

# Índice

---

<b>1 Matrices y operaciones con matrices</b>	<b>1</b>
1.1 Matrices	1
1.2 Adición de matrices	4
1.3 Multiplicación de un escalar por una matriz	7
1.4 El producto de matrices	8
1.5 Aplicación del producto matriz-vector	15
1.6 Ejercicios	19
<b>2 Resolución de sistemas de ecuaciones lineales con matrices</b>	<b>23</b>
2.1 Introducción	23
2.2 Resolución de sistemas lineales	26
2.3 Matrices escalonadas y rango de una matriz	35
2.3.1 Matrices escalonadas	35
2.3.2 Rango de una matriz	37
2.4 Sistemas lineales homogéneos	40
2.5 El Teorema de Rouché-Frobenius	43
2.6 Aplicación: el modelo input-output	45
2.7 Métodos iterativos	48
2.8 Códigos detectores y correctores de errores	53
2.9 Problemas	61
<b>3 Determinante de una matriz</b>	<b>65</b>
3.1 Introducción	65
3.2 Propiedades de los determinantes	70
3.3 Regla de Cramer	74
3.4 Aplicación: puntos para representar curvas y superficies.	77
3.4.1 Las órbitas de los asteroides.	79
3.5 El cifrado matricial de Hill	83
3.6 Problemas	90

<b>4 Valores y vectores propios de una matriz</b>	<b>93</b>
4.1 Introducción	93
4.2 Valores y vectores propios de una matriz	95
4.3 Propiedades de los valores y vectores propios	99
4.3.1 Valores y vectores propios de $A^2$	100
4.3.2 Descomposición de $A$ mediante valores y vectores propios	101
4.4 Diagonalización de matrices específicas	104
4.4.1 Diagonalización de matrices simétricas	104
4.4.2 Diagonalización de una matriz definida positiva	106
4.5 Valores y vectores propios de matrices $2 \times 2$ - representación gráfica	107
4.6 La factorización y algoritmo QR	111
4.7 El método de Lanczos para el cálculo de valores y vectores propios	116
4.8 Aplicación: las matrices de Markov	120
4.9 Problemas	127
<b>5 Grafos y matrices</b>	<b>131</b>
5.1 Algunos ejemplos de grafos	131
5.1.1 Tráfico aéreo	131
5.1.2 El grafo de un equipo de fútbol	133
5.1.3 El grafo de una red de caminos	136
5.1.4 La colección de grafos KONECT	137
5.1.5 Grafos contra el terrorismo	139
5.2 Breve introducción a los grafos	141
5.3 Matrices asociadas a un grafo	148
5.4 Propiedades espectrales de $A$ y $L$	154
5.5 Vector de Fiedler de la Laplaciana	156
5.6 Las ciudades representadas como grafos	156
5.7 El grafo dual de un grafo $G$	160
<b>6 Breve introducción a la teoría de redes</b>	<b>165</b>
6.1 Introducción	165
6.2 Algunos conceptos y medidas básicas en redes	167
6.2.1 Longitud media de caminos en una red	168
6.2.2 Coeficiente de comunidades o agrupamiento	168
6.2.3 Distribución de grados	171
6.3 La centralidad en redes complejas	183
6.4 El concepto PageRank	191
6.4.1 El PageRank	192
6.4.2 Cálculo del vector PageRank por sistemas de ecuaciones	194
6.4.3 Un ejemplo	195
6.4.4 Analizando una red terrorista	198
<i>Bibliografía</i>	203

# 1 Matrices y operaciones con matrices

---

El campo de las matrices estuvo sembrado antes de recoger el fruto. Los determinantes fueron estudiados a mediados del siglo XVIII. Un determinante contiene un cuadro de números que podía ser estudiado en sí mismo y manipulado para muchos propósitos. El cuadro de números es lo que llamamos *matriz*. La palabra matriz fue utilizada por primera vez por Sylvester (1814–1897) en 1850. Es cierto, como dice Cayley (1821–1895), que la idea de matriz es anterior a la de determinante, pero históricamente el orden fue el inverso. Cayley fue el primero en desarrollar de modo independiente el concepto de matriz en un artículo que se publicó en 1855. A partir de este artículo, los trabajos sobre matrices se dispararon.

## 1.1 Matrices

**Definición 1.1.1** Definimos una **matriz** de orden o dimensión  $m \times n$  a un conjunto cualquiera de  $mn$  elementos (generalmente números) distribuidos rectangularmente en  $m$  filas y  $n$  columnas y que delimitaremos entre corchetes. Cada uno de esos escalares constituye un **elemento** de la matriz.

---

**Ejemplo 1.1.1** Las matrices

$$\begin{bmatrix} 4 & -3 & 0 & 6 \end{bmatrix}, \quad \begin{bmatrix} 2 & -3 & 1 \\ 0 & 2 & -7 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 5 \\ -3 \\ 2 \\ 0 \\ 6 \end{bmatrix}$$

son matrices de tamaños  $1 \times 4$ ,  $2 \times 3$  y  $5 \times 1$  respectivamente.

**Definición 1.1.2** Diremos que dos matrices son iguales cuando tienen la misma dimensión y los elementos que ocupan el mismo lugar en ambas son iguales.

En adelante, siempre denotaremos las matrices por letras mayúsculas  $A, B, C, \dots$  y utilizaremos la notación de doble subíndice para referirnos a los elementos de una matriz. Así, denotaremos por  $a_{ij}$  al elemento de la matriz que ocupa la posición en la fila  $i$ -ésima y en la columna  $j$ -ésima, como se aprecia en la figura 1.1.

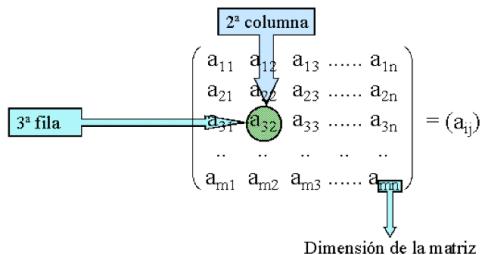


Figura 1.1 Representación gráfica de una matriz..

Con esta notación, escribimos una matriz general  $A$  de tamaño  $m \times n$  como

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

aunque en muchas ocasiones convendrá escribir la matriz anterior simplemente como

$$A = [a_{ij}]_{m \times n} \quad \text{o} \quad A = [a_{ij}]$$

utilizando esta última notación cuando el tamaño de la matriz está claro por el contexto.

**Definición 1.1.3** Sea  $A = [a_{ij}]_{m \times n}$  una matriz de tamaño  $m \times n$ . Llamamos **matriz traspuesta** de  $A$ , y escribimos  $A^T$ , a la matriz de tamaño  $n \times m$  cuyo elemento  $(i,j)$  coincide con el elemento  $(j,i)$  de  $A$ , es decir

$$A^T = [a_{ji}]_{n \times m}.$$

Notemos que, de acuerdo con esta definición, la  $i$ -ésima fila de  $A^T$  coincide con la  $i$ -ésima columna de  $A$ .

**Ejemplo 1.1.2** Para las matrices

$$A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 2 & 5 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & -1 \\ 2 & 3 \\ 5 & 0 \\ 1 & 2 \end{bmatrix}$$

tenemos que

$$A^T = \begin{bmatrix} 2 & 0 \\ 3 & 2 \\ -1 & 5 \end{bmatrix} \quad \text{y} \quad B^T = \begin{bmatrix} 4 & 2 & 5 & 1 \\ -1 & 3 & 0 & 2 \end{bmatrix}.$$

Claramente  $(A^T)^T = A$  para cualquier matriz  $A$ .

**Definición 1.1.4** Decimos que  $A = [a_{ij}]_{m \times n}$  es una **matriz cuadrada** si tiene el mismo número de filas que de columnas, es decir, si  $m = n$ .

Así,

$$\begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 3 & -1 \\ 1 & 2 & 3 \\ 5 & -4 & 1 \end{bmatrix}$$

son matrices cuadradas de tamaño  $2 \times 2$  y  $3 \times 3$  respectivamente.

**Definición 1.1.5** Decimos que una matriz cuadrada  $A = [a_{ij}]$  es **triangular superior** si  $a_{ij} = 0$ , cuando  $i > j$ , y **triangular inferior** si  $a_{ij} = 0$ , cuando  $i < j$ .

Decimos que una matriz es **triangular** si es triangular inferior o superior.

Las matrices

$$\begin{bmatrix} 1 & -2 & 5 \\ 0 & 3 & -1 \\ 0 & 0 & 6 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & 0 & 0 \\ 1 & 7 & 0 \\ 5 & -4 & 1 \end{bmatrix}$$

son triangular superior y triangular inferior respectivamente.

Notemos que la traspuesta de una matriz triangular superior es una matriz triangular inferior y viceversa.

**Definición 1.1.6** Para una matriz cuadrada  $A = [a_{ij}]_{n \times n}$  decimos que  $a_{11}, a_{22}, \dots, a_{nn}$  son los **elementos diagonales** de  $A$ .

Decimos que  $A$  es una **matriz diagonal** si los elementos no diagonales son todos nulos.

Así

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

son matrices diagonales de tamaño  $3 \times 3$  y  $4 \times 4$  respectivamente.

Notemos que una matriz diagonal es una matriz que es triangular superior e inferior al mismo tiempo. Así mismo, la traspuesta de una matriz diagonal coincide consigo misma.

**Definición 1.1.7** Llamamos **matriz fila** a una matriz de tamaño  $1 \times n$  y **matriz columna** a una matriz de tamaño  $m \times 1$ .

Es decir, una matriz fila es una matriz que tiene una única fila, mientras que una matriz columna es una matriz con una única columna.

Las matrices

$$\begin{bmatrix} 1 & -3 & 2 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 2 & 1 & 0 & -5 \end{bmatrix}$$

son matrices fila mientras que

$$\begin{bmatrix} 2 \\ 1 \\ 0 \\ 5 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -3 \\ 2 \\ 1 \\ -5 \\ 7 \end{bmatrix}$$

son matrices columna.

Claramente, la traspuesta de una matriz fila es una matriz columna y la traspuesta de una matriz columna es una matriz fila.

Finalmente, en algunas ocasiones escribiremos la  $n$ -tuple  $(u_1, u_2, \dots, u_n)$  como una matriz columna

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

de tamaño  $n \times 1$ , que no debemos confundir con la matriz fila

$$\begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$$

de tamaño  $1 \times n$ . Es decir

$$(u_1, u_2, \dots, u_n) = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}$$

pero

$$(u_1, u_2, \dots, u_n) \neq \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}.$$

## 1.2 Adición de matrices

**Definición 1.2.1** Sean  $A = [a_{ij}]$  y  $B = [b_{ij}]$  dos matrices de tamaño  $m \times n$ . La **adición** de las matrices  $A$  y  $B$  es la matriz  $A + B$  también de tamaño  $m \times n$ , llamada **suma** de  $A$  y  $B$ , cuyo elemento de la posición  $(i,j)$  es  $a_{ij} + b_{ij}$ , es decir

$$A + B = [a_{ij} + b_{ij}].$$

Así, la suma de dos matrices del mismo tamaño es la matriz, también del mismo tamaño, cuyos elementos se obtienen sumando los correspondientes elementos de cada una de ellas.

Notemos que la adición de  $n$ -tuplas (o vectores) es un caso particular de la adición de matrices.

**Ejemplo 1.2.1** Si

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 5 \\ 4 & -1 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} -7 & 0 \\ 1 & 2 \\ -3 & 1 \end{bmatrix}$$

entonces

$$A + B = \begin{bmatrix} 2 - 7 & -1 + 0 \\ 3 + 1 & 5 + 2 \\ 4 - 3 & -1 + 1 \end{bmatrix} = \begin{bmatrix} -5 & -1 \\ 4 & 7 \\ 1 & 0 \end{bmatrix}.$$

**Definición 1.2.2** Llamamos **matriz nula** de tamaño  $m \times n$ , y la denotamos por  $O$ , a la matriz que tiene todos los elementos iguales a 0.

Por ejemplo

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

son las matrices nulas de tamaños  $2 \times 2$  y  $3 \times 4$  respectivamente. A veces escribimos  $O_{m \times n}$  para referirnos a la matriz nula de tamaño  $m \times n$ .

**Definición 1.2.3** Llamamos **matriz opuesta** de la matriz  $A = [a_{ij}]$  de tamaño  $m \times n$ , y escribimos  $-A$ , a la matriz

$$-A = [-a_{ij}]$$

también de tamaño  $m \times n$ .

Es decir,  $-A$  es la matriz cuyos elementos son los opuestos de los correspondientes elementos de  $A$ .

El teorema siguiente recoge las propiedades más importantes de la adición de matrices, cuya demostración es simple.

**Teorema 1.2.1** Si  $A$ ,  $B$  y  $C$  son matrices de tamaño  $m \times n$ , entonces se satisfacen las propiedades siguientes:

1.  $(A + B) + C = A + (B + C)$  (propiedad asociativa)
2.  $A + B = B + A$  (propiedad conmutativa)
3.  $A + O = A = O + A$
4.  $A + (-A) = O = (-A) + A$

Podemos comprobar todas estas propiedades escribiendo los elementos de las matrices y utilizando la definición de adición de matrices junto con las propiedades de la adición de números reales.

La propiedad asociativa de la adición de matrices permite escribir  $A + B + C$  para denotar cualquiera de las sumas  $(A + B) + C$  o  $A + (B + C)$  tal como ocurre con los escalares y los vectores.

La introducción de la matriz opuesta nos permite definir la substracción de dos matrices en los siguientes términos.

**Definición 1.2.4** La **substracción** de las matrices  $A$  y  $B$  de tamaño  $m \times n$  es la matriz  $A - B$  también de tamaño  $m \times n$ , llamada **resta** de  $A$  y  $B$ , tal que

$$A - B = A + (-B).$$

Así, si  $A = [a_{ij}]$  y  $B = [b_{ij}]$ , entonces

$$A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}]$$

es la matriz que resulta de restar a cada elemento de  $A$  el correspondiente elemento de  $B$ .

Notemos que, de acuerdo con la definición anterior,  $A - B$  no define una nueva operación en  $M_{m \times n}$ ; se trata simplemente de una forma abreviada de escribir la adición de la matriz  $A$  con la opuesta de la matriz  $B$  (tal como ocurre con los escalares y los vectores).

Finalizamos esta sección con la siguiente propiedad cuya demostración es inmediata y se propone como ejercicio.

**Teorema 1.2.2** Si  $A$  y  $B$  son matrices de tamaño  $m \times n$ , entonces

$$(A + B)^T = A^T + B^T.$$

La matriz traspuesta se puede utilizar para definir un tipo muy importante de matrices.

**Definición 1.2.5** Decimos que una matriz  $A$  es **simétrica** si  $A^T = A$ . Decimos que una matriz  $A$  es **antisimétrica** si  $A^T = -A$ .

Evidentemente, la condición  $A^T = A$  o  $A^T = -A$  fuerza que  $A$  sea cuadrada.

**Ejemplo 1.2.2** Las matrices

$$\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & 4 \\ 0 & 4 & -5 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} -2 & 1 & 7 \\ 1 & 3 & 6 \\ 7 & 6 & -8 \end{bmatrix}$$

son simétricas, mientras que las matrices

$$\begin{bmatrix} 0 & -3 \\ 3 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & -2 & 5 \\ 2 & 0 & 7 \\ -5 & -7 & 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 & 6 & -5 \\ -6 & 0 & -3 \\ 5 & 3 & 0 \end{bmatrix}$$

son antisimétricas.

Notemos que si  $A$  es una matriz diagonal, entonces  $A$  es simétrica (¿por qué?).

### 1.3 Multiplicación de un escalar por una matriz

**Definición 1.3.1** Sea  $A = [a_{ij}]$  una matriz de tamaño  $m \times n$ . La **multiplicación del escalar  $\alpha$  por la matriz  $A$**  es la matriz  $\alpha A$  de tamaño  $m \times n$ , llamada **producto** de  $\alpha$  por  $A$ , cuyo elemento de la posición  $(i,j)$  es  $\alpha a_{ij}$ , es decir

$$\alpha A = [\alpha a_{ij}].$$

Así,  $\alpha A$  es la matriz del mismo tamaño que  $A$  cuyos elementos se obtienen multiplicando  $\alpha$  por los correspondientes elementos de  $A$  y, por tanto,  $M_{m \times n}$  es **cerrado** para la multiplicación por un escalar.

Notemos que la multiplicación de un escalar por una  $n$ -tupla (o vector) es un caso particular de la multiplicación de un escalar por una matriz.

**Ejemplo 1.3.1** Si

$$A = \begin{bmatrix} 5 & 4 & -1 \\ -2 & 8 & 7 \end{bmatrix}$$

entonces

$$3A = \begin{bmatrix} 3 \cdot 5 & 3 \cdot 4 & 3 \cdot (-1) \\ 3 \cdot (-2) & 3 \cdot 8 & 3 \cdot 7 \end{bmatrix} = \begin{bmatrix} 15 & 12 & -3 \\ -6 & 24 & 21 \end{bmatrix}.$$

El teorema siguiente recoge las propiedades más importantes de la multiplicación de un escalar por una matriz (comparad con las propiedades ??–?? del teorema ??).

**Teorema 1.3.1** Si  $A$  y  $B$  son matrices de tamaño  $m \times n$  y  $\alpha$  y  $\beta$  son escalares, entonces se satisfacen las propiedades siguientes:

1.  $\alpha(A + B) = \alpha A + \alpha B$
2.  $(\alpha + \beta)A = \alpha A + \beta A$
3.  $\alpha(\beta A) = (\alpha\beta)A$
4.  $1A = A$

Podemos comprobar todas estas propiedades escribiendo los elementos de las matrices y utilizando la definición de adición de matrices, multiplicación de un escalar por una matriz y las propiedades de la adición y multiplicación de números reales (¡hacedlo!).

Notemos que

$$OA = O = \alpha O$$

para cualquier matriz  $A$  y cualquier escalar  $\alpha$  con  $O$  la matriz nula del mismo tamaño que  $A$ .

Finalizamos esta sección con la siguiente propiedad cuya demostración es inmediata (¡hacela!).

**Teorema 1.3.2** *Si  $A$  es una matriz de tamaño  $m \times n$  y  $\alpha$  es un escalar, entonces*

$$(\alpha A)^T = \alpha A^T.$$

## 1.4 El producto de matrices

### Anécdotas matemáticas.

Una antiquísima leyenda cuenta que Sheram, príncipe de la India, quedó tan maravillado cuando conoció el juego del ajedrez, que quiso recompensar generosamente a Sessa, el inventor de aquel entretenimiento. Le dijo: *Pídeme lo que quieras*. Sessa le respondió: *Soberano, manda que me entreguen un grano de trigo por la primera casilla del tablero, dos por la segunda, cuatro por la tercera, ocho por la cuarta, y así sucesivamente hasta la casilla 64*.

El príncipe no pudo complacerle, porque el resultado de esa operación  $S = 1 + 2 + 4 + \dots + 2^{63}$  es aproximadamente 18 trillones de granos. Para obtenerlos habría que sembrar la Tierra entera 65 veces.

Circula por los círculos matemáticos un sorprendente final de la historia. Sheram, preocupado al haber empeñado su palabra, mandó llamar al matemático del reino, un tal Pepe Martínez Aroza, el cual razonó de la siguiente manera:

Alteza, puesto que no tenéis trigo suficiente para pagar la deuda contraída con Sessa, igual os daría deberle aún más. Sed, pues, magnánimo y aumentad vuestra recompensa a la cantidad  $S = 1 + 2 + 4 + 8 + \dots$  hasta el infinito. Observad que, a partir de la segunda casilla, todas las cantidades a sumar son pares, lo cual nos permite escribir  $S = 1 + 2 \times (1 + 2 + 4 + 8 + \dots)$ , o lo que es lo mismo,  $S = 1 + 2 \times S$ . Ahora, vos mismo podéis resolver esta sencilla ecuación de primer grado y, veréis que la única solución es  $S = -1$ . Podéis decir a Sessa que no solamente puede considerarse pagado con creces, ya que habéis aumentado enormemente vuestra recompensa, sino que actualmente os adeuda un grano de trigo.

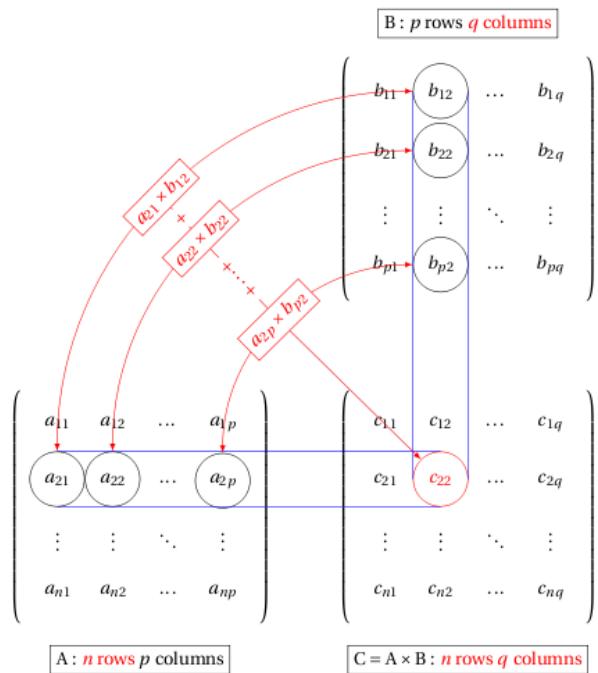
**Definición 1.4.1** Consideremos  $A = [a_{ij}]$  y  $B = [b_{ij}]$  dos matrices de tamaños  $m \times n$  y  $n \times p$ , respectivamente. La multiplicación de las matrices  $A$  y  $B$  es la matriz  $AB = [c_{ij}]$  de tamaño  $m \times p$ , llamada producto de  $A$  y  $B$ , tal que

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj},$$

para  $i = 1, 2, \dots, m$  y  $j = 1, 2, \dots, p$ .

Notemos, a partir de la definición anterior, que la condición para que pueda realizarse el producto de dos matrices es que el número de columnas de la primera coincida con el número de filas de la segunda.

La figura 1.2 muestran gráficamente el modo de multiplicar dos matrices.



**Figura 1.2** Representación gráfica del producto de dos matrices..

Veamos un ejemplo que clarifique la operación.

**Ejemplo 1.4.1** Dadas las matrices

$$A = \begin{bmatrix} 1 & 0 & 3 \\ -2 & 1 & 4 \\ 5 & -3 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & 4 \\ 5 & 0 \\ 6 & 1 \end{bmatrix}$$

calculamos  $AB$ .

Procedemos con el cálculo

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 0 & 3 \\ -2 & 1 & 4 \\ 5 & -3 & 0 \end{bmatrix} \begin{bmatrix} -2 & 4 \\ 5 & 0 \\ 6 & 1 \end{bmatrix} = \\ &\begin{bmatrix} 1(-2) + 0 \cdot 5 + 3 \cdot 6 & 1 \cdot 4 + 0 \cdot 0 + 3 \cdot 1 \\ -2 \cdot (-2) + 1 \cdot 5 + 4 \cdot 6 & -2 \cdot 4 + 1 \cdot 0 + 4 \cdot 1 \\ 5 \cdot (-2) - 3 \cdot 5 + 0 \cdot 6 & 5 \cdot 4 - 3 \cdot 0 + 0 \cdot 1 \end{bmatrix} = \begin{bmatrix} 16 & 7 \\ 33 & -4 \\ -25 & 21 \end{bmatrix}. \end{aligned}$$

Notemos que en el ejemplo que hemos visto anteriormente no podemos calcular el producto  $BA$  ya que  $B$  tiene dos columnas y  $A$  tiene tres filas. Esto nos lleva a una de las características más importantes del producto de matrices, como es que no cumple la propiedad conmutativa, es decir, en general, dadas dos matrices  $A$  y  $B$ , se cumple que  $AB \neq BA$ .

### Propiedades del producto de matrices

Como característica esencial del producto de matrices, cabe recordar que el producto de matrices no es, en general, conmutativo. Las propiedades más importantes del producto son:

1. **Propiedad asociativa:**  $(AB)C = A(BC)$ .
2. **Propiedad conmutativa:**  $AB \neq BA$ . No se cumple.
3.  $AI_n = I_nA = A$ , siendo  $I_n$  la matriz identidad de tamaño  $n$ .
4. **Propiedad distributiva respecto de la suma:**  $A(B+C) = AB+AC$ .

**Definición 1.4.2** Decimos que una matriz cuadrada  $A$  es **invertible** si existe una matriz cuadrada  $B$ , del mismo tamaño que  $A$ , tal que

$$AB = I = BA.$$

En tal caso, decimos que  $B$  es la **inversa** de  $A$  y escribimos  $B = A^{-1}$ .

A continuación enunciamos algunas propiedades fundamentales del concepto de inversa de una matriz cuadrada.

#### Teorema 1.4.1

1. Si  $A$  es una matriz invertible, entonces  $A^{-1}$  también lo es y

$$(A^{-1})^{-1} = A.$$

2. Si  $A$  es una matriz invertible, entonces  $A^T$  también lo es y

$$(A^T)^{-1} = (A^{-1})^T.$$

**3.** Si  $A$  es una matriz invertible y  $\alpha \neq 0$ , entonces  $\alpha A$  es invertible y

$$(\alpha A)^{-1} = (1/\alpha)A^{-1}.$$

**4.** Si  $A$  y  $B$  son matrices invertibles del mismo tamaño, entonces  $AB$  también lo es y

$$(AB)^{-1} = B^{-1}A^{-1}.$$

Aparentemente, el producto de matrices parece una operación sofisticada. Veamos un ejemplo en el que ponemos de manifiesto la utilidad de disponer de números en forma ordenada por filas y columnas.

Consideremos el problema que conocemos con el nombre de **relación de influencias** en una familia. Consideremos una familia formada por los siguientes miembros: Padre (P), Madre (M), Abuelo (Ab), Abuela (Aba), Hijo Mayor (HM), Hija (Ha), Hijo Pequeño (Ho).

El problema de las relaciones de influencia entre los miembros de esta familia consiste en determinar las personas que ejercen una influencia sobre otras. Por ejemplo, si el abuelo influye notablemente en las decisiones que toma el padre, podemos afirmar que ejerce una influencia sobre él. Intentamos modelizar esta situación por medio de la estructura de números que hemos definido en este capítulo y que llamamos matriz.

Para representar las relaciones de influencia en esta familia, disponemos a sus elementos, de forma ordenada, por filas y columnas, de la forma siguiente:

$$\begin{matrix} & (P) & (M) & (Ab) & (Aba) & (HM) & (Ha) & (Ho) \\ (P) & & & & & & & \\ (M) & & & & & & & \\ (Ab) & & & & & & & \\ (Aba) & & & & & & & \\ (HM) & & & & & & & \\ (Ha) & & & & & & & \\ (Ho) & & & & & & & \end{matrix}$$

Si, por ejemplo, el abuelo influye sobre el padre, pondremos un 1 en la fila correspondiente al abuelo (Ab) y la columna correspondiente al padre (P). Si miramos por columnas, tendremos los elementos de la familia que tienen una influencia sobre este elemento.

Construimos esta matriz completando sus elementos de acuerdo con el criterio establecido por filas donde el elemento  $ij = 1$  si existe una relación de influencia de  $i$  sobre  $j$ .

Supongamos que esta familia presenta las siguientes relaciones de influencia:

$$I = \begin{matrix} & \begin{matrix} (P) & (M) & (Ab) & (Aba) & (HM) & (Ha) & (Ho) \end{matrix} \\ \begin{matrix} (P) \\ (M) \\ (Ab) \\ (Aba) \\ (HM) \\ (Ha) \\ (HP) \end{matrix} & \left[ \begin{array}{ccccccc} 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right] \end{matrix}$$

Por ejemplo, si nos fijamos en la tercera fila (Ab), podemos ver las relaciones de influencia del abuelo. Ejerce influencias sobre el padre, la hija y el hijo pequeño. Una columna de unos significaría que todos ejercen influencia sobre el elemento que ocupa esa columna.

En esta matriz se tiene que  $I_{51} = 0$ , lo que significa que el hijo mayor no ejerce influencia sobre el padre. Supongamos que el hijo mayor quiere pedirle al padre que le aumente la paga semanal. Si plantea la cuestión directamente al padre, lo normal es que no consiga su propósito. Sin embargo, la cuestión que se plantea es si podría tener influencia sobre el padre a través de otra persona que sí tenga influencia sobre él. Es lo que se entiende por llegar de un elemento a otro mediante un salto. Esta es la cuestión que ahora nos ocupa.

Relacionando esta idea con la matriz construida  $I$ , debemos fijarnos en la fila del hijo mayor para saber dónde ejerce influencias y, a través de éstas, comprobar la posible influencia sobre el padre. El primer elemento que tenemos es  $I_{52} = 1$ , es decir, tiene una gran influencia sobre la madre. Ahora falta comprobar si la madre tiene influencia sobre el padre, es decir, nos fijamos en el elemento  $I_{21}$ , que es igual a 1, lo que significa que ya hemos encontrado una posible solución. Tenemos:

$$I_{52} = 1 + I_{21} = 1 \implies 5 \rightarrow 2 \rightarrow 1.$$

La pregunta que nos hacemos ahora es si ésta es la única solución o podemos encontrar otros elementos intermedios que nos conduzcan al destino final. La solución se basa en construir la matriz  $I^2$ , es decir, el producto de la matriz  $I$  por sí misma.

$$\begin{array}{c|cccccc|c|cccccc} & P & M & Ab & Aba & HM & Ha & Ho & P & M & Ab & Aba & HM & Ha & Ho \\ P & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ M & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ Ab & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ Aba & 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ HM & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ Ha & 1 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ HP & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{array}$$

Si hacemos el producto de la quinta fila (Abuelo) por la primera columna correspondiente al padre, obtenemos el elemento  $I_{51}$ , que viene determinado por

$$0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 + 1 \cdots 0 + 0 \cdot 1 + 1 \cdot 0 = 1.$$

Este número corresponde al elemento 51 de la matriz  $I^2$ , que es  $I_{51}^2 = 1$ . Este número es precisamente el que nos indica los caminos que hay de 5 a 1 a través de un salto. El único camino que tenemos en este caso es utilizar la influencia sobre la madre y la de la madre sobre el padre.

$$HM \longrightarrow M \longrightarrow P.$$

Calculamos la matriz  $I^2$

$$I^2 = \begin{bmatrix} P & M & Ab & Aba & HM & Ha & Ho \\ P & 2 & 1 & 2 & 1 & 2 & 1 & 2 \\ M & 3 & 2 & 3 & 2 & 3 & 2 & 1 \\ Ab & 3 & 1 & 4 & 1 & 2 & 2 & 2 \\ Aba & 2 & 4 & 3 & 4 & 2 & 2 & 4 \\ HM & 1 & 4 & 2 & 4 & 2 & 1 & 3 \\ Ha & 3 & 1 & 3 & 1 & 3 & 2 & 2 \\ HP & 2 & 3 & 3 & 3 & 1 & 2 & 3 \end{bmatrix}$$

que representa todas las influencias entre los miembros de esta familia con un salto, es decir, con una persona de intermediación. Notemos la utilidad de disponer un conjunto de números en forma ordenada, así como la utilidad de la operación del producto matricial de la forma en que se ha definido.

Análogamente, si construyéramos la matriz  $I^3$  tendríamos las relaciones de influencia entre los miembros de la familia con dos intermediarios. Así sucesivamente si siguiéramos con las sucesivas potencias de la matriz inicial.

Estudiemos otro ejemplo en que ponemos de manifiesto la importancia de trabajar con números de forma ordenada y cómo el producto de matrices puede resolver algunas situaciones reales sencillas.

---

**Ejemplo 1.4.2** La Universidad de Alicante dispone de cinco equipos de baloncesto: Filosofía y Letras (F), Derecho (D), Económicas (E), Ciencias (C) e Informática (I), que denotamos por  $F, D, E, C$  y  $I$  respectivamente. Cada cuatrimestre compiten en un torneo *todos contra todos* en el que cada equipo juega contra cada uno de los otros una sola vez.

Los resultados, en cuanto a ganar o perder en enfrentamientos directos, se pueden disponer en forma de una matriz ordenada donde por filas y columnas situamos ordenadamente los cinco equipos implicados en la competición. Así, por filas ponemos un 1 si el equipo de la fila  $i$ -ésima le ha ganado al equipo de la columna  $j$ -ésima.

Supongamos que el resultado de los partidos entre estos equipos se resumen en la matriz siguiente:

$$A = \begin{bmatrix} F & D & E & C & I \\ F & 0 & 1 & 0 & 1 & 0 \\ D & 0 & 0 & 1 & 0 & 0 \\ E & 1 & 0 & 0 & 0 & 0 \\ C & 0 & 1 & 1 & 0 & 1 \\ I & 1 & 1 & 1 & 0 & 0 \end{bmatrix}$$

Una forma de obtener la clasificación de los cinco equipos basándonos en los resultados de sus encuentros consiste en sumar el número de victorias de cada uno de ellos. Ahora bien, el número de victorias de cada equipo es precisamente la suma de los elementos que hay en la fila de la matriz de adyacencia correspondiente a cada equipo. Si denotamos por  $\mathbf{u}$  la  $n$ -tupla que tiene todas sus componentes iguales a 1, podemos calcular esta suma como  $A\mathbf{u}$ . En nuestro caso

$$A\mathbf{u} = \begin{bmatrix} F & D & E & C & I \\ F & 0 & 1 & 0 & 1 & 0 \\ D & 0 & 0 & 1 & 0 & 0 \\ E & 1 & 0 & 0 & 0 & 0 \\ C & 0 & 1 & 1 & 0 & 1 \\ I & 1 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \\ 3 \\ 3 \end{bmatrix}$$

que proporciona la clasificación siguiente:

- Primero: Ciencias, Informática (empatados)
- Segundo: Filosofía y Letras
- Tercero: Derecho, Económicas (empatados)

Surge ahora, de modo natural, la siguiente pregunta: ¿son los equipos que han quedado empatados igual de buenos? El equipo de Ciencias puede argumentar que se merece el primer lugar ya que derrotó al equipo de Informática. El equipo de Derecho podría utilizar el mismo argumento para romper el empate ya que éste derrotó al equipo de Económicas. Sin embargo, el equipo de Económicas podría argumentar que ellos tienen dos *victorias indirectas* ya que derrotaron al equipo de Filosofía y Letras (que a su vez derrotó al equipo de Derecho) y al equipo de Ciencias; además, puede hacer notar que el equipo de Derecho solamente tiene una *victoria indirecta* ya que derrotó al equipo de Económicas y éste al de Filosofía y Letras.

Puesto que las victorias indirectas corresponden a trayectorias de longitud 2 en el grafo dirigido y teniendo en cuenta que el número de estas trayectorias viene dado por  $A^2$ , para calcular tanto las victorias como las victorias indirectas de cada equipo necesitamos las

sumas de las filas de  $A + A^2$ , es decir,  $(A + A^2)\mathbf{u}$ . Así,

$$\begin{aligned}
 (A + A^2)\mathbf{u} &= \left( \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & 2 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 2 & 1 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \end{bmatrix} \right) \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} 0 & 2 & 2 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 2 & 3 & 0 & 1 \\ 2 & 2 & 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\
 &= \begin{bmatrix} 6 \\ 2 \\ 3 \\ 8 \\ 7 \end{bmatrix},
 \end{aligned}$$

con lo que

$$\begin{array}{ll}
 F & \begin{bmatrix} 6 \\ 2 \\ 3 \\ 8 \\ 7 \end{bmatrix} \\
 D & \\
 E & \\
 C & \\
 I &
 \end{array}$$

tenemos la siguiente clasificación de los equipos: Ciencias, Informática, Filosofía y Letras, Económicas y Derecho. Desafortunadamente, este enfoque no garantiza que se rompan los empates.

## 1.5 Aplicación del producto matriz-vector

Partimos del planteamiento inicial de un problema: la compra de una motocicleta. Realmente no parece que esto suponga un problema para aquellas personas muy decididas y poco dubitativas; sin embargo, existe un buen número de personas que, ante el hecho de tener que decidir por un determinado artículo cuando tienes la posibilidad de elegir entre una gama más o menos extensa del mismo, sufren una serie de indecisiones y cambios repentinos de opinión que hacen de su decisión un suplicio.

Tampoco parece que sea un problema en el que el álgebra matricial tenga mucho que decir. Veamos que esto no es así.

Planteamiento del problema. El problema ya lo conocemos, debemos elegir una moto para comprarnos, de entre una serie de modelos a la venta en el mercado. El tipo de moto elegida es la scouter y su cilindrada será al menos de 90 centímetros cúbicos.

La primera idea que se nos ocurre ante esta elección es la de comprar revistas, ver vídeos publicitarios, anuncios, preguntar en las tiendas, ver fotos, etc. Pero todo esto resulta un tanto desordenado. Parece claro que lo ideal quizás sería tratar el asunto desde el punto de vista numérico. Podemos tener en cuenta aquellos factores o características que consideremos importantes en la elección de una moto y tratar de asignarles valores cuantitativos, de manera que eso nos ayude a establecer una clasificación. En dicha clasificación, el modelo que ocupara la posición con una mayor puntuación sería la que en su conjunto presentara unas mejores características de acuerdo con nuestros gustos particulares.

Este es precisamente el objetivo: ser capaces de establecer una clasificación numérica de los modelos que más se acercan a nuestras necesidades y nuestros gustos. Antes de plantearnos un posible método de resolución siguiendo este esquema, nos planteamos ciertas dudas sobre la posibilidad de establecer unos criterios comunes determinados para todas las personas que se enfrenten ante el mismo problema. Por ejemplo,

- ¿Qué factores o características de una moto considero más importantes a la hora de mi elección?
- ¿Considero todas las características igualmente importantes?
- ¿Qué ocurre si los factores que yo considero importantes no coinciden con los del resto de la gente?
- ¿Cómo puedo cuantificar una característica cualitativa para establecer un orden numérico?

Estas y algunas otras cuestiones nos surgen al reflexionar sobre este problema. Aunque parece ciertamente complejo, podemos desarrollar un método de resolución.

La idea sobre la que se basa el método que proponemos es considerar una matriz de datos con los que vamos a trabajar. Por un lado, tendremos un determinado número de modelos de motocicletas; por otro lado, tendremos un conjunto de características que debemos cuantificar. El problema de la cuantificación será tratado posteriormente con mayor detenimiento.

Una vez dispongamos de esa matriz de datos, estableceremos los factores de corrección de cada característica. Con todo esto, mediante el producto de matrices obtendremos la solución final. Los factores de corrección tienen la función de ponderar las distintas características asociadas a los diferentes modelos analizados.

### Toma de datos

Supongamos que tienes la posibilidad de comprarte una de estas motos que tienes en la siguiente lista:

1. Aprilia Leonardo 125
2. Honda CH 125 Spacy
3. Honda Yupy NH 90
4. Honda SJ 100 Bali
5. Honda Scoopy SH100
6. Honda CN 250 Spazio 10

7. Kymco Heroism 125
8. Kymco Movie 125
9. Italjet Formula 125
10. PGO Galaxy 90
11. Peugeot SV 125 L
12. Piaggio Typhoon 125
13. Piaggio Sfera 125
14. Piaggio Vespa 125 ET/A
15. Suzuki AN 125 Vecstar

Ahora debes llenar algunas cuestiones de forma individual para completar la actividad.

1. Enumera los factores o características que consideres más importantes a la hora de tomar la decisión de elegir una motocicleta.
2. Si consideras que no todos los factores tienen la misma importancia, asigna a cada característica un factor, dependiendo de la importancia que le asignes. Cuanto mayor sea el factor, significa que mayor importancia le otorgas.

Ahora debemos seleccionar los factores más importantes de la lista. Una vez seleccionados, daremos una clasificación-puntuación a cada una de los modelos respecto de la característica que estamos estudiando.

Es decir, si estamos estudiando, por ejemplo, la característica **frenada**, tendremos que establecer una puntuación a cada motocicleta, donde la máxima puntuación será para la moto que mejor frenada posea, y así en orden decreciente.

Con esto, podremos establecer una matriz de datos, donde representamos por filas cada uno de los modelos de motocicletas y por columnas representamos cada una de las características que analizamos.

Análogamente, formamos un vector columna donde en cada posición aparece el factor de importancia que se le ha asignado a cada característica. El producto de la matriz de datos anterior por el vector con los factores de importancia nos proporciona una clasificación para los diferentes modelos donde el de mayor puntuación sería la primera opción y así sucesivamente las siguientes opciones.

Veamos un ejemplo concreto.

---

**Ejemplo 1.5.1** Disponemos de una lista de 15 motocicletas sobre las que se analizaron las siguientes características:

$C_1$  = Cilindrada  $C_2$  = Potencia  $C_3$  = Peso  $C_4$  = Depósito  $C_5$  = Frenada  $C_6$  = Estética  $C_7$  = Precio.

El factor de importancia que asignó a cada una de las características anteriores fue:

Cilindrada = 1.25; Potencia = 1.35; Peso = 1.1; Depósito = 1; Frenada = 1.5; Estética = 1.65; Precio = 2.

**Tabla 1.1** Puntuaciones de cada motocicleta respecto de las características evaluadas..

modelo	$C_1$	$C_2$	$C_3$	$C_4$	$C_5$	$C_6$	$C_7$
Aprilia Leonardo	4	6	6	5	7	4	3
Honda CH 125	4	3	5	3	1	2	4
Honda Yupy NH 90	1	1	2	2	4	1	6
Honda SJ 100 Bali	2	2	2	2	4	2	6
Honda Scoopy	3	2	2	1	5	1	6
Honda CN 250 Spazio	5	7	7	6	4	4	1
Kymco Heroism 125	4	3	4	1	3	4	5
Kymco Movie 125	4	3	4	2	3	3	5
Italjet Formula 125	4	5	4	5	6	5	2
PGO Galaxy 90	1	1	1	1	2	2	7
Peugeot SV 125 L	4	3	2	4	4	2	6
Piaggio Typhoon 125	4	4	2	3	1	3	6
Piaggio Sfera 125	4	4	3	3	5	2	5
Piaggio Vespa 125	4	4	4	4	5	1	4
Suzuki AN 125 Vecstar	4	4	3	3	2	4	4

Consultando comparativas de revistas especializadas en motocicletas, se estableció la tabla 1.1 donde se recogen las puntuaciones que cada modelo ha conseguido al valorar cada una de las características anteriormente reseñadas.

En consecuencia la matriz de datos es

$$M_d = \begin{bmatrix} 4 & 6 & 6 & 5 & 7 & 4 & 3 \\ 4 & 3 & 5 & 3 & 1 & 2 & 4 \\ 1 & 1 & 2 & 2 & 4 & 1 & 6 \\ 2 & 2 & 2 & 2 & 4 & 2 & 6 \\ 3 & 2 & 2 & 1 & 5 & 1 & 6 \\ 5 & 7 & 7 & 6 & 4 & 4 & 1 \\ 4 & 3 & 4 & 1 & 3 & 4 & 5 \\ 4 & 3 & 4 & 2 & 3 & 3 & 5 \\ 4 & 5 & 4 & 5 & 6 & 5 & 2 \\ 1 & 1 & 1 & 1 & 2 & 2 & 7 \\ 4 & 3 & 2 & 4 & 4 & 2 & 6 \\ 4 & 4 & 2 & 3 & 1 & 3 & 6 \\ 4 & 4 & 3 & 3 & 5 & 2 & 5 \\ 4 & 4 & 4 & 4 & 5 & 1 & 4 \\ 4 & 4 & 3 & 3 & 2 & 4 & 4 \end{bmatrix}$$

El vector con el factor multiplicativo es  $\mathbf{v}$ ,

$$[1.25 \quad 1.35 \quad 1.10 \quad 1.00 \quad 1.50 \quad 1.65 \quad 2.00]^T$$

Por lo que si multiplicamos  $M_d \cdot \mathbf{v}$  tendremos

$$\begin{bmatrix} 47.80 \\ 30.35 \\ 26.45 \\ 30.70 \\ 30.80 \\ 44.00 \\ 35.55 \\ 34.90 \\ 42.40 \\ 25.00 \\ 36.55 \\ 34.05 \\ 37.50 \\ 35.95 \\ 34.30 \end{bmatrix}$$

Este vector representa las puntuaciones obtenidas por cada modelo. Observamos que queda como ganador el modelo primero con 47.80 puntos, lo que significa que esa sería nuestra mejor elección.

## 1.6 Ejercicios

**1.1** Hallar todas las matrices  $2 \times 2$  que comuten

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

**1.2** Justifica que no es cierta la igualdad:  $(A + B)(A - B) = A^2 - B^2$  siendo  $A$  y  $B$  dos matrices cualesquiera.

**1.3** Encuentra la matriz  $A$  de la expresión

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

**1.4** Calcular el producto  $A \times B \times C$  con

$$A = \begin{pmatrix} 1 & 0 \\ -1 & 2 \\ 3 & 1 \end{pmatrix}, B = \begin{pmatrix} 0 & 2 \\ 1 & -1 \end{pmatrix} \text{ y } C = \begin{pmatrix} 2 \\ -22 \\ 41 \end{pmatrix}.$$

**1.5** Halla los valores de  $m$  para los cuales

$$X = \begin{pmatrix} m & 0 \\ 0 & 2 \end{pmatrix}$$

verifique

$$X^2 - \frac{5}{2}X + I = 0$$

**1.6** Calcula  $A^n$  y  $B^n$  siendo

$$A = \begin{pmatrix} 1 & 1/7 & 1/7 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ y } B = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

**1.7** Dada la matriz

$$A = \begin{pmatrix} 3 & 0 & 8 \\ 3 & -1 & 6 \\ -2 & 0 & -5 \end{pmatrix}$$

comprueba que  $(A + I)^2 = 0$  y expresa  $A^2$  como combinación lineal de  $A$  e  $I$ .

