementos factibles $(\Delta x, \Delta y, \Delta z)$ es un subespacio vectorial de $\mathbb{R}^3$.*

SOLUCIÓN: Para incrementar la distribución en cada mercado en cantidades $(\Delta x, \Delta y, \Delta z)$, es necesario un incremento de producción $\Delta x + \Delta y + \Delta z$. Como la empresa no quiere alterar su producción, tendrá que ser $\Delta x + \Delta y + \Delta z = 0$ (no olvidemos que los incrementos pueden ser negativos). Por lo tanto, el conjunto de incrementos factibles es

$$W = \{(\Delta x, \Delta y, \Delta z) \in \mathbb{R}^3 \mid \Delta x + \Delta y + \Delta z = 0\}.$$

Como se trata de un conjunto definido por una ecuación lineal homogénea, podemos asegurar que es un subespacio vectorial de $\mathbb{R}^3$. ■

3.2 Sistemas generadores

Un concepto muy útil para reconocer fácilmente espacios vectoriales es el de combinación lineal de vectores:

Definición Sea V un espacio vectorial. Diremos que un vector $\bar{x} \in V$ es *combinación lineal* de los vectores $\bar{v}_1, \dots, \bar{v}_n \in V$ si existen escalares $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ tales que

$$\bar{x} = \lambda_1 \bar{v}_1 + \cdots + \lambda_n \bar{v}_n.$$

A los escalares $\lambda_1, \lambda_2, \dots, \lambda_n$ se les llama *coeficientes* de la combinación lineal.

Teorema *Si V es un espacio vectorial y $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \in V$, el conjunto de todas las combinaciones lineales de estos vectores es un subespacio vectorial de V, llamado* envoltura lineal *o* subespacio generado *por $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$. Se representa por*

$$\langle \bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \rangle = \{\lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n \mid \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{R}\}.$$

Demostración: Comprobamos que $W = \langle \bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \rangle$ cumple las propiedades de espacio vectorial:

1. $\bar{0} \in W$, pues $\bar{0} = 0 \cdot \bar{v}_1 + 0 \cdot \bar{v}_2 \cdots + 0 \cdot \bar{v}_n$.

Dados dos vectores $\bar{v}, \bar{w} \in W$, tenemos que

$$\bar{v} = \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n, \quad \bar{w} = \mu_1 \bar{v}_1 + \mu_2 \bar{v}_2 + \cdots + \mu_n \bar{v}_n$$

2. La suma $\bar{v} + \bar{w}$ está en W, pues

$$\begin{aligned} \bar{v} + \bar{w} &= \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n + \mu_1 \bar{v}_1 + \mu_2 \bar{v}_2 + \cdots + \mu_n \bar{v}_n \\ &= (\lambda_1 + \mu_1) \bar{v}_1 + (\lambda_2 + \mu_2) \bar{v}_2 + \cdots + (\lambda_n + \mu_n) \bar{v}_n \in W \end{aligned}$$

3. El producto por un escalar, $\lambda \in \mathbb{R}$, está en W:

$$\begin{aligned} \lambda \bar{v} &= \lambda(\lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n) \\ &= \lambda\lambda_1 \bar{v}_1 + \lambda\lambda_2 \bar{v}_2 + \cdots + \lambda\lambda_n \bar{v}_n \in W \end{aligned}$$

■

Ejemplo (Solución alternativa al ejemplo de la página 26). Hay que probar que el conjunto

$$W = \{\Delta P \cdot (3, 5, 2) \mid \Delta P \in \mathbb{R}\}$$

es un subespacio vectorial de $\mathbb{R}^3$, pero esto es evidente, pues $W = \langle (3, 5, 2) \rangle$. ■

Definición Sea V un espacio vectorial y $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \in V$. Se dice que los vectores $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$ son un *sistema generador* de V si $V = \langle \bar{v}_1, \bar{v}_2, \dots, \bar{v}_n \rangle$, es decir, si todo vector de V se puede expresar como combinación lineal de los vectores $\bar{v}_1, \bar{v}_2, \dots, \bar{v}_n$.

Hemos visto que un sistema de ecuaciones homogéneas determina un espacio vectorial. En el contexto de los espacios vectoriales, resolver el sistema equivale a encontrar un generador de dicho espacio. Lo veremos mediante un ejemplo:

Ejemplo *Una empresa distribuye su producto en cuatro mercados A, B, C y D en cantidades x, y, z, w, respectivamente. El coste de transportar una unidad de producto a cada mercado es de 3, 1, 2 y 1€ respectivamente. La empresa estudia una modificación de su distribución, pero no quiere alterar la producción ni los costes de transporte. Comprueba que el conjunto de incrementos factibles $(\Delta x, \Delta y, \Delta z, \Delta w)$ es un subespacio vectorial de $\mathbb{R}^4$ y encuentra un sistema generador.*

Solución: Un vector de incrementos $(\Delta x, \Delta y, \Delta z, \Delta w)$ requiere un incremento de producción de $\Delta P = \Delta x + \Delta y + \Delta z + \Delta w$, y un incremento de los costes de transporte de $\Delta C = 3\Delta x + \Delta y + 2\Delta z + \Delta w$ €. Como además la empresa exige $\Delta P = \Delta C = 0$, el conjunto de incrementos factibles es

$$W = \{(\Delta x, \Delta y, \Delta z, \Delta w) \in \mathbb{R}^4 \mid \Delta x + \Delta y + \Delta z + \Delta w = 0,$$

$$3\Delta x + \Delta y + 2\Delta z + \Delta w = 0\}.$$

Como está definido por dos ecuaciones lineales homogéneas, es un subespacio vectorial de $\mathbb{R}^4$. Para encontrar un sistema generador, resolvemos el sistema:

$$\left.\begin{array}{rcl} \Delta x + \Delta y + \Delta z + \Delta w & = & 0 \\ 3\Delta x + \Delta y + 2\Delta z + \Delta w & = & 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} \Delta x + \Delta y + \Delta z + \Delta w & = & 0 \\ -2\Delta y - \Delta z - 2\Delta w & = & 0 \end{array}\right\}$$

Tomamos como parámetros $\lambda = \Delta y$, $\mu = \Delta w$, con lo que $\Delta z = -2\lambda - 2\mu$ y $\Delta x = -\Delta y - \Delta z - \Delta w = -\lambda + 2\lambda + 2\mu - \mu = \lambda + \mu$. La solución es

$$(\Delta x, \Delta y, \Delta z, \Delta w) = (\lambda + \mu, \lambda, -2\lambda - 2\mu, \mu) = (\lambda, \lambda, -2\lambda, 0) + (\mu, 0, -2\mu, \mu)$$

$$= \lambda(1, 1, -2, 0) + \mu(1, 0, -2, 1).$$

De este modo, podemos expresar W como

$$W = \{\lambda(1, 1, -2, 0) + \mu(1, 0, -2, 1) \mid \lambda,\ \mu \in \mathbb{R}\} = \langle (1, 1, -2, 0), (1, 0, -2, 1) \rangle .$$

Así pues, un sistema generador de W está formado por los vectores $(1, 1, -2, 0)$ y $(1, 0, -2, 1)$. ■

Expresar un espacio vectorial en términos de un sistema generador nos permite, entre otras cosas, comparar fácilmente dos espacios vectoriales (para determinar si son el mismo o si uno está contenido en otro). Esto se debe al teorema siguiente:

Teorema *Sean W_1 y W_2 dos subespacios vectoriales de un mismo espacio V y sea $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ un sistema generador de W_1. Entonces $W_1 \subset W_2$ si y sólo si $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \in W_2$, es decir, para garantizar que un subespacio está contenido en otro basta probar que los vectores de un sistema generador del primero pertenecen al segundo.*

Demostración:

[$\Rightarrow$] Supongamos que $W_1 \subset W_2$. Esto significa que todos los vectores de W_1 están en W_2, luego en particular, los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \in W_2$.

[$\Leftarrow$] Supongamos ahora que $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \in W_2$. Hemos de probar que $W_1 \subset W_2$. Para ello tomamos $\bar{v} \in W_1$ y hemos de probar que $\bar{v} \in W_2$.

Como $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ es un sistema generador de W_1, existen $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ tales que

$$\bar{v} = \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n.$$

Como W_2 es un subespacio vectorial, los vectores $\lambda \bar{v}_i \in W_2$ (por la propiedad 3) y $\bar{v} \in W_2$ (por la propiedad 2). ■

Ejemplo *Determina si los espacios vectoriales*

$$\begin{aligned} W_1 &= \{(x,y,z,t) \in \mathbb{R}^4 \mid -x-y+z=0,\ y-t=0\}, \\ W_2 &= \langle (1,1,2,1), (1,0,1,0) \rangle \end{aligned}$$

son iguales o no.

Solución: Es claro que los vectores $(1,-1,0,1)$ y $(1,0,1,0)$ verifican las ecuaciones que definen W_1, luego el teorema anterior nos asegura que $W_2 \subset W_1$. Para probar la otra inclusión calcularemos un sistema generador de W_1. Resolvemos el sistema

$$\left.\begin{aligned} -x-y+z &= 0 \\ y-t &= 0 \end{aligned}\right\} \quad \Rightarrow \quad (x,y,z,t) = (\alpha-\beta, \beta, \alpha, \beta), \quad \alpha,\ \beta \in \mathbb{R}.$$

Por lo tanto:

$$(x,y,z,t) = \alpha(1,0,1,0) + \beta(-1,1,0,1)$$

y $W_1 = \langle (1,0,1,0), (-1,1,0,1) \rangle$. Ahora hemos de estudiar si estos vectores están en W_2. Claramente, el vector $(1,0,1,0) \in W_2$. Para decidir si $(-1,1,0,1) \in W_2$ planteamos la ecuación

$$(-1,1,0,1) = \lambda(1,1,2,1) + \mu(1,0,1,0),$$

que tiene como solución $(\lambda, \mu) = (1,-2)$. Concluimos que $(-1,1,0,1)$ es combinación lineal de los generadores de W_2 y, por consiguiente, podemos aplicar el teorema anterior para concluir que $W_1 \subset W_2$. Esto prueba que, de hecho, $W_1 = W_2$. ■

Otra razón del interés de los sistemas generadores es que a menudo es más ilustrativo conocer la expresión de un vector como combinación lineal de un cierto generador que no conocer el vector mismo. Veamos un ejemplo:

Ejemplo *Una empresa utiliza tres inputs A, B y C en cantidades x, y, z para producir dos artículos P y Q. Las cantidades requeridas para la producción de una unidad de cada producto viene dada por la tabla siguiente:*

Artículos	Inputs		
	A	B	C
P	2	1	3
Q	1	2	2

La empresa estudia alterar su producción. Prueba que el conjunto de incrementos factibles $(\Delta x, \Delta y, \Delta z)$ *(que no produzcan excedentes) es el subespacio vectorial de* $\mathbb{R}^3$ *generado por los vectores* $\bar{v} = (2, 1, 3)$ *y* $\bar{w} = (1, 2, 2)$. *Determina si el vector* $(5, 7, 9)$ *determina un incremento factible de los inputs de la empresa.*

Solución: Llamemos r y s a las cantidades producidas de P y Q respectivamente. Para incrementar la producción de P en Δr unidades se necesita un incremento de los inputs de $(\Delta x, \Delta y, \Delta z) = (2\Delta r, \Delta r, 3\Delta r)$, y para incrementar la producción de Q en Δs unidades se necesita un incremento de inputs de $(\Delta x, \Delta y, \Delta z) = (\Delta s, 2\Delta s, 2\Delta s)$. Por consiguiente, un incremento arbitrario de la producción $(\Delta r, \Delta s)$ se corresponde con un incremento de los inputs consumidos de

$$(\Delta x, \Delta y, \Delta z) = (2\Delta r, \Delta r, 3\Delta r) + (\Delta s, 2\Delta s, 2\Delta s) = \Delta r\, \bar{v} + \Delta s\, \bar{w}.$$

El conjunto de incrementos factibles es, pues,

$$W = \{\Delta r\, \bar{v} + \Delta s\, \bar{w} \mid \Delta r,\ \Delta s \in \mathbb{R}\} = \langle \bar{v}, \bar{w} \rangle .$$

Para saber si $(5, 7, 9)$ es un vector factible, es decir, si $(5, 7, 9) \in \langle \bar{v}, \bar{w} \rangle$, hemos de estudiar si $(5, 7, 9)$ es combinación lineal de $\bar{v}$, $\bar{w}$, es decir, si existen λ y $\mu \in \mathbb{R}$ tales que

$$(5, 7, 9) = \lambda\, \bar{v} + \mu\, \bar{w} = \lambda(2, 1, 3) + \mu(1, 2, 2).$$

Al igualar componentes esto equivale al sistema de ecuaciones

$$\left.\begin{array}{rcl} 2\lambda + \mu & = & 5 \\ \lambda + 2\mu & = & 7 \\ 3\lambda + 2\mu & = & 9 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} \lambda + 2\mu & = & 7 \\ 2\lambda + \mu & = & 5 \\ 3\lambda + 2\mu & = & 9 \end{array}\right\} \Rightarrow \left.\begin{array}{rcr} \lambda + 2\mu & = & 7 \\ -3\mu & = & -9 \\ -4\mu & = & -12 \end{array}\right\}$$

de donde $\mu = 3$ y $\lambda = 7 - 6 = 1$. En conclusión

$$(7, 5, 11) = (2, 1, 3) + 3\ (1, 2, 2) = \bar{v} + 3\bar{w},$$

es decir, $(5, 7, 9)$ es el incremento de inputs necesario para incrementar la producción de P en 1 unidad y la de Q en 3 unidades. ■

Según comentábamos, viendo $(5, 7, 9)$ no sabemos si se trata de un vector factible o no, mientras que la expresión $\bar{v} + 3\bar{w}$ no sólo nos muestra que sí lo es, sino que además nos indica con qué incremento de producción se corresponde. Observemos que, en este ejemplo, si la empresa dispone de un incremento de inputs de $(5, 7, 9)$, la única forma que tiene de aprovecharlo sin dejar excedentes es aumentar en 1 unidad la producción de P y en 3 unidades la de Q. En términos abstractos, lo que tenemos es que el vector $(5, 7, 9)$ determina de forma única los coeficientes que lo expresan como combinación lineal del generador $\bar{v}$, $\bar{w}$. Sin embargo, esta unicidad no es cierta en general. Veamos un ejemplo:

Ejemplo *Una empresa utiliza tres inputs* A, B *y* C *en cantidades* x, y, z *para producir tres artículos* P, Q *y* R. *La cantidad de inputs que requiere la producción de una unidad de cada producto viene dada por la tabla siguiente:*

Artículos	Inputs		
	A	B	C
P	1	1	3
Q	2	2	1
R	3	3	4

La empresa estudia alterar su producción. Prueba que el conjunto de incrementos factibles $(\Delta x, \Delta y, \Delta z)$ *(que no produzcan excedentes) es el subespacio vectorial de* $\mathbb{R}^3$ *generado por los vectores* $\bar{v} = (1, 1, 3)$*,* $\bar{w} = (2, 2, 1)$ *y* $\bar{x} = (3, 3, 4)$*. Determina si el vector* $(7, 7, 11)$ *determina un incremento factible de los inputs de la empresa.*

Solución: Si llamamos r, s y t a las cantidades producidas de P, Q y R respectivamente, es claro que un incremento Δr en la producción de P requiere un incremento de inputs de $(\Delta r, \Delta r, 3\Delta r) = \Delta r\,\bar{v}$. Similarmente, un incremento Δs en la producción de Q requiere un incremento de inputs de $\Delta s\,\bar{w}$ y un incremento Δt en la producción de R requiere un incremento de inputs de $\Delta t\,\bar{x}$. Un incremento arbitrario de la producción $(\Delta r, \Delta s, \Delta t)$ requiere un incremento de inputs

$$(\Delta x, \Delta y, \Delta z) = \Delta r\,\bar{v} + \Delta r\,\bar{w} + \Delta r\,\bar{x},$$

luego el conjunto de incrementos factibles es

$$W = \{\Delta r\,\bar{v} + \Delta r\,\bar{w} + \Delta r\,\bar{x} \mid \Delta r,\ \Delta s,\ \Delta t \in \mathbb{R}\} = \langle \bar{v}, \bar{w}, \bar{x} \rangle .$$

Para saber si $(7, 7, 11) \in W$ hemos de determinar si es combinación lineal de $\bar{v}$, $\bar{w}$ y $\bar{x}$, es decir, si existen λ, μ, $\nu \in \mathbb{R}$ tales que

$$(7, 7, 11) = \lambda(1, 1, 3) + \mu(2, 2, 1) + \nu(3, 3, 4).$$

Igualando componentes, esto equivale a que el sistema siguiente tenga solución:

$$\left.\begin{array}{r} \lambda + 2\mu + 3\nu = 7 \\ \lambda + 2\mu + 3\nu = 7 \\ 3\lambda + \mu + 4\nu = 11 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \lambda + 2\mu + 3\nu = \quad 7 \\ -5\mu - 5\nu = -10 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \lambda + 2\mu + 3\nu = 7 \\ \mu + \nu = 2 \end{array}\right\}$$

Tomando como parámetro $\nu = \alpha$ queda $\mu = 2 - \alpha$ y $\lambda = 7 - 2(2 - \alpha) - 3\alpha$. En total:

$$(\lambda, \mu, \nu) = (3 - \alpha, 2 - \alpha, \alpha), \qquad \text{para todo } \alpha \in \mathbb{R}.$$

Como el sistema es compatible, $(7, 7, 11)$ es un vector factible de incrementos.

■

Vemos que en este caso, ante un incremento de inputs de $(7, 7, 11)$, la empresa tiene varias posibilidades para aprovecharlo sin excedentes. Por ejemplo, con $\alpha = 0$ tenemos que

$$(7, 7, 11) = 3\bar{v} + 2\bar{w} + 0 \cdot \bar{x},$$

es decir, que una posibilidad es incrementar en 3 unidades la producción de P y en 2 la de Q, pero con $\alpha = 1$ tenemos

$$(7, 7, 11) = 2\bar{v} + \bar{w} + \bar{x},$$

de modo que otra alternativa es incrementar en 2 unidades la producción de P, en 1 la de Q y en 1 la de R. En términos abstractos, tenemos que $\bar{v}$, $\bar{w}$, $\bar{x}$ es un generador del espacio W con la propiedad de que un mismo vector $(7, 7, 11)$ admite infinitas expresiones distintas como combinación lineal de sus vectores.

3.3 Dependencia e independencia lineal

La diferencia entre los dos últimos ejemplos se explica en términos de un nuevo concepto:

Definición Diremos que los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ de un espacio vectorial V son *linealmente dependientes* (o que forman un *sistema ligado*) si existen escalares $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}$ *no todos nulos* de modo que

$$\lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \bar{\lambda}_n\bar{v}_n = \bar{0}.$$

En caso contrario se dice que son *linealmente independientes* (o que forman un *sistema libre.*)

Conviene explicitar la definición de vectores independientes:

Los vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ de un espacio vectorial V son *linealmente independientes* (o bien que forman un *sistema libre*) si la *única* combinación lineal que cumple $\lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \bar{\lambda}_n\bar{v}_n = \bar{0}$ es aquella en la que todos los escalares $\lambda_1, \lambda_2, \ldots, \lambda_n$ son nulos.

Ejemplo: *Determina si los vectores* $(2, 1, 3)$ *y* $(1, 2, 2)$ *son linealmente dependientes o independientes.*

SOLUCIÓN: Hemos de estudiar si existen escalares λ y μ no ambos nulos tales que

$$\lambda(2, 1, 3) + \mu(1, 2, 2) = (0, 0, 0).$$

Igualando componentes vemos que esta ecuación equivale al sistema

$$\left.\begin{array}{rcl} 2\lambda + \mu & = & 0 \\ \lambda + 2\mu & = & 0 \\ 3\lambda + 2\mu & = & 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} \lambda + 2\mu & = & 0 \\ 2\lambda + \mu & = & 0 \\ 3\lambda + 2\mu & = & 0 \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} \lambda + 2\mu & = & 0 \\ -3\mu & = & 0 \\ -4\mu & = & 0 \end{array}\right\}$$

de donde $\mu = 0$ y, en consecuencia, $\lambda = 0$. Como la única posibilidad es $\lambda = \mu = 0$, los vectores son linealmente independientes. ∎

Es importante conocer esta caracterización:

Teorema *Un conjunto de vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ de un espacio vectorial es ligado si y sólo si uno de ellos es combinación lineal de los demás.*

DEMOSTRACIÓN: Veamos la doble implicación:

[⇒] Si el sistema es ligado, existen escalares $\lambda_1, \lambda_2, \dots, \lambda_n$ no todos nulos tales que $\lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \lambda_n\bar{v}_n = \bar{0}$. Podemos suponer sin pérdida de generalidad que $\lambda_1 \neq 0$. Entonces

$$\lambda_1\bar{v}_1 = -\lambda_2\bar{v}_2 - \cdots - \lambda_n\bar{v}_n \quad \Rightarrow \quad \bar{v}_1 = -\frac{\lambda_2}{\lambda_1}\,\bar{v}_2 - \cdots - \frac{\lambda_n}{\lambda_1}\,\bar{v}_n,$$

luego $\bar{v}_1$ es combinación lineal de $\bar{v}_2, \dots, \bar{v}_n$.

[⇐] Si un vector —por ejemplo $\bar{v}_1$— es combinación lineal del resto, entonces

$$\bar{v}_1 = \lambda_2\bar{v}_2 + \cdots + \bar{\lambda}_n\bar{v}_n, \qquad \text{para ciertos } \lambda_2, \dots, \lambda_n \in \mathbb{R}.$$

Por consiguiente, $(-1)\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \bar{\lambda}_n\bar{v}_n = \bar{0}$, y en esta combinación lineal no todos los coeficientes son nulos (ya que el primero es -1). Esto prueba que los vectores son linealmente dependientes. ∎

La demostración anterior nos muestra además cómo reconocer un vector dependiente del resto en un sistema ligado: basta encontrar una combinación lineal nula con coeficientes no todos nulos. Cualquier vector cuyo coeficiente no sea nulo será combinación lineal del resto.

Ejemplo *Estudia si los vectores* $(1,1,3)$, $(2,2,1)$, $(3,3,4)$ *son linealmente dependientes o independientes. Si son dependientes encuentra uno que sea combinación lineal de los otros.*

SOLUCIÓN: Los vectores serán dependientes o independientes según si la ecuación

$$\lambda(1,1,3) + \mu(2,2,1) + \nu(3,3,4) = (0,0,0)$$

tiene o no una solución con algún coeficiente no nulo. Igualando coordenadas, esta ecuación equivale al sistema de ecuaciones

$$\left.\begin{array}{r} \lambda + 2\mu + 3\nu = 0 \\ \lambda + 2\mu + 3\nu = 0 \\ 3\lambda + \mu + 4\nu = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \lambda + 2\mu + 3\nu = 0 \\ -5\mu - 5\nu = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{r} \lambda + 2\mu + 3\nu = 0 \\ \mu + \nu = 0 \end{array}\right\}$$

Tomando como parámetro $\nu = \alpha$ la solución es

$$(\lambda, \mu, \nu) = (-\alpha, -\alpha, \alpha), \qquad \text{para todo } \alpha \in \mathbb{R}.$$

Como hay solución los vectores son linealmente dependientes. En realidad, hay infinitas soluciones. Tomando, por ejemplo, $\alpha = 1$ obtenemos la solución

$$-1 \cdot (1,1,3) - 1 \cdot (2,2,1) + 1 \cdot (3,3,4) = (0,0,0).$$

Como los tres coeficientes son no nulos, cualquiera de los tres vectores es combinación de los demás (cualquiera puede ser despejado). Por ejemplo

$$(3,3,4) = (1,1,3) + (2,2,1).$$

∎

El hecho fundamental es que si un vector de un sistema generador es combinación lineal de los otros, puede ser eliminado:

Teorema *Sea $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ un sistema generador del espacio vectorial*

$$V = \langle \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \rangle .$$

Si uno de los generadores, por ejemplo $\bar{v}_n$, es combinación lineal de los demás, entonces el sistema que resulta de eliminarlo sigue siendo un sistema generador de V, es decir,

$$V = \langle \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_{n-1} \rangle .$$

Por consiguiente, si tenemos un sistema generador de un espacio vectorial, podemos ir eliminando vectores que sean dependientes de los demás hasta llegar a un sistema generador en el que ningún vector dependa de los otros, es decir, un generador linealmente independiente.

3.4 Bases y dimensión

Definición Una *base* de un espacio vectorial es un conjunto de vectores que es a la vez sistema generador y linealmente independiente.

Teorema *Todo espacio vectorial tiene una base. Además, dos bases cualesquiera de un mismo espacio tienen necesariamente el mismo número de vectores.*

Definición Se llama *dimensión* de un espacio vectorial V al número de vectores de cualquiera de sus bases. Se representa por $\dim V$.

Los razonamientos precedentes nos muestran cómo encontrar una base de un espacio vectorial una vez conocemos un sistema generador: basta ir eliminando los vectores que sean combinación lineal de los restantes hasta que no quede ninguno en esta situación.

Ejemplo *Encuentra una base del espacio vectorial*

$$W = \langle (0,0,0), (1,1,1), (2,2,2), (1,2,3), (0,1,2) \rangle .$$

SOLUCIÓN: Tenemos un sistema generador de W formado por 5 vectores. Hemos de estudiar si alguno es combinación lineal de los demás, y en tal caso eliminarlo. Cuando ya no quede ninguno en estas condiciones tendremos una base.

Ante todo, el vector $\bar{0}$ siempre es combinación lineal de otros vectores cualesquiera. Concretamente:

$$(0,0,0) = 0 \cdot (1,1,1) + 0 \cdot (2,2,2) + 0 \cdot (1,2,3) + 0 \cdot (0,1,2).$$

Por lo tanto, el vector $\bar{0}$ siempre se puede eliminar de un generador:

$$W = \langle (1,1,1), (2,2,2), (1,2,3), (0,1,2) \rangle .$$

A simple vista notamos que $(2,2,2) = 2 \cdot (1,1,1)$, luego también podemos eliminarlo:

$$W = \langle (1,1,1), (1,2,3), (0,1,2) \rangle .$$

Ahora estudiamos si los tres vectores que nos quedan son dependientes o independientes, para lo cual planteamos la ecuación

$$\lambda(1,1,1) + \mu(1,2,3) + \nu(0,1,2) = (0,0,0),$$

que equivale al sistema

$$\left.\begin{aligned} \lambda + \mu \qquad &= 0 \\ \lambda + 2\mu + \nu &= 0 \\ \lambda + 3\mu + 2\nu &= 0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} \lambda + \mu \qquad &= 0 \\ \mu + \nu &= 0 \\ 2\mu + 2\nu &= 0 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} \lambda + \mu \qquad &= 0 \\ \mu + \nu &= 0 \end{aligned}\right\}$$

Tomando como parámetro $\nu = \alpha$, la solución es $(\lambda, \mu, \nu) = (\alpha, -\alpha, \alpha)$, para $\alpha \in \mathbb{R}$. Por ejemplo, para $\alpha = 1$ tenemos

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

lo que muestra que, por ejemplo $(0,1,2) = -(1,1,1) + (1,2,3)$ es combinación lineal de los otros dos vectores. Por consiguiente

$$W = \langle (1,1,1), (1,2,3) \rangle .$$

Se comprueba fácilmente que la ecuación $\lambda(1,1,1) + \mu(1,2,3) = (0,0,0)$ sólo tiene la solución $(\lambda, \mu) = (0,0)$, con lo que hemos llegado a dos vectores linealmente independientes. Por lo tanto $(1,1,1)$ y $(1,2,3)$ son una base de W, que es, por consiguiente, un espacio vectorial de dimensión 2. ■

Este proceso de ir eliminando vectores puede resultar muy largo si no somos capaces de descartar rápidamente algunos de ellos. Por ello hay otros procedimientos que conviene conocer.

Rango Se llama *rango* de un conjunto de vectores $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ de un espacio vectorial V a la dimensión del subespacio que generan:

$$\operatorname{rang}\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\} = \dim \langle \bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n \rangle ,$$

es decir, el rango es el número de vectores que quedan después de eliminar todos los que dependen de los demás.

Puede probarse que si A es una matriz, entonces el rango de sus vectores fila coincide necesariamente con el rango de sus vectores columna, y a este rango común es simplemente el rango de A.

Para calcular el rango de una matriz basta triangularla. La matriz triangulada tendrá el mismo rango que la matriz original, y éste coincidirá con el número de filas no nulas que hayan quedado.

Si no alteramos el orden de las filas, los vectores que en la matriz original ocupen la misma posición que las filas no nulas de la matriz triangulada serán linealmente independientes.

Veamos un ejemplo:

Ejemplo *Un banco gestiona cuatro carteras A, B, C y D, cuyo capital se distribuye en acciones de tres compañías P, Q y R en la proporción que indica la tabla siguiente:*

Compañías	Carteras			
	A	B	C	D
P	20%	30%	25%	24%
Q	0%	60%	30%	24%
R	80%	10%	45%	52%

El banco desea ampliar el capital de sus carteras. Comprueba que los incrementos $(\Delta P, \Delta Q, \Delta R)$ de la inversión del banco en cada compañía que pueden distribuirse entre las cuatro carteras sin crear excedentes es un subespacio vectorial de $\mathbb{R}^3$. Calcula su dimensión.

SOLUCIÓN: Un incremento ΔA en el capital de la cartera A requiere que el banco compre (o venda) acciones por una cuantía de

$$(0.2\,\Delta A, 0, 0.8\,\Delta A) = \Delta A\,(0.2, 0, 0.8).$$

Similarmente, unos incrementos ΔB, ΔC y ΔD en las otras carteras se corresponden con incrementos del capital invertido en cada compañía dados por $\Delta B\,(0.3, 0.6, 0.1)$, $\Delta C\,(0.25, 0.3, 0.45)$ y $\Delta D\,(0.24, 0.24, 0.52)$, respectivamente. Entonces, un incremento arbitrario de las cuatro carteras $(\Delta A, \Delta B, \Delta C, \Delta D)$ requerirá una compra (o venta) de acciones dada por el vector

$$\begin{aligned}(\Delta P, \Delta Q, \Delta R) = \quad & \Delta A\,(0.2, 0, 0.8) + \Delta B\,(0.3, 0.6, 0.1) \\ & +\Delta C\,(0.25, 0.3, 0.45) + \Delta D\,(0.24, 0.24, 0.52).\end{aligned}$$

Por consiguiente, los incrementos factibles forman el espacio vectorial

$$W = \langle (0.2, 0, 0.8), (0.3, 0.6, 0.1), (0.25, 0.3, 0.45), (0.24, 0.24, 0.52) \rangle\,.$$

Tenemos un sistema generador de W. Hemos de estudiar si hay vectores dependientes y, en tal caso, eliminarlos. Para hacerlo rápidamente formamos la matriz que tiene por filas estos vectores y la triangulamos (primeramente multiplicamos y dividimos sus filas por números adecuados para simplificar los coeficientes):

$$A = \begin{pmatrix} 0.2 & 0 & 0.8 \\ 0.3 & 0.6 & 0.1 \\ 0.25 & 0.3 & 0.45 \\ 0.24 & 0.24 & 0.52 \end{pmatrix} \sim \begin{pmatrix} 2 & 0 & 8 \\ 3 & 6 & 1 \\ 25 & 30 & 45 \\ 24 & 24 & 52 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 4 \\ 3 & 6 & 1 \\ 5 & 6 & 9 \\ 6 & 6 & 13 \end{pmatrix}$$

$$\sim \begin{pmatrix} 1 & 0 & 4 \\ 0 & 6 & -11 \\ 0 & 6 & -11 \\ 0 & 6 & -11 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 4 \\ 0 & 6 & -11 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Concluimos que el rango de la matriz es 2 y —como no hemos cambiado el orden de las filas— que los vectores $(0.2, 0, 0.8)$ y $(0.3, 0.6, 0.1)$ son linealmente independientes. Así pues,

$$W = \langle (0.2, 0, 0.8), (0.3, 0.6, 0.1) \rangle ,$$

estos vectores forman una base y la dimensión de W es 2. ∎

Recordemos otra forma de calcular mediante determinantes: El rango de una matriz A es el mayor número n tal que A contiene una submatriz cuadrada $n \times n$ con determinante no nulo.

Ejemplo *Determina si los vectores* $(2, 3, -1, 4)$ *y* $(4, 6, 0, 1)$ *son linealmente dependientes o independientes.*

SOLUCIÓN: Como el rango de n vectores es el número de vectores independientes que hay entre ellos, resulta que n vectores son independientes si y sólo si su rango es n. En nuestro caso, los dos vectores dados serán independientes si y sólo si su rango es 2. Esto equivale a que la matriz

$$A = \begin{pmatrix} 2 & 3 & -1 & 4 \\ 4 & 6 & 0 & 1 \end{pmatrix}$$

tenga rango 2. Como el determinante formado por las dos últimas columnas vale $-1 \neq 0$, ciertamente el rango es 2 y los vectores son independientes. ∎

Propiedades de las bases Veamos algunos hechos adicionales sobre bases en un espacio vectorial. En primer lugar calculamos la dimensión de $\mathbb{R}^n$:

Teorema *La dimensión de* $\mathbb{R}^n$ *es* n*. Una base de* $\mathbb{R}^n$ *(la llamada* base canónica*) está formada por los* n *vectores que tienen una coordenada igual a* 1 *y las restantes nulas.*

Por ejemplo, la base canónica de $\mathbb{R}^2$ es la formada por los vectores $(0, 1)$ y $(1, 0)$, la base canónica de $\mathbb{R}^3$ es la formada por $(1, 0, 0)$, $(0, 1, 0)$ y $(0, 0, 1)$, etc.

El primero de los dos teoremas siguientes lo conocíamos ya, pero lo incluimos para compararlo con el segundo:

Teorema *Todo sistema generador de un espacio vectorial puede convertirse en una base eliminando parte de sus vectores.* (Concretamente, hay que eliminar vectores que sean linealmente dependientes de los demás hasta que no quede ninguno en estas condiciones.)

Teorema *Todo sistema libre de un espacio vectorial puede convertirse en una base añadiendo vectores.* (Concretamente, podemos llegar hasta una base añadiendo vectores de una base prefijada del espacio que sean linealmente independientes de los del sistema dado, y repitiendo el proceso hasta tener tantos vectores como requiera la dimensión del espacio.)

De estos teoremas se deduce a su vez una consecuencia interesante. En principio, para que un sistema de vectores sea base de un espacio V de dimensión n ha de cumplir dos condiciones: que sea un sistema generador y un sistema libre. Ahora bien, si el sistema tiene precisamente n vectores basta con que cumpla una de las dos condiciones para que necesariamente cumpla la otra: si es linealmente independiente tiene que ser base, porque de lo contrario podríamos convertirlo en una base añadiendo vectores, pero ya tiene los vectores justos para ser base y no es posible añadir más. Igualmente, si es un sistema generador tiene que ser base, porque podría convertirse en una base quitando vectores, pero no podemos quitar ninguno porque ya tiene el número justo para ser base. Esto debemos recordarlo:

Un sistema de n vectores en un espacio vectorial de dimensión n es una base si es un sistema libre o un sistema generador (basta que cumpla una de las dos propiedades para que necesariamente cumpla la otra).

Ejemplo *Determina si los vectores $(1,2,1)$, $(-1,0,2)$ y $(2,1,1)$ son una base del espacio vectorial $\mathbb{R}^3$.*

Solución: Como la dimensión de $\mathbb{R}^3$ es 3 y tenemos 3 vectores, no es necesario comprobar que es un sistema generador y libre, sino que basta comprobar una de las dos propiedades. Lo más fácil es comprobar que es libre, pues eso es tanto como ver que tienen rango 3, lo cual equivale a ver que la matriz

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

tiene rango 3. Para ello basta comprobar que su determinante es no nulo, pero $|A| = 7 \neq 0$. ■

Ejemplo *Comprueba que el vector $(4,3,0)$ pertenece al espacio*

$$V = \langle (2,1,-2), (1,1,1) \rangle$$

y encuentra una base de V que lo contenga.

Solución: En primer lugar observamos que los dos vectores que generan V son, de hecho, una base de V, es decir, son linealmente independientes. Para ello basta ver que el rango de la matriz

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

es 2, lo cual se cumple porque el determinante formado por las dos primeras columnas es no nulo. Así pues, $\dim V = 2$. El vector $(4,3,0)$ estará en V si es combinación lineal de la base, es decir, si el rango del sistema $(4,3,0)$, $(2,1,-2)$, $(1,1,1)$ sigue siendo 2. Para ello calculamos el rango de la matriz

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

Se comprueba que $|B| = 0$, por lo que, ciertamente $\operatorname{rang} B = 2$ y $(4, 3, 0)$ está en el espacio V.

Ahora queremos completar el sistema libre formado por $(4, 3, 0)$ a una base de V. Como la dimensión de V es 2, sabemos que nos falta añadir 1 vector. Nos servirá cualquiera de la base que tenemos de V con la condición de que sea linealmente independiente de $(4, 3, 0)$. Probamos con $(2, 1, -2)$: el rango de la matriz

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

es 2, pues el determinante formado por las dos primeras columnas es no nulo. Los vectores son independientes y, por consiguiente, una base de V es la formada por $(4, 3, 0)$ y $(2, 1, -2)$. ■

El método que hemos empleado para determinar si $(4, 3, 0)$ estaba o no en el espacio V es general:

Para comprobar si un vector de un espacio vectorial pertenece a un subespacio del que conocemos una base basta ver si al añadir el vector a la base el rango no aumenta.

Ejemplo *Una empresa distribuye tres productos A, B y C en cantidades x, y, z en dos mercados M y N. Los costes unitarios de distribución de cada producto en cada mercado vienen dados por la tabla siguiente:*

Mercados	Productos		
	A	B	C
M	1	2	1
N	1	3	1

La empresa estudia reestructurar su producción en cantidades $(\Delta x, \Delta y, \Delta z)$. Sean W_1 y W_2 los conjuntos de los incrementos de producción $(\Delta x, \Delta y, \Delta z)$ que no alteran los costes de distribución en el mercado M y N respectivamente, y sea W el conjunto de incrementos de producción que no altera los costes de distribución en ninguno de los dos mercados.

1. *Comprueba que W_1, W_2 y W son subespacios vectoriales de $\mathbb{R}^3$.*

2. *Calcula bases de cada uno de ellos.*

3. *Encuentra bases de W_1 y W_2 que completen la base encontrada para W.*

Solución: Si un incremento de producción $(\Delta x, \Delta y, \Delta z)$ se distribuye en el mercado M, genera un incremento de costes de distribución dado por

$$\Delta x + 2\Delta y + \Delta z,$$

luego $W_1 = \{(\Delta x, \Delta y, \Delta z) \in \mathbb{R}^3 \mid \Delta x + 2\Delta y + \Delta z = 0\}$. Similarmente concluimos que $W_2 = \{(\Delta x, \Delta y, \Delta z) \in \mathbb{R}^3 \mid \Delta x + 3\Delta y + 2\Delta z = 0\}$. Por último, para que no se alteren los costes de distribución en ambos mercados se han de cumplir ambas condiciones al mismo tiempo, luego

$$W = \{(\Delta x, \Delta y, \Delta z) \in \mathbb{R}^3 \mid \Delta x + 2\Delta y + \Delta z = 0,\ \Delta x + 3\Delta y + 2\Delta z = 0\}.$$

1. Como los tres conjuntos están definidos por ecuaciones lineales homogéneas, son subespacios vectoriales de $\mathbb{R}^3$.

2. Para encontrar un sistema generador de W_1 resolvemos $\Delta x + 2\Delta y + \Delta z = 0$. Tomamos como parámetros $\alpha = \Delta y$, $\beta = \Delta z$ y tenemos la solución

$$\Delta x + 2\Delta y + \Delta z = 0 \quad \Longrightarrow \quad (\Delta x, \Delta y, \Delta z) = (-2\alpha - \beta, \alpha, \beta), \quad \alpha, \beta \in \mathbb{R}.$$

Entonces,

$$(\Delta x, \Delta y, \Delta z) = (-2\alpha - \beta, \alpha, \beta) = (-2\alpha, \alpha, 0) + (-\beta, 0, \beta) = \alpha(-2, 1, 0) + \beta(-1, 0, 1).$$

Por consiguiente, los vectores de W_1 son las combinaciones lineales de $(-2, 1, 0)$ y $(-1, 0, 1)$, es decir, $W_1 = \langle(-2, 1, 0), (-1, 0, 1)\rangle$. Es claro que estos dos vectores son linealmente independientes, luego forman una base de W_1.

Razonamos igualmente con W_2: tenemos $\Delta x + 3\Delta y + 2\Delta z = 0$, hacemos $\Delta y = \alpha$, $\Delta z = \beta$ y obtenemos

$$(\Delta x, \Delta y, \Delta z) = (-3\alpha - 2\beta, \alpha, \beta) = \alpha(-3, 1, 0) + \beta(-2, 0, 1), \quad \alpha, \beta \in \mathbb{R}.$$

En conclusión, $W_2 = \langle(-3, 1, 0), (-2, 0, 1)\rangle$. Estos vectores también son linealmente independientes, luego son una base de W_2.

Para encontrar una base de W resolvemos:

$$\left.\begin{array}{r} \Delta x + 2\Delta y + \ \ \Delta z = 0 \\ \Delta x + 3\Delta y + 2\Delta z = 0 \end{array}\right\} \Rightarrow \quad \left.\begin{array}{r} \Delta x + 2\Delta y + \Delta z = 0 \\ \Delta y + \Delta z = 0 \end{array}\right\}$$

Hacemos $\Delta z = \alpha$, con lo que $\Delta y = -\alpha$, $\Delta x = 2\alpha - \alpha = \alpha$. La solución es

$$(\Delta x, \Delta y, \Delta z) = (\alpha, -\alpha, \alpha) = \alpha(1, -1, 1), \qquad \text{para todo } \alpha \in \mathbb{R}.$$

Por lo tanto $W = \langle(1, -1, 1)\rangle$.

3. Para extender $(1, -1, 1)$ a una base de W_1 hemos de añadir 1 vector. Nos sirve por ejemplo la base $W_1 = \langle(1, -1, 1), (-2, 1, 0)\rangle$, pues los dos vectores son linealmente independientes. Similarmente, $W_2 = \langle(1, -1, 1), (-3, 1, 0)\rangle$. ■

La propiedad fundamental que cumplen las bases y a la que deben en gran parte su interés es la siguiente:

Teorema *Si $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ es una base de un espacio vectorial V, entonces cada vector $\bar{v} \in V$ se expresa de forma única como combinación lineal*

$$\bar{v} = \lambda_1 \bar{v}_1 + \lambda_2 \bar{v}_2 + \cdots + \lambda_n \bar{v}_n.$$

Los coeficientes $\lambda_1, \lambda_2, \ldots, \lambda_n$ se llaman coordenadas *de $\bar{v}$ en la base $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$.*

Ejemplo *Una empresa fabrica dos artículos A y B a partir de dos materias primas P y Q. Cada unidad de producto requiere las cantidades que indica la tabla siguiente:*

Materias primas	Artículos A	Artículos B
P	2	3
Q	1	2

La empresa dispone de un stock de 21 unidades de P y 13 unidades de Q. Demostrar que los vectores $(2,1)$ y $(3,2)$ forman una base de $\mathbb{R}^2$, calcular en dicha base las coordenadas del vector $(21,13)$ y explicar la interpretación de dichas coordenadas.

Solución: Los vectores $(2,1)$ y $(3,2)$ son linealmente independientes porque

$$\begin{vmatrix} 2 & 1 \\ 3 & 2 \end{vmatrix} = 1 \neq 0.$$

Como son dos vectores y la dimensión de $\mathbb{R}^2$ es 2, podemos concluir que son una base. Las coordenadas de $(21,13)$ en dicha base son los escalares (λ,μ) que cumplen

$$\lambda(2,1) + \mu(3,2) = (21,13).$$

Igualando componentes esto equivale al sistema

$$\left.\begin{array}{r} 2\lambda + 3\mu = 21 \\ \lambda + 2\mu = 13 \end{array}\right\} \Rightarrow \left.\begin{array}{r} 2\lambda + 3\mu = 21 \\ \mu = 5 \end{array}\right\} \Rightarrow \lambda = 3.$$

Así pues, las coordenadas son $(\lambda,\mu) = (3,5)$, luego tenemos que

$$3 \cdot (2,1) + 5 \cdot (3,2) = (21,13),$$

luego vemos que, con su stock, la empresa puede fabricar 3 unidades de A y 5 unidades de B, y ésta es la única opción con la que no tiene excedentes. ∎

Subespacios definidos por ecuaciones Para determinar la dimensión de un subespacio W de $\mathbb{R}^n$ definido por ecuaciones lineales no es necesario calcular una base.

$$\boxed{\text{Dim } W = n - \textit{ número de ecuaciones independientes.}}$$

El número de ecuaciones independientes ha de entenderse como el rango de la matriz de coeficientes del sistema de ecuaciones.

Ejemplo *Tres empresas A, B y C que producen bienes b_1, b_2 y b_3 respectivamente, se intercambian la producción de la forma siguiente: A reparte su producción en partes iguales, B cede $2/5$ de su producción a A y se queda con el resto y C cede $1/4$ a A y utiliza el resto internamente. Demuestra que las producciones que permiten que ninguna empresa pague a las restantes es un espacio vectorial de dimensión 1.*

SOLUCIÓN: Si llamamos (x, y, z) a la producción de (b_1, b_2, b_3), la forma de distribuir la producción se expresa con el sistema siguiente:

$$\left.\begin{array}{rcl} \frac{1}{3}x + \frac{2}{5}y + \frac{1}{4}z & = & x \\ \frac{1}{3}x + \frac{3}{5}y \phantom{+ \frac{1}{4}z} & = & y \\ \frac{1}{3}x \phantom{+ \frac{2}{5}y} + \frac{3}{4}z & = & z \end{array}\right\} \Rightarrow \left.\begin{array}{rcl} -40x + 24y + 15z & = & 0 \\ 5x - \ 6y & = & 0 \\ 4x - 3z & = & 0 \end{array}\right\}$$

Entonces el conjunto de producciones que permiten que ninguna empresa pague al resto es

$$W = \{(x, y, z) \in \mathbb{R}^3 \mid -40x + 24y + 15z = 0,\ 5x - 6y = 0,\ 4x - 3z = 0\}.$$

Como está dado por tres ecuaciones lineales homogéneas es un subespacio vectorial. Para calcular la dimensión de W sólo necesitamos conocer el rango de la matriz del sistema

$$A = \begin{pmatrix} -40 & 24 & 15 \\ 5 & 6 & 0 \\ 4 & 0 & -3 \end{pmatrix}$$

Tenemos $|A| = 0$, por tanto el rango de A es a lo sumo 2. Como

$$\begin{vmatrix} -40 & 24 \\ 5 & -6 \end{vmatrix} = 120 \neq 0,$$

el rango de A es 2, es decir que hay 2 ecuaciones independientes. Aplicando el resultado anterior $\dim W = 3 - 2 = 1$. ∎

Ejemplo *Expresa el subespacio $W = \big\langle (1, 0, 0, 1), (1, -2, 3, 0), (1, 2, -3, 2) \big\rangle$ mediante ecuaciones.*

SOLUCIÓN: En primer lugar calculamos una base de W, es decir, eliminamos de $\{(1, 0, 0, 1), (1, -2, 3, 0), (1, 2, -3, 2)\}$ los vectores que sean ligados (si los hay). Para ello calculamos el rango de la matriz

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

Orden 1: $\quad |1| = 1 \neq 0 \quad \Longrightarrow \quad \operatorname{rang} A \geq 1,$

Orden 2: $\quad \begin{vmatrix} 1 & 0 \\ 1 & -2 \end{vmatrix} = 1 \neq 0 \quad \Longrightarrow \quad \operatorname{rang} A \geq 2,$

Orden 3: $\begin{vmatrix} 1 & 0 & 0 \\ 1 & -2 & 3 \\ 1 & 2 & -3 \end{vmatrix} = 0, \quad \begin{vmatrix} 1 & 0 & 1 \\ 1 & -2 & 0 \\ 1 & 2 & 2 \end{vmatrix} = 0.$

Entonces, la matriz A tiene rango 2 y sabemos que dim W=2. Como los vectores $(1,0,0,1),(1,-2,3,0)$ son libres forman una base de W.

Para expresar W con ecuaciones, tenemos que encontrar todos los vectores $(x,y,z,t)\in\mathbb{R}^4$ que pertenezcan a $\big\langle(1,0,0,1),(1,-2,3,0)\big\rangle$, es decir aquellos vectores que hacen que la matriz

$$\begin{pmatrix} x & y & z & t \\ 1 & 0 & 0 & 1 \\ 1 & -2 & 3 & 0 \end{pmatrix}$$

tenga rango 2. Y precisamente esta condición sobre el rango será la que proporcione las ecuaciones:

$$\begin{vmatrix} x & y & z \\ 1 & 0 & 0 \\ 1 & -2 & 3 \end{vmatrix} = 0 \quad\Longrightarrow\quad -3y-2z=0$$

$$\begin{vmatrix} x & y & t \\ 1 & 0 & 1 \\ 1 & -2 & 0 \end{vmatrix} = 0 \quad\Longrightarrow\quad 2x-y-2t=0$$

Entonces, $W=\{(x,y,z,t)\in\mathbb{R}^4 \mid -3y-2z=0, \quad 2x-y-2t=0\}$. ■

3.5 Ejercicios

1. Las existencias en un almacén están dadas por un vector $\bar{e}\in\mathbb{R}^n$, donde n es el número de productos distintos almacenados y cada componente e_i es la cantidad de producto i-ésimo almacenado. Si $\bar{p}\in\mathbb{R}^n$ es el vector de precios y $V\in\mathbb{R}$ es el valor total de la mercancía almacenada, expresa la relación matemática entre V, $\bar{e}$ y $\bar{p}$.

2. Una empresa fabrica dos artículos A y B. Las ventas del último semestre vienen dadas por la tabla siguiente:

Artículos	Meses					
	J	A	S	O	N	D
A	3	0	9	15	10	12
B	5	0	6	17	11	9

Determina, realizando siempre operaciones con vectores de $\mathbb{R}^2$,

(a) El vector $\bar{v}$ de ventas totales en el semestre.

(b) El vector $\bar{v}_m$ de ventas medias mensuales en dicho periodo.

(c) El vector $\Delta\bar{v}$ de incremento de ventas de noviembre a diciembre.

Si el vector de precios es $\bar{p} = (3, 5)$, calcula —siempre con operaciones vectoriales— el beneficio de la empresa en el semestre.

3. Determina si los conjuntos de vectores siguientes son libres o ligados

   (a) $\{(1, 2, 1), (0, 0, 0), (1, 2, 3)\}$.

   (b) $\{(1, 2, 0, 1), (0, 1, 0, 0), (1, 2, 3, 1)\}$.

   (c) $\{(1, 2, 0, 1), (0, 1, 0, 0), (1, 2, 3, 1), (1, 2, 4, 5), (1, 1, 1, 3)\}$.

   (d) $\{(1, 2, 0, 1), (0, 1, 0, 0), (1, 2, 3, 1), (0, 0, 0, 1)\}$.

4. Calcula los valores de z para los cuales el vector $(1, z, 3)$ es combinación lineal de los vectores $(1, 2, 1), (1, 0, 2)$. Determina los coeficientes de la combinación lineal.

5. Razona las afirmaciones siguientes:

   (a) $A = \{(x, y, z) \in \mathbb{R}^3 \mid x = 2z\}$ es subespacio vectorial de $\mathbb{R}^3$.

   (b) $B = \{(x, y, z, t) \in \mathbb{R}^4 \mid x^2 - y = 0\}$ no es subespacio vectorial de $\mathbb{R}^4$.

   (c) $C = \{(x, y, z) \in \mathbb{R}^3 \mid -y + z = 0, \quad x - z = 0\}$ es subespacio de $\mathbb{R}^3$.

   (d) $D = \{(x, y) \in \mathbb{R}^2 \mid x + y = 1\}$ no es subespacio vectorial de $\mathbb{R}^2$.

6. Comprueba si los vectores $(-1, 2, -5)$, $(2, 2, 3)$, $(-1, -4, 1)$ están en el subespacio $W =< (1, 2, 1), (1, 0, 3) >$.

7. Determina el valor de a para que el vector $(2, a, 1, 0)$ esté en el subespacio $W =< (1, 0, 2, 1), (1, 1, -1, -1) >$.

8. Expresa los siguientes subespacios vectoriales como conjuntos engendrados por familias de vectores:

   (a) $V_1 = \{(x, y, z) \in \mathbb{R}^3 \mid x = -z\}$.