**1.8** Halla las matrices  $X$  e  $Y$  que verifican el sistema

$$2X + Y = \begin{pmatrix} 1 & 4 \\ 2 & 0 \end{pmatrix}, X - Y = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}.$$

**1.9** Una empresa de automóviles dispone de dos plantas de fabricación localizadas una en España y otra en Gran Bretaña. En ellas se fabrica dos modelos de coches  $M_1$  y  $M_2$ , en tres colores  $x, y, z$ . Su capacidad de producción diaria en cada planta está dada por las siguientes matrices:

$$A = \begin{bmatrix} 300 & 95 \\ 250 & 100 \\ 200 & 100 \end{bmatrix} \quad B = \begin{bmatrix} 190 & 90 \\ 200 & 100 \\ 150 & 80 \end{bmatrix}.$$

1. Determina la representación matricial de la producción total por día.
2. Si se eleva la producción en España un 20% y se disminuye en Gran Bretaña un 10% ¿qué matriz representa la nueva producción total?

**1.10** Suponga que toda la industria de refresco produce dos colas: Coca Cola y Pepsi Cola. Cuando una persona ha comprado Coca Cola hay una probabilidad de 90% de que siga comprándola la vez siguiente. Si una persona compró Pepsi, hay 80% de que repita la vez siguiente. Se pide: a) Si una persona actualmente es comprador de Pepsi. ¿Cuál es la probabilidad de que compre Coca Cola pasadas dos compras a partir de hoy? b) Si en la actualidad una persona es comprador de Coca Cola. ¿Cuál es la probabilidad de que compre Coca Cola pasadas tres compras a partir de ahora? c) Supongamos que el 60% de toda la gente toma hoy Coca Cola y el 40% Pepsi. A tres compras a partir de ahora, ¿Qué fracción de los compradores estará tomando Coca Cola. d) Determinar el estado estable.

- 1.11** En una comunidad hay 3 supermercados ( $S_1, S_2, S_3$ ) existe la movilidad de un cliente de uno a otro. El 1 de septiembre,  $\frac{1}{4}$  de los clientes va al  $S_1$ ,  $\frac{1}{3}$  al  $S_2$  y  $\frac{5}{12}$  al  $S_3$  de un total de 10,000 personas. Cada mes el  $S_1$  retiene el 90 % de sus clientes y pierde el 10 % que se va al  $S_2$ . Se averiguó que el  $S_2$  solo retiene el 5 % y pierde el 85 % que va a  $S_1$  y el resto se va a  $S_3$ , el  $S_3$  retiene solo el 40 %, pierde el 50 % que va al  $S_1$  y el 10 % va al  $S_2$ .

1. Establecer la matriz de transición.
2. ¿Cuál es la proporción de clientes para los supermercados el 1 de noviembre?
3. Halle el vector de probabilidad estable.

- 1.12** Dos empresas A y B suministran televisión por cable en una ciudad de 100,000 hogares. El movimiento de suscripciones entre ellas varía de acuerdo con los siguientes parámetros: El 70 % de los abonados a A se quedan en la compañía al año siguiente; el 80 % de abonados de B siguen en B al próximo año; el 70 % de los hogares sin televisión por cable no contrata con ninguna compañía. El 20 % de abonados en A se pasan a B al año siguiente, mientras que el 15 % de abonados en B se pasan a A. El 10 % de abonados en A y el 5 % de abonados en B van a cancelar su televisión por cable (pasan a televisión TDT normal) el próximo año. De los hogares sin TV por cable (solo TDT), el 15 % contrata al año siguiente con A y el 15 % contrata con B. Actualmente A cuenta con 15,000 abonados y B con 20,000. a) Representa la matriz de transición de este problema. ¿Cuál es el vector inicial de estado? b) Cuántos abonados tendrá cada empresa el año próximo? ¿Y dentro de 3 años?

- 1.13** Resuelve la ecuación matricial  $XA = B + C$ , siendo

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix},$$

¿Cómo puedes asegurar que existe solución (antes de resolverla)?

- 1.14** Tres empresas  $E_1, E_2, E_3$ , necesitan cuatro materias primas  $P_1, P_2, P_3, P_4$ . El consumo mensual medio (en Tm.) de estas empresas se puede expresar mediante la matriz siguiente:

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

En el primer trimestre del año 2001, los precios de estas materias primas, expresados en euros por Tm., han sido

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

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.

# **2 Resolución de sistemas de ecuaciones lineales con matrices**

---

Los sistemas de ecuaciones lineales comenzaron a ser estudiados sistemáticamente por Leibniz y Cramer a mediados del siglo XVIII. Cramer expuso lo que hoy conocemos como regla de Cramer para los sistemas de orden 3. A mediados del siglo XIX fue Cayley, al estudiar las matrices, quien expuso la condición necesaria y suficiente para que un sistema cuadrado de ecuaciones lineales tuviera una solución única, basándose en la invertibilidad de la matriz de coeficientes.

Fue Frobenius quien introdujo la noción de rango de una matriz en 1879, aunque relacionado con los determinantes. Gauss había deducido a principios del siglo XIX su célebre método que resuelve cualquier sistema de ecuaciones lineales. Jordan dedujo un algoritmo para calcular la inversa de una matriz, dando origen al algoritmo que hoy utilizamos y llamamos Gauss-Jordan.

## **2.1 Introducción**

Alrededor de un 75 % de los problemas matemáticos que aparecen en el estudio de aplicaciones científicas o técnicas (de las que solamente veremos unas pocas a lo largo del curso), necesitan resolver, en alguna de sus etapas, un sistema de ecuaciones lineales. De ahí que éste sea uno de los problemas fundamentales de las matemáticas, cuya aplicación abarca todos los ámbitos de la misma.

**Definición 2.1.1** Llamamos **sistema lineal** de  $m$  ecuaciones en las **incógnitas** o **variables**  $x_1, x_2, \dots, x_n$  a un conjunto de  $m$  ecuaciones lineales

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{array} \right\} \quad (2.1)$$

donde los **coeficientes**  $a_{ij}$ , para  $i = 1, 2, \dots, m$  y  $j = 1, 2, \dots, n$ , y los **términos independientes**  $b_i$ , para  $i = 1, 2, \dots, m$ , son números conocidos.

**Ejemplo 2.1.1** El sistema lineal

$$\left. \begin{array}{rcl} x_1 + 2x_2 - x_3 + x_4 & = & 3 \\ -x_1 + x_2 + 4x_3 + 7x_4 & = & -5 \end{array} \right\}$$

tiene 2 ecuaciones y 4 incógnitas. En cambio, el sistema lineal

$$\left. \begin{array}{rcl} 2x_1 - x_2 + 5x_3 & = & 0 \\ -x_1 + 2x_3 & = & 2 \\ x_1 + 3x_2 - x_3 & = & -3 \end{array} \right\}$$

tiene 3 ecuaciones y 3 incógnitas; en la segunda ecuación de este sistema lineal no hemos escrito la incógnita  $x_2$  porque su coeficiente es 0. Finalmente,

$$x_1 - 3x_2 + 2x_3 = 6$$

es un sistema lineal con una ecuación y 3 incógnitas.

Las ecuaciones

$$x_1x_2 + 4x_3 = 1, \quad \frac{x_1}{x_2} + x_3 = 5 \quad \text{y} \quad 2x_1 - \cos x_2 + 5x_3 = \sqrt{2}$$

no son lineales.

Por tanto, las ecuaciones lineales no contienen productos, inversos u otras funciones de las variables; éstas se presentan únicamente en su primera potencia y están multiplicadas sólo por escalares.

**Ejemplo 2.1.2** En  $\mathbb{R}^2$  la ecuación general de una recta es de la forma

$$ax + by = c$$

con  $a, b, c \in \mathbb{R}$ ; por tanto, la ecuación de una recta en  $\mathbb{R}^2$  es una ecuación lineal.

En  $\mathbb{R}^3$  la ecuación general de un plano es de la forma

$$ax + by + cz = d$$

con  $a,b,c,d \in \mathbb{R}$ ; por tanto, la ecuación de un plano en  $\mathbb{R}^3$  es también una ecuación lineal. En cambio, un sistema de dos ecuaciones con tres incógnitas

$$\left. \begin{array}{l} a_1x + b_1y + c_1z = d_1 \\ a_2x + b_2y + c_2z = d_2 \end{array} \right\}$$

en el que los vectores  $(a_1, b_1, c_1)$  y  $(a_2, b_2, c_2)$  no son proporcionales, representa, como veremos más adelante, una recta en  $\mathbb{R}^3$ .

**Ejemplo 2.1.3** Consideremos el sistema lineal

$$\left. \begin{array}{rclcrclcrcl} 3x_1 & + & x_2 & - & 2x_3 & + & 2x_4 & = & 3 \\ -2x_1 & - & 4x_2 & - & x_3 & + & 2x_4 & = & 1 \\ & & x_2 & + & x_3 & - & x_4 & = & 0 \\ x_1 & + & x_2 & + & x_3 & + & x_4 & = & 6 \end{array} \right\}$$

Si tomamos  $x_1 = 2$ ,  $x_2 = -1$ ,  $x_3 = 3$  y  $x_4 = 2$ , entonces

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

$$(-1) + 3 - 2 = 0 \quad \text{y} \quad 2 + (-1) + 3 + 2 = 6,$$

con lo que los números  $x_1 = 2$ ,  $x_2 = -1$ ,  $x_3 = 3$  y  $x_4 = 2$  satisfacen las cuatro ecuaciones del sistema lineal.

Resolver un sistema lineal consiste en determinar el conjunto de valores de las incógnitas que satisfacen todas las ecuaciones del sistema. Si el sistema tiene  $n$  incógnitas debemos obtener  $n$  números, uno por cada incógnita. Así, en el ejemplo anterior, hemos determinado 4 valores porque el sistema tenía 4 incógnitas. En general, será útil expresar los  $n$  números de la solución de forma conjunta mediante una  $n$ -tupla o vector de  $\mathbb{R}^n$ .

**Definición 2.1.2** Decimos que una  $n$ -tupla  $\mathbf{u} = (u_1, u_2, \dots, u_n)$  es una **solución** del sistema lineal (2.1) si es solución de cada una de las ecuaciones lineales que constituyen dicho sistema. Es decir, si al sustituir  $x_i$  por  $u_i$ , para  $i = 1, 2, \dots, n$ , se satisfacen todas las ecuaciones del sistema.

**Ejemplo 2.1.4** La terna  $(4, 2, 1)$  es una solución del sistema lineal

$$\left. \begin{array}{rclcrclcrcl} x_1 & - & x_2 & + & 2x_3 & = & 4 \\ x_1 & + & 2x_2 & - & x_3 & = & 7 \\ x_1 & - & x_2 & + & x_3 & = & 3 \end{array} \right\}$$

ya que

$$4 - 2 + 2 \cdot 1 = 4, \quad 4 + 2 \cdot 2 - 1 = 7 \quad \text{y} \quad 4 - 2 + 1 = 3.$$

En cambio, la terna  $(7, -1, -2)$  no es una solución de dicho sistema ya que

$$7 - (-1) + 2 \cdot (-2) = 4 \quad \text{y} \quad 7 + 2 \cdot (-1) - (-2) = 7$$

pero

$$7 - (-1) + (-2) = 6 \neq 3.$$

## 2.2 Resolución de sistemas lineales

Veamos el siguiente ejemplo.

**Ejemplo 2.2.1** Consideremos el sistema lineal

$$\left. \begin{array}{rcl} x_1 - 3x_2 + 4x_3 & = & 2 \\ x_2 - 3x_3 & = & -2 \\ x_3 & = & 1 \end{array} \right\}$$

De la última ecuación tenemos que  $x_3 = 1$ ; sustituyendo dicho valor en la segunda ecuación obtenemos

$$x_2 - 3 \cdot 1 = -2$$

con lo que  $x_2 = 1$ . Ahora, sustituyendo los valores de  $x_3$  y  $x_2$  en la primera ecuación resulta

$$x_1 - 3 \cdot 1 + 4 \cdot 1 = 2$$

con lo que  $x_1 = 1$ .

Así pues, la solución del sistema es la terna  $(1, 1, 1)$ .

Llamamos **sustitución hacia atrás** al procedimiento empleado en la resolución del sistema lineal del ejemplo anterior.

**Ejemplo 2.2.2** Consideremos los sistemas lineales

$$\left. \begin{array}{rcl} x_1 - 4x_2 & = & 11 \\ 3x_1 - 2x_2 & = & 13 \end{array} \right\} \quad \text{y} \quad \left. \begin{array}{rcl} x_1 - 4x_2 & = & 11 \\ x_2 & = & -2 \end{array} \right\}$$

cuya primera ecuación es común a ambos sistemas.

Notemos que el segundo sistema es fácil de resolver por sustitución hacia atrás. De la segunda ecuación tenemos que  $x_2 = -2$  y sustituyendo en la primera ecuación de dicho sistema tenemos que

$$x_1 - 4(-2) = 11$$

con lo que  $x_1 = 3$ . Por tanto el par  $(3, -2)$  es una solución del segundo sistema. Además, es evidente que ésta es la única solución del mismo.

Puesto que

$$3 - 4(-2) = 11 \quad \text{y} \quad 3 \cdot 3 - 2(-2) = 13$$

tenemos que el par  $(3, -2)$  es también solución del primer sistema.

Supongamos ahora que el par  $(u_1, u_2)$  es una solución del primer sistema; entonces

$$u_1 - 4u_2 = 11 \quad \text{y} \quad 3u_1 - 2u_2 = 13$$

con lo que

$$3(u_1 - 4u_2) - (3u_1 - 2u_2) = 3 \cdot 11 - 13$$

es decir

$$-10u_2 = 20$$

y, simplificando,  $u_2 = -2$ . Por tanto, el par  $(u_1, u_2)$  también es solución del segundo sistema y, en consecuencia,  $(u_1, u_2) = (3, -2)$ , con lo que ésta es también la única solución del primer sistema.

Como hemos visto en el ejemplo anterior, existen sistemas lineales que tienen el mismo conjunto de soluciones, siendo uno de ellos más fácil de resolver que el otro.

**Definición 2.2.1** Decimos que dos sistemas lineales son **equivalentes** si ambos sistemas tienen el mismo conjunto de soluciones.

Seguidamente desarrollamos un procedimiento para resolver un sistema lineal que consiste en transformar el sistema lineal dado en otro equivalente cuya solución sea fácil de obtener; el patrón “triangular” o “escalonado” del ejemplo 2.2.1 (en el que la segunda ecuación tiene una incógnita menos que la primera y la tercera ecuación una incógnita menos que la segunda), es lo que intentaremos conseguir. La justificación del método la proporciona el teorema siguiente.

**Teorema 2.2.1** Consideremos el sistema lineal

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \end{array} \right\} \quad (2.2)$$

**1. El sistema lineal**

$$\left. \begin{array}{l} a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \end{array} \right\} \quad (2.3)$$

es equivalente al sistema lineal (2.2).

**2. Para cualquier escalar  $\alpha \neq 0$ , el sistema lineal**

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ \alpha a_{21}x_1 + \alpha a_{22}x_2 + \cdots + \alpha a_{2n}x_n = \alpha b_2 \end{array} \right\} \quad (2.4)$$

es equivalente al sistema lineal (2.2).

3. Para cualquier escalar  $\beta$ , el sistema lineal

$$\left. \begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ (a_{21} + \beta a_{11})x_1 + \cdots + (a_{2n} + \beta a_{1n})x_n = (b_2 + \beta b_1) \end{array} \right\} \quad (2.5)$$

es equivalente al sistema lineal (2.2).

No desarrollamos la demostración, quedando esta como ejercicio.

**Ejemplo 2.2.3** Consideremos el sistema lineal

$$\left. \begin{array}{l} x_1 - 3x_2 + 4x_3 = 2 \\ -2x_1 + 6x_2 + x_3 = 5 \\ x_1 - 2x_2 + x_3 = 0 \end{array} \right\}$$

Si sumamos a la segunda ecuación la primera ecuación multiplicada por 2, obtenemos el sistema equivalente

$$\left. \begin{array}{l} x_1 - 3x_2 + 4x_3 = 2 \\ 9x_3 = 9 \\ x_1 - 2x_2 + x_3 = 0 \end{array} \right\}$$

Si restamos a la tercera ecuación la primera, es decir, sumamos a la tercera ecuación la primera multiplicada por  $(-1)$ , obtenemos el sistema equivalente

$$\left. \begin{array}{l} x_1 - 3x_2 + 4x_3 = 2 \\ 9x_3 = 9 \\ x_2 - 3x_3 = -2 \end{array} \right\}$$

Si intercambiamos la posición de las dos últimas ecuaciones, obtenemos el sistema equivalente

$$\left. \begin{array}{l} x_1 - 3x_2 + 4x_3 = 2 \\ x_2 - 3x_3 = -2 \\ 9x_3 = 9 \end{array} \right\}$$

Finalmente, si dividimos la tercera ecuación por 9, es decir, multiplicamos la tercera ecuación por  $1/9$ , obtenemos el sistema equivalente

$$\left. \begin{array}{l} x_1 - 3x_2 + 4x_3 = 2 \\ x_2 - 3x_3 = -2 \\ x_3 = 1 \end{array} \right\}$$

En el ejemplo 2.2.1 hemos obtenido, por sustitución hacia atrás, que la solución del sistema lineal anterior es la terna  $(1,1,1)$ . Por tanto, de acuerdo con el teorema 2.2.1, la terna  $(1,1,1)$  es también la solución del sistema original.

Desde el punto de vista geométrico, el ejemplo anterior pone de manifiesto que los planos

$$x_1 - 3x_2 + 4x_3 = 2, \quad -2x_1 + 6x_2 + x_3 = 5 \quad \text{y} \quad x_1 - 2x_2 + x_3 = 0$$

de  $\mathbb{R}^3$  se cortan en el punto  $(1,1,1)$ .

De acuerdo con el teorema 2.2.1, y tal como hemos visto en el ejemplo 2.2.3, para resolver un sistema lineal aplicamos una sucesión de **operaciones elementales** sobre las ecuaciones del sistema inicial consistentes en:

- intercambiar la posición de dos ecuaciones (**operación elemental del tipo I**),
- multiplicar una ecuación por un escalar distinto de cero (**operación elemental del tipo II**),
- sumar a una ecuación otra ecuación multiplicada por un escalar (**operación elemental del tipo III**),

hasta obtener un sistema que se pueda resolver por sustitución hacia atrás o que no tenga solución.

Si observamos detenidamente el ejemplo 2.2.3, vemos que las operaciones elementales que hemos realizado hasta obtener el sistema equivalente del ejemplo 2.2.1 (que hemos resuelto por sustitución hacia atrás), son, en definitiva, operaciones sobre los coeficientes y términos independientes de las ecuaciones y no sobre las incógnitas. Por tanto, podemos ahorrarnos algo de escritura si escribimos los coeficientes y términos independientes de forma matricial como

$$\left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ -2 & 6 & 1 & 5 \\ 1 & -2 & 1 & 0 \end{array} \right]$$

donde cada fila representa una ecuación, de manera que las tres primeras posiciones contienen los coeficientes de las incógnitas en orden y la última posición contiene el término independiente; la línea vertical sirve para recordarnos los signos de igualdad de las ecuaciones. Llamamos **matriz ampliada del sistema** a dicha matriz. Análogamente, llamamos **matriz de coeficientes del sistema** a la matriz

$$\left[ \begin{array}{ccc} 1 & -3 & 4 \\ -2 & 6 & 1 \\ 1 & -2 & 1 \end{array} \right]$$

formada solamente por los coeficientes de las incógnitas (sin los términos independientes).

Podemos escribir la sucesión de operaciones elementales necesarias para obtener, a partir del sistema lineal del ejemplo 2.2.3 el sistema lineal equivalente del ejemplo 2.2.1,

como

$$\begin{array}{c}
 \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ -2 & 6 & 1 & 5 \\ 1 & -2 & 1 & 0 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 + 2F_1} \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 0 & 9 & 9 \\ 1 & -2 & 1 & 0 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 - F_1} \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 0 & 9 & 9 \\ 0 & 1 & -3 & -2 \end{array} \right] \\
 \xrightarrow{F_2 \leftrightarrow F_3} \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 9 & 9 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow (1/9)F_3} \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right]
 \end{array}$$

En la anterior sucesión de matrices hemos introducido una notación para las operaciones elementales que explicamos brevemente. Por  $F_k$  denotamos la  $k$ -ésima fila de la matriz; a la izquierda de la flecha escribimos la única fila que modificamos y a la derecha escribimos la operación elemental que realizamos sobre las filas de la matriz anterior para obtener la nueva matriz. Esta notación es bastante usual en informática y se interpreta como *almacenar en el lugar indicado por la flecha, la operación descrita a la derecha de la misma*.

Así, la primera de ellas  $F_2 \leftarrow F_2 + 2F_1$  significa que la segunda fila de la nueva matriz es el resultado de sumar a la segunda fila de la matriz anterior, la primera fila multiplicada por 2. Las operaciones correspondientes al intercambio de dos filas tienen una notación ligeramente diferente:  $F_i \leftrightarrow F_j$  significa que debemos intercambiar las filas  $i$ -ésima y  $j$ -ésima de la matriz anterior para obtener la nueva matriz.

El tío Petros y la conjectura de Goldbach es una novela de 1992 escrita por el griego Apostolos Doxiadis. El anciano tío Petros vive retirado de la vida social y familiar, entregado al cuidado de su jardín y a la práctica del ajedrez. Su sobrino, sin embargo, descubre motivado por la curiosidad que el tío Petros fue un matemático eminentemente, profesor en Alemania e Inglaterra, niño prodigo en esta disciplina y estudioso totalmente absorto en sus investigaciones científicas. Como irá descubriendo el sobrino, y el lector con él, la vida de Petros Papachristos ha girado durante años en torno a la famosa conjectura de Goldbach, un problema en apariencia sencillo, pero que durante más de dos siglos nadie ha conseguido resolver.

**Ejemplo 2.2.4** Consideremos el sistema lineal

$$\left. \begin{array}{rclclclcl} x_1 & - & 3x_2 & + & 3x_3 & + & x_4 & = & 1 \\ x_1 & + & x_2 & - & x_3 & + & x_4 & = & 5 \\ x_1 & - & x_2 & + & x_3 & + & x_4 & = & -3 \end{array} \right\}$$

Mediante operaciones elementales sobre la matriz ampliada tenemos que

$$\begin{array}{c}
 \left[ \begin{array}{cccc|c} 1 & -3 & 3 & 1 & 1 \\ 1 & 1 & -1 & 1 & 5 \\ 1 & -1 & 1 & 1 & -3 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - F_1} \left[ \begin{array}{cccc|c} 1 & -3 & 3 & 1 & 1 \\ 0 & 4 & -4 & 0 & 4 \\ 1 & -1 & 1 & 1 & -3 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 - F_1} \left[ \begin{array}{cccc|c} 1 & -3 & 3 & 1 & 1 \\ 0 & 4 & -4 & 0 & 4 \\ 0 & 2 & -2 & 0 & -4 \end{array} \right] \\
 \xrightarrow{F_2 \leftarrow (1/4)F_2} \left[ \begin{array}{cccc|c} 1 & -3 & 3 & 1 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 2 & -2 & 0 & -4 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 - 2F_2} \left[ \begin{array}{cccc|c} 1 & -3 & 3 & 1 & 1 \\ 0 & 1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -6 \end{array} \right]
 \end{array}$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 - 3x_2 + 3x_3 + x_4 = 1 \\ x_2 - x_3 = 1 \\ 0 = -6 \end{array} \right\}$$

que claramente no tiene solución pues  $0 \neq -6$ .

Desde el punto de vista geométrico, el ejemplo anterior pone de manifiesto que los hiperplanos

$$x_1 - 3x_2 + 3x_3 + x_4 = 1, \quad x_1 + x_2 - x_3 + x_4 = 5 \quad \text{y} \quad x_1 - x_2 + x_3 + x_4 = -3$$

de  $\mathbb{R}^4$  no tienen ningún punto en común.

**Ejemplo 2.2.5** Consideremos el sistema lineal

$$\left. \begin{array}{rcl} x_1 + x_2 - x_3 - 2x_4 = 3 \\ 2x_1 + x_2 + x_3 + x_4 = 2 \\ x_1 + 2x_3 + 3x_4 = -1 \end{array} \right\}$$

Mediante operaciones elementales sobre la matriz ampliada de dicho sistema tenemos que

$$\left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 2 & 1 & 1 & 1 & 2 \\ 1 & 0 & 2 & 3 & -1 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - 2F_1} \left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 0 & -1 & 3 & 5 & -4 \\ 1 & 0 & 2 & 3 & -1 \end{array} \right]$$

$$\xrightarrow{F_3 \leftarrow F_3 - F_1} \left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 0 & -1 & 3 & 5 & -4 \\ 0 & -1 & 3 & 5 & -4 \end{array} \right]$$

$$\xrightarrow{F_2 \leftarrow -F_2} \left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 0 & 1 & -3 & -5 & 4 \\ 0 & -1 & 3 & 5 & -4 \end{array} \right]$$

$$\xrightarrow{F_3 \leftarrow F_3 + F_2} \left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 0 & 1 & -3 & -5 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 + x_2 - x_3 - 2x_4 = 3 \\ x_2 - 3x_3 - 5x_4 = 4 \end{array} \right\}$$

donde hemos eliminado la tercera ecuación que se ha convertido en  $0 = 0$ .

Ahora podemos obtener la solución fácilmente. Si  $x_3$  y  $x_4$  son escalares cualesquiera, por ejemplo  $x_3 = \alpha$  y  $x_4 = \beta$ , donde  $\alpha$  y  $\beta$  son parámetros arbitrarios, entonces

$$\left. \begin{array}{rcl} x_1 + x_2 = 3 + \alpha + 2\beta \\ x_2 = 4 + 3\alpha + 5\beta \end{array} \right\}$$

y por sustitución hacia atrás, obtenemos

$$x_2 = 4 + 3\alpha + 5\beta \quad \text{y} \quad x_1 = -1 - 2\alpha - 3\beta$$

con lo que la solución general del sistema lineal es

$$\begin{aligned} (x_1, x_2, x_3, x_4) &= (-1 - 2\alpha - 3\beta, 4 + 3\alpha + 5\beta, \alpha, \beta) \\ &= (-1, 4, 0, 0) + \alpha(-2, 3, 1, 0) + \beta(-3, 5, 0, 1). \end{aligned}$$

Puesto que los vectores  $(-2, 3, 1, 0)$  y  $(-3, 5, 0, 1)$  no son proporcionales (¡comprobadlo!) tenemos que, desde el punto de vista geométrico, el ejemplo anterior pone de manifiesto que los hiperplanos

$$x_1 + x_2 - x_3 - 2x_4 = 3, \quad 2x_1 + x_2 + x_3 + x_4 = 2 \quad \text{y} \quad x_1 + 2x_3 + 3x_4 = -1$$

de  $\mathbb{R}^4$  se cortan en el plano cuya ecuación vectorial es

$$(x_1, x_2, x_3, x_4) = (-1, 4, 0, 0) + \alpha(-2, 3, 1, 0) + \beta(-3, 5, 0, 1).$$

Por tanto,  $(-1, 4, 0, 0)$  es un punto de paso de dicho plano mientras que  $(-2, 3, 1, 0)$  y  $(-3, 5, 0, 1)$  son los vectores directores.

Hasta ahora, los coeficientes y los términos independientes de todos los sistemas lineales que hemos encontrado eran números reales y, en consecuencia, las soluciones eran vectores de  $\mathbb{R}^n$  para algún valor de  $n$ . Sin embargo, a veces aparecen vectores cuyas componentes son elementos de  $\mathbb{Z}_p$  para algún primo  $p$ , que se comporta, por lo que se refiere a la adición y a la multiplicación como  $\mathbb{R}$ , por tanto, también podemos resolver sistemas lineales cuando los coeficientes y los términos independientes pertenecen a  $\mathbb{Z}_p$  para algún primo  $p$ .

**Ejemplo 2.2.6** Consideremos el sistema lineal, con coeficientes en  $\mathbb{Z}_7$ ,

$$\left. \begin{array}{rcl} x_1 + 2x_2 + 4x_3 = 2 \\ 2x_1 + 6x_2 + x_3 = 0 \\ 4x_1 + 6x_2 + x_3 = 4 \end{array} \right\}$$

Puesto que la adición y la multiplicación en  $\mathbb{Z}_7$  están definidas por las tablas

+	0	1	2	3	4	5	6		0	1	2	3	4	5	6
0	0	1	2	3	4	5	6		0	0	0	0	0	0	0
1	1	2	3	4	5	6	0		1	0	1	2	3	4	5
2	2	3	4	5	6	0	1		2	0	2	4	6	1	3
3	3	4	5	6	0	1	2		3	0	3	6	2	5	1
4	4	5	6	0	1	2	3		4	0	4	1	5	2	6
5	5	6	0	1	2	3	4		5	0	5	3	1	6	4
6	6	0	1	2	3	4	5		6	0	6	5	4	3	2

mediante operaciones elementales sobre la matriz ampliada de dicho sistema tenemos que

$$\begin{array}{c} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 2 & 6 & 1 & 0 \\ 4 & 6 & 1 & 4 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 + 5F_1} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 0 & 2 & 0 & 3 \\ 4 & 6 & 1 & 4 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 + 3F_1} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 0 & 2 & 0 & 3 \\ 0 & 5 & 6 & 3 \end{array} \right] \\ \xrightarrow{F_2 \leftarrow 4F_2} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 5 & 6 & 3 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 + 2F_2} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 6 & 6 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow 6F_3} \left[ \begin{array}{ccc|c} 1 & 2 & 4 & 2 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 1 \end{array} \right] \end{array}$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 + 2x_2 + 4x_3 & = & 2 \\ & x_2 & = 5 \\ & x_3 & = 1 \end{array} \right\}$$

Notemos que, de acuerdo con las tablas de la adición y la multiplicación en  $\mathbb{Z}_7$ , se satisfacen las relaciones

$$-2 = 5, \quad -4 = 3, \quad 2^{-1} = 4 \quad \text{y} \quad 6^{-1} = 6$$

utilizadas en las operaciones elementales anteriores.

Ahora resolvemos el sistema anterior mediante sustitución hacia atrás. De la tercera y segunda ecuaciones tenemos que  $x_3 = 1$  y  $x_2 = 5$  y sustituyendo dichos valores en la primera ecuación

$$x_1 + 2 \cdot 5 + 4 \cdot 1 = 2,$$

es decir

$$x_1 + 3 + 4 = 2$$

pues, de acuerdo con las tablas de la adición y la multiplicación en  $\mathbb{Z}_7$ ,  $2 \cdot 5 = 3$  y  $4 \cdot 1 = 4$ . Ahora, como  $3 + 4 = 0$ , concluimos que  $x_1 = 2$ . Por tanto, la terna  $(2, 5, 1)$  es la solución del sistema lineal dado.

De los ejemplos anteriores se derivan algunas consecuencias importantes.

En primer lugar, un sistema lineal puede no tener solución, tener una única solución o tener infinitas soluciones, si los coeficientes son números reales; si los coeficientes son elementos de  $\mathbb{Z}_p$  para algún primo  $p$ , entonces un sistema lineal puede no tener solución, tener una única solución o tener más de una solución —pero siempre un número finito (*¿por qué?*).

En segundo lugar, hemos visto que hay un método sistemático de eliminación de algunas incógnitas en todas las ecuaciones a partir de una ecuación determinada, de manera que el sistema lineal resultante (que por el teorema 2.2.1 es equivalente al sistema lineal inicial) tiene una forma simple que nos permite decidir, de forma inmediata, cuantas soluciones tiene y, mediante sustitución hacia atrás, obtener el conjunto de soluciones cuando éste sea no vacío.

Este procedimiento se conoce con el nombre de **método de eliminación de Gauss** o simplemente **método de Gauss**.

**Definición 2.2.2** Decimos que un sistema lineal es **incompatible** si no tiene ninguna solución; en caso contrario decimos que es **compatible**. En este último caso distinguimos entre **determinado** si tiene una única solución e **indeterminado** si tiene más de una solución.

## 2.3 Matrices escalonadas y rango de una matriz

### 2.3.1 Matrices escalonadas

Como hemos visto en los ejemplos de la sección anterior, en la resolución de un sistema lineal por el método de Gauss obtenemos unas matrices con unas determinadas características que definimos seguidamente.

**Definición 2.3.1** Decimos que una matriz es **escalonada** si satisface las tres condiciones siguientes:

- El primer elemento no nulo (por la izquierda) de cada fila es un 1 al que llamamos **uno principal**.
- Cada uno principal está a la derecha de los unos principales de las filas anteriores.
- Las filas nulas, si existen, están en la parte inferior de la matriz.

**Ejemplo 2.3.1** Las matrices siguientes son escalonadas:

$$\left[ \begin{array}{cccc} 1 & -1 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array} \right], \quad \left[ \begin{array}{ccccc} 1 & 2 & 3 & -5 & 2 \\ 0 & 0 & 1 & 2 & 7 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right], \quad \left[ \begin{array}{ccccc} 0 & 1 & -4 & 7 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

En cambio,

$$\left[ \begin{array}{cccc} 1 & -3 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{array} \right] \quad \text{y} \quad \left[ \begin{array}{ccccc} 1 & 6 & -3 & 7 & 2 \\ 0 & 0 & 2 & 2 & 8 \\ 0 & 0 & 0 & 0 & 1 \end{array} \right]$$

no son matrices escalonadas.

Existe una variación del método de Gauss, conocida con el nombre de **método de Gauss-Jordan**, que proporciona directamente la solución del sistema inicial, cuando éste es compatible. Veamos algunos ejemplos.

**Ejemplo 2.3.2** Consideremos de nuevo el sistema lineal del ejemplo 2.2.3 para el que hemos obtenido, en la página 30, la matriz ampliada

$$\left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

correspondiente al sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 - 3x_2 + 4x_3 = 2 \\ x_2 - 3x_3 = -2 \\ x_3 = 1 \end{array} \right\}$$

del ejemplo 2.2.1.

Si continuamos haciendo operaciones elementales sobre las filas de la matriz anterior, para anular los elementos que están por encima de los unos principales, obtenemos

$$\begin{array}{c} \left[ \begin{array}{ccc|c} 1 & -3 & 4 & 2 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right] \xrightarrow{F_1 \leftarrow F_1 + 3F_2} \left[ \begin{array}{ccc|c} 1 & 0 & -5 & -4 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right] \\ \xrightarrow{F_1 \leftarrow F_1 + 5F_3} \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & -3 & -2 \\ 0 & 0 & 1 & 1 \end{array} \right] \\ \xrightarrow{F_2 \leftarrow F_2 + 3F_3} \left[ \begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right] \end{array}$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 & = & 1 \\ x_2 & = & 1 \\ x_3 & = & 1 \end{array} \right\}$$

cuya solución, que coincide obviamente con la obtenida en el ejemplo 2.2.1, es directa.

**Ejemplo 2.3.3** Consideremos la matriz escalonada

$$\left[ \begin{array}{cccc|c} 1 & 1 & -1 & -2 & 3 \\ 0 & 1 & -3 & -5 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

correspondiente a un sistema lineal equivalente al sistema lineal del ejemplo 2.2.5. Mediante la operación elemental  $F_1 \leftarrow F_1 - F_2$  transformamos la matriz anterior en la matriz

$$\left[ \begin{array}{cccc|c} 1 & 0 & 2 & 3 & -1 \\ 0 & 1 & -3 & -5 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right]$$

que es la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 & + & 2x_3 & + & 3x_4 & = & -1 \\ x_2 & - & 3x_3 & - & 5x_4 & = & 4 \end{array} \right\}$$

Asignando los parámetros  $\alpha$  y  $\beta$  a las incógnitas  $x_3$  y  $x_4$  respectivamente, obtenemos que

$$x_1 = -1 - 2\alpha - 3\beta \quad \text{y} \quad x_2 = 4 + 3\alpha + 5\beta$$

con lo que la solución general del sistema lineal es

$$\begin{aligned}(x_1, x_2, x_3, x_4) &= (-1 - 2\alpha - 3\beta, 4 + 3\alpha + 5\beta, \alpha, \beta) \\ &= (-1, 4, 0, 0) + \alpha(-2, 3, 1, 0) + \beta(-3, 5, 0, 1)\end{aligned}$$

que coincide, evidentemente, con la obtenida en el ejemplo 2.2.5.

En los ejemplos 2.3.2 y 2.3.3 hemos obtenido unas matrices escalonadas que tienen la particularidad de que también son nulos los elementos que están por encima de los unos principales.

**Definición 2.3.2** Decimos que una matriz es **escalonada reducida** si es escalonada y cada columna que contiene un uno principal tiene nulos los restantes elementos.

Consideremos un sistema lineal compatible cuya matriz de coeficientes es una matriz escalonada o escalonada reducida.

Llamamos **incógnitas principales** a las incógnitas correspondientes a las columnas que contienen un uno principal e **incógnitas paramétricas** a las incógnitas no principales (si las hay). Éstas se asignan a parámetros en la solución general obteniendo, de esta forma, las incógnitas principales en función de dichos parámetros tal como hemos visto en el ejemplo 2.3.3.

Así, es posible predecir el número de incógnitas paramétricas antes de encontrar la solución explícita mediante sustitución hacia atrás. Asociamos un nombre a dicho número, a saber, el **rango de la matriz**.

Finalmente, desde el punto de vista computacional es más eficiente (en el sentido de que requiere un número menor de operaciones) reducir primero la matriz a una matriz escalonada para después, trabajando de derecha a izquierda, anular los elementos que están por encima de los unos principales hasta obtener la forma escalonada reducida. Sin embargo, para el cálculo manual, el alumno puede encontrar más fácil trabajar de izquierda a derecha y crear los unos principales y los ceros en sus columnas a medida que avanza de izquierda a derecha.

### 2.3.2 Rango de una matriz

**Definición 2.3.3** Llamamos **rango por filas** de la matriz  $A$  y lo denotaremos por  $rg_F(A)$  al número de unos principales de su forma escalonada reducida.

Notemos que si  $A$  es una matriz de tamaño  $m \times n$ , entonces  $rg_F(A) \leq m$ .

**Definición 2.3.4** Llamamos **rango** de una matriz  $A$ , y lo denotamos por  $rg(A)$ , al rango por filas o por columnas de  $A$ .

Notemos que si  $A$  es una matriz de tamaño  $m \times n$ , entonces

$$rg(A) \leq \min\{m, n\}.$$

Por la forma en que hemos definido el rango de una matriz, tenemos el siguiente resultado.

**Teorema 2.3.1** *Si  $A$  es una matriz cualquiera, entonces*

$$\operatorname{rg}(A) = \operatorname{rg}(A^T).$$

A partir del concepto establecido de rango de una matriz podemos dar un resultado donde caracterizamos la forma de averiguar si una matriz es invertible o no. Cuando definimos el concepto de matriz inversa, no dimos condiciones ni métodos para el cálculo de dicha matriz invertible.

**Teorema 2.3.2** *Sea  $A$  una matriz cuadrada de tamaño  $n \times n$ . Son equivalentes:*

1.  *$A$  es invertible.*
2. *Para cualquier  $n$ -tupla  $\mathbf{b}$  el sistema lineal  $A\mathbf{x} = \mathbf{b}$  es compatible determinado.*
3. *El sistema lineal homogéneo  $A\mathbf{x} = \mathbf{0}$  es compatible determinado.*
4.  $\operatorname{rg}(A) = n$ .
5. *La forma escalonada reducida de  $A$  es  $I_n$ .*

A partir de este teorema podemos construir una forma para calcular, utilizando el concepto de matriz escalonada reducida, la inversa de una matriz. Si dada la matriz  $A$  existe una matriz  $B$  tal que  $BA = I$  tenemos que  $A^{-1} = B$ .

Por tanto, la misma sucesión de operaciones elementales elementales que transforma la matriz  $A$  en la matriz  $I$  transforma la matriz  $I$  en la matriz  $A^{-1}$ . Veamos algunos ejemplos.

#### Una leyenda de Arquímedes.

Según la tradición, dentro de sus trabajos en la defensa de Siracusa, Arquímedes podría haber creado un sistema de espejos uestorios que reflejaban la luz solar concentrándola en los barcos enemigos y con la finalidad de incendiárlas. La credibilidad de esta historia ha sido objeto de debate desde el Renacimiento. En octubre de 2005 un grupo de estudiantes del Instituto Tecnológico de Massachusetts llevó a cabo un experimento con 127 espejos cuadrados de 30 cm de lado enfocados en una maqueta de madera de un barco a una distancia de 30 m. Brotaron llamas en una parte del barco, pero únicamente después de que el cielo se despejara y de que el barco permaneciera inmóvil alrededor de diez minutos. Se concluyó que el arma era un mecanismo viable bajo estas condiciones.

**Ejemplo 2.3.4** Para obtener la inversa de la matriz

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

añadimos a la derecha de  $A$  la matriz identidad y efectuamos operaciones elementales sobre las filas de la matriz  $[A : I]$ .

$$\begin{array}{c} \left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -2 & 0 & 1 & 0 \\ 2 & -1 & 0 & 0 & 0 & 1 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - F_1} \left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -1 & 1 & 0 \\ 2 & -1 & 0 & 0 & 0 & 1 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 - 2F_1} \left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -1 & 1 & 0 \\ 0 & 1 & -2 & -2 & 0 & 1 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 - F_2} \left[ \begin{array}{ccc|ccc} 1 & -1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right] \\ \xrightarrow{F_1 \leftarrow F_1 - F_3} \left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 2 & 1 & -1 \\ 0 & 1 & -3 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right] \\ \xrightarrow{F_2 \leftarrow F_2 + 3F_3} \left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 2 & 1 & -1 \\ 0 & 1 & 0 & -4 & -2 & 3 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right] \\ \xrightarrow{F_1 \leftarrow F_1 + F_2} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -1 & 2 \\ 0 & 1 & 0 & -4 & -2 & 3 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array} \right] \end{array}$$

Por tanto,

$$A^{-1} = \begin{bmatrix} -2 & -1 & 2 \\ -4 & -2 & 3 \\ -1 & -1 & 1 \end{bmatrix}.$$

El método utilizado en el ejemplo anterior proporciona  $A^{-1}$  si  $A$  es invertible. Si  $A$  no es invertible, la forma escalonada reducida de  $A$  será distinta de  $I$ , por tanto, podemos utilizar también dicho método para averiguar si una matriz dada es invertible o no.

**Ejemplo 2.3.5** Consideremos la matriz

$$A = \begin{bmatrix} 1 & 1 & -3 \\ 3 & 4 & -1 \\ -4 & -5 & 4 \end{bmatrix}$$

Procediendo como en el ejemplo anterior obtenemos

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & -3 & 1 & 0 & 0 \\ 3 & 4 & -1 & 0 & 1 & 0 \\ -4 & -5 & 4 & 0 & 0 & 1 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - 3F_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 8 & -3 & 1 & 0 \\ -4 & -5 & 4 & 0 & 0 & 1 \end{array} \right]$$

$$\xrightarrow{F_3 \leftarrow F_3 + 4F_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 8 & -3 & 1 & 0 \\ 0 & -1 & -8 & 4 & 0 & 1 \end{array} \right]$$

$$\xrightarrow{F_3 \leftarrow F_3 + F_2} \left[ \begin{array}{ccc|ccc} 1 & 1 & -3 & 1 & 0 & 0 \\ 0 & 1 & 8 & -3 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{array} \right]$$

Puesto que  $\text{rg}(A) = 2$  (¿por qué?), tenemos, por el teorema 2.3.2, que  $A$  no es invertible.

## 2.4 Sistemas lineales homogéneos

En esta sección introducimos los sistemas lineales homogéneos.

**Definición 2.4.1** Decimos que un sistema lineal es **homogéneo** si son nulos los términos independientes de todas las ecuaciones.

Así, un sistema homogéneo es un sistema de la forma

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

Notemos que un sistema homogéneo siempre es compatible ya que, al menos, admite la  $n$ -tuple nula  $\mathbf{0}_n$  como solución, que llamamos **solución trivial**. Cualquier otra solución, caso de que exista, decimos que es una **solución no trivial**.

Así, desde el punto de vista geométrico, un sistema lineal homogéneo representa la intersección de un conjunto de hiperplanos que pasan por el origen de coordenadas.

**Ejemplo 2.4.1** Consideremos el sistema lineal homogéneo

$$\left. \begin{array}{rcl} x_1 + 2x_2 + x_3 = 0 \\ 2x_1 + 5x_2 - 3x_3 = 0 \\ 3x_1 + 6x_2 + 4x_3 = 0 \end{array} \right\}$$

mediante operaciones elementales sobre la matriz ampliada de dicho sistema tenemos que

$$\left[ \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 5 & -3 & 0 \\ 3 & 6 & 4 & 0 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - 2F_1} \left[ \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & -5 & 0 \\ 3 & 6 & 4 & 0 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 - 3F_1} \left[ \begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & -5 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right]$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{rcl} x_1 + 2x_2 + x_3 = 0 \\ x_2 - 5x_3 = 0 \\ x_3 = 0 \end{array} \right\}$$

Ahora, podemos resolver el sistema lineal anterior mediante sustitución hacia atrás. De la última ecuación, claramente,  $x_3 = 0$ ; sustituyendo dicho valor en la segunda ecuación, resulta  $x_2 = 0$  y, sustituyendo ambos valores en la primera ecuación, tenemos que  $x_1 = 0$ . Por tanto, la solución trivial  $(0,0,0)$  es la única solución del sistema lineal homogéneo original.