   (b) $V_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y = 0, \quad x + y + z = 0\}$.

   (c) $V_3 = \{(x, y, z, t) \in \mathbb{R}^4 \mid x - 2y = 0, \quad x - z + t = 0\}$.

9. Expresa mediante conjuntos dados por ecuaciones los subespacios siguientes:

   (a) $A = \big\langle (1, -2, 1), (1, 0, 2) \big\rangle$.

   (b) $B = \big\langle (1, 0, 1, 3) \big\rangle$.

   (c) $B = \big\langle (1, 0, 1, -1), (0, 1, 1, 0) \big\rangle$.

10. Halla una base de $\mathbb{R}^3$ que contenga los vectores $(1, -2, 1), (1, 0, 2)$.

11. Calcula una base de $\big\langle (1, -2, 1), (1, 0, 2) \big\rangle$ que contenga al vector $(2, -2, 3)$.

12. En una economía formada por tres productores, A, B y C, se sabe que la matriz de input-output es

$$A = \begin{pmatrix} 0.25 & 0.25 & 0 \\ 0.25 & 0.25 & 0.2 \\ 0.5 & 0.5 & 0.8 \end{pmatrix}.$$

Si esta economía responde a un modelo cerrado de Leontief, razona que el conjunto de producciones que dan lugar al equilibrio es un subespacio vectorial.

13. Si M es la matriz input-output del modelo de Leontief razona si el conjunto de producciones que dan lugar al equilibrio son un subespacio vectorial.

4. Aplicaciones lineales

Las aplicaciones lineales expresan la dependencia más simple que puede darse desde un punto de vista matemático entre distintas magnitudes. En este tema veremos que muchos de los resultados que hemos visto sobre espacios vectoriales se expresan más claramente en términos de estas aplicaciones, y en capítulos posteriores veremos que el estudio de otras dependencias más complejas entre unas magnitudes dadas puede aproximarse en muchos casos mediante aplicaciones lineales, con lo que los resultados que veremos aquí serán igualmente aplicables.

4.1 Definición y propiedades básicas

Empezamos introduciendo la noción de aplicación lineal:

Definición Una aplicación $f : V \longrightarrow W$ entre dos espacios vectoriales es *lineal* si cumple las propiedades siguientes:

1. $f(\bar{u} + \bar{v}) = f(\bar{u}) + f(\bar{v})$, para todo $\bar{u}$, $\bar{v} \in V$,
2. $f(\lambda\bar{v}) = \lambda f(\bar{v})$, para todo $\lambda \in \mathbb{R}$, $\bar{v} \in V$.

Ejemplo *Comprueba que la aplicación $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por*

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

es una aplicación lineal.

SOLUCIÓN: Se trata de comprobar (1) y (2) de la definición.

Veamos que se verifica (1):

Consideramos $\bar{u} = (x, y), \bar{v} = (x', y')$ dos vectores arbitrarios de $\mathbb{R}^2$.

$$\begin{aligned} f(\bar{u} + \bar{v}) &= f\big((x, y) + (x', y')\big) \\ &= f\big(x + x', y + y'\big) = \big(x + x' + y + y', y + y', -(x + x') + 2(y + y')\big) \\ &= (x + x' + y + y', y + y', -x - x' + 2y + 2y') \end{aligned}$$

$$\begin{aligned} f(\bar{u}) &= f(x,y) = (x+y, y, -x+2y) \\ f(\bar{v}) &= f(x',y') = (x'+y', y', -x'+2y') \end{aligned}$$

Sumando las dos expresiones anteriores tenemos

$$\begin{aligned} f(\bar{u}) + f(\bar{v}) &= f(x,y) + f(x',y') = (x+y, y, -x+2y) + (x'+y', y', -x'+2y') \\ &= (x+y+x'+y', y+y', -x+2y-x'+2y') \\ &= (x+x'+y+y', y+y', -x-x'+2y+2y') \end{aligned}$$

Con lo cual se tiene que $f(\bar{u}+\bar{v}) = f(\bar{u}) + f(\bar{v})$.

Veamos que se verifica (2):

Dado un vector arbitrario $\bar{u} = (x,y) \in \mathbb{R}^2$ y cualquier escalar $\lambda \in \mathbb{R}$, se tiene

$$f(\lambda\bar{u}) = f\big(\lambda(x,y)\big) = f(\lambda x, \lambda y) = (\lambda x + \lambda y, \lambda y, -\lambda x + 2\lambda y).$$

Por otra parte tenemos que

$$f(\bar{u}) = f(x,y) = (x+y, y, -x+2y) \Longrightarrow \lambda f(\bar{u}) = (\lambda x + \lambda y, \lambda y, -\lambda x + \lambda 2y),$$

entonces $f(\lambda\bar{u}) = \lambda f(\bar{u})$. ∎

Recordemos un par de propiedades elementales de las aplicaciones lineales.

Propiedades Dada $f : V \longrightarrow W$ una aplicación se tiene:

a) $f(\bar{0}) = \bar{0}$

b) Las aplicaciones lineales conservan las combinaciones lineales, es decir, si

$$\bar{v} = \lambda_1\bar{v}_1 + \lambda_2\bar{v}_2 + \cdots + \lambda_n\bar{v}_n \in V,$$

entonces

$$f(\bar{v}) = \lambda_1 f(\bar{v}_1) + \lambda_2 f(\bar{v}_2) \cdots + \lambda_n f(\bar{v}_n) \in W.$$

El teorema siguiente nos muestra cómo reconocer fácilmente las aplicaciones lineales de $\mathbb{R}^n$ en $\mathbb{R}^m$:

Teorema *Una aplicación $f : V \longrightarrow W$ es lineal si y sólo si puede expresarse en la forma*

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n,$$

para ciertos números reales $a_1, a_2, \ldots a_n$. Es decir, si la imagen es combinación lineal de las variables

Demostración:

[⇒] Suponemos que f es lineal. Consideramos la base canónica $\bar{e}_1, \bar{e}_2, \ldots, \bar{e}_n$ de $\mathbb{R}^n$. Entonces, todo vector $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ se expresa como combinación lineal

$$(x_1, x_2, \ldots, x_n) = x_1\bar{e}_1 + x_2\bar{e}_2 + \cdots + x_n\bar{e}_n.$$

Como las aplicaciones lineales conservan las combinaciones lineales de vectores, se cumple que

$$f(x_1, x_2, \ldots, x_n) = x_1 f(\bar{e}_1) + x_2 f(\bar{e}_2) + \cdots + x_n f(\bar{e}_n),$$

y llamando $a_1 = f(\bar{e}_1)$, $a_2 = f(\bar{e}_2), \ldots, a_n = f(\bar{e}_n)$ queda

$$f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n.$$

[$\Leftarrow$] Si f es de la forma $f(x_1, x_2, \ldots, x_n) = a_1x_1 + a_2x_2 + \cdots + a_nx_n$, vamos a comprobar que es lineal mediante la definición.

Sean $\bar{x} = (x_1, x_2, \ldots, x_n)$, $\bar{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Hemos de probar:

(1) $f(\bar{x} + \bar{y}) = f(\bar{x}) + f(\bar{y})$. En efecto:

$$f(\bar{x} + \bar{y}) = f(x_1 + y_1, \ldots, x_n + y_n) = a_1(x_1 + y_1) + \cdots + a_n(x_n + y_n).$$

Por otra parte,

$$f(\bar{x}) + f(\bar{y}) = a_1x_1 + \cdots + a_nx_n + a_1y_1 + \cdots a_ny_n.$$

Es claro que ambas expresiones son iguales.

(2) $f(\lambda\bar{x}) = \lambda f(\bar{x})$. En efecto:

$$f(\lambda\bar{x}) = f(\lambda x_1, \ldots, \lambda x_n) = a_1\lambda x_1 + \cdots + a_n\lambda x_n.$$

Por otra parte,

$$\lambda f(\bar{x}) = \lambda(a_1x_1 + \cdots + a_nx_n) = \lambda a_1x_1 + \cdots + \lambda a_nx_n.$$

De nuevo vemos que ambas expresiones coinciden. ■

El caso general de una aplicación $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ se reduce al teorema anterior, pues

La aplicación $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es lineal si y sólo si es de la forma

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)),$$

donde cada $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una aplicación lineal.

Ejemplo *Indica si la aplicación $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ dada por*

$$f(x, y) = (x - 3y, 2x + y, 0)$$

es lineal y explica por qué.

Solución: Llamemos $f_1(x, y) = x - 3y$, $f_2(x, y) = 2x + y$, $f_3(x, y) = 0$. Las aplicaciones $f_i : \mathbb{R}^2 \longrightarrow \mathbb{R}$ son lineales porque las imágenes son combinación lineal de las variables. (Respecto a la tercera, notemos que $f_3(x, y) = 0 \cdot x + 0 \cdot y$.) Como $f(x, y) = (f_1(x, y), f_2(x, y), f_3(x, y))$, el teorema anterior nos permite concluir que f también es lineal. ■

Ejemplo *Supongamos que cada entrada de cine cuesta* 5 € *y cada libro vale* 6 €. *Sean* ΔC *y* ΔL *los incrementos de entradas de cine y libros vendidos en una cierta ciudad respecto del año pasado. Comprueba que la función* $f(\Delta C, \Delta L)$ *que da el incremento de gasto conjunto de estos bienes es lineal. ¿Cuál será el incremento de gasto si el número de libros vendidos ha pasado de* 1000 *a* 300 *y el número de entradas de cine ha pasado de* 50.000 *a* 80.000*?*

SOLUCIÓN: Por cada entrada que se compra de más, el gasto en cine aumenta en 5 €, luego el gasto total en cine aumenta en $5\Delta C$. Similarmente, el gasto en libros aumenta en $6\Delta L$ y el aumento conjunto de gasto en ambos bienes es

$$f(\Delta C, \Delta L) = 5\Delta C + 6\Delta L.$$

Tenemos que $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ es lineal, pues la imagen es combinación lineal de las variables.

La segunda parte del enunciado nos dice que $\Delta C = 80000 - 50000 = 30000$ € y que $\Delta L = 300 - 1000 = -700$ €. Por consiguiente,

$$f(\Delta C, \Delta L) = 5 \cdot 30000 + 6(-700) = 150000 - 4200 = 145800 \text{ €}.$$

■

El teorema siguiente muestra que una aplicación lineal $f : V \longrightarrow W$ está determinada por una base de V, una base de W y una matriz asociada:

Teorema *Supongamos que* $f : V \longrightarrow W$ *es una aplicación lineal,* $\{\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n\}$ *es una base de* V *y* $\{\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_m\}$ *es una base de* W. *Sea* A *la matriz* $m \times n$ *cuya columna i-ésima está formada por las coordenadas de* $f(\bar{v}_i)$ *en la base de* W. *Entonces, si* $\bar{v} \in V$ *es un vector de coordenadas* $\bar{x} = (\alpha_1, \alpha_2, \ldots, \alpha_m)$ *en la base de* V*, el vector de coordenadas* $\bar{y}$ *de* $f(\bar{v})$ *en la base de* W *es el determinado por la relación* $\beta^t = A\alpha^t$.

La matriz A del teorema anterior se llama *matriz* de la aplicación f en las bases dadas. El teorema puede resumirse así:

La matriz de una aplicación lineal en unas bases dadas se caracteriza por que cuando se multiplica por las coordenadas de un vector (puesto en columna) se obtienen las coordenadas de la imagen del vector.

En particular, si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una aplicación lineal y A es su matriz en las bases canónicas, entonces, teniendo en cuenta que un vector coincide con sus coordenadas en la base canónica, la propiedad anterior se reduce a que

$$f(\bar{v})^t = A\bar{v}^t, \qquad \text{para todo } \bar{v} \in \mathbb{R}^n.$$

Cuando se habla de la matriz de una aplicación lineal $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ sin especificar bases, se entiende que se trata de la matriz respecto de las bases canónicas. Así pues,

Si $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una aplicación lineal y A es su matriz, entonces f está determinada por la relación $f(\bar{v})^t = A\bar{v}^t$, para todo $\bar{v} \in \mathbb{R}^n$.

Ejemplo *Calcula la matriz de la aplicación lineal* $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *dada por* $f(x,y) = (x-2y, 0, 2x+y)$.

SOLUCIÓN: Hemos de entender que se pide la matriz de f respecto a las bases canónicas. Para obtenerla tomamos la base canónica de $\mathbb{R}^2$, que es $(1,0)$, $(0,1)$ y calculamos las imágenes de sus vectores:

$$f(1,0) = (1,0,2), \quad f(0,1) = (-2,0,1).$$

La matriz de f es la que tiene a estas imágenes por columnas:

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

$\blacksquare$

Operaciones entre aplicaciones lineales

a) Dadas dos aplicaciones lineales $f, g : V \longrightarrow W$, las aplicaciones $f+g$ y $f-g$ son también aplicaciones lineales. Además si A y B son las matrices de f y g para unas bases dadas, la matriz de $f+g$ en estas bases es $A+B$ y la de $f-g$ es $A-B$.

b) Si $f : V \longrightarrow W$ es aplicación lineal, dado $\alpha \in \mathbb{R}$ la aplicación αf también es lineal. La matriz de αf es αA, donde A es la matriz de f.

c) Dadas dos aplicaciones lineales $f : V \longrightarrow W$ y $g : W \longrightarrow U$ la aplicación $g \circ f$, dada por $(g \circ f)(\bar{x}) = g(f(\bar{x}))$, también es lineal. Además, si A y B son las matrices de f y g respectivamente, la matriz de $g \circ f$ es BA.

4.2 Núcleo e imagen de una aplicación lineal

En el tema anterior hemos visto dos formas de expresar un subespacio vectorial de $\mathbb{R}^n$: mediante un sistema de ecuaciones y mediante un sistema generador. Ambas formas están relacionadas con las aplicaciones lineales. Para verlo introducimos dos conceptos:

Definición Si $f : V \longrightarrow W$ es una aplicación lineal entre espacios vectoriales, se llama *núcleo* de f al conjunto

$$\ker f = \{\bar{v} \in V \mid f(\bar{v}) = \bar{0}\} \subset V.$$

(Se abrevia "ker" por el inglés kernel.) La *imagen* de f es el conjunto

$$\operatorname{Im} f = \{f(\bar{v}) \mid \bar{v} \in V\} \subset W.$$

Teorema *El núcleo y la imagen de una aplicación lineal $f : V \longrightarrow W$ son subespacios vectoriales de V y W, respectivamente.*

Por ejemplo, el núcleo de la aplicación lineal $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^2$ dada por

$$f(x, y, z, w) = (x - 2y + w, x + y - z - 2w)$$

es el espacio vectorial

$$V = \{(x, y, z, w) \in \mathbb{R}^4 \mid x - 2y + w = 0,\ x + y - z - 2w = 0\}.$$

En general, cualquier subespacio de $\mathbb{R}^n$ definido mediante un sistema de ecuaciones lineales homogéneas puede verse como el núcleo de la aplicación lineal definida por dichas ecuaciones. Paralelamente, todo espacio definido mediante un sistema generador puede interpretarse como la imagen de una aplicación lineal. Para verlo conviene conocer el teorema siguiente:

Teorema *Sean V y W dos espacios vectoriales, sea $\bar{v}_1, \bar{v}_2, \ldots, \bar{v}_n$ una base de V y sean $\bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n$ vectores cualesquiera de W. Entonces existe una única aplicación lineal $f : V \longrightarrow W$ tal que $f(\bar{v}_i) = \bar{w}_i$ para $i = 1, \ldots, n$. Además, se cumple que*

$$\operatorname{Im} f = \langle \bar{w}_1, \bar{w}_2, \ldots, \bar{w}_n \rangle .$$

Así pues, el espacio $W = \langle (1, -3, 5, 1), (1, 1, -1, 0) \rangle$ puede interpretarse como la imagen de la aplicación lineal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ dada por

$$f(1, 0) = (1, -3, 5, 1), \quad f(0, 1) = (1, 1, -1, 0).$$

Ejemplo *Halla una expresión explícita para la aplicación lineal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^4$ dada por*

$$f(1, 0) = (1, -3, 5, 1), \quad f(0, 1) = (1, 1, -1, 0).$$

Solución: Puesto que conocemos la imagen de la base canónica, la matriz de f es

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

y a partir de la matriz podemos obtener $f(x, y)$, pues

$$f(x, y)^t = A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + y \\ -3x + y \\ 5x - y \\ x \end{pmatrix},$$

es decir,

$$f(x, y) = (x + y, -3x + y, 5x - y, x).$$

■

Del teorema anterior se sigue que un generador de la imagen de una aplicación lineal $f : V \longrightarrow W$ está formado por las imágenes por f de los vectores de cualquier base de V.

Ejemplo Calcula una base de la imagen de la aplicación lineal $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por $f(x, y, z) = (x + y + z, -x + 2z)$.

SOLUCIÓN: Para obtener un generador de Imf basta calcular la imagen de los vectores de la base canónica de $\mathbb{R}^3$, formada por los vectores $(1, 0, 0), (0, 1, 0), (0, 0, 1)$.

$$f(1, 0, 0) = (1, -1), \quad f(0, 1, 0) = (1, 0), \quad f(0, 0, 1) = (1, 2).$$

Sabemos que $\mathrm{Im}f = \langle (1, -1), (1, 0), (1, 2) \rangle$. Ahora bien, estos vectores forman un sistema generador, pero no necesariamente una base. Para obtener una base hemos de eliminar los que sean combinación lineal de los demás. Observamos que

$$\mathrm{rang} \begin{pmatrix} 1 & -1 \\ 1 & 0 \\ 1 & 2 \end{pmatrix} = 2,$$

pues, por ejemplo, el determinante de las dos primeras filas es no nulo. Además esto nos dice que los vectores $(1, -1)$ y $(1, 0)$ son linealmente independientes. Por lo tanto, forman una base de Imf. ■

Citamos una última propiedad que relaciona el núcleo y la imagen de una aplicación lineal:

Teorema *Si $f : V \longrightarrow W$ es una aplicación lineal, entonces*

$$\dim V = \dim \ker f + \dim \mathrm{Im}\, f.$$

Ejemplo *Calcula las dimensiones de los subespacios* Imf *y* ker$f$ *asociados a la aplicación lineal $f(x, y, z) = (x - y + z, x + 2z)$.*

SOLUCIÓN: Se tiene $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, por tanto sabemos que la dimesión del espacio origen, $\mathbb{R}^3$, es 3. Entonces, para conocer las dimensiones de los subespacios núcleo e imagen basta conocer la dimensión de uno de ellos y aplicar el resultado anterior.

Para conocer la dimesión de Imf escribimos la matriz de f y calculamos su rango. Como

$$f(1, 0, 0) = (1, 1), \quad f(0, 1, 0) = (-1, 0), \quad f(0, 0, 1) = (1, 2),$$

la matriz de f es

$$\begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix} \Longrightarrow \mathrm{rang} \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 2.$$

Entonces, dim Im $f = 2$, y por el teorema anterior

$$\dim \ker f = \dim \mathbb{R}^3 - \dim \mathrm{Im} f = 3 - 2 = 1.$$

■

4.3 Valores propios y vectores propios

Definición Consideramos una aplicación lineal $f : V \longrightarrow V$. Diremos que un escalar $\lambda \in \mathbb{R}$ es un *valor propio* de f si existe un vector $\bar{v} \in V$, $\bar{v} \neq \bar{0}$, de manera que

$$f(\bar{v}) = \lambda \bar{v}.$$

A cualquier vector no nulo $\bar{v} \in V$ que verifique esta condición se le llama *vector propio* de f asociado al valor propio λ.

Para calcular valores y vectores propios escribimos la aplicación f en forma matricial. Supongamos que A es la matriz de f. Entonces,

$$A\bar{v}^t = \lambda \bar{v}^t \quad \Longrightarrow \quad A\bar{v}^t - \lambda \bar{v}^t = (A - \lambda I_n)\bar{v}^t = \bar{0}$$

Por tanto se trata de calcular el sistema de ecuaciones homogéneo $(A-\lambda I_n)\bar{v}^t = \bar{0}$. Este sistema tendrá solución distinta de la trivial si $|A - \lambda I_n| = 0$. Entonces, en primer lugar calcularemos los valores propios, que serán los valores de λ que verifican

$$|A - \lambda I_n| = 0.$$

Dado un valor propio λ_i, cualquier solución no nula del sistema $(A - \lambda_i I_n)\bar{v}^t = \bar{0}$ será el conjunto de vectores propios asociados.

Ejemplo *Calcula los valores y vectores propios de la aplicación* $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ *dada por* $f(x, y, z) = (x + y, x + z, y + z)$.

Solución: La imagen de la base canónica es

$$f(1,0,0) = (1,1,0), \quad f(0,1,0) = (1,0,1), \quad f(0,0,1) = (0,0,1),$$

y la matriz asociada a f en la base canónica es

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \Longrightarrow \quad A - \lambda I_3 = \begin{pmatrix} 1-\lambda & 1 & 0 \\ 1 & 0-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{pmatrix}.$$

Entonces, planteamos la ecuación

$$0 = |A - \lambda I_3| = \begin{vmatrix} 1-\lambda & 1 & 0 \\ 1 & 0-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{vmatrix} = -\lambda^3 + 2\lambda^2 + \lambda - 2$$

Sus soluciones son $\lambda_1 = 1, \quad \lambda_2 = -1, \quad \lambda_3 = 2.$

Vectores propios asociados a $\lambda_1 = 1$:

Planteamos el sistema de ecuaciones lineales asociado a $\lambda_1 = 1$:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (A - 1I_3)\bar{v}^t = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \left.\begin{array}{r} y = 0 \\ x - y + z = 0 \\ y = 0 \end{array}\right\}$$

La solución del sistema es $x = \alpha, \quad y = 0, \quad z = -\alpha, \quad \alpha \in \mathbb{R}$. Entonces los vectores propios asociados a $\lambda_1 = 1$ son

$$\{(\alpha, 0, -\alpha), \quad \alpha \in \mathbb{R}, \quad \alpha \neq 0\}.$$

Vectores propios asociados a $\lambda_2 = -1$:

Planteamos el sistema de ecuaciones lineales asociado a $\lambda_1 = -1$:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (A + 1I_3)\bar{v}^t = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \left.\begin{array}{r} 2x + y = 0 \\ x + y + z = 0 \\ y + 2z = 0 \end{array}\right\}$$

La solución del sistema es $x = \alpha, \quad y = -2\alpha, \quad z = \alpha, \quad \alpha \in \mathbb{R}$. Entonces los vectores propios asociados a $\lambda_2 = -1$ son

$$\{(\alpha, -2\alpha, \alpha), \quad \alpha \in \mathbb{R}, \quad \alpha \neq 0\}.$$

Vectores propios asociados a $\lambda_3 = 2$:

Planteamos el sistema de ecuaciones lineales asociado a $\lambda_1 = 2$:

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = (A - 2I_3)\bar{v}^t = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \left.\begin{array}{r} -x + y = 0 \\ x - 2y + z = 0 \\ y - z = 0 \end{array}\right\}$$

La solución del sistema es $x = \alpha, \quad y = \alpha, \quad z = \alpha, \quad \alpha \in \mathbb{R}$. Entonces los vectores propios asociados a $\lambda_3 = 2$ son

$$\{(\alpha, \alpha, \alpha), \quad \alpha \in \mathbb{R} \quad \alpha \neq 0\}.$$

■

4.4 Ejercicios

1. Indica si las aplicaciones siguientes son lineales o no:

   (a) $f(x, y, z) = (x, 2x + y + z, z)$,

   (b) $g(x, y) = (3x, x - 2y, x - 2)$,

   (c) $h(x, y, z) = x - y + z$,

   (d) $p(u, v) = (u - v, v, 2)$,

   (e) $q(s, t) = (s, t, 2t - 3s)$.

2. Escribe la matriz (para las bases canónicas) de las aplicaciones lineales siguientes:

   (a) $f(x, y, z) = (x, 2x + y + z, z)$,

   (b) $g(x, y, z, t) = (3x, x - 2y + z, x - y - 2t)$,

(c) $h(x, y, z) = (x - y + z, x - y + z)$,

(d) $p(u, v) = (u - v, v, u + v)$,

(e) $q(r, s, t) = 2r + 3s - 5t$.

3. Dadas las aplicaciones lineales del ejercicio anterior, escribe la matriz de las aplicaciones siguientes:

   (a) $3g$, $2p$,

   (b) $h \circ f$,

   (c) $q \circ g$,

   (d) $p \circ h \circ f$.

4. Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ la aplicación lineal que sobre la base B de $\mathbb{R}^2$ formada por los vectores $(1, 2)$, $(-2, 1)$ viene dada por

$$f(1, 2) = (1, 1, 1), \quad f(-2, 1) = (0, 1, 0)$$

   (a) Escribe la matriz de f en las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^3$,

   (b) Escribe la matriz de f en la base B de $\mathbb{R}^2$ y la base canónica de $\mathbb{R}^3$.

5. Calcula la aplicación lineal $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por

$$f(1, 0, 1) = (1, 1), \quad f(2, 1, 1) = (1, 0), \quad f(0, 1, 1) = (0, 1).$$

6. Calcula bases del núcleo y la imagen de las aplicaciones lineales siguientes:

   (a) $f(x, y) = (x, 2x + y)$,

   (b) $g(x, y, z, t) = (3x, x - 2y + z, x - y - 2t)$,

   (c) $h(x, y, z) = (x - 2y + z, x - 2y + z)$.

7. Una empresa produce cuatro tipos de artículos, p_1, p_2, p_3 y p_4, que requieren el uso de tres materias primas diferentes m_1, m_2 y m_3. Las unidades de materia prima que se emplean en elaborar cada artículo vienen expresadas en la tabla siguiente:

Materias primas	Artículos			
	p_1	p_2	p_3	p_4
m_1	1	0.2	0.1	1
m_2	0.2	0.7	0.8	0.9
m_3	0.2	2	0.9	1

   (a) Determina la relación que asocia cada producción con el uso de las materias primas necesarias.

   (b) ¿La relación anterior describe una aplicación lineal?

8. Una industria produce tres artículos A, B y C. Sabemos que la variación de los precios de estos estos productos, que representaremos como $\Delta\bar{p} = (\Delta p_A, \Delta p_B, \Delta p_C)$ depende de la variación de los salarios s y de la tasa de impuestos t, es decir que depende del vector (s,t).

   Sabemos que un aumento del 2% en los salarios provoca un aumento de los precios de A, B y C en un 3%, un 1% y un 2% respectivamente. Por otro lado se conoce que un aumento simultáneo de salarios y tasas en un 1% cada uno origina un aumento en los precios de A, B y C en un 4%, un 2% y un 3% respectivamente.

   Suponiendo que la relación entre las variaciones de precios es lineal, calcula la expresión general que relaciona estas variaciones. Si el próximo año se prevé una subida de salarios del 4% y una subida de las tasas del 2%, ¿cómo afectará a los precios de la producción? (Ayuda: conocemos la imagen de una aplicación lineal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ sobre la base $(1,0)$ y $(1,1)$).

9. Obtén los valores y vectores propios de los endomorfismos siguientes:

   (a) $f(x,y,z) = (5x + 3z, -y, 2y + 8z)$,

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

5. Límites y continuidad de funciones

El concepto de función es fundamental a la hora de estudiar matemáticamente la relación entre distintas magnitudes. Se dice que una magnitud M es función de otras magnitudes $x_1, \ldots, x_n$ si el valor que toman éstas determina completamente el valor de M. En tal caso escribimos que $M = M(x_1, \ldots, x_n)$. Por ejemplo, si depositamos un capital en un banco, el capital C disponible en un tiempo dado t es función de t, del capital inicial depositado C_0 y del tipo de interés i. Concretamente

$$C = C(C_0, i, t) = C_0(1 + i)^t.$$

Esta fórmula nos dice cómo calcular C si conocemos las variables C_0, i, t. En general, la relación entre una magnitud y otras no tiene por qué ser expresable mediante una única fórmula matemática. Por ejemplo, supongamos que un trabajador es contratado por tres años con un sueldo de 200.000 € anuales con una revisión anual del 4%. Entonces su salario S como función del tiempo $S = S(t)$ viene dado por

$$S(t) = \begin{cases} 200.000 & \text{si } 0 \leq t < 1, \\ 208.000 & \text{si } 1 \leq t < 2, \\ 216.320 & \text{si } 2 \leq t \leq 3. \end{cases}$$

Más en general, es frecuente que una magnitud pueda considerarse función de otras sin que conozcamos explícitamente la relación entre ellas. En tales casos entra en juego la modelización matemática, que nos lleva a proponer fórmulas aproximadas o razonables. Por ejemplo, se puede suponer que la demanda de un artículo es función de su precio, lo cual no es exacto, pues además del precio pueden influir otros muchos factores, si bien éstos pueden ser despreciados en la medida en la que puedan considerarse constantes. Aun así, una relación explícita entre demanda y precios siempre deberá considerarse como una aproximación teórica más o menos simplificada a una función que en realidad es muy compleja.

5.1 Funciones de varias variables

Definición Consideramos D un subconjunto de $\mathbb{R}^n$. Una *función escalar* de n variables reales es cualquier criterio que a cada punto $\bar{x} \in D$ le asigna un único número real $f(\bar{x})$. La función f se representa $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$.

El conjunto D sobre el que está definida la función f se llama *dominio* de f y, dado $\bar{x} \in D$, el número $f(\bar{x})$ se llama *imagen* de $\bar{x}$ por f.

Cuando $n = 1$, es decir, cuando $f : D \subset \mathbb{R} \longrightarrow \mathbb{R}$, se dice que f es una *función real de variable real.*

En general, para determinar el dominio de una función habrá que tener en cuenta que:

1. Los polinomios tienen por dominio $\mathbb{R}^n$.
2. El denominador de una fracción no puede anularse.
3. El argumento de un logaritmo ha de ser mayor que 0.
4. El dominio de la exponencial, el seno y el coseno es el mismo que el dominio del argumento.
5. El radicando de una raíz de índice par ha de ser mayor o igual que 0.
6. El dominio de una raíz de índice impar es el mismo que el del radicando.
7. La base de una potencia de exponente variable ha de ser mayor que 0.

Ejemplo *Calcula el dominio de la función* $f(x, y, z) = \ln(x + y + z)$.

SOLUCIÓN: Se trata de una función $f : D \subset \mathbb{R}^3 \longrightarrow \mathbb{R}$, definida sobre el conjunto

$$D = \{(x, y, z) \in \mathbb{R}^3 \mid x + y + z > 0\}.$$

■

No obstante, a menudo sucede que si una función tiene una interpretación económica, conviene considerarla definida en un conjunto menor que su dominio matemático. Por ejemplo, la función $C(C_0, i, t) = C_0(1+i)^t$ es una función escalar cuyo dominio matemático es

$$D = \{(C_0, i, t) \in \mathbb{R}^3 \mid i > -1\},$$

pues la única restricción (propiedad 7) es que la base $1+i$ de la potencia ha de ser mayor que 0. No obstante, cuando consideramos a C como la ley de capitalización correspondiente al interés compuesto hemos de considerar que su dominio es

$$D' = \{(C_0, i, t) \in \mathbb{R}^3 \mid C_0 \geq 0,\ i \geq 0\},$$

pues no tiene sentido considerar capitales iniciales negativos C_0 ni tampoco tipos de interés negativos. Incluso podemos desechar por trivial el caso de un capital o interés nulo. En determinados contextos convendrá considerar únicamente tiempos no negativos ($t \geq 0$), aunque en otros puede tener sentido que t sea arbitrario.

Definición Consideremos un subconjunto D de $\mathbb{R}^n$. Una *función vectorial* de n variables y m coordenadas es cualquier criterio que a cada n números reales del dominio D les asigna un único vector de m números reales. La función f se representa por

$$f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m.$$

Las funciones escalares son el caso particular de las funciones vectoriales que se da cuando $m = 1$. Toda función vectorial se expresa en la forma

$$f(x_1, \ldots, x_n) = (f_1(x_1, \ldots, x_n), \ldots, f_m(x_1, \ldots, x_n)),$$

donde las funciones $f_i : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ son funciones escalares con mismo dominio que f y se llaman *funciones coordenadas* de f.

Por ejemplo, la función $f(x, y, z) = (x^2/y,\ x + y + z)$ es una función vectorial de tres variables y dos funciones coordenadas, es decir, $f : D \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, donde

$$D = \{(x, y, z) \in \mathbb{R}^3 \mid y \neq 0\}$$

y sus funciones coordenadas son $f_1(x, y, z) = x^2/y$, $f_2(x, y, z) = x + y + z$.

5.2 Nociones de topología en $\mathbb{R}^n$

Definición La *norma* de un vector $\bar{x} \in \mathbb{R}^n$ se define como el escalar

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

Geométricamente se trata de la longitud del vector $\bar{x}$.

Teorema *Si $\bar{x}$ e $\bar{y}$ son vectores no nulos en $\mathbb{R}^n$, entonces*

$$\bar{x} \cdot \bar{y} = \|\bar{x}\|\|\bar{y}\| \cos\phi,$$

donde ϕ es el ángulo que forman.

Como el coseno toma valores entre -1 y 1, tenemos la *desigualdad de Schwarz* (válida incluso si algún vector es nulo):

$$|\bar{x} \cdot \bar{y}| \leq \|\bar{x}\|\|\bar{y}\|.$$

Las propiedades básicas de la norma vienen dadas por el teorema siguiente:

Teorema *Si $\bar{x}$, $\bar{y}$ son vectores de $\mathbb{R}^n$ y α es un número real, se cumple:*

1. $\|\bar{x}\| \geq 0$,
2. $\|\bar{x}\| = 0$ *si y sólo si* $\bar{x} = \bar{0}$,
3. $\|\alpha\bar{x}\| = |\alpha|\,\|\bar{x}\|$,
4. $\|\bar{x} + \bar{y}\| \leq \|\bar{x}\| + \|\bar{y}\|$.

Demostración La propiedad 1) es obvia, pues $\|\bar{x}\| = \sqrt{\bar{x}\cdot\bar{x}} \geq 0$ (las raíces cuadradas siempre son mayores o iguales que 0).

2) Se cumple $\|\bar{x}\| = 0$ si y sólo si $\sqrt{\bar{x}\cdot\bar{x}} = 0$, si y sólo si $\bar{x}\cdot\bar{x} = 0$ (pues el único número real con raíz cuadrada nula es 0), si y sólo si $x_1^2+\cdots+x_n^2 = 0$. Como todos los sumandos son mayores o iguales que 0 (porque son cuadrados) esto sucede si y sólo si cada $x_i^2 = 0$, es decir, si y sólo si cada $x_i = 0$, lo que a su vez equivale a que $\bar{x} = \bar{0}$.

3) $\|\alpha\bar{x}\| = \sqrt{(\alpha x_1)^2 + \cdots + (\alpha x_n)^2} = \sqrt{\alpha^2 x_1^2 + \cdots + \alpha^2 x_n^2}$

$$= \sqrt{\alpha^2(x_1^2 + \cdots + x_n^2)} = |\alpha|\sqrt{x_1^2 + \cdots + x_n^2} = |\alpha|\|\bar{x}\|.$$

4) Para demostrar que $\|\bar{x}+\bar{y}\| \leq \|\bar{x}\|+\|\bar{y}\|$ basta ver —elevando al cuadrado— que

$$\|\bar{x}+\bar{y}\|^2 \leq (\|\bar{x}\| + \|\bar{y}\|)^2 = \|\bar{x}\|^2 + \|\bar{y}\|^2 + 2\|\bar{x}\|\|\bar{y}\|.$$

Por la definición de norma, esto equivale a

$$(\bar{x}+\bar{y})\cdot(\bar{x}+\bar{y}) \leq \bar{x}\cdot\bar{x} + \bar{y}\cdot\bar{y} + 2\|\bar{x}\|\|\bar{y}\|.$$

Por las propiedades del producto escalar, esto equivale a su vez a que

$$\bar{x}\cdot\bar{x} + \bar{y}\cdot\bar{y} + 2\bar{x}\cdot\bar{y} \leq \bar{x}\cdot\bar{x} + \bar{y}\cdot\bar{y} + 2\|\bar{x}\|\|\bar{y}\|.$$

Simplificando los términos iguales, lo que hemos de probar es

$$2\bar{x}\cdot\bar{y} \leq 2\|\bar{x}\|\|\bar{y}\|.$$

Eliminamos los doses y queda $\bar{x}\cdot\bar{y} \leq \|\bar{x}\|\|\bar{y}\|$. Ahora bien, esto se cumple por la desigualdad de Schwarz:

$$\bar{x}\cdot\bar{y} \leq |\bar{x}\cdot\bar{y}| \leq \|\bar{x}\|\cdot\|\bar{y}\|.$$

■

Definición La *distancia* entre dos puntos de $\mathbb{R}^n$ se define como

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

Definición Dado $\epsilon > 0$ y $\bar{p} \in \mathbb{R}^n$, se define la *bola abierta* de centro $\bar{p}$ y radio ϵ como el conjunto

$$B_\epsilon(\bar{p}) = \{\bar{x} \in \mathbb{R}^n \mid \|\bar{x}-\bar{p}\| < \epsilon\}.$$

Así, para ϵ suficientemente pequeño, la bola $B_\epsilon(\bar{p})$ contiene todos los puntos de alrededor de $\bar{p}$, es decir, todos los puntos a los que podemos llegar mediante modificaciones pequeñas de $\bar{p}$.

Un subconjunto A de $\mathbb{R}^n$ es *abierto* si para todo punto $\bar{p} \in A$ existe un $\epsilon > 0$ tal que $B_\epsilon(\bar{p}) \subset A$.

En otras palabras, A es abierto si cuando estamos en un punto $\bar{p}$ de A podemos movernos en cualquier dirección sin salirnos de A.

Un subconjunto C de $\mathbb{R}^n$ es *cerrado* si su complementario $\mathbb{R}^n \sim C$ es abierto.

Un punto $\bar{p} \in \mathbb{R}^n$ es un *punto de acumulación* de un subconjunto A de $\mathbb{R}^n$ si para todo $\epsilon > 0$ existe un punto $\bar{x} \in A$, $\bar{x} \neq \bar{p}$, tal que $\|\bar{x} - \bar{p}\| < \epsilon$.

En otros términos, $\bar{p}$ es un punto de acumulación de A si hay puntos de A distintos del propio $\bar{p}$ tan cercanos a $\bar{p}$ como queramos o, también, si podemos acercarnos arbitrariamente a $\bar{p}$ desde puntos de A sin pasar por $\bar{p}$.

Si $\bar{p} \in A$ no es un punto de acumulación de A, se dice que $\bar{p}$ es un *punto aislado* de A.

Según esto, un punto de A es aislado si no tiene a su alrededor ningún otro punto de A.

5.3 Límites

Definición Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar y sea $\bar{p} \in \mathbb{R}^n$ un punto de acumulación de D. Se dice que un punto $l \in \mathbb{R}$ es el *límite* de $f(\bar{x})$ cuando $\bar{x}$ tiende a $\bar{p}$ (y se representa por $\lim\limits_{\bar{x} \to \bar{p}} f(\bar{x}) = l$) si para todo $\epsilon > 0$ existe $\delta > 0$ tal que si $\bar{x} \in D$, $\bar{x} \neq \bar{p}$ y $\|\bar{x} - \bar{p}\| < \delta$, entonces $\|f(\bar{x}) - l\| < \epsilon$.

Si ϵ es muy pequeño, la condición $\|f(\bar{x}) - l\| < \epsilon$ significa que $f(\bar{x})$ es casi igual a $\bar{l}$, luego la definición de límite puede parafrasearse así: Si tomamos puntos $\bar{x}$ suficientemente cercanos a $\bar{p}$ (concretamente, tan cercanos como para que $\|\bar{x} - \bar{p}\| < \delta$), que estén en D (es decir, tales que exista $f(\bar{x})$) y que no sean el propio $\bar{p}$, entonces $f(\bar{x})$ es casi igual a l.

En la práctica conviene recordar lo siguiente:

> La igualdad $\lim\limits_{\bar{x} \to \bar{p}} f(\bar{x}) = l$ significa que si tomamos puntos $\bar{x} \approx \bar{p}$ sobre los que esté definida la función f, entonces $f(\bar{x}) \approx l$, entendiendo que la última aproximación será mejor cuanto mejor sea la primera.

El ejemplo siguiente muestra un caso práctico de "paso al límite":

Aplicación: el interés continuo Supongamos que un banco nos ofrece una cuenta corriente con un interés simple de un 3% anual. Esto significa que los intereses generados en un tiempo t (expresado en años) serán $I = C_0 it$, donde C_0 es el capital depositado e i el tanto por uno anual. El capital total será entonces $C = C_0 + I = C_0(1 + it)$. Supongamos que depositamos 10000 € durante 3 años.

Según esto, al cabo de tres años dispondremos de un capital

$$C(3) = 10000(1 + 0.03 \cdot 3) = 10900 \text{ €}.$$

Ahora bien, supongamos que al cabo de un año ($t = 1$) vamos al banco y cancelamos la cuenta, con lo que el banco nos da

$$C(1) = 10000(1 + 0.03 \cdot 1) = 10300 \text{ €},$$

con los cuales abrimos una nueva cuenta al mismo tipo de interés (simple). Al cabo de un año tendremos

$$10300(1 + 0.03 \cdot 1) = 10609\,€,$$

y si los volvemos a ingresar, al tercer año acabaremos con un capital de

$$10609(1 + 0.03 \cdot 1) = 10927.27\,€.$$

Así pues, hemos ganado 27.27 € sin más que meter y sacar nuestro dinero cada año. La razón es clara: al forzar al banco a que nos pague los intereses del primer año e ingresarlos en la cuenta, estos intereses generan a su vez nuevos intereses, los cuales, a partir del segundo año generan nuevos intereses. Ahora bien, todavía podríamos obtener más dinero si cancelamos y reabrimos nuestra cuenta cada mes. En general, si forzamos al banco a capitalizar los intereses cada Δt años (por ejemplo, $\Delta t = 1/12$ sería cada mes), tenemos que al final de cada periodo nuestro capital se multiplica por $1 + i\Delta t$, y como en t años hay $t/\Delta t$ periodos, el capital final sería

$$C(\Delta t) = C_0(1 + i\Delta t)^{t/\Delta t}.$$

Con los datos anteriores, si capitalizamos cada mes obtendríamos

$$C = 10000\Big(1 + 0.03 \cdot \frac{1}{12}\Big)^{12\cdot 3} = 10940,51\,€,$$

con lo que hemos "arañado" 13 € más (40.27 más que al principio). La tabla siguiente muestra más casos:

Δt	$C(\Delta t)$
3 años	10900
1 año	10927.27
1 mes	10940.51
1 día	10941.70
1 hora	10941.74

En la práctica no tiene sentido ir más allá, pues las diferencias para valores más pequeños para Δt no llegan a un céntimo. De todos modos, desde un punto de vista puramente matemático, el capital obtenido con capitalizaciones de 1 hora es $C = 10941.74115\ldots$, mientras que si capitalizamos cada segundo resulta $C = 10941.74280\ldots$, con lo que ganamos más de una milésima de euro. Puede probarse que

$$\lim_{\Delta t \to 0} C_0(1 + i\Delta t)^{t/\Delta t} = C_0 e^{it}.$$

La fórmula

$$\boxed{C = C_0 e^{it}}$$

nos da el capital generado por un capital inicial C_0 en un tiempo t a un tanto por uno de *interés continuo* i.

En nuestro ejemplo $C = 10000e^{0.03\cdot 3} = 10941.74283\ldots$. El hecho de que

$$C = \lim_{\Delta t \to 0} C(\Delta t)$$

se interpreta (en el caso de nuestro ejemplo), como que si $\Delta t \approx 0$, entonces el capital generado por un interés simple anual del 3% durante 3 años capitalizando intereses cada Δt años es $C(\Delta t) \approx 10941.74283\ldots$ Puesto que podemos descartar las milésimas de euro, podemos decir que si Δt es a lo sumo de una hora, entonces el capital generado es prácticamente de 10941.74 €.

Esto explica —entre otras razones— por qué los bancos no ofrecen interés simple, pues ello sería invitar a sus clientes a meter y sacar su dinero cuantas más veces mejor. En la práctica se usa el interés compuesto (discreto), que consiste en pactar no sólo el tipo de interés, sino también los intervalos en que capitalizan los intereses. Como es sabido, un depósito de un capital C_0 a un tanto por uno efectivo anual i genera en un tiempo t un capital

$$C = C_0(1+i)^t,$$

lo cual es equivalente a un interés (simple) i en el que los intereses capitalizan cada año (sin necesidad de que vayamos al banco a cancelar y reabrir la cuenta). Podría pensarse que el interés continuo es una entelequia, en el sentido de que ningún banco nos va a capitalizar intereses cada segundo o cada hora, pero cualquiera familiarizado con el interés compuesto sabe que, en realidad, el "periodo de capitalización de intereses" carece de sentido con el interés continuo, pues un interés efectivo anual es equivalente a otro tipo de interés efectivo trimestral, o mensual, o diario, etc.; y del mismo modo es equivalente a un interés continuo. Para comprobarlo basta igualar las leyes de capitalización compuesta (discreta) y continua con dos tipos de interés distintos i e i_∞:

$$C_0(1+i)^t = C_0 e^{i_\infty t},$$

de donde, tomando logaritmos, obtenemos la relación

$$t\ln(1+i) = t i_\infty \ln e = t i_\infty \Rightarrow i_\infty = \ln(1+i).$$

En resumen:

> Un interés efectivo anual (discreto) i es equivalente a un interés continuo $i_\infty = \ln(1+i)$, en el sentido de que la ley de capitalización correspondiente al interés discreto i proporciona el mismo capital final que la ley de capitalización correspondiente al interés continuo[1] i_∞.

Así pues, hablar de una cuenta corriente con un tipo de interés continuo del 3% no es una abstracción matemática, sino únicamente una forma matemáticamente más cómoda y simple de expresar la ley de capitalización que emplean habitualmente los bancos. ■

Propiedades de los límites

1. Para que tenga sentido calcular $\lim\limits_{\bar{x}\to\bar{p}} f(\bar{x})$ es necesario que $\bar{p}$ sea un punto de acumulación del dominio de f, es decir, que nos podamos acercar a $\bar{p}$ por puntos donde esté definida f.

[1] Usaremos i_∞ para denotar el interés continuo en analogía con la notación usual i_{12} para el interés mensual, i_{365} para el interés diario, etc.

2. En la definición sólo se considera el valor de $f(\bar{x})$ los puntos $\bar{x}$ cercanos a $\bar{p}$ distintos del propio $\bar{p}$, por lo que el límite de una función en un punto $\bar{p}$ sólo depende de la definición de f en los puntos de **alrededor** de $\bar{p}$ distintos del propio $\bar{p}$.

3. Si una función tiene límite en un punto, el límite es único, es decir, una función no puede tener dos límites distintos en el mismo punto.

4. Una función vectorial $f = (f_1, \ldots, f_m)$ tiene límite en un punto $\bar{p}$ si y sólo si lo tienen todas las funciones coordenadas f_i, y entonces

$$\lim_{\bar{x}\to\bar{p}} f(\bar{x}) = \left(\lim_{\bar{x}\to\bar{p}} f_1(\bar{x}), \ldots, \lim_{\bar{x}\to\bar{p}} f_m(\bar{x})\right).$$

Teorema *Sean $f, g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ dos funciones escalares, $\alpha \in \mathbb{R}$ y $\bar{p}$ un punto de acumulación del dominio D. Entonces*

1. *Si f y g tienen límite en $\bar{p}$, también lo tienen $f \pm g$ y fg y se cumple que*

$$\lim_{\bar{x}\to\bar{p}}(f(\bar{x}) + g(\bar{x})) = \lim_{\bar{x}\to\bar{p}} f(\bar{x}) + \lim_{\bar{x}\to\bar{p}} g(\bar{x}),$$

$$\lim_{\bar{x}\to\bar{p}} f(\bar{x})g(\bar{x}) = \lim_{\bar{x}\to\bar{p}} f(\bar{x}) \lim_{\bar{x}\to\bar{p}} g(\bar{x}).$$

2. *Si f tiene límite en $\bar{p}$, también lo tiene αf y se cumple que*

$$\lim_{\bar{x}\to\bar{p}}(\alpha f(\bar{x})) = \alpha \lim_{\bar{x}\to\bar{p}} f(\bar{x}).$$

3. *Si f y g tienen límite en $\bar{p}$ y $\lim_{\bar{x}\to\bar{p}} g(\bar{x}) \neq 0$, entonces f/g tiene límite en $\bar{p}$ y*

$$\lim_{\bar{x}\to\bar{p}} \frac{f(\bar{x})}{g(\bar{x})} = \frac{\lim_{\bar{x}\to\bar{p}} f(\bar{x})}{\lim_{\bar{x}\to\bar{p}} g(\bar{x})}.$$

5.4 Continuidad

Definición Diremos que una función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es *continua* en un punto de acumulación $\bar{p} \in D$ si existe $\lim_{\bar{x}\to\bar{p}} f(\bar{x}) = f(\bar{p})$.

En la práctica, para comprobar que f es continua en $\bar{p}$ hemos de verificar las propiedades siguientes:

1. f está definida en $\bar{p}$ (es decir, existe $f(\bar{p})$).

2. Existe $\lim_{\bar{x}\to\bar{p}} f(\bar{x})$.

3. Se da la igualdad $f(\bar{p}) = \lim_{\bar{x}\to\bar{p}} f(\bar{x})$.

De las propiedades de los límites se deducen las siguientes propiedades de la continuidad:

Propiedades de las funciones continuas

1. La continuidad de una función f en un punto $\bar{p}$ depende únicamente de los valores que toma f **alrededor** de $\bar{p}$ (pero, al contrario de lo que sucede con los límites, teniendo en cuenta el valor que toma en $\bar{p}$).

2. Una función vectorial $f = (f_1, \dots, f_m)$ es continua en un punto $\bar{p}$ si y sólo si lo son todas sus funciones coordenadas f_i.

Hemos definido únicamente la continuidad de una función en un punto de acumulación de su dominio, lo cual deja sin considerar el caso de los puntos aislados. Se considera que una función es siempre continua en los puntos aislados de su dominio, pero este caso no nos va aparecer en ningún momento.

Teorema *Se cumple:*
1) Si f, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ son funciones continuas en un punto $\bar{p} \in D$ y $\alpha \in \mathbb{R}$, entonces también son continuas en $\bar{p}$ las funciones $f \pm g$, αf, fg y, si además $g(\bar{p}) \neq 0$, también es continua la función f/g.
2) Todo polinomio es una función continua en todos los puntos de su dominio.
3) La composición de funciones continuas es una función continua.
4) Las funciones e^x, $\ln x$, $\operatorname{sen} x$, $\cos x$, $\sqrt[n]{x}$ son continuas en todos los puntos de su dominio.

En definitiva, hemos de recordar:

Toda función construida a partir de polinomios mediante sumas, productos, cocientes y composición con las funciones usuales (exponencial, logaritmo, seno, coseno, etc.) es continua en todos los puntos de su dominio.

Por ejemplo, la función

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

es continua en todos los puntos de $\mathbb{R}^2$, pues es un cociente de polinomios (que son funciones continuas) y el denominador no se anula en ningún punto.

En particular

$$\lim_{(x,y)\to(1,0)} \frac{x + y}{x^2 + y^2 + 1} = f(1, 0) = \frac{1}{2}.$$

En general:

Siempre que tengamos que calcular el límite de una función en un punto donde sepamos que ésta es continua, el límite se obtendrá sin más que evaluar la función en el punto, por la propia definición de continuidad.

Ejemplo *Estudia la continuidad de la función*

$$f(x,y) = \begin{cases} x & si\ y = x^2 \\ y & si\ y \neq x^2 \end{cases}$$

en todo $\mathbb{R}^2$.

SOLUCIÓN: Hemos de tomar un punto arbitrario $(x_0, y_0) \in \mathbb{R}^2$ y estudiar si f es continua en (x_0, y_0). Distinguimos dos casos:
Si $y_0 \neq x_0^2$, entonces ALREDEDOR del punto (x_0, y_0) la función f está definida por y, que es un polinomio, luego una función continua.
Si $y_0 = x_0^2$ entonces alrededor del punto (x_0, y_0) hay puntos con definiciones diferentes, por lo que hemos de estudiar la continuidad mediante la definición.