El término Google tiene origen en la matemática, google viene de googol, que es el número  $10^{100}$ , es decir, el dígito 1 seguido de cien ceros. El googol tiene poca utilidad, sólo para explicar la diferencia entre un gran número y el infinito, y debido a su magnitud, los fundadores de Google decidieron adaptar el término para dar nombre a su empresa. El término fue acuñado por Milton Sirotta, sobrino del matemático estadounidense Edward Kasner, y se popularizó gracias al libro *Mathematics and the Imagination*, de Kasner y James Newman. Para la empresa, Google simboliza el objetivo de organizar la inmensa cantidad de información disponible en Internet.

Desde el punto de vista geométrico, el ejemplo anterior pone de manifiesto que el único punto en común de los planos

$$x_1 + 2x_2 + x_3 = 0, \quad 2x_1 + 5x_2 - 3x_3 = 0 \quad \text{y} \quad 3x_1 + 6x_2 + 4x_3 = 0$$

de  $\mathbb{R}^3$  es el punto  $(0,0,0)$ .

**Ejemplo 2.4.2** Consideremos el sistema lineal homogéneo

$$\left. \begin{array}{rcl} x_1 + 3x_2 + 2x_3 - 7x_4 + x_5 = 0 \\ -2x_1 - 6x_2 - 3x_3 + 9x_4 = 0 \\ 2x_1 + 6x_2 + 2x_3 - 4x_4 - 2x_5 = 0 \\ x_1 + 3x_2 + x_3 - 2x_4 - x_5 = 0 \end{array} \right\}$$

Mediante operaciones elementales sobre la matriz ampliada de dicho sistema tenemos que

$$\begin{array}{c}
 \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ -2 & -6 & -3 & 9 & 0 & 0 \\ 2 & 6 & 2 & -4 & -2 & 0 \\ 1 & 3 & 1 & -2 & -1 & 0 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 + 2F_1} \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ 0 & 0 & 1 & -5 & 2 & 0 \\ 2 & 6 & 2 & -4 & -2 & 0 \\ 1 & 3 & 1 & -2 & -1 & 0 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 - 2F_1} \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ 0 & 0 & 1 & -5 & 2 & 0 \\ 0 & 0 & -2 & 10 & -4 & 0 \\ 1 & 3 & 1 & -2 & -1 & 0 \end{array} \right] \\
 \xrightarrow{F_4 \leftarrow F_4 - F_1} \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ 0 & 0 & 1 & -5 & 2 & 0 \\ 0 & 0 & -2 & 10 & -4 & 0 \\ 0 & 0 & -1 & 5 & -2 & 0 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 + 2F_2} \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ 0 & 0 & 1 & -5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 5 & -2 & 0 \end{array} \right] \\
 \xrightarrow{F_4 \leftarrow F_4 + F_2} \left[ \begin{array}{ccccc|c} 1 & 3 & 2 & -7 & 1 & 0 \\ 0 & 0 & 1 & -5 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right]
 \end{array}$$

que corresponde a la matriz ampliada del sistema lineal equivalente

$$\left. \begin{array}{l} x_1 + 3x_2 + 2x_3 - 7x_4 + x_5 = 0 \\ x_3 - 5x_4 + 2x_5 = 0 \end{array} \right\}$$

Ahora, podemos resolver el sistema anterior mediante sustitución hacia atrás. Si  $x_2 = \alpha$ ,  $x_4 = \beta$  y  $x_5 = \gamma$ , donde  $\alpha$ ,  $\beta$  y  $\gamma$  son parámetros arbitrarios, entonces de la última ecuación tenemos que

$$x_3 = 5\beta - 2\gamma$$

y sustituyendo dicho valor en la primera ecuación, resulta que

$$x_1 = -3\alpha - 3\beta + 3\gamma.$$

Por tanto,

$$(x_1, x_2, x_3, x_4, x_5) = \alpha(-3, 1, 0, 0, 0) + \beta(-3, 0, 5, 1, 0) + \gamma(3, 0, -2, 0, 1)$$

es la solución general del sistema lineal homogéneo dado; es decir, la solución general del sistema lineal homogéneo es combinación lineal de las soluciones particulares

$$(-3, 1, 0, 0, 0), \quad (-3, 0, 5, 1, 0) \quad \text{y} \quad (3, 0, -2, 0, 1).$$

Notemos que en la resolución de un sistema lineal homogéneo, las distintas operaciones

elementales no afectan a la columna de términos independientes (que siempre son nulos), como hemos visto en los dos ejemplos anteriores.

## 2.5 El Teorema de Rouché-Frobenius

El teorema de Rouché-Frobenius establece la relación entre el conjunto de soluciones de un sistema lineal y el conjunto de soluciones del sistema lineal homogéneo asociado a dicho sistema lineal, lo cual no significa que el sistema lineal inicial sea compatible como ocurre con el sistema lineal homogéneo asociado que siempre es compatible.

En el ejemplo siguiente presentamos un sistema lineal incompatible cuyo sistema lineal homogéneo asociado es compatible indeterminado.

**Ejemplo 2.5.1** Consideremos el sistema lineal

$$\begin{array}{rcl} x_1 - x_2 & = & 1 \\ x_1 + x_2 + 2x_3 & = & 5 \\ x_2 + x_3 & = & 3 \end{array} \left. \right\}.$$

Mediante operaciones elementales sobre las filas de la matriz ampliada del sistema tenemos

$$\begin{array}{c} \left[ \begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 1 & 1 & 2 & 5 \\ 0 & 1 & 1 & 3 \end{array} \right] \xrightarrow{F_2 \leftarrow F_2 - F_1} \left[ \begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 0 & 2 & 2 & 4 \\ 0 & 1 & 1 & 3 \end{array} \right] \\ \xrightarrow{F_2 \leftarrow (1/2)F_2} \left[ \begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 1 & 1 & 3 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 - F_2} \left[ \begin{array}{ccc|c} 1 & -1 & 0 & 1 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array} \right] \\ \xrightarrow{F_1 \leftarrow F_1 - F_2} \left[ \begin{array}{ccc|c} 1 & 0 & 1 & 3 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{array} \right] \end{array}$$

Por tanto, se trata de un sistema incompatible. Sin embargo, la matriz ampliada del sistema lineal homogéneo asociado al sistema lineal dado es

$$\left[ \begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

cuya solución general es, obviamente,

$$\left[ \begin{array}{c} x_1 \\ x_2 \\ x_3 \end{array} \right] = \alpha \left[ \begin{array}{c} -1 \\ -1 \\ 1 \end{array} \right].$$

**Teorema 2.5.1 (Teorema de Rouché-Frobenius)** Supongamos que  $A$  es una matriz de tamaño  $m \times n$  y que  $\mathbf{b}$  es una  $m$ -tupla. El sistema lineal  $A\mathbf{x} = \mathbf{b}$  es compatible si y solo si  $\text{rg}(A) = \text{rg}([A\mathbf{b}])$ .

En particular, si  $r = \text{rg}(A) = \text{rg}([A\mathbf{b}])$ , entonces el sistema lineal  $A\mathbf{x} = \mathbf{b}$  es:

1. compatible determinado si y solo si  $r = n$ ,
2. compatible indeterminado si y solo si  $r < n$ .

Veamos un ejemplo de aplicación a un sistema concreto.

**Ejemplo 2.5.2** Consideremos el sistema lineal

$$\left. \begin{array}{rcl} ax_1 + x_2 + x_3 = 1 \\ x_1 + ax_2 + x_3 = a \\ x_1 + x_2 + ax_3 = a^2 \end{array} \right\}$$

Mediante operaciones elementales sobre la matriz ampliada de dicho sistema obtenemos que

$$\begin{array}{c} \left[ \begin{array}{ccc|c} a & 1 & 1 & 1 \\ 1 & a & 1 & a \\ 1 & 1 & a & a^2 \end{array} \right] \xrightarrow{F_1 \leftrightarrow F_3} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 1 & a & 1 & a \\ a & 1 & 1 & 1 \end{array} \right] \\ \xrightarrow{F_2 \leftarrow F_2 - F_1} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ a & 1 & 1 & 1 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 - aF_1} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ 0 & 1-a & 1-a^2 & 1-a^3 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow F_3 + F_2} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & a-1 & 1-a & a-a^2 \\ 0 & 0 & 2-a-a^2 & 1+a-a^2-a^3 \end{array} \right] \\ = B \end{array}$$

Si  $a = 1$ , entonces

$$B = \left[ \begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]$$

con lo que el sistema es compatible indeterminado, siendo la solución general

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}.$$

Si  $a \neq 1$ , entonces  $a - 1 \neq 0$  y mediante operaciones elementales, obtenemos que

$$\begin{array}{l} \text{B} \xrightarrow{F_2 \leftarrow \frac{1}{a-1}F_2} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & 1 & -1 & -a \\ 0 & 0 & (1-a)(2+a) & (1-a)(1+a)^2 \end{array} \right] \\ \xrightarrow{F_3 \leftarrow \frac{1}{2+a}F_3} \left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & 1 & -1 & -a \\ 0 & 0 & 2+a & (1+a)^2 \end{array} \right] \end{array}$$

Así, si  $a = -2$ , entonces la matriz anterior es

$$\left[ \begin{array}{ccc|c} 1 & 1 & -2 & 4 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 1 \end{array} \right]$$

con lo que el sistema es incompatible. En cambio, si  $a \neq -2$ , entonces  $2+a \neq 0$  y mediante la operación elemental  $F_3 \leftarrow \frac{1}{2+a}F_3$  obtenemos la matriz

$$\left[ \begin{array}{ccc|c} 1 & 1 & a & a^2 \\ 0 & 1 & -1 & -a \\ 0 & 0 & 1 & \frac{(1+a)^2}{2+a} \end{array} \right]$$

con lo que el sistema es compatible determinado. Mediante sustitución hacia atrás obtenemos, de la última ecuación, que

$$x_3 = \frac{(1+a)^2}{2+a},$$

sustituyendo ahora  $x_3$  en la segunda ecuación obtenemos

$$x_2 = -a + \frac{(1+a)^2}{2+a} = \frac{1}{2+a}.$$

Finalmente, sustituyendo  $x_2$  y  $x_3$  en la primera ecuación, obtenemos

$$x_1 = a^2 - \frac{1}{2+a} - a \frac{(1+a)^2}{2+a} = -\frac{1+a}{2+a}.$$

Así  $\left(-\frac{1+a}{2+a}, \frac{1}{2+a}, \frac{(1+a)^2}{2+a}\right)$  es la solución del sistema para  $a \neq -2$ .

## 2.6 Aplicación: el modelo input-output

Wassily Leontief recibió en 1973 el premio Nobel de Economía por su modelo *input-output* que describimos en esta sección. En líneas generales se trata de un modelo económico formado por diferentes sectores tales que cada uno produce un producto que es consumido

**Tabla 2.1** Distribución del trabajo (ejemplo 2.6.1).

		TRABAJO		
		FONTANERO	ÉLECTRICISTA	PINTOR
CASA	FONTANERO	70 %	40 %	20 %
	ELECTRICISTA	10 %	40 %	40 %
	PINTOR	20 %	20 %	40 %

por todos los sectores de forma que no hay pérdidas ni beneficios por parte de ninguno de los sectores. El ejemplo siguiente nos ayudará a entender el modelo.

**Ejemplo 2.6.1** Un fontanero, un electricista y un pintor deciden realizar los correspondientes trabajos de fontanería, electricidad y pintura en sus respectivas casas de forma que ninguno de ellos pague más de lo que ha de cobrar. ¿Cuánto debería cobrar cada uno de los trabajadores anteriores si la distribución del trabajo realizado por cada uno de ellos en cada una de las casas es la que se indica en la tabla 2.1?

Denotemos por  $x_1$ ,  $x_2$  y  $x_3$  las cantidades que cobrarán el fontanero, el electricista y el pintor respectivamente. De la tabla 2.1 tenemos que el fontanero deberá pagar

$$0.7x_1 + 0.4x_2 + 0.2x_3.$$

Ahora bien, como queremos que pague exactamente lo mismo que cobra, deberá cumplirse que

$$0.7x_1 + 0.4x_2 + 0.2x_3 = x_1.$$

Análogamente, para el electricista y el pintor tenemos que

$$\begin{aligned} 0.1x_1 + 0.4x_2 + 0.4x_3 &= x_2 \\ 0.2x_1 + 0.2x_2 + 0.4x_3 &= x_3 \end{aligned}$$

Podemos escribir las ecuaciones anteriores en forma matricial como

$$\begin{bmatrix} 0.7 & 0.4 & 0.2 \\ 0.1 & 0.4 & 0.4 \\ 0.2 & 0.2 & 0.4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix},$$

es decir

$$E\mathbf{x} = \mathbf{x}$$

o equivalentemente

$$(I - E)\mathbf{x} = \mathbf{0}$$

donde  $I$  es la matriz identidad de tamaño  $3 \times 3$ ,

$$E = \begin{bmatrix} 0.7 & 0.4 & 0.2 \\ 0.1 & 0.4 & 0.4 \\ 0.2 & 0.2 & 0.4 \end{bmatrix} \quad \text{y} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Resolviendo el sistema lineal homogéneo  $(I - E)\mathbf{x} = \mathbf{0}$  obtenemos la solución general

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2\alpha \\ \alpha \\ \alpha \end{bmatrix}$$

que pone de manifiesto que el electricista y el pintor han de cobrar la misma cantidad mientras que el fontanero ha de cobrar el doble que ellos.

En general tenemos la siguiente situación. Consideremos una **economía cerrada**, es decir, una economía en la que ni se importa ni se exporta ningún producto y tampoco se producen más productos de los que se consumen. Supongamos que esta economía está formada por  $n$  industrias de forma que cada una de ellas se abastece y es abastecida por las demás (y posiblemente por ella misma) y los ingresos de cada una de ellas igualan sus gastos. Decimos entonces que la economía está en **estado de equilibrio**. El problema consiste en determinar cuáles deben ser los ingresos (o porcentajes de los ingresos) totales anuales de cada industria para mantener el estado de equilibrio.

Para obtener el modelo matemático llamamos  $e_{ij}$  a la proporción consumida por la industria  $i$  del total del producto producido por la industria  $j$ . La matriz  $E = [e_{ij}]$  construida de esta forma decimos que es la **matriz input-output** de la economía en cuestión. Claramente

- $E$  es **no negativa**, es decir

$$e_{ij} \geq 0 \quad \text{para } i = 1, 2, \dots, n \text{ y } j = 1, 2, \dots, n.$$

- La suma de los elementos de cada columna es 1, es decir

$$\sum_{i=1}^n e_{ij} = 1 \quad \text{para } j = 1, 2, \dots, n.$$

De donde claramente

$$e_{ij} \leq 1 \quad \text{para } i = 1, 2, \dots, n \text{ y } j = 1, 2, \dots, n.$$

Llamamos **matrices estocásticas** a las matrices que satisfacen las condiciones anteriores. Las matrices estocásticas también aparecen en otras aplicaciones, como las cadenas de Markov que veremos en el capítulo de valores y vectores propios.

Si denotamos por  $x_i$  los ingresos anuales de la industria  $i$ , entonces la condición para que la economía esté en equilibrio se traduce en la ecuación

$$e_{i1}x_1 + e_{i2}x_2 + \cdots + e_{in}x_n = x_i.$$

Llamando  $\mathbf{x}$  a la matriz columna cuya  $i$ -ésima componente es  $x_i$ , las ecuaciones anteriores dan lugar a la ecuación matricial

$$E\mathbf{x} = \mathbf{x},$$

que llamamos **condición de equilibrio**. Finalmente, llamamos **estructura de precios en equilibrio** a las soluciones del sistema lineal homogéneo

$$(I - E)\mathbf{x} = \mathbf{0}$$

que se deduce de la ecuación matricial anterior.

## 2.7 Métodos iterativos

La mayoría de los sistemas lineales que aparecen en las aplicaciones no pueden resolverse a mano porque están formados por cientos o miles de ecuaciones e incógnitas. Necesitamos, por tanto, utilizar ordenadores con programas eficientes. Aunque los métodos de Gauss y de Gauss-Jordan estudiados en este capítulo son métodos muy importantes para resolver un sistema, no siempre son los más eficientes, aun cuando utilicemos un ordenador.

Aunque aparentemente el método de Gauss-Jordan es más eficiente que el de Gauss porque no requiere la sustitución hacia atrás, esto no es cierto ya que el método de Gauss-Jordan requiere más operaciones que el de Gauss. De hecho, para un sistema de  $n$  ecuaciones con  $n$  incógnitas, cuando  $n$  es grande, puede probarse que el método de Gauss requiere aproximadamente  $2n^3/3$  operaciones aritméticas (es decir, adiciones y multiplicaciones) mientras que el de Gauss-Jordan requiere aproximadamente  $n^3$  operaciones aritméticas. La demostración de este hecho está fuera de los objetivos de este curso, sin embargo, el alumno interesado puede encontrar una demostración del mismo en [?]. Así, cuando se utiliza el método de Gauss-Jordan para un sistema pequeño, por ejemplo, de 50 ecuaciones con 50 incógnitas, se necesitan aproximadamente 125 000 operaciones frente a las aproximadamente 83 000 operaciones requeridas por el método de Gauss.

Los métodos de Gauss y de Gauss-Jordan son **métodos directos** en el sentido de que, cuando el sistema es compatible, siempre se obtiene la solución, como acabamos de ver, mediante un número finito de operaciones.

Además de utilizar un método directo para resolver un sistema lineal, también podemos utilizar un **método iterativo**, que consiste en aproximar la solución de un sistema recurriendo a iteraciones que comienzan con un valor aproximado inicial. Si las iteraciones sucesivas se aproximan a la solución, decimos que el método es **convergente**, en caso contrario, decimos que es **divergente**. El proceso termina cuando en dos iteraciones sucesivas se obtiene la misma respuesta con una precisión prefijada. A diferencia de los métodos directos, el número de operaciones aritméticas necesarias para la resolución de un sistema lineal por un método iterativo no se conoce con antelación.

Seguidamente describimos los métodos de Jacobi y de Gauss-Seidel que se aplican a sistemas cuadrados, es decir, a sistemas que tienen el mismo número de ecuaciones que de incógnitas. Empecemos con un par de ejemplos.

**Ejemplo 2.7.1** Consideremos el sistema lineal

$$\left. \begin{array}{rcl} 4x_1 & + & x_2 & + & x_3 & = & 9 \\ -2x_1 & + & 4x_2 & - & x_3 & = & 3 \\ -x_1 & - & x_2 & + & 4x_3 & = & 9 \end{array} \right\}$$

Despejando la primera incógnita en la primera ecuación, la segunda incógnita en la segunda ecuación y la tercera incógnita en la tercera ecuación, podemos escribir el sistema anterior como

$$\left. \begin{array}{l} x_1 = \frac{9}{4} - \frac{1}{4}x_2 - \frac{1}{4}x_3 \\ x_2 = \frac{3}{4} + \frac{2}{4}x_1 + \frac{1}{4}x_3 \\ x_3 = \frac{9}{4} + \frac{1}{4}x_1 + \frac{1}{4}x_2 \end{array} \right\}$$

que da lugar al esquema iterativo

$$\left. \begin{array}{l} x_1^{(k+1)} = \frac{9}{4} - \frac{1}{4}x_2^{(k)} - \frac{1}{4}x_3^{(k)} \\ x_2^{(k+1)} = \frac{3}{4} + \frac{2}{4}x_1^{(k)} + \frac{1}{4}x_3^{(k)} \\ x_3^{(k+1)} = \frac{9}{4} + \frac{1}{4}x_1^{(k)} + \frac{1}{4}x_2^{(k)} \end{array} \right\} \quad \text{para } k \geq 0.$$

Ahora elegimos, de forma arbitraria, una aproximación de la solución del sistema  $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)})$ . Si no disponemos de ninguna información adicional podemos tomar como primera aproximación  $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}) = (0,0,0)$ . Así,

$$\left. \begin{array}{l} x_1^{(1)} = \frac{9}{4} - \frac{1}{4} \cdot 0 - \frac{1}{4} \cdot 0 = 2.25 \\ x_2^{(1)} = \frac{3}{4} + \frac{2}{4} \cdot 0 + \frac{1}{4} \cdot 0 = 0.75 \\ x_3^{(1)} = \frac{9}{4} + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 = 2.25 \end{array} \right\}$$

Ahora, sustituimos dichos valores para obtener una nueva aproximación,

$$\left. \begin{array}{l} x_1^{(2)} = \frac{9}{4} - \frac{1}{4} \cdot 0.75 - \frac{1}{4} \cdot 2.25 = 1.5 \\ x_2^{(2)} = \frac{3}{4} + \frac{2}{4} \cdot 2.25 + \frac{1}{4} \cdot 2.25 = 2.4375 \\ x_3^{(2)} = \frac{9}{4} + \frac{1}{4} \cdot 2.25 + \frac{1}{4} \cdot 0.75 = 3 \end{array} \right\}$$

Procediendo de esta forma, obtenemos tres sucesiones  $\{x_1^{(k)}\}$ ,  $\{x_2^{(k)}\}$  y  $\{x_3^{(k)}\}$ , cuyos primeros elementos se muestran en la tabla 2.2. En dicha tabla observamos que las sucesiones  $\{x_1^{(k)}\}$ ,  $\{x_2^{(k)}\}$  y  $\{x_3^{(k)}\}$  convergen a los números  $x_1 = 1$ ,  $x_2 = 2$  y  $x_3 = 3$  respectivamente, que constituyen la solución del sistema propuesto como podemos comprobar fácilmente mediante sustitución directa (¡hacedlo!).

En el ejemplo anterior hemos tomado como primera aproximación de la solución del sistema la terna  $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}) = (0,0,0)$  y a partir de ella hemos obtenido una nueva aproximación de la solución  $(x_1^{(1)}, x_2^{(1)}, x_3^{(1)})$ . Si examinamos detenidamente dicho paso observamos que hay una cierta ineficiencia en el método propuesto ya que, para calcular  $x_2^{(1)}$  utilizamos  $x_1^{(0)}$  en lugar del valor  $x_1^{(1)}$  que ya está calculado; lo mismo ocurre con el cálculo de  $x_3^{(1)}$  donde utilizamos  $x_1^{(0)}$  y  $x_2^{(0)}$  en lugar de los valores  $x_1^{(1)}$  y  $x_2^{(1)}$  que ya están disponibles. Dado que las iteraciones anteriores convergen a la solución del sistema, parece lógico pensar que si modificamos el método del ejemplo 2.7.1 de manera que en

**Tabla 2.2** Iteraciones del sistema lineal del ejemplo 2.7.1.

$k$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	2.2500	0.7500	2.2500
2	1.5000	2.4375	3.0000
3	0.8906	2.2500	3.2344
4	0.8789	2.0039	3.0352
5	0.9902	1.9482	2.9707
6	1.0203	1.9878	2.9846
7	1.0069	2.0063	3.0020
8	0.9979	2.0040	3.0033
9	0.9982	1.9998	3.0005
10	0.9999	1.9992	2.9995
11	1.0003	1.9998	2.9998
12	1.0001	2.0001	3.0000
13	1.0000	2.0001	3.0001
14	1.0000	2.0000	3.0000
15	1.0000	2.0000	3.0000

cada iteración utilizamos los valores más recientes, entonces la convergencia será más rápida.

**Ejemplo 2.7.2** Consideremos de nuevo el sistema lineal del ejemplo 2.7.1 y consideremos, de acuerdo con los comentarios anteriores, el esquema iterativo

$$\left. \begin{array}{l} x_1^{(k+1)} = \frac{9}{4} - \frac{1}{4}x_2^{(k)} - \frac{1}{4}x_3^{(k)} \\ x_2^{(k+1)} = \frac{3}{4} + \frac{2}{4}x_1^{(k+1)} + \frac{1}{4}x_3^{(k)} \\ x_3^{(k+1)} = \frac{9}{4} + \frac{1}{4}x_1^{(k+1)} + \frac{1}{4}x_2^{(k+1)} \end{array} \right\} \text{ para } k \geq 0.$$

Ahora, tomando la aproximación inicial  $(x_1^{(0)}, x_2^{(0)}, x_3^{(0)}) = (0,0,0)$  obtenemos la primera aproximación

$$\left. \begin{array}{l} x_1^{(1)} = \frac{9}{4} - \frac{1}{4} \cdot 0 - \frac{1}{4} \cdot 0 = 2.25 \\ x_2^{(1)} = \frac{3}{4} + \frac{2}{4} \cdot 2.25 + \frac{1}{4} \cdot 0 = 1.8750 \\ x_3^{(1)} = \frac{9}{4} + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 = 3.2813 \end{array} \right\}$$

la segunda aproximación

$$\left. \begin{array}{l} x_1^{(2)} = \frac{9}{4} - \frac{1}{4} \cdot 1.8750 - \frac{1}{4} \cdot 3.2813 = 0.9609 \\ x_2^{(2)} = \frac{3}{4} + \frac{2}{4} \cdot 0.9609 + \frac{1}{4} \cdot 3.2813 = 2.0508 \\ x_3^{(2)} = \frac{9}{4} + \frac{1}{4} \cdot 0.9609 + \frac{1}{4} \cdot 2.0508 = 3.0029 \end{array} \right\}$$

**Tabla 2.3** Iteraciones del sistema lineal del ejemplo 2.7.2.

$k$	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$
0	0	0	0
1	2.2500	1.8750	3.2813
2	0.9609	2.0508	3.0029
3	0.9866	0.9866	2.9951
4	1.0027	2.0001	3.0007
5	0.9998	2.0001	3.0000
6	1.0000	2.0000	3.0000
7	1.0000	2.0000	3.0000

y procediendo de esta forma, obtenemos las sucesiones cuyos primeros elementos se muestran en la tabla 2.3. Como vemos, las sucesiones  $\{x_1^{(k)}\}$ ,  $\{x_2^{(k)}\}$  y  $\{x_3^{(k)}\}$  también convergen a los números  $x_1 = 1$ ,  $x_2 = 2$  y  $x_3 = 3$  respectivamente, que, como hemos visto en el ejemplo 2.7.1, constituyen la solución del sistema propuesto. Sin embargo, ahora la convergencia es *más rápida* ya que el número de iteraciones es menor que en el método del ejemplo 2.7.1.

En general, para un sistema lineal de  $n$  ecuaciones con  $n$  incógnitas,

$$\left. \begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \qquad \vdots \qquad \vdots \qquad \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{array} \right\} \quad (2.6)$$

el método descrito en el ejemplo 2.7.1 se conoce con el nombre de **método de Jacobi**. Este método proporciona la  $(k+1)$ -ésima aproximación de  $x_i$ , para  $i = 1, 2, \dots, n$ , como

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j^{(k)} \right)$$

es decir, para calcular la  $(k+1)$ -ésima aproximación de  $x_i$  utilizamos la  $k$ -ésima aproximación de  $x_j$  para  $j = 1, 2, \dots, n$  con  $j \neq i$ .

En cambio, el método descrito en el ejemplo 2.7.2 se conoce con el nombre de **método de Gauss-Seidel**. Este método proporciona la  $(k+1)$ -ésima aproximación de  $x_i$ , para  $i = 1, 2, \dots, n$ , como

$$x_i^{(k+1)} = \frac{1}{a_{ii}} \left( b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^n a_{ij}x_j^{(k)} \right)$$

es decir, para calcular la  $(k+1)$ -ésima aproximación de  $x_i$  utilizamos la  $(k+1)$ -ésima

aproximación de  $x_j$ , para  $j = 1, 2, \dots, i-1$ , y la  $k$ -ésima aproximación de  $x_j$ , para  $j = i+1, i+2, \dots, n$ .

Hemos visto en los ejemplos 2.7.1 y 2.7.2 que el método de Gauss-Seidel requiere menos iteraciones que el método de Jacobi. Esto *parece* ser cierto en la mayoría de casos, pero no siempre. Desafortunadamente, no podemos conocer de antemano cuál es el método más eficiente.

El teorema siguiente, cuya demostración omitimos por estar fuera de los objetivos de este curso, establece una condición suficiente para que los métodos de Jacobi y de Gauss-Seidel sean convergentes.

**Teorema 2.7.1** *Si la matriz de coeficientes del sistema lineal (2.6) es diagonal dominante, entonces los métodos de Jacobi y de Gauss-Seidel son convergentes.*

**Definición 2.7.1** La matriz cuadrada  $A = [a_{ij}]$  de tamaño  $n \times n$  es **diagonal dominante** si

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad \text{para } i = 1, 2, \dots, n.$$

Por tanto, una matriz es diagonal dominante si para cada fila, el valor absoluto de cada elemento de la diagonal es mayor que la suma de los valores absolutos de los restantes elementos de la fila.

Así, la matriz del sistema del ejemplo 2.7.1

$$\begin{bmatrix} 4 & 1 & 1 \\ -2 & 4 & -1 \\ -1 & -1 & 4 \end{bmatrix}$$

es diagonal dominante ya que

- $|4| > |1| + |1|$ ,
- $|4| > |-2| + |-1|$ ,
- $|4| > |-1| + |-1|$ .

De este modo, por el teorema 2.7.1, tenemos que los métodos de Jacobi y Gauss-Seidel son convergentes tal como pusimos de manifiesto en los ejemplos 2.7.1 y 2.7.2 respectivamente.

En cambio, la matriz

$$\begin{bmatrix} 4 & 2 & -1 \\ 3 & -5 & 2 \\ 1 & -2 & 6 \end{bmatrix}$$

no es diagonal dominante ya que para la segunda fila tenemos que

$$|-5| = 5 \quad \text{no es mayor que} \quad |3| + |2| = 5.$$

Notemos que los métodos de Jacobi y de Gauss-Seidel pueden converger aun cuando la matriz de coeficientes del sistema no sea diagonal dominante (recordemos que, por el teorema 2.7.1, ésta es solamente una condición suficiente para la convergencia de los métodos). A veces, una reordenación de las ecuaciones proporciona como resultado una matriz de coeficientes diagonal dominante. Por ejemplo, la matriz de coeficientes del sistema

$$\left. \begin{array}{rcl} -x_1 & - & x_2 & + & 4x_3 & = & 9 \\ 4x_1 & + & x_2 & + & x_3 & = & 9 \\ -2x_1 & + & 4x_2 & - & x_3 & = & 3 \end{array} \right\}$$

no es diagonal dominante (*¿por qué?*). Sin embargo, si pasamos la segunda ecuación a la primera posición, la tercera ecuación a la segunda posición y la primera ecuación a la tercera posición, obtenemos el sistema del ejemplo 2.7.1 cuya matriz de coeficientes es diagonal dominante tal como hemos comentado anteriormente.

Finalmente, dado que con un método iterativo no es posible calcular en general la solución con un número finito de iteraciones, es necesario determinar criterios de parada. El más utilizado consiste en fijar una tolerancia  $\varepsilon > 0$  (suficientemente pequeña) y detener el proceso cuando

$$\left\| \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)} \right\| < \varepsilon$$

donde  $\mathbf{x}^{(k-1)}$  y  $\mathbf{x}^{(k)}$  son dos aproximaciones consecutivas de la solución.

## 2.8 Códigos detectores y correctores de errores

A lo largo de la historia, las personas han utilizado distintos *códigos* para transmitir información, siendo el código Morse, con su sistema de puntos y rayas, uno de los ejemplos más conocidos.

La aparición de los ordenadores a mediados del siglo xx generó la necesidad de transmitir cantidades inmensas de datos de forma rápida y precisa. Muchos avances tecnológicos dependen de los códigos, encontrándolos en la vida diaria, muchas veces sin tener conciencia de ello: comunicaciones vía satélite, reproductores de discos compactos, el código EAN (*European Article Number*) que aparece en los códigos de barras de la mayoría de productos que encontramos en el supermercado, el ISBN (*International Standard Book Number*) que se encuentra en cada libro, o el CCC (Código de Cuenta Cliente) utilizado por los bancos y cajas de ahorro, son sólo algunos ejemplos de códigos.

En esta sección utilizamos vectores para diseñar códigos que permiten detectar los posibles errores que se pueden presentar durante la transmisión de los datos.

Los vectores que aparecen en el estudio de los códigos no son los familiares de  $\mathbb{R}^n$  sino vectores cuyas componentes son elementos de un determinado conjunto finito. Además, se basan en una aritmética diferente a la de los números reales, llamada *aritmética modular*, que desarrollamos en esta sección.

Debido a que los ordenadores representan los datos mediante secuencias de 0 y 1 (es decir, secuencias de *bits*<sup>1</sup>), comenzamos estudiando los **códigos binarios** que consisten

---

<sup>1</sup> Un *bit*, abreviación de *binary digit*, es un 0 o un 1.

en vectores cuyas componentes pueden ser solamente 0 o 1. En este contexto, debemos modificar las reglas habituales de la aritmética, ya que el resultado de una operación con escalares debe ser un 0 o un 1. Así, la adición y multiplicación de escalares se efectúa de acuerdo con las tablas siguientes:

$+_2$	0	1
0	0	1
1	1	0

$\cdot_2$	0	1
0	0	0
1	0	1

Con estas operaciones, el conjunto de escalares  $\{0,1\}$  se denota por  $\mathbb{Z}_2$  y recibe el nombre de conjunto de los **enteros módulo 2**.

Así, en  $\mathbb{Z}_2$  tenemos que:

- $1 +_2 0 +_2 1 +_2 1 = 1$
- $(1 +_2 1 +_2 0 +_2 1) \cdot_2 (1 +_2 0 +_2 1) = 1 \cdot_2 0 = 0$

Denotamos por  $\mathbb{Z}_2^n$  el conjunto de  $n$ -tuplas binarias, es decir

$$\mathbb{Z}_2^n = \{(u_1, u_2, \dots, u_n) \mid u_i \in \mathbb{Z}_2 \text{ para } i = 1, 2, \dots, n\}.$$

Los elementos de  $\mathbb{Z}_2^n$  también se llaman **vectores binarios** de longitud  $n$ .

Claramente, el número de vectores binarios de longitud  $n$ , es decir, el número de elementos de  $\mathbb{Z}_2^n$ , es  $2^n$ .

La adición y multiplicación por un escalar se realizan componente a componente (módulo 2). Así, si  $\mathbf{u} = (1, 0, 1, 1, 0)$  y  $\mathbf{v} = (1, 1, 0, 1, 0)$  son dos vectores binarios de longitud 5, entonces

$$\begin{aligned}\mathbf{u} +_2 \mathbf{v} &= (1, 0, 1, 1, 0) +_2 (1, 1, 0, 1, 0) \\ &= (1 +_2 1, 0 +_2 1, 1 +_2 0, 1 +_2 1, 0 +_2 0) \\ &= (0, 1, 1, 0, 0).\end{aligned}$$

En cambio

$$0\mathbf{u} = \mathbf{0} \quad \text{y} \quad 1\mathbf{u} = \mathbf{u}.$$

Finalmente,

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle &= 1 \cdot_2 1 +_2 0 \cdot_2 1 +_2 1 \cdot_2 0 +_2 1 \cdot_2 1 +_2 0 \cdot_2 0 \\ &= 1 +_2 0 +_2 0 +_2 1 +_2 0 = 0.\end{aligned}$$

**Definición 2.8.1** Un **código binario**  $\mathcal{C}$  de longitud  $n$  es un subconjunto de  $\mathbb{Z}_2^n$ . Los elementos de  $\mathcal{C}$  se llaman **vectores código** o **palabras código**.

**Definición 2.8.2** Llamamos **codificación** al proceso por el que transformamos un mensaje (consistente en palabras, números o símbolos) en vectores código. Llamamos **decodificación** al proceso por el que recuperamos el mensaje a partir de los vectores código.

Supongamos que ya hemos codificado un mensaje como un conjunto de vectores código de un código binario. Ahora deseamos enviar dichos vectores a través de un canal (tal

**Tabla 2.4** Código binario de longitud 2 del ejemplo 2.8.1.

MENSAJE	norte	sur	este	oeste
CÓDIGO	(0,0)	(0,1)	(1,0)	(1,1)

**Tabla 2.5** Código binario de longitud 3 del ejemplo 2.8.1.

MENSAJE	norte	sur	este	oeste
CÓDIGO	(0,0,0)	(0,1,1)	(1,0,1)	(1,1,0)

como un radiotransmisor, una línea telefónica, un cable de fibra óptica, un láser de CD, etc.). Desafortunadamente, como consecuencia del *ruido* (interferencias eléctricas, suciedad, etc.) al que está sometido el canal, pueden aparecer errores, es decir, algún 0 se puede cambiar por un 1 o viceversa. El ejemplo siguiente muestra una forma de protegernos frente a tales problemas.

**Ejemplo 2.8.1** Deseamos codificar y transmitir un mensaje formado por una de las palabras siguientes: norte, sur, este y oeste. Para ello decidimos utilizar los cuatro vectores de  $\mathbb{Z}_2^2$  tal como indicamos en la tabla 2.4. Si el receptor también dispone de la tabla 2.4 y el mensaje codificado se transmite sin errores, entonces la decodificación es trivial.

Supongamos ahora que, como consecuencia del ruido, se produce un error en la transmisión, es decir, una de las componentes del vector código ha cambiado. En la práctica, la probabilidad de cometer errores es insignificantemente pequeña. Así, si el emisor envía el mensaje “norte” codificado como (0,0) pero el receptor recibe el vector (1,0), entonces decodifica el mensaje recibido como “este” que es incorrecto. Incluso en el caso de que el receptor pudiera saber que ha ocurrido un error, no sabría si el vector código correcto enviado fue (0,0) o (1,1).

Para poder detectar errores, codificamos los mensajes como vectores de  $\mathbb{Z}_2^3$ , es decir, mediante vectores binarios de longitud 3, tal como indicamos en la tabla 2.5. De esta forma, si el emisor envía el mensaje “norte” codificado como (0,0,0) y se produce un error en la transmisión, entonces el receptor recibirá uno de los vectores (0,0,1), (0,1,0) o (1,0,0) y, por tanto, sabrá que se ha producido un error ya que ninguno de estos vectores es un vector código. Sin embargo, no sabrá en qué componente se ha producido el error (¿por qué?), con lo que deberá solicitar la retransmisión del mensaje.

El código definido por la tabla 2.5 es un **código detector de un error** porque permite detectar un error, aunque no permite su corrección. Notemos que dicho código se ha obtenido añadiendo a cada una de las palabras código de la tabla 2.4 una componente extra de manera que el número total de 1 de la nueva palabra código sea par. Se trata pues, de un **código de control de paridad**.

En general, si el mensaje que queremos codificar corresponde al vector binario  $\mathbf{b} = (b_1, b_2, \dots, b_n) \in \mathbb{Z}_2^n$ , entonces el vector código de control de paridad es  $\mathbf{v} = (b_1, b_2, \dots, b_n, d) \in$

$\mathbb{Z}_2^{n+1}$ , donde el dígito de control  $d$  se elige de modo que

$$b_1 +_2 b_2 +_2 \cdots +_2 b_{n+1} +_2 d = 0 \quad \text{en } \mathbb{Z}_2$$

o equivalentemente

$$\langle \mathbf{1}, \mathbf{v} \rangle = 0,$$

donde  $\mathbf{1}$  es el vector de longitud  $n+1$  con todas las componentes iguales a 1. Decimos que  $\mathbf{1}$  es el **vector de control de paridad** ya que si el receptor recibe el vector  $\mathbf{u}$  y  $\langle \mathbf{1}, \mathbf{u} \rangle = 1$ , entonces podemos afirmar que ha ocurrido un error.

Podemos generalizar la aritmética módulo 2 a cualquier entero  $q$ . Así, el conjunto de los **enteros módulo  $q$**  es

$$\mathbb{Z}_q = \{0, 1, 2, \dots, q-1\}$$

con la adición y la multiplicación definidas como

$$a +_q b = (a + b) \text{ mód } q \quad \text{y} \quad a \cdot_q b = (a \cdot b) \text{ mód } q.$$

Así, la suma y el producto de dos elementos de  $\mathbb{Z}_q$  es también un elemento de  $\mathbb{Z}_q$  con lo que  $\mathbb{Z}_q$  es **cerrado** para dichas operaciones.

Por ejemplo, para  $q = 5$  las tablas de la adición y la multiplicación son

$+_q$	0	1	2	3	4	$\cdot_q$	0	1	2	3	4
0	0	1	2	3	4	0	0	0	0	0	0
1	1	2	3	4	0	1	0	1	2	3	4
2	2	3	4	0	1	2	0	2	4	1	3
3	3	4	0	1	2	3	0	3	1	4	2
4	4	0	1	2	3	4	0	4	3	2	1

Es fácil comprobar que si  $q$  es un número primo, entonces  $\mathbb{Z}_q$  se comporta en muchos aspectos como  $\mathbb{R}$ : podemos sumar, restar, multiplicar y dividir (por elementos no nulos); en otras palabras,  $(\mathbb{Z}_q, +_q, \cdot_q)$  satisface las propiedades de un cuerpo.

La extensión de la suma y la multiplicación de  $\mathbb{Z}_q$  a la suma de vectores y producto de un escalar por un vector es directa. Así, en  $\mathbb{Z}_5^4$  si

$$\mathbf{u} = (2, 0, 1, 3), \quad \mathbf{v} = (3, 1, 0, 4) \quad \text{y} \quad \alpha = 3$$

entonces

$$\mathbf{u} +_5 \mathbf{v} = (2 +_5 3, 0 +_5 1, 1 +_5 0, 3 +_5 4) = (0, 1, 1, 2),$$

$$\alpha \mathbf{u} = (3 \cdot_5 2, 3 \cdot_5 0, 3 \cdot_5 1, 3 \cdot_5 3) = (1, 0, 3, 4)$$

y

$$\langle \mathbf{u}, \mathbf{v} \rangle = 2 \cdot_5 3 +_5 0 \cdot_5 1 +_5 1 \cdot_5 0 +_5 3 \cdot_5 4 = 1 +_5 0 +_5 0 +_5 2 = 3.$$

En general, nos referiremos a los vectores de  $\mathbb{Z}_q^n$  como **vectores  $q$ -arios de longitud  $n$** .

Además, escribiremos  $+$  y  $\cdot$  en lugar de  $+_q$  y  $\cdot_q$  respectivamente, cuando no haya posibilidad de confusión con el entero  $q$  que estemos considerando.

Terminamos esta sección con la descripción de un código que permite detectar errores simples.

**Ejemplo 2.8.2** Cada libro tiene asociado un número llamado ISBN (*International Standard Book Number*) formado por diez dígitos. Los siguientes son algunos ejemplos de ISBN

0-13-067464-6, 0-444-85193-3, 3-540-97450-4, 3-540-00706-7,  
84-8454-081-2, 84-404-8586-7, 84-89727-35-X, 968-880-799-0.

El primer grupo representa el país de edición del libro, el segundo representa la editorial, el tercero es el número que la editorial asigna al libro y el último dígito es redundante. Los guiones tienen como única finalidad facilitar la lectura, pero son ignorados desde un punto de vista formal.

Para que

$$d_1d_2d_3d_4d_5d_6d_7d_8d_9d_{10}$$

sea un ISBN debe cumplir las dos condiciones siguientes:

- $0 \leq d_i \leq 9$  para  $i = 1, 2, 3, 4, 5, 6, 7, 8, 9$ .
- $d_{10} = (d_1 + 2d_2 + 3d_3 + 4d_4 + 5d_5 + 6d_6 + 7d_7 + 8d_8 + 9d_9) \bmod 11$ . Si  $d_{10} = 10$ , escribimos X en lugar de 10 ya que, para evitar confusiones, es preferible que cada componente de un ISBN sea un solo dígito.

Notemos que al ser  $-1 = 10 \bmod 11$  (¿por qué?), podemos escribir la última expresión como

$$d_1 + 2d_2 + 3d_3 + 4d_4 + 5d_5 + 6d_6 + 7d_7 + 8d_8 + 9d_9 + 10d_{10} = 0$$

con las operaciones realizadas en  $\mathbb{Z}_{11}$ , o equivalentemente

$$\langle (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8, d_9, d_{10}), (1, 2, 3, 4, 5, 6, 7, 8, 9, 10) \rangle = 0$$

es decir  $(1, 2, 3, 4, 5, 6, 7, 8, 9, 10)$  es el vector de control del código ISBN.

Ahora para averiguar si 84-08-04214-9 y 0-521-59374-1 corresponden a ISBN correctos, calculamos (en  $\mathbb{Z}_{11}$ )

$$\begin{aligned} & 1 \cdot 8 + 2 \cdot 4 + 3 \cdot 0 + 4 \cdot 8 + 5 \cdot 0 \\ & + 6 \cdot 4 + 7 \cdot 2 + 8 \cdot 1 + 9 \cdot 4 + 10 \cdot 9 = 0 \end{aligned}$$

y

$$\begin{aligned} & 1 \cdot 0 + 2 \cdot 5 + 3 \cdot 2 + 4 \cdot 1 + 5 \cdot 5 \\ & + 6 \cdot 9 + 7 \cdot 3 + 8 \cdot 7 + 9 \cdot 4 + 10 \cdot 1 = 2 \end{aligned}$$

con lo que el primer número representa un ISBN correcto, mientras que el segundo no lo es.

Hemos introducido los códigos detectores de errores. Ahora construimos códigos que pueden *corregir*, además de detectar, ciertos errores. En este caso nuestro mensaje es

un vector  $\mathbf{x} \in \mathbb{Z}_2^k$  para algún  $k$  que codificamos como  $\mathbf{c} = G\mathbf{x}$ , donde  $G$  es una matriz de tamaño  $n \times k$  de rango  $k$  con coeficientes en  $\mathbb{Z}_2$ . Decimos que  $G$  es la **matriz generadora** del código  $\mathcal{C} = \text{Col}(G)$  y que  $\mathcal{C}$  es un  $[n,k]$ -**código lineal**. Veamos un ejemplo sencillo que nos ayudará a entender este concepto.

**Ejemplo 2.8.3** Supongamos que el mensaje o información que queremos codificar es un bit 0 o 1, que codificamos como  $(0,0)$  y  $(1,1)$  respectivamente. El código  $\{(0,0), (1,1)\}$  puede detectar errores simples. Por ejemplo si el emisor envía  $(0,0)$  y ocurre un error en la transmisión de manera que el receptor recibe  $(0,1)$ , entonces éste sabe que se ha cometido un error pues  $(0,1)$  no es un vector código. Sin embargo, el receptor no puede corregir dicho error ya que  $(0,1)$  podría ser el resultado de un error en la primera componente del vector código  $(0,0)$  o en la segunda componente del vector código  $(1,1)$ .

Para solventar este problema, codificamos 0 y 1 como  $(0,0,0)$  y  $(1,1,1)$  respectivamente. Ahora, si el emisor envía  $(0,0,0)$  y se produce un error en la transmisión, entonces el receptor recibirá uno de los vectores  $(0,0,1)$ ,  $(0,1,0)$  o  $(1,0,0)$ . El receptor sabe entonces que se ha producido un error que, además puede corregir, es decir, el receptor sabe que el vector código enviado fue  $(0,0,0)$  ya que un único error en el vector código  $(1,1,1)$  no puede producir ninguno de los vectores  $(0,0,1)$ ,  $(0,1,0)$  y  $(1,0,0)$ .

Notemos que podemos conseguir este código mediante la matriz

$$G = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

ya que

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} 0 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} 1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Claramente

$$\mathcal{C} = \text{Col}(G) = \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

es un  $[3,1]$ -código lineal (¿por qué?).

Si el receptor recibe el vector  $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$  necesita que  $c_1 = c_2 = c_3$  para que sea un vector código. Podemos escribir las igualdades anteriores como  $c_1 = c_2$  y  $c_1 = c_3$ , o equivalentemente como el sistema lineal

$$\begin{array}{rcl} c_1 & + & c_2 & = & 0 \\ c_1 & & + & c_3 & = & 0 \end{array} \left. \right\}$$

con coeficientes en  $\mathbb{Z}_2$  (recordad las tablas de la adición y la multiplicación en  $\mathbb{Z}_2$  introducidas en la sección ??). El sistema anterior es equivalente a la ecuación matricial

$$H\mathbf{c} = \mathbf{0}$$

donde  $H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$  es la **matriz de control de paridad** del código. Notemos que

$$HG = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = O.$$

**Definición 2.8.3** Sea  $k < n$ . Decimos que la matriz  $G$  con elementos en  $\mathbb{Z}_2$  es **sistemática** si

$$G = \begin{bmatrix} I_k \\ A \end{bmatrix}$$

donde  $A$  es una matriz de tamaño  $(n - k) \times k$ .

**Teorema 2.8.1** *Sea  $k < n$ . Si*

$$G = \begin{bmatrix} I_k \\ A \end{bmatrix}$$

*es la matriz generadora de un  $[n,k]$ -código binario  $\mathcal{C}$ , entonces*

$$H = [ A \quad I_{n-k} ]$$

*es una matriz de control de paridad de  $\mathcal{C}$ .*

*Además,  $\mathcal{C}$  corrige errores (simples) si y solo si las columnas de  $H$  son no nulas y distintas dos a dos.*

Consideremos de nuevo el código del ejemplo 2.8.3. Supongamos que el emisor envía un vector código desconocido para el receptor y que éste recibe el vector  $\mathbf{u} = (1, 1, 0)$ . Puesto que

$$H\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

tenemos que  $\mathbf{u}$  no es un vector código. El hecho de que las columnas de  $H$  sean no nulas y distintas dos a dos junto con que  $H\mathbf{u}$  coincide con la tercera columna de  $H$  pone de manifiesto, de acuerdo con el teorema 2.8.1, que la componente errónea de  $\mathbf{u}$  es la tercera; por tanto, la palabra código enviada fue  $(1, 1, 1)$ .

**Ejemplo 2.8.4** Supongamos que queremos diseñar un código binario corrector de un error que utilice tres ecuaciones para el control de la paridad; entonces  $n - k = 3$  y, en consecuencia  $k = n - 3$ .

Los vectores de información son vectores de  $\mathbb{Z}_2^k$ , por tanto, necesitamos que  $k$ , y por tanto  $n$ , sean tan grandes como sea posible para que podamos transmitir la máxima información. Sabemos, de acuerdo con el teorema 2.8.1 que las  $n$  columnas de  $H$  han de ser no nulas y

distintas dos a dos, así que el máximo valor de  $n$  corresponde a los  $2^3 - 1 = 7$  vectores distintos de  $\mathbb{Z}_2^3$ . Así pues, podemos tomar

$$H = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix},$$

con lo que, por el teorema 2.8.1

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}$$

es una matriz generadora sistemática del código.

Supongamos ahora que disponemos del vector información  $\mathbf{x} = (0, 1, 1, 0)$  que codificamos como

$$\mathbf{c} = G\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Si el receptor recibe el vector anterior, puede afirmar que no se ha producido ningún error ya que

$$H\mathbf{c} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \mathbf{0}.$$

Supongamos ahora que el receptor recibe el vector  $\mathbf{u} = (0,1,1,0,1,0,0)$ . Entonces

$$H\mathbf{u} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \neq \mathbf{0}$$

con lo que el receptor detecta que se ha producido un error. Puesto que el vector anterior coincide con la sexta columna de  $H$ , concluimos que el error en el vector  $\mathbf{u}$  se encuentra en la sexta componente con lo que el vector código enviado es  $\mathbf{c} = (0,1,1,0,1,1,0)$ .

Ahora, como las primeras cuatro componentes de dicho vector corresponden al vector información original, tenemos que  $\mathbf{x} = (0,1,1,0)$ .

El código del ejemplo anterior recibe el nombre de **[7,4]-código de Hamming**.

## 2.9 Problemas

**2.1** Discute y resuelve, si es posible, el siguiente sistema.

$$\left. \begin{array}{rcl} x_1 + 2x_2 - x_3 + x_4 = 1 \\ x_1 - 3x_2 + x_3 - x_4 = 1 \\ x_1 + x_2 + x_3 + 2x_4 = 4 \end{array} \right\}$$

**2.2** Resuelve, si es posible, estos sistemas por el método de Gauss.

$$a) \left. \begin{array}{rcl} x_1 - x_2 + 2x_3 = 2 \\ -x_1 + 3x_2 + x_3 = 3 \\ x_1 + x_2 + 5x_3 = 7 \end{array} \right\} b) \left. \begin{array}{rcl} 2x_1 - x_2 + x_4 = 0 \\ x_1 - 2x_2 + x_3 = 0 \\ 5x_1 - x_2 + x_3 + x_4 = 0 \\ 5x_1 - 2x_2 - x_3 + 2x_4 = 0 \end{array} \right\}$$

$$c) \left. \begin{array}{rcl} 2x_1 - x_2 + x_4 = 9 \\ x_1 - 2x_2 + x_3 = 11 \\ 5x_1 - x_2 + x_3 + x_4 = 24 \\ 5x_1 - 2x_2 - x_3 + 2x_4 = 0 \end{array} \right\}$$

**2.3** Una inmobiliaria ha vendido un total de 65 plazas de garaje en tres urbanizaciones diferentes. Las ganancias obtenidas por la venta de una plaza de garaje en la urbanización A son de 2000 €, 4000 € por una en la urbanización B y 6000 por una en la urbanización C. Se sabe que se han vendido un 50 % más de plazas en la urbanización A que en la urbanización C. Calcula el número de plazas de garaje vendidas en cada urbanización sabiendo que el beneficio obtenido por las vendidas en la urbanización C es igual a la suma de los beneficios obtenidos por las vendidas en las urbanizaciones A y B.

**2.4** Resuelve, si es posible, el siguiente sistema en  $\mathbb{Z}_2$ :

$$\left. \begin{array}{l} x_1 + x_2 + x_3 + x_4 = 1 \\ x_1 + x_2 = 1 \\ x_2 + x_3 = 0 \\ x_3 + x_4 = 0 \\ x_1 + x_4 = 1 \end{array} \right\}$$

**2.5** Resuelve el siguiente sistema en  $\mathbb{Z}_2$ .

$$\begin{array}{l} x+y=1 \\ y+z=0 \\ x+z=1 \end{array}$$

**2.6** Resuelve, si es posible, el siguiente sistema sobre  $\mathbb{Z}_3$

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

**2.7** Resuelve, si es posible, el siguiente sistema sobre  $\mathbb{Z}_5$

$$\left. \begin{array}{l} 3x_1 + 2x_2 = 1 \\ x_1 + 4x_2 = 1 \end{array} \right\}$$

**2.8** En una heladería, por una copa de la casa, dos horchatas y cuatro batidos te cobran 34 ?. Unos días después, por cuatro copas de la casa y cuatro horchatas te cobran 44 ?. Un tercer día, te piden 26 ? por una horchata y cuatro batidos. ¿Alguno de los días se han equivocado en la cuenta?

**2.9** Consideremos el alfabeto siguiente:

$$A \rightarrow 1 \ B \rightarrow 2 \ C \rightarrow 3 \ D \rightarrow 4 \ E \rightarrow 5 \ F \rightarrow 6 \ G \rightarrow 7 \ H \rightarrow 8 \ I \rightarrow 9 \ J \rightarrow 10$$

$$K \rightarrow 11 \ L \rightarrow 12 \ M \rightarrow 13 \ N \rightarrow 14 \ O \rightarrow 15 \ P \rightarrow 16 \ Q \rightarrow 17 \ R \rightarrow 19$$

$$S \rightarrow 20 \ T \rightarrow 21 \ U \rightarrow 22 \ V \rightarrow 23 \ W \rightarrow 24 \ X \rightarrow 25 \ Y \rightarrow 26 \ Z \rightarrow 27 \ . \rightarrow 28$$

$$, \rightarrow 29 ; \rightarrow 30 : \rightarrow 31 \% \rightarrow 32 ! \rightarrow 33 \rightarrow 34 ? \rightarrow 35 @ \rightarrow 36 \text{ espacio} \rightarrow 0$$

Supongamos que el mensaje que queremos cifrar, una vez identificado numéricamente mediante el alfabeto anterior, es el siguiente:

$$\begin{matrix} 1 & 0 & 18 & 22 & 9 & 5 & 14 & 0 & 3 & 1 \\ 4 & 19 & 22 & 7 & 1 & 29 & 0 & 4 & 9 & 16 \\ 20 & 0 & 12 & 5 & 0 & 1 & 26 & 22 & 4 & 1 \end{matrix}$$

Agrupamos elementos de cinco en cinco de este vector de números, con el fin de aplicar posteriormente sobre estos vectores una matriz de transformación que nos permita la

codificación. Agrupamos los números anteriores de la siguiente manera:

$$\begin{array}{ccccc} 1 & 0 & 18 & 22 & 9 \\ 5 & 14 & 0 & 3 & 1 \\ 4 & 19 & 22 & 7 & 1 \\ 29 & 0 & 4 & 9 & 16 \\ 20 & 0 & 12 & 5 & 0 \\ 1 & 26 & 22 & 4 & 1 \end{array}$$

Consideremos la siguiente matriz de cifrado

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

Se pide:

1. Cifra el mensaje anterior utilizando la matriz  $A$ .
2. Calcula la matriz de descifrado.
3. Supongamos que recibes el mensaje siguiente:

S!W?,E: ,DRG,R XQGWDHMSZBOHCRC

¿Puedes descifrarlo?



# 3 Determinante de una matriz

---

La capacidad para efectuar rápidamente operaciones aritméticas mentales parece tener sólo una moderada correlación con la inteligencia general y menor aún con la intuición y creatividad matemáticas. Algunos de los matemáticos más sobresalientes han tenido dificultades al operar, y muchos *calculistas ultrarrápidos* profesionales (aunque no los mejores) han sido torpes en todas las demás capacidades mentales. Sin embargo, algunos grandes matemáticos han sido también diestros calculistas mentales. Carl Friedrich Gauss por ejemplo, podía llevar a cabo prodigiosas hazañas matemáticas en la mente. Le gustaba hacer alarde de que aprendió antes a calcular que a hablar. Se cuenta que en cierta ocasión su padre, de oficio albañil, estaba confeccionando la nómina general de sus empleados, cuando Friedrich, que entonces tenía 3 años, le interrumpió diciéndole: «Papá, la cuenta está mal...». Al volver a sumar la larga lista de números se comprobó que la suma correcta era la indicada por el niño. Nadie le había enseñado nada de aritmética. John von Neumann era un genio matemático que también estuvo dotado de este poder peculiar de computar sin usar lápiz ni papel. Robert Jungk habla en su libro Brighter than a Thousand Suns acerca de una reunión celebrada en Los Álamos, durante la Segunda Guerra Mundial, en la que von Neumann, Enrico Fermi, Edward Teller y Richard Feynman lanzaban continuamente ideas. Siempre que había que efectuar un cálculo matemático, Fermi, Feynman y von Neumann se ponían en acción. Fermi empleaba una regla de cálculo, Feynman una calculadora de mesa, y von Neumann su cabeza. «La cabeza», escribe Jungk (citando a otro físico), «terminaba normalmente la primera, y es notable lo próximas que estaban siempre las tres soluciones».

## 3.1 Introducción

Cada matriz cuadrada tiene asociado un escalar que llamamos determinante. La noción de determinante y su utilización para la resolución de sistemas lineales mediante la regla

de Cramer es (o al menos debería ser) conocida por los alumnos. Pero ésta no es la única utilidad de los determinantes. Posiblemente la propiedad más interesante sea la caracterización de la invertibilidad de las matrices cuadradas, así como la fórmula que permite calcular la inversa de una matriz invertible.

De entre las distintas formas que suelen utilizarse para definir el concepto de determinante utilizamos una que es inductiva, es decir, empezamos definiendo el determinante de una matriz de tamaño  $1 \times 1$  y a partir de él, definimos el determinante de cualquier matriz de tamaño  $n \times n$  en función del determinante de ciertas matrices de tamaño  $(n - 1) \times (n - 1)$ .

Si  $A = [a_{ij}]$  es una matriz de tamaño  $n \times n$ , denotamos por  $M_{ij}(A)$  la matriz de tamaño  $(n - 1) \times (n - 1)$  que obtenemos al eliminar la  $i$ -ésima fila y la  $j$ -ésima columna de  $A$ . Así, para

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

tenemos que

$$M_{11}(A) = [a_{22}], \quad M_{12}(A) = [a_{21}],$$

$$M_{21}(A) = [a_{12}], \quad M_{22}(A) = [a_{11}].$$

En cambio, si

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

tenemos que

$$M_{11}(A) = \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{12}(A) = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{13}(A) = \begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix},$$

$$M_{21}(A) = \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix}, \quad M_{22}(A) = \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix}, \quad M_{23}(A) = \begin{bmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{bmatrix},$$

$$M_{31}(A) = \begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix}, \quad M_{32}(A) = \begin{bmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{bmatrix}, \quad M_{33}(A) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

**Definición 3.1.1** Sea  $A = [a_{ij}]$  una matriz de tamaño  $n \times n$ . Llamamos **determinante** de  $A$ , y escribimos  $\det A$ , al escalar

$$\det A = \begin{cases} a_{11}, & \text{si } n = 1 \\ \sum_{j=1}^n a_{1j}(-1)^{1+j} \det M_{1j}(A), & \text{si } n > 1 \end{cases}$$

Antes de continuar, introducimos una notación que será útil en todo lo que sigue.

**Definición 3.1.2** Sea  $A = [a_{ij}]$  una matriz de tamaño  $n \times n$ . Llamamos **cofactor**  $(i,j)$  de  $A$ , y escribimos  $C_{ij}(A)$ , al escalar

$$C_{ij}(A) = (-1)^{i+j} \det M_{ij}(A).$$

Llamamos **menor**  $(i,j)$  de  $A$  al escalar  $\det M_{ij}(A)$ .

Así, de acuerdo con la definición anterior y la definición 3.1.1 tenemos que

$$\det A = \sum_{j=1}^n a_{1j} \mathcal{C}_{1j}(A).$$

Decimos que la expresión anterior, para  $n > 1$ , es la **expansión de Laplace** del determinante de  $A$  por los cofactores de la primera fila de  $A$ .

Veamos con detalle el cálculo de  $\det A$  para una matriz  $A = [a_{ij}]$  de tamaño  $n \times n$  para algunos valores de  $n$ .

Si  $n = 1$ , claramente  $\det A = a_{11}$ .

Si  $n = 2$ , entonces

$$\begin{aligned}\det A &= a_{11} \mathcal{C}_{11}(A) + a_{12} \mathcal{C}_{12}(A) \\ &= a_{11}(-1)^{1+1} \det M_{11}(A) + a_{12}(-1)^{1+2} \det M_{12}(A) \\ &= a_{11}a_{22} - a_{12}a_{21}\end{aligned}\tag{3.1}$$

por tanto, para calcular el determinante de una matriz de tamaño  $2 \times 2$ , basta con restar al producto de los elementos de la diagonal principal de la matriz, el producto de los elementos de la diagonal secundaria. Por ejemplo, si

$$A = \begin{bmatrix} 1 & -3 \\ 2 & 4 \end{bmatrix},$$

entonces  $\det A = 1 \cdot 4 - (-3) \cdot 2 = 10$ .

Otra notación habitual para el determinante de una matriz consiste en utilizar un par de líneas verticales para delimitar los elementos de la matriz, en lugar de los habituales corchetes. Así, podemos escribir los cálculos anteriores como

$$\left| \begin{array}{cc} 1 & -3 \\ 2 & 4 \end{array} \right| = 4 + 6 = 10.$$

Notemos que podemos escribir la expresión (3.1) como

$$\begin{aligned}\det A &= a_{21}(-a_{12}) + a_{22}a_{11} \\ &= a_{21}(-1)^{2+1} \det M_{21}(A) + a_{22}(-1)^{2+2} \det M_{22}(A) \\ &= a_{21} \mathcal{C}_{21}(A) + a_{22} \mathcal{C}_{22}(A).\end{aligned}$$

De hecho, podemos probar (queda como ejercicio) que también se satisfacen las igualdades siguientes:

$$\begin{aligned}\det A &= a_{11} \mathcal{C}_{11}(A) + a_{21} \mathcal{C}_{21}(A) \\ &= a_{12} \mathcal{C}_{12}(A) + a_{22} \mathcal{C}_{22}(A).\end{aligned}$$

Si  $n = 3$ , entonces

$$\begin{aligned}\det A &= a_{11} \mathcal{C}_{11}(A) + a_{12} \mathcal{C}_{12}(A) + a_{13} \mathcal{C}_{13}(A) \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.\end{aligned}$$

Esta fórmula, que permite calcular  $\det A$  para  $n = 3$ , se conoce con el nombre de **regla de Sarrus**.

Así, para calcular  $\det A$  con  $A = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 3 \\ 1 & 3 & 8 \end{bmatrix}$  escribimos

$$\begin{array}{c} 3 \quad 1 \quad 2 \quad 27 \quad -8 \\ -1 \quad 1 \quad 3 \quad 1 \quad 1 \\ 1 \quad 3 \quad 8 \quad 24 \quad 3 \quad -6 \end{array}$$

ahora sumamos los tres productos de la parte inferior (los correspondientes a las diagonales descendentes) y restamos los tres productos de la parte superior (los correspondientes a las diagonales ascendentes) con lo que obtenemos

$$\det A = (24 + 3 - 6) - (2 + 27 - 8) = 0.$$

Podemos probar (queda como ejercicio) que también se satisfacen las igualdades:

$$\begin{aligned}\det A &= a_{21} \mathcal{C}_{21}(A) + a_{22} \mathcal{C}_{22}(A) + a_{23} \mathcal{C}_{23}(A) \\ &= a_{31} \mathcal{C}_{31}(A) + a_{32} \mathcal{C}_{32}(A) + a_{33} \mathcal{C}_{33}(A) \\ &= a_{11} \mathcal{C}_{11}(A) + a_{21} \mathcal{C}_{21}(A) + a_{31} \mathcal{C}_{31}(A) \\ &= a_{12} \mathcal{C}_{12}(A) + a_{22} \mathcal{C}_{22}(A) + a_{32} \mathcal{C}_{32}(A) \\ &= a_{13} \mathcal{C}_{13}(A) + a_{23} \mathcal{C}_{23}(A) + a_{33} \mathcal{C}_{33}(A)\end{aligned}$$

Para determinantes de matrices de tamaño  $n \times n$  con  $n \geq 4$  no existen reglas prácticas como las anteriores.

Hemos visto que la expansión de Laplace del determinante de matrices de tamaños  $2 \times 2$  y  $3 \times 3$  es independiente de la fila que elijamos; más aún, en dichos casos, el valor del determinante coincidía con la expansión de Laplace por los elementos de cualquiera

de las columnas. El teorema siguiente, establece que esto también es cierto para matrices de tamaño  $n \times n$  en general.

**Teorema 3.1.1** *Sea  $A = [a_{ij}]$  una matriz de tamaño  $n \times n$ . Entonces*

$$1. \det A = \sum_{j=1}^n a_{ij} \mathcal{C}_{ij}(A), \text{ para } i = 1, 2, \dots, n.$$

$$2. \det A = \sum_{i=1}^n a_{ij} \mathcal{C}_{ij}(A), \text{ para } j = 1, 2, \dots, n.$$

El cálculo del determinante de una matriz de tamaño  $n \times n$  implica, de acuerdo con la definición 3.1.1, el cálculo del determinante de  $n$  matrices de tamaño  $(n - 1) \times (n - 1)$ . Puesto que, por el teorema anterior, el valor de dicho determinante es independiente de la fila o columna que elijamos para su expansión, podemos ahorrarnos un poco de trabajo eligiendo la fila o columna que más ceros contenga. Veamos un ejemplo.

**Ejemplo 3.1.1** Para calcular el determinante de la matriz

$$A = \begin{bmatrix} 2 & 3 & 0 & -1 & 6 \\ 3 & 8 & 0 & 0 & 2 \\ -1 & 7 & 0 & 0 & 4 \\ 0 & 3 & -2 & -4 & 9 \\ 0 & 5 & 0 & 0 & 0 \end{bmatrix}$$

elegimos la tercera columna que contiene cuatro ceros (podríamos haber elegido la quinta fila que también contiene cuatro ceros), con lo que

$$\det A = (-2)(-1)^{4+3} \begin{vmatrix} 2 & 3 & -1 & 6 \\ 3 & 8 & 0 & 2 \\ -1 & 7 & 0 & 4 \\ 0 & 5 & 0 & 0 \end{vmatrix}.$$

Ahora, para calcular el determinante anterior, elegimos la cuarta fila que contiene tres ceros (podríamos haber elegido la tercera columna que también contiene tres ceros), por tanto

$$\det A = 2 \cdot 5(-1)^{4+2} \begin{vmatrix} 2 & -1 & 6 \\ 3 & 0 & 2 \\ -1 & 0 & 4 \end{vmatrix}.$$

Puesto que la columna con más elementos nulos en el determinante anterior es la segunda, tenemos que

$$\det A = 2 \cdot 5 \cdot (-1)(-1)^{1+2} \begin{vmatrix} 3 & 2 \\ -1 & 4 \end{vmatrix} = 10(12 + 2) = 140.$$

Desafortunadamente no todos los determinantes se pueden reducir tan fácilmente como ha ocurrido en el ejemplo anterior. Así, la aplicación de la expansión de Laplace es, en general, inviable ya que requiere del orden de  $n!$  multiplicaciones. Un supercomputador que pudiera hacer un billón de operaciones por segundo tendría que trabajar durante más de 500 000 años para calcular el determinante de una matriz de tamaño  $25 \times 25$  ya que  $25! \approx 1.55 \times 10^{25}$ .

Por tanto, para calcular determinantes de matrices de tamaño mayor que 4, debemos utilizar las propiedades que introducimos en la sección siguiente, que permiten introducir algunos ceros en algunas posiciones de manera que hacen viable la aplicación de la expansión de Laplace.

## 3.2 Propiedades de los determinantes

En esta sección introducimos las propiedades que nos permitirán calcular el determinante de cualquier matriz cuadrada sin necesidad de calcular los cofactores de una fila o una columna.

**Teorema 3.2.1** *Si  $A$  es una matriz cuadrada con una fila o una columna nula, entonces  $\det A = 0$ .*

La demostración se basa en la expansión de Laplace por los cofactores de la fila o la columna nula resulta evidente que  $\det A = 0$ .

**Teorema 3.2.2**  $\det A = \det A^T$  para cualquier matriz cuadrada  $A$ .

Como consecuencia del teorema anterior tenemos que cualquier propiedad sobre determinantes que demostremos en términos de las filas de una matriz es automáticamente válida enunciada por columnas.

A continuación enunciamos algunas propiedades interesantes de los determinantes, quedando las demostraciones como ejercicio para el lector.

**Teorema 3.2.3** *Sea  $A$  una matriz cuadrada. Si  $A$  tiene dos filas o dos columnas idénticas, entonces  $\det A = 0$ .*

**Teorema 3.2.4** Si  $A = [a_{ij}]$  es una matriz de tamaño  $n \times n$ , entonces

$$\sum_{k=1}^n a_{pk} \mathcal{C}_{qk}(A) = \begin{cases} \det A, & \text{si } p = q \\ 0, & \text{si } p \neq q \end{cases}$$

Ahora, consecuencia inmediata del teorema anterior y del teorema ?? tenemos que si  $A = [a_{ij}]$  es una matriz de tamaño  $n \times n$ , entonces

$$\sum_{k=1}^n a_{kp} \mathcal{C}_{kq}(A) = \begin{cases} \det A, & \text{si } p = q \\ 0, & \text{si } p \neq q \end{cases}.$$

El ejemplo siguiente pone de manifiesto que, en general, para dos matrices cuadradas  $A$  y  $B$  del mismo tamaño, tenemos que

$$\det(A + B) \neq \det A + \det B.$$

**Ejemplo 3.2.1** Para las matrices

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

tenemos que

$$\det A = \begin{vmatrix} 1 & 1 \\ 0 & 2 \end{vmatrix} = 2 \quad \text{y} \quad \det B = \begin{vmatrix} 2 & 0 \\ -1 & 3 \end{vmatrix} = 6$$

mientras que

$$\det(A + B) = \begin{vmatrix} 3 & 1 \\ -1 & 5 \end{vmatrix} = 16.$$

Por tanto,  $\det(A + B) \neq \det A + \det B$ .

El siguiente resultado se basa en las propiedades de las matrices elementales.

**Teorema 3.2.5**

1. Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  multiplicando una fila cualquiera por  $\alpha$ , entonces  $\det B = \alpha \det A$ .
2. Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  sumando a una fila un múltiplo de otra, entonces  $\det B = \det A$ .

**3.** Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  intercambiando la posición de dos filas, entonces  $\det B = -\det A$ .

### Ejemplo 3.2.2

1. Consideremos las matrices

$$A = \begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 1 \\ 1 & 2 & 3 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 1 & 3 & 1 \\ 4 & 10 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Puesto que la segunda fila de  $B$  es 2 veces la segunda fila de  $A$ , tenemos que

$$\det B = 2 \det A.$$

Notemos que (¡comprobadlo!)

$$\det A = -3 \quad \text{y} \quad \det B = -6.$$

2. Consideremos ahora las matrices

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

Claramente la tercera fila de  $B$  se obtiene sumando a la tercera fila de  $A$  la primera fila de  $A$ , por tanto

$$\det A = \det B.$$

En este caso (¡comprobadlo!)

$$\det A = -4 = \det B.$$

3. Finalmente, consideremos las matrices

$$A = \begin{bmatrix} 0 & 4 & 1 \\ 1 & 2 & 3 \\ 4 & 1 & 5 \end{bmatrix} \quad \text{y} \quad B = \begin{bmatrix} 4 & 1 & 5 \\ 1 & 2 & 3 \\ 0 & 4 & 1 \end{bmatrix}.$$

Puesto que la matriz  $B$  se obtiene intercambiando la primera y la tercera fila de  $A$ , tenemos que

$$\det A = -\det B.$$

Ahora (¡comprobadlo!)

$$\det A = 31 \quad \text{y} \quad \det B = -31.$$

Consecuencia inmediata de los resultados anteriores, establecemos la conclusión siguiente. Si  $A$  es una matriz de tamaño  $n \times n$ , entonces

$$\det(\alpha A) = \alpha^n \det(A).$$

Así

$$\begin{aligned} \det \begin{bmatrix} 5 & 15 & 5 \\ 10 & 25 & 5 \\ 5 & 10 & 15 \end{bmatrix} &= \det \left( 5 \begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 1 \\ 1 & 2 & 3 \end{bmatrix} \right) \\ &= 5^3 \det \begin{bmatrix} 1 & 3 & 1 \\ 2 & 5 & 1 \\ 1 & 2 & 3 \end{bmatrix} = 125(-3) = -475. \end{aligned}$$

Los siguientes resultados son muy interesantes en el cálculo práctico de determinantes.

### Teorema 3.2.6

1. Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  multiplicando una columna cualquiera por  $\alpha$ , entonces  $\det B = \alpha \det A$ .
2. Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  sumando a una columna un múltiplo de otra, entonces  $\det B = \det A$ .
3. Si  $B$  es la matriz que se obtiene a partir de la matriz  $A$  intercambiando la posición de dos columnas, entonces  $\det B = -\det A$ .

Resulta sencillo comprobar que si la matriz cuadrada  $A$  tiene una fila (o columna) que es combinación lineal de las restantes filas (o columnas), entonces  $\det A = 0$ .

Además, si las filas (o columnas) de una matriz cuadrada  $A$  son linealmente dependientes, entonces  $\det A = 0$ .

Ahora tenemos ya todas las condiciones para establecer uno de los resultados más importantes de la teoría de determinantes.

### Teorema 3.2.7

Sea  $A$  una matriz cuadrada.  $A$  es invertible si y solo si  $\det A \neq 0$ .

**Teorema 3.2.8 (Teorema de Binet-Cauchy)** Si  $A$  y  $B$  son matrices cuadradas del mismo tamaño, entonces

$$\det(AB) = \det A \cdot \det B.$$

Ahora, del teorema anterior y del hecho de que para una matriz invertible  $A$  se satisface la igualdad  $AA^{-1} = I$ , tenemos que si  $A$  es invertible, entonces  $\det A \neq 0$  y

$$\det A^{-1} = \frac{1}{\det A}.$$

### 3.3 Regla de Cramer

En esta sección introducimos un método para calcular la inversa de una matriz invertible  $A$  y la conocida regla de Cramer que permite resolver el sistema lineal  $Ax = \mathbf{b}$  utilizando determinantes.

#### La regla de Cramer.

Gabriel Cramer, nacido en Ginebra (Suiza) en 1704, fue un profesor de matemáticas de la Universidad de Ginebra durante el período 1724-27. En 1750 ocupó la cátedra de filosofía en dicha universidad. En 1731 presentó ante la Academia de las Ciencias de París, una memoria sobre las múltiples causas de la inclinación de las órbitas de los planetas. Falleció en 1752.

La Regla de Cramer es un teorema en álgebra lineal, que da la solución de un sistema lineal de ecuaciones en términos de determinantes. Recibe este nombre en honor a Gabriel Cramer (1704 - 1752), quien publicó la regla en su *Introduction à l'analyse des lignes courbes algébriques* de 1750, aunque Colin Maclaurin también publicó el método en su *Treatise of Geometry* de 1748 (y probablemente conocía el método desde 1729).

**Definición 3.3.1** Sea  $A = [a_{ij}]$  una matriz de tamaño  $n \times n$ . Llamamos **matriz adjunta** de  $A$ , y escribimos  $\text{Adj}(A)$ , a la matriz cuyo elemento  $(i,j)$  es el cofactor  $(j,i)$  de  $A$ .

Notemos que para obtener  $\text{Adj}(A)$  debemos sustituir cada elemento de  $A$  por su cofactor y a continuación trasponer la matriz resultante. Veamos un ejemplo.

**Ejemplo 3.3.1** Para la matriz

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

tenemos que

$$\begin{aligned}\mathcal{C}_{11}(A) &= (-1)^{1+1} \begin{vmatrix} -1 & 2 \\ -3 & 0 \end{vmatrix} = 6, \\ \mathcal{C}_{12}(A) &= (-1)^{1+2} \begin{vmatrix} 3 & 2 \\ 1 & 0 \end{vmatrix} = 2, \\ \mathcal{C}_{13}(A) &= (-1)^{1+3} \begin{vmatrix} 3 & -1 \\ 1 & -3 \end{vmatrix} = -8, \\ \mathcal{C}_{21}(A) &= (-1)^{2+1} \begin{vmatrix} 0 & -2 \\ -3 & 0 \end{vmatrix} = 6, \\ \mathcal{C}_{22}(A) &= (-1)^{2+2} \begin{vmatrix} 1 & -2 \\ 1 & 0 \end{vmatrix} = 2, \\ \mathcal{C}_{23}(A) &= (-1)^{2+3} \begin{vmatrix} 1 & 0 \\ 1 & -3 \end{vmatrix} = 3, \\ \mathcal{C}_{31}(A) &= (-1)^{3+1} \begin{vmatrix} 0 & -2 \\ -1 & 2 \end{vmatrix} = -2, \\ \mathcal{C}_{32}(A) &= (-1)^{3+2} \begin{vmatrix} 1 & -2 \\ 3 & 2 \end{vmatrix} = -8, \\ \mathcal{C}_{33}(A) &= (-1)^{3+3} \begin{vmatrix} 1 & 0 \\ 3 & -1 \end{vmatrix} = -1,\end{aligned}$$

por tanto

$$\text{Adj}(A) = \begin{bmatrix} \mathcal{C}_{11}(A) & \mathcal{C}_{21}(A) & \mathcal{C}_{31}(A) \\ \mathcal{C}_{12}(A) & \mathcal{C}_{22}(A) & \mathcal{C}_{32}(A) \\ \mathcal{C}_{13}(A) & \mathcal{C}_{23}(A) & \mathcal{C}_{33}(A) \end{bmatrix} = \begin{bmatrix} 6 & 6 & -2 \\ 2 & 2 & -8 \\ -8 & 3 & -1 \end{bmatrix}.$$

**Teorema 3.3.1** Si  $A$  es una matriz cuadrada, entonces

$$A \text{Adj}(A) = (\det A) I_n = \text{Adj}(A) A. \quad (3.2)$$

Si además  $A$  es invertible, entonces

$$A^{-1} = \frac{1}{\det A} \text{Adj}(A).$$

Veamos un ejemplo.

**Ejemplo 3.3.2** La matriz  $A$  del ejemplo 3.3.1 es invertible ya que

$$\det A = 1\mathcal{C}_{11}(A) + 0\mathcal{C}_{12}(A) + (-2)\mathcal{C}_{13}(A) = 6 + 0 + 12 = 22 \neq 0,$$

por tanto,

$$A^{-1} = \frac{1}{\det A} \text{Adj}(A) = \frac{1}{22} \begin{bmatrix} 6 & 6 & -2 \\ 2 & 2 & -8 \\ -8 & 3 & -1 \end{bmatrix}.$$

Finalmente, introducimos la regla de Cramer para resolver un sistema lineal  $A\mathbf{x} = \mathbf{b}$  cuando  $A$  es una matriz invertible.

**Teorema 3.3.2 (Regla de Cramer)** *Sean  $A$  una matriz invertible de tamaño  $n \times n$  y  $\mathbf{b}$  una  $n$ -tupla. Si  $B_i$ , para  $i = 1, 2, \dots, n$ , es la matriz que se obtiene al sustituir la  $i$ -ésima columna de  $A$  por  $\mathbf{b}$ , entonces*

$$x_i = \frac{\det B_i}{\det A}$$

*es la  $i$ -ésima componente de la única solución del sistema lineal  $A\mathbf{x} = \mathbf{b}$ .*

Veamos un ejemplo.

**Ejemplo 3.3.3** Consideremos el sistema lineal

$$\left. \begin{array}{rcl} 2x_1 - x_2 + x_3 = 5 \\ -2x_1 + 3x_2 - x_3 = -1 \\ x_1 - 2x_2 + 5x_3 = 13 \end{array} \right\}.$$

Mediante la regla de Sarrus tenemos que

$$\det A = \begin{vmatrix} 2 & -1 & 1 \\ -2 & 3 & -1 \\ 1 & -2 & 5 \end{vmatrix} = (30 + 1 + 4) - (3 + 4 + 10) = 18 \neq 0$$

con lo que  $A$  es invertible. Ahora

$$x_1 = \frac{\begin{vmatrix} 5 & -1 & 1 \\ -1 & 3 & -1 \\ 13 & -2 & 5 \end{vmatrix}}{\det A} = \frac{(75 + 13 + 2) - (39 + 10 + 5)}{18} = 2,$$

$$x_2 = \frac{\begin{vmatrix} 2 & 5 & 1 \\ -2 & -1 & -1 \\ 1 & 13 & 5 \end{vmatrix}}{\det A} = \frac{(-10 - 5 - 26) - (-1 - 26 - 50)}{18} = 2,$$

$$x_3 = \frac{\begin{vmatrix} 2 & -1 & 5 \\ -2 & 3 & -1 \\ 1 & -2 & 13 \end{vmatrix}}{\det A} = \frac{(78 + 1 + 20) - (15 + 4 + 26)}{18} = 3.$$

**Tabla 3.1** Número de operaciones del método de Gauss y de la regla de Cramer.

	MULTIPLICACIONES	ADICIONES
MÉTODO DE GAUSS	$\frac{1}{3}n^3 + n^2 - \frac{1}{3}n$	$\frac{1}{3}n^3 + \frac{1}{2}n^2 - \frac{5}{6}n$
REGLA DE CRAMER	$\frac{1}{3}n^4 + \frac{1}{3}n^3 + \frac{2}{3}n^2 + \frac{1}{3}n - 1$	$\frac{1}{3}n^4 - \frac{1}{6}n^3 - \frac{1}{3}n^2 + \frac{1}{6}n$

Es importante que el alumno entienda la regla de Cramer para resolver sistemas de ecuaciones lineales cuya matriz de coeficientes es invertible. Sin embargo, debemos insistir en que la importancia de la regla de Cramer radica en que proporciona una fórmula para la solución del sistemas lineales compatibles determinados con el mismo número de ecuaciones que de incógnitas, pero no suele utilizarse en la práctica.

El método de Gauss es más eficiente que la regla de Cramer ya que requiere menos operaciones. La tabla 3.1 muestra el número de operaciones necesarias para resolver un sistema de  $n$  ecuaciones con  $n$  incógnitas utilizando ambos métodos. Esto significa, por ejemplo, que podemos resolver un sistema lineal de 4 ecuaciones con 4 incógnitas mediante 36 multiplicaciones y 26 adiciones por el método de Gauss que contrasta con las 115 multiplicaciones y 70 adiciones que necesitamos utilizando la regla de Cramer. Notemos que, para  $n$  grande, el número de operaciones para el método de Gauss es del orden de  $2n^3/3$  mientras que para la regla de Cramer es del orden de  $2n^4/3$ .

### 3.4 Aplicación: puntos para representar curvas y superficies.

Recordemos que un sistema de ecuaciones lineales homogéneo con tantas ecuaciones como incógnitas tiene una solución no trivial (distinta de cero) si el determinante de la matriz de coeficientes es cero.

Esta característica que estudiamos en el tema de sistemas de ecuaciones lineales es fundamental cuando nos planteamos el problema de determinar curvas y superficies por medio de un conjunto finito de puntos.

#### Ecuación de una línea.

Comenzamos con el caso más simple. Por todos es conocido que dos puntos del plano  $A(x_1, y_1)$  y  $B(x_2, y_2)$  determinan una única recta en el plano. La ecuación general de la recta en el plano viene determinada por la ecuación:

$$r \equiv ax + by + c = 0.$$

Pero los puntos  $A$  y  $B$  están en la recta, luego podemos asegurar que verifican su ecuación, luego tenemos que:

$$\left. \begin{array}{l} ax_1 + by_1 + c = 0 \\ ax_2 + by_2 + c = 0 \end{array} \right\}$$

Podemos escribir el sistema de tres ecuaciones

$$\left. \begin{array}{l} ax + by + c = 0 \\ ax_1 + by_1 + c = 0 \\ ax_2 + by_2 + c = 0 \end{array} \right\}$$

que constituye un sistema lineal homogéneo de tres ecuaciones con tres incógnitas  $a,b,c$ , que son distintas de cero. esto significa que el sistema debe ser compatible indeterminado para que tenga más de una solución. Esto nos lleva directamente a afirmar que su determinante debe ser nulo, es decir,

$$\left| \begin{array}{ccc} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{array} \right| = 0.$$

### Un círculo que pasa por tres puntos.

De geometría analítica básica sabemos que por tres puntos del plano no alineados pasa una circunferencia. La ecuación de la circunferencia en el plano es:

$$c \equiv a(x^2 + y^2) + bx + cy + d = 0.$$

Siguiendo un proceso similar al descrito para la recta, tomamos tres puntos  $A(x_1,y_1)$ ,  $B(x_2,y_2)$  y  $C(x_3,y_3)$  y los sustituimos en la ecuación general de la circunferencia.

$$\left. \begin{array}{l} a(x_1^2 + y_1^2) + bx_1 + cy_1 + d = 0 \\ a(x_2^2 + y_2^2) + bx_2 + cy_2 + d = 0 \\ a(x_3^2 + y_3^2) + bx_3 + cy_3 + d = 0 \end{array} \right\}$$

Como hemos hecho anteriormente escribimos el sistema de ecuaciones

$$\left. \begin{array}{l} a(x^2 + y^2) + bx + cy + d = 0 \\ a(x_1^2 + y_1^2) + bx_1 + cy_1 + d = 0 \\ a(x_2^2 + y_2^2) + bx_2 + cy_2 + d = 0 \\ a(x_3^2 + y_3^2) + bx_3 + cy_3 + d = 0 \end{array} \right\}$$

cuyas incógnitas son  $a,b,c,d$ . Como todas no pueden ser nulas, esto significa que este sistema tiene más soluciones además de la trivial, lo que significa que

$$\left| \begin{array}{cccc} x^2 + y^2 & x & y & 1 \\ x_1^2 + y_1^2 & x_1 & y_1 & 1 \\ x_2^2 + y_2^2 & x_2 & y_2 & 1 \\ x_3^2 + y_3^2 & x_3 & y_3 & 1 \end{array} \right| = 0.$$

El cálculo de este determinante nos permite determinar la ecuación de la circunferencia que pasa por los puntos  $A,B,C$ .

- ¿Qué se obtiene al intersectar dos circunferencias en el plano?

- ¿Y si se intersectan tres circunferencias?

### Una cónica pasa por cinco puntos.

Cuando hablamos de cónicas nos estamos refiriendo a la parábola, la hipérbola o la elipse, así como a las formas degeneradas de estas curvas. La ecuación general de una cónica viene dada por la ecuación

$$C \equiv ax^2 + bxy + cy^2 + dx + ey + f = 0.$$

La primera cuestión que nos planteamos es ¿cuántos puntos necesitamos para determinar la ecuación de la sección cónica? En principio, si nos fijamos en la ecuación diríamos que necesitamos resolver 6 coeficientes, por lo que son necesarios 6 puntos. Sin embargo, podemos hacer una pequeña simplificación algebraica para reducir este número.

Sabemos que todos los coeficientes  $a, b, c, d, e, f$  no pueden ser cero, es decir, existe al menos un coeficiente distinto de 0, sea éste coeficiente  $b$ . Entonces podemos dividir por este coeficiente y obtendremos

$$C \equiv \frac{a}{b}x^2 + xy + \frac{c}{b}y^2 + \frac{d}{b}x + \frac{e}{b}y + \frac{f}{b} = 0,$$

con lo que hemos pasado a tener cinco coeficientes. Esto significa que basta con 5 puntos para determinar la ecuación de la cónica. Sean estos puntos  $A(x_1, y_1)$ ,  $B(x_2, y_2)$ ,  $C(x_3, y_3)$ ,  $D(x_4, y_4)$ ,  $E(x_5, y_5)$ .