1. En primer lugar observamos que existe $f(x_0, y_0) = x_0$.
2. En segundo lugar estudiamos el límite $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y)$. Como podemos acercarnos a (x_0, y_0) por puntos de los dos subdominios donde f tiene definiciones distintas, hemos de calcular por separado

$$\lim_{\substack{(x,y)\to(x_0,y_0) \\ y=x^2}} f(x,y) = \lim_{\substack{(x,y)\to(x_0,y_0) \\ y=x^2}} x = x_0,$$

$$\lim_{\substack{(x,y)\to(x_0,y_0) \\ y\neq x^2}} f(x,y) = \lim_{\substack{(x,y)\to(x_0,y_0) \\ y\neq x^2}} y = y_0 = x_0^2.$$

   El límite existirá solamente si $x_0 = x_0^2$, es decir, $x_0^2 - x_0 = 0 \Rightarrow x_0(x_0 - 1) = 0$, lo que sucede si y sólo si $x_0 = 0$ o $x_0 = 1$. En definitiva, el límite existe en los puntos $(0,0)$ y $(1,1)$. Concretamente,

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0, \qquad \lim_{(x,y)\to(1,1)} f(x,y) = 1.$$

3. Como $f(0,0) = 0$ y $f(1,1) = 1$, en estos puntos el valor del límite coincide con el de la función, la cual es, por tanto, continua.

En definitiva, f es continua en el conjunto $\{(x,y) \in \mathbb{R}^2 \mid y \neq x^2\} \cup \{(0,0),(1,1)\}$. ∎

Ejemplo *Estudia la continuidad de la función*

$$f(x,y,z) = (x^2 y, z + y + z, y \operatorname{sen} z).$$

SOLUCIÓN: La función f es vectorial, luego será continua en los puntos donde los sean su tres funciones coordenadas. Las dos primeras son polinomios, luego son continuas en $\mathbb{R}^3$. La tercera es continua en $\mathbb{R}^3$ porque es producto del polinomio y por la función continua $\operatorname{sen} z$. Por consiguiente, f es continua en $\mathbb{R}^3$. ∎

5.5 Ejercicios

1. Calcula el dominio de las funciones siguientes:

   (a) $\dfrac{x+y}{x-y}$,

   (b) $\sqrt{x^2+y^2}$,

   (c) $e^{y/\sqrt{x}}$,

   (d) x^y,

   (e) $\dfrac{x\ln(x+y+1)}{x^2+y^2}$,

   (f) $\dfrac{\sqrt[3]{x^2-2y}}{\sqrt[4]{x-y^2}}$,

   (g) $\operatorname{sen}(x^2+e^y)$,

   (h) $f(x,y)=\begin{cases}\dfrac{1}{x+y} & \text{si } x<0\\ y\ln x & \text{si } x\geq 0.\end{cases}$

2. Calcula el dominio de las funciones siguientes:

   (a) $g(x,y)=\begin{cases}\ln(x+y) & \text{si } x+y>0\\ \frac{1}{\sqrt{-x-y}} & \text{si } x+y\leq 0,\end{cases}$

   (b) $h(x,y,z)=(x\ln(y+z), e^{1/y}, x+y^2-3z)$,

   (c) $p(u,v)=\Big(\dfrac{3}{u^2+v^2}, \dfrac{e^u}{u^3+v^3}\Big)$.

3. Calcula:

$$\lim_{(x,y)\to(1,2)} 3e^{(x-1)^2+(2-y)}, \qquad \lim_{(x,y)\to(1,0)} (x^2+3xy, x-y, x+y^4-3).$$

4. Pon un ejemplo de función $f:\mathbb{R}^3\longrightarrow\mathbb{R}^2$ cuyas funciones coordenadas sean polinomios.

5. Calcula el dominio de las funciones siguientes y estudia su continuidad en $(3,3)$ y en $(5,5)$:

$$f(x,y)=\begin{cases}\sqrt{xy} & \text{si } x=3,\\ \sqrt{3}\sqrt{y} & \text{si } x\neq 3.\end{cases} \qquad g(x,y)=\frac{x^2+1}{x^2+3y^2}.$$

6. Utiliza las funciones

$$f(x,y)=\begin{cases}x+y+4 & \text{si } (x,y)\neq(1,2)\\ 7 & \text{si } (x,y)=(1,2)\end{cases}$$

   y

$$g(x,y)=\begin{cases}x+y+4 & \text{si } (x,y)\neq(1,2)\\ 3 & \text{si } (x,y)=(1,2)\end{cases}$$

   como ejemplo para explicar esta afirmación:

La existencia de límite de una función en un punto y —en su caso— el valor de éste no dependen del valor de la función en el punto, mientras que la continuidad sí que depende de dicho valor.

7. Estudia la continuidad de la función

$$f(x,y) = \begin{cases} 3x^2 + y & \text{si } x \neq 1 \\ -2y^2 & \text{si } x = 1 \end{cases}$$

en los puntos $(1,1)$ y $(3,3)$.

8. Estudia la continuidad de la función

$$f(x,y) = \begin{cases} x^2 + y & \text{si } x \geq 0 \\ 3 - x & \text{si } x < 0 \end{cases}$$

en los puntos $(-2,3)$, $(0,3)$ y en un punto (x_0, y_0) tal que $x_0 > 0$.

9. Estudia la continuidad de la función

$$f(x,y) = \begin{cases} x^3 & \text{si } x \geq y \\ 2y - 1 & \text{si } x < y. \end{cases}$$

10. Estudia la continuidad de las funciones

$$f(x,y) = \begin{cases} 1 & \text{si } xy \neq 0 \\ 0 & \text{si } xy = 0, \end{cases} \qquad g(x,y) = \begin{cases} xy & \text{si } xy \neq 0 \\ 0 & \text{si } xy = 0. \end{cases}$$

11. Se estima que la demanda de un producto en un mercado está en función del precio p y de la renta media r de los consumidores:

$$D(r,p) = \frac{20rp^2}{r^3 + p^3}\, e^{r/p}.$$

Actualmente la renta es $r = 5$ u.m. y el precio es $p = 2$ u.m. Determina el dominio matemático de la función D y el subdominio con sentido económico. Estudia la continuidad en uno y otro. Calcula la demanda actual del producto, así como el incremento esperado si $\Delta p = 1$. Ídem si $\Delta r = 1$. Interpreta el resultado. ¿Cuál sería el incremento si tanto la renta como el precio se duplicaran, es decir, si se pasara a $(r,p) = (10,4)$? Interprétalo. Calcula $D(5, 0.05)$ e interpreta el resultado.

12. Se estima que la función de costes de una empresa es

$$C(x,y) = 150 \ln(3 + x + 2y),$$

donde x e y son las cantidades producidas de los dos artículos que fabrica la empresa. La producción actual es $(x,y) = (100,100)$. Determina el dominio matemático de la función C así como el subdominio con sentido económico. Estudia la continuidad de C en uno y otro. Determina los costes fijos de la empresa y la función de costes variables. Calcula el incremento ΔC que experimentan los costes si la producción se incrementa en $\Delta x = 4$. Ídem si $\Delta y = 4$. Ídem si, simultáneamente, $\Delta x = 3$ y $\Delta y = 2$. Escribe el vector de incremento de la producción en cada uno de los tres casos. Repite el problema suponiendo que el vector de producción actual es $(x,y) = (50,30)$.

13. Una editorial distribuye una enciclopedia y una historia del arte a través de un equipo de 10 vendedores. Cada uno de ellos tiene un salario anual de 7000 € más una comisión de 30 € por cada enciclopedia vendida y 35 € por cada historia del arte. Además, si las ventas anuales conjuntas superan las 2000 unidades el equipo recibe un suplemento de 5000 € repartido en proporción a las ventas de cada miembro.

    Calcula la función de costes anuales $C(x, y)$ que le supone a la editorial su equipo de ventas, donde x e y son las unidades vendidas de la enciclopedia y la historia, respectivamente. Indica su dominio económico y estudia en él su continuidad. El beneficio de la editorial (sin contar el coste del equipo) es de 130 € por cada enciclopedia vendida y de 140 € por cada historia del arte. Calcula la función de beneficio neto $B(x, y)$ (es decir, descontando el coste del equipo), estudia su continuidad, calcula $B(15, 20)$ e interpreta el resultado.

14. La función de costes de una empresa en un instante t (expresado en años) es

    $$C(x, y, t) = (100 + 20x + 10y)\, e^{0.01t},$$

    donde x e y son las cantidades producidas de cada uno de los dos artículos que fabrica.

    Calcula el dominio matemático de la función C y el subdominio con sentido económico. Se entiende que $t = 0$ es el año actual, pero la empresa ya existía en años anteriores.

    Suponiendo que la producción no se altera de un año al siguiente, ¿en qué porcentaje aumentará el coste de la producción? (Es decir, se pregunta el porcentaje de aumento que supone pasar de t a $t+1$.) En el año actual $t = 0$ la producción ha sido $(x, y) = (50, 30)$, mientras que para el año siguiente se prevé $(x, y) = (51, 32)$. Calcula ΔC para este periodo.

15. Depositamos un capital de 30000 € durante 5 años a un 8% de interés continuo anual. Calcula el capital final y el interés efectivo anual. Calcula el incremento medio anual del capital.

16. Un banco nos ofrece comprar unos bonos por 5.000 € por los que dentro de 3 años nos dará 7000 €, mientras que otro nos ofrece bonos por valor de 4000 €, por los que dentro de 5 años nos dará 7000 €. Determina el tipo de interés continuo que nos ofrece cada banco. ¿Qué inversión es más rentable?

17. En la teoría del consumidor se considera una *función de utilidad* que expresa la satisfacción que obtiene un consumidor con cada adquisición posible. Por ejemplo, si consideramos un mercado con cuatro productos A, B, C, D y representamos las adquisiciones posibles como cuádruplas $\bar{c} = (x, y, z, w)$ (donde x es la cantidad adquirida de A, y la cantidad de B, etc.) entonces una función de utilidad podría ser

    $$U(x, y, z, w) = x(y + z)(2 - e^{-w}).$$

Determina el dominio matemático de U y el subdominio con sentido económico. Estudia la continuidad de U en uno y otro. ¿Qué utilidad tiene para el consumidor una compra con $x = 0$?, ¿cómo se interpreta esto? Dos de los bienes son sustitutivos, ¿cuáles? Justifica la respuesta. Supongamos que un consumidor considera satisfactoria la utilidad proporcionada por $x = 10$, $y = 5$, $z = 1$ y que las centésimas de utilidad le resultan inapreciables. ¿Cuál es la máxima cantidad de D que le interesaría adquirir? (Observa que desde un punto de vista estrictamente matemático la respuesta sería "infinito".)

18. Un trabajador empieza cobrando un salario de 500 € mensuales con un incremento anual del 4%. ¿Cuál de las dos funciones siguientes refleja más fielmente el salario $S(t)$ en función del tiempo durante los primeros 5 años?

$$S_1(t) = \begin{cases} 500 & \text{si } 0 \le t < 1 \\ 540 & \text{si } 1 \le t < 2 \\ 540.80 & \text{si } 2 \le t < 3 \\ 562.43 & \text{si } 3 \le t < 4 \\ 584.93 & \text{si } 4 \le t < 5, \end{cases} \qquad S_2(t) = 500(1.04)^t.$$

¿Qué diferencia hay entre ambas funciones desde un punto de vista matemático? Calcula el porcentaje de error relativo que cometemos al considerar que $S(3.5) = S_2(3.5)$.

6. Derivación

En economía, no sólo es importante determinar magnitudes que reflejen una situación dada, sino también estudiar cómo varían estas magnitudes y cómo influyen sobre unas las variaciones de otras. Así, por ejemplo, la inflación es una medida de la variación de los precios a lo largo del tiempo; si dos países tienen la misma tasa de paro, pero la de uno está creciendo y la de otro decreciendo, entonces sus economías son diferentes; un empresario puede estar interesado en estimar la variación de sus beneficios que ocasionaría un incremento en la producción, o un incremento en el gasto publicitario, etc. En este tema introduciremos las herramientas matemáticas básicas para cuantificar la variación de una magnitud dada y estudiar la relación entre las variaciones de unas magnitudes y las de otras.

6.1 Incrementos parciales

Definición Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función escalar definida en un abierto D y $\bar{x} \in D$, llamaremos *función de incrementos parciales* de f respecto de la variable x_i a la función

$$\Delta_{x_i} f(\bar{x}) = f(x_1, \ldots, x_i + \Delta x_i, \ldots, x_n) - f(x_1, \ldots, x_i, \ldots, x_n).$$

De este modo:

> La función de incrementos parciales de una función f respecto a x_i es otra función cuyas variables son las de f más la nueva variable Δx_i, y nos permite calcular el incremento que experimenta f cuando la variable x_i se incrementa en la cantidad Δx_i. Si particularizamos a un punto $\bar{x}$, obtenemos una función cuya única variable es Δx_i.

Ejemplo *Una empresa fabrica dos productos A y B, de modo que su función de beneficios es*

$$B(x, y) = x^2 + 3y^2 - xy - 20,$$

donde x e y son, respectivamente, las cantidades producidas de A y B.

a) *Calcula las funciones de incrementos parciales $\Delta_x B$ e $\Delta_y B$.*
b) La producción actual de la empresa es $(x, y) = (200, 150)$. *Calcula las funciones de incrementos parciales para esta producción.*

c) La empresa tiene la posibilidad de incrementar en 2 unidades la producción de A o de B. Determina cuál de las dos opciones es más conveniente.

SOLUCIÓN:

a) El incremento respecto de x es

$$\begin{aligned}
\Delta_x B(x,y) &= B(x+\Delta x, y) - B(x,y) \\
&= (x+\Delta x)^2 + 3y^2 - (x+\Delta x)y - 20 - (x^2+3y^2-xy-20) \\
&= x^2 + 2x\Delta x + \Delta x^2 + 3y^2 - xy - y\Delta x - 20 - x^2 - 3y^2 + xy + 20,
\end{aligned}$$

luego en definitiva:

$$\Delta_x B(x,y) = (2x-y)\Delta x + \Delta x^2.$$

Similarmente:

$$\begin{aligned}
\Delta_y B(x,y) &= B(x, y+\Delta y) - B(x,y) \\
&= x^2 + 3(y+\Delta y)^2 - x(y+\Delta y) - 20 - (x^2+3y^2-xy-20) \\
&= x^2 + 3(y^2 + 2y\Delta y + \Delta y^2) - xy - x\Delta y - 20 - x^2 - 3y^2 + xy + 20 \\
&= x^2 + 3y^2 + 6y\Delta y + 3\Delta y^2 - xy - x\Delta y - 20 - x^2 - 3y^2 + xy + 20,
\end{aligned}$$

luego

$$\Delta_y B(x,y) = (6y-x)\Delta y + 3\Delta y^2.$$

b) Sustituimos $x = 200$ e $y = 150$ en las expresiones anteriores:

$$\Delta_x B(200,150) = 250\Delta x + \Delta x^2, \qquad \Delta_y B(200,150) = 700\Delta y + 3\Delta y^2.$$

c) Al aumentar la producción de A (es decir, $\Delta x = 2$), el beneficio se incrementa en

$$\Delta_x B(200,150)(2) = 504 \text{ u.m.},$$

mientras que si $\Delta y = 2$ entonces el beneficio se incrementa en

$$\Delta_y B(200,150)(2) = 1412 \text{ u.m.}$$

Por consiguiente es preferible incrementar la producción de B. ■

6.2 Derivadas parciales

Las funciones de incrementos suelen ser complicadas. Sin embargo, si estamos dispuestos a renunciar a calcular el valor exacto de los incrementos de una función y sustituirlo por una aproximación razonable, podemos obtener una fórmula relativamente sencilla.

Definición Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar definida en un abierto D. Para cada punto $\bar{p} \in D$, definimos la *derivada parcial* de f respecto de la variable x_i en el punto $\bar{p}$ como

$$\left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}} = \lim_{\Delta x_i \to 0} \frac{\Delta_{x_i} f(\bar{p})}{\Delta x_i} = \lim_{\Delta x_i \to 0} \frac{f(p_1, \ldots, p_i + \Delta x_i, \ldots, p_n) - f(p_1, \ldots, p_n)}{\Delta x_i}.$$

Observemos ante todo que una derivada parcial es un límite y, por lo tanto, no tiene por qué existir. En el supuesto de que exista, de acuerdo con la interpretación del límite vista en el tema anterior, tenemos que si Δx_i es suficientemente pequeño, entonces

$$\frac{\Delta_{x_i} f(\bar{p})}{\Delta x_i} \approx \left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}},$$

luego

$$\boxed{\Delta_{x_i} f(\bar{p}) \approx \left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}} \Delta x_i.}$$

Esta fórmula contiene la interpretación de las derivadas parciales. Para enunciarla matemáticamente conviene introducir el vocabulario usual en economía:

> Una *unidad marginal* de una magnitud es una unidad pequeña en comparación con el valores que toma dicha magnitud. Un *incremento marginal* de una magnitud es un incremento pequeño en comparación con el valor que toma dicha magnitud.

En la mayoría de los ejemplos que vamos a considerar será fundamental que las unidades consideradas sean marginales. Por ejemplo, si la inversión de una empresa en producción es del orden de varios miles de euros, podremos tomar como unidad monetaria (marginal) para estudiarla un centenar de euros, mientras que si queremos estudiar el gasto mensual de cierta familia (por ejemplo, del orden de 300 €), un centenar de euros no servirá como unidad marginal, pero sí servirá, en cambio, un euro. En estos términos, la interpretación de la fórmula anterior es:

> La derivada parcial de una función f respecto de una variable x en un punto $\bar{p}$ representa el incremento que experimenta f por cada unidad marginal que aumenta la variable x alrededor del punto $\bar{p}$ (suponiendo que las demás variables no se alteran).

Función derivable Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ una función vectorial definida en un abierto D. dado $\bar{p} \in D$, decimos que f es derivable en $\bar{p}$ si existen las derivadas parciales respecto de todas las variables de sus funciones coordenadas, es decir existen

$$\frac{\partial f_j}{\partial x_i}(\bar{p}), \quad i = 1, \ldots n, \quad j = 1, \ldots m.$$

En tal caso,

$$\frac{\partial f}{\partial x_i}(\bar{p}) = \left(\frac{\partial f_1}{\partial x_i}(\bar{p}), \ldots, \frac{\partial f_m}{\partial x_i}(\bar{p})\right).$$

Si una función tiene derivada parcial respecto de una variable en todos los puntos de su dominio, entonces la derivada parcial es otra función con las mismas variables:

Definición Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función derivable respecto de x_i en todos[1] los puntos del abierto D, entonces definimos la *función derivada parcial* respecto de x_i

$$\frac{\partial f}{\partial x_i} : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$$

como la función que a cada punto $\bar{p} \in D$ le asigna la derivada parcial de f respecto de x_i en dicho punto.

Ejemplo *Calcula las derivadas parciales de la función de beneficios*

$$B(x, y) = x^2 + 3y^2 - xy - 20$$

considerada en la sección anterior.

Solución:

$$\frac{\partial B}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta_x B}{\Delta x} = \lim_{\Delta x \to 0} \frac{(2x - y)\Delta x + \Delta x^2}{\Delta x} = \lim_{\Delta x \to 0} 2x - y + \Delta x = 2x - y,$$

$$\frac{\partial B}{\partial y} = \lim_{\Delta y \to 0} \frac{\Delta_y B}{\Delta y} = \lim_{\Delta y \to 0} \frac{(6y - x)\Delta y + 3\Delta y^2}{\Delta y} = \lim_{\Delta y \to 0} 6y - x + 3\Delta y = 6y - x.$$

■

En particular,

$$\left.\frac{\partial B}{\partial x}\right|_{(200,150)} = 250, \qquad \left.\frac{\partial B}{\partial y}\right|_{(200,150)} = 700.$$

De acuerdo con la interpretación de la derivada, un incremento de 2 unidades en la producción de A produce un incremento de beneficios de

$$\Delta_x B(200, 150)(2) \approx \left.\frac{\partial B}{\partial x}\right|_{(200,150)} \Delta x = 250 \cdot 2 = 500 \text{ u.m.},$$

mientras que el mismo incremento en la producción de B da lugar a un incremento de los beneficios de

$$\Delta_y B(200, 150)(2) \approx \left.\frac{\partial B}{\partial y}\right|_{(200,150)} \Delta y = 700 \cdot 2 = 1400 \text{ u.m.}$$

Estas estimaciones no son exactas, pues los incrementos exactos los calculamos en la sección anterior y eran de 504 y 1412 u.m. respectivamente, pero observamos que la aproximación con las derivadas parciales es buena y, como veremos enseguida, es mucho más fácil de calcular.

[1] En realidad no es necesario que exista la derivada en todo el dominio. Puede existir en un subconjunto del mismo, y en este caso el dominio de la derivada sería este subconjunto.

Reglas de derivación Al comparar la definición de derivada parcial con la definición de derivada de una función de una variable se llega fácilmente a la siguiente conclusión:

> La derivada parcial de una función f respecto de una variable x_i puede calcularse con las mismas reglas de derivación válidas para funciones de una variable sin más que considerar como constantes a las demás variables de f.

Por ejemplo, la derivada de $3x^2$ es $3\cdot 2x = 6x$ y, por la misma regla, la derivada respecto de x de la función yx^2 es $y2x = 2yx$. Hemos tratado a y como si fuera la constante 3. Para derivarla respecto de y consideramos a x^2 como si fuera una constante, y así, al igual que la derivada de $y5$ es 5, la derivada de yx^2 es x^2.

Ejemplo *La función de beneficios de una empresa es $B(x, y) = 108\sqrt{x}\,\sqrt[3]{y}$, donde x, y son las cantidades invertidas respectivamente en la producción de dos artículos A y B. La producción actual es $(x, y) = (4, 27)$, pero la empresa dispone de 0.3 u.m. para aumentar la producción. ¿Le convendrá más destinarlas al artículo A o al artículo B?*

Solución: Calculamos las derivadas parciales

$$\frac{\partial B}{\partial x} = 54\frac{1}{\sqrt{x}}\,\sqrt[3]{y}, \qquad \frac{\partial B}{\partial y} = 36\sqrt{x}\,\frac{1}{\sqrt[3]{y^2}}.$$

Para la producción actual tenemos

$$\left.\frac{\partial B}{\partial x}\right|_{(4,27)} = 81, \qquad \left.\frac{\partial B}{\partial y}\right|_{(4,27)} = 8.$$

Esto significa que por cada unidad que la empresa aumentara la inversión en el producto A los beneficios aumentarían aproximadamente en 81 u.m., mientras que la misma inversión en el producto B produciría un incremento aproximado de 8 u.m. Más concretamente, con un incremento de 0.3 u.m. los beneficios aumentarían aproximadamente en

$$\Delta_x B \approx \left.\frac{\partial B}{\partial x}\right|_{(4,27)} \Delta x = 81\cdot 0.3 = 24.3 \text{ u.m.} \qquad \Delta_y B \approx \left.\frac{\partial B}{\partial y}\right|_{(4,27)} \Delta y = 8\cdot 0.3 = 2.4.$$

Así pues, es preferible incrementar la producción del artículo A. ■

Ejemplo *Calcula los incrementos exactos de la función de beneficios correspondientes al ejemplo anterior y compáralos con las aproximaciones que hemos obtenido con las derivadas parciales.*

Solución: Los incrementos exactos de la función de beneficios son:

$$\Delta_x B(4, 27)(0.3) = B(4.3, 27) - B(4, 27) = 671.86 - 648 = 23.86,$$

$$\Delta_y B(4, 27)(0.3) = B(4, 27.3) - B(4, 27) = 650.39 - 648 = 2.39.$$

Vemos que, efectivamente, los incrementos exactos del beneficio son muy similares a los que hemos calculado aproximadamente. ■

6.3 Aplicaciones de las derivadas parciales

Magnitudes marginales En economía es frecuente referirse a la derivada de una magnitud añadiéndole a esta el calificativo "marginal". Por ejemplo, si $C(x, y)$ es una función de costes, donde x e y son las cantidades producidas de dos artículos, la derivada

$$\frac{\partial C}{\partial x}$$

es el *coste marginal* respecto de x, es decir, el (incremento del) coste que ocasionaría aumentar en una unidad la producción del primer artículo. Igualmente, si $B(t)$ son beneficios de una empresa en un tiempo t, entonces el *beneficio marginal* es

$$\frac{dB}{dt},$$

que se interpreta como el beneficio que se obtiene al pasar una unidad (marginal) de tiempo. Si $U(x, y)$ es la utilidad que obtiene un consumidor al adquirir cantidades x e y de dos productos A y B, entonces la *utilidad marginal* respecto de y es la derivada

$$\frac{\partial U}{\partial y},$$

es decir, (el incremento de) la utilidad que obtendría el consumidor al gastar una unidad monetaria más en el producto B, etc. Es importante señalar que éstas y todas las interpretaciones particulares del adjetivo "marginal" en su uso en economía son casos particulares de la interpretación general de las derivadas parciales.[2]

A menudo se usa la palabra "acumulado" por oposición a "marginal". Por ejemplo, si decimos que el beneficio acumulado por una empresa en un tiempo t (expresado en años) es

$$B(t) = 50000 \ln(1 + t) \text{ €}$$

esto significa que en su primer año obtuvo $B(1) = 34657$ €, en los dos primeros años $B(2) = 54930$ € (con lo que el beneficio acumulado durante el segundo año únicamente fue de $B(2) - B(1) = 20273$ €, etc. Por otra parte, el beneficio marginal será

$$B_m(t) = \frac{50000}{1 + t} \text{ €/año},$$

lo cual significa que la empresa comenzó acumulando beneficios a un ritmo de $B_m(0) = 50000$ €/año, pero al final de su primer año la tasa de incremento de los beneficios se había reducido a $B_m(1) = 25000$ €/año, etc. Es frecuente que no se especifique si una función corresponde a cantidades acumuladas o marginales, pues esto puede deducirse de las unidades: si nos hablan de unos beneficios de $5000t$ € hay que entender que son beneficios acumulados, pero si nos dicen $5000t$ €/año entonces han de ser beneficios marginales.

[2]En algunos libros se encuentra otra definición distinta de "marginal". Por ejemplo, "la utilidad marginal respecto de un bien A es la utilidad que el consumidor obtiene de la última unidad monetaria que invierte en A." Esta interpretación corresponde en realidad a la matemática discreta y no a la matemática continua que aquí estamos considerando. No obstante, para cuestiones en las que la distinción matemática continua/matemática discreta no es relevante, ambas interpretaciones son equiparables, en cuanto que llevan a las mismas consecuencias.

En general, las unidades de una derivada

$$\frac{\partial f}{\partial x}$$

son unidades de f/unidad de x. Por ejemplo, un coste marginal o una utilidad marginal se expresa en unidades monetarias/unidad de producto, etc.

El signo de las derivadas Recordemos que un incremento puede representar un aumento o una disminución según que su signo sea positivo o negativo. Como consecuencia:

> La derivada de una función f respecto de una variable x_i será positiva si un aumento de x da lugar a un aumento de f, y será negativa si, por el contrario, un aumento de x da lugar a una disminución de f.

Por ejemplo, en condiciones normales, la derivada de la demanda de un producto respecto de su precio será negativa, pues un aumento del precio da lugar a una disminución de la demanda. Sin embargo, la derivada de la demanda respecto al gasto en publicidad será positiva, pues un aumento del gasto publicitario da lugar a un aumento en la demanda.

Incrementos porcentuales Si una función F depende (entre otras) de la variable x, la derivada

$$\frac{\partial F}{\partial x}$$

representa el incremento *absoluto* que produce en F un incremento de una unidad en x, aunque a menudo es más útil determinar el incremento relativo al valor de F, es decir, es más informativo saber que el paro ha descendido en un 2% que no saber cuántos parados menos hay. Para obtener un incremento porcentual basta observar que tomando el valor de F como el 100%, la derivada anterior supone un porcentaje de

$$\Delta_{\%}F \approx \frac{100}{F}\frac{\partial F}{\partial x}.$$

Si no ponemos el 100 tenemos el tanto por uno de incremento. Por ejemplo, la tasa de crecimiento porcentual de un capital sujeto a un interés continuo i respecto del tiempo, es decir, la tasa de crecimiento porcentual de la función $C = C_0e^{it}$ respecto de t, es

$$\frac{100}{C}\frac{\partial C}{\partial t} = \frac{100}{C_0e^{it}}iC_0e^{it} = 100i.$$

Si no ponemos el 100, concluimos que i es el tanto por uno de incremento que experimenta el capital en cada instante. Más concretamente:

> El interés continuo es el interés generado por cada unidad de capital en una unidad marginal de tiempo.

Elasticidad Hay casos en que la disparidad entre las unidades de la función y de la variable hacen aconsejable considerar otro índice de crecimiento relativo. Por ejemplo, si D es la demanda de un producto y p es su precio, un valor grande (normalmente negativo) para

$$\frac{\partial D}{\partial p}$$

indicará que una pequeña variación en el precio provoca una gran variación en la demanda, mientras que un valor pequeño de la derivada indica que la demanda apenas se altera por la variación del precio. Ahora bien, hemos de tener presente que no podemos estimar si la derivada es grande o pequeña sin tener en cuenta las unidades que estamos considerando, principalmente porque la demanda vendrá dada en unidades de producto y el precio en unidades monetarias.

Para evitar este inconveniente lo usual es considerar incrementos relativos tanto de la variable como de la función. Es decir, supongamos que el precio del artículo se incrementa en un 1%. Esto significa que

$$\Delta p = \frac{p}{100}.$$

De acuerdo con la interpretación de la derivada, el incremento que esto ocasionará en la demanda será aproximadamente de

$$\Delta D \approx \frac{\partial D}{\partial p}\,\frac{p}{100}.$$

Ahora bien, ¿qué porcentaje de la demanda supone este incremento? Será

$$\frac{100\,\Delta D}{D} \approx \frac{100}{D}\,\frac{\partial D}{\partial p}\,\frac{p}{100} = \frac{p}{D}\,\frac{\partial D}{\partial p}.$$

El valor

$$E(D,p) = \frac{p}{D}\,\frac{\partial D}{\partial p}$$

recibe el nombre de *elasticidad* de la función D respecto de la variable p (hay quien lo define con un signo negativo para que sea positiva en el caso de la demanda/precio) y, según acabamos de ver, representa el porcentaje en que se incrementa la función D por cada 1% que se incrementa la variable p. Este valor es independiente de las unidades en que se expresen D y p.

Por ejemplo, la elasticidad respecto del tiempo de un capital sujeto a un interés continuo i es

$$E(C,t) = \frac{t}{C}\,\frac{\partial C}{\partial t} = it.$$

Esto se interpreta como que, cuanto más tiempo pasa, el incremento del capital se vuelve más sensible al paso del tiempo.

6.4 Conceptos relacionados con las derivadas

En esta sección introduciremos varias nociones matemáticas que nos serán de utilidad más adelante, junto con algunos resultados relacionados con ellas. Empezamos por la posibilidad de derivar varias veces una misma función.

Derivadas sucesivas Si las derivadas parciales de una función derivable son también funciones derivables, entonces podemos calcular las llamadas *derivadas segundas,* para las cuales se usa la notación que ilustra el ejemplo siguiente:

Sea $f(x,y) = x^2 \operatorname{sen} y$. Las derivadas primeras de f son

$$\frac{\partial f}{\partial x} = 2x \operatorname{sen} y, \qquad \frac{\partial f}{\partial y} = x^2 \cos y.$$

Si volvemos a derivar $\frac{\partial f}{\partial x}$ otra vez respecto de x obtenemos

$$\frac{\partial^2 f}{\partial x^2} = 2 \operatorname{sen} y.$$

Si, en cambio, tomamos $\frac{\partial f}{\partial x}$ y la derivamos respecto de y obtenemos

$$\frac{\partial^2 f}{\partial x \partial y} = 2x \cos y.$$

Así, vemos que el exponente superior indica el número de veces que hemos derivado en total, y en la parte inferior indicamos las variables respecto a las que hemos derivado en el orden en que lo hemos hecho. Similarmente se calculan

$$\frac{\partial^2 f}{\partial y \partial x} = 2x \cos y, \qquad \frac{\partial^2 f}{\partial y^2} = -x^2 \operatorname{sen} y.$$

Por ejemplo, si $f(x,y,z)$ es una función de tres variables, la notación

$$\frac{\partial^5 f}{\partial x \partial y^2 \partial z^2}$$

representa a la función que resulta de derivar f cinco veces: primero respecto de x, luego dos veces respecto de y y luego dos veces respecto de z. Las funciones obtenidas derivando varias veces una función se llaman *derivadas parciales sucesivas.*

Definición Una función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ definida en un abierto D es de clase C^n en D, donde $n \geq 1$ es un número natural, si existen las derivadas parciales de f hasta orden n y todas ellas son continuas en D.

Si f tiene derivadas parciales de todos los órdenes posibles y todas ellas son continuas, entonces se dice que f es de clase C^∞ en D.

La mayoría de las funciones definidas analíticamente son de clase C^∞ en su dominio, como consecuencia del teorema siguiente:

Teorema
1) *Si f, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ son funciones de clase C^∞ en el abierto D y $\alpha \in \mathbb{R}$, entonces también son C^∞ en D las funciones $f \pm g$, αf, fg y f/g es C^∞ donde g no se anula.*
2) *Todo polinomio es una función C^∞.*
3) *La composición de funciones C^∞ es una función C^∞.*
4) *Las funciones e^x, $\ln x$, $\operatorname{sen} x$, $\cos x$ son C^∞ en todos los puntos de su dominio, mientras que $\sqrt[n]{x}$ es C^∞ en todos los puntos del dominio distintos de 0.*

Vector gradiente Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función escalar derivable en el abierto D, definimos su *vector gradiente* como el vector formado por sus derivadas parciales:

$$\nabla f(\bar{x}) = \left(\frac{\partial f}{\partial x_1}(\bar{x}), \ldots, \frac{\partial f}{\partial x_n}(\bar{x}) \right)$$

Ejemplo *Calcula el gradiente de la función $f(x, y, z) = x^2 y z^3$.*

SOLUCIÓN: El gradiente es

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = (2xyz^3, x^2 z^3, 3x^2 y z^2).$$

∎

Matriz jacobiana Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es una función vectorial derivable en el abierto D se define la *matriz jacobiana* de f como la matriz formada por todas las derivadas parciales primeras de sus funciones coordenadas $f_1, \ldots, f_m$ del modo siguiente:

$$Jf(\bar{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\bar{x}) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(\bar{x}) & \cdots & \frac{\partial f_m}{\partial x_n}(\bar{x}) \end{pmatrix}$$

Ejemplo *Calcula la matriz jacobiana de la función*

$$f(x, y, z) = (x^2 yz, x + y, xyz)$$

SOLUCIÓN: Se tiene

$$Jf(x, y, z) = \begin{pmatrix} 2xyz & x^2 z & x^2 y \\ 1 & 1 & 0 \\ yz & xz & xz \end{pmatrix}$$

∎

Matriz hessiana Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función escalar dos veces derivable respecto de todas las variables en el abierto D, definimos su *matriz hessiana* como la matriz formada por sus derivadas parciales de orden 2 del modo siguiente:

$$Hf(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\bar{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\bar{x}) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\bar{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\bar{x}) \end{pmatrix}$$

Ejemplo *Calcula la matriz hessiana de la función* $f(x,y) = x^2 \operatorname{sen} y$.

SOLUCIÓN:

$$Hf(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 2 \operatorname{sen} y & 2x \cos y \\ 2x \cos y & -x^2 \operatorname{sen} y \end{pmatrix}.$$

■

La matriz hessiana de una función de clase C^2 es siempre una matriz simétrica, como se deduce del teorema siguiente:

Teorema de Schwarz *Si* $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *es una función escalar de clase* C^2 *en el abierto* D*, entonces, para* $i,\ j = 1, \ldots, n$ *se cumple que*

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Con este teorema se tiene que el orden de la derivación no afecta al resultado.

6.5 Algunas demostraciones

Los resultados como el teorema de Schwarz tienen pruebas complejas, pero las propiedades básicas de las derivadas requieren únicamente aplicar la definición y las propiedades de los límites. Veremos algunas demostraciones que nos puedan ayudar a familiarizarnos con la noción de derivada:

Teorema *Si* $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ *es la función constante* $f(\bar{x}) = c$*, entonces*

$$\left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}} = 0.$$

DEMOSTRACIÓN: Por simplificar la notación supondremos que $n = 2$ y que $x_i = x$. El caso general es formalmente análogo. Consideramos un punto arbitrario $\bar{p} = (x,y) \in \mathbb{R}^2$. Se tiene

$$\frac{\partial f}{\partial x}(\bar{p}) = \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x,y)}{\Delta x} = \lim_{\Delta x \to 0} \frac{c - c}{\Delta x} = \lim_{\Delta x \to 0} \frac{0}{\Delta x} = \lim_{\Delta x \to 0} 0 = 0.$$

■

Teorema *Sea* $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *una función derivable en un punto* $\bar{p} \in D$*, donde* D *abierto, y* $\alpha \in \mathbb{R}$*. Entonces,* αf *es derivable en* $\bar{p}$ *y*

$$\left.\frac{\partial \alpha f}{\partial x_i}\right|_{\bar{p}} = \alpha \left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}}.$$

DEMOSTRACIÓN: Como en el teorema anterior, tomaremos $n = 2$, $x_i = x$ y $\bar{p} = (x, y)$. El caso general es formalmente análogo.

$$\left.\frac{\partial \alpha f}{\partial x}\right|_{\bar{p}} = \lim_{\Delta x \to 0} \frac{\alpha f(x + \Delta x, y) - \alpha f(x, y)}{\Delta x} = \lim_{\Delta x \to 0} \alpha \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x}$$

$$= \alpha \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} = \alpha \left.\frac{\partial f}{\partial x}\right|_{\bar{p}}.$$

■

Teorema *Si f, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ son funciones derivables en un punto $\bar{p} \in D$, donde D abierto, entonces $f + g$ es derivable en $\bar{p}$ y*

$$\left.\frac{\partial (f + g)}{\partial x_i}\right|_{\bar{p}} = \left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}} + \left.\frac{\partial g}{\partial x_i}\right|_{\bar{p}}.$$

DEMOSTRACIÓN: Como en los teoremas anteriores, podemos suponer $n = 2$, $\bar{p} = (x, y)$, $x_i = x$. El caso general es formalmente análogo.

$$\begin{aligned}
\left.\frac{\partial (f + g)}{\partial x}\right|_{\bar{p}} &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) + g(x + \Delta x, y) - f(x, y) - g(x, y)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y) + g(x + \Delta x, y) - g(x, y)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left(\frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} + \frac{g(x + \Delta x, y) - g(x, y)}{\Delta x} \right) \\
&= \lim_{\Delta x \to 0} \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} + \lim_{\Delta x \to 0} \frac{g(x + \Delta x, y) - g(x, y)}{\Delta x} \\
&= \left.\frac{\partial f}{\partial x}\right|_{\bar{p}} + \left.\frac{\partial g}{\partial x}\right|_{\bar{p}}.
\end{aligned}$$

■

6.6 Ejercicios

1. Calcula las derivadas parciales en el punto $(1, 2)$ de las funciones

$$f(x, y) = \begin{cases} x^2 + y & \text{si } x > y, \\ x + y^2 & \text{si } x \le y. \end{cases} \qquad g(x, y) = \begin{cases} x^2 + y & \text{si } y \ge x + 1, \\ x + y^2 & \text{si } y < x + 1. \end{cases}$$

2. Calcula las derivadas parciales de las funciones siguientes:

   (a) $f_1(x, y, z) = 3x^5 y + xy^4 z + y^5 + 2z^2 + 5$,

   (b) $f_2(x, y) = x^5 \cos 3y^4$,

   (c) $f_3(x, y, z) = (x^2 + 2xy + yz^5)e^{x+2y-z+1}$,

(d) $f_4(x,y) = \cos^5(x+2y)$,

(e) $f_5(x,y) = \cos(x+2y)^5$,

(f) $f_6(x,y) = \sqrt{x^2y^3}$,

(g) $f_7(x,y) = \sqrt[3]{2^{x+y}}$,

(h) $f_8(x,y) = \dfrac{x^2y - y^3}{x+2y-3}$,

(i) $f_9(x,y,z) = \dfrac{x\cos y}{\sqrt{y+2z}}$,

(j) $f_{10}(x,y) = \ln^3(x/y)$.

3. Calcula el vector gradiente de las funciones siguientes:

(a) $f(x,y,z) = x^5y - 3x^2z + xyz - 3y + 2$,

(b) $g(x,y) = x/y^3$,

(c) $h(x,y,z) = \sqrt{x}\,\mathrm{sen}(x^2+y^2+z^2)$,

(d) $p(u,v) = \dfrac{u+v}{u^2+v^2}$,

(e) $q(s,t) = e^t\ln(s^2+t+1)$.

4. Calcula la matriz hessiana de las funciones siguientes:

(a) $f(x,y) = x^2 - 2xy^3 + 3x$,

(b) $g(x,y,z) = e^{x^2}\,\mathrm{sen}(y+z)$,

(c) $h(u,v,w) = u + 2v - w$,

(d) $r(a,b) = \dfrac{1}{a+b^5}$.

5. Calcula la matriz jacobiana de las funciones siguientes:

(a) $f(x,y,z) = (x+y^2, xyz)$,

(b) $g(x,y) = (x, y, \ln(x+y), x\,\mathrm{sen}\,y)$,

(c) $h(u,v,w) = (5, u^4+v, e^{vw^2})$,

(d) $C(p) = (p^3, 1/p, \sqrt{p})$.

6. Calcula la matriz jacobiana de la función $f(x,y,z) = (x^2 - yz, x - z^3, xyz)$ en el punto $(1,0,0)$.

7. Calcula el vector gradiente de la función $f(x,y,z) = x^2 - \mathrm{sen}(yz)$ en el punto $(1,0,0)$.

8. Calcula la matriz hessiana en el punto $(3,1)$ de las funciones cuyos gradientes son los indicados:

(a) $\nabla f(x,y) = (3x^3 + 3y^3 - 2, 9xy^2)$,

(b) $\nabla g(x,y) = (y/x, \ln x)$

(c) $\nabla h(x, y) = (2x, 2y)$

9. ¿Existe una función cuyo vector gradiente sea $\nabla f = (x^2 + 2y, y^2 + 3x)$? En caso afirmativo, calcula su matriz hessiana.

10. Calcula las funciones de incrementos parciales de las funciones siguientes:

    (a) $f(x, y) = x^2 y$,
    (b) $g(x, y, z) = x(y + z)$,
    (c) $h(u, v, w) = u - 3v + w + 4$.

    Utiliza las expresiones obtenidas para calcular las derivadas parciales.

11. Indica el signo que tendrán en condiciones normales las derivadas siguientes:

    (a) La derivada del salario de un trabajador respecto al tiempo.
    (b) La derivada parcial de la demanda de un artículo respecto de su precio.
    (c) La derivada parcial del volumen de ventas de una empresa respecto de su inversión en publicidad.
    (d) La derivada parcial del ahorro medio de los habitantes de un país respecto del índice de precios.
    (e) La derivada respecto al tiempo de la población de un país en el que cada familia tiene una media de 1.8 hijos.
    (f) La derivada del índice general de la bolsa de Madrid respecto del tiempo.

12. Si $P(t)$ es el producto interior bruto de un país en un tiempo t, ¿qué es la derivada $\frac{dP}{dt}$?

    (a) la inflación del país en un tiempo t,
    (b) el crecimiento económico del país en un tiempo t,
    (c) no tiene interpretación económica.

13. El precio del petróleo es una función P que depende —entre otras variables— de la oferta x de crudo en el mercado. ¿Cuál será —en condiciones normales— el signo de la derivada $\frac{\partial P}{\partial x}$?

14. Una editorial A es una de las principales suministradoras de libros a una pequeña ciudad, aunque tiene una única competidora B. La empresa estima que la demanda de sus libros en la ciudad depende del precio medio al que los vende p_1, del precio medio a que vende los libros la editorial B y del precio medio de los artículos de primera necesidad. Si la función de demanda de los libros de A es $D(p_1, p_2, p_3)$ y la empresa estima que, para los precios actuales $\bar{p}_0$, se tiene

    $$\left.\frac{\partial D}{\partial p_1}\right|_{\bar{p}_0} = -2, \qquad \left.\frac{\partial D}{\partial p_2}\right|_{\bar{p}_0} = -1, \qquad \left.\frac{\partial D}{\partial p_3}\right|_{\bar{p}_0} = 2,$$

    ¿Cuál de las variables p_2, p_3 representa —presumiblemente— a los precios de la editorial B y cuál a los precios de los artículos de primera necesidad? ¿Qué efecto tendría para la editorial una rebaja media de sus precios de 0.8 unidades monetarias?

15. Sea $C(x,y)$ la función de costes de una empresa, donde x e y son las cantidades producidas de dos artículos A y B. Explica la diferencia de interpretación entre
$$\left.\frac{\partial C}{\partial x}\right|_{(400,200)} \quad \text{y} \quad \left.\frac{\partial C}{\partial x}\right|_{(40,20)}.$$

16. El IPC de un cierto país en un instante t (expresado en años) viene dado por la fórmula
$$P = e^{\sqrt{(1+t/50)^3}}.$$
   (a) Calcula la inflación del país, es decir, el porcentaje de aumento de los precios:
   $$I = \frac{100}{P}\frac{\partial P}{\partial t}.$$
   ¿Qué tanto por ciento de inflación se tiene en $t = 0$?
   (b) Estudia el comportamiento de la inflación en $t = 0$. ¿Está aumentando o disminuyendo?
   (c) Calcula la tasa de incremento de la inflación en el país, es decir,
   $$T = \frac{100}{I}\frac{\partial I}{\partial t}.$$
   ¿Cuánto vale en $t = 0$?
   (d) ¿Cómo varía la tasa de incremento de la inflación del país?, ¿crece o decrece?
   (e) Explica la frase siguiente:

   > *En la crisis de 1972 el presidente Nixon anunció que la tasa de incremento de la inflación estaba descendiendo. Ésta fue la primera vez que un presidente usó la tercera derivada como argumento para su reelección.* Hugo Rossi, Notices of the AMS, v. 43, nº 10, octubre 1996.

17. La tasa de paro (porcentual) de dos países A y B viene dada por las funciones
$$P_A(t) = 8(1.1)^t \qquad \text{y} \qquad P_B(t) = 8(0.9)^t.$$
Determina la tasa de paro en la actualidad ($t = 0$). Estudia la evolución del paro en ambos países. ¿Cuál es más favorable?

18. Una empresa desea estudiar su función de costes, cuya expresión estimada es
$$C(x,y) = 3x^2 - xy + 2y^2 - y + 30,$$
donde x e y son las cantidades producidas de dos artículos. ¿A cuánto ascienden los costes fijos? Determina la expresión más simple posible para las funciones de incrementos parciales $\Delta_x C$ y $\Delta_y C$. Utiliza estas expresiones para calcular, por la definición de derivada parcial, las derivadas
$$\frac{\partial C}{\partial x} \quad \text{y} \quad \frac{\partial C}{\partial y}.$$

Calcula estas derivadas para una producción $(x, y) = (5, 3)$ e interprétalas. Calcula el incremento de coste que provocaría un incremento de la producción $\Delta y = 0.6$. Compáralo con la aproximación que proporcionan las derivadas.

19. Una empresa fabrica un artículo X a partir de dos factores de producción A y B. La función de producción es $P(x, y) = x + y + 0.005xy^2$ unidades de X, donde x e y son las cantidades de los factores de producción. Calcula la producción actual si se están empleando 100 unidades de A y 80 unidades de B. Calcula la producción marginal respecto de y para la producción actual. Indica su interpretación y las unidades en que viene expresada. Calcula el incremento que producción que puede obtenerse si la cantidad empleada del factor B pasa a ser de 82 unidades. Haz el cálculo exacto y el cálculo aproximado a partir de la producción marginal y compáralos. Ídem si se utilizan 200 unidades de B. Compara los resultados en ambos casos.

20. Sea $C(x)$ la función de costes de una empresa, donde x es la cantidad producida de un artículo.

    (a) Explica la diferencia de interpretación entre
    $$\left.\frac{\partial C}{\partial x}\right|_{10} \qquad \text{y} \qquad \left.\frac{\partial C}{\partial x}\right|_{1000}.$$
    (b) ¿Cuál es el signo que cabría esperar en estas derivadas?

    (c) ¿Cuál de las dos cabe esperar que sea mayor?

21. Sea $U(x)$ la función de utilidad de un consumidor, donde x es la cantidad adquirida de un artículo.

    (a) Explica la diferencia de interpretación entre
    $$\left.\frac{\partial U}{\partial x}\right|_{10} \qquad \text{y} \qquad \left.\frac{\partial U}{\partial x}\right|_{1000}.$$
    (b) ¿Cuál es el signo que cabría esperar en estas derivadas?

    (c) ¿Cuál de las dos cabe esperar que sea mayor?

22. Una empresa cuenta actualmente con un capital C. En cada instante consigue de su capital una rentabilidad r (tanto por uno). Supongamos que esta rentabilidad es mayor o igual que la rentabilidad media del mercado k. En cada instante, la empresa reinvierte un tanto por uno b de sus beneficios y abona el resto como dividendos a sus accionistas. Bajo el supuesto de que $0 < b < k/r$, un modelo económico afirma que el valor actual de las acciones de la empresa (calculado como el valor actual de los dividendos que la empresa generará en el futuro) viene dado por
    $$V = \frac{C(1-b)}{k - rb}.$$
    Determina el dominio matemático y el dominio económico de la función $V(C, k, r, b)$. Estudia el efecto que tiene sobre el valor V de las acciones una

reducción del porcentaje de beneficios destinado a dividendos, es decir, un aumento del tanto por uno reinvertido b.

23. Depositamos 500 € en un banco a un 3% de interés continuo anual. Utiliza la relación
$$\Delta C \approx \left.\frac{\partial C}{\partial t}\right|_0 \Delta t$$
para determinar aproximadamente sin usar calculadora nuestro saldo al cabo de 2 años. Usa la calculadora para determinar el error cometido. Repite el problema para 10 años y compara los errores. Explica la diferencia.

24. La función de demanda de un artículo es $D(p,r) = \ln\left(1 + \frac{2r}{p}\right)$, donde p es el precio y r la renta media de los consumidores. Determina el dominio matemático y el subdominio económico de la función D. Calcula la elasticidad (respecto al precio) para $(p,r) = (2,100)$. Interprétala.

25. El capital de una empresa durante un periodo de diez años $[0,10]$ viene dado por la función
$$C(t) = 500e^{3e^{0.01t}-3}.$$
Determina el capital con que contaba la empresa al principio del periodo y el capital final. Determina el beneficio marginal $B_m(t)$ de la empresa en cada instante t. Determina la rentabilidad de la empresa en cada instante t, es decir, el beneficio generado por cada unidad de capital disponible en un instante t:
$$R(t) = \frac{B_m(t)}{C(t)}.$$
Calcula la rentabilidad inicial y final de la empresa. Determina el tanto por ciento de incremento anual de la rentabilidad de la empresa.

26. La función de utilidad de un consumidor respecto de dos productos A y B es
$$U(x,y) = \ln(1+xy),$$
donde x e y son las cantidades de producto que adquiere. Supongamos que actualmente consume $(x,y) = (10,10)$.

    (a) Calcula la utilidad marginal respecto del producto A. Interpreta su signo.

    (b) Justifica matemáticamente esta afirmación: "Por cada unidad que aumenta el consumo de A, la utilidad marginal disminuye, es decir, el consumidor obtiene cada vez menos satisfacción adicional al incrementar su consumo de A".

    (c) Justifica matemáticamente esta afirmación: "Por cada unidad que aumenta el consumo de B la utilidad marginal de A aumenta, es decir, si el consumidor aumenta el consumo de B, entonces le es más útil aumentar el consumo de A."

    (d) Pon un ejemplo de dos productos para los que estas propiedades sean razonables.

7. Diferenciabilidad

En el tema anterior hemos estudiado cómo se comporta una función derivable cuando modificamos *una* de sus variables. Ahora vamos a ocuparnos de lo que sucede cuando modificamos varias variables simultáneamente.

7.1 Incrementos totales

Definición Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función escalar definida en un abierto D y $\bar{x} \in D$, llamaremos *función de incrementos* de f a la función

$$\Delta f(\bar{x}) = f(x_1 + \Delta x_1, \dots, x_n + \Delta x_n) - f(x_1, \dots, x_n).$$

De este modo:

La función de incrementos de una función f es otra función cuyas variables son las de f más las variables Δx_i, y nos permite calcular el incremento que experimenta f cuando cada variable x_i se incrementa en la cantidad Δx_i. Si particularizamos a un punto $\bar{p}$ queda una función de las variables Δx_i.