Siguiendo un razonamiento similar al descrito en los casos anteriores llegamos que el determinante

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ x_1^2 & x_1y_1 & y_1^2 & x_1 & y_1 & 1 \\ x_2^2 & x_2y_2 & y_2^2 & x_2 & y_2 & 1 \\ x_3^2 & x_3y_3 & y_3^2 & x_3 & y_3 & 1 \\ x_4^2 & x_4y_4 & y_4^2 & x_4 & y_4 & 1 \\ x_5^2 & x_5y_5 & y_5^2 & x_5 & y_5 & 1 \end{vmatrix} = 0.$$

#### 3.4.1 Las órbitas de los asteroides.

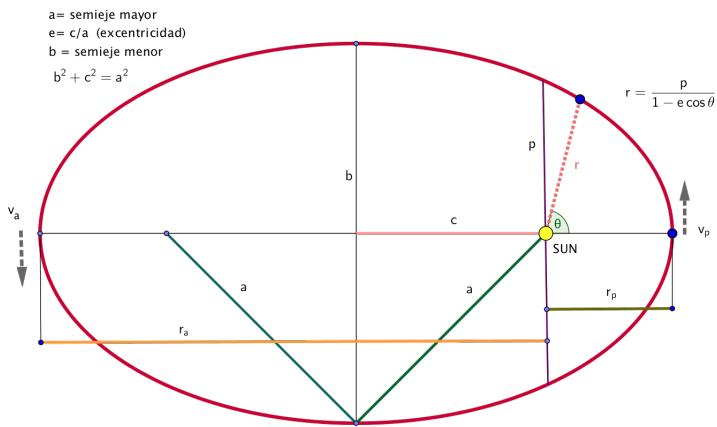
Los asteroides y los cometas son considerados restos de la gigantesca nube de gas y polvo que se condensó para crear el sol, los planetas y las lunas hace unos 4.500 millones de años. Hoy en día, la mayoría de los asteroides orbitan el sol en un cinturón apretado entre Marte y Júpiter. Los cometas son relegados a una nube o cinturón en el sistema de franjas solares. Diversos factores como las colisiones orbitales ocasionalmente alteran un asteroide o un cometa en un camino descarrilado.

Los asteroides son esencialmente trozos de roca que miden en tamaño de unos pocos pies a varios kilómetros de diámetro. (Los asteroides pequeños se llaman meteoroides.) El asteroide más grande, Ceres, tiene unos 950 kilómetros de ancho. Como la mayoría de los asteroides, se encuentra en el cinturón de asteroides entre Marte y Júpiter. Muchos astrónomos creen que el cinturón es un material primordial que nunca se proyectó en un planeta debido a la atracción gravitatoria de Júpiter. Otros astrónomos dicen que el cinturón es un planeta que se rompió durante una colisión.

El movimiento de los planetas y la mayoría de asteroides y cometas del sistema solar se rige por las conocidas leyes de Kepler del movimiento planetario, que afirman lo siguiente:

- Todos los planetas se mueven en torno al sol describiendo órbitas elípticas, con el sol en uno de sus focos.
- El radio vector que une un planeta y el sol barre áreas iguales en tiempos iguales.
- Para cualquier planeta, el cuadrado de su período orbital es directamente proporcional al cubo de la longitud del semieje mayor de su órbita elíptica.

Nos interesa la primera ley de Kepler donde se afirma que los planetas describen órbitas elípticas en torno al sol. En la figura 3.1 vemos una de estas elipses que describe un cuerpo que está en la posición  $r$  en torno al sol.



**Figura 3.1** Representación gráfica de la órbita elíptica de un asteroide..

En el valor 0, es decir,  $\theta = 0$ , la distancia mínima es

$$r_{min} = \frac{p}{1 + e}.$$

Con ángulo de  $\theta = 180$ , la distancia máxima es

$$r_{max} = \frac{p}{1 - e}.$$

El semieje mayor, denotado por  $a$ , es la media aritmética entre  $r_{min}$  y  $r_{max}$ , es decir,

$$a = \frac{r_{min} + r_{max}}{2} = \frac{p}{1 - e^2},$$

por lo que el radio  $r$  puede calcularse de la forma

$$r = \frac{a(1 - e^2)}{1 + e \cos \theta}. \quad (3.3)$$

El semieje menor, denotado por  $b$ , es la media geométrica entre  $r_{min}$  y  $r_{max}$ , es decir,

$$b = \sqrt{r_{min} r_{max}} = \frac{P}{\sqrt{1 - e^2}}.$$

Los asteroides siguen órbitas en el sistema solar que son cónicas, es decir, elipses, hipérbolas o parábolas. Supongamos que tenemos los datos de un cierto asteroide donde se recogen 5 de sus posiciones en la órbita que sigue en torno al Sol. Los cinco puntos que representan sus posiciones, en coordenadas cartesianas son:

$$A = (1.84, -0.37)$$

$$B = (1.68, -0.53)$$

$$C = (1.55, -0.61)$$

$$D = (1.34, -0.71)$$

$$E = (0.77, -0.87)$$

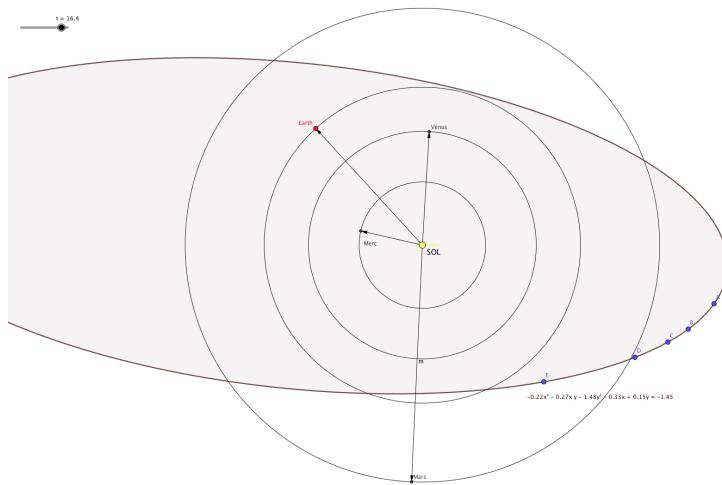
Tenemos que resolver el determinante:

$$\begin{vmatrix} x^2 & xy & y^2 & x & y & 1 \\ (1.84)^2 & (1.84)(-0.37) & (-0.37)^2 & 1.84 & -0.37 & 1 \\ (1.68)^2 & (1.68)(-0.53) & (-0.53)^2 & 1.68 & -0.53 & 1 \\ (1.55)^2 & (1.55)(-0.61) & (-0.61)^2 & 1.55 & -0.61 & 1 \\ (1.34)^2 & (1.34)(-0.71) & (-0.71)^2 & 1.34 & -0.71 & 1 \\ (0.77)^2 & (0.77)(-0.87) & (-0.87)^2 & 0.77 & -0.87 & 1 \end{vmatrix} = 0.$$

La solución es:

$$-0.22x^2 - 0.27xy - 1.48y^2 - 0.33x + 0.15y = -1.45$$

que representa la ecuación de una elipse. Lo podemos representar en la imagen 3.2.



**Figura 3.2** Representación gráfica de la órbita de un asteroide..

Los llamados *Near-Earth Objects* (NEOs) u *Objetos cercanos a la Tierra* son cometas y asteroides que han sido empujados por la atracción gravitacional de planetas cercanos en órbitas que les permiten entrar en el vecindario de la Tierra. Estos objetos son susceptibles de estudio y de un seguimiento profundo por parte de astrónomos de todo el planeta con el fin de vigilar cualquier posible alteración de sus órbitas y su trayectoria.

CNEOS (<https://cneos.jpl.nasa.gov/about/cneos.html>) calcula soluciones de órbitas de alta precisión para NEOs, utilizando conjuntos completos de observaciones astrométricas, métodos sofisticados de ponderación de datos, modelos detallados de las fuerzas perturbadoras y métodos numéricos de alta precisión.

### La órbita del asteroide Eros.

El 13 de agosto de 1898 el astrónomo Carl Gustav Witt descubrió el asteroide Eros desde el observatorio Urania de Copenhague. Eros es un asteroide relativamente pequeño y de forma irregular. Mide alrededor de 34 kilómetros de largo por 11 de ancho. La forma de papa se debe a que su gravedad no es suficientemente intensa como para haberle dado la forma esférica típica de cuerpos celestes más grandes. La menor distancia que ha alcanzado respecto a la Tierra fue de 23 millones de km en 1932 y los astrónomos aprovecharon este acercamiento para determinar el valor de la unidad astronómica (u.a.) que es de 140 millones de km, la distancia media de la Tierra al Sol. En el punto más cercano al sol (perihelio) la órbita de Eros se acerca mucho a la de la Tierra, pero no la cruza.

Su órbita presenta las siguientes características:

- Tiene una excentricidad de 0.223.

- Su órbita está contenida en un plano que forma un ángulo de casi 11 grados con el plano de la eclíptica.
- Eros da una vuelta completa alrededor de su eje en cinco horas y dieciséis minutos y completa una revolución alrededor del Sol en 642 días.

#### **La órbita del cometa C/2002 B2 LINEAR.**

El cometa C/2002 B2 LINEAR constituye un ejemplo de cometa que describe una órbita parabólica, cuya excentricidad es 1.000.

La órbita de C/2002 B2 LINEAR está contenida en un plano que forma un ángulo de casi 153 grados con el que contiene a la órbita de la Tierra (eclíptica).

En la página [www.perihelio.org](http://www.perihelio.org) hay una enorme cantidad de información sobre los cometas, sus características, órbitas, ... en la que podemos reproducir simulaciones orbitales para un buen número de cometas. Además existen enlaces a otros programas sobre órbitas celestes. Para ejecutar una animación del cometa B2 LINEAR podemos entrar en la página [www.perihelio.org/c2002t7orb.htm](http://www.perihelio.org/c2002t7orb.htm).

#### **La órbita del cometa C/2002 E2 Snyder-Murakami.**

El cometa C/2002 E2 Snyder-Murakami presenta la característica de describir una órbita hiperbólica de excentricidad 1.000468. Además, la órbita de C/2002 E2 está contenida en un plano prácticamente perpendicular al que contiene a la órbita de la Tierra (eclíptica). Podemos encontrar los parámetros orbitales de este cometa y una animación completa de su órbita en la dirección [www.astrosurf.com/cometas-obs/C2002E2/orbita.htm](http://www.astrosurf.com/cometas-obs/C2002E2/orbita.htm).

## **3.5 El cifrado matricial de Hill**

En 1994, **crackers** americanos (piratas informáticos de Internet que se introducen en sistemas con el fin de causar algún daño), se hacen vía Internet desde Mallorca con 140.000 números de tarjetas de crédito telefónicas de los Estados Unidos. Usuarios de todo el mundo llaman a cuenta de las víctimas. El fraude llega a los 140 millones de dólares perdidos por compañías de teléfonos americanas.

Ese mismo año desde San Petesburgo, el matemático ruso Vladimir Levin usa la red para sustraer del Citibank en Nueva York 10 millones de dólares en distintas veces. Opera por la noche y envía el dinero a bancos de siete países. Los cómplices recogen el dinero al día siguiente.

En el año 1995, el FBI detiene a Kelvin Mitnick, *el Condor*, uno de los crackers más célebres de EEUU. Había robado más de 20.000 números de tarjetas de crédito.

Como vemos, el tema de la información y la protección de datos y mensajes es de vital importancia en una sociedad como la actual en la que la interacción humano-máquina cada vez es mayor.

La transmisión de grandes cantidades de información por medios electrónicos hace necesario preservar su privacidad: operaciones bancarias, mensajes entre empresas, redes informáticas, etc. Si el mensaje solamente debe ser entendido por un conjunto limitado de receptores estas comunicaciones deben trabajar con sistemas de cifrado (para evitar el acceso a todos los usuarios) y posteriormente de descifrado (para permitirlo a los conocedores de las claves).

La criptología es la disciplina que se ocupa de los problemas teóricos relacionados con la seguridad de la transmisión de mensajes en clave, de un emisor a un receptor, a través de un canal de comunicación. La criptografía es la parte que se ocupa del cifrado de los mensajes en clave y del diseño de los llamados criptosistemas.

Un criptosistema se puede definir como una cuaterna  $(A, K, E, D)$  donde  $A$  es un conjunto llamado alfabeto (frecuentemente  $A$  es el conjunto  $Z_q$ , de los enteros módulo  $q$ , para algún  $q \in N$ );  $K$  es otro conjunto finito llamado *espacio de las claves*;  $E$  es una familia de aplicaciones  $E_k$  del conjunto  $A$  en sí mismo ( $k \in K$ ) llamadas *transformaciones de cifrado* y, por último,  $D$  es la familia de aplicaciones  $D_k$  del conjunto  $A$  en sí mismo ( $k \in K$ ) llamadas transformaciones de descifrado y tales que  $D_k(E_k(x)) = x$ , para cada  $x \in A$  y cada  $k \in K$ .

Existen, básicamente, dos tipos de criptosistemas:

- criptosistemas de clave privada,
- criptosistemas de clave pública.

Los de clave privada utilizan la misma clave para el cifrado y para el descifrado. Nos centraremos en este tipo de criptosistemas. Estudiaremos el cifrado matricial de Hill.

En 1929 Lester Hill desarrolló un método de cifrado que toma  $m$  caracteres consecutivos del texto plano y los sustituye por  $m$  caracteres del texto cifrado. La sustitución se determina mediante  $m$  ecuaciones lineales

$$\begin{aligned} c_1 &= (k_{11}p_1 + k_{12}p_2 + \cdots + k_{1m}p_m) \text{ mód } 26 \\ c_2 &= (k_{21}p_1 + k_{22}p_2 + \cdots + k_{2m}p_m) \text{ mód } 26 \\ &\vdots \\ c_m &= (k_{m1}p_1 + k_{m2}p_2 + \cdots + k_{mm}p_m) \text{ mód } 26 \end{aligned}$$

en las que a cada carácter (del alfabeto inglés) se le asigna un valor numérico de acuerdo con la [tabla 3.2](#).

Las operaciones se realizan módulo 26.

**Tabla 3.2** Codificación de las letras del alfabeto inglés.

LETRA	CÓDIGO	LETRA	CÓDIGO	LETRA	CÓDIGO	LETRA	CÓDIGO
A	0	H	7	O	14	V	21
B	1	I	8	P	15	W	22
C	2	J	9	Q	16	X	23
D	3	K	10	R	17	Y	24
E	4	L	11	S	18	Z	25
F	5	M	12	T	19		
G	6	N	13	U	20		

Podemos describir el método como la ecuación matricial

$$\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1m} \\ k_{21} & k_{22} & \cdots & k_{23} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \cdots & k_{mm} \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix} \text{ mód } 26$$

que abreviamos como

$$\mathbf{c} = K\mathbf{p} \text{ mód } 26$$

donde

$$\mathbf{p} = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_m \end{bmatrix} \quad \text{y} \quad \mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix}$$

son los vectores o matrices columna de  $m$  componentes correspondientes al texto plano y al texto cifrado respectivamente y

$$K = \begin{bmatrix} k_{11} & k_{12} & \cdots & k_{1m} \\ k_{21} & k_{22} & \cdots & k_{23} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \cdots & k_{mm} \end{bmatrix}$$

es una matriz de tamaño  $m \times m$  que representa la clave de cifrado que solamente conocen el emisor y el receptor del mensaje.

**Ejemplo 3.5.1** Supongamos que Alicia quiere cifrar y enviar el siguiente mensaje a Bernardo

HE WILL ARRIVE AT THE AIRPORT AT NOON

Previamente Alicia y Bernardo han elegido la clave

$$K = \begin{bmatrix} 16 & 19 & 5 \\ 4 & 15 & 2 \\ 5 & 5 & 10 \end{bmatrix}$$

para cifrar y descifrar sus mensajes.

A continuación, Alicia divide el mensaje en bloques de longitud 3 (sin considerar los espacios en blanco)

HEW ILL ARR IVE ATT HEA IRP ORT ATN OON

y codifica cada bloque, de acuerdo con la tabla 3.2, obteniendo los vectores

$$\begin{array}{ccccc} (7,4,22) & (8,11,11) & (0,17,17) & (8,21,4) & (0,19,19) \\ (7,4,0) & (8,17,15) & (14,17,19) & (0,19,13) & (14,14,13) \end{array}$$

que escribe matricialmente como

$$P = \begin{bmatrix} 7 & 8 & 0 & 8 & 0 & 7 & 8 & 14 & 0 & 14 \\ 4 & 11 & 17 & 21 & 19 & 4 & 17 & 17 & 19 & 14 \\ 22 & 11 & 17 & 4 & 19 & 0 & 15 & 19 & 13 & 13 \end{bmatrix}.$$

A continuación calcula, en  $\mathbb{Z}_{26}$ , el producto

$$\begin{aligned} C &= KP \\ &= \begin{bmatrix} 16 & 19 & 5 \\ 4 & 15 & 2 \\ 5 & 5 & 10 \end{bmatrix} \begin{bmatrix} 7 & 8 & 0 & 8 & 0 & 7 & 8 & 14 & 0 & 14 \\ 4 & 11 & 17 & 21 & 19 & 4 & 17 & 17 & 19 & 14 \\ 22 & 11 & 17 & 4 & 19 & 0 & 15 & 19 & 13 & 13 \end{bmatrix} \\ &= \begin{bmatrix} 12 & 2 & 18 & 1 & 14 & 6 & 6 & 18 & 10 & 9 \\ 2 & 11 & 3 & 17 & 11 & 10 & 5 & 11 & 25 & 6 \\ 15 & 23 & 21 & 3 & 25 & 3 & 15 & 7 & 17 & 10 \end{bmatrix} \end{aligned}$$

que corresponde, de acuerdo con la tabla 3.2, al texto cifrado

MCP CLX SDV BRD OLZ GKD GFP SLH KZR JGK

Ahora, para evitar que una tercera persona intercepte el mensaje y obtenga información sobre el tamaño de la clave, elimina los espacios. De esta manera, el mensaje que envía a Bernardo es

MCPCLXSDVBRDOLZGKDGFPSLHKZRJGK

El descifrado de los mensajes requiere la utilización de la matriz  $K^{-1}$ . Así, cuando el receptor recibe el mensaje cifrado  $C$  calcula, en  $\mathbb{Z}_{26}$ , la matriz

$$K^{-1}C$$

que coincide con el mensaje original ya que

$$K^{-1}C = K^{-1}(KP) = (K^{-1}K)P = IP = P.$$

**Ejemplo 3.5.2** Continuando con el ejemplo 3.5.1. Cuando Bernardo recibe el texto cifrado

MCPCLXSDVBRDOLZGKDGFPSLHKZRJGK

lo separa en bloques de longitud 3

MCP CLX SDV BRD OLZ GKD GFP SLH KZR JGK

que codifica, de acuerdo con la tabla 3.2, como la matriz

$$C = \begin{bmatrix} 12 & 2 & 18 & 1 & 14 & 6 & 6 & 18 & 10 & 9 \\ 2 & 11 & 3 & 17 & 11 & 10 & 5 & 11 & 25 & 6 \\ 15 & 23 & 21 & 3 & 25 & 3 & 15 & 7 & 17 & 10 \end{bmatrix}.$$

Ahora, calcula el producto

$$\begin{aligned} K^{-1}C &= \begin{bmatrix} 22 & 1 & 7 \\ 12 & 11 & 10 \\ 9 & 7 & 2 \end{bmatrix} \begin{bmatrix} 12 & 2 & 18 & 1 & 14 & 6 & 6 & 18 & 10 & 9 \\ 2 & 11 & 3 & 17 & 11 & 10 & 5 & 11 & 25 & 6 \\ 15 & 23 & 21 & 3 & 25 & 3 & 15 & 7 & 17 & 10 \end{bmatrix} \\ &= \begin{bmatrix} 7 & 8 & 0 & 8 & 0 & 7 & 8 & 14 & 0 & 14 \\ 4 & 11 & 17 & 21 & 19 & 4 & 17 & 17 & 19 & 14 \\ 22 & 11 & 17 & 4 & 19 & 0 & 15 & 19 & 13 & 13 \end{bmatrix} \end{aligned}$$

que corresponde al texto plano original (véase el ejemplo 3.5.1).

Barnaby Jack, un famoso pirata informático de 35 años murió en 2013 en San Francisco. Su especialidad era encontrar fallos en los diminutos ordenadores dentro de equipamiento de médico y de máquinas de bancos. Llegó a comentar en alguna ocasión que era capaz de matar a un hombre a 9 metros de distancia hackeando su dispositivo marcapasos.

Además, es conocido por ser el primer *hacker* que supo romper la seguridad de un cajero automático. Uno de sus grandes éxitos fue encontrar errores en pequeños ordenadores integrados en los equipos tales como dispositivos médicos y cajeros automáticos. De hecho, había planeado para demostrar sus técnicas introducirse en marcapasos y desfibriladores. En 2010, se «conectó» a una máquina dispensadora de lingotes de oro en un casino en Abu Dhabi asegurando que tenía permiso del director de hotel para intentar «hackear» la máquina, pero el sistema de seguridad intervino y resultó que la compañía no disponía, en realidad, de ninguna la máquina similar.

El criptoanalista (o espía que desea conocer el significado del texto cifrado) debe calcular la inversa de la clave  $K$  a partir de la información disponible para poder recuperar el texto plano. Para ello deberá *probar* todas las posibles claves. Por tanto, cuanto mayor sea el tamaño de la matriz  $K$ , que es desconocido para el criptoanalista, mayor número de operaciones tendrá que hacer para determinar  $K^{-1}$ .

Tal como comentamos anteriormente la aritmética en  $\mathbb{Z}_p$ , con  $p$  primo, tiene las mismas propiedades que la aritmética en  $\mathbb{R}$ . Sin embargo, en el cifrado de Hill se utiliza la aritmética de  $\mathbb{Z}_{26}$ . Puesto que 26 no es un número primo, tenemos que  $\mathbb{Z}_{26}$  no es un cuerpo con la adición y la multiplicación módulo 26 y, en consecuencia, algunos resultados de las secciones anteriores no son válidos en  $\mathbb{Z}_{26}$ . Por ejemplo, la matriz

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

no admite forma escalonada reducida, ya que 2 y 4 no son invertibles en  $\mathbb{Z}_{26}$ .

En matemáticas, la aritmética modular es un sistema aritmético para clases de equivalencia de números enteros llamadas *clases de congruencia*. La aritmética modular tiene su origen en un libro de Friedrich Gauss (conocido como el *Príncipe de las Matemáticas*), publicado en 1801 llamado *Disquisitiones Arithmeticae*.

La aritmética modular es conocida popularmente como la *aritmética del reloj*, ya que los números van dando vueltas después de alcanzar un valor que llamamos módulo. El ámbito de aplicación de la aritmética modular es muy amplio y variado, como son la teoría de números, el álgebra abstracta, la criptografía, así como en las artes visuales y musicales.

En música, debido a la equivalencia de octavas y equivalencia armónica (los pasos en razones 1/2 o 2/1 son equivalentes), la aritmética modular se utiliza cuando consideramos la escala de doce tonos igualmente temperada, especialmente en el dodecafismo. En artes visuales esta aritmética puede utilizarse para la creación de patrones artísticos basados en las tablas de multiplicación módulo  $n$ . Sin embargo, la gran importancia de la aritmética modular la solemos asociar con la computación numérica. Las operaciones aritméticas que hacen hoy en día las computadoras son aritméticas modulares, donde el módulo es  $2^b$ , siendo  $b$  el número de bits de los valores sobre los que operamos. Esto se aprecia claramente en la compilación de lenguajes de programación, como el C, por ejemplo. En C, todas las operaciones sobre enteros se toman módulo  $2^{32}$ , en la mayoría de computadoras.

En general, para un  $q$  no necesariamente primo, tenemos que

$$a \in \mathbb{Z}_q \text{ es invertible si y solo si } \text{mcd}(a,q) = 1.$$

Teniendo en cuenta este resultado, para calcular la matriz  $K^{-1}$  del ejemplo 3.5.2 procedemos de la siguiente forma:

$$\begin{array}{c}
 \left[ \begin{array}{ccc|ccc} 16 & 19 & 5 & 1 & 0 & 0 \\ 4 & 15 & 2 & 0 & 1 & 0 \\ 5 & 5 & 10 & 0 & 0 & 1 \end{array} \right] \xrightarrow{F_1 \leftrightarrow F_3} \left[ \begin{array}{ccc|ccc} 5 & 5 & 10 & 0 & 0 & 1 \\ 4 & 15 & 2 & 0 & 1 & 0 \\ 16 & 19 & 5 & 1 & 0 & 0 \end{array} \right] \\
 \xrightarrow{F_1 \leftarrow 21F_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 4 & 15 & 2 & 0 & 1 & 0 \\ 16 & 19 & 5 & 1 & 0 & 0 \end{array} \right] \\
 \xrightarrow{F_2 \leftarrow F_2 + 22F_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 0 & 11 & 20 & 0 & 1 & 20 \\ 16 & 19 & 5 & 1 & 0 & 0 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 + 10F_1} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 0 & 11 & 20 & 0 & 1 & 20 \\ 0 & 3 & 25 & 1 & 0 & 2 \end{array} \right] \\
 \xrightarrow{F_2 \leftarrow 19F_2} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 0 & 1 & 16 & 0 & 19 & 16 \\ 0 & 3 & 25 & 1 & 0 & 2 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow F_3 + 23F_2} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 0 & 1 & 16 & 0 & 19 & 16 \\ 0 & 0 & 3 & 1 & 21 & 6 \end{array} \right] \\
 \xrightarrow{F_3 \leftarrow 9F_3} \left[ \begin{array}{ccc|ccc} 1 & 1 & 2 & 0 & 0 & 21 \\ 0 & 1 & 16 & 0 & 19 & 16 \\ 0 & 0 & 1 & 9 & 7 & 2 \end{array} \right] \\
 \xrightarrow{F_1 \leftarrow F_1 + 25F_2} \left[ \begin{array}{ccc|ccc} 1 & 0 & 12 & 0 & 7 & 5 \\ 0 & 1 & 16 & 0 & 19 & 16 \\ 0 & 0 & 1 & 9 & 7 & 2 \end{array} \right] \\
 \xrightarrow{F_1 \leftarrow F_1 + 14F_3} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 22 & 1 & 7 \\ 0 & 1 & 16 & 0 & 19 & 16 \\ 0 & 0 & 1 & 9 & 7 & 2 \end{array} \right] \\
 \xrightarrow{F_2 \leftarrow F_2 + 10F_3} \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & 22 & 1 & 7 \\ 0 & 1 & 0 & 12 & 11 & 10 \\ 0 & 0 & 1 & 9 & 7 & 2 \end{array} \right]
 \end{array}$$

Por tanto,

$$K^{-1} = \begin{bmatrix} 22 & 1 & 7 \\ 12 & 11 & 10 \\ 9 & 7 & 2 \end{bmatrix}.$$

Finalmente, podemos adaptar el cifrado de Hill para que incluya los caracteres del alfabeto castellano, así como el espacio en blanco y algunos signos de puntuación, de forma que el número total de caracteres del alfabeto a considerar sea un número primo, por ejemplo 37.

### 3.6 Problemas

**3.1** Calcula el determinante de las matrices

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 2 & 3 & 2 \\ 2 & 3 & 4 & 0 \\ 1 & 2 & -1 & -2 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 0 & 3 \\ 4 & 7 & 1 & 1 \\ -1 & 0 & 2 & 1 \\ 2 & 1 & -1 & 0 \end{bmatrix}$$

Calcula  $\det(AB)$ .

**3.2** Calcula el determinante de las matrices siguientes, utilizando el concepto de cofactor.

$$A = \begin{bmatrix} 3 & 0 & 4 \\ 2 & 3 & 2 \\ 0 & 5 & -1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 5 & -2 & 4 \\ 0 & 3 & -5 \\ 2 & -4 & 7 \end{bmatrix}$$

**3.3** Calcula, si existe, la inversa de la matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 4 & 9 \end{bmatrix}$$

**3.4** Calcula  $\det A^5$ , si

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

**3.5** Calcula las inversas de las matrices reales siguientes

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

**3.6** Sea  $A$  una matriz cuadrada de tamaño  $n$ . Supongamos que  $\det(A)$  es no nulo.

1. ¿Qué condición podemos establecer para afirmar que  $A^{-1}$  es una matriz de números enteros?
2. Establece y explica con un ejemplo un procedimiento para construir matrices invertibles cuya inversa sea de números enteros.

**3.7** Sea  $A$  una matriz cuadrada de tamaño 2 de la forma

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

siendo una matriz distinta de la matriz nula. ¿Es invertible? Justifica tu respuesta. En caso afirmativo, halla  $A^{-1}$ .

**3.8** Sea  $A$  una matriz cuadrada de tamaño 2 de la forma

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

siendo una matriz distinta de la matriz nula y sea  $k$  un escalar. Encuentra una expresión que relacione  $\det(kA)$  con  $k\det A$ .

**3.9** Sea  $A$  una matriz cuadrada tal que  $\det A^4 = 0$ . Explica razonadamente por qué  $A$  no es invertible.

**3.10** En el cifrado matricial, para cifrar el texto en claro se utiliza la llamada *matriz de clave*. Si el texto se identifica numéricamente de acuerdo con un alfabeto previamente establecido de  $p$  caracteres, para cifrar hacemos

$$K * P \pmod p,$$

donde  $K$  es la matriz de clave y  $P$  es la matriz formada con el texto a cifrar. Si obtenemos  $C = K * P \pmod p$ , para descifrar el mensaje basta calcular

$$K^{-1} * C \pmod p.$$

- ¿Qué características debe tener  $K$ ? Razona y justifica la respuesta.
- ¿Existe alguna restricción para  $p$ ? Justifica la respuesta.
- Supongamos que vamos a utilizar un alfabeto compuesto por 37 caracteres entre letras y símbolos. Construye una matriz de claves  $K$  de tamaño  $3 \times 3$  de manera que pueda ser utilizada como matriz de claves en el cifrado matricial. ¿Qué procedimiento has utilizado?



# 4 Valores y vectores propios de una matriz

---

La palabra alemana **eigen**, que se traduce en propio, se usó por primera vez en este contexto por David Hilbert en 1904 (aunque Helmholtz la usó previamente con un significado parecido). Eigen se ha traducido también como inherente, característico o el prefijo auto-, donde se aprecia el énfasis en la importancia de los valores propios para definir la naturaleza única de una determinada transformación lineal. Las denominaciones vector y valor característicos también se utilizan habitualmente.

Las transformaciones lineales en el espacio (como por ejemplo la rotación o la reflexión), pueden interpretarse mediante el efecto que ocasionan sobre vectores. Así, un vector propio de una rotación en tres dimensiones es un vector situado en el eje de rotación sobre el cuál se realiza la rotación. El valor propio correspondiente es 1 y el espacio propio es el eje de giro.

Se estudian los valores y vectores propios de una matriz por la gran aplicación que tienen estos conceptos sobre diversos problemas de la ingeniería, las matemáticas, la física y otras ramas de la ciencia. Así, a modo de ejemplo, tenemos que en mecánica, los vectores propios del momento de inercia definen los ejes principales de un cuerpo rígido. En teoría espectral de grafos, se calculan los vectores propios de la matriz de adyacencia del mismo (o de la Laplaciana) para el cálculo de la centralidad de los nodos del mismo. Un ejemplo lo tenemos en el algoritmo de búsqueda de páginas web que emplea Google. En teoría de sistemas dinámicos los valores propios determinan la estabilidad del sistema.

## 4.1 Introducción

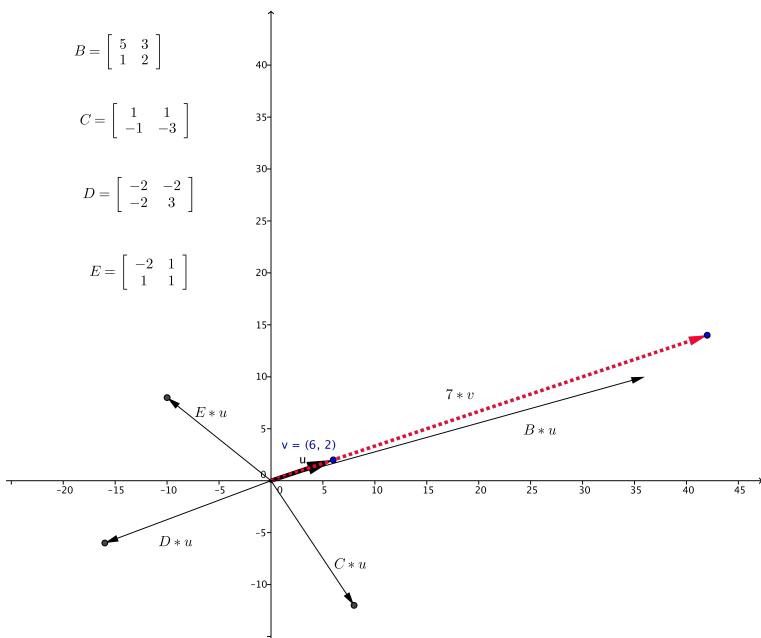
Intentamos en esta introducción que el lector comprenda el significado geométrico asociado al concepto de valor y vector propio de una matriz cuadrada. Para ello, comenzamos

preguntándonos qué ocurre si multiplicamos un vector por una escalar.

Supongamos que tenemos el vector  $\mathbf{u} = (6, 2)$  y el escalar  $\alpha = 7$ . Si hacemos el producto

$$7 \cdot \mathbf{u} = 7 \cdot \begin{bmatrix} 6 \\ 2 \end{bmatrix} = \begin{bmatrix} 42 \\ 14 \end{bmatrix}$$

lo que obtenemos es otro vector de la misma dirección, que constituye una transformación lineal del vector  $\mathbf{u}$ . La siguiente cuestión es, ¿qué ocurre cuando multiplicamos un vector por una matriz?



**Figura 4.1** El producto de una matriz por un vector..

Supongamos que tenemos el vector anterior  $\mathbf{u}$ , que podemos ver dibujado en la [figura 4.1](#), y tomamos algunas matrices cuadradas de tamaño  $2 \times 2$

$$B = \begin{bmatrix} 5 & 3 \\ 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 1 \\ -1 & -3 \end{bmatrix}, \quad D = \begin{bmatrix} -2 & -2 \\ -2 & 3 \end{bmatrix}, \quad E = \begin{bmatrix} -2 & 1 \\ 1 & 1 \end{bmatrix}.$$

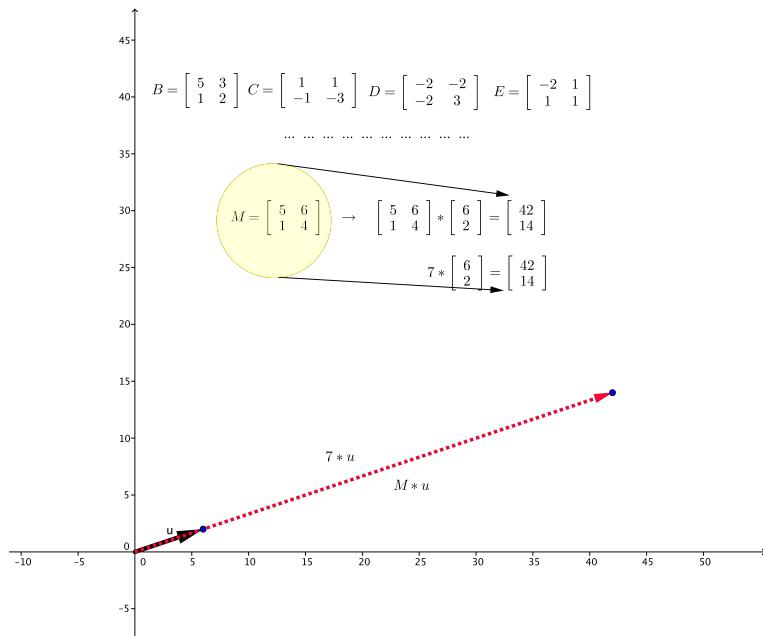
En la [figura 4.1](#) se observa el resultado geométrico de multiplicar cada una de las matrices anteriores por el vector  $\mathbf{u} = (6, 2)$ . Al multiplicar una matriz por un vector obtenemos un nuevo vector que cambia de dirección. En los casos anteriores, el producto matriz-vector no era una transformación lineal del vector puesto que cambiaba la dirección del mismo.

Supongamos ahora que tomamos la matriz

$$M = \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix}.$$

En la [figura 4.2](#) tenemos el resultado geométrico del producto  $M \cdot \mathbf{u}$ . Resulta que

$$M \cdot \mathbf{u} = \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 6 \\ 2 \end{bmatrix} = 7 \cdot \begin{bmatrix} 6 \\ 2 \end{bmatrix}.$$



**Figura 4.2** Representación geométrica del concepto de valor y vector propio..

Este ejemplo nos muestra que hemos encontrado un vector,  $\mathbf{u} = (6,2)$ , un escalar 7 y una matriz cuadrada  $M$  tales que

$$M \cdot \mathbf{u} = 7 \cdot \mathbf{u}.$$

Veremos en la siguiente sección que el escalar 7 es un valor propio de la matriz  $M$  asociado al vector propio  $\mathbf{u}$ .

## 4.2 Valores y vectores propios de una matriz

Comenzamos formalizando la propiedad estudiada en la sección anterior.

**Definición 4.2.1 Valores y vectores propios.** Dada una matriz  $A$  cuadrada de dimensión  $n$ , existen como máximo  $n$  escalares  $\lambda$  y vectores  $\mathbf{x}$  tales que

$$A\mathbf{x} = \lambda\mathbf{x}. \quad (4.1)$$

A los escalares que verifican esta ecuación se les llama **valores propios** de la matriz  $A$ , mientras que los vectores no nulos  $\mathbf{x}$ , que satisfacen la ecuación reciben el nombre de **vectores propios** asociados al valor propio  $\lambda$ .

Notemos que el número de valores propios que puede tener una matriz es siempre menor o igual que su tamaño, mientras que por cada valor propio tiene, al menos, un vector propio asociado al mismo. También se verifica siempre que  $A\mathbf{0} = \lambda\mathbf{0}$ , por lo que el vector cero no se considera un vector propio.

**Definición 4.2.2 Espectro de  $A$  y radiopectral.** Dada una matriz  $A$  cuadrada de dimensión  $n$ , al conjunto de todos los valores propios de  $A$  lo llamamos espectro de  $A$  y lo denotamos por  $\sigma(A)$ .

El radiopectral de  $A$  se define como el módulo del valor propio mayor y se denota por  $\rho(A)$ , es decir,

$$\rho(A) = \max |\lambda|, \text{ tales que } \lambda \in \sigma(A).$$

De la propia definición (5.1) tenemos

$$A\mathbf{x} = \lambda\mathbf{x} \Leftrightarrow (A - \lambda I)\mathbf{x} = 0,$$

lo que significa que la matriz  $A - \lambda I$  es singular. Esto significa que los valores propios  $\lambda$  son los valores que hacen nulo el determinante de la matriz  $A - \lambda I$ . En otras palabras, los valores propios de  $A$  son las raíces de la ecuación

$$\det(A - \lambda I) = 0.$$

**Definición 4.2.3 Polinomio característico de  $A$ .** El polinomio

$$\det(A - \lambda I) = b_0 + b_1\lambda + b_2\lambda^2 + \dots + b_n\lambda^n$$

se llama **polinomio característico** de  $A$  y lo denotamos por  $p_A(\lambda)$ , mientras que la ecuación  $\det(A - \lambda I) = 0$  se conoce como **ecuación característica**.

Volviendo a la idea geométrica desarrollada en la sección anterior, podemos decir que los valores propios  $\lambda_i$ , para  $i \leq n$ , de  $A$  nos indican si el vector propio asociado al valor propio se contrae ( $\lambda_i \leq 1$ ), se expande ( $\lambda_i \geq 1$ ) o, si por el contrario, permanece inalterado ( $\lambda_i = 1$ ).

#### Cálculo de los valores propios de $A$ .

Sabemos que  $\lambda$  es un valor propio de  $A$  si  $A\mathbf{x} = \lambda\mathbf{x}$ , por lo que para calcular los valores propios de una matriz  $A$  basta resolver el determinante

$$\det(A - \lambda I) = 0.$$

Esto significa que, aunque  $A$  tiene  $n$  valores propios puesto que son las raíces del polinomio característico, algunas pueden ser complejas (aunque las entradas de  $A$  sean números reales) y algunos pueden estar repetidos, como ocurre con las raíces de cualquier polinomio. Si  $A$  es de números reales, entonces sus valores propios complejos se dan por pares conjugados.

**Ejemplo 4.2.1** Calcula los valores y vectores propios de la matriz

$$A = \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix}$$

Para calcular los valores propios, resolvemos el determinante

$$\det(A - \lambda I) = 0,$$

es decir,

$$\det \left| \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right| = 0.$$

El polinomio característico es

$$\det \begin{bmatrix} 5 - \lambda & 6 \\ 1 & 4 - \lambda \end{bmatrix} = 0 \rightarrow (5 - \lambda)(4 - \lambda) - 6 = 0,$$

de donde se obtiene que  $\lambda_1 = 7, \lambda_2 = 2$  son los valores propios.

Ahora calculamos los vectores propios asociados a cada valor propio.

- Vector propio asociado a  $\lambda_1 = 7$ .

Ahora se trata de resolver el sistema

$$(A - \lambda_1 I) \mathbf{x}_1 = 0 \rightarrow \begin{bmatrix} -2 & 6 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(1) \\ \mathbf{x}_1(2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuya solución es

$$\mathbf{x}_1 = \alpha \begin{bmatrix} 3 \\ 1 \end{bmatrix}.$$

Esto significa que  $\mathbf{x}_1 = (3, 1)$  y cualquier vector proporcional al mismo son vectores propios asociados al valor propio  $\lambda_1 = 7$ , aunque el vector  $\mathbf{x}_1 = (3, 1)$  es el único independiente.

- Vector propio asociado a  $\lambda_2 = 2$ .

Ahora se trata de resolver el sistema

$$(A - \lambda_2 I) \mathbf{x}_2 = 0 \rightarrow \begin{bmatrix} 3 & 6 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_2(1) \\ \mathbf{x}_2(2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

cuya solución es

$$\mathbf{x}_2 = \alpha \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

Esto significa que  $\mathbf{x}_2 = (1, -0.5)$  y cualquier vector proporcional al mismo son vectores propios asociados al valor propio  $\lambda_2 = 2$ , aunque el vector  $\mathbf{x}_2 = (1, -0.5)$  es el único independiente.

**Ejemplo 4.2.2** Considera la matriz

$$A = \begin{bmatrix} 1 & -1 & 4 \\ 3 & 2 & -1 \\ 2 & 1 & -1 \end{bmatrix}.$$

1. Demuestra que  $\lambda_0 = 1$  es un valor propio de  $A$ .
2. Calcula los vectores propios asociados a  $\lambda_0 = 1$ .

Procedemos a resolver el ejemplo.

1. Comenzamos demostrando que  $\lambda_0 = 1$  es un valor propio de  $A$ .

$$\det(A - \lambda_0 I) = \begin{bmatrix} 1 - \lambda & -1 & 4 \\ 3 & 2 - \lambda & -1 \\ 2 & 1 & -1 - \lambda \end{bmatrix} = 0.$$

Resolviendo este determinante obtenemos el polinomio característico de  $A$ , que es

$$-\lambda^3 + 2\lambda^2 + 5\lambda - 6 = 0.$$

Aplicando la regla de Ruffini podemos determinar las raíces del polinomio característico  $-\lambda^3 + 2\lambda^2 + 5\lambda - 6 = 0$ . Basta sustituir  $\lambda_0 = 1$  en el polinomio para comprobar que  $\lambda_0 = 1$  es un valor propio de  $A$ .

2. Calculamos ahora los vectores propios asociados a  $\lambda_0 = 1$ .

Ahora se trata de resolver el sistema

$$(A - \lambda_0 I)\mathbf{x}_0 = 0 \rightarrow \begin{bmatrix} 0 & -1 & 4 \\ 3 & 1 & -1 \\ 2 & 1 & -2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_0(1) \\ \mathbf{x}_0(2) \\ \mathbf{x}_0(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

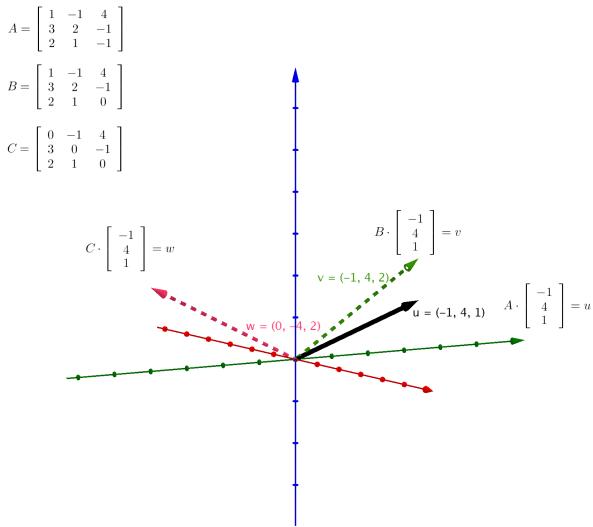
Resolviendo este sistema, las soluciones vienen dadas por los vectores de la forma

$$\left\{ \alpha \begin{bmatrix} -1 \\ 4 \\ 1 \end{bmatrix}, \alpha \in \mathbb{R} \right\},$$

por lo que el vector  $\mathbf{x}_0 = (-1, 4, 1)$  es un vector propio independiente de  $A$  asociado al valor propio  $\lambda_0 = 1$ .

En este último ejemplo, la matriz  $A$  era de dimensión tres, por lo que los vectores estaban en el espacio euclídeo. Siguiendo con el razonamiento geométrico del concepto de valores

y vectores propios que se ha hecho para matrices de dimensión dos y vectores en el plano, la figura 4.3 nos muestra la transformación lineal del vector  $\mathbf{u} = (-1, 4, 1)$ , respecto de la matriz  $A$ . Notemos que en este caso, como  $\lambda = 1$ , el vector queda inalterado al multiplicar por la matriz.