Ejemplo *Una empresa fabrica dos productos A y B, de modo que su función de beneficios es $B(x, y) = x^2 + 3y^2 - xy - 20$, donde x e y son, respectivamente, las cantidades producidas de A y B.*
a) Calcula la función de incrementos de B.
b) La producción actual de la empresa es $(x, y) = (200, 150)$. Calcula la función de incrementos para esta producción.
c) La empresa tiene la posibilidad de incrementar en 2 unidades la producción de A y en 1 unidad la de B. Determina el incremento de beneficio que obtendrá.

SOLUCIÓN:

a) Aplicacmos la definición:

$$\begin{aligned}\Delta B(x, y) &= B(x + \Delta x, y + \Delta y) - B(x, y)\\ &= (x + \Delta x)^2 + 3(y + \Delta y)^2 - (x + \Delta x)(y + \Delta y) - 20 - (x^2 + 3y^2 - xy - 20)\end{aligned}$$

$$
\begin{aligned}
&= x^2 + 2x\Delta x + \Delta x^2 + 3(y^2 + 2y\Delta y + \Delta y^2) - (xy + x\Delta y + y\Delta x + \Delta x\Delta y) \\
&\quad - 20 - x^2 - 3y^2 + xy + 20 \\
&= x^2 + 2x\Delta x + \Delta x^2 + 3y^2 + 6y\Delta y + 3\Delta y^2 - xy - x\Delta y - y\Delta x - \Delta x\Delta y \\
&\quad - 20 - x^2 - 3y^2 + xy + 20
\end{aligned}
$$

b) La producción actual de la empresa es $(x, y) = (200, 150)$. La función de incrementos es

$$\Delta B(200, 150) = 250\Delta x + 700\Delta y + \Delta x^2 + 3\Delta y^2 - \Delta x\Delta y.$$

c) Si el vector de incrementos es $(\Delta x, \Delta y) = (2, 1)$, el beneficio se incrementa en

$$\Delta B(200, 150)(2, 1) = 250 \cdot 2 + 700 \cdot 1 + 2^2 + 3 \cdot 1^2 - 2 \cdot 1 = 1205 \text{ u.m.}$$

■

Llegar a este resultado nos ha obligado a calcular explícitamente la función de incrementos de B, que es complicada. A través del concepto de diferenciabilidad que vamos a estudiar podremos obtener resultados aproximados mediante el uso de derivadas. Veamos lo que podemos decir únicamente con lo que sabemos del tema anterior. Ante todo,

$$\frac{\partial B}{\partial x} = 2x - y, \qquad \frac{\partial B}{\partial y} = 6y - x.$$

En particular,

$$\left.\frac{\partial B}{\partial x}\right|_{(200,150)} = 250, \qquad \left.\frac{\partial B}{\partial y}\right|_{(200,150)} = 700.$$

Por lo tanto, las funciones de incrementos parciales de B son aproximadamente

$$\Delta_x B(200, 150) \approx 250\,\Delta x, \qquad \Delta_y B(200, 150) \approx 700\,\Delta y.$$

Esto nos permite interpretar la forma de la función de incrementos totales:

$$\Delta B(200, 150) = 250\Delta x + 700\Delta y + \Delta x^2 + 3\Delta y^2 - \Delta x\Delta y.$$

Vemos que consta de tres partes: la aproximación del incremento parcial respecto de x, la del incremento parcial respecto de y más un residuo

$$\Delta x^2 + 3\Delta y^2 - \Delta x\Delta y$$

que es mucho menor que las otras partes. Más concretamente, un incremento $\Delta x = 2$ produce por sí solo un incremento de beneficios de $250 \cdot 2 = 500$ u.m., un incremento de $\Delta y = 1$ produce un incremento de beneficios de 700 u.m., mientras que los dos incrementos a la vez producen un incremento de beneficios de

$$\Delta B \approx 500 + 700 + (2^2 + 3 - 2) = 500 + 700 + 5 = 1205 \text{ u.m.},$$

como ya habíamos calculado, pero ahora vemos que el incremento de B es aproximadamente la suma de los incrementos que produce cada incremento parcial por separado. Como veremos a continuación, esto es consecuencia de que la función B es diferenciable.

7.2 Funciones diferenciables

Supongamos que una función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ tiene derivadas parciales en un punto $\bar{p} \in D$. Si incrementamos la variable x_i en una cantidad Δx_i, sabemos que el incremento que experimenta f es aproximadamente

$$\Delta_{x_i} f(\bar{p}) \approx \left.\frac{\partial f}{\partial x_i}\right|_{\bar{p}} \Delta x_i.$$

Si consideramos incrementos para cada una de las variables, la suma de los incrementos parciales es

$$\left.\frac{\partial f}{\partial x_1}\right|_{\bar{p}} \Delta x_1 + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{p}} \Delta x_n = \left(\left.\frac{\partial f}{\partial x_1}\right|_{\bar{p}}, \ldots, \left.\frac{\partial f}{\partial x_n}\right|_{\bar{p}} \right) (\Delta x_1, \ldots, \Delta x_n) = \nabla f(\bar{p}) \cdot \Delta \bar{x}.$$

Las funciones diferenciables son las funciones para las que esta suma de incrementos parciales es una buena aproximación del incremento total de la función.

Definición Una función escalar $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ definida en un abierto D es *diferenciable* en un punto $\bar{p} \in D$ si tiene derivadas parciales de $\bar{x}$ y existe

$$\lim_{\Delta \bar{x} \to \bar{0}} \frac{\Delta f(\bar{p})(\Delta \bar{x}) - \nabla f(\bar{p}) \cdot \Delta \bar{x}}{\|\Delta \bar{x}\|} = \lim_{\Delta \bar{x} \to \bar{0}} \frac{f(\bar{p} + \Delta \bar{x}) - f(\bar{p}) - \nabla f(\bar{p}) \cdot \Delta \bar{x}}{\|\Delta \bar{x}\|} = 0.$$

Diremos que f *es diferenciable* si lo es en todos los puntos de su dominio.

De acuerdo con la interpretación de los límites, si f es diferenciable en $\bar{p}$ tenemos que, para un vector de incrementos marginales $\Delta \bar{x}$, se cumple

$$\frac{\Delta f(\bar{p})(\Delta \bar{x}) - \nabla f(\bar{p}) \cdot \Delta \bar{x}}{\|\Delta \bar{x}\|} \approx 0,$$

de modo que también $\Delta f(\bar{p})(\Delta \bar{x}) - \nabla f(\bar{x}) \cdot \Delta \bar{x} \approx 0$ y, por consiguiente,

$$\Delta f(\bar{p})(\Delta \bar{x}) \approx \nabla f(\bar{p}) \cdot \Delta \bar{x} = \left.\frac{\partial f}{\partial x_1}\right|_{\bar{p}} \Delta x_1 + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{p}} \Delta x_n.$$

Al miembro derecho se le llama diferencial de f en $\bar{x}$. Más precisamente:

Definición Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función diferenciable en un punto $\bar{x}$, se llama *diferencial* de f en $\bar{x}$ a la aplicación $df(\bar{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

$$df(\bar{x})(\Delta \bar{x}) = \nabla f(\bar{x}) \cdot \Delta \bar{x} = \left.\frac{\partial f}{\partial x_1}\right|_{\bar{x}} \Delta x_1 + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{x}} \Delta x_n.$$

La notación usual para la expresión de la diferencial de una función se obtiene observando que si f es simplemente la función $f(\bar{x}) = x_i$, entonces la expresión anterior se reduce a $dx_i(\Delta \bar{x}) = \Delta x_i$, por lo que, para una función arbitraria f, la expresión anterior equivale a

$$df(\bar{x})(\Delta \bar{x}) = \left.\frac{\partial f}{\partial x_1}\right|_{\bar{x}} dx_1(\Delta \bar{x}) + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{x}} dx_n(\Delta \bar{x}).$$

Si f es diferenciable en su dominio esto es cierto para todo punto $\bar{x}$ y todo vector de incrementos $\Delta\bar{x}$, por lo que podemos expresar esta igualdad como una igualdad de funciones

$$\boxed{df = \frac{\partial f}{\partial x_1}dx_1 + \cdots + \frac{\partial f}{\partial x_n}dx_n,}$$

que es la expresión usual para la diferencial de una función escalar.

La definición de función diferenciable resulta poco operativa para decidir si una función dada lo es. En la práctica nos bastará con esta condición suficiente:

Teorema *Toda función de clase C^1 en un abierto es diferenciable en dicho abierto.*

Ejemplo *Dada la función $f(x,y,z) = x^2y + 2yz$ calcula*

$$df, \quad df(2,3,-1) \quad y \quad df(2,3,-1)(1,0,3).$$

Solución: La función f es diferenciable en $\mathbb{R}^3$, pues es un polinomio y los polinomios son de clase C^1. Su diferencial es

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz = 2xy\,dx + (y+2z)dy + 2y\,dz.$$

La diferencial en el punto $(x,y,z) = (2,3,-1)$ es

$$df(2,3,-1) = 12\,dx + dy + 6\,dz.$$

Esta diferencial actuando sobre el vector de incrementos $(\Delta x, \Delta y, \Delta z) = (1,0,3)$ es

$$df(2,3,-1)(1,0,3) = 12 \cdot 1 + 0 + 6 \cdot 3 = 30.$$

■

Aunque teóricamente no es exacto, en la práctica podemos pensar que df depende de las mismas variables que f (en el ejemplo anterior x, y, z) y de las diferenciales de las variables de f (en el ejemplo dx, dy, dz), es decir:

Si f es una función diferenciable en un abierto, entonces df es una función cuyas variables son las variables x_i de f y las variables dx_i. Cuando particularizamos la diferencial en un punto $\bar{x}$, entonces $df(\bar{x})$ es una función de las variables dx_i. Su interpretación es que para incrementos marginales dx_i se cumple que df es una aproximación de la función de incrementos Δf.

Ejemplo *Calcula la diferencial de la función* $f(x,y) = \sqrt{xy}$ *en el punto* $(3,12)$.

Solución: Simplemente, $\frac{\partial f}{\partial x} = \frac{y}{2\sqrt{xy}}$, $\frac{\partial f}{\partial y} = \frac{x}{2\sqrt{xy}}$, luego

$$df = \frac{y}{2\sqrt{xy}}\,dx + \frac{x}{2\sqrt{xy}}\,dy.$$

Si la queremos en el punto $(3,12)$ será

$$df(3,12) = dx + \frac{1}{4}\,dy.$$

■

Ejemplo *Calcula aproximadamente sin calculadora el valor de* $f(3.1, 11.8)$*, donde* f *es la función del ejemplo anterior.*

Solución: Es fácil calcular $f(3,12) = \sqrt{36} = 6$, con lo que podemos expresar

$$f(3.1, 11.8) = f(3 + 0.1, 12 - 0.2) \approx f(3,12) + df(3,12)(0.1, -0.2)$$

$$= 6 + 0.1 + \frac{1}{4}(-0.2) = 6.1 - 0.05 = 6.05.$$

El valor real (que da la calculadora) es 6.048, luego nos hemos equivocado en 2 milésimas. ■

Interpretación marginal de la diferenciabilidad Ahora podemos precisar la idea con la que hemos introducido la noción de diferenciabilidad. Consideremos una función diferenciable f. Por simplicidad supongamos que tiene dos variables $f(x,y)$. El incremento que experimenta tras un incremento marginal de sus variables $(\Delta x, \Delta y)$ es aproximadamente

$$\Delta f = df(x,y)(\Delta x, \Delta y) = \frac{\partial f}{\partial x}\,\Delta x + \frac{\partial f}{\partial y}\,\Delta y.$$

Ahora bien, el primer sumando es aproximadamente $\Delta_x f$, y el segundo $\Delta_y f$. Esto es cierto en general:

La diferenciabilidad de una función en un punto se traduce en que el incremento que experimenta al incrementar marginalmente sus variables es aproximadamente igual a la suma de los incrementos producidos por el incremento de cada variable por separado.

Por ejemplo, supongamos que una empresa fabrica dos productos A y B y se sabe que el coste marginal respecto a cada uno de ellos para la producción actual $(x,y) = (300, 500)$ es

$$\left.\frac{\partial C}{\partial x}\right|_{(300,500)} = 5 \text{ u.m./unidad de A}, \qquad \left.\frac{\partial C}{\partial y}\right|_{(300,500)} = 3 \text{ u.m./unidad de B}.$$

La empresa va a incrementar su producción en $(\Delta x, \Delta y) = (2, -1)$. Con estos datos podemos afirmar que si únicamente incrementara la producción de A en $\Delta x = 2$ el incremento de coste sería

$$\Delta_x C \approx \left.\frac{\partial C}{\partial x}\right|_{(300,500)} \cdot \Delta x = 5 \cdot 2 = 10 \text{ u.m.},$$

mientras que si sólo incrementara la producción de B en $\Delta y = -1$ el incremento del coste sería

$$\Delta_y C \approx \left.\frac{\partial C}{\partial y}\right|_{(300,500)} \cdot \Delta y = 3 \cdot (-1) = -3 \text{ u.m.}$$

No tenemos datos para estimar el incremento de coste que se produce con el incremento simultáneo $(\Delta x, \Delta y)$ a menos que sepamos que la función de costes $C(x, y)$ es diferenciable, en cuyo caso el incremento total será aproximadamente la suma de los incrementos parciales:

$$\Delta C \approx \Delta_x C + \Delta_y C \approx 10 - 3 = 7 \text{ u.m.}$$

Teorema *Toda función diferenciable es continua.*

Demostración: Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función diferenciable en un punto $\bar{p} \in D$. Entonces existe

$$\lim_{\Delta\bar{x}\to\bar{0}} \frac{f(\bar{p} + \Delta\bar{x}) - f(\bar{p}) - \nabla f(\bar{p}) \cdot \Delta\bar{x}}{\|\Delta\bar{x}\|} = 0.$$

Puesto que $\lim\limits_{\Delta\bar{x}\to\bar{0}} \|\Delta\bar{x}\| = 0$ (la norma es continua) al multiplicar por $\|\Delta\bar{x}\|$ el límite será $0 \cdot 0 = 0$, es decir:

$$\lim_{\Delta\bar{x}\to\bar{0}} f(\bar{p} + \Delta\bar{x}) - f(\bar{p}) - \nabla f(\bar{p}) \cdot \Delta\bar{x} = 0.$$

Ahora bien,

$$\lim_{\Delta\bar{x}\to\bar{0}} \nabla f(\bar{p}) \cdot \Delta\bar{x} = \nabla f(\bar{p}) \cdot \bar{0} = 0,$$

luego

$$\lim_{\Delta\bar{x}\to\bar{0}} f(\bar{p} + \Delta\bar{x}) - f(\bar{p}) = 0$$

o, lo que es lo mismo, existe

$$\lim_{\Delta\bar{x}\to\bar{0}} f(\bar{p} + \Delta\bar{x}) = f(\bar{p}).$$

Si llamamos $\bar{x} = \bar{p} + \Delta\bar{x}$ es claro que cuando $\Delta x \to \bar{0}$ entonces $\bar{x} \to \bar{p}$ y viceversa, luego

$$\lim_{\bar{x}\to\bar{p}} f(\bar{x}) = f(\bar{p}).$$

Esto significa que f es continua en $\bar{p}$. ∎

El teorema anterior nos proporciona una condición necesaria de diferenciabilidad. Es decir, si una función no es continua en $\bar{p}$, podemos asegurar que no es diferenciable en $\bar{p}$.

Diferenciabilidad de funciones vectoriales La definición de función diferenciable se extiende a funciones vectoriales $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ sin más que sustituir el vector gradiente $\nabla f(\bar{p})$ por la matriz jacobiana $Jf(\bar{p})$, es decir, la función f es diferenciable en un punto $\bar{p} \in D$ (D abierto) si

$$\lim_{\Delta\bar{x}\to\bar{0}} \frac{f(\bar{p}+\Delta\bar{x}) - f(\bar{p}) - (Jf(\bar{p})\Delta\bar{x}^t)^t}{\|\Delta\bar{x}\|} = 0.$$

Notemos que para multiplicar la matriz jacobiana por $\Delta\bar{x}$ hemos de trasponer éste último y el resultado es un vector columna, por lo que hemos de volver a trasponerlo para restarlo de $f(\bar{p}+\Delta\bar{x}) - f(\bar{p})$.

La diferencial de f en $\bar{x}$ se define como la aplicación $df(\bar{p}) : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ dada por $df(\bar{p})(\Delta\bar{x}) = (Jf(\bar{p})\Delta\bar{x}^t)^t$. En la práctica basta tener en cuenta el teorema siguiente:

Teorema Una función vectorial $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ es diferenciable en un punto $\bar{p} \in D$ si y sólo si lo son todas las funciones coordenadas f_i, y en tal caso $df(\bar{p}) = (df_1(\bar{p}), \ldots, df_m(\bar{p}))$.

Ejemplo *Comprueba que la función $f(x,y) = (x^2y, x\operatorname{sen} y, e^{xy})$ es diferenciable en $\mathbb{R}^2$ y calcula su diferencial.*

Solución: Es fácil ver que todas las funciones coordenadas son de clase C^1 y por consiguiente son diferenciables. Entonces, f es diferenciable y su diferencial es

$$df = (2xt\,dx + x^2\,dy,\ \operatorname{sen} y\,dx + x\cos y\,dy,\ ye^{xy}\,dx + xe^{xy}\,dy).$$

■

Direcciones de crecimiento máximo, mínimo y nulo

Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función diferenciable en un punto $\bar{x} \in D$, podemos plantearnos la siguiente cuestión: ¿Qué vector de incrementos $\Delta\bar{x}$ producirá el mayor incremento de la función f, es decir, hará que $f(\bar{x}+\Delta\bar{x}) - f(\bar{x})$ sea máximo? En general esta pregunta no está bien planteada, pues puede ocurrir —por ejemplo— que cuanto mayores sean los incrementos de las variables mayor sea el incremento que experimenta la función, por lo que no haya un incremento máximo. Para que la pregunta tenga sentido podemos considerar únicamente incrementos unitarios, es decir, tales que $\|\Delta\bar{x}\| = 1$. La pregunta correcta es, pues, ¿qué vector de incrementos unitario produce el mayor incremento de la función? Si f es diferenciable, sabemos que $\Delta f(\bar{x})(\Delta\bar{x}) \approx df(\bar{x})(\Delta\bar{x})$, por lo que el problema equivale a determinar el incremento unitario que maximiza la diferencial. Con esto determinamos más exactamente la dirección en la que la función crece más rápidamente. Tenemos que

$$df(\bar{x})(\Delta\bar{x}) = \nabla f(\bar{x}) \cdot \Delta\bar{x} = \|\nabla f(\bar{x})\|\|\Delta\bar{x}\|\cos\alpha = \|\nabla f(\bar{x})\|\cos\alpha,$$

donde α es el ángulo entre $\Delta\bar{x}$ y el gradiente de f.

Notemos que aquí hemos de suponer que $\nabla f(\bar{x}) \neq \bar{0}$. Si el gradiente se anula no podemos continuar nuestro análisis. Esto puede deberse a que la función tenga un máximo o un mínimo en $\bar{x}$ (con lo que no hay direcciones de máximo crecimiento o bien hay infinitas) o porque sea necesario considerar derivadas de órdenes superiores.

La diferencial será máxima cuando $\cos\alpha = 1$, lo cual sucede si $\alpha = 0$, es decir, si $\Delta\bar{x}$ tiene la dirección del gradiente y es, por consiguiente,

$$\Delta\bar{x} = \frac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|}.$$

Esta dirección se llama *dirección de máximo crecimiento* de f en $\bar{x}$.

Asimismo, la diferencial es mínima cuando $\cos\alpha = -1$, o sea, $\alpha = \pi$, lo que se traduce en que $\Delta\bar{x}$ tiene la dirección opuesta a la anterior, es decir,

$$\Delta\bar{x} = -\frac{\nabla f(\bar{x})}{\|\nabla f(\bar{x})\|}.$$

Esta dirección se llama *dirección de máximo decrecimiento* de f en $\bar{x}$.

Finalmente, las direcciones para las que $\cos\alpha = 0$, es decir, las direcciones perpendiculares al gradiente, son aquellas en las que la variación de f es menos apreciable, y se llaman *direcciones de crecimiento nulo* de f en $\bar{x}$. Están determinadas por la relación $\Delta\bar{x}\nabla f(\bar{x}) = 0$.

Ejemplo *Calcula direcciones de crecimiento máximo, mínimo y nulo de la función*

$$f(x, y, z) = x^2yz^2$$

en el punto $(1, 1, 1)$.

SOLUCIÓN: El vector gradiente de f es

$$\nabla f(x, y, z) = (2xyz^2, x^2z^2, 2x^2yz) \quad \Rightarrow \quad \nabla f(1, 1, 1) = (2, 1, 2).$$

Entonces, la dirección de máximo crecimiento es

$$u = \frac{\nabla f(1,1,1)}{||\nabla f(1,1,1)||} = (\frac{2}{3}, \frac{1}{3}, \frac{2}{3}),$$

la dirección de mínimo crecimiento es

$$-u = -\frac{\nabla f(1,1,1)}{||\nabla f(1,1,1)||} = (-\frac{2}{3}, -\frac{1}{3}, -\frac{2}{3}).$$

Para obtener las direcciones de crecimiento nulo, calculamos los vectores (a, b, c) perpendiculares a $\nabla f(1, 1, 1)$:

$$(a, b, c)(2, 1, 2) = 2a + b + 2c = 0.$$

Por tanto, cualquier vector no nulo del conjunto

$$\{(a, b, c) \mid 2a + b + 2c = 0, \quad a, b, c \in \mathbb{R}\}$$

es una dirección de crecimiento nulo. ■

7.3 Derivadas direccionales

Definición: Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar definida en un abierto D. Se define la derivada direccional de f en un punto $\bar{p} \in D$ en la dirección $\bar{v} \in \mathbb{R}^n$, $\bar{v} \neq \bar{0}$ como

$$D_{\bar{v}} f(\bar{p}) = \lim_{h \to 0} \frac{f(\bar{p} + h\bar{v}) - f(\bar{p})}{h}.$$

Teorema Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función diferenciable en el abierto D. Entonces, f tiene derivadas direccionales en cualquier dirección y, además, dado $\bar{p} \in D$

$$D_{\bar{v}} f(\bar{p}) = df(\bar{p})(\bar{v}).$$

Ejemplo *Calcula la derivada direccional de la función* $f(x, y) = \sqrt{xy}$ *en el punto* $(1, 1)$ *y en la dirección* $(2, 1)$*.*

SOLUCIÓN: Es fácil comprobar que f es diferenciable en $(1, 1)$ y, por tanto, la derivada direccional puede calcularse mediante la diferencial, es decir, tenemos que

$$df = \frac{y}{2\sqrt{xy}}\, dx + \frac{x}{2\sqrt{xy}}\, dy$$

y, por consiguiente,

$$D_{(2,1)} f(1, 1) = df(1, 1)(2, 1) = \frac{1}{2} \cdot 2 + \frac{1}{2} \cdot 1 = 1.5.$$

■

Teorema *Sea* $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *una función escalar definida en el abierto* D*. Si existen las derivadas parciales en* $\bar{p} \in D$*, entonces*

$$\frac{\partial f}{\partial x_i}(\bar{p}) = D_{\bar{e}_i} f(\bar{p}),$$

donde $\bar{e}_i$ *es el vector i-ésimo de la base canónica de* $\mathbb{R}^n$*.*

DEMOSTRACIÓN: La expresión de la derivada direccional en la direcció $\bar{e}_i$ es

$$D_{\bar{e}_i} f(\bar{p}) = \lim_{h \to 0} \frac{f(\bar{p} + h\bar{e}_i) - f(\bar{p})}{h} = \lim_{h \to 0} \frac{f(p_1, \ldots, p_i + h, \ldots, p_n) - f(\bar{p})}{h}$$
$$= \frac{\partial f}{\partial x_i}(\bar{p}).$$

■

De los dos teoremas anteriores se deduce que toda función diferenciable en un punto es derivable en dicho punto.

Relación entre continuidad, derivabilidad y diferenciabilidad Se cumplen los siguientes hechos:

1. Toda función con derivadas parciales continuas es diferenciable.
2. Toda función diferenciable es continua.
3. Toda función diferenciable es derivable.
4. Toda función diferenciable tiene derivadas direccionales.

Las condiciones 2, 3 y 4 anteriores son condiciones necesarias de diferenciabilidad. Por tanto, basta con que alguna no se verifique para que la función no sea diferenciable. Sin embargo, el hecho de que se verifiquen no garantiza la diferenciabilidad de la función.

Ejemplo *Estudia la diferenciabilidad de la función*

$$f(x,y) = \begin{cases} x+2y, & x \geq y \\ x-2y, & x < y \end{cases}$$

en el punto $(1,1)$.

Solución: Es fácil comprobar que la función no es continua en $(1,1)$ porque no existe el límite. Por tanto no es diferenciable. ■

Cualquier otra implicación entre estos conceptos es falsa en general: hay funciones continuas que no son diferenciables, funciones derivables que no son diferenciables, etc.

7.4 El polinomio de Taylor

Hemos visto que si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función diferenciable en un punto $\bar{p}$ del abierto D entonces, para incrementos marginales $\Delta\bar{x}$ se cumple

$$\Delta f = f(\bar{p}+\Delta\bar{x}) - f(\bar{p}) \approx \left.\frac{\partial f}{\partial x_1}\right|_{\bar{p}} \Delta x_1 + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{p}} \Delta x_n.$$

Equivalentemente, para todo punto $\bar{x} = \bar{p} + \Delta\bar{x}$ cercano a $\bar{p}$ se cumple

$$f(\bar{x}) \approx f(\bar{p}) + \left.\frac{\partial f}{\partial x_1}\right|_{\bar{p}} (x_1 - p_1) + \cdots + \left.\frac{\partial f}{\partial x_n}\right|_{\bar{p}} (x_n - p_n).$$

El miembro derecho es un polinomio de grado 1. Si estamos dispuestos a considerar polinomios de grado mayor (y derivadas de orden superior) podemos obtener aproximaciones polinómicas mucho mejores de una función. Concretamente, la mejor aproximación polinómica de un grado dado n a una función f de clase C^n se conoce como polinomio de Taylor de grado n de f alrededor de un punto dado $\bar{p}$. Este polinomio está determinado por el teorema siguiente:

Teorema *Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función de clase C^n en un punto $\bar{p}$ del abierto D, existe un único polinomio $P_n f(\bar{p})(\bar{x})$ de grado $\leq n$ cuyas derivadas parciales de orden $\leq n$ en $\bar{p}$ coinciden con las de f. Dicho polinomio se llama* polinomio de Taylor *de f en $\bar{p}$ y para valores de $\bar{x}$ cercanos a $\bar{p}$ se tiene la aproximación*

$$f(\bar{x}) \approx P_n f(\bar{p})(\bar{x}).$$

Esta aproximación es mejor cuanto mayor es n.

Existe una fórmula para calcular el polinomio de Taylor de grado n de una función de clase C^n, pero para funciones de varias variables el cálculo es laborioso, así que nos limitaremos a considerar polinomios de grado 1 y 2, que son los que después emplearemos.

La fórmula para el polinomio de Taylor de grado 1 es la que hemos visto antes, que no es sino la aproximación por la diferencial:

$$P_1 f(\bar{p})(\bar{x}) = f(\bar{p}) + \nabla f(\bar{p})(\bar{x} - \bar{p}).$$

Para el polinomio de grado 2 hay que añadir otro término:

$$P_2 f(\bar{p})(\bar{x}) = f(\bar{p}) + \nabla f(\bar{p})(\bar{x} - \bar{p}) + \frac{1}{2}(\bar{x} - \bar{p}) H f(\bar{p})(x - \bar{p})^t.$$

Ejemplo *Calcula el polinomio de Taylor de grado* 2 *de la función* $\operatorname{sen}(xy)$ *en el punto* $\bar{p} = (1, 1)$.

SOLUCIÓN:
El término de grado 0 es $f(1, 1) = \operatorname{sen} 1$.
El término de grado 1 se construye con el gradiente:

$$\nabla f(x, y) = (y \cos(xy), x \cos(xy)), \qquad \nabla f(1, 1) = (\cos 1, \cos 1),$$

con lo que el término de grado 1 es

$$(\cos 1, \cos 1)(x - 1, y - 1) = (\cos 1)(x - 1) + (\cos 1)(y - 1).$$

El término de grado 2 se construye con la hessiana:

$$Hf(x, y) = \begin{pmatrix} -y^2 \operatorname{sen}(xy) & \cos(xy) - xy \operatorname{sen}(xy) \\ \cos(xy) - xy \operatorname{sen}(xy) & -x^2 \operatorname{sen}(xy) \end{pmatrix},$$

$$Hf(1, 1) = \begin{pmatrix} -\operatorname{sen} 1 & \cos 1 - \operatorname{sen} 1 \\ \cos 1 - \operatorname{sen} 1 & -\operatorname{sen} 1 \end{pmatrix},$$

con lo que el término de grado 2 es

$$\frac{1}{2}(x - 1, y - 1) \begin{pmatrix} -\operatorname{sen} 1 & \cos 1 - \operatorname{sen} 1 \\ \cos 1 - \operatorname{sen} 1 & -\operatorname{sen} 1 \end{pmatrix} \begin{pmatrix} x - 1 \\ y - 1 \end{pmatrix}$$

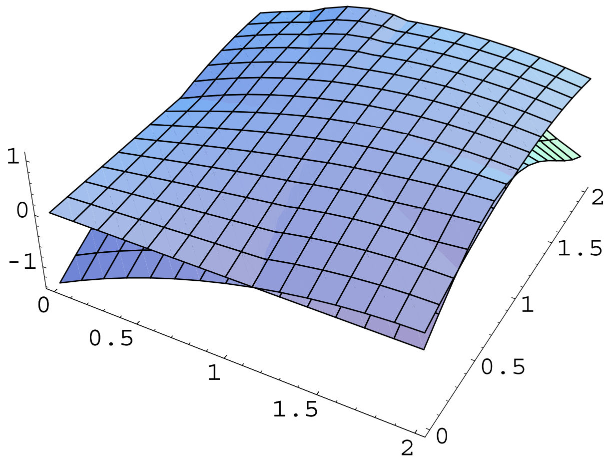
$$= -\frac{1}{2}(\operatorname{sen} 1)(x - 1)^2 + (\cos 1 - \operatorname{sen} 1)(x - 1)(y - 1) - \frac{1}{2}(\operatorname{sen} 1)(y - 1)^2.$$

El polinomio de grado 2 es, por tanto:

$$\begin{aligned} P_2 f(1,1)(x,y) &= 0.84 + 0.54(x-1) + 0.54(y-1) \\ &\quad -0.42(x-1)^2 - 0.30(x-1)(y-1) - 0.42(y-1)^2. \end{aligned}$$

■

La figura muestra la función sen xy y el polinomio que acabamos de calcular. La función sen xy es la que está debajo en el punto $(2,2)$.



Vemos que alrededor de $(1,1)$ son muy parecidas, aunque luego se separan. Esto hace que para estudiar diversas propiedades de una función f de clase C^2 (por ejemplo, si tiene un máximo o un mínimo en un punto) en muchos casos podemos sustituir f por el polinomio de Taylor, sin que ello modifique las conclusiones.

7.5 Ejercicios

1. Calcula la diferencial de las funciones siguientes:

   (a) $P(x,y) = x^2 - 3xy + y^4$,

   (b) $Q(x,y) = 3x - 5y + 7$,

   (c) $z(x,y) = \dfrac{1}{xy}$,

   (d) $C(u,y,t) = e^{2u+3y-t}$,

   (e) $U(x,y) = \operatorname{sen}^4(x^3y^2)$,

   (f) $t(x,p) = \sqrt[3]{x^2+p^3-2}$.

2. Dada la función $V(p,q) = p^3 + pq + 2$, calcula $dV(3,1)$ y $dV(2,1)(2,3)$.

3. Dada la función $T(x,y,z) = e^{2x+y}z$ calcula $dT(1,-2,5)$.

4. Calcula la dirección de máximo crecimiento, máximo decrecimiento y las direcciones de crecimiento nulo de las funciones siguientes en los puntos indicados:
   (a) $f(x, y) = x^2 + 2xy$ en $(3, 4)$,
   (b) $g(u, v, w) = we^{u+v}$ en $(5, 1, -1)$,
   (c) $h(p, q) = 3p + 4q$ en $(1, 1)$,
   (d) $j(a, b) = ab$ en $(3, 1)$.
5. Calcula el polinomio de Taylor de grado 1 de las funciones siguientes en los puntos indicados:
   (a) $f(x, y) = \cos(x + y)$ en $(a, b) = (\pi/2, \pi/2)$.
   (b) $f(x, y, z) = x^2y + 2xz^3$ en $(a, b, c) = (1, -2, 1)$.
   (c) $f(u, v) = 3u - v$ en $(a, b) = (3, 2)$.
6. Calcula el polinomio de Taylor de grado 2 de las funciones siguientes en los puntos indicados:
   (a) $T(p, q, r) = pq^3 - 3qr + 2p$ en $(a, b, c) = (1, -1, 1)$.
   (b) $S(x, z) = \sqrt{x + 2z}$ en $(a, b) = (5, 2)$.
   (c) $f(x, y) = x \operatorname{sen} y$ en $(a, b) = (1, \pi)$.
7. Calcula aproximadamente (sin usar calculadora) $1.2e^{0.05}$ usando el polinomio de Taylor de grado 2 de la función $f(x, y) = xe^y$ en el punto adecuado. Compara el resultado con el que da la calculadora. ¿Cuál es el porcentaje de error?
8. Calcula los polinomios de Taylor de grado 1 y de grado 2 de la función $f(u, v) = x \ln y$ en el punto $(2, 0)$. Calcula con ellos las aproximaciones correspondientes de $1.9 \ln 0.3$ y el porcentaje de error de cada una de ellas.
9. Una empresa fabrica dos productos A y B en cantidades x e y. Los beneficios que obtiene con su producción vienen dados por una cierta función $B(x, y)$. Actualmente los beneficios ascienden a 200 u.m., pero la empresa tiene más demanda de la que realmente está cubriendo, por lo que se plantea aumentar su producción. Sus recursos le permiten un aumento de 10 unidades de producto. La empresa estima que, para la producción actual $\bar{p} = (x, y)$ se cumple
$$\left.\frac{\partial B}{\partial x}\right|_{\bar{p}} = 2\text{u.m./unidad de A}, \qquad \left.\frac{\partial B}{\partial y}\right|_{\bar{p}} = 3\text{u.m./unidad de B}.$$
   (a) ¿Cuál es exactamente la interpretación de estas derivadas en este contexto concreto?
   (b) ¿Qué beneficios pasaría a obtener la empresa si aumentara en 5 unidades la producción de A? (La pregunta es *qué beneficios*, no *qué incremento de beneficios*.) ¿Y si aumenta en 5 unidades la producción de B?

(c) Para estimar con estos datos los beneficios de la empresa en el supuesto de que aumente simultáneamente 5 unidades la producción de A y 5 la de B necesitamos una hipótesis sobre la función B, ¿cuál?

(d) Con dicha hipótesis, ¿cuáles pasarían a ser los beneficios de la empresa?

10. Si $f : D \subset \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ es diferenciable en el punto $(3, 1, 2) \in D$. Di qué clase de objeto es $df(3, 1, 2)$ (un número, un vector, una matriz, una aplicación. . .) Precisa todo lo posible la respuesta.

11. Una función $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ es diferenciable en el punto $(2, 3)$. Indica qué relación hay entre $df(2, 3)$ y $\nabla f(2, 3)$.

12. Una empresa estima que sus beneficios $B(p, x)$ dependen del precio medio de sus materias primas p y de la cantidad de producto que fabrica x. Actualmente sus beneficios son de 100 u.m. y corresponden a una producción de $x = 5$ unidades y a unos precios de $p = 1$ u.m. Así mismo considera que

$$\left.\frac{\partial B}{\partial p}\right|_{(1,5)} = -3 \text{ u.m./u.m}, \qquad \left.\frac{\partial B}{\partial x}\right|_{(1,5)} = 2 \text{ u.m./u. prod.}$$

a) Interpreta las derivadas, especialmente su signo.

b) ¿Qué beneficios cabría esperar si los precios aumentan a 1.3 u.m.?

c) ¿Y si, además de dicho aumento de precios, la empresa aumenta su producción en 3 unidades?

d) ¿Hace falta alguna hipótesis matemática sobre la función B para justificar la respuesta a c)?

13. Sea $D(p, r, t)$ la función de demanda de un artículo en un mercado, donde p es el precio (en u.m.), r la renta media de los consumidores (en u.m.) y t el tiempo en años. Actualmente ($t = 0$) se tiene $(p, r) = (5, 15)$ y $D(5, 15, 0) = 200$. Además

$$\left.\frac{\partial D}{\partial t}\right|_{(5,15,0)} = 20, \qquad \left.\frac{\partial D}{\partial p}\right|_{(5,15,0)} = -15, \qquad \left.\frac{\partial D}{\partial r}\right|_{(5,15,0)} = 10.$$

(a) Interpreta estas derivadas e indica las unidades en que vienen expresadas.

(b) ¿Qué demanda cabría esperar dentro de un año si la renta ha pasado de $r = 16$ u.m. y el precio a $p = 4.5$ u.m.?, ¿qué hipótesis sobre D es necesaria para responder a esta pregunta con los datos disponibles?

14. Una empresa ha hecho una estimación de la repercusión en sus beneficios de su inversión en producción x y de su inversión en publicidad P a través de una función $B(x, P)$ (que suponemos diferenciable). Concretamente estima que, para los valores actuales

$$\left.\frac{\partial B}{\partial x}\right|_{(1000,100)} = 27, \qquad \left.\frac{\partial B}{\partial P}\right|_{(1000,100)} = 36.$$

(a) Interpreta estas derivadas. Indica las unidades en que se expresan.

(b) Calcula $dB(1000, 100)$.

(c) Calcula la dirección de máximo crecimiento de B en $(1000, 100)$.

(d) Si la empresa dispone de 5 unidades monetarias marginales para invertir, ¿que cantidad le convendrá invertir en producción y qué cantidad en publicidad para conseguir el mayor incremento de beneficios?

15. Explica por qué las respuestas siguientes son incorrectas.

(a) Sea $f(x, y) =$ una función cualquiera. Calcula $df(1, 2)$.
RESPUESTAS:

a) $df(1, 2) = 5$, *b*) $df(1, 2) = 3x^2 + y$, *c*) $df(1, 2) = 3x\,dx - dy$.

(b) Sea $f(x, y, z) =$ una función normal y corriente. Calcula $Hf(1, 0, -1)$.
RESPUESTAS:

a) $Hf(1, 0, -1) = \begin{pmatrix} 2yz & 2xz & 2xy \\ 2xz & 0 & x^2 \\ 2xy & x^2 & 0 \end{pmatrix}$,

b) $Hf(1, 0, -1) = \begin{pmatrix} 3 & 2 & 4 \\ 1 & 0 & 3 \\ 1 & -1 & 1 \end{pmatrix}$.

(c) Sea $f : \mathbb{R} \longrightarrow \mathbb{R}^2$ la función dada por $f(t) =$ (una función, otra función). Calcula la matriz jacobiana de f en $t = 3$.
RESPUESTAS:

a) $Jf(3) = 0$, *b*) $Jf(3) = (2t, 3t^2)$, *c*) $Jf(3) = (2, 3)$.

(d) Calcula la dirección de máximo crecimiento de la función $f(x, y) =$ "lo que sea" en el punto $(3, 4)$.
RESPUESTAS:

a) $DMC = (2x, y)$, *b*) $DMC = (2, 2)$.

8. Funciones compuestas y homogéneas

8.1 Composición de funciones

Definición Si $f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ y $g : B \subset \mathbb{R}^m \longrightarrow \mathbb{R}^k$ son funciones tales que $f[A] \subset B$, definimos la *función compuesta* $g \circ f : A \subset \mathbb{R}^n \longrightarrow \mathbb{R}^k$ como la función dada por $(g \circ f)(\bar{x}) = g(f(\bar{x}))$.

En otras palabras, $g \circ f$ es la función que a cada punto $\bar{x}$ le asigna el resultado de aplicar f a $\bar{x}$ y después aplicar g al resultado $f(\bar{x})$.

Ejemplo *La demanda de una empresa está dada por la función*

$$D(p_1, p_2) = 50/(p_1 p_2)$$

donde p_1 y p_2 son los precios a los que vende sus dos artículos. La empresa fija p_1 y p_2 en función de los precios q_1 y q_2 de las materias primas que emplea en su fabricación según las relaciones

$$p_1 = 3q_1 + q_2, \quad p_2 = q_1 + 2q_2.$$

Calcula la demanda en función de los precios de las materias primas.

SOLUCIÓN:

La composición de estas funciones es la función $D(q_1, q_2)$ que nos da la demanda de la empresa en términos de los precios q_1 y q_2 de las materias primas. En este caso se tiene

$$D(p_1, p_2) = \frac{50}{p_1 p_2} = \frac{50}{(3q_1 + q_2)(q_1 + 2q_2)}.$$

■

En la práctica, calcular una composición de funciones consiste en sustituir unas funciones en otras. Es muy importante no confundir la función $D(p_1, p_2)$ con la función compuesta $D(q_1, q_2)$. Es frecuente que se use el mismo nombre para ambas (en este caso D), y entonces se distinguen por las variables.

Para ver la relación entre el ejemplo anterior y la definición previa hemos de considerar las funciones q_1 y q_2 como las componentes de una función vectorial $f : A \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, de modo que

$$f(q_1, q_2) = (3q_1 + q_2, q_1 + 2q_2).$$

El dominio es $A = \{(q_1, q_2) \in \mathbb{R}^2 \mid q_1 > 0,\ q_2 > 0\}$, pues no tiene sentido considerar precios menores o iguales que 0. A su vez, la función de demanda $D(p_1, p_2)$ es una función $D : B \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$, donde $B = \{(p_1, p_2) \in \mathbb{R}^n \mid p_1 > 0,\ p_2 > 0\}$. La función $D(q_1, q_2)$ es, en los términos de la definición de función compuesta, la función $D \circ f$. En efecto, para calcular $D(q_1, q_2)$ aplicamos primero f a (q_1, q_2) para obtener (p_1, p_2) y luego $D(p_1, p_2)$ para obtener la demanda correspondiente.

Regla de la cadena *Sean $f : A \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ y $g : B \subset \mathbb{R}^n \longrightarrow \mathbb{R}^k$ dos funciones definidas en abiertos A y B tales que $f[A] \subset B$. Supongamos que f es diferenciable en un punto $\bar{p} \in A$ y que g es diferenciable en $f(\bar{p}) \in B$. Entonces $g \circ f$ es diferenciable en $\bar{p}$ y*

$$d(g \circ f)(\bar{p}) = dg(f(\bar{p})) \circ df(\bar{p}).$$

Suele ser más cómodo trabajar con las matrices jacobianas en lugar de con las diferenciales. Para ello usamos la siguiente consecuencia de la regla de la cadena:

Teorema *Sean $f : A \subset \mathbb{R}^m \longrightarrow \mathbb{R}^n$ y $g : B \subset \mathbb{R}^n \longrightarrow \mathbb{R}^k$ dos funciones definidas en abiertos A y B tales que $f[A] \subset B$. Supongamos que f es diferenciable en un punto $\bar{p} \in A$ y que g es diferenciable en $f(\bar{p}) \in B$. Entonces*

$$J(g \circ f)(\bar{p}) = Jg(f(\bar{p})) \cdot Jf(\bar{p}).$$

Demostración: Por la regla de la cadena sabemos que

$$d(g \circ f)(\bar{p}) = dg(f(\bar{p})) \circ df(\bar{p}).$$

Por lo tanto, si $\bar{v} \in \mathbb{R}^m$ tendremos que

$$d(g \circ f)(\bar{p})(\bar{v}) = dg(f(\bar{p}))(df(\bar{p})(\bar{v})).$$

Como la diferencial se calcula multiplicando por la matriz jacobiana, el miembro izquierdo es

$$J(g \circ f)(\bar{p}) \cdot \bar{v}^t,$$

mientras que el miembro derecho se obtiene aplicando primero $df(\bar{p})$ y luego $dg(f(\bar{p}))$, es decir:

$$\bar{v} \stackrel{df(\bar{p})}{\mapsto} Jf(\bar{p}) \cdot \bar{v}^t \stackrel{dg(f(\bar{p}))}{\mapsto} Jg(f(\bar{p})) \cdot Jf(\bar{p}) \cdot \bar{v}^t.$$

Así pues, tenemos que

$$J(g \circ f)(\bar{p}) \cdot \bar{v}^t = Jg(f(\bar{p})) \cdot Jf(\bar{p}) \cdot \bar{v}^t.$$

Como esto es cierto para todo vector $\bar{v}$, se ha de dar la igualdad:

$$J(g \circ f)(\bar{p}) = Jg(f(\bar{p})) \cdot Jf(\bar{p}).$$

■

Ejemplo *Aplica el teorema a la función de demanda del ejemplo anterior suponiendo un consumo de materias primas* $(q_1, q_2) = (1, 2)$. *Escribe la diferencial en este punto.*

Solución: Hemos visto anteriormente que podemos expresar la demanda en función de las cantidades de materias primas como la $D(q_1, q_2) = (D \circ f)(q_1, q_2)$ donde $D : B \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ es

$$D(p_1, p_2) = 50/(p_1 p_2)$$

y $f : A \subset \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ es

$$f(q_1, q_2) = (3q_1 + q_2, q_1 + 2q_2).$$

Calculamos las matrices jacobianas de D y de f:

$$JD(p_1, p_2) = \nabla D(p_1, p_2) = \Big(-\frac{50}{p_1^2 p_2}, -\frac{50}{p_1 p_2^2} \Big), \quad Jf(q_1, q_2) = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$$

Sabemos que

$$JD(q_1, q_2) = J(D \circ f)(q_1, q_2) = JD(f(q_1, q_2)) \cdot Jf(q_1, q_2).$$

Calculemos las matrices en los puntos correspondientes:

$$JD(1, 2) = J(D \circ f)(1, 2) = JD(f(1, 2)) \cdot Jf(1, 2) = JD(5, 5) \cdot Jf(1, 2) =$$

$$= \Big(-\frac{2}{5} \quad -\frac{2}{5} \Big) \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix} = \Big(-\frac{8}{5} \quad -\frac{6}{5} \Big).$$

Una vez conocida $JD(1, 2)$, para obtener la diferencial en $(1, 2)$ multiplicamos por el vector de incrementos (dq_1, dq_2):

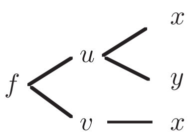
$$dD(1, 2) = \Big(-\frac{8}{5} \quad -\frac{6}{5} \Big) \begin{pmatrix} dq_1 \\ dq_2 \end{pmatrix} = -\frac{8}{5} dq_1 - \frac{6}{5} dq_2$$

■

Cuando interesa calcular una derivada parcial en concreto de una función compuesta, puede ser más rápido utilizar una fórmula explícita que se deduce del teorema anterior al hacer explícito el producto de matrices.

Ejemplo *Sean* $f(u,v) = u^2 - v^2 + 2uv$, $u = x \operatorname{sen} y$, $v = x^2$. *Calcula* $\dfrac{\partial f}{\partial x}$.

Solución: En primer lugar conviene hacer un esquema de la dependencia de las variables que aparecen:



Para calcular la derivada de f respecto de x hemos de considerar todos los caminos que llevan desde f hasta la variable x. En este caso tenemos dos: $f-u-x$ y $f-v-x$. La fórmula tiene un sumando para cada camino:

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = (2u+2v)\operatorname{sen} y + (-2v+2u)2x \\
&= (2x \operatorname{sen} y + 2x^2)\operatorname{sen} y + (-2x^2 + 2x \operatorname{sen} y)2x \\
&= 2x \operatorname{sen}^2 y + 6x^2 \operatorname{sen} y - 4x^2.
\end{aligned}$$

■

La regla de la cadena permite calcular derivadas de funciones compuestas siempre que se conozcan, o se puedan calcular, todas las derivadas de las funciones componentes, sin que sea necesario tener la expresión explícita de todas las funciones.

Ejemplo *Una empresa estima que sus beneficios vienen dados por la función*

$$B(t,p) = \frac{4+0.2t}{\sqrt{p^2-5}},$$

donde el numerador es una estimación de la demanda futura en función del tiempo t *y el denominador es una corrección en función del IPC* p. *El tiempo actual es* $t=1$ *y el IPC es* $p=3$ *u.m. No hay ninguna previsión fiable de la evolución del IPC, pero la empresa estima que en la actualidad*

$$\left.\frac{dp}{dt}\right|_1 = 0.2.$$

Según estas estimaciones, ¿los beneficios de la empresa van a aumentar o a disminuir a corto plazo?

Solución: Puesto que la función de beneficios depende del tiempo y del IPC el cual, a su vez, depende del tiempo, podemos asegurar que existe una función (compuesta) en la que los beneficios dependen del tiempo como única variable. En consecuencia, podemos establecer una función compuesta del modo siguiente:

$$\begin{array}{ccccccc}
B\circ f: & \mathbb{R} & \xrightarrow{f} & \mathbb{R}^2 & \xrightarrow{B} & \mathbb{R} \\
 & t & \mapsto & (t,p) & \mapsto & B(t,p)
\end{array}$$

donde $(B\circ f)(t) = B(t)$ es la función que buscamos.

Conocemos explícitamente la expresión de $B(t,p)$ y sabemos que el momento actual se corresponde con $t = 1$ y $p = 3$. Sin embargo, en la expresión de f sólo conocemos la primera función coordenada. De la segunda función coordenada no conocemos su expresión pero sabemos que

$$\left.\frac{dp}{dt}\right|_1 = 0.2 \qquad \text{y} \qquad p(1) = 3.$$

Para responder a la pregunta del problema necesitamos conocer $\left.\frac{dB}{dt}\right|_1$, es decir, la derivada de la función compuesta. Claramente:

$$\frac{dB}{dt}(1) = \frac{\partial B}{\partial t}(1,3)\frac{dt}{dt}(1) + \frac{\partial B}{\partial p}(1,3)\frac{dp}{dt}(1)$$

Sustituyendo los valores de las derivadas

$$\frac{dB}{dt}(1) = 0.1 \cdot 1 - 1.575 \cdot 0.2 = -0.215.$$

Por tanto, a corto plazo los beneficios de la empresa disminuirán. ■

Si hubiésemos calculado $\frac{\partial B}{\partial t}(t,p)$ no estaríamos respondiendo a la pregunta porque esta derivada nos indica la variación de los beneficios con el tiempo suponiendo constante el IPC. Puesto que en el problema nos aseguran que el IPC depende del tiempo, ésta última hipótesis es falsa.

8.2 Funciones homogéneas

Definición Una función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ definida en un abierto D es *homogénea* de grado $m \in \mathbb{R}$ si para todo $\bar{x} \in D$ y todo $\lambda > 0$ tal que $\lambda\bar{x} \in D$ se cumple que $f(\lambda\bar{x}) = \lambda^m f(\bar{x})$.

Ejemplo *Comprueba si las funciones siguientes son homogéneas*

1. $f(x,y) = x/y^2$

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

SOLUCIÓN:

1. Consideramos un $\lambda > 0$ se cumple

$$f(\lambda x, \lambda y) = \frac{\lambda x}{\lambda^2 y^2} = \frac{x}{\lambda y^2} = \lambda^{-1}\frac{x}{y^2} = \lambda^{-1} f(x,y).$$

Entonces f es homogénea de grado $m = -1$.

2. Consideramos un $\lambda > 0$,

$$f(\lambda x, \lambda y) = (\lambda x)^2 + \lambda y \neq \lambda^m f(x,y) \quad \text{para ningún} \quad m \in \mathbb{R},$$

por tanto no es homogénea. ■

Teorema *Sean f, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ funciones definidas en un abierto D. Entonces*

1. *Si f y g son homogéneas de grado m, entonces $f+g$ también es homogénea de grado m.*
2. *Si f es homogénea de grado m y $\alpha \in \mathbb{R}$, entonces αf también es homogénea de grado m.*
3. *Si f es homogénea de grado m y g es homogénea de grado r, entonces fg es homogénea de grado $m+r$.*
4. *Si f es homogénea de grado m, g es homogénea de grado r y g no se anula en D, entonces f/g es homogénea de grado $m-r$.*
5. *Si f es de clase C^1 en D y es homogénea de grado m, entonces sus derivadas parciales son homogéneas de grado $m-1$.*

DEMOSTRACIÓN:

1. Si f y g son homogéneas de grado m, $\bar{x} \in D$ y $\lambda > 0$ cumple que $\lambda\bar{x} \in D$, entonces

$$(f+g)(\lambda\bar{x}) = f(\lambda\bar{x})+g(\lambda\bar{x}) = \lambda^m f(\bar{x})+\lambda^m g(\bar{x}) = \lambda^m(f(\bar{x})+g(\bar{x})) = \lambda^m(f+g)(\bar{x}),$$

luego $f+g$ es homogénea de grado m.