**Figura 4.3** Representación geométrica del valor propio de  $A$ .

¿Cuáles son los valores propios de una matriz  $A$  triangular?

### 4.3 Propiedades de los valores y vectores propios

Una vez estudiada la forma de determinar los valores y vectores propios de una matriz cuadrada, analizamos brevemente algunas propiedades y consecuencias que se obtienen de estos conceptos básicos del álgebra lineal.

**Definición 4.3.1 Trraza de  $A$ .** Dada una matriz cuadrada  $A$  de tamaño  $n$  cuyos elementos de la diagonal denotamos por  $a_{ii}$ , para  $i = 1, 2, \dots, n$ , definimos la traza de  $A$  y la denotamos por  $tr(A)$  a la suma de sus elementos diagonales, es decir

$$tr(A) = \sum_{i=1}^n a_{ii}.$$

Se pueden establecer dos propiedades interesantes relacionadas con el conjunto de valores propios de una matriz cuadrada.

**Teorema 4.3.1** *Sea una matriz cuadrada  $A$  de tamaño  $n$  y  $\sigma(A)$  su conjunto de valores propios. Entonces*

- *La traza de  $A$  es la suma de sus valores propios, es decir*

$$\operatorname{tr}(A) = \sum_{i=1}^n \lambda_i, \quad \lambda_i \in \sigma(A).$$

- *El determinante de  $A$  es igual al producto de los valores propios, es decir*

$$\operatorname{tr}(A) = \prod_{i=1}^n \lambda_i, \quad \lambda_i \in \sigma(A).$$

### 4.3.1 Valores y vectores propios de $A^2$

Vemos, a partir del siguiente ejemplo, cómo son los valores y vectores propios del cuadrado de una matriz.

Siguiendo con el ejemplo se la sección 1, tomamos la matriz  $M$  como

$$M = \begin{bmatrix} 5 & 6 \\ 1 & 4 \end{bmatrix} \rightarrow M^2 = \begin{bmatrix} 31 & 54 \\ 9 & 22 \end{bmatrix}.$$

Los valores propios de  $M$  vimos que eran  $\sigma(M) = \{\lambda_1 = 2, \lambda_2 = 7\}$ , mientras que los vectores propios son

$$\left\{ \mathbf{x}_1 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}.$$

Ahora calculamos los valores propios de  $M^2$ .

$$\det(M^2 - \lambda I) = 0 \rightarrow \det \begin{bmatrix} 31 - \lambda & 54 \\ 9 & 22 - \lambda \end{bmatrix} = 0 \rightarrow (31 - \lambda)(22 - \lambda) - 486 = 0,$$

El polinomio característico de  $M^2$  es  $p_{M^2}(\lambda) = \lambda^2 - 53\lambda + 196$ , cuyas raíces son  $\sigma(M^2) = \{\lambda_1 = 49, \lambda_2 = 4\}$ . Si calculamos los vectores propios asociados a estos valores propios, obtenemos los vectores propios

$$\left\{ \mathbf{x}_1 = \begin{bmatrix} 6 \\ 2 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \right\}.$$

El ejemplo anterior nos lleva a la conclusión de que dada una matriz cuadrada  $A$ , los vectores propios de la matriz  $A^2$  son los mismos que los de la matriz  $A$ , mientras que sus valores propios se elevan al cuadrado.

Ahora se plantea la siguiente cuestión.

La matriz

$$A = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$

tiene valores propios  $\sigma(M) = \{\lambda_1 = 1, \lambda_2 = 0\}$ . ¿Es una matriz idempotente, es decir  $A^2 = A$ ? ¿Qué consecuencias tiene sobre el espectro de la misma? ¿Qué implica para la matriz  $A$  que tenga un valor propio nulo?

### 4.3.2 Descomposición de $A$ mediante valores y vectores propios

Las matrices diagonales tienen una estructura muy simple. La cuestión que surge es si podemos establecer alguna relación entre matrices cuadradas y matrices diagonales. Para determinar esta relación es fundamental el concepto de semejanza.

**Definición 4.3.2 Semejanza de matrices.** Dadas dos matrices  $A$  y  $B$  cuadradas de dimensión  $n$ , se dice que  $A$  y  $B$  son **semejantes** si existe una matriz  $P$  no singular tal que  $P^{-1}AP = B$ . El producto  $P^{-1}AP = B$  se llama **transformación de semejanza** de  $A$ .

El problema fundamental consiste en reducir una matriz cuadrada  $A$  cualquiera a su forma más simple por medio de transformaciones de semejanza. Aquí nos encontramos con el concepto de matriz diagonalizable.

**Definición 4.3.3 Matriz diagonalizable.** Dada una matriz  $A$  cuadradas de dimensión  $n$ , se dice que  $A$  es **diagonalizable** cuando  $A$  es semejante a una matriz diagonal.

Cuando calculamos los vectores propios de una matriz cuadrada de tamaño  $n$ , a partir de sus  $n$  valores propios, puede suceder que obtengamos un conjunto de  $n$  vectores que forman un conjunto de vectores linealmente independiente. En este caso se puede establecer una caracterización de matriz diagonalizable, mediante el teorema siguiente.

**Teorema 4.3.2** *Dada una matriz cuadrada  $A$  de tamaño  $n$  diremos que  $A$  es diagonalizable si y solo si  $A$  tiene un conjunto de  $n$  vectores propios linealmente independientes.*

Así, en el caso en que tenemos una matriz cuadrada  $A$  de tamaño  $n$  con  $n$  vectores propios  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  linealmente independientes, podemos afirmar que existe una matriz  $P$  cuadrada de tamaño  $n$  tal que

$$P^{-1}AP = D,$$

donde  $P$  es una matriz tal que sus columnas son los  $n$  vectores propios de  $A$  y  $D$  es una matriz diagonal de tamaño  $n$  tal que

$$D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n),$$

es decir, sus elementos diagonales son los valores propios de  $A$ . Es importante tener en cuenta que debemos ordenar los valores y vectores propios en las matrices  $P$  y  $D$ , de manera que  $(\lambda_j, P_{*j})$  es el par vector propio  $\mathbf{x}_j$  asociado al valor propio  $\lambda_j$ .

**Ejemplo 4.3.1** Diagonaliza, si es posible, la matriz

$$A = \begin{bmatrix} 1 & -4 & -4 \\ 8 & -11 & -8 \\ -8 & 8 & 5 \end{bmatrix}.$$

Para ver si es diagonalizable, debemos calcular el conjunto de valores y vectores propios de  $A$  para comprobar si  $A$  tiene 3 vectores propios linealmente independientes. En dicho caso, podremos afirmar que existe la relación de semejanza con una matriz diagonal formada por los valores propios.

Calculamos los valores propios mediante el procedimiento habitual, resolviendo el polinomio característico de  $A$ ,

$$\det(A - \lambda I) = \begin{bmatrix} 1 - \lambda & -4 & -4 \\ 8 & -11 - \lambda & -8 \\ -8 & 8 & 5 - \lambda \end{bmatrix} = 0 \rightarrow \lambda^3 + 5\lambda^2 + 3\lambda - 9 = 0.$$

Factorizamos el polinomio característico para determinar sus raíces

$$\lambda^3 + 5\lambda^2 + 3\lambda - 9 = (\lambda - 1)(\lambda + 3)^2 = 0.$$

En consecuencia, los valores propios son  $\{\lambda_1 = 1, \lambda_2 = -3\}$ . Notemos que el valor propio  $\lambda_1 = 1$  es simple y tiene multiplicidad algebraica 1, el valor propio  $\lambda_1 = -3$  es doble y tiene multiplicidad algebraica 2.

**Vectores propios asociados a  $\lambda = 1$ .** Ahora se trata de resolver el sistema

$$(A - I)\mathbf{x}_1 = 0 \rightarrow \begin{bmatrix} 0 & -4 & -4 \\ 8 & -12 & -8 \\ -8 & 8 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(1) \\ \mathbf{x}_1(2) \\ \mathbf{x}_1(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Resolviendo este sistema, las soluciones vienen dadas por los vectores de la forma

$$\left\{ \alpha \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \alpha \in \mathfrak{R} \right\},$$

por lo que el vector independiente

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$$

es el vector propio asociado al valor propio  $\lambda = 1$ .

**Vectores propios asociados a  $\lambda = -3$ .** Ahora se trata de resolver el sistema

$$(A - 3I)\mathbf{x}_2 = 0 \rightarrow \begin{bmatrix} -2 & -4 & -4 \\ 8 & -14 & -8 \\ -8 & 8 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1(1) \\ \mathbf{x}_1(2) \\ \mathbf{x}_1(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Resolviendo este sistema, las soluciones vienen dadas por los vectores de la forma

$$\left\{ \alpha \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \alpha, \beta \in \mathbb{R} \right\},$$

por lo que los vectores

$$\mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

son los vectores asociados al valor propio  $\lambda = -3$ .

Es fácil comprobar que el conjunto de vectores

$$\left\{ \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

constituyen un conjunto linealmente independiente de vectores. En consecuencia,  $A$  es diagonalizable y existe una matriz  $P$  dada por los vectores propios de  $A$ ,

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

y una matriz  $D$  diagonal

$$D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

tales que  $P^{-1}AP = D$ .

**Ejercicio.** Comprueba si es diagonalizable la matriz

$$A = \begin{bmatrix} -1 & -1 & -2 \\ 8 & -11 & -8 \\ -10 & 11 & 7 \end{bmatrix}.$$

El siguiente resultado resume algunas de las propiedades que comparten las matrices semejantes entre sí.

**Teorema 4.3.3** Dos matrices  $A$  y  $B$  semejantes verifican que

1. Tienen el mismo polinomio característico.
2. Tienen los mismos valores propios con las mismas multiplicidades.
3. La traza de ambas matrices coincide.
4. Tienen el mismo determinante.

Este teorema tiene una gran importancia puesto que establece que dos matrices semejantes tienen los mismos valores propios.

**Ejemplo 4.3.2** Prueba que la matriz

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

no es diagonalizable.

Como  $A$  es triangular, sus valores propios son los elementos de la diagonal. Así, el único valor propio es  $\lambda_1 = 1$  con multiplicidad 2. La matriz  $|I - A|$  tiene la forma escalonada reducida

$$I - A = \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

esto significa que el conjunto de soluciones es

$$\left\{ \alpha \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \alpha \in \mathbb{R} \right\}.$$

Por lo tanto,  $A$  no tiene dos vectores propios linealmente independientes, luego no es diagonalizable.

**Ejercicio.** Demuestra que si una matriz  $A$  es diagonalizable,  $A^T$  es diagonalizable.

## 4.4 Diagonalización de matrices específicas

Vamos a estudiar la diagonalización de algunas matrices especiales como son las matrices simétricas (la matriz coincide con su traspuesta) y las matrices definidas positivas.

### 4.4.1 Diagonalización de matrices simétricas

Las matrices simétricas tienen una gran importancia en la teoría de matrices puesto que muchas aplicaciones reales se pueden modelizar mediante matrices simétricas. Por ejemplo,

las cónicas o las aproximaciones por mínimos cuadrados. Tampoco debemos olvidar que en teoría de grafos, la matriz de adyacencia del grafo, que define las conexiones entre sus nodos, es una matriz simétrica.

La gran ventaja que tienen las matrices simétricas es que son diagonalizables y todos sus valores propios son reales.

Sabemos que una matriz  $A$  cuadrada y simétrica de tamaño 2 tiene dos valores propios  $\lambda_1, \lambda_2$  cumpliendo

$$\begin{aligned} A\mathbf{x}_1 &= \lambda_1 \mathbf{x}_1 \\ A\mathbf{x}_2 &= \lambda_2 \mathbf{x}_2 \end{aligned}$$

Como  $A$  es simétrica,

$$(A\mathbf{x}_1)^T = (\lambda_1 \mathbf{x}_1)^T \rightarrow (\mathbf{x}_1)^T A = \lambda_1 (\mathbf{x}_1)^T$$

por lo que

$$(\mathbf{x}_1)^T A \mathbf{x}_2 = \lambda_1 (\mathbf{x}_1)^T \mathbf{x}_2.$$

Como  $A\mathbf{x}_2 = \lambda_2 \mathbf{x}_2$ , entonces

$$\lambda_1 (\mathbf{x}_1)^T A \mathbf{x}_2 = \lambda_2 (\mathbf{x}_1)^T \mathbf{x}_2,$$

por lo que

$$(\lambda_1 - \lambda_2) (\mathbf{x}_1)^T A \mathbf{x}_2 = 0. \quad (4.2)$$

Si en la ecuación (4.2) se verifica que  $\lambda_1 \neq \lambda_2$ , entonces

$$(\mathbf{x}_1)^T A \mathbf{x}_2 = 0,$$

lo que significa que los vectores  $\mathbf{x}_1$  y  $\mathbf{x}_2$  son ortogonales, es decir, los vectores propios son ortogonales entre sí.

Lo escribimos en forma matricial.

Si llamamos

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad X = [\mathbf{x}_1 \quad \mathbf{x}_2]$$

entonces  $A$  verifica que

$$AX = X\Lambda.$$

Si suponemos ahora que los vectores propios están normalizados, podemos afirmar que

$$(\mathbf{x}_1)^T A \mathbf{x}_1 = 1, \quad (\mathbf{x}_2)^T A \mathbf{x}_2 = 1, \quad (4.3)$$

mientras que por la ortogonalidad se cumple que

$$(\mathbf{x}_1)^T A \mathbf{x}_2 = 0, \quad (\mathbf{x}_2)^T A \mathbf{x}_1 = 0. \quad (4.4)$$

Las ecuaciones (4.3) y (4.4) pueden escribirse como

$$X^T X = I,$$

o, lo que es lo mismo

$$X^{-1} = X^T,$$

lo que significa que  $X$  es una matriz ortogonal.

**Teorema 4.4.1** Una matriz  $A$  cuadrada simétrica es diagonalizable por una matriz ortogonal formada por los vectores propios de  $A$ .

Si la matriz  $A$  es simétrica de tamaño 2, como la matriz  $X$  es ortogonal, significa que  $X$  es una rotación, una reflexión o una combinación de ambas. Se mantienen ángulos y longitudes.

#### 4.4.2 Diagonalización de una matriz definida positiva

En álgebra lineal, una matriz definida positiva es una matriz Hermitiana (matriz de números complejos tal que es igual a su traspuesta conjugada) que es análoga a los números reales positivos. Comenzamos con la definición de matriz definida positiva.

**Definición 4.4.1 Matriz definida positiva.** Dada una matriz  $A$  cuadrada real de dimensión  $n$ , se dice que  $A$  es **definida positiva** si

$$\mathbf{x}^T A \mathbf{x} > 0, \quad (4.5)$$

para cualquier vector no nulo  $\mathbf{x}$  en  $\mathbb{R}^2$ .

Para hacernos una idea del significado geométrico del concepto, supongamos que  $\mathbf{x}$  es unitario. La condición (4.5) significa, desde el punto de vista geométrico, que el ángulo entre  $\mathbf{x}$  y  $A\mathbf{x}$  está entre  $-90$  grados y  $+90$  grados.

El concepto de matriz definida positiva es difícil de utilizar para demostrar que una matriz concreta lo es. Por ello, establecemos una caracterización que nos permita establecer condiciones más sencillas en cuestiones prácticas.

Recordemos que los menores principales de una matriz cuadrada a los determinantes de las sucesivas submatrices cuadradas superior izquierda de orden  $1, 2, \dots, n$  de forma que el primer menor principal es la primera componente  $a_{11}$  y el último menor principal es el propio determinante de la matriz de toda la matriz.

**Teorema 4.4.2** Dada una matriz  $A$  cuadrada real de dimensión  $n$ , se dice que  $A$  es **definida positiva** si y solo si todos sus menores principales son positivos.

Podemos establecer una caracterización, a partir de este teorema, para determinar cuándo una matriz cuadrada de tamaño 2 es definida positiva.

**Teorema 4.4.3** Una matriz  $A$  cuadrada de tamaño 2 de la forma

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

es definida positiva si y solo si  $a > 0$  y  $ac - b^2 > 0$ .

Una de las propiedades más interesantes de las matrices definidas positivas viene reflejada en el teorema siguiente.

**Teorema 4.4.4** Sea una matriz  $A$  cuadrada Hermitiana de tamaño  $n$ . Entonces  $A$  se dice definida positiva si cumple que todos los valores propios de  $A$  son positivos, es decir,  $\lambda, \lambda \in \sigma(A)$  verifican que  $\lambda > 0$ .

## 4.5 Valores y vectores propios de matrices $2 \times 2$ - representación gráfica

Dada una matriz cuadrada de la forma

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

podemos calcular sus valores propios mediante su ecuación característica, como

$$A\mathbf{x} = \lambda \mathbf{x} \rightarrow \det \left( \begin{bmatrix} a-\lambda & b \\ c & d-\lambda \end{bmatrix} \right) = 0$$

es decir,

$$(a-\lambda)(d-\lambda) - bc = 0 \rightarrow \lambda^2 + k_1\lambda + k_2 = 0.$$

Consideremos el ejemplo de la matriz

$$\begin{bmatrix} 3 & 6 \\ 6 & 3 \end{bmatrix}.$$

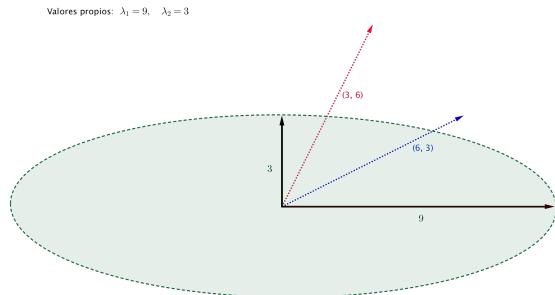
Si calculamos el polinomio característico de esta matriz obtenemos

$$\lambda^2 - 6\lambda - 27 = 0,$$

lo que significa que sus valores propios son  $\lambda_1 = 9, \lambda_2 = -3$ .

Ahora dibujamos los vectores de  $A$  por filas, mientras que los valores propios son solución de una ecuación de segundo grado, luego representan la magnitud o longitud de

los ejes mayor y menor de una elipse. Lo vemos en la figura 4.4.



**Figura 4.4** Representación geométrica del concepto de valor y vector propio de matrices  $2 \times 2$ .

Supongamos ahora que giramos la elipse de forma que pase por los puntos  $(3,6)$  y  $(6,3)$ .

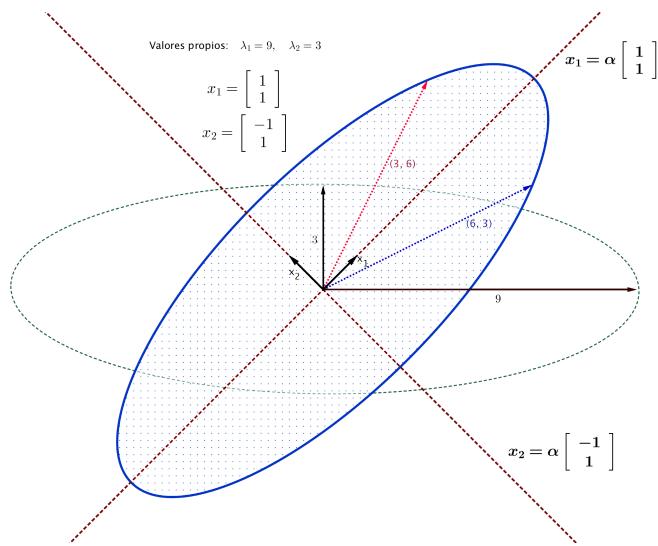
Si calculamos los vectores propios asociados a los valores propios, obtenemos

$$\lambda_1 = 9 \rightarrow \mathbf{x}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \lambda_2 = -3 \rightarrow \mathbf{x}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Entonces, una vez calculados los valores y vectores propios, sabemos que:

- El vector propio asociado al valor propio más grande representa la dirección del vector del eje mayor de la elipse.
- El vector propio asociado al segundo valor propio más grande representa la dirección del vector del eje menor de la elipse.

Vemos la interpretación geométrica de los vectores y valores propios en la figura 4.5.



**Figura 4.5** Representación geométrica del concepto de valor y vector propio de matrices  $2 \times 2$ .

Hemos visto, por medio de este ejemplo, la idea geométrica que hay detrás del concepto de vector y valor propio en matrices de tamaño 2.

### Ejercicios.

1. Calcula los valores y vectores propios de la matriz

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

y de la matriz  $A^2$ .

2. Calcula los valores y vectores propios de las matrices

$$A = \begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} -2 & -7 \\ 1 & 2 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

3. Consideremos una matriz  $A$  de tamaño  $2 \times 2$ . Comprueba que la ecuación característica de  $A$  puede expresarse de la forma

$$\lambda^2 - \text{tr}(A)\lambda + \det(A) = 0.$$

- a) Encuentra una expresión algebraica que nos proporcione las soluciones de la ecuación característica, en función de los elementos de la matriz.

**b)** ¿Cuándo se podrá afirmar que la matriz  $A$  tendrá dos valores propios reales distintos?

**c)** ¿Cuándo se podrá afirmar que una matriz  $A$  tendrá un único valor propio real doble?

**4.** El teorema de Cayley-Hamilton establece que toda matriz cuadrada  $A$  verifica su ecuación caracetrística

$$p_A(\lambda) = \det(A - \lambda I) = 0.$$

Verifica el teorema de Cayley-Hamilton para la matriz

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & -5 & 4 \end{bmatrix}.$$

**5.** Calcula los vectores y valores propios de la matriz

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

**6.** Probad que la matriz

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

no es diagonalizable.

**7.** Probad que la matriz

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

es diagonalizable y calculad una matriz  $P$  tal que  $P^{-1}AP$  sea diagonal.

Resuelve las cuestiones que se plantean a continuación.

### Cuestiones.

**1.** ¿Qué debe verificar una matriz triangular para que sea singular?

**2.** Se dice que una matriz cuadrada  $A$  es nilpotente si existe un entero positivo  $k$  tal que  $A^k = 0$ . ¿Cuáles son los posibles valores propios de una matriz nilpotente?

3. Si los valores propios de

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

son  $\lambda_1 = 0$  y  $\lambda_2 = 1$ . ¿Cuáles son los posibles valores de  $a$  y  $d$ ?

4. ¿Cómo deben ser  $a$  y  $d$  para que la matriz

$$A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix}$$

sea definida positiva?

## 4.6 La factorización y algoritmo QR

La factorización *QR* de una matriz  $A$ , cuyas columnas son linealmente independientes, es una descomposición de la misma como producto de una matriz ortogonal  $Q$  por una triangular superior  $R$ . La descomposición *QR* es la base del algoritmo *QR* utilizado para el cálculo de los vectores y valores propios de una matriz.

Se va a calcular la factorización *QR* mediante el método de **ortogonalización de Gram-Schmidt**. El método de ortogonalización de Gram-Schmidt construye una base ortonormal de vectores en un espacio vectorial  $V$  de dimensión  $n$  a partir de una base cualquiera de vectores. Pierre-Simon Laplace ya estudió un algoritmo de este tipo en el siglo XVIII, aunque el algoritmo que se utiliza en la actualidad se debe a los matemáticos del siglo XIX Jorgen Gram y Erhard Schmidt. Dicho algoritmo fue uno de los primeros en el campo del álgebra lineal computacional.

Inicialmente se escribe la matriz  $A$  por columnas,  $A = (\mathbf{a}_1 | \mathbf{a}_2 | \cdots | \mathbf{a}_n)$ . Entonces, si aplicamos a los vectores columna de  $A$  el algoritmo de Gram-Schmidt tendremos que

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\ \mathbf{u}_2 &= \mathbf{a}_2 - \text{proj}_{\mathbf{e}_1} \mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\ \mathbf{u}_3 &= \mathbf{a}_3 - \text{proj}_{\mathbf{e}_1} \mathbf{a}_3 - \text{proj}_{\mathbf{e}_2} \mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \\ &\vdots \\ \mathbf{u}_k &= \mathbf{a}_k - \sum_{j=1}^{k-1} \text{proj}_{\mathbf{e}_j} \mathbf{a}_k, & \mathbf{e}_k &= \frac{\mathbf{u}_k}{\|\mathbf{u}_k\|} \end{aligned}$$

Despejando  $\mathbf{a}_j$  se tiene

$$\begin{aligned} \mathbf{a}_1 &= \mathbf{e}_1 \|\mathbf{u}_1\| \\ \mathbf{a}_2 &= \text{proj}_{\mathbf{e}_1} \mathbf{a}_2 + \mathbf{e}_2 \|\mathbf{u}_2\| \\ \mathbf{a}_3 &= \text{proj}_{\mathbf{e}_1} \mathbf{a}_3 - \text{proj}_{\mathbf{e}_2} \mathbf{a}_3 + \mathbf{e}_3 \|\mathbf{u}_3\| \\ &\vdots \\ \mathbf{a}_k &= \sum_{j=1}^{k-1} \text{proj}_{\mathbf{e}_j} \mathbf{a}_k + \mathbf{e}_k \|\mathbf{u}_k\| \end{aligned}$$

Se puede escribir matricialmente estas ecuaciones lineales.

$$(\mathbf{e}_1 | \mathbf{e}_2 | \cdots | \mathbf{e}_n) \begin{bmatrix} \|\mathbf{u}_1\| & \mathbf{e}_1 \cdot \mathbf{a}_2 & \mathbf{e}_1 \cdot \mathbf{a}_3 & \cdots \\ 0 & \|\mathbf{u}_2\| & \mathbf{e}_2 \cdot \mathbf{a}_3 & \cdots \\ 0 & 0 & \|\mathbf{u}_3\| & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Este producto de matrices es igual a la matriz  $A$  con la que se inicia la factorización. De esta forma, se puede escribir que

$$A = QR,$$

donde  $Q = (\mathbf{e}_1 | \mathbf{e}_2 | \cdots | \mathbf{e}_n)$  es una matriz ortogonal, ya que está formada por los vectores  $\mathbf{e}_j$ , para  $j = 1, \dots, n$  obtenidos mediante el algoritmo de Gram-Schmidt. La matriz  $R$

$$R = \begin{bmatrix} \|\mathbf{u}_1\| & \mathbf{e}_1 \cdot \mathbf{a}_2 & \mathbf{e}_1 \cdot \mathbf{a}_3 & \cdots \\ 0 & \|\mathbf{u}_2\| & \mathbf{e}_2 \cdot \mathbf{a}_3 & \cdots \\ 0 & 0 & \|\mathbf{u}_3\| & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

es triangular superior.

Nótese que se tiene otra forma de calcular la matriz  $R$  como

$$R = Q^T A.$$

Esto es debido a que como  $Q$  es ortogonal, entonces su inversa coincide con su traspuesta. Veamos un ejemplo que clarifique los pasos que seguimos.

#### Ejemplo 4.6.1 Determina una factorización $QR$ para la matriz

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

El primer paso sería aplicar el algoritmo de Gram-Schmidt a los vectores que forman las columnas de  $A$ , utilizando las expresiones

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\ \mathbf{u}_2 &= \mathbf{a}_2 - proj_{\mathbf{e}_1} \mathbf{a}_2, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\ \mathbf{u}_3 &= \mathbf{a}_3 - proj_{\mathbf{e}_1} \mathbf{a}_3 - proj_{\mathbf{e}_2} \mathbf{a}_3, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \end{aligned}$$

Desarrollando estas ecuaciones para  $A$  tenemos los vectores:

$$\mathbf{e}_1 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \mathbf{e}_3 = \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

Los vectores  $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$  constituyen los vectores de las columnas de la matriz ortogonal  $Q$ . Así,

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{bmatrix}.$$

Ahora solo falta calcular la matriz  $R$ . Sabemos que  $R = Q^T A$ , por lo que haciendo los cálculos correspondientes tendremos que

$$\begin{aligned} R &= Q^T A \\ &= \left[ \begin{array}{ccc} \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} \end{array} \right]^T \cdot \left[ \begin{array}{ccc} 1 & -2 & 1 \\ -1 & 3 & 2 \\ 1 & -1 & -4 \end{array} \right] \\ &= \left[ \begin{array}{ccc} \sqrt{3} & -2\sqrt{3} & -\frac{5}{3}\sqrt{3} \\ 0 & \sqrt{2} & -2\sqrt{3} \\ 0 & 0 & \frac{1}{3}\sqrt{96} \end{array} \right]. \end{aligned}$$

A continuación veamos otro ejemplo para el cálculo de la factorización  $QR$  en el que únicamente tendremos en cuenta el proceso de orto normalización de Gram-Schmidt.

**Ejemplo 4.6.2** Supongamos que queremos obtener la factorización  $QR$  de la matriz

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

El primer paso, al igual que en el caso anterior, sería aplicar el algoritmo de Gram-Schmidt a los vectores que forman las columnas de  $A$ , utilizando las expresiones habituales. Sin embargo, teniendo en cuenta la expresión de las proyecciones de un vector sobre otro, estas ecuaciones las podemos escribir de otra forma, es decir,

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1, & \mathbf{e}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\ \mathbf{u}_2 &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1, & \mathbf{e}_2 &= \frac{\mathbf{u}_2}{\|\mathbf{u}_2\|} \\ \mathbf{u}_3 &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2} \mathbf{u}_2, & \mathbf{e}_3 &= \frac{\mathbf{u}_3}{\|\mathbf{u}_3\|} \end{aligned}$$

Hemos especificado un poco más detalladamente el cálculo que se realiza. Por cuestiones de comodidad para la parte final de la factorización, podemos renombrar los escalares que aparecen en las expresiones anteriores, de manera que

$$\alpha_{21} = \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1, \quad \alpha_{31} = \frac{\mathbf{a}_3 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1, \quad \alpha_{32} = \frac{\mathbf{a}_3 \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2} \mathbf{u}_2.$$

Procedemos con los cálculos escribiendo los vectores columna de  $A$ , por comodidad en la notación, como

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = (1,1,0), \quad \mathbf{a}_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = (1,,0,1), \quad \mathbf{a}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = (0,1,1)$$

En consecuencia,

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{a}_1(1,1,0), \quad \|\mathbf{u}_1\| = \sqrt{2}. \\ \mathbf{u}_2 &= \mathbf{a}_2 - \frac{\mathbf{a}_2 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 = (1,0,1) - \frac{1}{2}(1,1,0) = (1/2, -1/2, 1). \|\mathbf{u}_2\| = \sqrt{\frac{3}{2}} \\ \mathbf{u}_3 &= \mathbf{a}_3 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_1}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 - \frac{\mathbf{a}_3 \cdot \mathbf{u}_2}{\|\mathbf{u}_2\|^2} \mathbf{u}_2 \\ &= (0,1,1) - \frac{1}{2}(1,1,0) - \frac{1/2}{3/2}(1/2, -1/2, 1) = \left(-\frac{2}{3}, \frac{2}{3}, \frac{2}{3}\right), \|\mathbf{u}_3\| = \frac{2}{\sqrt{3}}. \end{aligned}$$

A partir de Gram-Schmidt podemos escribir

$$\begin{aligned} A &= (\mathbf{a}_1 \quad \mathbf{a}_2 \quad \mathbf{a}_3) \\ &= (\mathbf{u}_1 \quad \alpha_{21}\mathbf{u}_1 + \mathbf{u}_2 \quad \alpha_{31}\mathbf{u}_1 + \alpha_{32}\mathbf{u}_2 + \mathbf{u}_3) \\ &= (\mathbf{u}_1 \quad \mathbf{u}_2 \quad \mathbf{u}_3) \begin{bmatrix} 1 & \alpha_{21} & \alpha_{31} \\ 0 & 1 & \alpha_{32} \\ 0 & 0 & 1 \end{bmatrix} \\ &= (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3) \begin{bmatrix} \|\mathbf{u}_1\| & \alpha_{21}\|\mathbf{u}_1\| & \alpha_{31}\|\mathbf{u}_1\| \\ 0 & \|\mathbf{u}_2\| & \alpha_{32}\|\mathbf{u}_2\| \\ 0 & 0 & \|\mathbf{u}_3\| \end{bmatrix}. \end{aligned}$$

En consecuencia, denotando por

$$Q = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3), \quad R = \begin{bmatrix} \|\mathbf{u}_1\| & \alpha_{21}\|\mathbf{u}_1\| & \alpha_{31}\|\mathbf{u}_1\| \\ 0 & \|\mathbf{u}_2\| & \alpha_{32}\|\mathbf{u}_2\| \\ 0 & 0 & \|\mathbf{u}_3\| \end{bmatrix}$$

se tiene que  $A = QR$ , con  $Q$  ortonormal y  $R$  triangular superior.

Tendremos que

$$Q = (\mathbf{e}_1 \quad \mathbf{e}_2 \quad \mathbf{e}_3) = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Análogamente,

$$R = \begin{bmatrix} \|\mathbf{u}_1\| & \alpha_{21}\|\mathbf{u}_1\| & \alpha_{31}\|\mathbf{u}_1\| \\ 0 & \|\mathbf{u}_2\| & \alpha_{32}\|\mathbf{u}_2\| \\ 0 & 0 & \|\mathbf{u}_3\| \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{2}{\sqrt{3}} \end{bmatrix}.$$

También podemos comprobar que  $R = Q^T A$ .

$$R = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} \\ 0 & \frac{\sqrt{2}}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}^T \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{\sqrt{2}}{\sqrt{3}} \end{bmatrix}.$$

### Algoritmo QR.

A partir de la factorización  $QR$  descrita, podemos diseñar un algoritmo que nos permita determinar los valores y vectores propios de  $A$ .

Los pasos son los siguientes:

- Se hace  $A_1 = A$ .
- Para  $i = 1$ , se calcula la descomposición QR de  $A_1$ , es decir,  $A_1 = Q_1 R_1$  y hacemos  $A_2 = R_1 Q_1$ .
- Para  $i = 2$ , se calcula la descomposición QR de  $A_2$ , es decir,  $A_2 = Q_2 R_2$  y hacemos  $A_3 = R_2 Q_2$ .
- Para  $i = 3$ , se calcula la descomposición QR de  $A_3$ , es decir,  $A_3 = Q_3 R_3$  y hacemos  $A_4 = R_3 Q_3$ .
- Para  $i = 4, 5, \dots$  se sigue repitiendo los mismos cálculos que en los pasos anteriores, hasta que se produzca la convergencia.

Mediante estos pasos, se obtienen dos secuencias de matrices

$$A_n = (A_1, A_2, A_3, \dots, A_n, A_{n+1}, \dots)$$

$$Q_n = (Q_1, Q_2, Q_3, \dots, Q_n, Q_{n+1}, \dots).$$

La secuencia  $A_n$  converge a una matriz diagonal  $D$  donde aparecen los valores propios de  $A$  en la diagonal. En cuanto a los vectores propios, éstos son las columnas de la matriz

$$S = \prod_i Q_i.$$

**Ejemplo 4.6.3** Queremos aplicar el algoritmo  $QR$  a la matriz

$$A = \begin{bmatrix} 8 & 7 \\ 1 & 2 \end{bmatrix}.$$

Tomamos  $A_1 = A$  y calculamos una factorización  $QR$  de  $A$ . Realizando los cálculos oportunos obtenemos

$$\begin{aligned} A_1 &= Q_1 R_1 \\ &= \begin{bmatrix} 0.9922 & -0.1240 \\ 0.1240 & 0.9922 \end{bmatrix} \begin{bmatrix} 8.0622 & 7.1940 \\ 0.0000 & 1.1163 \end{bmatrix} \\ &= \begin{bmatrix} 8.8923 & 6.1384 \\ 0.1384 & 1.1076 \end{bmatrix} \end{aligned}$$

Ahora hacemos  $A_2 = R_1 Q_1$  y volvemos a calcular una factorización  $QR$  de  $A_2$ .

$$\begin{aligned} A_2 &= Q_2 R_2 \\ &= \begin{bmatrix} 0.9998 & -0.1556 \\ 0.0155 & 0.9998 \end{bmatrix} \begin{bmatrix} 8.8933 & 6.1549 \\ 0.0000 & 1.0119 \end{bmatrix} \\ &= \begin{bmatrix} 8.9881 & 6.0157 \\ 0.0157 & 1.0118 \end{bmatrix} \end{aligned}$$

Ahora hacemos  $A_3 = R_2 Q_2$  y volvemos a calcular una factorización  $QR$  de  $A_3$ .

$$\begin{aligned} A_3 &= Q_3 R_3 \\ &= \begin{bmatrix} 0.9999 & -0.0017 \\ 0.0017 & 0.9999 \end{bmatrix} \begin{bmatrix} 8.9881 & 6.0175 \\ 0.0000 & 1.0013 \end{bmatrix} \\ &= \begin{bmatrix} 8.9986 & 6.0107 \\ 0.0017 & 1.0013 \end{bmatrix} \end{aligned}$$

Tenemos

$$A_1 = \begin{bmatrix} 8.8923 & 6.1384 \\ 0.1384 & 1.1076 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 8.9881 & 6.0157 \\ 0.0157 & 1.0118 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 8.9986 & 6.0107 \\ 0.0017 & 1.0013 \end{bmatrix}, \dots$$

La secuencia  $A_n$  converge a una matriz diagonal  $D$  donde aparecen los valores propios de  $A$  en la diagonal. En cuanto a los vectores propios, éstos son las columnas de la matriz

$$S = \prod_i Q_i.$$

Parece claro a la vista de esta sucesión de matrices que los valores propios de  $A$  son aproximadamente 9 y 1.

## 4.7 El método de Lanczos para el cálculo de valores y vectores propios

El algoritmo de Lanczos para matrices  $A$  simétricas calcula, mediante un proceso iterativo, los valores y vectores propios de matrices grandes y dispersas. Constituye el algoritmo más

eficiente, desde el punto de vista computacional, cuando se trabaja con matrices grandes, simétricas y con pocos elementos distintos de cero.

El  $n$ -ésimo paso del algoritmo transforma la matriz  $A$  inicial, de dimensión  $n$  en una matriz  $T$  tridiagonal tal que esta matriz es semejante a  $A$ . Por definición de matrices semejantes, se puede afirmar que existe una matriz  $V$  ortonormal tal que se verifica que

$$T = V^T A V, \quad \longrightarrow \quad VT = AV.$$

Sea dicha matriz  $T$  de la forma

$$T = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \beta_3 & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \beta_{n-1} & \alpha_n \end{bmatrix}, \quad V = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n).$$

La ecuación  $T = V^T A V$  puede escribirse como  $AV = VT$ , que matricialmente es:

$$A(\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) = (\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n) \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \beta_3 & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \beta_{n-1} & \alpha_n \end{bmatrix}.$$

Desarrollando esta expresión,

$$\begin{aligned} A\mathbf{v}_1 &= \alpha_1\mathbf{v}_1 + \beta_1\mathbf{v}_2 \\ A\mathbf{v}_2 &= \beta_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \beta_2\mathbf{v}_3 \\ A\mathbf{v}_3 &= \beta_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3 + \beta_3\mathbf{v}_4 \\ &\vdots \end{aligned}$$

es decir,

$$A\mathbf{v}_j = \beta_{j-1}\mathbf{v}_{j-1} + \alpha_j\mathbf{v}_j + \beta_j\mathbf{v}_{j+1}, \quad (4.6)$$

para  $j = 1, 2, \dots, n-1$ .

Suponiendo que  $\beta_0 = \mathbf{v}_0 = 0$ , entonces

$$\begin{aligned} \mathbf{v}_j^T A \mathbf{v}_j &= \beta_{j-1}\mathbf{v}_j^T \mathbf{v}_{j-1} + \alpha_j\mathbf{v}_j^T \mathbf{v}_j + \beta_j\mathbf{v}_j^T \mathbf{v}_{j+1} \\ &= 0 + \alpha_j \|\mathbf{v}_j\|^2 + 0 \\ &= \alpha_j, \end{aligned}$$

luego

$$\alpha_j = \mathbf{v}_j^T A \mathbf{v}_j, \quad j = 1, 2, \dots, n. \quad (4.7)$$

A partir de las expresiones ([expresión \(4.6\)](#)) y ([expresión \(4.7\)](#)) se puede construir el algoritmo de Lanczos.

**Algoritmo 4.7.1 (Algoritmo de Lanczos)** **Entrada:** Matriz  $A$  simétrica y  $\mathbf{v}_1$  unitario.  
**Salida:**  $T$  tridiagonal y  $V$  ortogonal.

**Step 1** Inicializar  $\mathbf{v}_1$  y  $\mathbf{v}_0 \leftarrow 0$ ,  $\beta_0 \leftarrow 0$ .

**Step 2** Se realiza el bucle:

```
for i = 1,2,...,n
     $\mathbf{w}_j \leftarrow A\mathbf{v}_j$ 
     $\alpha_j \leftarrow \mathbf{w}_j^T \mathbf{v}_j$ 
     $\mathbf{w}_j \leftarrow \mathbf{w}_j - \alpha_j \mathbf{v}_j - \beta_{j-1} \mathbf{v}_{j-1}$ 
     $\beta_j \leftarrow \|\mathbf{w}_j\|$ 
     $\mathbf{v}_{j+1} \leftarrow \mathbf{w}_j / \beta_j$ 
endfor
```

**Step 3** Se hace  $\mathbf{w}_n \leftarrow A\mathbf{v}_n$ .

**Step 4** Se hace  $\alpha_n \leftarrow \mathbf{w}_n^T \mathbf{v}_n$ .

Un ejemplo es:

**Ejemplo 4.7.1 (Ejemplo del algoritmo de Lanczos)** Se considera

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}.$$

Aplicando el algoritmo de Lanczos para determinar los valores y vectores propios de esta matriz.

Se inicializa  $\mathbf{v}_1 = (1,0,0)^T$  y  $\mathbf{v}_0 \leftarrow 0$ ,  $\beta_0 \leftarrow 0$ .

$j = 1$  Se ejecutan los pasos del algoritmo para  $j = 1$ .

$$\mathbf{w}_1 = A\mathbf{v}_1 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

$$\alpha_1 = \mathbf{w}_1^T \mathbf{v}_1 = (1,2,3) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1.$$

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - 1 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - 10 = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix}.$$

$$\beta_1 = \|\mathbf{w}_1\| = \sqrt{13} = 3.6056.$$

$$\mathbf{v}_2 = \mathbf{w}_1 / \beta_1 = \begin{bmatrix} 0 \\ 0.5547 \\ 0.8321 \end{bmatrix}.$$

$j = 2$  Se ejecutan los pasos del algoritmo para  $j = 2$ .

$$\mathbf{w}_2 = A\mathbf{v}_2 = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix} \begin{bmatrix} 0 \\ 0.5547 \\ 0.8321 \end{bmatrix} = \begin{bmatrix} 3.6056 \\ 4.9923 \\ 6.3791 \end{bmatrix}.$$

$$\alpha_2 = \mathbf{w}_2^T \mathbf{v}_2 = 8.0769.$$

$$\mathbf{w}_2 = \mathbf{w}_2 - \alpha_2 \mathbf{v}_2 - \beta_2 \mathbf{v}_1 = \begin{bmatrix} 0 \\ 0.5120 \\ -0.3414 \end{bmatrix}.$$

$$\beta_2 = \|\mathbf{w}_2\| = 0.6154.$$

$$\mathbf{v}_3 = \mathbf{w}_2 / \beta_2 = \begin{bmatrix} 0 \\ 0.8321 \\ -0.5547 \end{bmatrix}.$$

Se calcula finalmente  $\mathbf{w}_3$  y  $\alpha_3$ .

$$\mathbf{w}_3^T A \mathbf{v}_3 = \begin{bmatrix} 0 \\ 0.2774 \\ 0.5547 \end{bmatrix} \longrightarrow \alpha_3 = \mathbf{w}_3^T \mathbf{v}_3 = -0.0769.$$

$$T_3 = \begin{bmatrix} \alpha_1 & \beta_1 & 0 \\ \beta_1 & \alpha_2 & \beta_2 \\ 0 & \beta_2 & \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 3.6056 & 0 \\ 3.6056 & 8.0769 & 0.6134 \\ 0 & 0.6154 & -0.0769 \end{bmatrix}$$

$$V = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.5547 & 0.8321 \\ 0 & 0.8321 & -0.5547 \end{bmatrix}.$$

Como se puede comprobar en el ejemplo estudiado, se cumple que  $T$  es una matriz semejante a  $A$ , por lo que se verifica que

$$AV = VT.$$

Lo más importante de este algoritmo es que se obtiene una matriz tridiagonal  $T$  cuyos vectores propios constituyen una buena aproximación de los vectores propios de la matriz  $A$ .

El problema ahora es obtener los valores y vectores propios de una matriz tridiagonal  $T$  de la forma

$$T = \begin{bmatrix} \alpha_1 & \beta_1 & & & \\ \beta_1 & \alpha_2 & \beta_2 & & \\ & \beta_2 & \alpha_3 & \beta_3 & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & \beta_{n-1} & \alpha_n \end{bmatrix}.$$

Pero este problema queda resuelto aplicando el algoritmo  $QR$  basado en la factorización  $QR$  estudiada en el apartado anterior.

## 4.8 Aplicación: las matrices de Markov

Un proceso de sucesión de eventos que se desarrolla en el tiempo en el cual el resultado en cualquier etapa contiene todo el elemento que depende del azar es denominado proceso aleatorio o proceso estocástico. Por ejemplo, un problema del estocástico podría ser las condiciones del tiempo en Paraná en una serie de días consecutivos: el tiempo cambia día un día de una manera que en apariencia es algo aleatorio. O bien, podría consistir en los precios de las acciones que cotizan en la bolsa y otra vez interviene cierto grado de aleatoriedad.

Las **cadenas de Markov** representan una herramienta para analizar el comportamiento y el gobierno de determinados tipos de procedimientos estocásticos, esto es, procesos que evolucionan de forma no determinista a lo largo del tiempo a partir de un conjunto de estados. Estas cadenas reciben su nombre del matemático ruso Andrei Andreevitch Markov (1856-1922). Estas cadenas tienen memoria, recuerdan el último evento y eso condiciona las posibilidades de los eventos futuros. Esto justamente las distingue de una serie de eventos independientes como el hecho de tirar una moneda.

Una cadena de Markov representa, por tanto, un sistema que varía su estado a través del tiempo, representando cada cambio una transición del sistema. Dichos cambios no están predeterminados, aunque sí lo está la probabilidad del próximo estado en función de los estados anteriores, probabilidad que es constante a lo largo del tiempo. Eventualmente, en una transición, el nuevo estado puede ser el mismo que el anterior, y es posible que exista la posibilidad de influir en las probabilidades de transición actuando adecuadamente sobre el sistema.

Centramos nuestra atención en las llamadas cadenas de Markov finitas, en las que el número de estados de transición es finito.

**Definición 4.8.1** Formalmente, podemos definir una cadena de Markov a partir de los siguientes elementos:

- Un conjunto de estados del sistema.
- La definición de transición.
- La ley de probabilidad condicional, que defina la probabilidad del nuevo estado en función de los estados anteriores.

Los estados son una caracterización de la situación en que se halla el sistema en un instante dado, pudiendo ser dicha caracterización tanto cuantitativa como cualitativa. Desde un punto de vista eminentemente práctico, la pregunta que podemos formularnos para entender lo que significa un estado concreto en un sistema es ¿Cómo están las cosas?

Más formalmente, el estado de un sistema en un instante  $t$  es una variable cuyos valores sólo pueden pertenecer al conjunto de estados del sistema. El sistema modelizado por la cadena, por lo tanto, es una variable que cambia de valor con el tiempo, cambio al que llamamos **transición**.

Dicho de otro modo, se trata de un conjunto de variables indexadas en el tiempo  $E_t$ , donde  $t$  denota intervalos significativos de tiempo para el fenómeno estudiado. Los valores de  $E_t$  se toman de un conjunto de categorías mutuamente excluyentes denominados estados del sistema. Por tratarse de un proceso estocástico, se conocerá únicamente la probabilidad asociada a cada uno de los estados.

Podemos considerar un ejemplo de cadena de Markov con dos estados  $A$  y  $B$ ; en general, tendremos un número finito  $n$  de estados. En cada etapa el proceso puede estar en cualquiera de los dos estados; en la etapa siguiente, el proceso puede permanecer en su estado actual o cambiar al otro estado. El estado al que el proceso se mueve en la etapa siguiente y la probabilidad de hacerlo así, dependen solamente del estado actual y no de los estados anteriores del proceso. Estas probabilidades decimos que son probabilidades de transición y suponemos que son constantes (es decir, la probabilidad de moverse del estado  $i$  al estado  $j$  es siempre la misma). Veamos un ejemplo.

**Ejemplo 4.8.1** Supongamos que en una zona geográfica, que comprende varias ciudades, se venden dos marcas diferentes de leche, que identificamos como  $A$  y  $B$ . Sabemos que debido al precio, a la publicidad, etc., el 80% de las personas que durante un mes compraron la marca  $A$ , continuaron comprándola al mes siguiente, mientras que el 20% cambió a la marca  $B$ ; análogamente, el 70% de las personas que durante un mes compraron la marca  $B$ , continuaron comprándola al mes siguiente, mientras que el 30% cambió a la marca  $A$ .

Elegimos una muestra de 1000 personas de las que 640 compran la marca  $A$  mientras que las 360 restantes compran la marca  $B$ . Al cabo de un mes, el número de personas que compran la marca  $A$  será

$$0.80 \cdot 640 + 0.30 \cdot 360 = 620$$

es decir, el 80% de las que inicialmente compraban la marca  $A$  más el 30% de las que compraban la marca  $B$  que cambiaron a la marca  $A$ . Análogamente, al cabo de un mes, el número de personas que compran la marca  $B$  será

$$0.20 \cdot 650 + 0.70 \cdot 350 = 380$$

es decir, el 20% de las que compraban la marca  $A$  que cambiaron a la marca  $B$  más el 70% de las que inicialmente compraban la marca  $B$ . Podemos escribir las expresiones anteriores en forma matricial como

$$\begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix} \begin{bmatrix} 640 \\ 360 \end{bmatrix} = \begin{bmatrix} 620 \\ 380 \end{bmatrix}$$

o, escrito de forma análoga,

$$P\mathbf{x}_0 = \mathbf{x}_1,$$

siendo

$$\mathbf{x}_0 = \begin{bmatrix} 640 \\ 360 \end{bmatrix}, \quad \mathbf{x}_1 = \begin{bmatrix} 620 \\ 380 \end{bmatrix},$$

los vectores correspondientes al número de personas que inicialmente y al cabo de un mes, respectivamente, compraban las marcas  $A$  y  $B$  (en dicho orden).

La matriz

$$P = \begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix}$$

decimos que es la **matriz de transición** de la cadena de Markov. Podemos pensar en las columnas de  $P$  como los estados actuales y en las filas de  $P$  como los estados siguientes.

Para determinar el número de personas que compran cada marca al cabo de dos meses, simplemente aplicamos el mismo razonamiento, comenzando con  $\mathbf{x}_1$  en lugar de  $\mathbf{x}_0$ . Así,

$$vecx_2 = P \cdot \mathbf{x}_1 = \begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix} \begin{bmatrix} 620 \\ 380 \end{bmatrix} = \begin{bmatrix} 610 \\ 390 \end{bmatrix}$$

Notemos que  $P$  es una matriz estocástica (la suma de sus columnas es igual a 1).

En general, si tenemos una cadena de Markov con  $n$  estados, entonces la matriz de transición  $P$  es una matriz estocástica de tamaño  $n \times n$  y los vectores  $\mathbf{x}_k$  decimos que son los vectores de estados. Así, para una cadena de Markov tenemos la relación

$$\mathbf{x}_{k+1} = P \cdot \mathbf{x}_k, \quad k = 0, 1, 2, \dots$$

que nos permite calcular un vector de estados arbitrario de forma iterativa una vez que conocemos  $\mathbf{x}_0$  y  $P$ . Es decir, una cadena de Markov está completamente determinada por su matriz de transición y su estado inicial. Supongamos que, en el ejemplo no queremos conocer el número exacto de personas que compra cada una de las marcas de leche sino que, en su lugar, estamos interesados en las proporciones de personas que compran cada marca.

Podemos convertir los datos en proporciones dividiendo entre 1000. De este modo, comenzaríamos con el vector

$$\begin{bmatrix} \frac{620}{1000} \\ \frac{380}{1000} \end{bmatrix} = \begin{bmatrix} 0.64 \\ 0.36 \end{bmatrix}$$

que nos proporciona la información de que el 64% de la población compra inicialmente la marca  $A$  mientras que el 36% se decanta por la marca  $B$ . Así,

$$\mathbf{x}_1 = P \cdot \mathbf{x}_0 = \begin{bmatrix} 0.62 \\ 0.38 \end{bmatrix}.$$

Los vectores de este tipo, cuyas componentes no son negativas y que suman 1 se llaman **vectores de probabilidad**. Por otro lado sabemos que

$$\mathbf{x}_k = P^k \cdot \mathbf{x}_0, \quad k = 1, 2, \dots$$

lo cual nos sugiere examinar las potencias de  $P$ .

Examinamos  $P^2$ ,

$$P^2 = \begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix} \begin{bmatrix} 0.80 & 0.30 \\ 0.20 & 0.70 \end{bmatrix} = \begin{bmatrix} 0.70 & 0.45 \\ 0.30 & 0.55 \end{bmatrix}$$

que vuelve a ser una matriz estocástica.

Consideremos, por ejemplo, el elemento  $p_{12}^2 = 0.45$  del ejemplo anterior. Podemos construir un diagrama de árbol que describa el origen de este elemento. Si inicialmente el proceso está en el estado  $B$ , hay cuatro cambios de estado posibles que pueden ocurrir a lo largo de dos meses, los cuales corresponden a las cuatro trayectorias de longitud 2 del diagrama de árbol. Una persona que inicialmente compraba la marca  $B$  puede terminar, dos meses después, comprando la marca  $A$ , de dos formas diferentes:

- puede continuar comprando  $B$  después de un mes y posteriormente comprar  $A$ , con probabilidad  $0.30 \cdot 0.80 = 0.24$ .
- puede cambiar a  $A$  después de un mes y entonces quedarse con  $A$ , con probabilidad  $0.70 \cdot 0.30 = 0.21$ .

La suma de ambas probabilidades da una probabilidad total de 0.45. Ahora bien, estos cálculos son precisamente los que hacemos para calcular  $p_{12}^2$ . Así pues,  $p_{12}^2 = 0.45$  es la probabilidad de moverse del estado 2 (comprar la marca  $B$ ) al estado 1 (comprar la marca  $A$ ) en dos transiciones.

Notemos, finalmente, que el orden de los subíndices es el contrario del que se podía esperar. Podemos generalizar el argumento del ejemplo anterior para obtener el siguiente resultado.

**Teorema 4.8.1** Si  $P$  es la matriz de transición de una cadena de Markov y  $P^k$  representa el producto de la matriz  $P = [p_{ij}]$  por ella misma  $k$  veces, entonces el elemento  $p_{ij}^k$  representa la probabilidad de moverse del estado  $j$  al estado  $i$  en  $k$  transiciones.

Volviendo al ejemplo anterior,

$$\mathbf{x}_0 = \begin{bmatrix} 0.64 \\ 0.36 \end{bmatrix} \quad \mathbf{x}_1 = \begin{bmatrix} 0.62 \\ 0.38 \end{bmatrix} \quad \mathbf{x}_2 = \begin{bmatrix} 0.61 \\ 0.39 \end{bmatrix}$$

$$\mathbf{x}_3 = \begin{bmatrix} 0.605 \\ 0.395 \end{bmatrix} \quad \mathbf{x}_4 = \begin{bmatrix} 0.6025 \\ 0.3975 \end{bmatrix} \quad \mathbf{x}_5 = \begin{bmatrix} 0.6012 \\ 0.3987 \end{bmatrix}$$

$$\mathbf{x}_6 = \begin{bmatrix} 0.6006 \\ 0.3994 \end{bmatrix} \quad \mathbf{x}_7 = \begin{bmatrix} 0.6003 \\ 0.3997 \end{bmatrix} \quad \mathbf{x}_8 = \begin{bmatrix} 0.6002 \\ 0.3998 \end{bmatrix}$$

Lo que podemos deducir de esta sucesión es que, en el futuro, el 60% de las personas comprarán la marca *A* mientras que el 40% elegirán la marca *B*. Es fácil comprobar que

$$\mathbf{x}_k = \begin{bmatrix} 0.6 \\ 0.4 \end{bmatrix},$$

para  $k \geq 10$ .

Observamos que se cumple una propiedad destacada:  $P \cdot \mathbf{x} = \mathbf{x}$ . Un vector verificando esta propiedad se llama **vector de estado estacionario**. En términos matemáticos decimos que 1 es un valor propio de  $P$  y que  $\mathbf{x}$  es un vector propio asociado al valor propio 1.

### Ejemplo 4.8.2 La formación de Huracanes

Los huracanes son un tipo de ciclones tropicales caracterizados por vientos intensos y lluvias abundantes, que se originan sobre los océanos tropicales en ambos lados del ecuador.

Los huracanes pueden llegar a tener un radio de entre 300 y 500 km., mientras que el centro del mismo, llamado ojo del huracán, tiene un diámetro del orden de 20 – 30 km., de forma circular o elíptica.

Diversos factores son imprescindibles para la formación de un huracán, como son:

- La principal fuente de energía para la formación de un huracán es el calor que extrae del océano; el agua del océano deberá estar por encima de los 26.6°C para que se evapore el agua necesaria. Cuando se desplaza por temperaturas menores, la pérdida de energía es grande y se disipa rápidamente.
- Los vientos deben encontrarse cerca de la superficie.
- El efecto Coriolis.
- Condiciones inestables en la atmósfera.
- El aire por encima de los 5486 km. necesitará estar húmedo al unirse a la tormenta.

El movimiento de un huracán es como el de un trompo. En sus primeras fases de formación, cuando se le llama tormenta tropical, se desplaza muy lentamente (de 15 a 20 km/h). Conforme el huracán madura aumenta su velocidad (de 30 a 45 km/h). Puede pasar un momento en que se detenga para después acelerarse por vientos del oeste y alcanzar velocidades de 80 a 100 km/h.

Veamos la relación existente entre la formación de huracanes y el álgebra lineal. Sabemos que la variabilidad de la temperatura es un factor bastante importante en la generación de los mismos.

La temperatura tanto de la atmósfera como del océano no puede ser determinada con certeza; sin embargo, la probabilidad de que se encuentre a una temperatura en algún determinado momento y lugar puede ser especificada si se han realizado observaciones previas. Este proceso constituye un ejemplo de cadena de Markov. En muestreos oceánicos o por imágenes vía satélite, se podrían registrar en tiempos diferentes (cada día, mes, etc.) tres estados térmicos en el agua de la superficie del océano:

- CÁLIDO (C).

**Tabla 4.1** Rango de los distintos estados térmicos..

Estado térmico	Rango temperaturas	Formación huracán
Frío (F)	(24.30,26.02)	NO
Moderado 1 (M1)	(26.03,26.59)	NO
Moderado 2 (M2)	(26.60,27.74)	SI
Caliente (C)	(27.75,29.47)	SI

- MODERADO (M).
- FRÍO (F).

Si a cada estado térmico se le asigna una probabilidad asociada a cambiar de un estado a otro o permanecer en el mismo, entonces podremos generar lo que hemos denominado una matriz de transición,  $P$ .

Entonces, con estos datos podemos conocer la probabilidad de que tengamos temperaturas ideales para la formación de huracanes.

Veamos un ejemplo concreto.

Con los datos existentes de la temperatura en una determinada región ( $10^{\circ}$  N,  $98,5^{\circ}$  W), desde el año 1968 podemos proceder a generar una matriz de transición.

Necesitamos inicialmente establecer los rangos de los distintos estados térmicos, que pueden ser los que aparecen en la tabla 4.1.

El estado del tiempo moderado se dividió en dos intervalos para conocer las temperaturas de ese rango que ayudan a la formación de huracanes y las que no. La matriz de transición generada a partir de los datos tomados es la siguiente

	F	M1	M2	C
F	0.77	0.60	0.18	0.00
M1	0.15	0.20	0.27	0.00
M2	0.08	0.20	0.55	0.25
C	0.00	0.00	0.00	0.75

Las columnas representan el estado actual, mientras que las filas representan el estado siguiente. Así, por ejemplo, la probabilidad que actualmente la temperatura sea moderada 2 y pase a fría en el instante siguiente es de 0.18.

Si a esta matriz de transición la multiplicamos por el vector de probabilidades para el tiempo actual, entonces obtendremos la probabilidad de que la temperatura superficial del océano para el siguiente mes sea de

$$\begin{array}{c|cccc} & F & M1 & M2 & C \\ \hline F & 0.77 & 0.60 & 0.18 & 0.00 \\ M1 & 0.15 & 0.20 & 0.27 & 0.00 \\ M2 & 0.08 & 0.20 & 0.55 & 0.25 \\ C & 0.00 & 0.00 & 0.00 & 0.75 \end{array} \begin{bmatrix} 0.18 \\ 0.32 \\ 0.35 \\ 0.15 \end{bmatrix} = \begin{bmatrix} 0.3936 \\ 0.1855 \\ 0.3084 \\ 0.1125 \end{bmatrix}.$$

Interpretando estos datos llegamos a la conclusión de que existe una mayor probabilidad, 0.3936, de que la temperatura superficial caiga dentro del rango de temperatura fría y, por

consiguiente, no se formen huracanes. Se advierte que existe una probabilidad de 0.3084 de que la temperatura caiga en el rango de moderado 1. Los otros dos estados son mucho menos probables (del 18% y del 11% aproximadamente).

Estos datos se realizan teniendo en cuenta que estamos simplificando al máximo el problema, ya que solamente contamos como factor para la formación de huracanes la temperatura superficial del océano. Además, sólo estamos midiendo los datos en un punto geográfico determinado.

#### Cuestiones

1. Intenta predecir para los siguientes meses las probabilidades de cada uno de los estados térmicos.
2. ¿Qué probabilidades podemos calcular de aquí a tres años? ¿Y a cinco?
3. Haz un estudio similar variando ligeramente los valores del vector de probabilidad actual. ¿Qué conclusiones obtienes?

#### The Hurricane Hunters

Los Cazadores de Huracanes vuelan en sus aviones hasta el ojo del huracán. Son el Escuadrón 53 de Reconocimiento Meteorológico de la Reserva de la Fuerza Aérea de Estados Unidos y forman parte de la base militar Keesler de Biloxi, Mississippi. Las fotografías más sorprendentes tomadas en el interior de los huracanes pertenecen a los Cazadores de Huracanes.

La misión principal es el reconocimiento de tormentas tropicales, tal y como establece el Plan de Operaciones de Huracanes. Este plan especifica que el Escuadrón 53 llevará a cabo operaciones 24 horas al día y será capaz de estudiar hasta 3 tormentas a la vez con un tiempo de respuesta de 16 horas. Para poder llevar a cabo estos objetivos, cuentan con un personal de vuelo de 10 personas a tiempo completo y 10 a tiempo parcial.

Cuando una tormenta empieza a formarse, el Centro Nacional de Huracanes manda a los Cazadores a investigar si los vientos soplan en sentido contrario a las agujas del reloj, lo que indica un sistema cerrado. Una vez que esto se determina, la misión cambia, y los Cazadores se dirigen al ojo del huracán. Para poder tomar distintos datos, los pilotos procuran entrar y salir por distintos sitios. Durante el vuelo, se recolecta información climática continuamente y se transmite al Centro Nacional de Huracanes vía satélite. Mientras se introducen en el ojo de la tormenta, sueltan un pequeño aparato con un paracaídas que ayudará a determinar el viento en la superficie y después, ya dentro del ojo, sueltan otro aparato para detectar la presión más baja. Una vez que salen del ojo, transmiten un mensaje que incluye toda esta información, así como la latitud y la longitud exactas del centro y las temperaturas máximas. ¿Y todo esto para qué? Pues porque con toda esta valiosísima información es posible determinar el recorrido, la fuerza y las dimensiones de los huracanes, lo que permite, entre otras cosas, saber qué áreas deben ser evacuadas y cuáles no.

## 4.9 Problemas

**4.1** ¿Podemos afirmar que

$$\begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix}.$$

es un vector propio de

$$A = \begin{bmatrix} 3 & 6 & 7 \\ 3 & 2 & 7 \\ 5 & 6 & 4 \end{bmatrix}$$

? Si lo fuera, encuentra el valor propio.

**4.2** ¿Podemos afirmar que  $\lambda = 1$  es un valor propio de

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

? Si lo fuera, encuentra el vector propio asociado a ese valor propio.

**4.3** Sin realizar ningún cálculo relacionado con el polinomio característico de  $A$ , encuentra un valor propio y dos vectores linealmente independientes de

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}$$

. Justifica tu respuesta.

**4.4** Sea  $A$  una matriz cuadrada de tamaño 2

$$\begin{bmatrix} 3 & 0 \\ 8 & -1 \end{bmatrix}.$$

Calcula todos sus valores y vectores propios.

¿Qué representa geométricamente el concepto de valor propio?

**4.5** Encuentra una base del vector propio correspondiente a cada uno de los valores propios que se listan.

$$A = \begin{bmatrix} 4 & 0 & 1 \\ -2 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}, \quad \lambda = 1, 2, 3.$$

$$A = \begin{bmatrix} 4 & 0 & -1 \\ 3 & 0 & 3 \\ -2 & -2 & 5 \end{bmatrix}, \quad \lambda = 3.$$

$$A = \begin{bmatrix} -4 & 1 & 1 \\ 2 & -3 & 2 \\ 3 & 3 & -2 \end{bmatrix}, \quad \lambda = -5.$$

**4.6** Consideremos  $A$

$$A = \begin{bmatrix} 4 & 3 & 3 \\ -4 & -3 & -4 \\ 2 & 2 & 3 \end{bmatrix}.$$

- Demuestra que  $(3, -4, 2)^T$  es un vector propio de  $A$ . ¿Cuál es su correspondiente valor propio?
- Determina todos los valores propios de  $A$  y sus vectores propios.

**4.7** Consideremos la matriz

$$A = \begin{bmatrix} -6 & 3 & -2 \\ -7 & 5 & -1 \\ 8 & -3 & 4 \end{bmatrix}$$

1. Calcula los valores y vectores propios de  $A$ .
2. ¿Existe una matriz diagonal semejante a  $A$ ? Razona la respuesta.

**4.8** Consideremos la matriz  $A = \begin{bmatrix} -6 & 3 & -2 \\ -7 & 5 & -1 \\ 8 & -3 & 4 \end{bmatrix}$

1. Calcula los valores y vectores propios de  $A$ .
2. ¿Existe una matriz diagonal semejante a  $A$ ? Razona la respuesta.

**4.9** Utiliza la teoría espectral para calcular el determinante de la matriz

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

**4.10** La matriz

$$A = \begin{bmatrix} 2 & -1 & -1 \\ 1 & 4 & 1 \\ -1 & -1 & 2 \end{bmatrix}$$

puede factorizarse de la forma  $PDP^{-1}$  como

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

Encuentra los valores y vectores propios de  $A$ .

**4.11** Diagonaliza las matrices

$$\begin{bmatrix} 2 & -1 \\ 1 & 4 \end{bmatrix}, \quad \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}.$$

**4.12** Diagonaliza la matriz

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

si es posible.

- 4.13**
- ¿Qué podemos decir de los valores propios de una matriz  $A$  que verifica que  $A^2 = 0$ ? Justifica la respuesta.
  - ¿Por qué una matriz  $A$  de tamaño  $n \times n$  puede tener como máximo  $n$  valores propios? ¿Deben ser todos distintos?



# 5 Grafos y matrices

---

## 5.1 Algunos ejemplos de grafos

El comienzo de la teoría de grafos se asocia a un artículo de Euler en 1736 tratando de dar solución al conocido problema de los puentes de Königsberg (actual Kalinigrado). El problema consistía en demostrar si era posible encontrar una ruta en la ciudad que recorriera los siete puentes, cruzando cada uno de ellos una única vez y regresando al punto de partida. Euler demostró que no era posible.

Luego, en 1847, Gustav Kirchhoff utilizó la teoría de grafos para el análisis de redes eléctricas publicando sus leyes de los circuitos para calcular el voltaje y la corriente en los circuitos eléctricos, conocidas como leyes de Kirchhoff, considerado la primera aplicación de la teoría de grafos a un problema de ingeniería.

Los grafos son perfectamente adecuados para modelar las interrelaciones entre un conjunto finito de objetos (carreteras que conectan varios pueblos o ciudades, enlaces de comunicaciones entre satélites, relaciones depredador-presa en un ecosistema, etc.) y las matrices son una herramienta útil para su estudio.

Veamos algunos ejemplos de grafos reales que representan problemas concretos o situaciones reales.

### 5.1.1 Tráfico aéreo

El grafo en cuestión que planteamos está relacionado con el tráfico aéreo en España. Hemos seleccionado los principales aeropuertos españoles y hemos estudiado las conexiones entre ellos mediante vuelos regulares. Aunque no figuran algunos aeropuertos, el tráfico aéreo de los que no se encuentran en esta lista es realmente insignificante respecto al de estos seleccionados y no resulta, en ningún caso, significativo.

Cada aeropuerto, tanto nacional como internacional se identifica mediante un código de tres letras mayúsculas. Los aeropuertos españoles con los que vamos a trabajar se identifican con las siguientes letras:

Aeropuerto	Identificador	Aeropuerto	Identificador
La Coruña	LCG	Madrid	MAD
Alicante	ALC	Almería	LEI
Asturias	OVD	Badajoz	BJZ
Barcelona	BCN	Bilbao	BIO
Fuerteventura	FUE	Las Palmas	LPA
Granada	GRX	Ibiza	IBZ
La Palma	SPC	Lanzarote	ACE
Málaga	AGP	Mahón	MAH
Murcia	MJV	Palma Mallorca	PMI
San Sebastián	EAS	Santiago	SCQ
Santander	SDR	Sevilla	SVQ
Tenerife N	TFN	Tenerife S	TFS
Valencia	VLC	Vigo	VGO
Zaragoza	ZAZ		

La USCA es la Unión Sindical de Controladores Aéreos que agrupa a más del 90% de los controladores aéreos españoles. En su página web <http://usca.es> se reflejan las incidencias sobre el tráfico aéreo en tiempo real. Se especifican las regulaciones aéreas, las demoras y sus causas. En cuanto al aspecto relacionado con la meteorología, es posible consultar la previsión meteorológica que nos ofrece AEMET, así como consultar la previsión de Turbulencias en vuelo en Europa (TURB).

Casi todos los aviones cuentan con una tecnología que les permite detectarlas poco antes de aproximarse a ellas. Además, los pilotos están permanentemente en comunicación con otros aviones que navegan en la misma ruta, intercambiando información sobre las condiciones atmosféricas. Basado en este sistema de detección de turbulencias se ha creado la web *Turbulence Forecast*, (<https://www.turbulenceforecast.com/>), el cual nos permitirá conocer un poco mejor el estado de la atmósfera y las posibles turbulencias que nos podemos encontrar por el camino. La forma de entender las previsiones son muy sencillas. Las turbulencias se muestran a través de unas manchas de colores en el mapa. Cada color corresponde a la intensidad de la turbulencia.

La web *Turbulence Forecast* permite visualizar el mapa de turbulencias de un periodo de tiempo determinado, concretamente el espacio que va entre las 6 y las 48 horas siguientes. El susto no nos lo quita nadie, pero al menos podemos hacernos una idea de como será nuestro viaje con este mapa de previsión de turbulencias para vuelos.

Se han determinado, a través de la página web de AENA, las rutas de los vuelos comerciales regulares entre los distintos aeropuertos que figuran en la tabla. A continuación se muestran los destinos nacionales desde cada aeropuerto, tomando como referencia sus siglas.

Id.	<b>Conexiones nacionales</b>
LCG	→ MAD, BCN, BIO, SVQ, VLC
MAD	→ LCG, ALC, LEI, OVD, BJZ, BCN, BIO, FUE, LPA, GRX, IBZ, SPC, ACE, AGP, MAH, MJV, PMI, EAS, SCQ, SDR, SVQ, TFS, VLC, VGO
ALC	→ MAD, OVD, BCN, BIO, LPA, IBZ, PMI, SCQ, SDR, SVQ, TFS
LEI	→ MAD, BCN, BIO, LPA, PMI, SCQ, SVQ
OVD	→ MAD, ALC, BCN, ACE, AGP, PMI, SVQ, TFS, VLC
BJZ	→ MAD, BCN
BCN	→ LCG, MAD, ALC, LEI, OVD, BJZ, BIO, FUE, LPA, GRX, IBZ, SPC, ACE, AGP, MAH, PMI, EAS, SCQ, SDR, SVQ, TFS, VLC, VGO
BIO	→ LCG, MAD, ALC, LEI, BCN, FUE, LPA, IBZ, ACE, AGP, PMI, SCQ, SVQ, TFS, VLC, VGO
FUE	→ MAD, BCN, BIO, LPA, ACE, AGP, SCQ, TFS
LPA	→ MAD, ALC, LEI, BCN, BIO, FUE, ACE, AGP, SCQ, SVQ, TFS, VLC
GRX	→ MAD, BCN, PMI,
IBZ	→ MAD, ALC, BCN, BIO, PMI, SCQ, SVQ, VLC, ZAZ
SPC	→ MAD, BCN, LPA, TFS
ACE	→ MAD, BCN, BIO, FUE, LPA, AGP, SCQ, SVQ, TFS, VLC
AGP	→ MAD, OVD, BCN, BIO, FUE, LPA, ACE, PMI, SCQ, SDR, TFS, VLC, VGO, ZAZ
MAH	→ MAD, BCN, PMI, SCQ, VLC, ZAZ
MJV	→ MAD
PMI	→ MAD, ALC, LEI, OVD, BCN, BIO, GRX, IBZ, AGP, MAH, SCQ, SVQ, VLC, ZAZ
EAS	→ MAD, BCN
SCQ	→ MAD, ALC, LEI, BCN, BIO, FUE, LPA, ACE, AGP, MAH, PMI, SVQ, TFS, VLC
SDR	→ MAD, ALC, BCN, AGP, SVQ, TFS
SVQ	→ LCG, MAD, ALC, LEI, OVD, BCN, BIO, LPA, IBZ, ACE, PMI, SCQ, SDR, TFS, VLC
TFN	→ MAD, ALC, OVD, BCN, BIO, FUE, LPA, ACE, AGP, SCQ, SVQ, VLC
TFS	→ MAD, BCN, FUE, LPA, AGP, SCQ, SDR, SVQ
VLC	→ LCG, MAD, OVD, BJZ, LPA, IBZ, AGP, MAH, PMI, SCQ, SVQ, TFS, VLC
VGO	→ MAD, BCN, BIO, AGP, VLC
ZAZ	→ IBZ, AGP, MAH, PMI

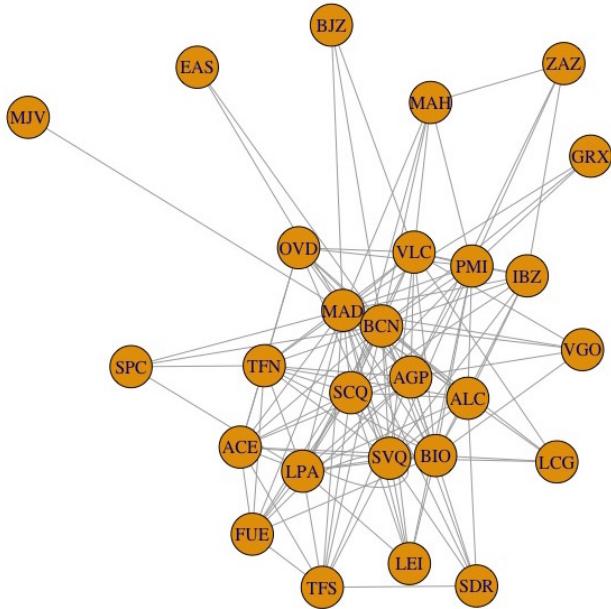
Notemos que los datos mostrados en la tabla anterior se refieren a vuelos nacionales; el ranking cambia si consideramos las conexiones internacionales, debido a la importancia del turismo y lo que ello representa para el tráfico aéreo.

Consideremos el grafo formado por las conexiones nacionales del tráfico aéreo de estos aeropuertos seleccionados. Los nodos serán cada uno de los aeropuertos y representaremos una arista entre dos nodos cuando exista una ruta comercial regular entre ambos aeropuertos. En consecuencia, tendremos un grafo no dirigido formado por 27 nodos. Vamos a utilizar, a modo de ejemplo, este grafo en lo que resta de libro. Podemos ver el grafo correspondiente en la [figura 5.1](#).

### 5.1.2 El grafo de un equipo de fútbol

En los deportes de equipo con un cierto número de jugadores por equipo, como por ejemplo el fútbol, el rugby, y otros, existen ciertas asociaciones entre algunos miembros del grupo mucho más fuertes que otras. Se oyen muchos comentarios por parte de los profesionales y expertos sobre lo bien que se entienden dos jugadores del equipo o lo bien que combinan entre sí.

Consideraremos el grafo representado en la [figura 5.2](#). Dicho grafo representa las posiciones de los jugadores del equipo de fútbol FC Barcelona y su disposición táctica



**Figura 5.1** Ejemplo del grafo de las conexiones aéreas..

sobre el terreno de juego. Se han representado los nodos con los números de cada uno de los jugadores que ocupan esa posición. En la parte inferior tenemos los nombres de los jugadores con el número que les corresponde. En el grafo hemos representado aristas entre aquéllos jugadores que más se asocian dentro del campo en el transcurso de un partido.

Hacemos una identificación de los nodos respecto al número de los jugadores, de manera que asociamos los vértices:

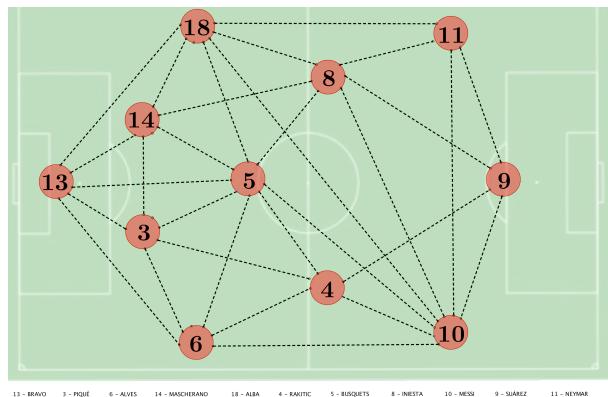
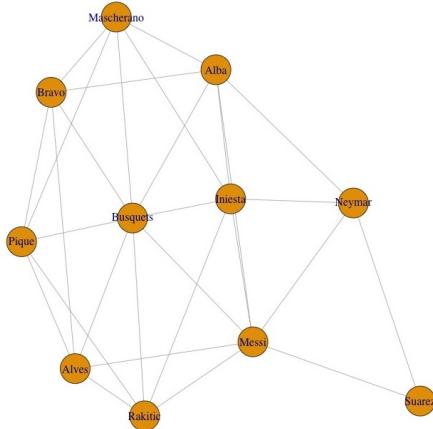
$$V = \{v_1, v_2, v_3, \dots, v_{11}\}$$

$$v_1 \rightarrow 13, v_2 \rightarrow 6, v_3 \rightarrow 3, v_4 \rightarrow 14, v_5 \rightarrow 18, v_6 \rightarrow 5,$$

$$v_7 \rightarrow 4, v_8 \rightarrow 8, v_9 \rightarrow 10, v_{10} \rightarrow 9, v_{11} \rightarrow 11.$$

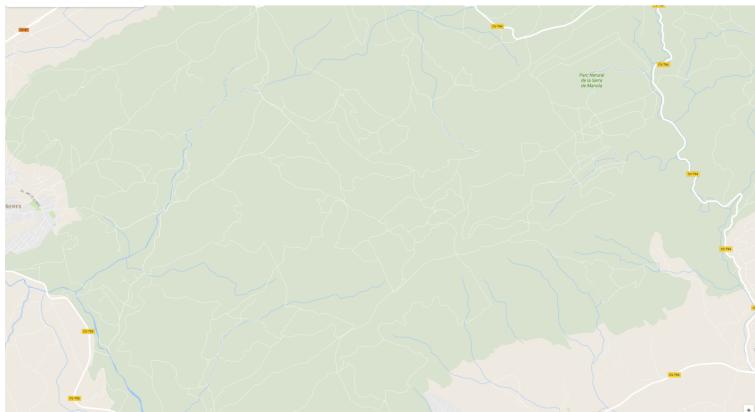
Disponemos de una hoja de datos, que hemos llamado *fcb.csv* que es la que hemos diseñado con 0 y 1 indicando cuándo se produce una conexión importante entre dos jugadores. Podemos tomar el criterio de los pases que se dan entre los jugadores para decidir si establecemos una conexión entre dos jugadores. Tenemos en la grafo en la [figura 5.3](#).

Veremos más adelante algunas propiedades relacionadas con estos grafos o redes.

**Figura 5.2** Grafo de ejemplo..**Figura 5.3** Grafo del equipo del Barcelona..

### Base de datos en Grafo.

Una Base de Datos en Grafo (BDG, a partir de ahora) es una base de datos que tiene como propósito almacenar estructuras de datos que tienen topología de grafo, es decir, que la información que se almacena se puede representar por medio de nodos y aristas entre ellos. Neo4j es un software libre de Base de datos orientada a grafos, implementado en Java. Si nosotros quisiéramos llevar al límite sus capacidades, tendríamos que superar un volumen total de 34,000 millones de nodos (datos), 34,000 millones de relaciones entre esos datos, 68,000 millones de propiedades y 32,000 tipos de relaciones.



**Figura 5.4** Mapa del entorno natural de Sierra de Mariola (Alicante)..

### 5.1.3 El grafo de una red de caminos

Durante el verano de 2012 se produjeron en todo el territorio nacional una serie de devastadores incendios que han arrasado miles y miles de hectáreas de bosque en zonas, la mayoría de ellas, de alto valor ecológico. Muchos parques naturales y reservas se ven afectadas por estos incendios que destruyen nuestro patrimonio natural con las terribles consecuencias que esto supone. Uno de estos incendios se produjo en el paraje natural conocido como la sierra de Mariola, que se encuentra en la provincia de Alicante próximo a las ciudades de Alcoy y Banyeres de Mariola.

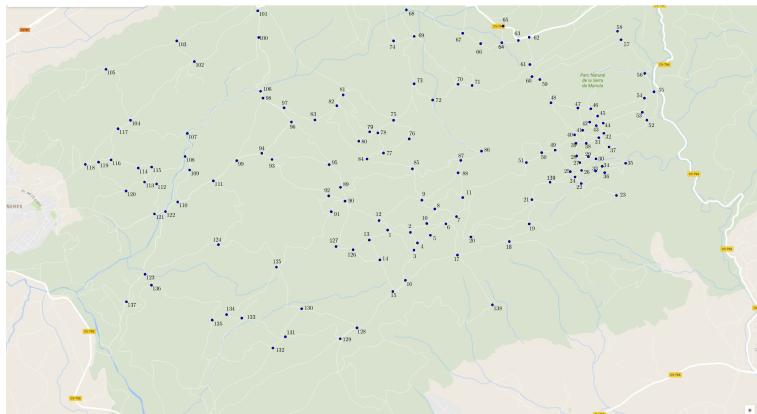
En la figura [figura 5.4](#) tenemos un mapa de una parte importante de la región del parque natural de la sierra de Mariola en el que podemos ver y apreciar diferentes caminos, carreteras y pistas forestales que atraviesan esta zona. Debido al alto valor de la flora y la fauna del lugar es necesario proteger la misma de este tipo de catástrofes.

Para proteger la zona, se desea establecer cinco bases con servicios básicos que cubran toda el área y en la que se puedan establecer unos puestos de control con material humano y técnico para la prevención de incendios, auxilio en caso de accidentes u otro tipo de incidencias. Así pues, necesitamos determinar cinco puntos de esta zona que estén bien comunicados con el máximo de puntos posible de la región.

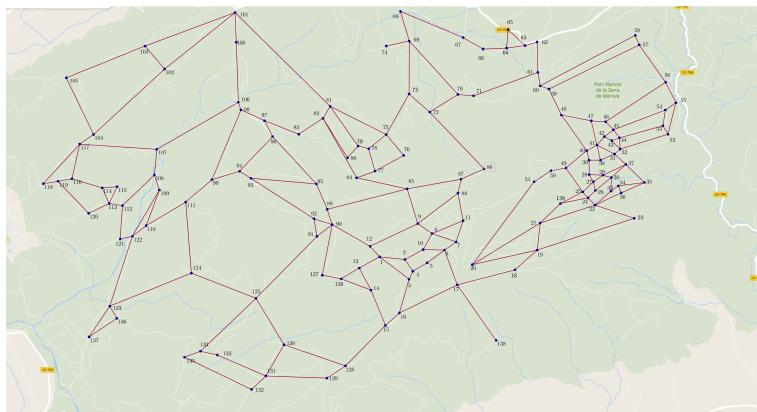
Podemos plantear este problema bajo la perspectiva de la modelización a través de un grafo que represente esta red de caminos. Para ello, debemos crear una red que simule el entramado de carreteras y caminos que tenemos en el mapa y, posteriormente, podríamos aplicar los conceptos o algoritmos básicos en grafos o teoría de redes.

En la figura [figura 5.5](#) vemos el conjunto de nodos del grafo que construimos reproduciendo la red de caminos. Cada nodo se ha dibujado junto al identificador numérico que tiene asociado. Tenemos un total de 139 nodos, teniendo en cuenta que nuestro objetivo no es el de representar fielmente y con gran nivel de detalle la red de caminos, sino demostrar cómo los conceptos de la teoría de redes tienen un amplio espectro de aplicaciones útiles.

Si unimos los nodos que tenemos mediante aristas siguiendo las trayectorias de los caminos o carreteras existentes en esta zona natural, obtenemos un grafo que representa,



**Figura 5.5** Mapa del entorno natural de Sierra de Mariola (Alicante) con los nodos dibujados..



**Figura 5.6** Grafo del entorno natural de Sierra de Mariola (Alicante)..

más o menos fielmente, una zona geo-localizada sobre la que podemos actuar a partir de su representación geométrica. Dicho grafo se muestra en la [figura 5.6](#).

#### 5.1.4 La colección de grafos KONECT

KONECT (Koblenz Network Collection, <http://konect.uni-koblenz.de>) es un proyecto para recopilar grandes conjuntos de datos de red de todo tipo con el fin de realizar investigaciones en ciencias de redes y campos relacionados, recopilados por el Instituto de Web Science and Technologies de la Universidad de Koblenz-Landau.

KONECT [1] contiene varios cientos de conjuntos de datos de red de varios tipos, incluidas redes dirigidas, no dirigidas, bipartitas, ponderadas, no ponderadas. Las redes de

KONECT cubren muchas áreas diversas, tales como redes sociales, redes de hipervínculos, redes de autoría, redes físicas, redes de interacción y redes de comunicación. El proyecto KONECT ha desarrollado herramientas de análisis de red de software libre que se utilizan para calcular estadísticas de red, dibujar gráficos e implementar varios algoritmos de predicción de enlaces.

Vamos a tomar como ejemplo una red del conjunto de 261 redes que han construido y analizado hasta el momento. Concretamente tomamos la red **Protein**, que consta de 1870 nodos (proteínas) y 2277 aristas (interacción entre proteínas). Antes de mostrar esta red, hablamos muy brevemente sobre algunos tipos de redes biológicas.

Se pueden representar diferentes tipos de información en forma de redes para modelar la célula. Los diferentes tipos de datos también producirán diferentes características generales de red en términos de conectividad, complejidad y estructura, donde los bordes y nodos potencialmente transmiten múltiples capas de información.

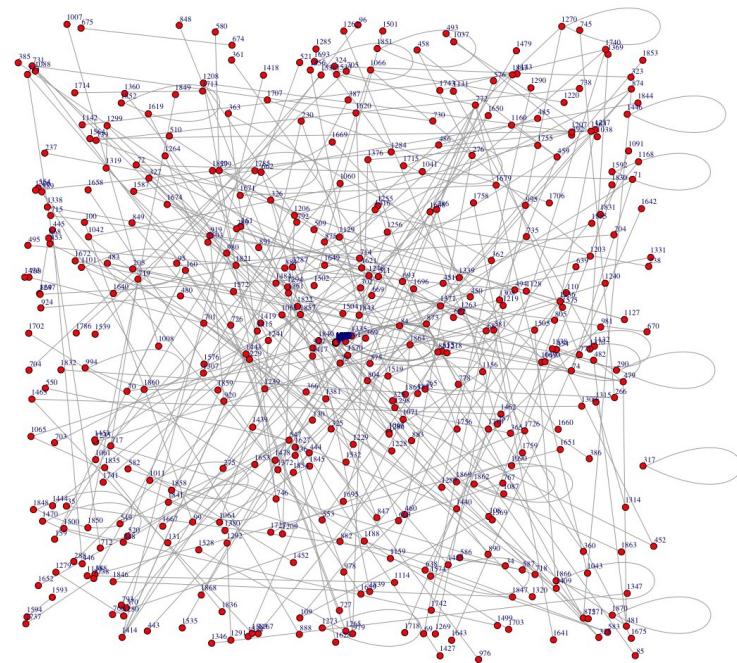
Algunos de los tipos más comunes de redes biológicas son:

- Redes de interacción proteína-proteína.
- Redes metabólicas.
- Redes de interacción genética.
- Redes normativas genéticas/transcripcionales.

Las interacciones proteína-proteína (PPI) son esenciales para casi todos los procesos en una célula, por lo que la comprensión de los PPI es crucial para comprender la fisiología celular en estados normales y de enfermedad. También es esencial en el desarrollo de fármacos, ya que los medicamentos pueden afectar los PPI. Las redes de interacción proteína-proteína (PPIN) son representaciones matemáticas de los contactos físicos entre proteínas en la célula.

Las proteínas raramente actúan solas ya que sus funciones tienden a regularse. Muchos procesos moleculares dentro de una célula se llevan a cabo mediante máquinas moleculares que se construyen a partir de una gran cantidad de componentes proteicos organizados por sus PPI. Estas interacciones conforman la llamada interactomía del organismo, mientras que los PPIs aberrantes son la base de múltiples enfermedades, como la enfermedad de Alzheimer.

Esta red no dirigida que vamos a estudiar contiene interacciones de proteínas contenidas en la levadura. La investigación demostró que las proteínas con un alto grado eran más importantes para la supervivencia de la levadura que otras. Un nodo representa una proteína y un borde representa una interacción metabólica entre dos proteínas. La red contiene bucles.