2. Si f es homogénea de grado m, $\bar{x} \in D$ y $\lambda > 0$ cumple que $\lambda\bar{x} \in D$, entonces

$$\alpha f(\lambda\bar{x}) = \alpha\lambda^m f(\bar{x}) = \lambda^m(\alpha f(\bar{x})),$$

luego αf es homogénea de grado m.

3. Si f es homogénea de grado m y g es homogénea de grado r, tomamos $\bar{x} \in D$ y $\lambda > 0$ tal que $\lambda\bar{x} \in D$. Entonces

$$(f \cdot g)(\lambda\bar{x}) = f(\lambda\bar{x})g(\lambda\bar{x}) = \lambda^m f(\bar{x})\lambda^r g(\bar{x}) = \lambda^{m+r} f(\bar{x})g(\bar{x}) = \lambda^{m+r}(f \cdot g)(\bar{x}),$$

luego fg es homogénea de grado $m+r$.

4. Si f es homogénea de grado m, g es homogénea de grado r y no se anula en D, tomamos $\bar{x} \in D$ y $\lambda > 0$ tal que $\lambda\bar{x} \in D$. Entonces

$$\frac{f}{g}(\lambda\bar{x}) = \frac{f(\lambda\bar{x})}{g(\lambda\bar{x})} = \frac{\lambda^m f(\bar{x})}{\lambda^r g(\bar{x})} = \lambda^{m-r}\frac{f(\bar{x})}{g(\bar{x})} = \lambda^{m-r}\frac{f}{g}(\bar{x}),$$

luego f/g es homogénea de grado $m-r$.

5. Si f es homogénea de grado m y es de clase C^1, tomamos $\bar{x} \in D$ y $\lambda > 0$ tal que $\lambda\bar{x} \in D$. Consideramos la función $g(\bar{x}) = f(\lambda\bar{x}) = \lambda^m f(\bar{x})$. Derivando ambas expresiones para g vemos que

$$\frac{\partial g}{\partial x_i} = \lambda\frac{\partial f}{\partial x_i}(\lambda\bar{x}) = \lambda^m\frac{\partial f}{\partial x_i}(\bar{x}),$$

de donde, despejando,

$$\frac{\partial f}{\partial x_i}(\lambda\bar{x}) = \lambda^{m-1}\frac{\partial f}{\partial x_i}(\bar{x}),$$

luego la derivada es homogénea de grado $m-1$. ∎

Teorema de Euler *Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función de clase C^1 en un abierto D, entonces f es homogénea de grado m si y sólo si*

$$x_1\frac{\partial f}{\partial x_1} + \cdots + x_n\frac{\partial f}{\partial x_n} = mf(\bar{x})$$

Ejemplo Dada la función $f(x,y) = \text{sen}(\frac{x^2-y^2}{x^2+y^2})$ calcula

$$A = x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y}$$

SOLUCIÓN: Es fácil comprobar que f es una función de clase C^1. Además, dado $\lambda > 0$ se tiene

$$f(\lambda x, \lambda y) = \text{sen}\left(\frac{\lambda^2(x^2-y^2)}{\lambda^2(x^2+y^2)}\right) = \text{sen}(\frac{x^2-y^2}{x^2+y^2}) = \lambda^0 f(x,y).$$

Entonces f es homogénea de grado 0 y por el teorema de Euler A=0. ∎

Aplicación 1: La ilusión monetaria Supongamos que ante un aumento en su salario, un trabajador piensa que ha aumentado su renta y esto provoca un incremento en su consumo. Sin embargo, si los precios han aumentado en la misma proporción (o mayor) que el salario, la renta real del trabajador no ha aumentado y en este caso se dice que está afectado por la *ilusión monetaria.*

Más formalmente, supongamos un mercado con n bienes de precios $p_1, p_2, \ldots p_n$. Para un consumidor determinado con renta R, la demanda del bien i viene dada por la función

$$x_i = x_i(p_1, p_2, \ldots, p_n, R).$$

Si los precios y la renta varían en una proporción λ, los consumidores no sufrirán ilusión monetaria si

$$x_i(\lambda p_1, \lambda p_2, \ldots, \lambda p_n, \lambda R) = x_i(p_1, p_2, \ldots, p_n, R) = \lambda^0 x_i(p_1, p_2, \ldots, p_n, R)$$

Eso significa que si las funciones de demanda son homogéneas de grado 0 no se da la ilusión monetaria.

Aplicación 2: Rendimientos a escala Consideramos una función de producción

$$Q(x_1, x_2, \ldots, x_n),$$

donde x_i es el i-ésimo factor productivo.Supongamos que Q es una función homogénea de grado m, es decir se cumple que

$$Q(\lambda x_1, \lambda x_2, \ldots, \lambda x_n) = \lambda^m Q(x_1, x_2, \ldots, x_n),$$

donde λ representa la proporción en la que varían los factores productivos y se denomina *factor de escala.* Es claro que

1. Si $m = 1$ la producción y los factores productivos aumentan en la misma proporción. En este caso decimos que hay rendimientos a escala constantes.

2. Si $m > 1$ la producción aumenta más que los factores productivos. En este caso decimos que hay rendimientos a escala crecientes.

3. Si $m < 1$ la producción aumenta menos que los factores productivos. En este caso decimos que hay rendimientos a escala decrecientes.

8.3 Ejercicios

1. Dadas $f(u, v) = u^2 - v^2$, $u = t^3 + 1$ y $v = (t - 2)^2$. Calcula la función compuesta $f(t)$.

2. Dadas $p(u, v, t) = u^2 + v + 2t$, $u = t^3$ y $v = 5t^6$. Calcula la función compuesta $p(t)$.

3. Dadas las funciones $f(x, y) = x^2 + y$, $x = y^2$ obtén la función compuesta $f(y)$. Calcula
$$\frac{\partial f(x, y)}{\partial y} \quad \text{y} \quad \frac{df(y)}{dy}.$$

4. Consideremos las funciones $f(x, y, t) = x^2 e^{yt}$ y $x = y^2 + t^3$.
   (a) Calcula df en términos de dx, dy, dt.
   (b) Calcula df en términos de dy, dt.

5. Consideremos la función $f(x, y) = x^2 + y^2$, donde x e y dependen de t por las relaciones $x = \operatorname{sen} t$, $y = \cos t$. Comprueba mediante la regla de la cadena que
$$\frac{df}{dt} = 0.$$
¿Cómo se interpreta este hecho?

6. Dadas las funciones $f(u, v) = u^2 + uv^3$ y $u = xv$,
   (a) Calcula la diferencial de la función $f(u, v)$.
   (b) Calcula, por la regla de la cadena, la diferencial de $f(v, x)$.
   (c) Calcula la función compuesta $f(v, x)$ y calcula directamente su diferencial. Comprueba que el resultado es el mismo que el del apartado anterior.

7. El coste de producción de una empresa está en función del precio de cada uno de los dos inputs que utiliza, $C(x, y) = 2 + 3x + 5y$ u.m. Por otra parte, el precio de los inputs varía con el tiempo según las funciones
$$x(t) = 1 + t \text{ u.m.}, \qquad y(t) = 1 + 2t \text{ u.m.},$$
donde t es el tiempo expresado en años.

(a) Calcula los precios de los inputs y el coste correspondiente en $t = 0$.

(b) Calcula la función $C(t)$ y el incremento de costes correspondiente al primer año (periodo $[0, 1]$).

(c) Calcula las derivadas

$$\left.\frac{\partial C}{\partial x}\right|_{(1,1)}, \quad \left.\frac{\partial C}{\partial y}\right|_{(1,1)}, \quad \left.\frac{dC}{dt}\right|_{0}$$

derivando directamente cada función (indica las unidades correspondientes). Interpreta las derivadas.

(d) Calcula la última derivada del apartado anterior mediante la regla de la cadena.

8. Sea $B(p, p')$ la función de beneficios de una empresa, donde p es el precio de su producto y p' el precio medio de la competencia. Para los precios actuales $p = 21$, $p' = 20$ se estima que

$$\left.\frac{\partial B(p,p')}{\partial p}\right|_{(21,20)} = -3, \qquad \left.\frac{\partial B(p,p')}{\partial p'}\right|_{(21,20)} = 2.$$

Supongamos que la competencia ajusta sus precios según los de la empresa, de modo que $p' = p - 1$. Calcula

$$\left.\frac{dB(p)}{dp}\right|_{21}$$

y explica la diferencia entre esta derivada y la anterior, desde un punto de vista matemático y en cuanto a su interpretación económica. ¿Cuál de ellas nos indica el efecto que tendría sobre los beneficios una disminución del precio p de 2 u.m.?

9. Una empresa fabrica un producto que vende a un precio de 6 u.m. por unidad. La producción tiene unos costes fijos de 50 u.m., y unos costes variables de 2 u.m. Además la empresa destina una cantidad P a publicidad. Con estos datos, la función de beneficios de la empresa es

$$B(x, D, P) = 6D - 2x - P - 50,$$

donde x es la cantidad de producto que fabrica y D la demanda del producto en el mercado. Justifica esta afirmación.

(a) Calcula las derivadas parciales de B e interprétalas.

(b) Supongamos que la empresa ajusta su producción a su demanda, es decir, considera a x como función de D en la forma más simple posible: $x = D$. Calcula la función compuesta $B(D, P)$ así como sus derivadas. Explica las diferencias respecto de las derivadas anteriores.

(c) El signo de $\partial B/\partial P$ es negativo, ¿como se interpreta esto?, ¿es razonable?

(d) La demanda de la empresa depende de su inversión en publicidad, es decir, tenemos una función $D(P)$. La empresa no conoce esta función, pero estima que, para la inversión actual P_0, se cumple

$$\left.\frac{\partial D}{\partial P}\right|_{P_0} = 1.$$

¿Es esto razonable?

(e) No podemos calcular la función compuesta $B(P)$, pero sí que podemos calcular

$$\left.\frac{dB}{dP}\right|_{P_0}.$$

Calcula esta derivada e interprétala. ¿Le conviene a la empresa aumentar su inversión en publicidad?

10. Sea $D(p, r, t)$ la función de demanda de un artículo en un mercado, donde p es el precio, r la renta media de los consumidores y t el tiempo en años. Actualmente ($t = 0$) se tiene $(p, r) = (5, 15)$ y $D(5, 15, 0) = 200$. Además

$$\left.\frac{\partial D}{\partial t}\right|_{(5,15,0)} = 20, \qquad \left.\frac{\partial D}{\partial p}\right|_{(5,15,0)} = -15, \qquad \left.\frac{\partial D}{\partial r}\right|_{(5,15,0)} = 10.$$

(a) Interpreta estas derivadas.

(b) ¿Qué demanda cabría esperar dentro de un año si la renta ha pasado a $r = 16$ u.m. y el precio a $p = 4.5$ u.m.?, ¿qué hipótesis sobre D es necesaria para responder a esta pregunta con los datos disponibles?

(c) Supongamos que $r = r(t)$ y $p = p(t)$, de modo que

$$\left.\frac{dr}{dt}\right|_0 = 0.2, \qquad \left.\frac{dp}{dt}\right|_0 = 0.1.$$

Calcula

$$\left.\frac{dD(t)}{dt}\right|_0.$$

(d) Interpreta esta última derivada explicando la diferencia con la interpretación de

$$\left.\frac{\partial D(p, r, t)}{\partial t}\right|_{(5,15,0)}.$$

11. Estudia la homogeneidad de las funciones siguientes:

(a) $f(x, y) = \sqrt[5]{xy^2}$,

(b) $P(r, s) = r + s$,

(c) $Q(K, L) = K^3 L^5$,

(d) $g(a, b, c) = \dfrac{a^2 c + b^3}{a - b + c}$.

(e) $g(x, y, z) = xy + x^2 - y$.

12. Comprueba que las funciones del ejercicio anterior cumplen el teorema de Euler.

13. Comprueba que la función $f(x, y, z) = x^2 \operatorname{sen}(y/z)$ cumple el teorema de Euler.

14. Calcula mediante el teorema de Euler el grado de homogeneidad de la función
$$f(x, y) = \frac{1}{x + y}.$$

15. Sea $f(x, y) = xy^3\, e^{x/y} \operatorname{sen} \sqrt{\cos(x/y)}$. Calcula el valor de
$$\frac{\partial f}{\partial x}x + \frac{\partial f}{\partial y}y.$$
¿Es homogénea la función $\frac{\partial f}{\partial x}$?

16. Dada la función
$$f(x, y, z) = \sqrt[3]{\frac{x \operatorname{sen}(y/z)}{y^6 z}}$$
Determina si las derivadas parciales de f son homogéneas y, en caso afirmativo, calcula su grado.

17. Consideramos un mercados con tres bienes cuyos precios son p_1, p_2, p_3. Para un consumidor con renta R las funciones de demanda son:
$$x_i(p_1, p_2, p_3, R) = \frac{Rp_i^3}{p_1^4 + p_2^4 + p_3^4 + R^4}, \quad i = 1, 2, 3.$$
Comprueba que el modelo está libre de ilusión monetaria.

18. Comprueba que cualquier función de producción del tipo Cobb-Douglas
$$Q(K, L) = AK^\alpha L^\beta$$
es una función homogénea. Explica la relación que debe darse para que los rendimientos a escala sean decrecientes.

19. La función de producción de una empresa es
$$Q(K, L) = A + \beta \left(\frac{K}{L}\right)^\alpha,$$
donde $A \geq 0$, $\alpha > 0$ y $\beta > 0$.
    (a) Estudia la homogeneidad de Q en función del valor de los parámetros A, α y β.
    (b) Para los valores en los que la función sea homogénea, calcula el tipo de rendimientos a escala que presenta.

9. Convexidad

Estudiamos en este tema los conjuntos y las funciones convexas. Se trata de unas familias de conjuntos y funciones cuyo comportamiento es especialmente sencillo desde un punto de vista matemático e intervienen en la modelización de muchas situaciones económicas.

9.1 Conjuntos convexos

Para definir los conjuntos convexos necesitamos algunos conceptos previos:

Si $\bar{p} \neq \bar{q}$ son dos puntos de $\mathbb{R}^n$, la recta que pasa por ellos está formada por los puntos de la forma

$$\bar{x} = (1-\lambda)\bar{p} + \lambda\bar{q}, \qquad \lambda \in \mathbb{R}.$$

Al variar λ obtenemos los distintos puntos de la recta. Concretamente, para $\lambda = 0$ pasamos por $\bar{x} = \bar{p}$ y para $\lambda = 1$ pasamos por $\bar{x} = \bar{q}$.

Si exigimos que λ tome sólo valores ≥ 0 o sólo valores ≤ 0 obtenemos los puntos de las dos semirrectas de extremo $\bar{p}$ contenidas en la recta anterior. La primera es la que contiene a $\bar{q}$, y la segunda es la que no lo contiene.

Si exigimos que λ varíe entre 0 y 1 obtenemos los puntos del segmento de extremos $\bar{p}$ y $\bar{q}$.

Definición Se dice que un conjunto $C \subset \mathbb{R}^n$ es *convexo* si cuando $\bar{x}$, $\bar{y} \in C$ y $0 \leq \lambda \leq 1$, entonces $(1-\lambda)\bar{x} + \lambda\bar{y} \in C$.

Es decir, C es convexo si cuando contiene dos puntos contiene también a todos los puntos del segmento que los une. Se considera que el conjunto vacío y los conjuntos con un solo punto son convexos. También es claro que $\mathbb{R}^n$ es convexo.

Ejemplo *Estudia analíticamente la convexidad de*

$$C_1 = \{(x,y) \in \mathbb{R}^2 \mid 2x + y = 4\}, \qquad C_2 = \{(0,0),(1,2),(2,2)\}.$$

SOLUCIÓN: Gráficamente se comprueba que C_1 es una recta y, por consiguiente, es convexo. No obstante, como nos piden una solución analítica aplicamos la definición: tomamos (x_1, y_1), $(x_2, y_2) \in C_1$ y $\lambda \in [0, 1]$. Hemos de comprobar que $(1-\lambda)(x_1, y_1) + \lambda(x_2, y_2) \in C_1$. Para ello operamos:

$$(1-\lambda)(x_1, y_1) + \lambda(x_2, y_2) = ((1-\lambda)x_1 + \lambda x_2, (1-\lambda)y_1 + \lambda y_2).$$

Este punto estará en C_1 si cumple la ecuación $2x + y = 4$. Calculamos:

$$\begin{aligned} 2((1-\lambda)x_1 + \lambda x_2) + (1-\lambda)y_1 + \lambda y_2 &= (1-\lambda)(2x_1 + y_1) + \lambda(2x_2 + y_2) \\ &= (1-\lambda)4 + \lambda 4 = 4, \end{aligned}$$

donde hemos usado que

$$\begin{aligned} (x_1, y_1) \in C_1 \Rightarrow 2x_1 + y_1 = 4, \\ (x_2, y_2) \in C_1 \Rightarrow 2x_2 + y_2 = 4. \end{aligned}$$

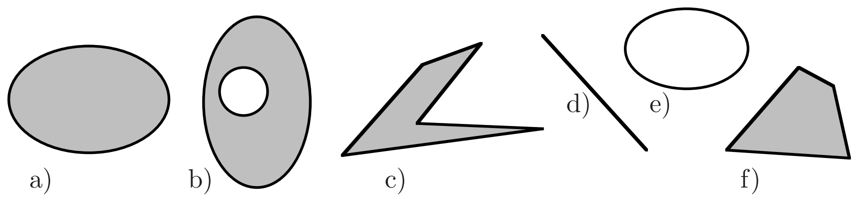
Por tanto C_1 es convexo.

Respecto a C_2, vemos que está formado por tres puntos, y el segmento que une dos de ellos contiene otros puntos que no pertenecen a C_2. Analíticamente, tomamos por ejemplo $(x_1, y_1) = (0, 0)$ y $(x_2, y_2) = (1, 2)$ y $\lambda = 0.5$, con lo que

$$(1-\lambda)(x_1, y_1) + \lambda(x_2, y_2) = 0.5(0, 0) + 0.5(1, 2) = (0.5, 1) \notin C_2.$$

Por tanto C_2 no es convexo. ∎

Ejemplo *Determina mediante argumentos geométricos intuitivos cuáles de los conjuntos siguientes son convexos:*



SOLUCIÓN: Son convexos a), d) y f). ∎

Definición Un *hiperplano* en $\mathbb{R}^n$ es un subconjunto de la forma

$$H = \{\bar{x} \in \mathbb{R}^n \mid c_1x_1 + \cdots + c_nx_n = \alpha\},$$

donde c_i, $\alpha \in \mathbb{R}$. Si llamamos $\bar{c}$ al vector de componentes c_i, podemos escribir, más brevemente, $H = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} = \alpha\}$. Por ejemplo,

$$\{(x, y, z) \in \mathbb{R}^3 \mid x - 2y + 3z = 5\}$$

es un hiperplano.

Teorema *Todo hiperplano es un conjunto convexo.*

Demostración: Si $H = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} = \alpha\}$ es un hiperplano, donde $\bar{c} \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, tomamos dos puntos $\bar{x}$, $\bar{y} \in H$ y $0 \leq \lambda \leq 1$. Hemos de comprobar que $(1-\lambda)\bar{x} + \lambda\bar{y} \in H$. En efecto,

$$\bar{c}((1-\lambda)\bar{x} + \lambda\bar{y})) = (1-\lambda)\bar{c}\bar{x} + \lambda\bar{c}\bar{y} = (1-\lambda)\alpha + \lambda\alpha = \alpha,$$

donde hemos usado que $\bar{c}\bar{x} = \bar{c}\bar{y} = \alpha$ porque ambos puntos están en H. ∎

Definición Un *semiespacio* en $\mathbb{R}^n$ es un subconjunto de la forma

$$S = \{\bar{x} \in \mathbb{R}^n \mid c_1x_1 + \cdots + c_nx_n \leq \alpha\},$$

donde c_i, $\alpha \in \mathbb{R}$.

Llamando $\bar{c}$ al vector de componentes c_i, podemos expresar S en la forma $S = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} \leq \alpha\}$. Notemos que si cambiamos $\leq$ por $\geq$ seguimos teniendo un semiespacio, pues $\bar{c}\bar{x} \geq \alpha$ equivale a $-\bar{c}\bar{x} \leq -\alpha$. Cuando la desigualdad es estricta, se dice que S es un *semiespacio abierto.*

Por ejemplo, los conjuntos

$$\{(x,y,z) \in \mathbb{R}^3 \mid x - 2y + 3z \leq 5\} \quad \text{y} \quad \{(x,y,z) \in \mathbb{R}^3 \mid x - 2y + 3z \geq 5\}$$

son semiespacios.

Teorema *Todo semiespacio es convexo.*

Demostración: Si $S = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} \leq \alpha\}$ es un semiespacio, donde $\bar{c} \in \mathbb{R}^n$ y $\alpha \in \mathbb{R}$, tomamos dos puntos $\bar{x}$, $\bar{y} \in S$ y $0 \leq \lambda \leq 1$. Hemos de comprobar que $(1-\lambda)\bar{x} + \lambda\bar{y} \in S$. En efecto,

$$\bar{c}((1-\lambda)\bar{x} + \lambda\bar{y})) = (1-\lambda)\bar{c}\bar{x} + \lambda\bar{c}\bar{y} \leq (1-\lambda)\alpha + \lambda\alpha = \alpha,$$

donde hemos usado que $\bar{c}\bar{x} \leq \alpha$, $\bar{c}\bar{y} \leq \alpha$ porque ambos puntos están en S, así como que $\lambda \geq 0$ y $1 - \lambda \geq 0$ para que se conserven las desigualdades al multiplicar. La prueba para subespacios abiertos es idéntica. ∎

Observemos que cada hiperplano $H = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} = \alpha\}$ divide a $\mathbb{R}^n$ en dos semiespacios $H^+ = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} \geq \alpha\}$ y $H^- = \{\bar{x} \in \mathbb{R}^n \mid \bar{c}\bar{x} \leq \alpha\}$. Es obvio que H es la intersección de los dos semiespacios H^+ y H^-.

Teorema *La intersección de conjuntos convexos es convexa.*

Demostración: Por simplicidad consideramos el caso de dos conjuntos convexos C_1 y C_2. Si son más de dos se razona igualmente. Para probar que la intersección $C_1 \cap C_2$ es convexa comprobamos que satisface la definición. Tomamos dos puntos $\bar{x}$, $\bar{y} \in C_1 \cap C_2$ y $0 \leq \lambda \leq 1$. Hemos de probar que

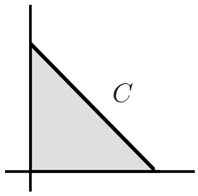
$$(1-\lambda)\bar{x} + \lambda y \in C_1 \cap C_2.$$

Del hecho de que $\bar{x}$, $\bar{y} \in C_1 \cap C_2$ se sigue que $\bar{x}$, $\bar{y} \in C_1$ y $\bar{x}$, $\bar{y} \in C_2$. Como C_1 es convexo se cumple que $(1-\lambda)\bar{x} + \lambda\bar{y} \in C_1$ e igualmente $(1-\lambda)\bar{x} + \lambda\bar{y} \in C_2$, lo cual significa que $(1-\lambda)\bar{x} + \lambda\bar{y} \in C_1 \cap C_2$, como queríamos probar. ∎

Definición Un *polítopo* es una intersección de un número finito de semiespacios. Si además está acotado se dice que es un *poliedro.* Por ejemplo, el conjunto

$$C = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq 1,\ x \geq 0,\ y \geq 0\}$$

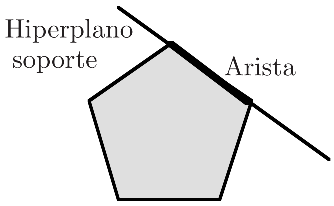
es un poliedro.

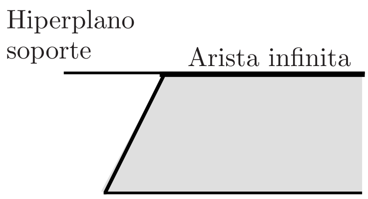


Según los teoremas anteriores los polítopos, y en particular los poliedros, son convexos.

Definición Si C es un conjunto convexo, se dice que un punto $\bar{x} \in C$ es un *punto extremo* de C si cuando $\bar{x}$ pertenece a un segmento con extremos en C, necesariamente es uno de sus extremos.

Un hiperplano H es un *hiperplano soporte* de un polítopo P si contiene puntos de P y P está contenido en uno de los semiespacios determinados por H. Una *cara* de P es la intersección de P con uno de sus hiperplanos soporte. Si $P \subset \mathbb{R}^2$, las caras se llaman *aristas.* Una arista puede ser finita o infinita. Si es finita, es un segmento que une puntos extremos del polítopo.





Ejemplo *Una empresa tiene capacidad para producir hasta un máximo de* 400 *unidades de un producto* A *y hasta un máximo de* 200 *unidades de un producto* B*. El coste unitario de* A *es de* 2 *u.m., y el de* B *es de* 1 *u.m. Además hay unos costes fijos de* 100 *u.m. El presupuesto disponible de la empresa es de* 700 *u.m. Representa gráficamente el conjunto de oportunidades de la empresa, es decir, el conjunto de todos los pares* (x, y) *correspondientes a una producción posible de* x *unidades de producto* A *e* y *unidades de producto* B*.*

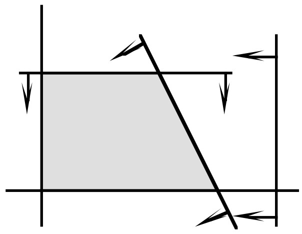
a) *Estudia su convexidad gráfica y analíticamente.*

b) *Calcula sus puntos extremos, sus aristas y sus aristas infinitas.*

Solución: El conjunto de oportunidades es

$$\begin{aligned} S &= \{(x, y) \in \mathbb{R}^2 \mid 2x + y + 100 \leq 700,\ 0 \leq x \leq 400,\ 0 \leq y \leq 200\} \\ &= \{(x, y) \in \mathbb{R}^2 \mid 2x + y \leq 600,\ 0 \leq x \leq 400,\ 0 \leq y \leq 200\}. \end{aligned}$$

Su representación gráfica es la siguiente:



a) Se ve claramente que cualquier segmento con extremos en S está contenido en S. Por lo tanto el conjunto es convexo.

Analíticamente es claro que S es un polítopo, puesto que es intersección de cinco semiespacios. Por lo tanto es convexo. De hecho, la figura muestra que está acotado y, por consiguiente, es un poliedro.

b) Como se ve en la figura, los puntos extremos de S son el $(0,0)$, la intersección de las dos rectas $y = 200$ y $2x + y = 600$, la intersección con el eje X de la recta $2x + y = 600$ y la intersección con el eje Y de la recta $y = 200$:

$$\left.\begin{aligned} 2x + y &= 600 \\ y &= 200 \end{aligned}\right\} \Rightarrow (x, y) = (200, 200),$$

$$\left.\begin{aligned} 2x + y &= 600 \\ y &= 0 \end{aligned}\right\} \Rightarrow (x, y) = (300, 0),$$

$$\left.\begin{aligned} y &= 200 \\ x &= 0 \end{aligned}\right\} \Rightarrow (x, y) = (0, 200).$$

Así pues, los puntos extremos de S son $(0,0)$, $(0,200)$, $(300,0)$ y $(200,200)$. Las aristas son los segmentos de ecuaciones

$$\begin{aligned} &(1-\lambda)(0,0) + \lambda(0,200), \\ &(1-\lambda)(0,200) + \lambda(200,200), \\ &(1-\lambda)(200,200) + \lambda(300,0), \\ &(1-\lambda)(300,0) + \lambda(0,0), \end{aligned}$$

para $0 \leq \lambda \leq 1$. Como S es un poliedro, no tiene aristas infinitas. ■

9.2 Funciones cóncavas y convexas

Definición Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función definida sobre un conjunto convexo D.

1. Diremos que f es *convexa* si cuando $\bar{x}, \bar{y} \in D$ y $0 \leq \lambda \leq 1$, se cumple que

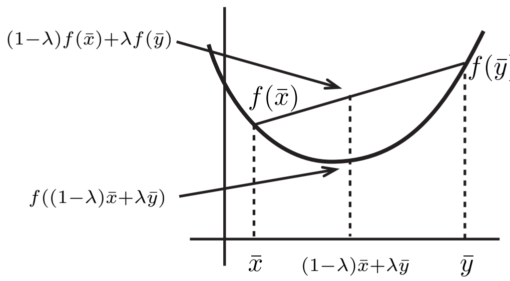
$$f((1-\lambda)\bar{x} + \lambda\bar{y}) \leq (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

2. Diremos que f es *cóncava* si cuando $\bar{x}, \bar{y} \in D$ y $0 \leq \lambda \leq 1$, se cumple que

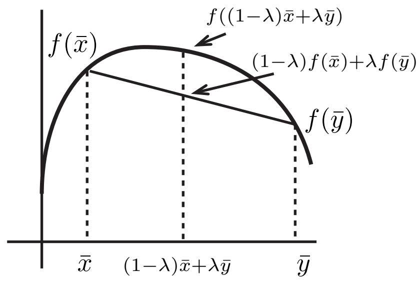
$$f((1-\lambda)\bar{x} + \lambda\bar{y}) \geq (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

3. Diremos que f es *estrictamente convexa* si cuando $\bar{x}$, $\bar{y} \in D$ cumplen $\bar{x} \neq \bar{y}$ y $0 < \lambda < 1$, entonces
$$f((1-\lambda)\bar{x} + \lambda\bar{y}) < (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

4. Diremos que f es *estrictamente cóncava* si cuando $\bar{x}$, $\bar{y} \in D$ cumplen $\bar{x} \neq \bar{y}$ y $0 < \lambda < 1$, entonces
$$f((1-\lambda)\bar{x} + \lambda\bar{y}) > (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

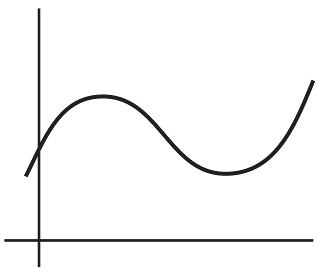


Función convexa



Función cóncava

Observemos que estas definiciones no son excluyentes, es decir, una función puede ser a la vez cóncava y convexa, cóncava y estrictamente cóncava, etc. Además, una función puede no ser ni cóncava ni convexa. Por otra parte, la concavidad o convexidad de una función puede depender del dominio D sobre el que la consideramos, en el sentido de que una misma función puede ser cóncava o convexa sobre un conjunto D y no serlo sobre otro.



Función ni cóncava ni convexa.

Ejemplo *Comprueba que una función $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ de la forma $f(\bar{x}) = \bar{c}\bar{x}$, donde $\bar{c} \in \mathbb{R}^n$ es a la vez cóncava y convexa.*

SOLUCIÓN: Si $\bar{x}$, $\bar{y} \in \mathbb{R}^n$ y $0 \leq \lambda \leq 1$, entonces
$$f((1-\lambda)\bar{x} + \lambda\bar{y}) = \bar{c}((1-\lambda)\bar{x} + \lambda\bar{y}) = (1-\lambda)\bar{c}\bar{x} + \lambda\bar{c}\bar{y} = (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

Al darse la igualdad, se cumple tanto la desigualdad $\leq$ como $\geq$, luego f es cóncava y convexa. ■

Teorema *Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^2 sobre un abierto convexo D.*

1. *La función f es convexa en D si y sólo si $Hf(\bar{x})$ es semidefinida positiva en todo punto $\bar{x} \in D$.*
2. *La función f es cóncava en D si y sólo si $Hf(\bar{x})$ es semidefinida negativa en todo punto $\bar{x} \in D$.*
3. *Si $Hf(\bar{x})$ es definida positiva en todo $\bar{x} \in D$ entonces f es estrictamente convexa.*
4. *Si $Hf(\bar{x})$ es definida negativa en todo $\bar{x} \in D$ entonces f es estrictamente cóncava.*

Ejemplo *Estudia al convexidad de la función $f(x, y) = x^2 - y^3$ sobre los conjuntos*

$$D_1 = \{(x, y) \in \mathbb{R}^2 \mid y < 0\} \quad y \quad D_2 = \{(x, y) \in \mathbb{R}^2 \mid y > 0\}.$$

Solución: Los conjuntos D_1 y D_2 son convexos porque son semiespacios, y la función f es de clase C^2 porque es un polinomio. Por lo tanto podemos aplicar el teorema. La hessiana de f es

$$Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -6y \end{pmatrix}.$$

Como es diagonal, para clasificarla basta observar el signo de sus coeficientes. Sobre D_1 tenemos que $2 > 0$ y $-6y > 0$, luego $Hf(x, y)$ es definida positiva en todo punto y f es estrictamente convexa.

En cambio, sobre D_2 se cumple que $2 > 0$ y $-6y < 0$, luego $Hf(x, y)$ es indefinida y la función no es ni cóncava ni convexa. ∎

Propiedades de las funciones cóncavas y convexas Consideremos una función $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, donde D es convexo. Se cumple:

1. f es estrictamente convexa si y sólo si $-f$ es estrictamente cóncava.
2. f es convexa si y sólo si $-f$ es cóncava.
3. Si f es estrictamente convexa (respectivamente, estrictamente cóncava) entonces f es convexa (respectivamente, cóncava), pero el recíproco no es cierto.
4. Si f es convexa y $\alpha \in \mathbb{R}$, el conjunto de nivel inferior

   $$D_\alpha = \{\bar{x} \in D \mid f(\bar{x}) \leq \alpha\}$$

   es convexo.

5. Si f es cóncava y $\alpha \in \mathbb{R}$, el conjunto de nivel superior

$$D^\alpha = \{\bar{x} \in D \mid f(\bar{x}) \geq \alpha\}$$

es convexo.

Demostración:
1) Supongamos que f es estrictamente convexa. Para probar que $-f$ es estrictamente cóncava tomamos dos puntos $\bar{x}$, $\bar{y} \in D$ y un número real $0 < \lambda < 1$. Como f es estrictamente convexa tenemos que

$$f((1-\lambda)\bar{x} + \lambda\bar{y}) < (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

Multiplicando ambos miembros por -1 queda

$$(-f)((1-\lambda)\bar{x} + \lambda\bar{y}) > (1-\lambda)(-f)(\bar{x}) + \lambda(-f)(\bar{y}),$$

por lo que $-f$ cumple la definición de función estrictamente cóncava. Invirtiendo el razonamiento obtenemos el recíproco.

2) es análogo a 1).

3) Supongamos que f es estrictamente convexa y vamos a probar que es convexa. Para ello tomamos dos puntos $\bar{x}$, $\bar{y} \in D$ y $0 \leq \lambda \leq 1$. Hemos de probar que

$$f((1-\lambda)\bar{x} + \lambda\bar{y}) \leq (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

Distinguimos dos casos:

A) Si $\bar{x} = \bar{y}$ entonces

$$(1-\lambda)\bar{x} + \lambda\bar{y} = (1-\lambda)\bar{x} + \lambda\bar{x} = \bar{x} - \lambda\bar{x} + \lambda\bar{x} = \bar{x}$$

e igualmente $(1-\lambda)f(\bar{x}) + \lambda f(\bar{y}) = f(\bar{x})$, por lo que la desigualdad que hemos de probar se reduce a $f(\bar{x}) \leq f(\bar{x})$, que es obviamente cierta.

B) Si $\bar{x} \neq \bar{y}$ distinguimos a su vez tres subcasos:

B1) Si $0 < \lambda < 1$ entonces tenemos

$$f((1-\lambda)\bar{x} + \lambda\bar{y}) < (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}).$$

por la definición de función estrictamente convexa, luego también tenemos la desigualdad anterior.

B2) Si $\lambda = 0$ la desigualdad que hemos de probar se reduce a $f(\bar{x}) \leq f(\bar{x})$, que obviamente es cierta.

B3) Si $\lambda = 1$, la desigualdad que hemos de probar se reduce a $f(\bar{y}) \leq f(\bar{y})$, que también es cierta.

En cualquier caso tenemos la desigualdad buscada. El caso para funciones cóncavas es análogo.

4) Supongamos que f es convexa. Para probar que D_α es convexo tomamos dos puntos $\bar{x}$, $\bar{y} \in D_\alpha$ y un número real $0 \leq \lambda \leq 1$. Hemos de probar que $(1-\lambda)\bar{x} + \lambda\bar{y} \in D_\alpha$, para lo cual hemos de comprobar que $f((1-\lambda)\bar{x} + \lambda\bar{y}) \leq \alpha$. Ahora bien, como $f(\bar{x}) \leq \alpha$ y $f(\bar{y}) \leq \alpha$ y f es convexa, se cumple:

$$f((1-\lambda)\bar{x} + \lambda\bar{y}) \leq (1-\lambda)f(\bar{x}) + \lambda f(\bar{y}) \leq (1-\lambda)\alpha + \lambda\alpha = \alpha - \lambda\alpha + \lambda\alpha = \alpha.$$

5) es análogo a 4). ■

Ejemplo *Estudia la convexidad de los conjuntos*

$$C_1 = \{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 \leq 5\}, \qquad C_2 = \{(x,y) \in \mathbb{R}^2 \mid -3x^2-y^2+xy \geq 3\}.$$

Solución: El conjunto C_1 es un conjunto de nivel inferior correspondiente a la función $f_1(x,y,z) = x^2 + y^2 + z^2$. Por el teorema anterior, basta probar que f_1 es convexa. Para ello calculamos su matriz hessiana:

$$Hf_1(x,y,z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Como es diagonal y los coeficientes de la diagonal son positivos, Hf_1 es definida positiva, luego f_1 es convexa.

Similarmente, C_2 es un conjunto de nivel superior correspondiente a la función $f_2(x,y) = -3x^2 - y^2 + xy$. Su matriz hessiana es

$$Hf_2(x,y) = \begin{pmatrix} -6 & 1 \\ 1 & -2 \end{pmatrix}.$$

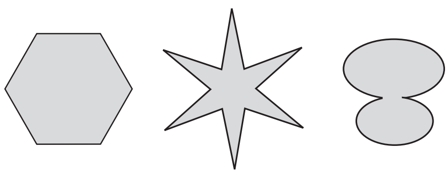
Para clasificarla, calculamos en primer lugar $|Hf_2(x,y)| = 11 \neq 0$. Por consiguiente, basta estudiar los signos de los menores principales conducentes:

$$A_1 = -6 < 0, \quad A_2 = 11 > 0.$$

Por la regla de Jacobi, $Hf_2(x,y)$ es definida negativa, luego f_2 es cóncava y, por el teorema anterior, C_2 es convexo. ■

9.3 Ejercicios

1. Estudia gráficamente la convexidad de los conjuntos siguientes:



2. Estudia la convexidad de los conjuntos siguientes:

   (a) $C_1 = \{(x, y) \in \mathbb{R}^2 \mid y = 2\}$,
   (b) $C_2 = \{(x, y, z) \in \mathbb{R}^3 \mid x + y = 2,\ z = 5\}$,
   (c) $C_3 = \{(x, y) \in \mathbb{R}^2 \mid xy \geq 5\}$,
   (d) $C_4 = \{(-1, 3), (1, 1), (2, 3)\}$.

3. Estudia si las funciones siguientes son cóncavas o convexas en los dominios indicados:

   (a) $f(x) = e^x$ en $D = \mathbb{R}$.
   (b) $u(s, t) = s^2 + 2st + t^2$ en $D = \mathbb{R}^2$.
   (c) $R(a, b) = \dfrac{1}{ab}$, en $D = \{(x, y) \in \mathbb{R}^2 \mid x > 0,\ y > 0\}$.
   (d) $K(x, y) = xe^y$ en $D = \mathbb{R}^2$.
   (e) $M(u, v) = uv - u^3 - v^3$ en $D = \{(u, v) \in \mathbb{R}^n \mid u > 1,\ v > 1\}$.
   (f) $P(x, y, t) = e^{x+y} + t^3$ en $D = \{(x, y, t) \in \mathbb{R}^3 \mid t > 0\}$.
   (g) $Q(x, y, t) = e^{x+y} + t^4$ en $D = \mathbb{R}^3$.
   (h) $L(x, y, z) = x^2y - 3xz^2$ en $D = \mathbb{R}^2$.
   (i) $C(r, s, t) = 5r^4 + s + 2t^6$ en $D = \mathbb{R}^3$.
   (j) $T(p, q, r) = p^3r + q^3r - 5r$ en $D = \{(p, q, r) \in \mathbb{R}^3 \mid p > 0,\ q > 0,\ r > 0\}$.
   (k) $T(p, q, r) = p^3r + q^3r - 5r$ en $D = \{(p, q, r) \in \mathbb{R}^3 \mid p < 0,\ q < 0,\ r > 0\}$.
   (l) $S(u, v, w) = \ln(u+v+w)$ en $D = \{(u, v, w) \in \mathbb{R}^3 \mid u > 0,\ v > 0,\ w > 0\}$.
   (m) $g(u, v, w) = u \ln(vw)$ en $D = \{(u, v, w) \in \mathbb{R}^3 \mid u > 0,\ v > 0,\ w > 0\}$.
   (n) $h(x, y, z, w) = 3x - 5y + 2z - w$ en $D = \mathbb{R}^4$.

4. El dominio de una función convexa ha de ser un conjunto convexo, por definición. ¿Son convexos los dominios de las funciones del problema anterior?

5. Estudia si los conjuntos siguientes son convexos. Indica cuáles de ellos son hiperplanos, semiespacios o polítopos:

   (a) $C_1 = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - 5 + z \leq 5,\ x < 16\}$.
   (b) $C_2 = \{(x, y, z, w) \in \mathbb{R}^4 \mid x - 3y = x + 2w\}$.
   (c) $C_3 = \mathbb{R}^2 \setminus \{(0, 0)\}$.
   (d) $C_4 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 16\}$.
   (e) $C_5 = \{(x, y) \in \mathbb{R}^2 \mid x + y \leq 16\}$.
   (f) $C_6 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^4 \geq 4\}$.
   (g) $C_7 = \{(x, y) \in \mathbb{R}^2 \mid 2xy - x^2 - y^2 \geq 4\}$.
   (h) $C_8 = \{(x, y) \in \mathbb{R}^2 \mid x \neq 3\}$.
   (i) $C_9 = \{(x, y, z) \in \mathbb{R}^3 \mid x < 0,\ y > 0\}$.
   (j) $C_{10} = \{(x, y, z) \in \mathbb{R}^3 \mid x < y + z\}$.

(k) $C_{11} = \{(x, y, z) \in \mathbb{R}^3 \mid \ln(x + y + z) \geq 0,\ x + y + z > 0\}$.

6. Indica si las afirmaciones siguientes son verdaderas o falsas. Si son verdaderas explica por qué, y si son falsas pon un ejemplo que lo muestre:

   (a) Una función puede ser estrictamente convexa y cóncava.

   (b) Una función puede ser estrictamente convexa y estrictamente cóncava.

   (c) Una función puede ser estrictamente cóncava y convexa.

   (d) Una función no puede ser convexa y cóncava a la vez.

   (e) Una función puede no ser cóncava ni convexa.

   (f) Toda función cóncava es estrictamente cóncava.

   (g) Un conjunto puede ser cóncavo y convexo a la vez.

10. Optimización clásica

10.1 Conceptos de programación matemática

La programación matemática se plantea calcular el máximo o el mínimo de una función de una o varias variables sujetas a un conjunto de restricciones. Supongamos que la función de utilidad de un consumidor es

$$U(x_1, x_2) = \ln(1 + x_1\, x_2),$$

donde x_1 y x_2 son las cantidades adquiridas de dos bienes A y B. La función $U(x_1, x_2)$ existe cuando $1 + x_1 x_2 > 0$. Como, además, se trata de una función de utilidad, podemos considerarla definida en $D = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \geq 0,\ x_2 \geq 0\}$.

Si nos preguntamos qué cantidades (x_1, x_2) maximizan la utilidad, es obvio que cuanto mayor sea el consumo mayor será la utilidad, es decir, la función $U(x_1, x_2)$ no tiene un máximo global. Sin embargo, en la práctica, el consumidor no podrá conseguir cualquier nivel de utilidad que pretenda, porque ello podría suponer un gasto mayor del que puede permitirse. Supongamos que el precio de A es de 2€, el precio de B es de 1€ y el consumidor dispone de un presupuesto de 12€. En este caso las únicas cantidades que tiene sentido considerar son las que satisfacen la restricción presupuestaria $2x + y \leq 12$. Ahora sí tiene sentido preguntarse la máxima utilidad que podemos conseguir para unos niveles de consumo (x, y). En los términos anteriores, buscamos el máximo global de la función U relativo al conjunto $S = \{(x, y) \in D \mid 2x+y \leq 12\}$. La forma usual de plantear matemáticamente este ejemplo es

$$\begin{array}{rl} \text{Max} & \ln(1 + x_1 x_2) \\ \text{s.a.} & 2x_1 + x_2 \leq 12 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{array}$$

Formulación general de un problema de programación matemática La forma general de plantear estos problemas es

$$\begin{array}{rl} \text{Max} & f(\bar{x}) \\ \text{s.a.} & g(\bar{x}) \leq \bar{b} \end{array} \quad \text{o} \quad \begin{array}{rl} \text{Min} & f(\bar{x}) \\ \text{s.a.} & g(\bar{x}) \geq \bar{b} \end{array}$$

donde $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ y $\bar{b} \in \mathbb{R}^m$.

En esta formulación aparecen los siguientes elementos:

1. La *función objetivo* $f(\bar{x})$. Describe matemáticamente lo que se quiere maximizar (Max) o minimizar (Min), esto es, lo que se quiere optimizar. Cualquier problema de maximización puede transformarse en uno de minimización, o viceversa, mediante la relación $\text{Min} f(x) = -\text{Max}(-f(x))$, lo cual quiere decir que $\bar{x}^*$ es un mínimo de $f(\bar{x})$ si y sólo si $\bar{x}^*$ es un máximo de $-f(\bar{x})$ y, obviamente, $f(\bar{x}^*) = -(-f(\bar{x}^*))$.

2. Las *restricciones* $g(\bar{x}) \leq \bar{b}$ o $g(\bar{x}) \geq \bar{b}$. La expresión $g(\bar{x}) \leq \bar{b}$ significa que existen m desigualdades entre las funciones coordenadas de g y las componentes de $\bar{b}$, es decir, $g_i(x) \leq b_i$, para $i = 1, \dots, m$. Decimos que cada una de ellas es una *restricción* del problema. Si es necesario, se puede suponer que todas las restricciones de un problema dado son del mismo tipo ($\leq$, $\geq$ o $=$), puesto que una restricción cualquiera se puede convertir en una (o dos) de otro tipo matemáticamente equivalente. Estas transformaciones se estudiarán, no obstante, en cursos de optimización más avanzados puesto que todas las restricciones que aparecerán en este capítulo serán de igualdad.

3. Las *variables de decisión* son $\bar{x} = (x_1, \dots x_n)$. Resolver el problema es encontrar unos valores de $(x_1, \dots x_n)$ *sujetos a* las restricciones del problema (esto se abrevia como "s.a." en el problema) para los que la función objetivo alcance el óptimo.

4. El *conjunto de oportunidades,* que representaremos por S, está formado por los puntos de $\mathbb{R}^n$ que cumplen les restriccions y pertenecen al dominio de la función f. A cualquier punto del conjunto de oportunidades se le llama *solución factible.*

Ejemplo *Consideramos el problema*

$$\begin{array}{ll} Max & 2x_1 + x_2 + x_3^2 - 4x_4 \\ s.a. & x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq 1 \\ & x_1, \ x_2 \geq 0 \end{array}$$

a) Escribe la función objetivo, las restricciones y las variables de decisión.
b) Calcula el conjunto de oportunidades y escribe una solución factible y una no factible.

Solución: Tenemos que

a) La función objetivo es $f : \mathbb{R}^4 \longrightarrow \mathbb{R}$ dada por

$$f(x_1, x_2, x_3, x_4) = 2x_1 + x_2 + x_3^2 - 4x_4.$$

Nótese que el dominio de f es $\mathbb{R}^4$ al tratarse de una función polinómica.

Las restricciones son $x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq 1$, $x_1 \geq 0$ y $x_2 \geq 0$. Observemos que las restricciones se pueden expresar con la función vectorial $g : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ dada por $g(x_1, x_2, x_3, x_4) = (x_1^2 + x_2^2 + x_3^2 + x_4^2, x_1, x_2)$ y el vector $\bar{b} = (1, 0, 0)$. Puesto que todas las funciones coordenadas son polinomios el dominio de g es $\mathbb{R}^4$.

Las variables de decisión son (x_1, x_2, x_3, x_4).

b) Como los dominios de f y g son $\mathbb{R}^4$, el conjunto de oportunidades está definido sólo por las restricciones y es:

$$S = \{(x_1, x_2, x_3, x_4) \in R^4 : x_1^2 + x_2^2 + x_3^2 + x_4^2 \leq 1, x_1 \geq 0, x_2 \geq 0\}.$$

Las soluciones factibles son los puntos del conjunto de oportunidades, esto es, los puntos que cumplen las restricciones. Por ejemplo, $(0, 0, 0, 0)$. En cambio, el punto $(-1, 0, 0, 0)$ es una solución no factible puesto que $-1 < 0$. ■

Óptimos globales y locales Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función escalar definida en un abierto D y sea $\bar{x}^* \in D$.

1. Se dice que $\bar{x}^*$ es un *máximo global estricto absoluto* de f si para todo $\bar{x} \in D$, $\bar{x} \neq \bar{x}^*$, se cumple $f(\bar{x}^*) > f(\bar{x})$.

2. Se dice que $\bar{x}^*$ es un *mínimo global estricto absoluto* de f si para todo $\bar{x} \in D$, $\bar{x} \neq \bar{x}^*$, se cumple $f(\bar{x}^*) < f(\bar{x})$.

3. Se dice que $\bar{x}^*$ es un *máximo global no estricto absoluto* de f si para todo $\bar{x} \in D$ se cumple $f(\bar{x}^*) \geq f(\bar{x})$.

4. Se dice que $\bar{x}^*$ es un *mínimo global no estricto absoluto* de f si para todo $\bar{x} \in D$ se cumple $f(\bar{x}^*) \leq f(\bar{x})$.

5. Se dice que $\bar{x}^*$ es un *máximo local estricto absoluto* de f si existe un $\epsilon > 0$ tal que si $\bar{x} \in D$, $\bar{x} \neq \bar{x}^*$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*) > f(\bar{x})$.

6. Se dice que $\bar{x}^*$ es un *mínimo local estricto absoluto* de f si existe un $\epsilon > 0$ tal que si $\bar{x} \in D$, $\bar{x} \neq \bar{x}^*$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*) < f(\bar{x})$.

7. Se dice que $\bar{x}^*$ es un *máximo local no estricto absoluto* de f si existe un $\epsilon > 0$ tal que si $\bar{x} \in D$ y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*)\ f(\bar{x})$.

8. Se dice que $\bar{x}^*$ es un *mínimo local no estricto absoluto* de f si existe un $\epsilon > 0$ tal que si $\bar{x} \in D$ y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*) \leq f(\bar{x})$.

Además, si $S \subset D$ es un subconjunto arbitrario entonces:

1. Se dice que $\bar{x}^*$ es un *máximo global estricto relativo a* S de f si $\bar{x}^* \in S$ y para todo $\bar{x} \in D \cap S$, $\bar{x} \neq \bar{x}^*$, se cumple $f(\bar{x}^*) > f(\bar{x})$.

2. Se dice que $\bar{x}^*$ es un *mínimo global estricto relativo a* S de f si $\bar{x}^* \in S$ y para todo $\bar{x} \in D \cap S$, $\bar{x} \neq \bar{x}^*$, se cumple $f(\bar{x}^*) < f(\bar{x})$.

3. Se dice que $\bar{x}^*$ es un *máximo global no estricto relativo a* S de f si $\bar{x}^* \in S$ y para todo $\bar{x} \in D \cap S$ se cumple $f(\bar{x}^*) \geq f(\bar{x})$.

4. Se dice que $\bar{x}^*$ es un *mínimo global no estricto relativo a* S de f si $\bar{x}^* \in S$ y para todo $\bar{x} \in D \cap S$ se cumple $f(\bar{x}^*) \leq f(\bar{x})$.

5. Se dice que $\bar{x}^*$ es un *máximo local estricto relativo a S* de f si $\bar{x}^* \in S$ y existe un $\epsilon > 0$ tal que para todo $\bar{x} \in D \cap S$, $\bar{x} \neq \bar{x}^*$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*) > f(\bar{x})$.

6. Se dice que $\bar{x}^*$ es un *mínimo local estricto relativo a S* de f si $\bar{x}^* \in S$ y existe un $\epsilon > 0$ tal que para todo $\bar{x} \in D \cap S$, $\bar{x} \neq \bar{x}^*$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$ se cumple $f(\bar{x}^*) < f(\bar{x})$.

7. Se dice que $\bar{x}^*$ es un *máximo local no estricto relativo a S* de f si $\bar{x}^* \in S$ y existe un $\epsilon > 0$ tal que para todo $\bar{x} \in D \cap S$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$, se cumple $f(\bar{x}^*) \geq f(\bar{x})$.

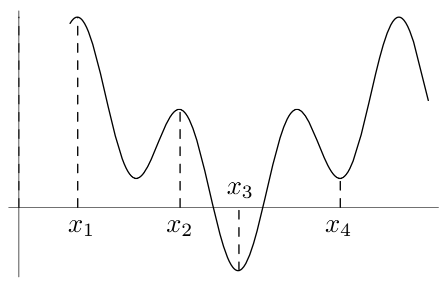
8. Se dice que $\bar{x}^*$ es un *mínimo local no estricto relativo a S* de f si $\bar{x}^* \in S$ y existe un $\epsilon > 0$ tal que para todo $\bar{x} \in D \cap S$, y $\|\bar{x} - \bar{x}^*\| < \epsilon$, se cumple $f(\bar{x}^*) \leq f(\bar{x})$.

A los puntos máximos o mínimos de cualquiera de estos tipos se les llama conjuntamente *extremos* de la función f.

Las definiciones anteriores son fáciles de recordar si se tiene en cuenta que:

Absoluto indica que consideramos la función sobre todo el abierto en el que está definida.
Relativo indica que consideramos únicamente los valores que toma la función sobre un cierto subconjunto S.
Global indica que la función toma en el punto un valor mayor (o menor) que en todos los demás puntos donde la estamos considerando.
Local indica que la función toma en el punto un valor mayor (o menor) que en los puntos de alrededor.
Estricto indica que el valor que toma la función en el punto no es igualado por el que toma en los demás puntos considerados.
No estricto indica que puede haber otros puntos donde la función tome el mismo valor.

Ejemplo *La función* $f :]3, 10[\subset \mathbb{R} \longrightarrow \mathbb{R}$ *dada por* $f(x) = 2 - \dfrac{\text{sen}(4x)}{\text{sen}(x)}$ *se representa gráficamente como*



Utilizando la gráfica clasifica los puntos x_1, x_2, x_3 *y* x_4.

SOLUCIÓN:

1. El punto x_1 es un máximo global no estricto. En este punto la función vale más (por eso es máximo) que en cualquier otro punto del dominio (por eso es global). No es estricto porque el valor $f(x_1)$ se repite, es decir hay otro punto del dominio para el que la función toma el mismo valor.

2. El punto x_2 es un máximo local estricto. Es un máximo porque en este punto la función vale más que en cualquier otro punto del dominio cercano a él y es local porque hay otros puntos, por ejemplo x_1, donde la función es mayor. Es estricto porque alrededor de él no hay otro punto que tome su misma imagen.

3. El punto x_3 es un mínimo global estricto. En él la función toma un valor más bajo (mínimo) que en cualquier otro punto del dominio (global). Es estricto porque no hay otro punto en el dominio que tome su misma imagen.

4. El punto x_4 es un mínimo local estricto. Es un mínimo porque en este punto la función vale menos que en cualquier otro punto del dominio cercano a él y es local porque hay otros puntos, por ejemplo x_1, donde la función es menor.

■

10.2 Optimización sin restricciones

Diremos que un problema es de *programación clásica sin restricciones* o de *optimización libre* cuando es de la forma

$$\mathrm{Opt} f(\bar{x}),$$

donde $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$.

El resto de la sección estará dedicado a ver los métodos mediante los cuales se puede resolver este problema, esto es, como se pueden calcular los máximos y los mínimos de una función escalar. Empezamos enunciando una condición necesaria de primer orden que nos permite calcular los puntos que pueden serlo:

Teorema *Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^1 en un abierto D. Si $\bar{x}^* \in D$ un extremo local absoluto de f, entonces $\nabla f(\bar{x}^*) = \bar{0}$.*

Definición Un *punto crítico* de una función f de clase C^1 es un punto $\bar{x}^*$ de su dominio tal que $\nabla f(\bar{x}^*) = \bar{0}$.

Ejemplo *Calcula los puntos críticos de las funciones siguientes:*

1. $f(x) = \ln x$.

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

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

SOLUCIÓN:

1. Calculamos $f'(x) = 1/x$ y planteamos la ecuación $1/x = 0$. Como esta ecuación no tiene solución, f no tiene puntos críticos.

2. El vector gradiente es $\nabla f(x, y) = (2x-2, 2y+2)$. Al igualar el vector gradiente al vector nulo, obtenemos el sistema

$$\begin{aligned} 2x - 2 &= 0 \\ 2y + 2 &= 0 \end{aligned}$$

cuya solución es $x = 1$, $y = -1$. Por tanto, $(1, -1)$ es el único punto crítico de f.

3. El gradiente es

$$\nabla f(x, y, z) = (-2x + 2z, -2y, 3z^2 + 2x),$$

que al igualarlo al vector nulo nos proporciona el sistema:

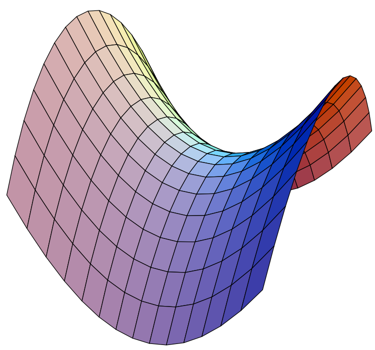
$$\begin{aligned} -2x + 2z &= 0 \\ -2y &= 0 \\ 3z^2 + 2x &= 0 \end{aligned}$$

A partir de la segunda ecuación, es obvio que $y = 0$. Despejamos $z = x$ de la primera ecuación y sustituimos en la tercera, con lo que obtenemos $3x^2 + 2x = 0$. Sacando factor común, $x(3x + 2) = 0$. Por tanto, $x = 0 = z$ ó $x = -2/3 = z$ y los puntos críticos son

$$(0, 0, 0) \quad \text{y} \quad (-\frac{2}{3}, 0, -\frac{2}{3}).$$

■

Según hemos visto, una condición necesaria para que un punto $\bar{x}^*$ sea extremo local de una función es que sea un punto crítico, es decir, si un punto no es crítico no puede ser un extremo local. Desafortunadamente esta condición no es suficiente y puede haber puntos críticos que no sean máximos ni mínimos. Un punto crítico que no es ni máximo ni mínimo se denomina *punto de silla*. El nombre proviene de la forma que adopta la gráfica de una función de dos variables alrededor de uno de estos puntos, similar a una silla de montar.



Punto de silla

El problema de decidir si un punto crítico de una función es máximo, mínimo o punto de silla se llama *clasificación* del punto crítico. El siguiente teorema, que es una condición suficiente de segundo orden para óptimos locales, nos proporciona una herramienta para resolverlo.

Teorema *Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^2 en el abierto D y sea $\bar{x}^* \in D$ un punto crítico. Entonces:*

1. *Si $Hf(\bar{x}^*)$ es definida positiva, $\bar{x}^*$ es un mínimo local absoluto de f.*
2. *Si $Hf(\bar{x}^*)$ es definida negativa, $\bar{x}^*$ es un máximo local absoluto de f.*
3. *Si $Hf(\bar{x}^*)$ es indefinida, $\bar{x}^*$ es un punto de silla de f.*

Ejemplo *Clasifica, si es posible, los puntos críticos de las funciones*

1. $f(x, y) = x^2 + y^2 - 2x + 2y + 2$.
2. $f(x, y, z) = -x^2 - y^2 + z^3 + 2xz$.

SOLUCIÓN:

1. En el ejemplo anterior hemos visto que el único punto crítico de la función

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

es $(1, -1)$. La matriz hessiana de f es

$$Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Con lo que

$$Hf(1, -1) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Calculamos los menores principales conducentes:

$$A_1 = |2| = 2,$$

$$A_2 = \begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4.$$

Puesto que $A_1 > 0$ y $A_2 > 0$, la forma cuadrática asociada a $Hf(1, -1)$ es definida positiva y el punto $(1, -1)$ es un mínimo local absoluto de f.

2. En el ejemplo anterior hemos calculado los puntos críticos de

$$f(x, y, z) = -x^2 - y^2 + z^3 + 2xz,$$

que son $(0, 0, 0)$ y $(-\frac{2}{3}, 0, -\frac{2}{3})$. Calculamos la matriz hessiana:

$$Hf(x, y, z) = \begin{pmatrix} -2 & 0 & 2 \\ 0 & -2 & 0 \\ 2 & 0 & 6z \end{pmatrix},$$

y ahora debemos hacer los cálculos para cada punto.

En el punto $(0,0,0)$,

$$Hf(0,0,0) = \begin{pmatrix} -2 & 0 & 2 \\ 0 & -2 & 0 \\ 2 & 0 & 0 \end{pmatrix},$$

cuyos menores principales conducentes son

$$A_1 = |-2| = -2,$$

$$A_2 = \begin{vmatrix} -2 & 0 \\ 0 & -2 \end{vmatrix} = 4$$

y

$$A_3 = \begin{vmatrix} -2 & 0 & 2 \\ 0 & -2 & 0 \\ 2 & 0 & 0 \end{vmatrix} = 8$$

Por el criterio de Jacobi, la forma cuadrática asociada a $Hf(0,0,0)$ es indefinida. Por tanto, el punto $(0,0,0)$ es un punto de silla de f, esto es, el punto $(0,0,0)$ no es un extremo de f.

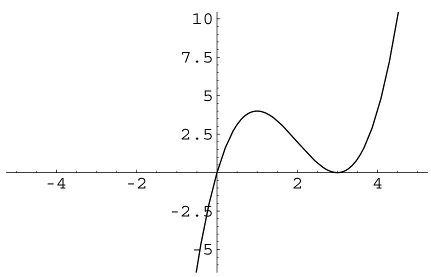
Análogamente, en el punto $(-\frac{2}{3}, 0, -\frac{2}{3})$, la matriz hessiana vale

$$Hf(-\frac{2}{3}, 0, -\frac{2}{3}) = \begin{pmatrix} -2 & 0 & 2 \\ 0 & -2 & 0 \\ 2 & 0 & -4 \end{pmatrix},$$

cuyos menores principales conducentes son $A_1 = -2$, $A_2 = 4$ y $A_3 = -8$. La forma cuadrática asociada a $Hf(-\frac{2}{3}, 0, -\frac{2}{3})$ es definida negativa y, en consecuencia, el punto $(-\frac{2}{3}, 0, -\frac{2}{3})$ es un máximo local absoluto.

■

La condición suficiente anterior nos permite calcular óptimos locales. Sin embargo, el hecho de tener un óptimo local no garantiza que la función en este punto tome su mejor valor, entendiendo por mejor que no existe ningun otro punto en el que el valor de la función sea mayor, si estamos maximizando, o menor, si estamos minimizando. Esto sólo se puede garantizar si el óptimo es global. Por ejemplo, consideremos la función $f(x) = x^3 - 6x^2 + 9x$ cuya gráfica es la siguiente:



Como se observa en la gráfica esta función tiene un máximo local absoluto en el punto $x = 1$ y un mínimo local absoluto en $x = 3$, hecho que es fácilmente verificable aplicando la condición necesaria y la condición suficiente. Pero $f(1) = 4$ mientras que existen otros puntos (por ejemplo, $f(5) = 20$) donde la función es mayor. Lo mismo ocurre con el mínimo. Este mismo ejemplo nos demuestra que ni siquiera el "mejor" óptimo local es el óptimo global. Por todo ello, necesitamos la siguiente condición suficiente para óptimos globales:

Teorema *Sea $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^2 en el convexo abierto D y sea $\bar{x}^* \in D$ un punto crítico. Entonces:*

1. *Si $f(\bar{x})$ es una función convexa en D, $\bar{x}^*$ es un mínimo global absoluto de f. Además, si la función es estrictamente convexa el mínimo es estricto.*
2. *Si $f(\bar{x})$ es una función cóncava en D, $\bar{x}^*$ es un máximo global absoluto de f. Además, si la función es estrictamente cóncava el máximo es estricto.*

Ejemplo *Calcula el mínimo valor que puede tomar la función $f(x, y) = x^2 + y^2$.*

Solución: Según el razonamiento visto en los párrafos anteriores, nos están pidiendo el valor que toma f en un mínimo global. Calculemos los puntos críticos de f para obtener los puntos que pueden ser óptimos. El gradiente de la función es $\nabla f(x, y) = (2x, 2y)$. Al igualarlo al vector nulo y resolver el sistema resultante, vemos que la función tiene un único punto crítico, $(x, y) = (0, 0)$.

Para clasificarlo le aplicamos la condición suficiente anterior. Puesto que f es un polinomio, su dominio es $\mathbb{R}^2$ que es un convexo abierto y es una función de clase C^2. La matriz hessiana de f es:

$$Hf(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

cuyos menores principales conducentes son $A_1 = 2$ y $A_2 = 4$. La forma cuadrática asociada a $Hf(x, y)$ es definida positiva y, en consecuencia, f es una función estrictamente convexa. Así pues, $(0, 0)$ es un mínimo global absoluto, que además es único, y el menor valor que toma la función es $f(0, 0) = 0$. ■

En resumen,

Para calcular los óptimos de una función, primero se aplica la condición necesaria con lo que obtenemos los puntos críticos que son los candidatos a óptimo. Para clasificarlos podemos utilizar la condición suficiente para óptimos locales, si son estos los que nos interesan, o la condición suficiente para óptimos globales.

Es importante observar que todos los resultados de esta sección están enunciados para funciones con dominios abiertos. En el caso en que el dominio no sea un abierto, los métodos anteriores pueden encontrar los óptimos si están en el interior del dominio pero son incapaces de estudiar los puntos de la frontera. En algunos casos, como en el ejemplo que veremos a continuación, un estudio directo de la función nos permite resolver el problema, pero en otros necesitaríamos otros métodos que están fuera del alcance de este texto.

Ejemplo *La función de costes de una empresa es*

$$C(x,y) = 2x^2 + 3y^2 + x + 2y + 75,$$

donde x, y son las cantidades producidas de los dos artículos que fabrica la empresa. Calcula la combinación de inputs que minimiza los costes.

SOLUCIÓN: Tenemos que calcular los mínimos de la función de costes $C(x,y)$, cuyo dominio económico es el conjunto cerrado $D = \{(x,y) : x \geq 0, y \geq 0\}$. Para obtener los puntos críticos planteamos el sistema

$$\nabla C(x,y) = (4x+1, 6y+2) = (0,0).$$

La solución es el punto $(-\frac{1}{4}, -\frac{1}{3})$ que, obviamente, no está en D y, por tanto, no es un punto crítico. Al no existir puntos críticos podríamos pensar que la función $C(x,y)$ no tiene mínimos. Sin embargo, si $x \geq 0$ e $y \geq 0$ entonces,

$$C(x,y) = 2x^2 + 3y^2 + x + 2y + 75 \geq 75 = C(0,0).$$

Por tanto $(0,0)$, que es un punto frontera, es un mínimo global de $C(x,y)$ y el coste mínimo se obtiene cuando no se produce nada.

■

10.3 Optimización con restricciones

Diremos que un problema es de *programación clásica con restricciones* o de *optimización condicionada* cuando es de de la forma

$$\begin{array}{ll} \text{Opt} & f(\bar{x}) \\ \text{s.a.} & g(\bar{x}) = \bar{b} \end{array}$$

donde $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$, $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$, $\bar{b} \in \mathbb{R}^m$ y $n > m$. En el resto de la sección nos referiremos a este problema como (PR) y supondremos que las restricciones son independientes, esto es, que el rango de $Jg(\bar{x})$ es m.

Para obtener métodos similares a los de la sección anterior necesitamos introducir el concepto de *función lagrangiana* de un problema restringido.

Función lagrangiana Definimos *función lagrangiana* o *lagrangiano* asociada a (PR) como la función $\mathcal{L} : D \times \mathbb{R}^m \longrightarrow \mathbb{R}$ dada por

$$\mathcal{L}(\bar{x}, \lambda) = f(\bar{x}) + \lambda(\bar{b} - g(\bar{x})),$$

es decir,

$$\mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots, \lambda_m) = f(\bar{x}) + \lambda_1(b_1 - g_1(\bar{x})) + \cdots + \lambda_m(b_m - g_m(\bar{x})).$$

Vemos que en la función lagrangiana aparecen las variables de la función f más una variable nueva λ_i por cada restricción $g_i(\bar{x}) = b_i$ que tenga el problema. Estas variables adicionales λ_i se llaman *multiplicadores de Lagrange*.

Ejemplo *Plantea la función lagrangiana del problema*

$$\begin{array}{ll} Min & x^2 + 141y^2 + 20xz \\ s.a. & x + 2y = 5 \\ & 4x^2 - z = -2 \end{array}$$

SOLUCIÓN: Como el problema tiene dos restricciones, tendrá dos multiplicadores de Lagrange y el lagrangiano será:

$$\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = x^2 + 141y^2 + 20xz + \lambda_1(5 - x - 2y) + \lambda_2(-2 - 4x^2 + z).$$

■

La condición necesaria de primer orden que utilizaremos para un problema restringido es:

Teorema *Si $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ es una función de clase C^1 en un abierto D y $\bar{x}^*$ es un extremo local de f relativo a un sistema de restricciones determinado por una función $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ de clase C^1, entonces existe un vector de multiplicadores $\bar{\lambda}^* \in \mathbb{R}^m$ tal que $(\bar{x}^*, \bar{\lambda}^*)$ es un punto crítico de la función lagrangiana $\mathcal{L}$ correspondiente al problema, es decir, $\nabla\mathcal{L}(\bar{x}^*, \bar{\lambda}^*) = \bar{0}$.*

Ejemplo *Calcula los puntos críticos del problema*

$$\begin{array}{ll} Min & x^2 + 141y^2 + 20xz \\ s.a. & x + 2y = 5 \\ & 4x^2 - z = -2 \end{array}$$

SOLUCIÓN: Hemos visto en el ejemplo anterior que su función lagrangiana era:

$$\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = x^2 + 141y^2 + 20xz + \lambda_1(5 - x - 2y) + \lambda_2(-2 - 4x^2 + z).$$

Para poder aplicar la condición necesaria calculamos el gradiente del lagrangiano y lo igualamos al vector nulo:

$$\begin{aligned} &\nabla\mathcal{L}(x, y, z, \lambda_1, \lambda_2) = \\ &\qquad = (2x + 20z - \lambda_1 - 8x\lambda_2, 282y - 2\lambda_1, 20x + \lambda_2, 5 - x - 2y, -2 - 4x^2 + z) \\ &\qquad = (0, 0, 0), \end{aligned}$$

con lo que obtenemos un sistema de cinco ecuaciones con cinco incógnitas. Si despejamos λ_2 de la tercera ecuación, y de la cuarta y z de la quinta ecuación, y sustituimos en las dos primeras, nos queda un sistema (no lineal) de dos ecuaciones con dos incógnitas. Lo resolvemos, sustituimos y obtenemos que el problema tiene dos puntos críticos que son: $(1, 2, 6)$ con multiplicadores de Lagrange asociados

$\lambda_1 = 282$, $\lambda_2 = -20$ y $(-1.3, 3.15, 8.78)$ con multiplicadores asociados $\lambda_1 = 444.3$, $\lambda_2 = 26.04$.

■

Al igual que en el caso de optimización libre, necesitamos un criterio para poder clasificar los puntos críticos.

Consideremos $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función de clase C^2 en un abierto D y $\bar{x}^*$ un punto crítico de f relativo a un sistema de restricciones determinado por una función $g : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ de clase C^1. Sea $\bar{\lambda}^*$ el vector de multiplicadores asociado. Consideramos la lagrangiana $\mathcal{L}(\bar{x}, \bar{\lambda}^*)$ como función de $\bar{x}$ únicamente, es decir, considerando $\bar{\lambda}^*$ fijo. Queremos saber qué ocurre ante un incremento marginal de $\bar{x}^*$, si L aumenta siempre, disminuye siempre, o si aumenta o disminuye según la dirección del incremento. En el primer caso $\bar{x}^*$ será un mínimo local, en el segundo un máximo local y en el tercero un punto de silla. En lugar de estudiar $\mathcal{L}(\bar{x}, \bar{\lambda}^*)$ estudiamos su polinomio de Taylor de grado 2 en $\bar{x}^*$, que será de la forma

$$p(\bar{x}) = \mathcal{L}(\bar{x}^*, \bar{\lambda}^*) + \nabla\mathcal{L}(\bar{x}^*, \bar{\lambda}^*) \cdot (\bar{x} - \bar{x}^*) + \frac{1}{2}(\bar{x} - \bar{x}^*)H_{\bar{x}}\mathcal{L}(\bar{x}^*, \bar{\lambda}^*)(\bar{x} - \bar{x}^*)^t.$$

Sabemos que $\mathcal{L}(\bar{x}^*, \bar{\lambda}^*) = C$, donde C es una constante, y $\nabla\mathcal{L}(\bar{x}^*, \bar{\lambda}^*) = \bar{0}$ porque $(\bar{x}^*, \bar{\lambda}^*)$ es un punto crítico. El polinomio de Taylor es entonces

$$p(\bar{x}) = C + 0 + \frac{1}{2}(\bar{x} - \bar{x}^*)H_{\bar{x}}\mathcal{L}(\bar{x}^*, \bar{\lambda}^*)(\bar{x} - \bar{x}^*)^t.$$

Como sólo nos interesa saber si $\mathcal{L}$ sufre variaciones positivas o negativas, podemos olvidar el factor $1/2$ del término de grado 2 y analizar la variación de

$$(\bar{x} - \bar{x}^*)H_{\bar{x}}\mathcal{L}(\bar{x}^*, \bar{\lambda}^*)(\bar{x} - \bar{x}^*)^t.$$

Si llamamos $d\bar{x} = \bar{x} - \bar{x}^*$, hemos de estudiar la forma cuadrática asociada a la matriz hessiana

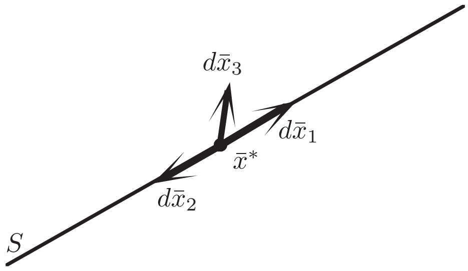
$$q(d\bar{x}) = d\bar{x}H_{\bar{x}}\mathcal{L}(\bar{x}^*, \bar{\lambda}^*)d\bar{x}^t.$$

- Si q es definida positiva, esto significa que cualquier incremento marginal $d\bar{x}$ provoca incrementos positivos en $\mathcal{L}$, luego $\bar{x}^*$ es un mínimo local relativo.

- Si q es definida negativa cualquier incremento marginal $d\bar{x}$ provoca incrementos negativos en $\mathcal{L}$, luego $\bar{x}^*$ es un es un máximo local relativo.

- Si q es indefinida todavía **no** podemos afirmar que $\bar{x}^*$ sea un punto de silla, pues en realidad *sólo debemos considerar el efecto producido por incrementos $d\bar{x}$ que no nos hagan salir del conjunto de oportunidades S*, por tanto debemos completar el estudio:

Por ejemplo, si S es la recta de la figura, el punto $\bar{x}^*$ será un mínimo local si los incrementos $d\bar{x}_1$ y $d\bar{x}_2$ producen incrementos positivos en la función objetivo, mientras que el efecto del incremento $d\bar{x}_3$ es irrelevante, ya que sólo queremos saber si $f(\bar{x}^*)$ es mayor o menor que las imágenes de los puntos de S cercanos a él.



La cuestión entonces es qué incrementos $d\bar{x}$ hemos de considerar. Tenemos que $\bar{x}^*$ es un punto de S, es decir, cumple $g(\bar{x}^*) = \bar{b}$, y queremos considerar incrementos que nos lleven a puntos $\bar{x}$ que cumplan esto mismo, es decir, que dejen constante a $\bar{g}$. En realidad, puesto que estamos considerando incrementos marginales, basta exigir que $dg = \bar{0}$.

La ecuación $dg = \bar{0}$ se traduce en la práctica en un sistema de m ecuaciones lineales (m es el número de restricciones) con n incógnitas $dx_1, \ldots, dx_n$. Las soluciones vendrán expresadas en términos de parámetros (en general $n-m$). Al sustituir estas soluciones en la forma cuadrática $q(d\bar{x})$ obtenemos otra forma cuadrática con menos variables $\tilde{q}$ que nos da el comportamiento de la función objetivo ante incrementos aceptables de las variables (es decir, incrementos compatibles con las restricciones del problema).

- Si $\tilde{q}$ es definida positiva $\bar{x}^*$ es un mínimo local relativo.
- Si $\tilde{q}$ es definida negativa $\bar{x}^*$ es un es un máximo local relativo.
- Si $\tilde{q}$ es indefinida $\bar{x}^*$ es un punto de silla.
- Si $\tilde{q}$ es semidefinida no podemos concluir nada.

Teorema *Consideramos el problema (PR) y la función lagrangiana $\mathcal{L}(x, \lambda)$. Suponemos que f y g son funciones de clase C^2. Sea (x^*, λ^*) un punto crítico de $\mathcal{L}$. Entonces:*

1. *Si $H_{\bar{x}}\mathcal{L}(x^*, \lambda^*)$ es definida positiva, el punto x^* es un mínimo local de (PR).*
2. *Si $H_{\bar{x}}\mathcal{L}(x^*, \lambda^*)$ es definida negativa, el punto x^* es un máximo local de (PR).*

Nótese que el signo de $H_{\bar{x}}\mathcal{L}(x^*, \lambda^*)$ es el de la forma cuadrática restringida.

A modo de resumen presentamos los pasos a seguir:

$$\begin{array}{ll} \text{Opt} & f(\bar{x}) \\ \text{s.a.} & g_1(\bar{x}) = \bar{b}_1, \\ & \cdots \\ & g_m(\bar{x}) = \bar{b}_m \end{array} \quad \Rightarrow \quad \begin{array}{l} n = \text{número de variables de decisión} \\ m = \text{número de restricciones} \\ r = m+1, \cdots, n \end{array}$$

1. Calculamos la función la función lagrangiana

$$\mathcal{L}(\bar{x},\lambda) = f(\bar{x}) + \lambda_1(\bar{b}_1 - g_1(\bar{x})) + \cdots + \lambda_k(\bar{b}_m - g_m(\bar{x}))$$

2. Obtenemos de puntos críticos Calculamos las soluciones de

$$\nabla\mathcal{L}(\bar{x}^*,\bar{\lambda}^*) = \bar{0} \quad \Rightarrow \quad \{\bar{x}_1^*, \bar{x}_2^*, \cdots, \bar{x}_p^*\} \quad \text{puntos críticos}$$

3. Clasificamos de los puntos críticos
   (*a*) Calculamos la matriz hessiana de $\mathcal{L}$, derivando sólo respecto de las variables originales del problema, es decir $\bar{x}$, esto es

   $$H_{\bar{x}}\mathcal{L}(\bar{x},\bar{\lambda})$$

   (*b*) Dado un punto crítico $(\bar{x},\bar{\lambda}^*)$ clasificamos la forma cuadrática asociada a la matriz $H_{\bar{x}}\mathcal{L}(\bar{x},\bar{\lambda}^*)$:

   - Si $H_{\bar{x}}\mathcal{L}(\bar{x},\bar{\lambda}^*)$ es definida positiva, en $\bar{x}$ hay un mínimo local y su multiplicador de Lagrange es $\bar{\lambda}^*$.
   - Si $H_{\bar{x}}\mathcal{L}(\bar{x},\bar{\lambda}^*)$ es definida negativa, en $\bar{x}_i$ hay un máximo local y su multiplicador de Lagrange es $\bar{\lambda}^*$.
   - Si $H_{\bar{x}}\mathcal{L}(\bar{x},\bar{\lambda}^*)$ no es definida debemos continuar: (*)

(*) En este punto tenemos dos posibilidades alternativas:

- Orlar la matriz hessiana y clasificar la forma cuadrática restringida (a este procedimiento le llamaremos MÉTODO I)

- Aplicar directamente lo que hemos hecho en la descripción teórica (a este procedimiento le llamaremos MÉTODO II)

Método I

(3.b.1.) Escribimos la matriz hessiana orlada en el punto crítico.
(3.b.2.) Aplicamos el método de clasificación de los menores orlados:

(a) Si $(-1)^m H_r > 0$ $\forall r = m+1,\ldots,n$ entonces q restringida a S es definida positiva $\Rightarrow$ en $\bar{x}^*$ hay un mínimo local.

(b) Si $(-1)^r H_r > 0$ $\forall r = m+1,\ldots,n$ entonces q restringida a S es definida negativa $\Rightarrow$ en $\bar{x}^*$ hay un máximo local.

Método II

(3.b.1.) Escribimos la forma cuadrática asociada a la matriz hessiana $H_{\bar{x}}\mathfrak{L}(\bar{x}^*, \bar{\lambda}^*)$, es decir

$$q(dx) = d\bar{x} H_{\bar{x}}\mathfrak{L}(\bar{x}, \bar{\lambda}^*)(d\bar{x})^t$$

(3.b.2.) Exigimos que las variaciones queden dentro del conjunto de oportunidades, es decir

$$\left.\begin{array}{ll} dg_1(\bar{x}^*, \bar{\lambda}^*) & = 0 \\ \dots & \\ dg_m(\bar{x}^*, \bar{\lambda}^*) & = 0 \end{array}\right\}$$

(3.b.3.) Sustituimos las soluciones del sistema dado en (3.b.2) en q y obtenemos una nueva forma cuadrática restringida $\tilde{q}$.

(3.b.4.) Clasificamos la forma cuadrática $\tilde{q}(\bar{x}^*, \bar{\lambda}^*)$

- Si es **definida positiva**, en $\bar{x}^*$ hay un mínimo local
- Si es **definida negativa**, en $\bar{x}^*$ hay un máximo local
- Si es **indefinida**, en $\bar{x}^*$ hay un punto de silla
- Si es **semidefinida** no podemos clasificarlo

Ejemplo *Resuelve el problema*

$$\begin{array}{ll} Max & -x^2 - 3y^2 + z^2 + 2xy + 2yz + 100 \\ s.a. & y - x = -2, \\ & x + z = 4. \end{array}$$

SOLUCIÓN: **1.** La función lagrangiana asociada al problema es:

$$\begin{array}{rcl} \mathfrak{L}(x, y, z, \lambda_1, \lambda_2) & = & -x^2 - 3y^2 + z^2 + 2xy + 2yz + 100 \\ & & +\lambda_1(-2 - y + x) + \lambda_2(4 - x - z). \end{array}$$

2. Para calcular los puntos críticos calculamos las derivadas parciales del lagrangiano:

$$\frac{\partial \mathfrak{L}}{\partial x} = -2x + 2y + \lambda_1 - \lambda_2, \quad \frac{\partial \mathfrak{L}}{\partial y} = -6y + 2x + 2z - \lambda_1,$$

$$\frac{\partial \mathfrak{L}}{\partial z} = 2z + 2y - \lambda_2, \quad \frac{\partial \mathfrak{L}}{\partial \lambda_1} = -2 - y + x,$$

$$\frac{\partial \mathfrak{L}}{\partial \lambda_2} = 4 - x - z,$$

y resolvemos el sistema:

$$\left.\begin{array}{rrrrrcr} -2x & +2y & & \lambda_1 & -\lambda_2 & = & 0 \\ 2x & -6y & +2z & -\lambda_1 & & = & 0 \\ & 2y & +2z & & -\lambda_2 & = & 0 \\ -x & +y & & & & = & -2 \\ x & & +z & & & = & 4 \end{array}\right\}$$

El único punto crítico es $(\bar{x}^*, \bar{\lambda}^*) = (x, y, z, \lambda_1, \lambda_2) = (2, 0, 2, 8, 4)$.

3. Clasificamos el punto $(2, 0, 2, 8, 4)$:

3.a) La matriz hessiana respecto las variables (x, y, z) es

$$H_{(x,y,z)}\mathcal{L}(\bar{x}, \bar{\lambda}) = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -6 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

3.b) La matriz hesiana en el punto crítico es

$$H_{(x,y,z)}\mathcal{L}(2, 0, 2, 8, 4) = \begin{pmatrix} -2 & 2 & 0 \\ 2 & -6 & 2 \\ 0 & 2 & 2 \end{pmatrix}$$

cuyos menores principales conducentes son

$$|A_1| = -2 < 0, \quad |A_2| = 8 > 0, \quad |A_3| = 8 > 0.$$

Por tanto, la forma cuadrática asociada es indefinida y de momento no podemos llegar a ninguna conclusión.

Clasificación con el Método I:

3.b.1) La matriz hessiana orlada es

$$H\mathcal{L}(0, 0, 100, 200, 200) = \left(\begin{array}{ccc|cc} -2 & 2 & 0 & 1 & -1 \\ 2 & -6 & 2 & -1 & 0 \\ 0 & 2 & 2 & 0 & -1 \\ \hline 1 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \end{array}\right)$$

3.b.2) Como n=3 y m=1, entonces r=3. Calculamos $H_3 = -6$ y se tiene que $(-1)^r H_r > 0$, es decir $(-1)^3 H_3 = 6 > 0$. Por tanto la hessiana restringida es definida negativa y el punto $(x, y, z) = (2, 0, 2)$ es un máximo local.

Clasificación con el Método II:

3.b.1) Calculamos la forma cuadrática asociada a la matriz hessiana:

$$q(dx, dy, dz) = -2dx^2 - 6dy^2 + 2dz^2 + 4dxdy + 4dydz$$

3.b.2) Exigimos que $dg_1(\bar{x}^*, \bar{\lambda}^*) = 0, \quad dg_2(\bar{x}^*, \bar{\lambda}^*) = 0$:

$$\left.\begin{array}{rrrcl} -dx & +dy & & = & 0 \\ & dx & +dz & = & 0 \end{array}\right\}$$

la solución es $dy = dx, \quad dz = -dx$.

3.b.3) Sustituyendo los resultados anteriores, la forma cuadrática restringida es:

$$\tilde{q}(dx) = -6dx^2.$$

3.b.4) Clasificamos la forma cuadrática $\tilde{q}(\bar{x}^*, \bar{\lambda}^*)$:

La matriz de $\tilde{q}$ es $A = [-6]$ que claramente es definida negativa. Por tanto, en $(2, 0, 2)$ hay un máximo local. ■

Del mismo modo que en programación clásica no restringida, la condición anterior nos proporciona una herramienta para determinar qué puntos críticos son óptimos locales, pero no asegura su globalidad. Vamos a introducir una condición suficiente de segundo orden para óptimos globales:

Teorema *Consideremos el problema (PR) y supongamos que D es un abierto convexo y $\bar{x}*$ es un punto crítico del problema. Entonces:*

1. *Si $f(\bar{x})$ es una función convexa y todas las restricciones son lineales, $\bar{x}^*$ es un mínimo global relativo de f. Además, si la función es estrictamente convexa el mínimo es estricto.*

2. *Si $f(\bar{x})$ es una función cóncava y todas las restricciones son lineales, $\bar{x}^*$ es un máximo global relativo de f. Además, si la función es estrictamente cóncava el máximo es estricto.*

Ejemplo *Resuelve el problema*

$$\begin{array}{ll} Max & -x^2 - 3y^2 \\ s.a. & y - x = 2 \end{array}$$

SOLUCIÓN: Al tratarse de polinomios, las funciones f y g están definidas sobre $\mathbb{R}^2$ que es un abierto convexo.

Calculemos los puntos críticos. La función lagrangiana del problema es:

$$\mathcal{L}(x, y\lambda) \quad = \quad -x^2 - 3y^2 + \lambda\,(2 - y + x).$$

Planteamos el gradiente del lagrangiano, lo igualamos al vector nulo y obtenemos el sistema:

$$\left.\begin{array}{rrrcr} -2x & & +\lambda & = & 0 \\ & -6y & -\lambda & = & 0 \\ -x & +y & & = & 2 \end{array}\right\}$$

cuya solución es $(-\frac{3}{2}, \frac{1}{2}, -3)$. Al tratarse del único punto crítico, es el único candidato a máximo que tenemos. Vamos a aplicar el teorema para saber si realmente es un máximo global. Es obvio que todas las restricciones (en este caso, sólo hay una) son lineales. Estudiemos la convexidad de la función objetivo $f(x, y) = -x^2 - 3y^2$. La matriz hessiana de f es:

$$Hf(x, y) = \begin{pmatrix} -2 & 0 \\ 0 & -6 \end{pmatrix}$$

Puesto que se trata de una matriz diagonal y todos los elementos de la diagonal principal son negativos, la forma cuadrática asociada a $Hf(x,y)$ es definida negativa y, en consecuencia, f es estrictamente cóncava. Por tanto, el punto crítico $(-\frac{3}{2}, \frac{1}{2})$ es una máximo global relativo que, además, es único. ∎

10.4 Interpretación de los multiplicadores de Lagrange

Los multiplicadores de Lagrange aportan una información valiosa sobre el problema considerado. Si $\bar{x}^*$ es un óptimo local de un problema con restricciones

$$\begin{array}{rl} \text{Opt} & f(\bar{x}) \\ \text{s.a.} & g(\bar{x}) = \bar{b}, \end{array}$$

y sus multiplicadores asociados son $\bar{\lambda}^* = (\lambda_1^*, \ldots, \lambda_m^*)$, una pequeña modificación de uno de los términos independientes b_i de las restricciones da lugar a un nuevo problema con una solución ligeramente distinta con un valor ligeramente distinto de la función objetivo. Si llamamos $f(b_i)$ al valor de la función objetivo óptima cuando el término independiente i-ésimo toma el valor b_i, se puede probar que

$$\lambda_i^* = \frac{\partial f}{\partial b_i},$$

y de acuerdo con la interpretación de la derivada parcial,

> El multiplicador de Lagrange λ_i^* indica aproximadamente lo que se modificará la función objetivo óptima por cada unidad marginal que variemos el término independiente b_i de la restricción i-ésima.

Ejemplo *Consideremos una empresa que fabrica tres artículos A, B y C en cantidades x, y, z respectivamente. La empresa fija los precios de sus artículos según unas funciones decrecientes en la cantidad producida del siguiente modo: un artículo A vale* $200 - 4x$ *unidades monetarias, un artículo B vale* $200 - 3y$ *u.m. y, por último, el precio de un artículo C es* $100 - z$ *u.m. Además, la empresa ha calculado empíricamente que su coste en función de las cantidades producidas puede aproximarse por la función* $x^2 + 2y^2 + z^2 + 100z + 100$. *En la actualidad, el nivel de producción total es de 59 unidades, pero la empresa considera que puede aumentarlo en una unidad sin incumplir sus restricciones técnicas. Calcula los precios óptimos y el beneficio óptimo actual y razona si a la empresa le conviene aumentar su producción.*

Solución: Veamos que este problema se puede plantear como un problema de programación clásica con restricciones. Para ello, pensemos que el objetivo de la empresa consiste en maximizar sus beneficios y supongamos que la empresa vende toda su producción. En este caso, y en ausencia de otras fuentes de financiación, la función de ingresos de la empresa vendrá dada por:

$$I(x,y,z) = (200-4x)x+(200-3y)y+(100-z)z = -4x^2-3y^2-z^2+200x+200y+100z,$$

mientras que la función de costes nos la proporciona el enunciado y es:

$$C(x, y, z) = x^2 + 2y^2 + z^2 + 100z + 100$$

Con todo ello, la función de beneficios podemos calcularla como:

$$B(x, y, z) = I(x, y, z) - C(x, y, z) = -5x^2 - 5y^2 - 2z^2 + 200x + 200y - 100$$

Además, sabemos que el nivel de producción total es de 59 unidades, es decir, debe cumplirse la restricción $x + y + z = 59$.

Por tanto, con la información de la que disponemos podemos decir que el problema que tiene que resolver la empresa es:

$$\begin{array}{ll} \text{Max} & -5x^2 - 5y^2 - 2z^2 + 200x + 200y - 100 \\ \text{s.a.} & x + y + z = 59 \end{array}$$

Con los métodos estudiados en epígrafes anteriores, es fácil calcular que el nivel de producción óptimo viene dado por $(x, y, z) = (24, 24, 10)$ puesto que $(24, 24, 10)$ con multiplicador asociado $\lambda = -40$ es el único máximo global relativo de B. Sustituyendo vemos que los precios óptimos de la empresa son $p_A = 200 - 4 \times 24 = 104\,\text{u.m.}$, $p_B = 200 - 3 \times 24 = 128\,\text{u.m.}$ y $p_C = 100 - 10 = 90\,\text{u.m.}$ Su beneficio máximo es $B(24, 24, 10) = 3540\,\text{u.m.}$

Si el nivel de producción total aumenta una unidad marginal, el término independiente b de la restricción aumenta en una unidad, y como

$$\frac{dB^*}{db} = \lambda = -40,$$

resulta que el beneficio máximo B^* de la empresa disminuiría aproximadamente en 40 u.m. Por tanto a la empresa no le conviene aumentar la producción. ■

10.5 Algunas aplicaciones

Son muchos los problemas económicos que se pueden plantear como un problema de programación matemática, como hemos visto en el ejemplo anterior. En esta sección veremos algunas aplicaciones. Siempre hemos de tener en cuenta que en las aplicaciones de la programación matemática a la economía nos interesa, normalmente, encontrar óptimos globales, aunque a veces es mejor tener un óptimo local que ninguna información en absoluto. Pensemos, por ejemplo, que si $\bar{x}^*$ es un máximo local de una función de beneficios esto no quiere decir más que alrededor de $\bar{x}^*$ no existe otro punto donde los beneficios sean mayores. Pero podría ser que en el dominio de la función hubiese otra combinación de inputs que proporcionase mayores beneficios. En tal caso, la solución $\bar{x}^*$ será de poca utilidad. Sin embargo, ante la ausencia de cualquier tipo de información, un óptimo local podría guiarnos en la toma de decisiones.

Ejemplo *Una empresa dispone de una almacen desde el que distribuye su producto a dos zonas comerciales diferentes. La empresa ha calculado que sus costes variables de transporte vienen dados por la función* $3x_1^2+2x_2^2+4x_1x_2-10x_1-8x_2$*, donde* x_i *representa la cantidad del producto, medida en miles de unidades, enviada a la zona* i*. Además, la empresa tiene que hacer frente a unos costes fijos de infraestructura de 14 u.m. Calcula cuantas unidades del artículo enviará la empresa a cada zona para que sus costes sean mínimos.*

SOLUCIÓN: El problema es minimizar los costes totales de transporte de la empresa que vienen dados por la suma de los costes variables y los costes fijos. Esto es, se trata de resolver el problema:

$$\text{Min } C(x_1, x_2) = 3x_1^2 + 2x_2^2 + 4x_1x_2 - 10x_1 - 8x_2 + 14$$

Para ello calculamos los puntos críticos del problema, igualando el gradiente de la función de costes al vector nulo y resolviendo el sistema así obtenido:

$$\nabla C(x_1, x_2) = (6x_1 + 4x_2 - 10, 4x_2 + 4x_1 - 8) = (0, 0).$$

El único punto crítico es $(1, 1)$. Vamos a clasificarlo:

$$HC(x_1, x_2) = \begin{pmatrix} 6 & 4 \\ 4 & 4 \end{pmatrix}.$$

Los menores conduentes son $A_1 = 6$ y $A_2 = 8$, con lo que la forma cuadrática asociada a $HC(x_1, x_2)$ es definida positiva y, por tanto, C es una función estrictamente convexa. En consecuencia, $(1, 1)$ es el único máximo global de la función y la respuesta al problema es que la empresa debe enviar 1000 unidades de producto a la primera zona comercial y 1000 unidades de producto a la segunda. ■

Ejemplo *Un empresario produce dos artículos en cantidades* x*,* y *respectivamente. La empresa tiene un coste fijo de 20 € y sus costes variables unitarios vienen dados por* x *€ para el primer artículo y* $x+2y$ *€ para el segundo. En la actualidad, la empresa tiene una producción total de 100 unidades. Calcula el coste mínimo para el nivel actual de producción y razona si al empresario le interesaría, caso de ser posible, aumentar o disminuir el nivel de producción.*

SOLUCIÓN: El problema es

$$\begin{array}{ll} \text{Min} & C(x, y) \\ \text{s.a.} & Q(x, y) = 100 \end{array}$$

Para calcular el coste sabemos que el coste variable por unidad del primer artículo es x. Si se producen x unidades, el coste variable total del primer artículo será $x \cdot x = x^2$. Del mismo modo, el coste variable total para el segundo artículo es $(x + 2y)y = xy + 2y^2$. Si añadimos el coste fijo, obtenemos el coste total

$$C(x, y) = x^2 + 2y^2 + xy + 20 \text{ €}.$$

Por otro lado, $Q(x,y) = x + y$ unidades. Por tanto debemos resolver

$$\begin{array}{ll} \text{Min} & x^2 + 2y^2 + xy + 20 \\ \text{s.a.} & x + y = 100 \end{array}$$

La función lagrangiana es

$$\mathcal{L}(x,y,\lambda) = x^2 + 2y^2 + xy + 20 + \lambda(100 - x - y).$$

Calculamos los puntos críticos:

$$\left.\begin{array}{rrrcl} 2x & +y & -\lambda & = & 0 \\ x & +4y & -\lambda & = & 0 \\ x & +y & & = & 100 \end{array}\right\}$$

Resolviendo el sistema vemos que el único punto crítico del problema es $(75,25)$ con multiplicador asociado $\lambda = 175$.

Puesto que la restricción del problema es lineal, vamos a estudiar la convexidad de la función objetivo $C(x,y) = x^2 + 2y^2 + xy + 20$.

$$HC(x,y) = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}.$$

Sus menores principales conducentes valen $A_1 = 2$ y $A_2 = 7$, con lo que la forma cuadrática asociada a $HC(x,y)$ es definida positiva y C es una función estrictamente convexa. Por tanto, $(75,25)$ es un mínimo global estricto relativo del problema y el coste mínimo es de 8770 €.

Si el empresario puede cambiar el nivel de producción actual, es decir, modificar el termino independente b de la restricción, el comportamiento local de los costes vendrá dado por el multiplicador λ:

$$\frac{\partial C^*}{\partial b} = \lambda = 175.$$

El signo positivo indica que por cada unidad marginal que el empresario aumente (disminuya) el nivel de producción, el coste aumentará (disminuirá) 175 unidades marginales. Luego al empresario le convendría disminuir el nivel de producción total. ∎

Ejemplo *Una empresa fabrica dos artículos en cantidades x, y. Su función de costes viene dada por $C(x,y) = x^2 + 2y^2 + xy + 20$€. Calcula el máximo nivel de producción de la empresa sabiendo que el coste total son 8770€. Razona qué ocurrirá con la producción si la empresa decide tener un coste total de 8769 €.*

Solución: El problema es

$$\begin{array}{ll} \text{Max} & Q(x,y) \\ \text{s.a.} & C(x,y) = 8770. \end{array}$$

Por tanto, debemos resolver

$$\begin{array}{ll} \text{Max} & x + y \\ \text{s.a.} & x^2 + 2y^2 + xy = 8750. \end{array}$$

La función lagrangiana es

$$\mathcal{L}(x, y, \lambda) = x + y + \lambda(8750 - x^2 - 2y^2 - xy).$$

Los puntos críticos son las soluciones del sistema

$$\left.\begin{array}{rcl} 1 - 2x\lambda - y\lambda & = & 0 \\ 1 - 4y\lambda - x\lambda & = & 0 \\ x^2 + 2y^2 + xy & = & 8750 \end{array}\right\}$$

Despejando λ de las dos primeras ecuaciones e igualando las expresiones resultantes, obtenemos $x = 3y$. Sustituyendo en la tercera ecuación, $14y^2 = 8750$, con lo que $y = 25$ o $y = -25$. Rechazamos este último resultado por carecer de interpretación económica y sustituimos en los resultados anteriores, con lo que al final obtenemos un punto crítico para este problema que es $(75, 25)$ con multiplicador asociado $\lambda = \frac{1}{175}$.

Puesto que la restricción no es lineal, no podemos clasificar el punto crítico mediante el teorema para óptimos globales. Vamos a utilizar, pues, la condición suficiente de segundo orden para óptimos locales. Tenemos que la matriz hessiana respecto de las variables x, y es:

$$H_{(x,y)}\mathcal{L}(x, y, \lambda) = \begin{pmatrix} -2\lambda & -\lambda \\ -\lambda & -4\lambda \end{pmatrix}$$

que calculada en el punto crítico es:

$$H_{(x,y)}\mathcal{L}(75, 25, \frac{1}{175}) = \begin{pmatrix} -\frac{2}{175} & -\frac{1}{175} \\ & \\ -\frac{1}{175} & -\frac{4}{175} \end{pmatrix}$$

Los menores conducentes valen:

$$A_1 = -\frac{1}{175} < 0$$

y

$$A_2 = \frac{1}{4375} > 0$$

con lo que la forma cuadrática asociada es definida negativa y el punto $(75, 25)$ con multiplicador asociado $\lambda = \frac{1}{175}$ es un máximo local relativo del problema.

Así pues, el máximo nivel de producción que puede alcanzar la empresa es $Q(75, 25) = 100$ unidades de producto, aunque hemos de tomar este resultado con precaución puesto que el óptimo calculado es local y, por tanto, no podemos asegurar que no existan otras combinaciones de inputs para las cuales la producción sea mayor con el mismo coste.

Si la empresa decide reducir su coste total en una unidad, esto es, pasar de 8770 a 8769 €, el término independiente del problema se reduce en una unidad que podemos considerar marginal. Por la interpretación del multiplicador de Lagrange, sabemos que

$$\frac{\partial Q^*}{\partial b} = \lambda = \frac{1}{175}.$$

El signo positivo indica que por cada unidad marginal que el empresario disminuya el coste, el nivel de producción disminuirá $\frac{1}{175}$ unidades marginales. ■

Ejemplo *La función de utilidad de un consumidor es* $U(x, y) = \ln(1+xy)$*, donde* x*,* y *son la unidades consumidas de los bienes* A *y* B*, respectivamente, cuyos precios son ambos de 1 € por unidad. El consumidor dispone de una renta de* 4 €*. Calcula las cantidades de ambos bienes que maximizan la utilidad suponiendo que el consumidor gasta toda la renta. Interpreta el multiplicador de Lagrange.*

SOLUCIÓN: Con los datos que tenemos podemos plantear la restricción presupuestaria $x + y = 4$. El problema es

$$\begin{array}{ll} \text{Max} & \ln(1 + xy) \\ \text{s.a.} & x + y = 4 \end{array}$$

La función lagrangiana es

$$\mathcal{L}(x, y, \lambda) = \ln(1 + xy) + \lambda(4 - x - y).$$

Calculamos los puntos críticos:

$$\left.\begin{array}{ccccc} \dfrac{\partial \mathcal{L}}{\partial x} & = & \dfrac{y}{1 + xy} - \lambda & = & 0 \\[2ex] \dfrac{\partial \mathcal{L}}{\partial y} & = & \dfrac{x}{1 + xy} - \lambda & = & 0 \\[2ex] \dfrac{\partial \mathcal{L}}{\partial \lambda} & = & 4 - x - y & = & 0 \end{array}\right\}$$

Despejando λ de las dos primeras ecuaciones e igualando obtenemos

$$y + xy^2 = x + x^2y.$$

Despejando y de la tercera ecuación del sistema y sustituyendo en la anterior, tenemos la ecuación de tercer grado en x:

$$x^3 - 6x^2 + 7x + 2 = 0$$

cuyas soluciones son $x = 2$, $x = 2 + \sqrt{5}$ y $x = 2 - \sqrt{5}$. Rechazamos este último resultado, por no tener sentido económico, y calculamos, para $x = 2$, $y = 4-x = 2$; para $x = 2+\sqrt{5}$, $y = 4-x = 2-\sqrt{5}$. Este último resultado tampoco tiene sentido y, por tanto no lo consideramos.

Con todo lo anterior, el problema que estamos intentando resolver tiene un punto crítico que es $(2, 2)$ con multiplicador asociado $\lambda = \frac{2}{5}$.

Puesto que la restricción es lineal, vamos a ver si podemos estudiar la convexidad de la función objetivo y, así, aplicarle la condición suficiente de segundo orden para óptimos globales. La matriz hessiana de la función U es:

$$HU(x, y) = \begin{pmatrix} -\frac{y^2}{(1+xy)^2} & \frac{1}{(1+xy)^2} \\[2ex] \frac{1}{(1+xy)^2} & -\frac{x^2}{(1+xy)^2} \end{pmatrix}.$$

Sus menores conducentes valen:

$$A_1 = -\frac{y^2}{(1+xy)^2} < 0$$

y

$$A_2 = \frac{x^2y^2 - 1}{(1+xy)^4},$$

cuyo signo puede variar dependiendo de x, y. Así pues, no podemos concluir que existe un máximo global. Empleemos la condición suficiente de segundo orden para óptimos locales. Primero estudiaremos la matriz $H_{(x,y)}\mathcal{L}(x, y, \lambda)$ que en este caso coincide con $HU(x, y)$. Calculada en el punto crítico tenemos que:

$$H_{(x,y)}\mathcal{L}(2, 2, \frac{2}{5}) = \begin{pmatrix} -\frac{4}{25} & \frac{1}{25} \\ \frac{1}{25} & -\frac{4}{25} \end{pmatrix}.$$

Sus menores principales conducentes son

$$A_1 = -\frac{4}{25} < 0 \text{ y } A_2 = \frac{3}{125} > 0$$

con lo que la forma cuadrática asociada es definida negativa y, por tanto, el punto $(2, 2)$ con multiplicador asociado $\lambda = \frac{2}{5}$ es un máximo local.

En conclusión, si el consumidor desea maximizar su utilidad utilizando toda su renta, debe consumir 2 unidades del bien A y 2 unidades del bien B.

El multiplicador de Lagrange se interpreta como:

$$\frac{\partial U^*}{\partial b} = \lambda = \frac{2}{5}.$$

Puesto que el termino independiente representa la renta total del consumidor, el significado de λ es que cuando su renta aumenta (disminuye) una unidad marginal, la utilidad aumenta (disminuye) 0.4 unidades marginales.

10.6 Ejercicios

1. Razona la respuesta a estas cuestiones:

   (a) ¿Una función puede tener más de un máximo global estricto? ¿Puede no tener ninguno? ¿Y máximos globales no estrictos?

   (b) ¿Un máximo global de una función es necesariamente un máximo local?

   (c) ¿Un mínimo local de una función es necesariamente un mínimo global?

   (d) ¿Un mínimo local estricto de una función es necesariamente un mínimo local no estricto? ¿Y al revés?

   (e) ¿Un máximo local absoluto de una función es necesariamente un máximo local relativo a un conjunto de restricciones S? ¿Y al revés?

   (f) ¿Un punto de silla de una función puede ser un máximo relativo?

(g) ¿Un máximo global de una función que satisfaga un sistema de restricciones puede ser un máximo relativo a las mismas? ¿Puede no serlo? ¿Y si es local?

(h) ¿Un mínimo global relativo de una función puede ser un mínimo global absoluto? ¿Puede no serlo?

(i) ¿Un máximo global de una función puede ser también un mínimo global? ¿Y si es estricto?

(j) Si una función tiene un máximo global absoluto y un máximo global relativo a un sistema de restricciones, ¿en cuál de los dos será mayor el valor de la función objetivo? ¿Puede valer lo mismo? Si la respuesta es afirmativa, ¿pueden no darse en el mismo punto?

2. ¿Tiene solución local el problema Min $x^2 - y^2$? ¿Y si lo restringimos a $2x - y = 3$? Cuando la respuesta sea afirmativa, ¿es global la solución?

3. Calcula los extremos locales de las siguientes funciones:

   (a) $f(x, y) = e^{8x^2+3y^2-4xy-8x}$,

   (b) $f(x, y, z) = 2x^2 + y^2 + 3z^2 + xy + yz - 35x - y - z$,

   (c) $f(x, y) = x^2 + xy^3 + x$,

   (d) $f(x, y) = y^2 + 4x^2y$ sujeta a la restricción $xy = 2$,

   (e) $f(x, y, z) = \ln(xyz)$ sujeta a la restricción $x + y + z^2 = 5$,

   (f) $f(x, y, z) = -3x^2 + 2y^2 - z^2 + 2xy - 4xz + 70$ sujeta a las restricciones $x + y = -5$, $2x + z = 6$.

4. Calcula los extremos globales de las funciones del ejercicio anterior.

5. Una empresa exporta un producto a tres países en cantidades x, y, z, respectivamente. La empresa tiene unos costes variables de transporte de $x - y$ unidades monetarias por cada unidad del producto enviada al primer país, $3y - x - 2z$ u.m. por cada unidad enviada al segundo país y $4z - 2y$ u.m. por cada unidad transportada hasta el tercero. Además, la empresa ha calculado que sus costes fijos son de 800 u.m. Su cuota de exportación es de 1500 unidades en total.

   (a) Calcula las cantidades que se exportarán a cada país si el objetivo de la empresa es minimizar sus costes totales de transporte.

   (b) Razona el efecto que produciría una pequeña disminución de la cuota de exportación.

6. Un inversor desea comprar dos activos cuyos rendimientos son del 10% y el 6% respectivamente. Si las cantidades invertidas son x e y, el inversor ha aproximado el riesgo por la siguiente forma cuadrática:

$$R(x, y) = 0.025x^2 + 0.005y^2 + 0.04xy \,€^2.$$

   (a) Si el inversor está dispuesto a asumir un riesgo de 1331 €^{2}, ¿qué cantidad debe invertir en cada activo para maximizar el rendimiento total?

(b) Si el inversor quiere obtener una rentabilidad total de 22 €, ¿qué cantidad debe invertir en cada activo para minimizar el riesgo?

(c) Analiza la relación entre los problemas de los dos apartado anteriores.

11. La integral definida

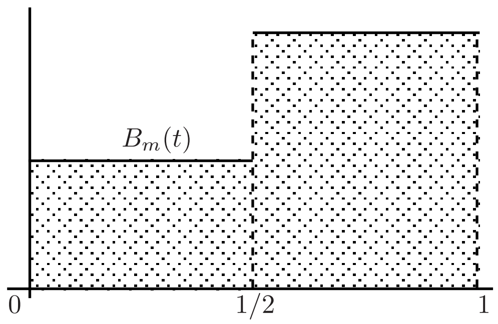
11.1 La integral de Riemann

El concepto de integral definida Supongamos que una empresa ha tenido a lo largo de un año unos beneficios marginales constantes de 200 u.m./año. Así, durante cada mes (1/12 de año) ganó 200/12 = 16.6 u.m., lo cual tiene sentido. También podemos decir que cada segundo estuvo ganando 0.00000038 u.m., lo cual no tiene sentido económico, pero esto no importa, pues las hipótesis serán aceptables si, efectivamente, cada mes la empresa incrementó sus beneficios en 16.6 u.m. Así, si una empresa tiene unos beneficios marginales constantes de 200 u.m./año, podemos decir que al cabo de un año sus beneficios acumulados pasan a ser de 200 u.m.

Supongamos ahora que los beneficios marginales no han sido constantes, sino que fueron de 100 u.m./año el primer semestre y de 200 u.m/año el segundo semestre. Más concretamente:

$$B_m(t) = \begin{cases} 100 & \text{si } 0 \leq t < 1/2, \\ 200 & \text{si } 1/2 \leq t \leq 1, \end{cases}$$

donde t es el tiempo expresado en años.



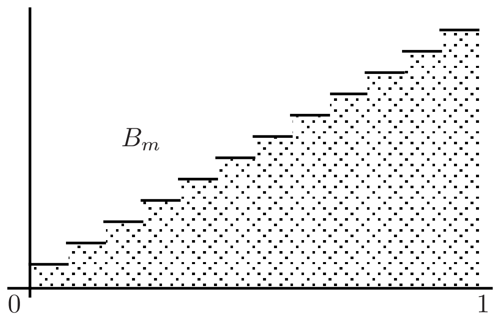
En este caso, los beneficios acumulados son

$$B = 100 \cdot \frac{1}{2} + 200 \cdot \frac{1}{2} = 150 \text{ u.m.}$$

Conviene observar que los beneficios coinciden con el área que queda por debajo de la gráfica de la función de beneficios marginales (la zona sombreada en la figura).

Supongamos ahora que los beneficios marginales no han sido constantes durante cada semestre, sino que han variado cada mes. Digamos que han sido de 20 u.m./año en enero, de 40 u.m./año en febrero, etc. Para expresar matemáticamente este caso, o bien distinguimos doce casos en la definición de B_m, o bien usamos la función parte entera $E(x)$:

$$B_m(t) = 20(E(12t) + 1).$$



Para calcular el beneficio acumulado hemos de descomponer el año en los doce periodos (los doce meses) en los que el beneficio marginal es constante. En enero el beneficio acumulado es de 20/12 u.m., en febrero acumulamos 40/12 u.m. que hay que sumar a las de enero, y así sucesivamente, con lo que el beneficio acumulado al terminar diciembre resulta ser

$$B = \frac{20}{12} + \frac{40}{12} + \cdots + \frac{240}{12} = 130 \text{ u.m.}$$

El resultado es también el área que deja bajo su gráfica la función de beneficio marginal.

La expresión del beneficio marginal es complicada porque matemáticamente es complicado tratar con saltos bruscos, y aquí estamos suponiendo que el beneficio marginal aumenta a saltos. Matemáticamente es más fácil trabajar si suponemos que el beneficio marginal varía de forma continua con el tiempo, lo cual es tan artificial como toda variación continua, pero a menudo un modelo matemático con este tipo de hipótesis se ajusta bien a la realidad en sus predicciones y es mucho más fácil de tratar.

Supongamos ahora que el beneficio marginal de la empresa viene dado por la función

$$B_m(t) = 240t.$$

Esto significa que, por ejemplo, a finales de febrero la empresa tenía una tasa de beneficios de $B_m(2/12) = 40$ u.m./año, igual que en el caso anterior, pero la diferencia es que esto no fue constante durante todo el mes de febrero, sino que a principio de mes sus beneficios eran de 20 u.m./año y fueron ascendiendo gradualmente hasta llegar a las 40 u.m./año a fin de mes. Para calcular el beneficio acumulado en este caso necesitaremos el cálculo integral, por lo que conviene empezar a introducir algunos de los conceptos en los que se basa.

Definición Diremos que una función $f : [a, b] \longrightarrow \mathbb{R}$ está *acotada* en el intervalo $[a, b]$ si existen números reales $m \leq M$ tales que para todo $x \in [a, b]$ se cumple $m \leq f(x) \leq M$. En tal caso se dice que m es una *cota inferior* de f y que M es una *cota superior.*

Una *partición* de un intervalo $[a, b]$ es un conjunto finito P de puntos del intervalo ordenados de forma creciente:

$$P = \{a = x_0 < x_1 < \cdots < x_n = b\}.$$

Para una partición dada P, llamaremos $\Delta x_i = x_i - x_{i-1}$. Llamaremos *norma* de la partición P al máximo de estas longitudes Δx_i. Se representa por $\|P\|$.

Definimos

$$m_i(f) = \inf_{x \in [x_{i-1}, x_i]} f(x), \qquad M_i(f) = \sup_{x \in [x_{i-1}, x_i]} f(x).$$

Aunque la definición del ínfimo y el supremo de una función en un conjunto presenta ciertas sutilezas, es bastante aproximado decir que $m_i(f)$ es el menor valor que toma f en el intervalo $[x_{i-1}, x_i]$ y $M_i(f)$ es el máximo valor que toma en dicho intervalo.

En nuestro ejemplo estamos estudiando la función $f = B_m$, definida en el intervalo $[0, 1]$ en el que, ciertamente, está acotada. Una cota inferior es $m = 0$ y una cota superior es $M = 240$. Para calcular el beneficio acumulado consideramos una partición del intervalo, por ejemplo la partición en meses:

$$P = \{0 < \frac{1}{12} < \frac{2}{12} < \frac{3}{12} < \frac{4}{12} < \frac{5}{12} < \frac{6}{12} < \frac{7}{12} < \frac{8}{12} < \frac{9}{12} < \frac{10}{12} < \frac{11}{12} < 1\}$$

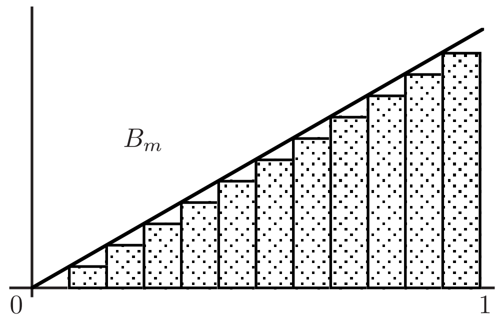
Es decir, $t_0 = 0$, $t_1 = 1/12$, $t_2 = 2/12$, etc. Los incrementos son todos iguales $\Delta t_i = 1/12$ y la norma de P es, por lo tanto, $\|P\| = 1/12$.

En el intervalo $[t_0, t_1]$, la función B_m pasa de 0 a 20, luego tenemos que $m_1(B_m) = 0$ y $M_1(B_m) = 20$. Similarmente $m_2(B_m) = 20$ y $M_2(B_m) = 40$, etc.

No sabemos calcular el beneficio acumulado cuando el beneficio marginal es B_m, pero sí que podemos resolver un problema más simple: ¿Cuál sería el beneficio acumulado si en cada intervalo $[t_{i-1}, t_i]$ (o sea, en cada mes) el beneficio marginal se mantuviera constante igual a $m_i(B_m)$? Entonces el beneficio en enero sería nulo (pues $m_1(B_m) = 0$), el de febrero sería $20/12$, pues $m_2(B_m) = 20$, etc. En total tendríamos

$$s(B_m, P) = \sum_{i=1}^{12} m_i(B_m) \frac{1}{12} = 110 \text{ u.m.}$$

Esto no es el beneficio acumulado realmente, pues hemos redondeado hacia abajo el beneficio marginal en cada intervalo. Por ejemplo, en enero hemos considerado que no se acumula beneficio porque el primer segundo de enero era así, pero a lo largo del mes el beneficio marginal crece y eso no lo hemos tenido en cuenta. Similarmente, hemos considerado que el beneficio marginal de febrero era de 20 u.m./año, cuando eso sólo es así el primer segundo y no hemos tenido en cuenta este crecimiento.

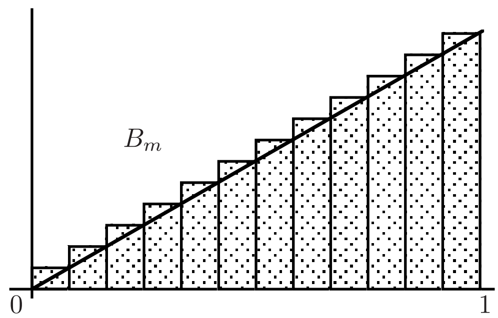


La figura muestra la diferencia entre lo que hemos calculado y lo que queremos calcular. Hemos calculado la región sombreada y lo que queremos calcular es el área que deja la función B_m bajo su gráfica. Encima de cada columna de área que hemos contado hay un triángulo que corresponde a los beneficios acumulados como consecuencia del aumento de B_m que no hemos considerado. Tenemos, pues, que la suma que hemos calculado queda por debajo del beneficio acumulado B que queremos calcular:

$$110 = s(B_m, P) \leq B.$$

Por otra parte, también sabemos calcular el beneficio acumulado si suponemos que el beneficio marginal en cada mes $[t_{i-1}, t_i]$ permanece constantemente igual a $M_i(B_m)$. En tal caso sería

$$S(B_m, P) = \sum_{i=1}^{12} M_i(B_m) \frac{1}{12} = 130 \text{ u.m.}$$



Ahora nuestro cálculo supera el valor B que queremos calcular, pues, por ejemplo, hemos supuesto que durante el mes de enero hemos tenido unos beneficios marginales de 20 u.m./año cuando esto sólo ha sido así a final de mes. La figura muestra unos triángulos de exceso de la suma respecto al valor que buscamos. En resumen tenemos que

$$110 = s(B_m, P) \leq B \leq S(B_m, P) = 130.$$

No hemos calculado el beneficio acumulado pero ahora sabemos que está entre 110 y 130 u.m. Generalicemos:

Definición Si $f : [a, b] \longrightarrow \mathbb{R}$ es una función acotada y P es una partición de $[a, b]$, se define la *suma superior de Riemann* y la *suma inferior de Riemann* de f respecto a P como

$$s(f, P) = \sum_{i=1}^{n} m_i(f)\Delta x_i, \qquad S(f, P) = \sum_{i=1}^{n} M_i(f)\Delta x_i.$$

Si la función $f(x)$ representa el incremento marginal de una cierta magnitud M (es decir, lo que aumenta M por cada unidad marginal que aumenta x en un punto dado), entonces las sumas $s(f, P)$ y $S(f, P)$ son aproximaciones por defecto y por exceso al incremento acumulado de M en el intervalo $[a, b]$. Más concretamente, la suma inferior es la aproximación que resulta de suponer que en cada intervalo $[x_{i-1}, x_i]$ el incremento marginal f permanece constantemente igual al menor valor que realmente toma en él. La suma superior es la aproximación que resulta de tomar como valor constante para f el mayor valor que realmente toma.

La clave para llegar al valor exacto del beneficio acumulado es que cuando tomamos particiones de norma cada vez más pequeña el error que cometemos se hace cada vez menor. Por ejemplo, vamos a ver qué ocurre si dividimos el intervalo $[0, 1]$ en n partes iguales de longitud $1/n$ (hasta ahora hemos trabajado con $n = 12$. Entonces $t_0 = 0$, $t_1 = 1/n$, $t_2 = 2/n$, etc. Claramente

$$m_i(B_m) = B_m(t_{i-1}) = \frac{240(i-1)}{n}, \qquad M_i(B_m) = B_m(t_i) = \frac{240i}{n}.$$

Por lo tanto

$$\begin{aligned} s(B_m, P) &= \sum_{i=1}^{n} m_i(B_m)\Delta t_i = \sum_{i=1}^{n} \frac{240(i-1)}{n}\frac{1}{n} = \frac{240}{n^2}(0 + 1 + \cdots + n - 1) \\ &= \frac{240}{n^2}\frac{(n-1)n}{2} = 120\frac{n-1}{n} = 120\left(1 - \frac{1}{n}\right). \end{aligned}$$

Similarmente

$$\begin{aligned} S(B_m, P) &= \sum_{i=1}^{n} M_i(B_m)\Delta t_i = \sum_{i=1}^{n} \frac{240i}{n}\frac{1}{n} = \frac{240}{n^2}(1 + 2 + \cdots + n) \\ &= \frac{240}{n^2}\frac{n(n+1)}{2} = 120\frac{n+1}{n} = 120\left(1 + \frac{1}{n}\right). \end{aligned}$$

Así pues, el beneficio acumulado B tiene que cumplir

$$120\left(1 - \frac{1}{n}\right) \leq B \leq 120\left(1 + \frac{1}{n}\right)$$

para todo número natural n. Como el límite cuando n tiende a infinito de los dos extremos es 120, ha de ser $B = 120$.

Definición Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función acotada. Se definen la *integral inferior de Darboux* y la *integral superior de Darboux* de f en $[a, b]$ como

$$\underline{\int_a^b} f(x)\, dx = \sup_P s(f, P), \qquad \overline{\int_a^b} f(x)\, dx = \inf_P S(f, P),$$

donde P recorre todas las particiones posibles de $[a, b]$. En otras palabras, la integral inferior es el mayor valor al que podemos acercarnos mediante sumas inferiores y la integral superior es el menor valor al que podemos acercarnos mediante sumas superiores.

Se dice que la función f es *integrable Riemann* en $[a, b]$ si ambas integrales coinciden, y entonces al valor común se le representa por

$$\int_a^b f(x)\, dx = \underline{\int_a^b} f(x)\, dx = \overline{\int_a^b} f(x)\, dx.$$

Por ejemplo, hemos comprobado que la función $B_m(t) = 240t$ es integrable Riemann en $[0, 1]$ y

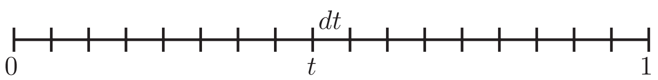
$$\int_0^1 240t\, dt = 120.$$

En conclusión, la interpretación marginal de la integral de Riemann es la siguiente:

$$\boxed{\text{Si la función } f(x) \text{ representa el incremento marginal de una función } F \text{ para cada valor de } x, \text{ entonces } \int_a^b f(x)\,dx \text{ representa el incremento acumulado de } F \text{ en el intervalo } [a,b], \text{ es decir, es igual a } F(b) - F(a).}$$

En la práctica, para determinar cuándo y de qué manera debemos calcular integrales para plantear y resolver un problema podemos emplear razonamientos infinitesimales como éste:

Volviendo a nuestro ejemplo, si los beneficios marginales de una empresa son de $240t$ u.m./año, esto significa que en un incremento de tiempo marginal de dt años nuestro beneficio se incrementará en $dB = 240t\, dt$ u.m. Si dt es un incremento de tiempo muy pequeño, entonces dB representa un incremento de beneficios muy pequeño.



Toda la discusión precedente se resume en que la suma de los infinitos incrementos infinitamente pequeños dB es matemáticamente la integral de esta función sobre el intervalo de tiempo en que se acumulan los beneficios. En este caso, $\int_0^1 240t\, dt$. Conviene tener siempre en cuenta lo siguiente:

> Para que una integral $\int_a^b f(x)\,dx$ tenga una interpretación económica el integrando $f(x)$ ha de ser siempre una magnitud marginal respecto de la variable dx, es decir, si x es tiempo, entonces el integrando tendrá que ser un beneficio marginal respecto del tiempo, o un ahorro marginal respecto del tiempo, etc. En particular tendrá que estar expresado en "unidades de lo que sea"/ unidades de tiempo. Si x es la cantidad producida de un bien, el integrando tendrá que ser el coste marginal respecto a la producción, o el beneficio marginal respecto de la producción, etc.

En la práctica, el único criterio que veremos para reconocer funciones integrables es el siguiente:

Teorema *Si una función acotada $f : [a, b] \longrightarrow \mathbb{R}$ es continua en todo $[a, b]$ salvo a lo sumo en un número finito de puntos entonces es integrable Riemann en $[a, b]$. En particular toda función continua en un intervalo $[a, b]$ es integrable Riemann.*

Para el cálculo explícito de integrales, basta aplicar el teorema siguiente:

Regla de Barrow *Si $f : [a, b] \longrightarrow \mathbb{R}$ es una función continua en $[a, b]$ entonces existe una función $F : [a, b] \longrightarrow \mathbb{R}$ continua en $[a, b]$ y derivable en $]a, b[$ tal que $F' = f$ y si F es cualquier función que cumpla esto entonces*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Cualquier función F en las condiciones de la regla de Barrow se llama *primitiva* de la función f.

Respecto al cálculo de integrales en la práctica, observemos lo siguiente: En las condiciones del ejemplo que hemos estado estudiando, llamemos $B(t)$ a la función que nos da el beneficio acumulado en el instante t. Supongamos que es derivable en todo punto. Entonces $B'(t)$ es el aumento que experimenta el beneficio acumulado en t por cada unidad marginal de tiempo que pasa, es decir, es el beneficio marginal en t. Con la notación que empleábamos: $B_m(t) = B'(t)$. Según hemos visto, $\int_a^b B_m(t)\,dt$ es el beneficio acumulado desde el instante a hasta el instante b, es decir,

$$\int_a^b B'(t)\,dt = B(b) - B(a).$$

Por ejemplo, para calcular $\int_0^1 240t\,dt$ basta observar que $F(t) = 120t^2$ es una primitiva de $240t$, luego

$$\int_0^1 240t\,dt = F(1) - F(0) = 120 - 0 = 120.$$

Un poco más en general, se cumple lo siguiente:

Teorema *Sea $f : [a, b] \longrightarrow \mathbb{R}$ una función integrable Riemann en $[a, b]$. Entonces la función $F : [a, b] \longrightarrow \mathbb{R}$ dada por*

$$F(t) = \int_a^t f(x)\,dx$$

es continua en $[a, b]$ y si f es continua entonces F es derivable y $F' = f$.

La interpretación de este teorema es que si f es el incremento marginal de una magnitud, entonces F es el incremento acumulado de la misma. Por ejemplo, si $f(t) = 240t$ es el beneficio marginal de una empresa, su beneficio acumulado (desde $t = 0$) en un instante t será

$$B(t) = \int_0^t 240x\,dx = [120x^2]_0^t = 120t^2.$$

A veces tiene interés determinar el valor medio de una función en el sentido del teorema siguiente:

Teorema de la media *Si $f : [a, b] \longrightarrow \mathbb{R}$ es una función integrable Riemann que toma valores comprendidos entre m y M, entonces se cumple*

$$m \leq \mu = \frac{\int_a^b f(x)\,dx}{b-a} \leq M.$$

Este valor μ se llama valor medio de f en $[a, b]$. Si f es continua entonces existe un punto $c \in [a, b]$ tal que $\mu = f(c)$.

Ejemplo *Supongamos que los beneficios marginales de una empresa vienen dados por $B_m(t) = 240t$ u.m./año. Calcula el beneficio medio del periodo $[1, 3]$.*

Solución: Según el teorema anterior,

$$\mu = \frac{\int_1^3 240t\,dt}{3-1} = \frac{1}{2}[120t^2]_1^3 = 480 \text{ u.m./año.}$$

Como el integrando es continuo, tiene sentido buscar en qué instante se alcanza la media. Concretamente, $240t = 480$ nos da $t = 2$. ■

Terminamos con algunas propiedades adicionales de la integral de Riemann.

Teorema *Se cumplen las propiedades siguientes:*

1. *Si f y g son funciones integrables Riemann en $[a, b]$ entonces $f + g$ también lo es y*
$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$
2. *Si f es una función integrable Riemann en $[a, b]$ y $\alpha \in \mathbb{R}$, entonces αf también lo es y*
$$\int_a^b \alpha f(x)\,dx = \alpha \int_a^b f(x)\,dx.$$

3. *Si f es una función integrable Riemann en $[a,b]$ entonces $|f|$ también lo es y*
$$\left|\int_a^b f(x)\,dx\right| \leq \int_a^b |f(x)|\,dx.$$

4. *Si f y g son funciones integrables Riemann en $[a,b]$ entonces fg también lo es, aunque no hay ninguna relación sencilla entre la integral del producto y las integrales de los factores.*

5. *Si f y g son funciones integrables Riemann en $[a,b]$ y para todo $x \in [a,b]$ se cumple $f(x) \leq g(x)$, entonces*
$$\int_a^b f(x)\,dx \leq \int_a^b g(x)\,dx.$$
En particular la integral de una función no negativa es no negativa y la integral de una función no positiva es no positiva.

6. *Si $f : [a,b] \longrightarrow \mathbb{R}$ es una función acotada en $[a,b]$ y $a < c < b$, entonces f es integrable Riemann en $[a,b]$ si y sólo si lo es en $[a,c]$ y en $[c,b]$, y en tal caso*
$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

La última propiedad permite calcular integrales de funciones discontinuas, descomponiéndolas en intervalos donde sean continuas y se pueda aplicar la regla de Barrow.

Ejemplo *Calcula la integral*
$$\int_1^3 e^{2x}\,dx.$$

Solución: El integrando es ciertamente integrable Riemann en $[1,3]$ puesto que es una función continua. Para calcular la integral observamos que si $F(x) = e^{2x}$, entonces $F'(x) = 2e^{2x}$, luego nos haría falta un 2 en el integrando para poder aplicar la regla de Barrow. Usando las propiedades de la integral concluimos que
$$\int_1^3 e^{2x}\,dx = \frac{1}{2}\int_1^3 2e^{2x}\,dx = \frac{1}{2}[e^{2x}]_1^3 = \frac{1}{2}(e^6 - e^2).$$
■

Ejemplo *Calcula la integral $\displaystyle\int_{-2}^2 f(x)\,dx$, donde*
$$f(x) = \begin{cases} x^2 - 1 & \text{si } x < 0, \\ \operatorname{sen} x & \text{si } x \geq 0. \end{cases}$$

Solución: Como las funciones $x^2 - 1$ y $\operatorname{sen} x$ son continuas en $\mathbb{R}$, son acotadas en todo intervalo acotado. Por lo tanto, la función f está acotada y a lo sumo es discontinua en $x = 0$. No es necesario estudiar si lo es de hecho o no porque aunque fuera discontinua en un punto, no por ello dejaría de ser integrable. Para calcular la integral no podemos aplicar directamente la regla de Barrow, sino que hemos de descomponer la integral en dos tramos:

$$\int_{-2}^{2} f(x)\,dx = \int_{-2}^{0} f(x)\,dx + \int_{0}^{2} f(x)\,dx = \int_{-2}^{0} (x^2-1)\,dx + \int_{0}^{2} \operatorname{sen} x\,dx$$

$$= \left[\frac{x^3}{3} - x\right]_{-2}^{0} + [-\cos x]_0^2 = -\frac{2}{3} - \cos 2 + 1 = \frac{1}{3} - \cos 2.$$

■

Aunque, como ya hemos visto, calcular integrales definidas se reduce en la práctica a buscar primitivas, para comprender la presencia de integrales en las fórmulas en las que aparecen hemos de tener presente la interpretación marginal de la integral de Riemann.

Aplicación: Rentas continuas Un problema sencillo de la matemática financiera es valorar una renta discreta. Supongamos que queremos hacer un depósito en un banco para que éste, al cabo de T_0 años nos pague una renta anual de C u.m. hasta el año T. El problema es determinar la cantidad que hemos de depositar sabiendo el tipo de interés efectivo i.

Si hemos de disponer de un capital C dentro de j años tendremos que depositar hoy un capital $C(1+i)^{-j}$ Por consiguiente, el valor actual de la renta será

$$V_0 = \sum_{j=T_0}^{T} C(1+i)^{-j}.$$

Una renta continua es una renta que no se cobra año a año, o mes a mes, o día a día, sino de forma continua, es decir, cada instante que pasa, por pequeño que éste sea, el capital reembolsado aumenta ligeramente, con un aumento menor cuanto menor sea el tiempo considerado. Naturalmente esto es ficción, pero nos lleva a considerar integrales en lugar de sumas enormes, y en muchos casos los resultados son lo suficientemente aproximados como para que, al menos, en estudios teóricos, convenga usar modelos continuos.

Una renta continua está determinada por la función de reembolso marginal $R(t)$ (medida, por ejemplo, en u.m./año) que nos indica lo que aumenta el capital reembolsado por cada unidad marginal de tiempo transcurrida. Más concretamente, si en un instante t dejamos pasar un tiempo marginal dt, nuestro capital habrá aumentado en $R(t)\,dt$. El valor actual (en $t_0 = 0$) de este capital que recibimos en t, será $R(t)(1+i)^{-t}\,dt$. El valor actual de la renta será la suma de los infinitos valores actuales de los incrementos infinitesimales de capital que recibimos en el periodo $[T_0, T]$, lo que matemáticamente se expresa mediante la integral

$$V_0 = \int_{T_0}^{T} R(t)(1+i)^{-t}\,dt.$$

Naturalmente, si en lugar de considerar el tipo de interés anual i consideramos interés continuo, el factor de descuento sería e^{-it} en lugar de $(1+i)^{-t}$. ■

Ejemplo *Calcula el valor inicial de una renta continua durante un periodo de* 5 *años por valor de* 3600 € *anuales el primer año con un incremento del* 4% *anual, considerando un tanto de descuento del* 7% *(continuo anual).*

SOLUCIÓN: El enunciado afirma que el reembolso marginal empieza siendo de 3600 € /año y se incrementa en un factor de 1.04 cada año. Por lo tanto

$$R_m(t) = 3600(1.04)^t \text{ € /año.}$$

El reembolso en un tiempo dt es $3600(1.04)^t\, dt$ €. El valor actual de este reembolso infinitesimal es $3600(1.04)^t e^{-0.07t}\, dt$ €. El valor inicial de toda la renta será la suma de estas cantidades infinitesimales cuando t varía entre 0 y 5, es decir,

$$V_0 = \int_0^5 3600(1.04)^t e^{-0.07t}\, dt.$$

Para calcular la integral conviene agrupar las dos exponenciales. Para ello usamos que

$$(1.04)^t = e^{t\ln(1.04)} \approx e^{0.04t}.$$

Por lo tanto

$$V_0 = \int_0^5 3600 e^{0.04t} e^{-0.07t} dt = 3600 \int_0^5 e^{-0.03t} dt =$$

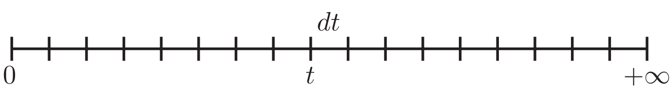
$$3600 \left[\frac{-1}{0.03} e^{-0.03t}\right]_0^5 \approx 16715.04 \text{ €}.$$

■

11.2 La integral impropia

Consideremos ahora el problema de valorar una renta continua perpetua. Concretamente, un consumidor abre un sobrecito de Nescafé y resulta premiado con un sueldo de 600 € al mes durante toda la vida. La pregunta es ¿cuánto dinero ha tenido que depositar Nescafé en un banco para que éste asuma el pago de dicha renta, supuesto un interés efectivo anual del 5%?

En la práctica podríamos hacer los cálculos bajo la hipótesis de que el consumidor no va a vivir más de 100 años, pero teóricamente resulta mucho más simple asegurar que el banco pagará la renta a perpetuidad (no resulta mucho más caro).



Concretamente, si expresamos el tiempo en años, 600 € /mes son 7200 € /año. Éste es el reembolso marginal, luego, fijado un tiempo t, el reembolso en un intervalo dt será $dR = 7.200\, dt$. El valor actual de este reembolso infinitesimal será

$dV = 7.200(1.05)^{-t}\,dt$, luego para asegurar la renta hasta un tiempo T tendremos que depositar:

$$\int_0^T 7.200(1.05)^{-t}\,dt = 7.200\left[-\frac{(1.05)^{-t}}{\ln(1.05)}\right]_0^T = \frac{7.200}{\ln(1.05)}(1-(1.05)^{-T}).$$

Si hacemos tender T a infinito queda

$$\int_0^{+\infty} 7.200(1.05)^{-t}\,dt = \frac{7.200}{\ln(1.05)} \approx 147\,570.73\,€.$$

(La diferencia con el cálculo a 100 años son $1\,122.20$ €. Vemos que, como habíamos comentado antes, no es muy grande en términos relativos.)

Acabamos de calcular lo que se conoce como una integral impropia.

Definición Sea $f : [a,+\infty[\longrightarrow \mathbb{R}$ una función integrable Riemann en cada intervalo $[a,b]$ con $b > a$. Se define la *integral impropia de primera especie*

$$\int_a^{+\infty} f(x)\,dx = \lim_{t\to+\infty}\int_a^t f(x)\,dx.$$

Si el límite no existe se dice que la integral es *divergente*. En caso contrario es *convergente*.

Similarmente, si $f :]-\infty,b] \longrightarrow \mathbb{R}$ es integrable en cada intervalo $[a,b]$, con $a < b$, se define

$$\int_{-\infty}^b f(x)\,dx = \lim_{t\to-\infty}\int_t^b f(x)\,dx.$$

Por último, si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es integrable en todo intervalo $[a,b]$, se define

$$\int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{a} f(x)\,dx + \int_a^{+\infty} f(x)\,dx,$$

donde $a \in \mathbb{R}$ es cualquier número prefijado, entendiendo que la integral es convergente si y sólo si lo son las dos integrales impropias de la derecha. Puede probarse que el valor de la integral completa no depende de la elección de a.

Ejemplo *Calcula*

$$\int_1^{+\infty}\frac{dx}{x^2}.$$

Solución: Tenemos que

$$\int_1^{+\infty}\frac{dx}{x^2} = \lim_{t\to+\infty}\int_1^t\frac{dx}{x^2} = \lim_{t\to+\infty}\left[-\frac{1}{x}\right]_1^t = \lim_{t\to+\infty}\left(-\frac{1}{t}+1\right) = 1.$$

■

Ejemplo *Calcula*

$$\int_{-\infty}^{+\infty} \frac{dx}{1+x^2}.$$

SOLUCIÓN: Partimos la integral por $x = 0$ y calculamos

$$\int_{-\infty}^{0} \frac{dx}{1+x^2} = \lim_{t \to -\infty} \int_{t}^{0} \frac{dx}{1+x^2} = \lim_{t \to -\infty} [\arctan x]_t^0 = \lim_{t \to -\infty} -\arctan t = \frac{\pi}{2}.$$

Similarmente

$$\int_{0}^{+\infty} \frac{dx}{1+x^2} = \lim_{t \to +\infty} \arctan t = \frac{\pi}{2},$$

luego

$$\int_{-\infty}^{+\infty} \frac{dx}{1+x^2} = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$

■

En este último ejemplo habríamos llegado al mismo resultado si en lugar de partir la integral en dos hubiéramos calculado directamente

$$\begin{aligned} \lim_{t \to +\infty} \int_{-t}^{t} \frac{dx}{1+x^2} &= \lim_{t \to +\infty} [\arctan x]_{-t}^{t} = \lim_{t \to +\infty} (\arctan t - \arctan -t) \\ &= \frac{\pi}{2} - \left(-\frac{\pi}{2}\right) = \pi. \end{aligned}$$

Ahora bien, esto no es correcto salvo que sepamos a priori que la integral es convergente. En efecto, si $f : \mathbb{R} \longrightarrow \mathbb{R}$ es integrable en todo intervalo finito, el límite

$$VP \int_{-\infty}^{+\infty} f(x)\, dx = \lim_{t \to +\infty} \int_{-t}^{t} f(x)\, dx$$

se llama *valor principal* de la integral $\int_{-\infty}^{+\infty} f(x)\, dx$, y coincide con la integral impropia cuando ésta existe, pero también puede existir aunque la integral sea divergente.

Ejemplo *Estudia la existencia de*

$$\int_{-\infty}^{+\infty} x^3\, dx \quad y \quad VP \int_{-\infty}^{+\infty} x^3\, dx.$$

SOLUCIÓN: Por una parte,

$$\begin{aligned} VP \int_{-\infty}^{+\infty} x^3\, dx &= \lim_{t \to +\infty} \int_{-t}^{t} x^3\, dx = \lim_{t \to +\infty} \left[\frac{x^4}{4}\right]_{-t}^{t} \\ &= \lim_{t \to +\infty} \frac{t^4}{4} - \frac{(-t)^4}{4} = \lim_{t \to +\infty} 0 = 0. \end{aligned}$$

Sin embargo, aunque vemos que existe el valor principal de la integral, no podemos decir que la integral valga 0. Al contrario, no existe, como se ve al partirla:

$$\int_0^{+\infty} x^3\,dx = \lim_{t\to+\infty}\int_0^t x^3\,dx = \lim_{t\to+\infty}\left[\frac{x^4}{4}\right] = \lim_{t\to+\infty}\frac{t^4}{4} = \infty.$$

■

Definición Si $f : [a,b[\longrightarrow \mathbb{R}$ es una función integrable Riemann en cada intervalo $[a,t]$, con $a < t < b$, entonces se define la *integral impropia de segunda especie*

$$\int_a^b f(x)\,dx = \lim_{t\to b^-}\int_a^t f(x)\,dx.$$

Como en el caso de las integrales de primera especie, se dice que la integral es convergente si existe el límite y divergente si no existe.

Así mismo, si $f :]a,b] \longrightarrow \mathbb{R}$ es integrable en los intervalos $[t,b]$, con $a < t < b$ se define

$$\int_a^b f(x)\,dx = \lim_{t\to a^+}\int_t^b f(x)\,dx.$$

Si $f :]a,b[\longrightarrow \mathbb{R}$ es integrable en cada intervalo $[t_1,t_2]$, con $a < t_1 < t_2 < b$, definimos

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

donde c es cualquier punto tal que $a < c < b$, entendiendo que la integral es convergente si y sólo si lo son las dos integrales impropias de la derecha.

Más en general, siempre que el dominio de una función f pueda descomponerse en intervalos donde f sea integrable impropia de primera o segunda especie alrededor de uno sólo de los extremos, se define la integral impropia de f como la suma de las integrales de f en cada uno de estos dominios.

Ejemplo *Estudia la integral*

$$\int_0^9 \frac{dx}{\sqrt[3]{(x-1)^2}}.$$

Solución: Observamos que el integrando es continuo en $[0,9]$ excepto en $x=1$, por lo que tiene sentido estudiar las integrales de segunda especie

$$\int_0^1 \frac{dx}{\sqrt[3]{(x-1)^2}}, \qquad \text{y} \qquad \int_1^9 \frac{dx}{\sqrt[3]{(x-1)^2}}.$$

Concretamente,

$$\begin{aligned}\int_0^1 \frac{dx}{\sqrt[3]{(x-1)^2}} &= \lim_{t\to1^-}\int_0^t (x-1)^{-2/3}\,dx = \lim_{t\to1^-} 3[(x-1)^{1/3}]_0^t \\ &= \lim_{t\to1^-} 3((t-1)^{1/3}-(-1)) = 3.\end{aligned}$$

Por otra parte,

$$\begin{aligned}\int_1^9 \frac{dx}{\sqrt[3]{(x-1)^2}} &= \lim_{t\to 1^+} \int_t^9 (x-1)^{-2/3}\,dx = \lim_{t\to 1^+} 3[(x-1)^{1/3}]_t^9 \\ &= \lim_{t\to 1^-} 3(2-(t-1)^{1/3}) = 6,\end{aligned}$$

luego la integral es convergente y

$$\int_0^9 \frac{dx}{\sqrt[3]{(x-1)^2}} = 9.$$

■

Ejemplo *Estudia la convergencia de la integral*

$$\int_0^{+\infty} \frac{dx}{x^2}.$$

SOLUCIÓN: La integral es impropia de primera especie porque uno de los límites de integración es infinito, pero también es impropia de segunda especie porque no está acotada alrededor del 0. Por lo tanto tendremos que descomponerla en dos partes:

$$\int_0^{+\infty} \frac{dx}{x^2} = \int_0^1 \frac{dx}{x^2} + \int_1^{+\infty} \frac{dx}{x^2}.$$

Analizamos cada una por separado. La integral será convergente si y sólo si lo son las dos partes.

$$\int_0^1 \frac{dx}{x^2} = \lim_{t\to 0^+} \int_t^1 \frac{dx}{x^2} = \lim_{t\to 0^+} \left[-\frac{1}{x}\right]_t^1 = \lim_{t\to 0^+} -1 + \frac{1}{t} = \infty.$$

Vemos, pues, que la primera parte es divergente, por lo que ya no es necesario estudiar la segunda. La integral del enunciado es divergente. ■

11.3 La integral múltiple

Estudiamos ahorala integración de funciones de varias variables. La diferencia fundamental, desde el punto de vista matemático, respecto de la integración de funciones de una variable es que el dominio en el cual se integra una función ya no tiene por qué ser algo tan simple como un intervalo $[a, b]$, sino que puede ser (casi) cualquier subconjunto acotado de $\mathbb{R}^n$. Por lo demás, la construcción de la integral de Riemann de varias variables es una generalización natural de la de una variable (en lugar de partir el dominio de la función en intervalos arbitrariamente pequeños se lo divide en cubos arbitrariamente pequeños y se definen igualmente las sumas superiores, inferiores, etc.) No obstante, no vamos a entrar en los detalles de la construcción, sino que nos limitaremos a explicar cómo se calcula en la práctica una integral múltiple.

Admitiremos, pues, que para cada subconjunto acotado $D \subset \mathbb{R}^n$ hay definida una familia de funciones $f : D \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ llamadas *funciones integrables Riemann* en D, para cada una de las cuales está definido el número real representado por $\int_D f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n$ y al que se le llama la *integral de Riemann* de la función f en el recinto D.

El teorema siguiente nos indica cómo calcular una integral cuando el recinto es un *cubo,* es decir, un producto de intervalos $D = [a_1, b_1] \times \cdots \times [a_n, b_n]$.

Teorema *Sea $D = [a_1, b_1] \times \cdots \times [a_n, b_n]$ un cubo en $\mathbb{R}^n$ y sea $f : D \longrightarrow \mathbb{R}$ una función integrable Riemann en D. Entonces*

$$\int_D f(x_1, \ldots, x_n)\, dx_1 \cdots dx_n = \int_{a_1}^{b_1} \cdots \left(\int_{a_n}^{b_n} f(x_1, \ldots, x_n)\, dx_n \right) \cdots dx_1$$

Por claridad vamos a escribir el caso particular correspondiente a funciones de tres variables:

$$\int_D f(x, y, z)\, dxdydz = \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \left(\int_{a_3}^{b_3} f(x, y, z)\, dz \right) dy \right) dx,$$

es decir, la integral se calcula integrando primero respecto de z la función $f(x, y, z)$ (donde consideramos a x e y como constantes), luego integrando el resultado respecto de y (considerando a x constante) y, por último, integrando el resultado respecto de x. Puede probarse así mismo que el orden de integración es irrelevante, es decir, también podríamos integrar primero respecto de x, luego respecto de z y luego respecto de y.

Ejemplo *Calcula $\int_D (x^2 y + z)\, dxdydz$, donde $D = [1, 3] \times [-1, 2] \times [0, 1]$.*

Solución: Se cumple que

$$\int_D (x^2 y + z)\, dxdydz = \int_1^3 \int_{-1}^2 \int_0^1 (x^2 y + z) dzdydx$$

$$= \int_1^3 \int_{-1}^2 \left[x^2 yz + \frac{z^2}{2} \right]_0^1 dydx = \int_1^3 \int_{-1}^2 \left(x^2 y + \frac{1}{2} - 0 \right) dydx$$

$$= \int_1^3 \left[\frac{x^2 y^2}{2} + \frac{y}{2} \right]_{-1}^2 dx = \int_1^3 \left(2x^2 + 1 - \frac{x^2}{2} + \frac{1}{2} \right) dx$$

$$= \int_1^3 \left(\frac{3x^2}{2} + \frac{3}{2} \right) dx = \left[\frac{x^3}{2} + \frac{3x}{2} \right]_1^3 = 18 - 2 = 16.$$

■

En el ejemplo anterior hemos admitido tácitamente que el polinomio $x^2 y + z$ es integrable Riemann en el cubo D. Esto es consecuencia del teorema siguiente, que recoge las propiedades básicas de la integral múltiple:

Teorema *Propiedades de la integral de Riemann:*

1. Toda función continua es integrable en todo recinto cerrado y acotado.

2. Si $D = D_1 \cup D_2 \subset \mathbb{R}^n$ y $D_1 \cap D_2 = \varnothing$, entonces una función $f : D \longrightarrow \mathbb{R}$ es integrable Riemann en D si y sólo si lo es en D_1 y en D_2, y en tal caso
$$\int_D f\,dx_1 \cdots dx_n = \int_{D_1} f\,dx_1 \cdots dx_n + \int_{D_2} f\,dx_1 \cdots dx_n.$$

3. Si f y g son integrables Riemann en D, entonces $f + g$ también lo es y
$$\int_D (f+g)\,dx_1 \cdots dx_n = \int_D f\,dx_1 \cdots dx_n + \int_D g\,dx_1 \cdots dx_n.$$

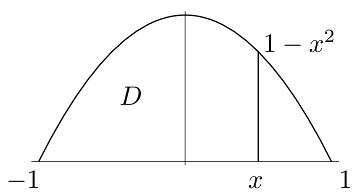
4. Si f es integrable Riemann en D y $\alpha \in \mathbb{R}$, entonces αf también lo es y
$$\int_D \alpha f\,dx_1 \cdots dx_n = \alpha \int_D f\,dx_1 \cdots dx_n.$$

5. Si f y g son integrables Riemann en D y $f(\bar{x}) \leq g(\bar{x})$ para todo $\bar{x} \in D$, entonces
$$\int_D \alpha f\,dx_1 \cdots dx_n \leq \int_D g\,dx_1 \cdots dx_n.$$

Para calcular una integral $\int_D f(x_1, \ldots, x_n)\,dx_1 \cdots dx_n$ cuando el recinto acotado D no es un cubo, consideramos un cubo C que contenga a D y la función $f^* : C \longrightarrow \mathbb{R}$ que coincide con f sobre D y vale 0 en los demás puntos de C. Se cumple que la función nula es integrable en $C \setminus D$ y su integral es nula, por lo que las propiedades del teorema anterior nos dan que
$$\int_D f(x_1, \ldots, x_n)\,dx_1 \cdots dx_n = \int_C f^*(x_1, \ldots, x_n)\,dx_1 \cdots dx_n,$$
y la segunda integral sabemos calcularla.

Ejemplo *Calcula $\int_D xy\,dxdy$, donde D es el recinto limitado por la parábola $1 - x^2$ y el eje X.*

SOLUCIÓN: Un cubo que contiene a D es $C = [-1, 1] \times [0, 1]$.



Hemos de integrar la función f^* que vale xy en D pero vale 0 en los demás puntos de C. Así,

$$\int_D xy\,dxdy = \int_{-1}^{1}\left(\int_0^1 f^*(x,y)\,dy\right)dx.$$

Ahora bien, para cada valor de x, la función $f^*(x,y)$ sólo vale xy cuando la y varía entre 0 y $1-x^2$, luego

$$\int_D xy\,dxdy = \int_{-1}^{1}\left(\int_0^{1-x^2} xy\,dy\right)dx = \int_{-1}^{1}\left[\frac{xy^2}{2}\right]_0^{1-x^2} dx$$

$$= \int_{-1}^{1}\frac{x(1-x^2)^2}{2}dx = \frac{1}{2}\int_{-1}^{1}(x-2x^3+x^5)dx = \frac{1}{2}\left[\frac{x^2}{2}-\frac{x^4}{2}+\frac{x^6}{6}\right]_{-1}^{1}$$

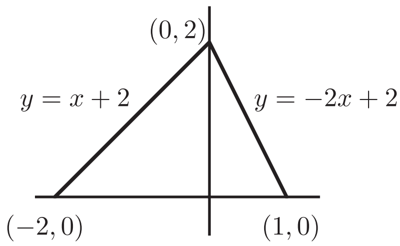
$$= \frac{1}{2}\left(\frac{1}{2}-\frac{1}{2}+\frac{1}{6}-\left(\frac{1}{2}-\frac{1}{2}+\frac{1}{6}\right)\right) = 0.$$

■

Para calcular una integral doble podemos integrar primero respecto de x y luego respecto de y o al revés. El resultado será el mismo, aunque el cálculo puede ser más sencillo de una forma que de otra.

Ejemplo *Calcula la integral $\int_T e^{x+y}\,dxdy$, donde T es el triángulo de vértices $(-2,0)$, $(1,0)$ y $(0,2)$.*

SOLUCIÓN: Necesitamos las ecuaciones de los lados del triángulo. Es claro que son las indicadas en la figura:



Para recorrer todo el triángulo la x ha de variar entre -2 y 1 y, fijado un valor de x, la y varía entre 0 y $x+2$ cuando $x \leq 0$ y entre 0 y $-2x+2$ cuando $x \geq 0$. Para considerar los dos casos hemos de descomponer la integral en dos:

$$\int_T e^{x+y}dxdy = \int_{T_1} e^{x+y}dxdy + \int_{T_2} e^{x+y}dxdy,$$

donde T_1 es el triángulo de vértices $(-2,0)$, $(0,0)$ y $(0,2)$ y T_2 es el triángulo de vértices $(0,0)$, $(1,0)$ y $(0,2)$.

$$\int_{T_1} e^{x+y}dxdy = \int_{-2}^{0}\int_0^{x+2} e^{x+y}dydx = \int_{-2}^{0}\left[e^{x+y}\right]_0^{x+2} dx = \int_{-2}^{0}(e^{2x+2}-e^x)\,dx$$

$$= \left[\frac{1}{2}e^{2x+2} - e^x\right]_{-2}^{0} = \frac{1}{2}e^2 - 1 - \frac{1}{2}e^{-2+e^{-2}} = \frac{1}{2}e^2 + \frac{1}{2}e^{-2} - 1,$$

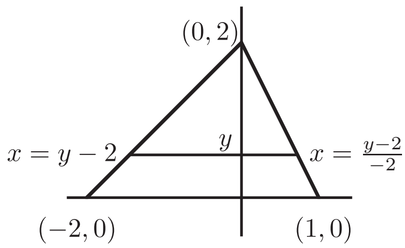
$$\int_{T_2} e^{x+y}dxdy = \int_0^1 \int_0^{-2x+2} e^{x+y}dydx = \int_0^1 [e^{x+y}]_0^{-2x+2}dx$$

$$= \int_0^1 (e^{-x+2} - e^x)\,dx = [-e^{-x+2} - e^x]_0^1 = -e - e - e^2 + 1 = -2e + e^2 + 1.$$

Por lo tanto, la integral completa es

$$\int_T e^{x+y}dxdy = \frac{1}{2}e^2 + \frac{1}{2}e^{-2} - 1 - 2e + e^2 + 1 = \frac{3}{2}e^2 - 2e + \frac{1}{2}e^{-2}.$$

Hemos calculado la integral integrando primero respecto de y y luego respecto de x. Una solución alternativa consiste en integrar primer respecto de x y luego respecto de y. Para recorrer el triángulo la y ha de variar entre 0 y 2 y, para un valor fijo de y, la x varía entre $y-2$ y $\frac{y-2}{-2}$:



Por consiguiente podemos calcular la integral sin necesidad de separar en dos sumandos:

$$\int_T e^{x+y}dxdy = \int_0^2 \int_{y-2}^{\frac{y-2}{-2}} e^{x+y}dxdy = \int_0^2 [e^{x+y}]_{y-2}^{\frac{y-2}{-2}}\,dy$$

$$= \int_0^2 (e^{(y+2)/2} - e^{2y-2})dy = \left[2e^{(y+2)/2} - \frac{1}{2}e^{2y-2}\right]_0^2$$

$$= 2e^2 - \frac{1}{2}e^2 - 2e + \frac{1}{2}e^{-2} = \frac{3}{2}e^2 - 2e + \frac{1}{2}e^{-2}.$$

Vemos que de esta forma el cálculo ha sido más corto. ■

11.4 Ejercicios

1. Calcula

$$\int_0^{\sqrt{\pi}} x \operatorname{sen} x^2\,dx, \quad \int_{10}^{17} \sqrt{x-1}\,dx, \quad \int_0^5 xe^x\,dx, \quad \int_0^5 x^3e^{x^2}\,dx$$

$$\int_0^{\pi} \operatorname{sen}^5 x \cos x\,dx, \quad \int_0^5 \frac{dx}{3x+2}, \quad \int_1^3 xe^x\,dx,$$

$$\int_0^1 xe^{x^2}\,dx, \quad \int_0^{\pi} (x^3 + x \operatorname{sen} x^2)\,dx.$$

2. Considera la función $f : \mathbb{R} \longrightarrow \mathbb{R}$ dada por

$$f(x) = \begin{cases} 7 & \text{si } x \leq 2, \\ 5 & \text{si } x > 2. \end{cases}$$

¿Es integrable en $[0, 8]$?, ¿por qué? Calcula analítica y geométricamente

$$\int_0^8 f(x)\, dx$$

3. Repite el problema anterior (excepto el cálculo geométrico de la integral), con la función $g : [0, +\infty[\longrightarrow \mathbb{R}$ dada por

$$g(x) = \begin{cases} (2-x)e^x & \text{si } x \leq 2, \\ \dfrac{1}{x+3} & \text{si } x > 2. \end{cases}$$

4. La función de costes marginales de una empresa es

$$C_m(x) = \frac{1}{(x+5)^2}.$$

Calcula el coste variable medio de fabricar 10 unidades de producto. Si los costes fijos son de 10 u.m., calcula la función de coste total.

5. Los costes fijos de una empresa son de 20 u.m., y la función de costes marginales es

$$C_m(x) = \frac{10}{(x+1)^2}.$$

Calcula la función de coste total $C(x)$. ¿Cuál es la función de coste variable?

6. La función de costes marginales de una empresa es

$$C(x) = 2 + \frac{1}{x+10} \text{ u.m.},$$

donde x es el número de unidades producidas. (Esto significa que cada unidad producida tiene un coste fijo de 2 u.m. y un coste variable que es menor cuanto más unidades se fabrican.) Además la producción tiene un coste fijo de 20 u.m. Calcula la función de costes de la empresa. Determina concretamente el coste de una producción de 90 unidades de producto.

7. Si el coste marginal de una empresa es $C_m(x) = e^{-x}$, ¿cuál es la función de coste variable?

8. El ahorro de una persona viene dado por $A(t) = 1200 + 6t$ €/año. Si el banco le da un interés efectivo anual de $i = 0.05$, calcula el ahorro acumulado en los primeros 5 años.

9. La función de reembolso de una renta percibida en un periodo de 15 años (a partir de $t = 0$) es $R(t) = 3000t^2$ € (donde el tiempo t está en años). Calcula la cantidad total percibida y el valor final de la renta, considerando un interés continuo del 7%.

10. Determina si las integrales siguientes son convergentes y en caso afirmativo calcula la integral:

$$\int_0^{+\infty} \frac{dx}{x^4}, \quad \int_{-2}^{6} (x-2)^{5/3}\, dx, \quad \int_{-\infty}^{-1} e^x\, dx, \quad \int_1^{+\infty} e^x\, dx, \quad \int_0^1 \frac{dx}{1-x}.$$

11. Un trabajador está estudiando dos ofertas de planes de pensiones. El banco A le ofrece un sistema de pagos mensuales de modo que el día previsto para su jubilación habrá generado un capital de 6.700.000 €, por el cual el banco le ofrece una pensión de 350.000 €/año. El banco B le propone un plan de depósitos al mismo tipo de interés con el que generará un capital de 5.800.000 € y a cambio le ofrece una pensión de 300.000 €/año. Suponiendo, por simplicidad, que las pensiones son perpetuas, ¿qué plan ofrece mayor rentabilidad? (En otros términos, la pregunta es qué tipo de interés i aplica cada banco para valorar su propuesta de pensión. Este valor recibe el nombre de Tasa de Rendimiento Interno (TIR) de la inversión.)

12. La función de beneficios marginales de una empresa viene dada por

$$B_m(t) = \operatorname{sen}(2\pi t) \text{ u.m./año.}$$

Calcula el beneficio acumulado durante el primer año e interpreta el resultado. Calcula el beneficio medio del primer semestre.

13. Se estima que los dividendos marginales de una empresa van a ser

$$D_m(t) = e^{0.1t}\text{u.m./año.}$$

Calcula el valor de las acciones de la empresa en $t = 0$, es decir, el valor actual de los dividendos que la empresa producirá en el periodo $[0, +\infty[$, actualizados con un interés continuo del 7%.

14. Calcula el valor inicial de una renta continua con función de reembolso

$$C(t) = 216.000 + 720t \ \text{€/año}$$

a percibir dentro de 1 año y durante un periodo de 3 años valorándola con un factor de descuento continuo de $i_\infty = 8\%$. Calcula también el reembolso acumulado en dicho periodo.

15. Una empresa estudia una inversión de 200 u.m. en maquinaria, de la que espera obtener un rendimiento de $100\, e^{-0.04}t$ u.m./año durante 5 años. Determina si la inversión es rentable calculando su VAN, es decir, calcula la diferencia entre el valor inicial de los rendimientos menos la inversión. Para el descuento considera un interés continuo de $i_\infty = 0.06$. Calcula también el rendimiento medio anual de la inversión.

16. Estudia la integral
$$\int_0^{+\infty} \frac{e^{-2x}}{1+e^{-2x}}\,dx$$

17. Calcula las integrales siguientes:

(a) $\displaystyle\int_D x^2(y-1)dxdy$, donde $D = [1,2]\times[1,3]$.

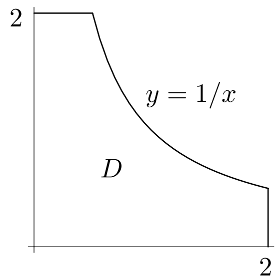
(b) $\displaystyle\int_D x\cos xy\,dxdy$, donde $D = [1,2]\times[0,\pi]$.

(c) $\displaystyle\int_D (xy+z)\,dxdydz$, donde $D = [0,1]\times[0,1]\times[-1,1]$.

(d) $\displaystyle\int_D (x+z^2)\,dxdydz$, donde $D = [-1,1]\times[0,1]\times[2,3]$.

(e) $\displaystyle\int_D \frac{xy}{(x^2+y^2)^2}dxdy$, donde $D = [0,1]\times[1,2]$.

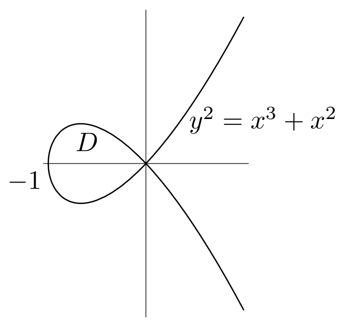
(f) $\displaystyle\int_D y\,dxdy$, donde D es el recinto indicado en la figura:



(g) $\displaystyle\int_D y\,dxdy$, donde $D = \{(x,y)\in\mathbb{R}^2 \mid x^2+y^2\le 1,\ y\ge 0\}$.

(h) $\displaystyle\int_D \frac{1}{x-10}\,dxdy$, donde D es el cuadrado de vértices $(1,0)$, $(0,1)$, $(-1,0)$ y $(0,-1)$.

(i) $\displaystyle\int_D (3x^2+2x)dxdy$, donde D es el recinto limitado por $y^2 = x^3+x^2$.



12. Ecuaciones diferenciales

Una ecuación diferencial ordinaria de primer grado es una ecuación de la forma $y' = f(x, y)$, donde $y = y(x)$ es una función desconocida. Resolver la ecuación es encontrar todas las funciones y que la satisfacen. Por ejemplo, la ecuación diferencial

$$y' = xy$$

tiene entre sus soluciones a la función $y(x) = e^{x^2/2}$. En efecto, si y es esta función, entonces $y' = xe^{x^2/2} = xy$, tal y como exige la ecuación.

En este tema veremos algunos métodos para resolver ecuaciones diferenciales de este tipo. Hay que advertir que es posible plantear y resolver ecuaciones diferenciales mucho más generales: una ecuación ordinaria de grado n es una ecuación en la que aparecen las derivadas de la función incógnita hasta el orden n. El adjetivo "ordinaria" se opone a "ecuación en derivadas parciales", que es una ecuación cuya función incógnita tiene varias variables y en la que aparecen las derivadas parciales de la misma. También es posible plantear sistemas de ecuaciones diferenciales con varias incógnitas. Pero aquí nos ocuparemos únicamente del caso que hemos descrito.

En general, una ecuación diferencial tiene infinitas soluciones, de modo que la familia de todas ellas depende de una constante de integración. Por ejemplo, la *solución completa* de la ecuación anterior es $y = Ce^{x^2/2}$, donde C es un número real arbitrario. Para cada valor de C tenemos una solución distinta.

Una solución concreta de una ecuación diferencial queda determinada si se especifica el valor en un punto de la función incógnita. Por ejemplo, la única solución de la ecuación anterior que cumple $y(0) = 3$ es la función $y = 3e^{x^2/2}$.

A menudo es frecuente usar la notación $\frac{dy}{dx}$ en lugar de y' en las ecuaciones diferenciales. Por ejemplo, la ecuación anterior puede expresarse también como

$$\frac{dy}{dx} = xy,$$

o incluso

$$dy = xy\,dx$$

12.1 Ecuaciones con variables separables

Las ecuaciones más fáciles de resolver son las de variables separables. Son aquellas que pueden escribirse como

$$u(y)\,dy = v(x)\,dx,$$

donde u y v son dos funciones. En tal caso, basta integral los dos miembros

$$\int u(y)\,dy = \int v(x)\,dx$$

para encontrar la solución general.

Ejemplo *Resuelve la ecuación diferencial $y' = xy$.*

SOLUCIÓN: La ecuación es de variables separables, pues se puede escribir en la forma

$$\frac{dy}{y} = x\,dx.$$

Para resolverla hacemos

$$\int \frac{dy}{y} = \int x\,dx,$$

lo que nos da

$$\ln y = \frac{x^2}{2} + k.$$

Para despejar la y calculamos la exponencial de ambos miembros:

$$y = e^{x^2/2+k} = e^k e^{x^2/2}.$$

Si llamamos $C = e^k$ queda la ecuación que habíamos indicado anteriormente: $y = Ce^{x^2/2}$. ■

Aplicación: Capitalización con interés variable Si invertimos un capital C_0, éste variará con el tiempo, es decir, tendremos una función $C(t)$ con $C(0) = C_0$. El interés instantáneo de la inversión es el tanto por uno (o, equivalentemente) tanto por ciento de incremento del capital en un instante, es decir, (en tanto por uno)

$$i_\infty(t) = \frac{1}{C}\frac{dC}{dt}.$$

Si conocemos el interés tenemos una ecuación diferencial que nos permite calcular el capital en un instante dado. Si el interés es constante obtenemos la fórmula usual de capitalización continua:

$$\frac{dC}{C} = i_\infty\,dt \quad\Rightarrow\quad \ln C = i_\infty t + K, \quad\Rightarrow\quad C = e^K e^{i_\infty t} \quad\Rightarrow\quad C = C_0 e^{i_\infty t}.$$

No obstante, la ecuación diferencial puede resolverse igualmente para intereses variables.

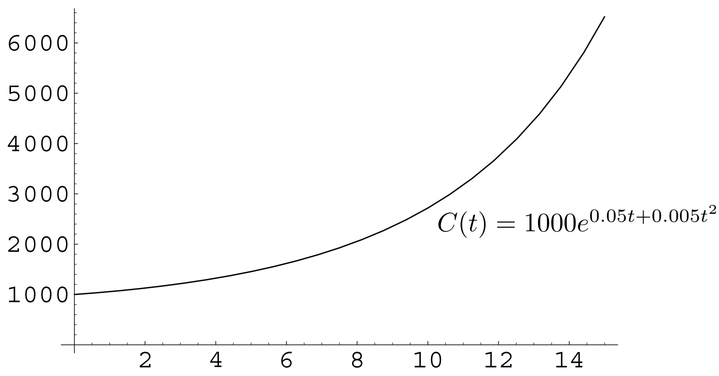
Por ejemplo, si depositamos un capital de 1000 € a un interés continuo variable, que ha resultado ser $i_\infty(t) = 0.05 + 0.01t$, el capital se obtiene como

$$C(t) = C_0 e^{(0.05+0.01t)t} = C_0 e^{(0.05t+0.01t^2)}$$

Como en el instante inicial se depositan 1000 € se tiene $C_0 = C(0) = 1000$. Entonces, el capital se puede expresar en función del tiempo como

$$C(t) = 1000e^{(0.05t+0.01t^2)}$$

La gráfica de la solución que hemos obtenido es la siguiente:



12.2 Ecuaciones lineales

Son las ecuaciones de la forma

$$y' + a(x)y = b(x).$$

Se resuelven haciendo el cambio de variable $y(x) = u(x)v(x)$.

Ejemplo *Resuelve la ecuación diferencial lineal $y' + 2xy = x$.*

Solución: Hacemos el cambio $y = uv$, con lo que $y' = u'v + uv'$. Resulta:

$$u'v + uv' + 2xuv = x.$$

Sacamos factor común u y queda

$$(*) \qquad u(v' + 2xv) + u'v - x = 0.$$

Ahora elegimos v de modo que anule el paréntesis, es decir,

$$v' + 2xv = 0.$$

Esta ecuación es de variables separables

$$dv = -2xv\,dx \quad \Rightarrow \quad \frac{dv}{v} = -2x\,dx,$$

$$\int \frac{dv}{v} = -\int 2x\,dx,$$

$$\ln v = -x^2 \quad \Rightarrow \quad v = e^{-x^2}.$$

No ponemos ninguna constante de integración porque nos basta una solución. Sustituyendo v en $(*)$ queda

$$u'e^{-x^2} - x = 0.$$

Esta ecuación también es de variables separables:

$$du = xe^{x^2}\,dx \quad \Rightarrow \quad u = \int xe^{x^2}\,dx = \frac{1}{2}e^{x^2} + c.$$

La solución es, por tanto,

$$y = uv = e^{-x^2}\left(\frac{1}{2}e^{x^2} + c\right).$$

■

Aplicación: El precio como función del tiempo *Supongamos que las funciones de oferta y demanda de un bien son*

$$Q_d = 12 - p, \qquad Q_s = -6 + 2p.$$

La tasa de cambio del precio respecto al tiempo (en días) es un tercio de la demanda excedente $Q_d - Q_s$.

1. *Calcula la función* $p(t)$.
2. *Calcula el precio esperado dentro de 2 días si el precio actual es* $p(0) = 4$ €.
3. *Calcula el precio esperado dentro de 10 días suponiendo* $p(0) = 4$ € *y, alternativamente,* $p(0) = 8$ €. *Interpreta el resultado.*

Solución:
1) La tasa de cambio de precio es su derivada respecto al tiempo, luego el dato es que

$$p' = \frac{1}{3}(Q_d - Q_s) = \frac{1}{3}(12 - p + 6 - 2p) = 6 - p.$$

Tenemos, pues, una ecuación diferencial lineal. Para resolverla hacemos $p = uv$, con lo que $p' = u'v + uv'$ y al sustituir queda

$$u'v + uv' = 6 - uv \quad \Rightarrow \quad u'v + uv' + uv = 6 \quad \Rightarrow \quad u(v' + v) + u'v = 6.$$

Resolvemos $v' + v = 0$, es decir,

$$\frac{dv}{dt} = -v \quad \Rightarrow \quad \frac{dv}{v} = -dt \quad \Rightarrow \int \frac{dv}{v} = -\int dt \quad \Rightarrow \quad \ln v = -t.$$

Despejando, $v = e^{-t}$. Para este valor de v, la ecuación se reduce a $u'e^{-t} = 6$, luego $u' = 6e^t$,

$$u = \int 6e^t\,dt = 6e^t + C.$$

La función precio será $p(t) = uv = (6e^t + C)e^{-t} = 6 + Ce^{-t}$.

2) Sabiendo el precio inicial $p(0) = 4$ podemos calcular la constante C:

$$4 = p(0) = 6 + Ce^0 = 6 + C \quad \Rightarrow \quad C = -2.$$

La función precio es $p(t) = 6 - 2e^{-t}$ y para $t = 2$ queda $p(5) = 6 - 2e^5 = 5.72$ €.

3) Para $p(0) = 4$ hemos visto que el precio es $p(t) = 6 - 2e^{-t}$, de donde obtenemos que $p(10) = 5,99991 \approx 6$ €.

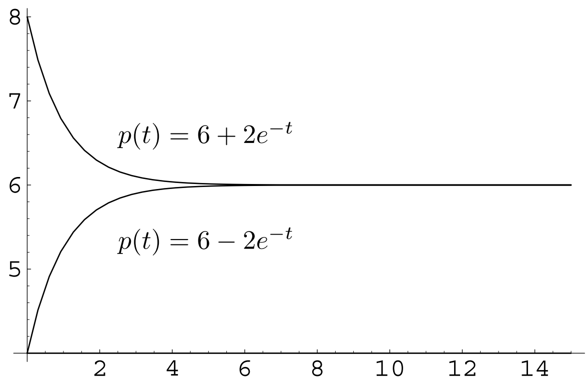
Para $p(0) = 8$ la constante C es

$$8 = 6 + Ce^{-0} \quad \Rightarrow \quad C = 2,$$

luego $p(t) = 6 + 2e^{-t}$ y $p(10) = 6,00009 \approx 6$ €.

Se obtiene aproximadamente el mismo resultado porque el precio tiende al precio de equilibrio, que, como es fácil comprobar, es de 6 €. ■

Si representamos gráficamente las funciones $p(t) = 6 + 2e^{-t}$ y $p(t) = 6 - 2e^{-t}$ se ve claramente que ambas tienden al precio de equilibrio $p = 6$ €:



12.3 Ejercicios

1. Comprueba que las funciones dadas son soluciones de las ecuaciones dadas:

   (a) $y' + 2y = 0, \qquad y(x) = e^{-2x}, \qquad y(x) = 5e^{-2x}$.

   (b) $y' + xy = 0, \qquad y(x) = e^{-x^2/2}$.

   (c) $y' + y = \operatorname{sen} x, \qquad y(x) = e^{-x} - \dfrac{1}{2}\cos x + \dfrac{1}{2}\operatorname{sen} x$.

2. Comprueba que las familias de funciones dadas son soluciones de las ecuaciones dadas. Encontrar la única solución que cumple en cada caso la condición que se indica.

(a) $y' + 2y = 0, \qquad y(x) = Ae^{-2x}, \qquad y(0) = 2.$

(b) $y' + y = \operatorname{sen} x, \quad y(x) = Ae^{-x} - \dfrac{1}{2}\cos x + \dfrac{1}{2}\operatorname{sen} x, \quad y(0) = -1.$

(c) $y' + 2y = x^2, \quad y = \dfrac{1}{4} - \dfrac{1}{2}x + \dfrac{1}{2}x^2 + Ae^{-2x}, \quad u(0) = 1.$

3. Resuelve:

(a) $\dfrac{dy}{dx} = 2y,$

(b) $\dfrac{dy}{dx} = -\dfrac{x}{y},$

(c) $(1 + e^x)y\,\dfrac{dy}{dx} = e^x, \quad y(0) = 1,$

(d) $(1 + y^2)\,dx + (1 + x^2)\,dy = 0,$

(e) $(1 + y^2)\,dx + xy\,dy = 0,$

(f) $\dfrac{dy}{dx}\operatorname{sen} x = y\cos x,$

(g) $x\sqrt{1 - x^2}\,dx + y\sqrt{1 - y^2}\,dy = 0, \qquad y(0) = 1,$

(h) $y \ln y\,dx + x\,dy = 0, \qquad y(1) = 1.$

4. Resuelve:

(a) $\dfrac{dy}{dx} + 2xy = 2xe^{-x^2},$

(b) $\dfrac{dy}{dx} + 2y = e^{-x},$

(c) $\dfrac{dy}{dx} + 2xy = 2x, \qquad y(0) = -1,$

(d) $\dfrac{dy}{dx} + xy = x^3,$

(e) $\dfrac{dy}{dx} + \dfrac{e^x}{e^x + 1}\,y = \dfrac{x}{e^x + 1}, \qquad y(0) = 1,$

(f) $\dfrac{dy}{dx} + \dfrac{y}{x} = \dfrac{\cos x}{x}, \qquad y(\pi/2) = 4/\pi,$

(g) $\dfrac{d\theta}{dt} = e^{2t} + 2\theta.$

5. Sea p el precio de un bien, y supongamos que la oferta y la demanda vienen dadas por

$$S(p) = 2p, \qquad D(p) = 100 - 8p.$$

Calcula el precio de equilibrio. Supongamos que el precio p varía con el tiempo $p = p(t)$ y que $p' = 0.5(D(p) - S(p))$. Interpreta económicamente esta condición. Calcula $p(t)$ y comprueba que $\lim\limits_{t\to+\infty} p(t)$ es el precio de equilibrio.

6. La población de cierto país aumenta proporcionalmente al número de habitantes. Si después de dos años la población se ha duplicado y después de tres años es de 20.000 habitantes, calcula la población inicial.

7. Depositamos un capital de 1000 u.m. durante 10 años a un interés continuo variable, que ha resultado ser $i_\infty(t) = 0.04 + 0.009t$. Calcula el capital final.

8. Se nos plantea la posibilidad de invertir un capital por un periodo de tres años. De entre las distintas expectativas sobre la rentabilidad de la inversión, la menos favorable pronostica que la evolución del interés será $i_\infty(t) = 5 + 16t - 3t^2$ %. Determina el mínimo capital que debemos invertir para asegurarnos un capital final de 1000 €.

A. Formas cuadráticas

Definición Si A es una matriz simétrica $n \times n$ de números reales, se llama *forma cuadrática* determinada por A a la aplicación $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por $q(\bar{x}) = \bar{x}A\bar{x}^t$.

Ejemplo *Calcula la forma cuadrática determinada por la matriz simétrica*

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

Solución: Se trata de la aplicación $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ dada por

$$\begin{aligned} q(x,y,z) &= (x,y,z) \begin{pmatrix} 2 & 3 & 1 \\ 3 & -1 & -2 \\ 1 & -2 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= (2x+3y+z, 3x-y-2z, x-2y+4z) \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= 2x^2+3xy+xz+3xy-y^2-2yz+xz-2yz+4z^2 \\ &= 2x^2-y^2+4z^2+6xy+2xz-4yz. \end{aligned}$$

■

En la práctica hay una regla simple para obtener directamente el resultado sin hacer ningún cálculo: los coeficientes de x^2, y^2, z^2 son siempre los que aparecen en la diagonal de la matriz, y los coeficientes de xy, xz, yz son los que aparecen por encima de la diagonal multiplicados por 2. Por ejemplo, para calcular el coeficiente de xy (primera variable por segunda variable) miramos la primera fila segunda columna, donde hay un 3, y lo multiplicamos por 2, con lo que resulta un 6. Aplicando este criterio al revés podemos obtener la matriz que determina una forma cuadrática dada.

Ejemplo *Calcula la matriz asociada a la forma cuadrática*

$$q(x,y,z) = 3x^2 - 5y^2 + z^2 + 4xy - 3yz.$$

Solución: La matriz es

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

■

De este modo, una forma cuadrática viene determinada indistintamente por una matriz simétrica o por su expresión explícita $q(\bar{x})$. Nos referiremos a ellas como la *expresión matricial* y la *expresión analítica* de la forma cuadrática.

La razón por la que nos interesan las formas cuadráticas es que el término de segundo grado del polinomio de Taylor en un punto $\bar{p}$ de una función f de clase C^2 es (salvo un factor $1/2$) la forma cuadrática determinada por la matriz hessiana $Hf(\bar{p})$. Por ello los resultados que vamos a obtener sobre formas cuadráticas se aplicarán al estudio de funciones C^2 a través de la matriz hessiana.

Definición Una forma cuadrática $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ es:

1. *Definida positiva* si $q(\bar{x}) > 0$ siempre que $\bar{x} \neq \bar{0}$.

2. *Definida negativa* si $q(\bar{x}) < 0$ siempre que $\bar{x} \neq \bar{0}$.

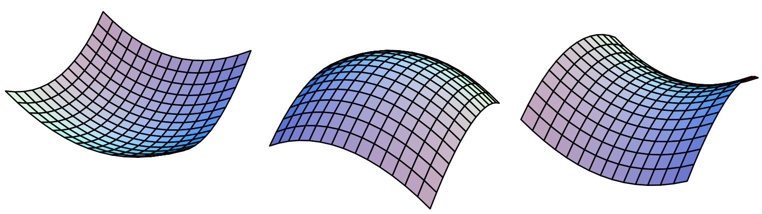
3. *Semidefinida positiva* si $q(\bar{x}) \geq 0$ para todo $\bar{x}$.

4. *Semidefinida negativa* si $q(\bar{x}) \leq 0$ para todo $\bar{x}$.

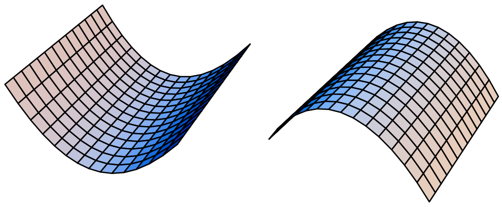
5. *Indefinida* si $q(\bar{x}) > 0$ en algunos puntos y $q(\bar{x}) < 0$ en otros puntos.

6. *Nula* si $q(\bar{x}) = 0$ en todo punto $\bar{x}$.

Por ejemplo, la forma cuadrática $q(x, y) = x^2 + y^2$ es definida positiva, mientras que $q(x, y) = -x^2 - y^2$ es definida negativa. La forma $q(x, y) = x^2 - y^2$ es indefinida, porque $q(1, 0) = 1 > 0$ y $q(0, 1) = -1 < 0$. La forma $q(x, y) = x^2$ es semidefinida positiva y $q(x, y) = -x^2$ es semidefinida negativa. Las gráficas de estas funciones son las siguientes:



$x^2 + y^2$ $\qquad$ $-x^2 - y^2$ $\qquad$ $x^2 - y^2$



x^2 $\qquad\qquad$ $-x^2$

Todas las formas cuadráticas son esencialmente como una de estas funciones, es decir, una forma cuadrática definida positiva tiene forma de copa, toda forma cuadrática semidefinida positiva tiene forma de valle, etc. Por lo tanto, si somos capaces de *clasificar* una forma cuadrática, tenemos una información cualitativa sobre cómo es su gráfica.

Hay un caso en el que es especialmente simple decidir de qué tipo es una forma cuadrática:

Teorema *Si A es una matriz diagonal, entonces la forma cuadrática asociada a A es:*

1. *definida positiva si todos los coeficientes de la diagonal son > 0,*
2. *definida negativa si todos los coeficientes de la diagonal son < 0,*
3. *semidefinida positiva si todos los coeficientes de la diagonal son ≥ 0,*
4. *semidefinida negativa si todos los coeficientes de la diagonal son ≤ 0,*
5. *indefinida si tiene coeficientes > 0 y < 0.*

En general, para cualquier matriz daremos un criterio que involucra los conceptos siguientes:

Definición Una matriz cuadrada es *regular* si su determinante es no nulo y es *singular* si su determinante es nulo.

Los *menores principales* de orden k de una matriz $n \times n$ (con $k \leq n$) son los determinantes de las matrices formadas por k filas de A (en orden) y las mismas k columnas.

El *menor principal conducente* de orden k es el menor principal formado con las k primeras filas y las k primeras columnas de la matriz.

Ejemplo *Calcula los menores principales y los menores principales conducentes de la matriz*

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

Solución: Sus menores principales de orden 1 son

$$A^1 = 2, \quad A^2 = -1, \quad A^3 = 4.$$

Sus menores principales de orden 2 son

$$A^{12} = \begin{vmatrix} 2 & 3 \\ 3 & -1 \end{vmatrix} = -11, \quad A^{13} = \begin{vmatrix} 2 & 1 \\ 1 & 4 \end{vmatrix} = 7,$$

$$A^{23} = \begin{vmatrix} -1 & -2 \\ -2 & 4 \end{vmatrix} = -8.$$

El único menor principal de orden 3 es

$$A^{123} = \begin{vmatrix} 2 & 3 & 1 \\ 3 & -1 & -2 \\ 1 & -2 & 4 \end{vmatrix} = -63.$$

Los menores principales conducentes son

$$A_1 = |2| = 2, \qquad A_2 = \begin{vmatrix} 2 & 3 \\ 3 & -1 \end{vmatrix} = -11, \qquad A_3 = -63.$$

■

El interés de los menores principales consiste en que son fáciles de calcular y nos permiten clasificar cualquier forma cuadrática:

Teorema (Criterio de Jacobi) *Sea $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ una forma cuadrática cuya matriz asociada es A. Entonces*

1. *En el caso en que A sea regular:*
   - (a) *Si los menores principales conducentes son todos > 0 entonces q es definida positiva.*
   - (b) *Si los menores principales conducentes tienen signos alternados*

     $$A_1 < 0, \quad A_2 > 0, \quad A_3 < 0, \ldots$$

     entonces q es definida negativa. Nótese que el primero debe ser negativo.
   - (c) *En otro caso q es indefinida.*
2. *En el caso en que A sea singular:*
   - (a) *Si todos los menores principales son ≥ 0 entonces A es semidefinida positiva.*
   - (b) *Si los menores principales de orden impar son ≤ 0 y los de orden par son ≥ 0 entonces A es semidefinida negativa.*
   - (c) *En otro caso q es indefinida.*

Para aplicar este criterio, en primer lugar hemos de calcular el determinante de A para saber si es regular o singular. Si es regular sólo necesitaremos calcular los menores principales conducentes, mientras que si es singular tendremos que estudiar el signo de todos los menores principales, conducentes o no.

Ejemplo *Clasifica las formas cuadráticas siguientes:*

1. $q(x, y, z) = x^2 - y^2 - 2z^2 + xy$.
2. $q(x, y, z) = 2x^2 + y^2 + z^2 - xy$.

SOLUCIÓN:

1. La matriz asociada a q es

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

En primer lugar calculamos su determinante: $|A| = \frac{5}{2}$. Como es no nulo basta estudiar los menores principales conducentes, que son:

$$A_1 = 1, \qquad A_2 = \begin{vmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & -1 \end{vmatrix} = -\frac{5}{4}, \qquad A_3 = \frac{5}{2}.$$

Según la regla de Jacobi, la forma cuadrática es indefinida.

2. La matriz asociada a q es

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

Como $|A| = \frac{7}{4} \neq 0$, sólo necesitamos estudiar los menores principales conducentes:

$$A_1 = 2, \qquad A_2 = \begin{vmatrix} 2 & -\frac{1}{2} \\ -\frac{1}{2} & 1 \end{vmatrix} = \frac{7}{4}, \qquad A_3 = \frac{7}{4}.$$

Como todos son positivos, la regla de Jacobi asegura que la forma cuadrática es definida positiva. ■

Formas cuadráticas restringidas

Definición Sea $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ una forma cuadrática y $S \subset \mathbb{R}^n$ un subespacio vectorial de $\mathbb{R}^n$ dado por las ecuaciones $B\bar{x}^t = \bar{0}$, donde B es una matriz $m \times n$ con $m < n$. La restricción de q a S es la función $\tilde{q} : S \longrightarrow \mathbb{R}$ dada por

$$\tilde{q}(\bar{x}) = q(\bar{x}).$$

Es decir, $\tilde{q}$ es la misma función q pero actuando sobre los puntos de un subespacio. Esta diferencia es importante si queremos estudiar el signo de $\tilde{q}$:

$\tilde{q}$ es *definida positiva* si sobre $S \sim \{0\}$ sólo toma valores positivos
$\tilde{q}$ es *definida negativa* si sobre $S \sim \{0\}$ sólo toma valores negativos
$\tilde{q}$ es *semidefinida positiva* si sobre S toma valores positivos o nulos.
$\tilde{q}$ es *semidefinida negativa* si sobre S toma valores negativos o nulos.
$\tilde{q}$ es *indefinida* si toma valores positivos y negativos en S.

Cuando la forma cuadrática q es definida positiva o negativa, continuará siéndolo al restringirla a cualquier subespacio S. Sin embargo, puede que q sea indefinida pero que $\tilde{q}$ sea definida positiva o negativa. Por tanto, necesitamos un método que nos proporcione el signo de la forma cuadrática restringida.

Matriz orlada asociada a un subespacio Sea A una matriz simétrica de orden n y $S \subset \mathbb{R}^n$ un subespacio vectorial dado por las ecuaciones $B\bar{x}^t = \bar{0}$, donde B es una matriz $m \times n$ con $m < n$. Definimos la *matriz orlada asociada* a A y a B como

$$H = \left(\begin{array}{c|c} A & B^t \\ \hline B & 0 \end{array} \right).$$

Ejemplo *Consideramos la matriz*

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

y el subespacio $S = \{(x_1, x_2, x_3, x_4) : x_1 + x_2 - x_4 = 0, \quad 2x_2 - x_3 = 0\}$. *Calcula la matriz orlada de A asociada a S.*

Solución: El subespacio S está definido por el sistema de ecuaciones siguiente:

$$\begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 2 & -1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ t \end{pmatrix} = 0,$$

entonces, la matriz $B = \begin{pmatrix} 1 & 1 & 0 & -1 \\ 0 & 2 & -1 & 0 \end{pmatrix}$. Por tanto,

$$H = \left(\begin{array}{c|c} A & B^t \\ \hline B & 0 \end{array} \right) = \left(\begin{array}{cccc|cc} 1 & -1 & 2 & 0 & 1 & 0 \\ -1 & 2 & 0 & 1 & 1 & 2 \\ 2 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ \hline 1 & 1 & 0 & -1 & 0 & 0 \\ 0 & 2 & -1 & 0 & 0 & 0 \end{array} \right)$$

■

Menores principales orlados Definimos el *menor principal orlado* de orden r de una matriz orlada

$$H = \left(\begin{array}{c|c} A & B^t \\ \hline B & 0 \end{array} \right)$$

como el determinante

$$H_r = \left| \begin{array}{c|c} A_r & B_r^t \\ \hline B_r & 0 \end{array} \right|,$$

donde A_r es el menor principal conducente de orden r de la matriz A y B_r es la matriz formada por las r primeras columnas de la matriz B (por tanto, B_r^t es la matriz formada por las r primeras filas de la matriz B^t).

Ejemplo *Dada la matriz*

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

y el subespacio

$$S = \{(x_1, x_2, x_3, x_4) \in R^4 : x_1 + x_2 + 2x_4 = 0,\ x_1 - x_3 = 0\}.$$

Calcula H_2 y H_4.

SOLUCIÓN: La matriz asociada al subespacio S es

$$B = \begin{pmatrix} 1 & 1 & 0 & 2 \\ 1 & 0 & -1 & 0 \end{pmatrix}.$$

Por tanto, la matriz orlada es

$$\left(\begin{array}{rrrr|rr} 1 & -2 & 1 & 0 & 1 & 1 \\ -2 & 0 & 2 & 1 & 1 & 0 \\ 1 & 2 & -2 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 2 & 0 \\ \hline 1 & 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \end{array}\right)$$

Entonces,

$$H_2 = \left|\begin{array}{c|c} A_2 & B_2^t \\ \hline B_2 & 0 \end{array}\right| = \begin{vmatrix} 1 & -2 & 1 & 1 \\ -2 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{vmatrix} = 1,$$

$$H_4 = \left|\begin{array}{c|c} A_4 & B_4^t \\ \hline B_4 & 0 \end{array}\right| = \begin{vmatrix} 1 & -2 & 1 & 0 & 1 & 1 \\ -2 & 0 & 2 & 1 & 1 & 0 \\ 1 & 2 & -2 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 2 & 0 \\ 1 & 1 & 0 & 2 & 0 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 \end{vmatrix} = -4.$$

■

Clasificación mediante menores orlados Consideramos una forma cuadrática q con matriz asociada A, de orden n, restringida a un subespacio

$$S = \{\bar{x} \in R^n : B\bar{x}^t = \bar{0}\},$$

donde B es una matriz de orden $m \times n$. Suponemos que el determinante de las m primeras columnas de B es no nulo. Se cumple:

1. Si $(-1)^m H_r > 0 \quad \forall r = m+1, \ldots, n.$, entonces q restringida a S es definida positiva .

2. Si $(-1)^r H_r > 0 \quad \forall r = m+1, \ldots, n.$, entonces q restringida a S es definida negativa.

Ejemplo *Clasifica la forma cuadrática*

$$q(x,y,z) = -2x^2 - y^2 + z^2 + xz$$

restringida al subespacio $S = \{(x,y,z) : x + z = 0\}$.

SOLUCIÓN: La matriz asociada a q es

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

La ecuación que define S es

$$(1,0,1)\begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0,$$

es decir, $B = (1,0,1)$. Por tanto, la matriz orlada es

$$H = \left(\begin{array}{c|c} A & B^t \\ \hline B & 0 \end{array}\right) = \left(\begin{array}{ccc|c} -2 & 0 & \frac{1}{2} & 1 \\ 0 & -1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 1 \\ \hline 1 & 0 & 1 & 0 \end{array}\right).$$

En este ejemplo,

$$\begin{array}{rcl} n & = & \text{número de variables de } q = \text{orden de } A = 3, \\ m & = & \text{número de filas de } B = \text{número de ecuaciones de } S = 1, \\ r & = & m+1,\dots,n = 2,3. \end{array}$$

Por tanto, tenemos que calcular los menores orlados H_2, H_3.

$$H_2 = \left|\begin{array}{c|c} A_2 & B_2^t \\ \hline B_2 & 0 \end{array}\right| = \begin{vmatrix} -2 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = 1,$$

$$H_3 = \left|\begin{array}{c|c} A_3 & B_3^t \\ \hline B_3 & 0 \end{array}\right| = \begin{vmatrix} -2 & 0 & \frac{1}{2} & 1 \\ 0 & -1 & 0 & 0 \\ \frac{1}{2} & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{vmatrix} = -2.$$

Como $(-1)^2 H_2 = 1 > 0$ y $(-1)^3 H_3 = 2 > 0$ se cumple que $(-1)^r H_r > 0$ para $r = 2,3$ y en consecuencia $\tilde{q}$ es definida negativa. ∎

Ejemplo Clasifica la forma cuadrática

$$q(x,y,z,t) = 2x^2 + 5y^2 - t^2 + 2zt$$

restringida al subespacio $S = \{(x,y,z,t) : 2y - z + t = 0, \quad x + 3y - 3z + 2t = 0\}$.

Solución: La matriz asociada a q es

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

El sistema de ecuaciones que define a S es

$$\begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & 3 & -3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0.$$

Como $B = \begin{pmatrix} 0 & 2 & -1 & 1 \\ 1 & 3 & -3 & 2 \end{pmatrix}$, la matriz orlada es

$$H = \left(\begin{array}{c|c} A & B^t \\ \hline B & 0 \end{array}\right) = \left(\begin{array}{cccc|cc} 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 5 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & -1 & 1 & 2 \\ \hline 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & 3 & -3 & 2 & 0 & 0 \end{array}\right).$$

Por tanto, tenemos

$$\begin{array}{rcl} n & = & \text{número de variables de } q = \text{orden de } A = 4, \\ m & = & \text{número de filas de } B = \text{número de ecuaciones de } S = 2, \\ r & = & m+1, \ldots, n = 3, 4. \end{array}$$

Entonces, debemos calcular los menores orlados H_3 y H_4:

$$H_3 = \left|\begin{array}{c|c} A_3 & B_3^t \\ \hline B_3 & 0 \end{array}\right| = \begin{vmatrix} 2 & 0 & 0 & 0 & 1 \\ 0 & 5 & 0 & 2 & 3 \\ 0 & 0 & 0 & -1 & -3 \\ 0 & 2 & -1 & 0 & 0 \\ 1 & 3 & -3 & 0 & 0 \end{vmatrix} = 23,$$

$$H_4 = \left(\begin{array}{c|c} A_4 & B_4^t \\ \hline B_4 & 0 \end{array}\right) = \begin{vmatrix} 2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 5 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 1 & -1 & -3 \\ 0 & 0 & 1 & -1 & 1 & 2 \\ 0 & 2 & -1 & 1 & 0 & 0 \\ 1 & 3 & -3 & 2 & 0 & 0 \end{vmatrix} = 5.$$

Como $(-1)^2 H_3 = 23 > 0$ y $(-1)^2 H_4 = 5 > 0$ se cumple que $(-1)^m H_r > 0$ para $r = 3, 4$ y en consecuencia q es definida positiva. ■

Otro método Para acabar este apéndice expondremos otro método que permite clasificar formas cuadráticas restringidas a subespacios a partir de la forma explícita de la forma cuadrática restringida. Dada la forma cuadrática q restringida al subespacio $S = \{\bar{x} \in \mathbb{R}^n \mid B\bar{x}^t = \bar{0}\}$.

1. Obtenemos la forma explícita de los vectores de S resolviendo el sistema $B\bar{x}^t = \bar{0}$.
2. Calculamos la expresión de la forma cuadrática restringida $\tilde{q}$ sustituyendo la solución anterior en q.
3. Clasificamos $\tilde{q}$ con el criterio Jacobi.

Ejemplo *Clasifica las formas cuadráticas de los dos ejemplos anteriores utilizando la forma explícita de las formas cuadráticas restringidas.*

SOLUCIÓN:

1. Tenemos $q(x,y,z) = -2x^2 - y^2 + z^2 + xz$ restringida a $S = \{(x,y,z) : x+z = 0\}$. Resolvemos el sistema

$$x + z = 0 \quad \Rightarrow \quad z = -x.$$

Sustituyendo esta solución en q obtenemos la forma cuadrática restringida:

$$\tilde{q} = q(x,y,-x) = -2x^2 - y^2 + x^2 - x^2 = -2x^2 - y^2$$

Ahora debemos clasificar $\tilde{q}$ con el criterio de Jacobi, pero, en este caso, es evidente que es $\tilde{q}$ es definida negativa porque $-2x^2 - y^2 \leq 0$ ∎

2. Tenemos $q(x,y,z,t) = 2x^2 + 5y^2 - t^2 + 2zt$ restringida al subespacio vectorial $S = \{(x,y,z,t) : 2y - z + t = 0, \quad x + 3y - 3z + 2t = 0\}$.
Resolvemos el sistema de ecuaciones que define a S:

$$\left.\begin{array}{r} 2y - \ z + \ t = 0 \\ x + 3y - 3z + 2t = 0 \end{array}\right\} \quad \Rightarrow \quad x = z + y, \quad t = z - 2y, \quad y, z \in \mathbb{R}.$$

Entonces, la forma cuadrática restringida es

$$\begin{array}{rcl} \tilde{q}(y,z) & = & q(z+y, y, z, z-2y) = 2(z+y)^2 + 5y^2 - (z-2y)^2 + 2z(z-2y) \\ & = & 3y^2 + 3z^2 + 4yz. \end{array}$$

Para estudiar el signo de $\tilde{q}$ aplicacmos la regla de Jacobi a su matriz asociada:

$$C = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}.$$

Como $A_1 = 3$ y $A_2 = 5$, todos sus menores principales conducentes son positivos, por tanto $\tilde{q}$ es definida positiva. ∎

Ejercicios

1. Clasifica las formas cuadráticas siguientes:
   (a) $f_1(x,y,z) = 24x^2 + 5y^2 + 11z^2 + 20xy + 6yz$,
   (b) $f_2(x,y,z) = -45x^2 - 5y^2 - 11z^2 + 6xy + 14xz + 6yz$,

(c) $f_3(x, y, z) = 3x^2 + 2y^2 + 2xy$,

(d) $f_4(x, y) = 3x^2 + 2y^2 + 2xy$,

(e) $f_5(x, y, z) = -4x^2 - 4xy + 2xz - 4y^2 + 10yz - 7z^2$,

(f) $f_6(x, y, z) = -2x^2 + y^2 + z^2 + 2xy$,

(g) $f_7(x, y, z) = -x^2 - 3y^2 - 2z^2$,

(h) $f_8(x, y, z) = 2xy - 4yz + 3x^2 + 3y^2 + 6z^2$,

(i) $f_9(x, y, z) = 5x^2 + 5y^2 + 12z^2 + 2xy + 2xz - 14yz$,

(j) $f_{10}(x, y, z) = 9x^2 + 3y^2 + 6z^2 - 6xy + 4xz - 8yz$,

(k) $f_{11}(x, y, z) = 2x^2 + y^2 - z^2 + 8xy + 2xz + 4yz$,

(l) $f_{12}(x, y, z) = -x^2 + y^2 + 2z^2$,

(m) $f_{13}(x, y, z) = 13x^2 - 2xy - 10yz + 4y^2 + 7z^2$,

(n) $f_{14}(x, y) = -x^2 + 20y^2 + xy$,

(o) $f_{15}(x, y, z) = yz$,

(p) $f_{16}(x, y, z) = -y^2$,

(q) $f_{17}(y) = -y^2$,

(r) $f_{18}(x, y) = 5x^2 + 3y^2$,

(s) $f_{19}(x, y, z) = 2x^2 + 4z^2$.

2. Consideramos la forma cuadrática

$$f(x, y, z) = 3x^2 + 5y^2 - 3z^2 + 2xy - xz + 6yz.$$

Calcula su matriz asociada y su matriz hessiana.

3. Clasifica las formas cuadráticas restringidas dadas por las funciones y subespacios siguientes:

(a) $f_1(x, y, z) = -x^2 + y^2 + z^2$ y $S = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 0\}$

(b) $f_2(x, y, z) = x^2 + 2y^2 + z^2$ restringida a

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x + 2y - z = 0, \quad x - z = 0\}$$

(c) $f_3(x, y, z) = -x^2 - 2y^2 + z^2$ y $S = \{(x, y, z) \in \mathbb{R}^3 \mid x - z = 0\}$

(d) $f_4(x, y, z) = x^2 - 2y^2 - z^2$ y $S = \{(x, y, z) \in \mathbb{R}^3 \mid 2x - 3y + 4z = 0\}$

B. Tablas

Tabla de derivadas más usuales

Función	Derivada
$y = f(x)^m, \quad m \in \mathbb{R}$	$y' = m\, f(x)^{m-1}\, f'(x)$
$y = \ln f(x)$	$y' = \dfrac{f'(x)}{f(x)}$
$y = \log_a f(x)$	$y' = \dfrac{f'(x)}{f(x)} \log_a(e)$
$y = a^{f(x)}$	$y' = f'(x) a^{f(x)} \ln(a)$
$y = e^{f(x)}$	$y' = f'(x) e^{f(x)}$
$y = \operatorname{sen} f(x)$	$y' = f'(x) \cos f(x)$
$y = \cos f(x)$	$y' = -f'(x) \operatorname{sen} f(x)$
$y = \tan f(x)$	$y' = f'(x) \dfrac{f'(x)}{\cos^2 f(x)} = \big(1 + \tan^2 f(x)\big)$
$y = \operatorname{cotan} f(x)$	$y' = \dfrac{-f'(x)}{\operatorname{sen}^2 f(x)} = -f'(x)\big(1 + \operatorname{cotan}^2 f(x)\big)$
$y = \operatorname{arcsen} f(x)$	$y' = \dfrac{f'(x)}{\sqrt{1 - [f(x)]^2}}$
$y = \arccos f(x)$	$y' = \dfrac{-f'(x)}{\sqrt{1 - [f(x)]^2}}$
$y = \arctan f(x)$	$y' = \dfrac{f'(x)}{1 + [f(x)]^2}$
$y = \operatorname{arc\,cotan} f(x)$	$y' = \dfrac{-f'(x)}{1 + [f(x)]^2}$

Tabla de primitivas inmediatas

Dada una función $f :]a, b[\longrightarrow \mathbb{R}$ derivable en $]a, b[$, recurriendo a las reglas de derivación se obtiene la siguiente tabla de integrales indefinidas:

$$\int [f(x)]^n f'(x)\, dx = \frac{[f(x)]^{n+1}}{n+1} + C; \quad n \neq -1$$

$$\int \frac{f'(x)}{f(x)}\, dx = \ln |f(x)| + C$$

$$\int e^{f(x)} f'(x)\, dx = e^{f(x)} + C$$

$$\int a^{f(x)} f'(x)\, dx = \frac{1}{\ln a}\, a^{f(x)} + C; \quad a > 0, \quad a \neq 1$$

$$\int f'(x) \operatorname{sen} f(x)\, dx = -\cos f(x) + C$$

$$\int f'(x) \cos f(x)\, dx = \operatorname{sen} f(x) + C$$

$$\int f'(x)\big(1 + \tan^2 f(x)\big)\, dx = \tan f(x) + C$$

$$\int \frac{f'(x)}{\cos^2 f(x)}\, dx = \tan f(x) + C$$

$$\int f'(x)\big(1 + \operatorname{cotan}^2 f(x)\big)\, dx = -\operatorname{cotan} f(x) + C$$

$$\int \frac{f'(x)}{\operatorname{sen}^2 f(x)}\, dx = -\operatorname{cotan} f(x) + C$$

$$\int \frac{f'(x)}{\sqrt{1 - [f(x)]^2}}\, dx = \operatorname{arcsen} f(x) + C$$

$$\int \frac{f'(x)}{1 + [f(x)]^2}\, dx = \operatorname{arctan} f(x) + C$$

Bibliografía

Aguiló Pons, I. y Matas Andreu, R. (1998): *Àlgebra lineal. Aplicacions en economia,* Palma de Mallorca, Universitat de les Illes Balears.

Alegre, P. *et al.* (1991): *Ejercicios resueltos de Matemáticas Empresariales 1,* Madrid, AC

Ballbás, A. *et al.* (1989): *Análisis matemático para la Economía I. Cálculo diferencial,* Madrid, AC.

Barbolla, R. y Sanz, P. (1998): *Algebra lineal y teoría de matrices,* Madrid, Prentice Hall.

Caballero, R. *et al.* (1992): *Métodos matemáticos para la economía,* Madrid, McGraw-Hill.

Caballero, R. *et al.* (1993): *Matemáticas aplicadas a la economía y a la empresa,* Madrid, Pirámide.

Canós, M. J. e Ivorra, C. (1999): *Matemàtiques per a economistes. Càlcul diferencial,* Valencia, Universitat de València.

Casany, J. *et al.* (1991): *Cálculo integral,* Valencia, Nau Llibres.

Casasús, T. *et al.* (1991): *Matemáticas empresariales,* Valencia, Nau Llibres.

Chiang, A. (1987): *Métodos fundamentales Economía Matemática,* Madrid, McGraw-Hill.

del Castillo Ballesteros, V. *et al.* (1990): *Matemàtiques 2 COU,* ant Adrià del Besòs, Bruño.

Gil Martos, J. *et al.* (1981): *Matemática,* Madrid, Santillana.

Grafe, J. (1991): *Matemáticas para economistas,* Madrid, McGraw-Hill.

Larson, J. *et al.* (1991): *Cálculo,* vol 2, Madrid, McGraw-Hill.

López Cachero, M. y Vegas Pérez, A. (1994): *Curso básico de matemáticas para la economía y dirección de empresas II. Ejercicios,* Madrid, Pirámide.

Muñoz, F. *et al.* (1988): *Manual de álgebra lineal,* Barcelona, Ariel Economía.

Sydsaeter, K. y Hammond, P.J. (1996): *Matemáticas para el análisis económico,* Madrid, Prentice Hall.