**Figura 5.7** Grafo de la red de proteínas..

La [figura 5.7](#) nos muestra un gráfico de la red de 1870 proteínas que nos sirve de ejemplo en este capítulo.

### 5.1.5 Grafos contra el terrorismo

Los acontecimientos del 11 de septiembre en Nueva York cambiaron las percepciones de las personas a nivel mundial, el hecho de usar las palabras *terrorismo* y *red* era un indicador de miedo para sociedades que se habían mantenido al margen de conflictos bélicos.

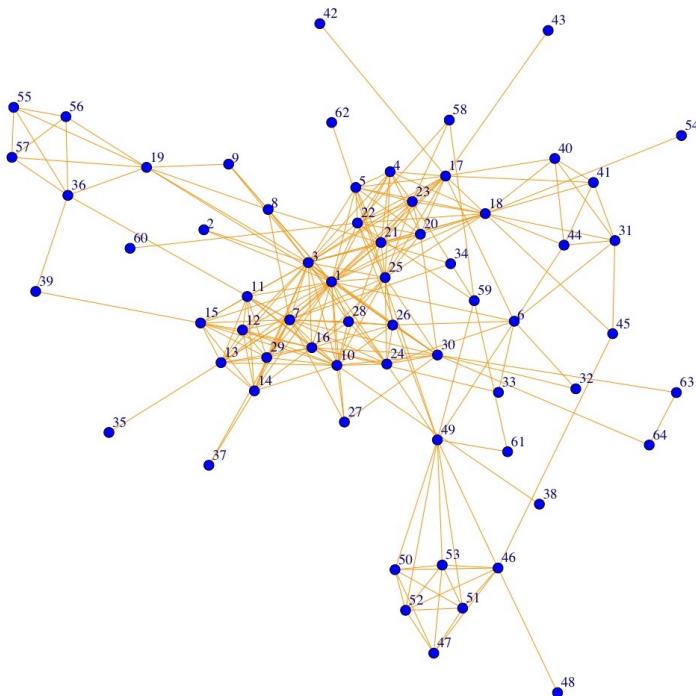
En la guerra convencional, es importante tener las ubicaciones específicas de los actores ya que es importante entender el terreno donde se libraran las batallas. Una característica importante en la guerra contra el terrorismo es que no hay una geo-localización del terreno de batalla. Los terroristas no tienen un campo geográfico de batalla, más bien operan en redes con miembros distribuidos en todo el mundo. En matemáticas se pueden analizar este tipo de redes, mediante los llamados grafos, que analizan las conexiones entre puntos conectados por diversas ramificaciones.

Posteriormente, en marzo de 2004, se produjo en Madrid el atentado más importante llevado a cabo por Al Qaeda después del de las torres gemelas de Nueva York. Cuatro trenes de cercanías sufrieron ataques perpetrados por teroristas que hicieron explotar bombas en el interior de los trenes.

José A. Rodríguez de la Universidad de Barcelona creó una red de personas involucradas en el ataque a los trenes suburbanos en Madrid el 11 de marzo de 2004. Rodríguez usó notas de prensa en los dos principales diarios españoles (*El País* y *El Mundo*) para reconstruir la red terrorista. Los nombres incluidos fueron de aquellas personas sospechosas de haber participado y sus relatos. Rodríguez especificó cuatro tipos de vínculos que vinculan a las personas involucradas:

- Confianza - amistad (contacto, parentesco, enlaces en el centro telefónico).
- Lazos con Al Qaeda y Osama Bin Laden.
- Coparticipación en campos de entrenamiento y / o guerras.
- Coparticipación en ataques terroristas previos.

Estos cuatro fueron agregados juntos proporcionando un índice de *fuerza de conexión* que va de 1 a 4. Para más detalles sobre esta cuestión, véase por ejemplo [2].



**Figura 5.8** Grafo de la red de terroristas del atentado de Madrid el 11 de marzo de 2004..

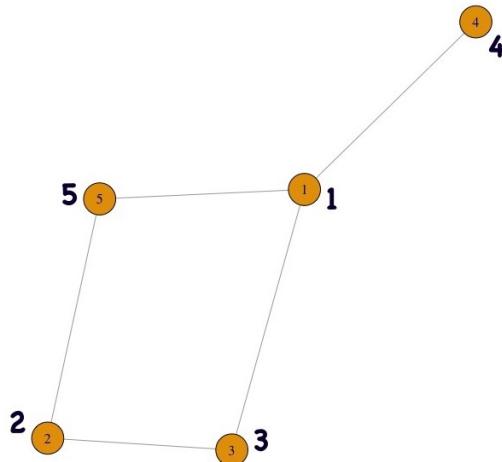
La figura 5.8 nos muestra el grafo construido para representar las conexiones entre las personas involucradas en los actos terroristas en los trenes de cercanías de Madrid. Tenemos 64 nodos que representan los terroristas y los 243 contactos entre ellos.

## 5.2 Breve introducción a los grafos

Los grafos son perfectamente adecuados para modelar las interrelaciones entre un conjunto finito de objetos (carreteras que conectan varios pueblos o ciudades, enlaces de comunicaciones entre satélites, relaciones depredador-presa en un ecosistema, etc.) y las matrices son una herramienta útil para su estudio.

El primer artículo científico relativo a los grafos fue escrito por el matemático suizo Leonhard Euler en 1736. Euler se basó en su artículo en el problema de los puentes de Königsberg. La ciudad de Kaliningrado, originalmente Königsberg, es famosa por sus siete puentes que unen ambas márgenes del río Pregel con dos de sus islas. Dos de los puentes unen la isla mayor con la margen oriental y otros dos con la margen occidental. La isla menor está conectada a cada margen por un puente y el séptimo puente une ambas islas. El problema planteaba lo siguiente: ¿es posible dar un paseo comenzando desde cualquiera de estas regiones, pasando por todos los puentes, recorriendo solo una vez cada uno y regresando al mismo punto de partida?

Un **grafo** está formado por un conjunto finito de puntos llamados **vértices o nodos** y un conjunto finito de **arcos o aristas**, cada una de las cuales conecta dos vértices (no necesariamente distintos). La figura 5.9 muestra la representación de un grafo con 5 vértices y 5 aristas.



**Figura 5.9** Representación gráfica de un grafo..

En adelante supondremos que  $G = (V(G), E(G))$  es un grafo con  $V(G) = \{1, 2, 3, \dots, n\}$  vértices y  $l$  aristas  $E(G) = \{e_1, e_2, e_3, \dots, e_l\}$ .

La primera de las matrices que vamos a estudiar asociadas a los grafos es la que llamamos **matriz de adyacencia** y su principal característica es que nos muestra la conectividad del grafo, es decir, su topología. Es fundamental en toda la teoría de grafos y redes.

**Definición 5.2.1** Sea  $G$  un grafo con  $n$  vértices. Llamamos **matriz de adyacencia** del grafo  $G$  a la matriz  $A = (a_{ij})_{i,j=1}^n$  tal que

$$a_{ij} = \begin{cases} 1, & \text{si existe un arco que une el vértice } i \text{ con el vértice } j, \\ 0, & \text{en otro caso.} \end{cases}$$

Así, la matriz de adyacencia del grafo de la figura 5.9 es

$$A = \begin{bmatrix} 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \end{bmatrix}.$$

Notemos que la matriz de adyacencia de este grafo es simétrica. ¿Por qué?

Continuamos con dos conceptos básicos a la hora de hablar de vértices y aristas.

**Definición 5.2.2** Una **trayectoria** es una sucesión de aristas que conecta dos vértices del grafo. Llamamos **longitud** de una trayectoria al número de aristas que contiene.

Así, para el grafo de la figura 5.9 tenemos que  $4 \rightarrow 1 \rightarrow 3 \rightarrow 2$  es una trayectoria de longitud 3 que conecta los vértices 4 y 2, mientras que  $1 \rightarrow 3 \rightarrow 2 \rightarrow 5 \rightarrow 1$  es una trayectoria de longitud 3 que conecta el vértice 1 consigo mismo.

**Definición 5.2.3** Llamamos **circuitos** a las trayectorias que empiezan y terminan en el mismo vértice.

Podemos utilizar las potencias de la matriz de adyacencia de un grafo para obtener información sobre las trayectorias de diferentes longitudes en el grafo. Así, para la matriz de adyacencia  $A$  del grafo de la figura 5.9 calculamos y analizamos el significado de la matriz  $A^2$ .

$$A^2 = \left[ a_{ij}^{(2)} \right]_{i,j=1}^n = \begin{bmatrix} 3 & 2 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 2 \end{bmatrix}.$$

Fijémonos en un elemento cualquiera de dicha matriz, por ejemplo  $a_{35}^{(2)}$  que, de acuerdo con la definición de multiplicación de matrices, se calcula como

$$a_{35}^{(2)} = a_{31}a_{15} + a_{32}a_{25} + a_{33}a_{35} + a_{34}a_{45} + a_{35}a_{55}$$

Para que este elemento sea un número distinto de cero necesitamos que alguno de los productos  $a_{3k}a_{k5}$  sea distinto de cero. Pero

$$a_{3k}a_{k5} \neq 0 \quad \text{si y solo si} \quad a_{3k} \neq 0 \text{ y } a_{k5} \neq 0$$

es decir, en el grafo tiene que haber un arco del vértice 3 al vértice  $k$  y otro del vértice  $k$  al vértice 5; equivalentemente, tiene que haber una trayectoria de longitud 2 que une el vértice 3 con el vértice 5. En nuestro ejemplo esto ocurre para  $k = 1$  y  $k = 2$ , de manera que

$$a_{35}^{(2)} = 1 \cdot 1 + 1 \cdot 1 + 0 \cdot 0 + 0 \cdot 0 + 0 \cdot 0 = 2$$

es el número de trayectorias de longitud 2 que conectan el vértice 3 con el vértice 5.

Análogamente, puesto que  $a_{12}^{(2)} = 2$ , tenemos que hay dos trayectorias de longitud 2 que conectan el vértice 1 con el vértice 2. Éstas son:  $1 \rightarrow 3 \rightarrow 2$  y  $1 \rightarrow 5 \rightarrow 2$ .

De  $a_{11}^{(2)} = 3$ , tenemos que hay tres circuitos de longitud 2 que conectan el vértice 1 consigo mismo: 1,3,1; 1,4,1 y 1,5,1.

En general, enunciamos el teorema siguiente cuya demostración omitimos por quedar fuera de los objetivos de este curso.

**Teorema 5.2.1** Si  $A$  es la matriz de adyacencia de un grafo  $G$  y

$$A^k = \left[ a_{ij}^{(k)} \right]_{i,j=1}^n,$$

entonces  $a_{ij}^{(k)}$  es el número de trayectorias de longitud  $k$  que unen el vértice  $i$  con el vértice  $j$ .

A modo de ejemplo, nos planteamos el cálculo del número de trayectorias de longitud

4 que unen el vértice 3 con el vértice 5. Para ello, necesitamos calcular el elemento  $a_{35}^{(4)}$ .

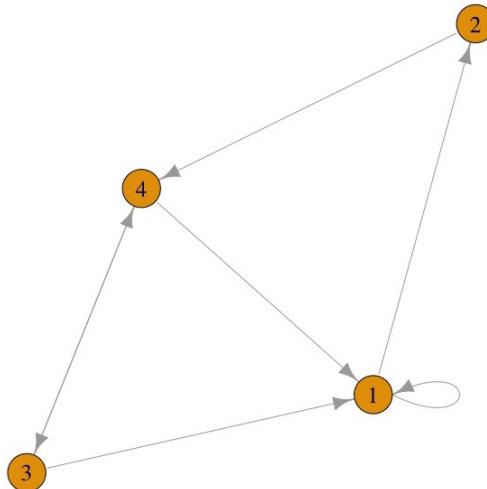
Puesto que

$$A^4 = \begin{bmatrix} 13 & 10 & 0 & 0 & 0 \\ 10 & 8 & 0 & 0 & 0 \\ 0 & 0 & 9 & 5 & 9 \\ 0 & 0 & 5 & 3 & 5 \\ 0 & 0 & 9 & 5 & 9 \end{bmatrix},$$

tenemos que  $a_{35}^{(4)} = 9$ , por tanto hay 9 trayectorias de longitud 4 que conectan el vértice 3 con el vértice 5.

En algunas aplicaciones que pueden ser modeladas mediante un grafo, los vértices están ordenados mediante algún tipo de relación que impone una dirección a los arcos. Por ejemplo, los arcos pueden ser utilizados para representar rutas de un solo sentido en un grafo que modele una red de transportes. Un grafo en el que los arcos están dirigidos decimos que es un **digrafo**.

La figura 5.10 muestra la representación gráfica de un digrafo con 4 vértices y 8 aristas, cuya matriz de adyacencia es

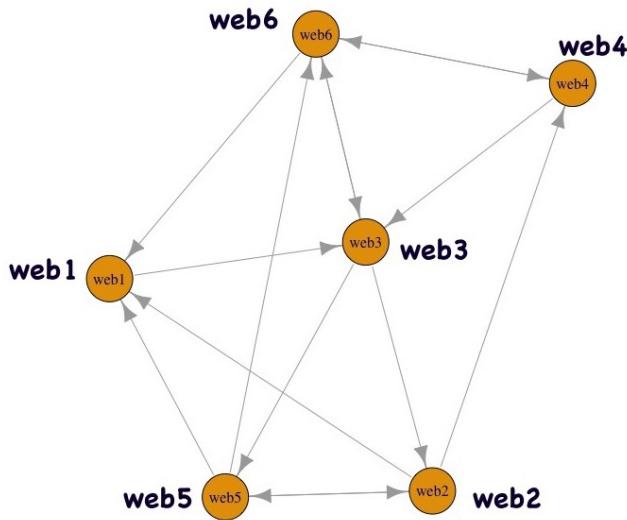


**Figura 5.10** Representación gráfica de un digrafo..

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}.$$

Como es lógico, la matriz de un digrafo no es, en general, simétrica. Además,  $A^k$  contiene los números de **trayectorias dirigidas** de longitud  $k$  entre los vértices del grafo. En una trayectoria dirigida todos los arcos de la trayectoria fluyen en la misma dirección. Así, para el digrafo de la figura 5.10 las trayectorias  $1 \rightarrow 2 \rightarrow 4 \rightarrow 3 \rightarrow 4 \rightarrow 1$  y  $3 \rightarrow 1 \rightarrow 2 \rightarrow 4$  son dirigidas.

**Ejemplo 5.2.1** Supongamos que tenemos 6 páginas web que nos aportan información sobre un cierto tema de interés para nosotros y queremos saber cuál de ellas es más importante o relevante; es decir, queremos establecer un ranking dentro de este conjunto de páginas. Para ello, decidimos que podemos crear una red entre ellas con los links o enlaces de una página a otra. Esto es lo que se conoce como un *webgraph*. Dibujaremos una conexión entre las páginas  $i$ -ésima y  $j$ -ésima cuando la página  $i$  tenga un enlace a la página  $j$  (decimos que referencia la página  $j$ ).



**Figura 5.11** Webgraph de 6 páginas web interrelacionadas..

El grafo que hemos dibujado en la figura 5.11 refleja las conexiones y referencias entre las seis páginas web. No tenemos en cuenta los enlaces de una página a ella misma, ni los enlaces exteriores de estas seis páginas, ni duplicamos los enlaces. Tampoco suponemos que existe una página que no tiene ningún enlace a otras páginas (no se producen nodos aislados en el grado).

La matriz de adyacencia  $A$  de este webgraph es aquella que en la entrada  $ij$  (elemento  $a_{ij}$ ) es 1 si la página  $j$ -ésima tiene un link hacia la página  $i$ -ésima y tiene un 0 en cualquier otro caso.

Así,

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

es la matriz de adyacencia del grafo. La matriz de adyacencia de este webgraph tiene un 1 en el elemento  $a_{ij}$  si la página  $j$ -ésima tiene un enlace a la página  $i$ . En otro caso,  $a_{ij} = 0$ . Debido a la definición de esta matriz de adyacencia, tenemos que:

- La suma de las entradas de una fila, por ejemplo, la fila  $i$  es el número de enlaces de entrada que tiene la página  $i$  desde otras páginas, es decir, el número de páginas que la mencionan.

- La suma de las entradas de la columna  $j$  es el número de enlaces desde la página  $j$  a otras páginas.

Supongamos que en un webgraph vamos *navegando* por las páginas que se van enlazando entre ellas. El **vector de estado**  $\mathbf{x}^k$  representa un vector columna cuya componente  $i$ -ésima representa la probabilidad de que un navegante esté en la página  $i$  después de  $k$  movimientos por las páginas del webgraph.

Supongamos que se conoce con certeza que un navegante está en la página  $j$ -ésima después de  $k$  clics con el ratón. En este caso, la  $j$ -ésima entrada del vector  $\mathbf{x}^{(k)}$  es 1 y todas las demás entradas son 0. Entonces, el producto  $A\mathbf{x}^{(k)}$  es el vector  $j$ -ésimo de  $A$ . Por ejemplo, si sabemos con certeza que un navegante comienza desde la página 2, entonces su vector de estado  $\mathbf{x}^{(0)}$  y el producto  $A\mathbf{x}^{(0)}$  será:

$$\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow A\mathbf{x}^{(0)} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}.$$

Las entradas unitarias que aparecen en el vector  $A\mathbf{x}^{(0)}$  nos indican que desde la página 2, el navegador tiene la opción de ir a las páginas 1, 4 o 5, ya que son las que están enlazadas con la 2.

Asociada a la matriz de adyacencia de un grafo, tenemos una matriz que llamamos *matriz de transición de probabilidades*, que definimos de la siguiente manera.

**Definición 5.2.4** Sea  $A = (a_{ij})$  la matriz de adyacencia de un grafo  $G$ . La matriz de **transición de probabilidades**  $B = (b_{ij})$  asociada a la matriz  $A$  es la matriz que se obtiene dividiendo cada entrada de  $A$  por la suma de las entradas de la columna en la que se encuentra, es decir,

$$b_{ij} = \frac{a_{ij}}{\sum_k a_{kj}}.$$

A partir de la definición de la matriz de transición de probabilidades, podemos calcular  $B$  como

$$B = \begin{bmatrix} 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 1 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix}.$$

Notemos que la suma de las columnas es 1 y todas sus entradas son 0 o positivas, lo que significa que es una matriz estocástica por columnas. Recordemos los contenidos estudiados en la sección 4.8 referentes a las cadenas de Markov.

Así, ahora realizamos los cálculos con la matriz  $B$ .

$$\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow B\mathbf{x}^{(0)} = \begin{bmatrix} 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 \\ 1 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & 0 & 0 & 0 & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{2} & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ 0 \\ 0 \\ \frac{1}{3} \\ \frac{1}{3} \\ 0 \end{bmatrix}.$$

El vector  $B\mathbf{x}^{(0)}$  nos indica que la probabilidad de navegar desde la página 2 a las páginas 1, 4 o 5 es igual a  $\frac{1}{3}$ .

Los vectores de estado de la navegación por las páginas relacionadas son

$$\mathbf{x}^{(0)} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \mathbf{x}^{(1)} = \begin{bmatrix} 0.3333 \\ 0 \\ 0 \\ 0.3333 \\ 0.3333 \\ 0 \end{bmatrix}, \mathbf{x}^{(2)} = \begin{bmatrix} 0.1111 \\ 0.1111 \\ 0.5000 \\ 0 \\ 0 \\ 0.2778 \end{bmatrix}, \dots, \mathbf{x}^{(15)} = \begin{bmatrix} 0.1544 \\ 0.1365 \\ 0.2727 \\ 0.1090 \\ 0.1365 \\ 0.1910 \end{bmatrix}$$

Podemos calcular la probabilidad de navegar entre las páginas de esta webpage a largo plazo, luego establecemos una clasificación de las páginas web en función de su conectividad.

Calculamos ahora los valores y vectores propios de la matriz  $B$ , es decir, las soluciones de la ecuación característica:

$$\det(B - \lambda I) = 0,$$

lo que produce las soluciones

$$\sigma(B) = \begin{bmatrix} 1 \\ -0.2977 + 0.5522i \\ -0.2977 - 0.5522i \\ -0.2554 \\ 0.1842 \\ -0.3333 \end{bmatrix}.$$

Calculamos el vector propio asociado al valor propio real mayor, en módulo, que es  $\lambda = 1$ . El vector propio es

$$\mathbf{x}(\lambda = 1) = \begin{bmatrix} 0.3605 \\ 0.3181 \\ 0.6361 \\ 0.2545 \\ 0.3181 \\ 0.4453 \end{bmatrix}. \quad (5.1)$$

El vector  $\mathbf{x}(\lambda = 1)$  constituye un vector donde todas sus componentes son números positivos. Pero, además, este vector propio constituye una ordenación de los nodos, en este

caso de páginas web. Esta ordenación nos dice que, dentro del ranking que ha quedado establecido, la página más importante o relevante en esta red es la número 3, mientras que la segunda en importancia es la página 6, quedando en tercera posición la página 1.

Si efectuamos una comparación entre el vector  $\mathbf{x}(\lambda = 1)$  y el vector de estado  $\mathbf{x}^{(15)}$  vemos la absoluta coherencia entre los datos que nos proporcionan ambos vectores. El vector  $\mathbf{x}(\lambda = 1)$  nos indica que la página más importante es la 3, mientras que el vector de estado  $\mathbf{x}^{(15)}$  nos dice que después de un número de saltos entre las páginas web de la red, la página 3 es la que presenta la mayor probabilidad de que nos encontremos en la misma. Pero, si observamos detenidamente los resultados que nos proporcionan ambos vectores notamos que esta coincidencia para el nodo o página 3 no es una casualidad; ocurre que la ordenación que nos ofrece  $\mathbf{x}(\lambda = 1)$  se mantiene siguiendo el vector  $\mathbf{x}(\lambda = 1)$  el mismo orden en sus probabilidades.

### 5.3 Matrices asociadas a un grafo

Recordemos que un grafo  $G = (V(G), E(G))$  está formado por un conjunto  $V(G)$  de nodos o vértices y  $E$  de aristas uniendo pares de vértices del conjunto  $V(G)$ . Suponemos que tenemos un conjunto de  $n$  vértices y  $l$  aristas. Dependiendo de que las aristas sean dirigidas o no dirigidas tendremos un **digrafo** (aristas dirigidas). También supondremos que el grafo  $G$  es conexo, es decir, que no existen vértices aislados o con grado cero.

Estudiamos diversas matrices asociadas a los grafos y digrafos.

**Definición 5.3.1 (Matriz de Adyacencia)** Dado un grafo  $G$  con  $n$  vértices sin bucles, definimos la matriz de adyacencia  $A = (a_{ij})_{i,j=1}^n$  como la matriz donde

$$a_{ij} = \begin{cases} 1 & \text{si hay una arista uniendo los vértices } i, j. \\ 0 & \text{en otro caso.} \end{cases}$$

La matriz de adyacencia es una matriz binaria que nos proporciona información sobre la conectividad del grafo que estamos estudiando, ya que podemos ver inmediatamente las conexiones entre los nodos. Además, cumple una propiedad que es muy importante en la teoría espectral, como es que es simétrica, es decir,  $a_{ij} = a_{ji}$ , para todo  $i, j = 1, \dots, n$ .

**Definición 5.3.2 (Matriz de Incidencia de un digrafo)** Dado un grafo  $G$  dirigido ( $G$  digrafo) con  $n$  vértices sin bucles, definimos la matriz de incidencia  $B = (b_{ij})_{i,j=1}^{n,l}$  como la matriz donde

$$b_{ij} = \begin{cases} 1 & \text{si la arista } j \text{ entra en el vértice } i \\ -1 & \text{si la arista } j \text{ sale del vértice } i \\ 0 & \text{en otro caso.} \end{cases}$$

Notemos que, mientras que la matriz de adyacencia es cuadrada y simétrica, la matriz de incidencia de un digrafo no es cuadrada, en general, tampoco es simétrica y ni siquiera sus entradas son positivas. La matriz de incidencia tiene por filas cada uno de los  $n$  nodos y por columnas cada una de las  $l$  aristas.

Si tenemos un grafo no dirigido también se puede establecer una definición de matriz de incidencia, teniendo en cuenta que podemos tomar una arista no dirigida entre el vértice  $v_i$  y  $v_j$  como dos aristas dirigidas, de  $v_i \rightarrow v_j$  y de  $v_j \rightarrow v_i$ .

La matriz de incidencia de un digrafo cumple la siguiente propiedad que no vamos a demostrar.

**Teorema 5.3.1** *Si  $B$  es la matriz de incidencia de un grafo dirigido  $G$  con  $n$  vértices, el rango de  $B$  es  $n - 1$ .*

La demostración de esta propiedad se basa en el hecho de que la suma de todas las columnas de la matriz suman cero y que el grafo no puede tener varias componentes puesto que es conexo. De hecho, si el garfo no fuera conex y tuviera  $k$  componentes, entonces el rango sería  $n - k$ .

Vamos a estudiar una matriz relacionada con la matriz de incidencia de un grafo y que tiene su importancia en la teoría general de redes, especialmente en el cálculo de las centralidades de los nodos. Dicha matriz recibe el nombre de **matriz de incidencia reducida**.

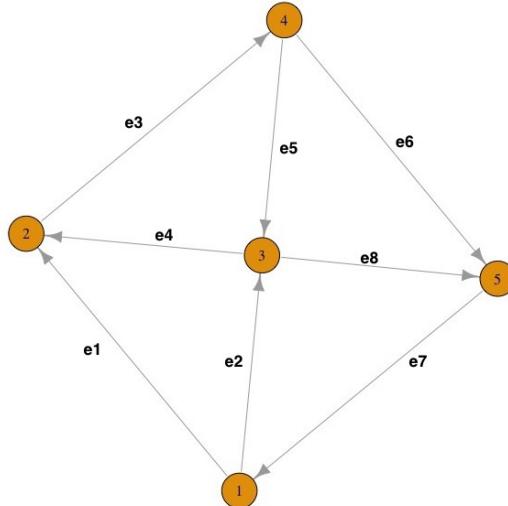
Si consideramos un grafo  $G$  conectado con  $n$  vértices y  $l$  aristas, es evidente que el orden de la matriz de incidencia  $B$  es de  $n \times m$ . Sabemos que el rango de  $B$  es  $n - 1$ . Si nosotros eliminamos cualquier fila de  $B$ , la submatriz resultante será de tamaño  $(n - 1) \times l$ . Sin embargo, su rango seguirá siendo  $n - 1$ . Esto se debe a que hay  $n - 1$  vectores linealmente independientes. Esto demuestra un hecho muy curioso con la matriz de incidencia y es que solamente son necesarias  $n - 1$  filas de vectores de la matriz para especificar el grafo correspondiente, ya que la matriz formada por una fila menos nos proporciona la misma información que la total. Esto nos da pie a las siguientes definiciones.

**Definición 5.3.3** Esta matriz de tamaño  $(n - 1) \times l$ , que resulta cuando eliminamos una fila cualquiera de la matriz de incidencia  $B$ , se llama **matriz de incidencia reducida** y la podemos denotar por  $B_f$ . El vértice correspondiente a la fila eliminada se llama **vértice referencia**.

Veamos este hecho tan notable con un ejemplo.

---

**Ejemplo 5.3.1** Consideremos el grafo que se muestra en la [figura 5.13](#).



**Figura 5.12** Grafo de ejemplo..

Dicho grafo dirigido consta de 5 vértices y de 8 aristas. En consecuencia, su matriz de incidencia tiene 5 filas correspondientes a cada uno de los nodos y 8 columnas para las distintas aristas, en orden.

Su matriz de incidencia es

$$B = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & -1 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & -1 \end{bmatrix}$$

En primer lugar comprobamos que el rango de esta matriz es  $n - 1 = 4$ . Si calculamos su matriz escalonada para ver el número de unos principales de la matriz podremos fácilmente determinar su rango. Efectuando las operaciones elementales sobre  $B$ , tendremos que la escalonada es:

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

Notamos que hemos hecho la última fila de ceros, por lo que tenemos cuatro unos principales, lo que significa que el rango de  $B$  es 4, es decir,  $n - 1$ . Así, si elimináramos la

fila 5, la matriz de incidencia reducida sería

$$B_f = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 1 & 1 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 & 1 \end{bmatrix}.$$

Lo más importante de todo es que la matriz  $B_f$  me proporciona las características del grafo.

**Definición 5.3.4 (Matriz de Incidencia de un grafo)** Dado un grafo  $G$  no dirigido con  $n$  vértices, definimos la matriz de incidencia  $\tilde{B} = (\tilde{b}_{ij})_{i,j=1}^{n,l}$  como la matriz donde

$$\tilde{b}_{ij} = \begin{cases} 1 & \text{si la arista } j \text{ conecta con el vértice } i \\ 0 & \text{en otro caso.} \end{cases}$$

La matriz de incidencia de un grafo no dirigido tiene las siguientes propiedades:

- Cada columna de  $B$  tiene dos unos, ya que cada arista incide sobre dos vértices.
- El número de unos en cada fila es el grado del vértice correspondiente.
- Una fila de ceros representa que ese vértice está aislado, por lo que el grafo no está conectado.
- Dos aristas conectando los mismos vértices producen dos columnas iguales en la matriz.
- Permutar dos filas o dos columnas de la matriz de incidencia significa renombrar los vértices o las aristas del grafo.

Antes de definir la matriz Laplaciana de una grafo, debemos introducir la **matriz de grados** del grafo. Sabemos que el grado de un nodo es el número de aristas que inciden en él. En consecuencia, la suma de las componentes de la fila  $i$  de la matriz de adyacencia coincide con el grado  $d_i$  del correspondiente nodo  $v_i$ .

**Definición 5.3.5 (Matriz de nodos de un grafo)** Dado un grafo  $G$  no dirigido con  $n$  vértices, la matriz de nodos  $\Delta$  es una matriz diagonal donde sus elementos son los grados de los nodos  $d_i$ :

$$\Delta = \begin{bmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{bmatrix},$$

donde  $d_i = \sum_{j=1}^n a_{ij}$ .

En el ejemplo anterior necesitamos construir la llamada matriz estocástica del grafo. Definimos formalmente esta matriz a partir de la propiedad siguiente.

**Teorema 5.3.2** Consideremos un grafo  $G$  no dirigido y su matriz de adyacencia  $A$ . Entonces, la suma de todas las componentes de la matriz de adyacencia suma el doble del número de aristas, es decir,

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} = \sum_{i=1}^n d_i = 2l.$$

Definimos la **matriz estocástica** asociada a un grafo.

**Definición 5.3.6 (Matriz estocástica de un grafo)** Dado un grafo  $G$  no dirigido con  $n$  vértices, definimos la matriz estocástica  $P$  como la matriz  $P = A\Delta^{-1}$ .

Definida de esta forma, todas las componentes de la matriz estocástica  $P$  se encuentran entre 0 y 1, así como la suma, por columnas, de sus componentes es 1. Como ya vimos, las componentes de  $P$  podían interpretarse en términos probabilísticos como la probabilidad de pasar de un nodo a otro a través de una arista.

La siguiente propiedad nos dará la posibilidad de definir la matriz Laplaciana y relaciona varias de las matrices definidas hasta ahora, como son la matriz de adyacencia, incidencia y de grados.

**Teorema 5.3.3** Consideremos un grafo  $G$  no dirigido y su matriz de adyacencia  $A$ , de incidencia  $B$  y de grados  $\Delta$ . Entonces,

$$BB^T = \Delta - A.$$

Si  $G$  es un digrafo, entonces

$$\tilde{B}\tilde{B}^T = \Delta + A.$$

Así, definimos la **matriz Laplaciana** de un grafo.

**Definición 5.3.7 (Matriz Laplaciana de un grafo)** Dado un grafo  $G$  no dirigido con  $n$  vértices, definimos la matriz Laplaciana  $L$  como la matriz

$$L = BB^T = \Delta - A.$$

Algunas de las propiedades más importantes de estas matrices se resumen en los siguientes puntos:

- Por definición de matriz de adyacencia, la suma de la fila  $i$ -ésima representa el grado del nodo  $v_i$ .

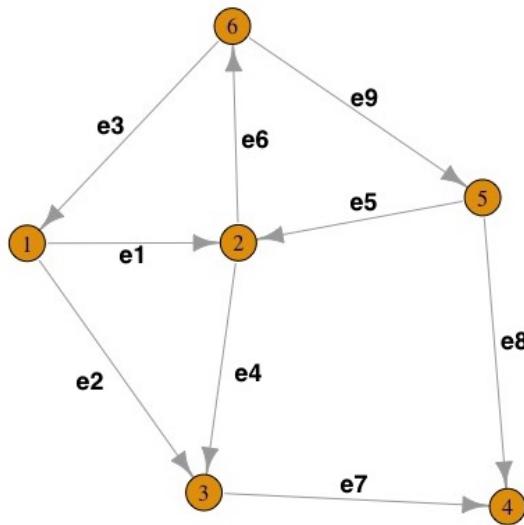
- Cada fila de la matriz Laplaciana  $L$  suma 0, es decir,

$$\sum_{k=1}^n q_{ik} = 0.$$

La consecuencia inmediata de esta propiedad es que el determinante de la matriz Laplaciana es 0, es decir,  $\det(L) = 0$ .

- La matriz  $P = A\Delta^{-1}$  es estocástica por columnas, lo que significa que la suma de los elementos por columnas es 1 y todos sus elementos son positivos entre  $[0,1]$ .
- Si el grafo  $G$  está conectado, entonces el rango de  $B$  es  $n - 1$ . Si el grafo tiene  $k$  componentes conexas, entonces el rango de  $B$  es  $n - k$ .

Vemos a continuación un ejemplo de un grafo dirigido con 6 nodos y las matrices asociadas al mismo (véase la [figura 5.13](#)).



**Figura 5.13** Grafo de ejemplo compuesto por 6 nodos y 9 aristas..

Dado el grafo que aparece en la figura [figura 5.13](#), calculamos su matriz de adyacencia y de incidencia asociadas al mismo.

**Matriz de adyacencia.**

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

**Matriz de incidencia.**

$$B = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & -1 \end{bmatrix}$$

## 5.4 Propiedades espectrales de $A$ y $L$

La teoría espectral de grafos está relacionada con las propiedades espectrales de la matriz de adyacencia y Laplaciana de un grafo. Antes de comenzar el estudio de cómo son los valores y vectores propios de la matriz de adyacencia y Laplaciana de un grafo, recordamos algunas propiedades básicas sobre los valores propios.

- Un vector propio de la matriz  $A$  es un vector tal que  $A\mathbf{x}$  es paralelo a  $\mathbf{x}$ .
- La traza de  $A$  es la suma de los valores propios de  $A$ .
- Si la matriz  $A$  es simétrica, todos sus valores propios son reales. Además, existe una base ortonormal de vectores formada por los vectores propios, de manera que sus correspondientes valores propios son las raíces del polinomio característico.
- Si una matriz  $A$  cuadrada es semidefinida positiva, todos sus valores propios son nonnegativos.

En el capítulo anterior hemos visto distintas matrices asociadas a un grafo. Estas son:

**Matriz de Adyacencia** Esta matriz está asociada a la conectividad de los nodos y nos proporciona las características topológicas del grafo. La denotamos por  $A$ .

**Matriz de incidencia** Esta matriz está asociada a las aristas del grafo y la denotamos por  $B$ .

**Matriz de nodos** Es una matriz diagonal donde se recogen los grados de los nodos. La denotamos por  $\Delta$ .

**Matriz estocástica** La definimos como  $P = A\Delta^{-1}$ .

**Matriz Laplaciana** La matriz Laplaciana la denotamos por  $L$  y la definimos como  $L = \Delta - A$ .

La teoría espectral de grafos está relacionada con las propiedades espectrales de la matriz Laplaciana. Debemos continuar recordando algunas de las propiedades más interesantes de la matriz Laplaciana, como son:

- $L$  es siempre semidefinida positiva.
- El número de veces que aparece un 0 como autovalor, es el número de componentes conexas de un grafo.

El siguiente resultado es fundamental en la teoría espectral y constituye la base teórica de un buen número de aplicaciones informáticas actuales, como por ejemplo el buscador Google, cuya tarea esencial consiste en realizar una ordenación de las páginas web existentes.

**Teorema 5.4.1 (Teorema de Perron-Frobenius)** *Sea  $A$  una matriz cuadrada de tamaño  $n$  con todas sus entradas no negativas. Entonces tiene un valor propio real no negativo  $\lambda$  que es el máximo en valor absoluto de todos los valores propios. Este valor propio  $\lambda$  tiene un vector propio real no negativo.*

*Si, además, la matriz no admite una descomposición triangular por bloques, entonces  $\lambda$  tiene multiplicidad 1 y el vector propio correspondiente es positivo.*

Las matrices de adyacencia y la Laplaciana de un grafo son **matrices simétricas**, lo que significa que sus valores propios son reales. Sean estos valores propios

$$\begin{array}{ll} A_G & \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \\ L_G & \mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \end{array}$$

El teorema de Perron-Frobenius implica necesariamente que si  $G$  es conexo, entonces el mayor valor propio  $\lambda_{\max}$  de la matriz de adyacencia tiene multiplicidad 1.

Notemos que si consideramos la matriz de transición de probabilidades que habíamos definido anteriormente, entonces podemos afirmar que el valor propio más grande es  $\lambda = 1$ . Esto es, si consideramos la matriz Laplaciana  $L$ , ocurre con el valor propio más pequeño, que es  $\lambda = 0$ , ya que la Laplaciana es singular y semidefinida positiva.

Esta característica nos permite enunciar una propiedad muy interesante en los grafos.

**Teorema 5.4.2** *Dado un grafo general  $G$ , consideramos su matriz asociada Laplaciana  $L$ . Dentro del espectro de  $L$  está el valor propio  $\lambda = 0$ . Entonces, la multiplicidad del valor propio 0 de la Laplaciana es igual al número de componentes conexas del grafo.*

Desde el inicio de la teoría espectral, el estudio de los valores y vectores propios de la matriz Laplaciana ha tenido un buen número de implicaciones y se han desarrollado aplicaciones en química. Los valores propios se asociaron a la estabilidad de las moléculas. también se ha utilizado en varios problemas de física teórica y mecánica cuántica. Podríamos citar un gran número de áreas en las que resulta muy útil conocer el espectro de una matriz.

## 5.5 Vector de Fiedler de la Laplaciana

El concepto de vector de Fiedler está asociado con la idea de conectividad algebraica de un grafo y tiene una interesante aplicación sobre los grafos, como discutiremos a continuación. Sabemos que la Laplaciana de un grafo siempre tiene un valor propio nulo y que la multiplicidad algebraica del valor nulo coincide con el número de componentes conexas del grafo. Si el grafo es conexo, podemos asegurar que el segundo valor propio más pequeño es mayor que cero.

Esta propiedad es la que nos va a permitir determinar estructuras clave en el grafo. El espectro de la Laplaciana motiva la definición de **conectividad algebraica** de una red, que se asigna al valor  $\lambda_{n-1}$ .

El vector propio  $\lambda_{n-1}$  tiene una propiedad muy interesante.

**Teorema 5.5.1** *Supongamos que  $G$  es un grafo conexo con Laplaciana  $L$ , cuyo segundo valor propio más pequeño es  $\lambda_{n-1}$ . Sea  $\mathbf{x}$  el vector propio asociado al valor propio  $\lambda_{n-1}$ . Sea  $r \in \mathbb{R}$  y hagamos una partición de  $V$  en los dos conjuntos siguientes*

$$V_1 = \{i \in V \text{ tal que } x_i \geq r\}, V_2 = \{i \in V \text{ tal que } x_i < r\}.$$

*Entonces los subgrafos de  $G$  inducidos por los conjuntos  $V_1$  y  $V_2$  son conexos.*

Como consecuencia del resultado anterior, se puede definir un método para obtener una partición de un grafo en dos partes, de forma que ambas partes sean conexas. Cada una de estas particiones puede, a su vez, ser dividida en otras dos particiones conexas y así sucesivamente. De esta forma, es posible conseguir clusters en la red. Eligiendo  $r$  de manera que sea la media de las componentes de  $\mathbf{x}$ , conseguimos dos partes de igual tamaño.

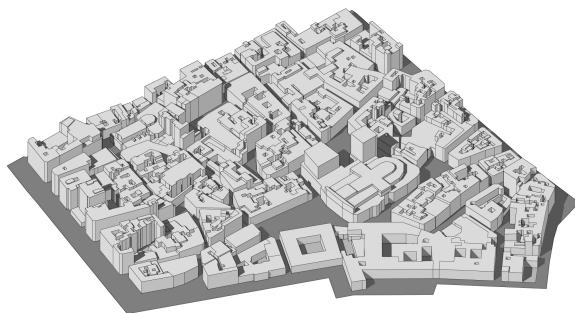
Otra de las elecciones más comunes para  $r$  es  $r = 0$ . Esta forma de dividir un grafo en partes se llama *clustering espectral*. El vector  $\mathbf{x}$  asociado al valor propio  $\lambda_{n-1}$  recibe el nombre de **vector de Fiedler**.

## 5.6 Las ciudades representadas como grafos

Existen numerosas páginas web o portales relacionados con la geografía y los datos de las ciudades. Existe un portal de datos especialmente interesante que se llama **City-Data** (<http://www.city-data.com/>), que recoge una cantidad enorme de datos sobre las ciudades de Estados Unidos y Canadá. Podemos obtener de cualquier ciudad americana fotografías, datos meteorológicos, datos diversos relacionados con la población y la procedencia étnica de sus habitantes, datos sobre la estructura geográfica de la ciudad, por ZIPs, datos económicos sobre la vivienda, el precio de las casas y los apartamentos. También disponemos de mapas de las ciudades con sus condados y sus alrededores. Además de proporcionarnos datos de las ciudades, obtenemos información sobre educación (colegios, tanto infantiles como High School).

Como complemento a toda esta información, disponemos de una base de datos extraordinaria sobre criminalidad en cualquiera de las ciudades americanas en las que hay registros. Así, podemos comprobar que en el año 2016 se produjeron en la ciudad de Anchorage (Alaska) un total de 28 asesinatos, 2 más que en el año 2015. De acuerdo con la población de Anchorage, este dato representa un total de 9.4 asesinatos por cada 100,000 habitantes. También se recoge un dato bastante interesante como es un índice global de criminalidad. Podemos ver que el índice global de Anchorage es de 621, siendo la media de todo el país 280.5.

La [figura 5.14](#) representa una parte del centro histórico de la ciudad de Murcia.



**Figura 5.14** Parte del centro histórico de la ciudad de Murcia..

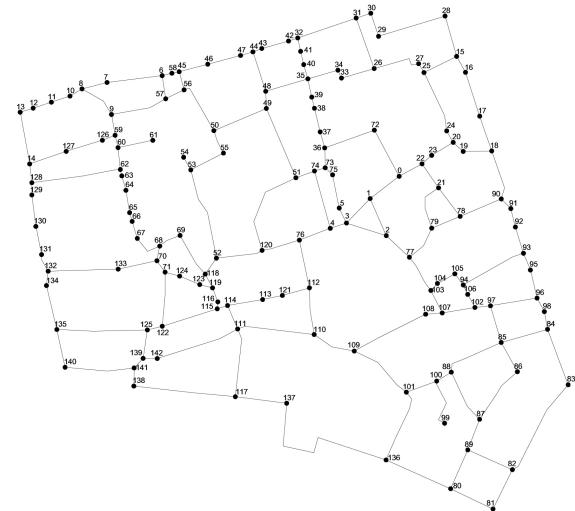
Nos planteamos representar matemáticamente esta estructura de calles y espacio público mediante alguna estructura que nos permita su análisis y visualización de una forma sencilla. Dicha estructura puede ser un grafo, es decir un conjunto de nodos y aristas. Ahora bien, debemos determinar con precisión qué van a ser los nodos y las aristas del mismo.

En una primera aproximación, podemos pensar en una estructura de nodos y aristas

donde los nodos son las intersecciones de las calles y las aristas son las propias calles del trazado urbano. Un grafo de este tipo es lo que llamamos un **primal graph** o *grafo primario*. El grafo primario se utiliza en diversos tipos de redes espaciales y el objetivo fundamental que pretende es describir la topología del trazado urbano. Este modelo se utiliza actualmente y es perfectamente válido para representar la topología de las ciudades. Además, es muy flexible y simple para definir la topología urbana, así como para descubrir las propiedades topológicas y de centralidad de la ciudad.

Si construimos un grafo primario que represente el trazado urbano que se muestra en la figura 5.14, obtenemos un grafo como el que vemos en la figura 5.15.

La figura 5.14 representa una parte del centro histórico de la ciudad de Murcia.



**Figura 5.15** Grafo primario del centro histórico de la ciudad de Murcia..

Notemos que en el grafo que hemos dibujado en la figura 5.15 aparecen algunos nodos en aristas que no se corresponden con intersecciones de calles. A veces, según las características del trazado urbano que estamos estudiando, podemos introducir nodos auxiliares cuando existen aristas muy largas.

En la figura figura 5.16 resumimos gráficamente el proceso de construcción del grafo primario a partir del trazado urbano objeto de estudio. Vemos perfectamente cómo vamos construyendo los nodos a partir de las diversas intersecciones, aunque a menudo resulta complicado por el propio trazado de los centros históricos de la mayoría de las ciudades españolas, con calles no lineales y plazas y espacios públicos de formas geométricas complejas.



**Figura 5.16** Grafo primario de otra parte del centro histórico de la ciudad de Murcia..

Cuando nos fijamos en el grafo de la figura 5.15 notamos una característica de este grafo que lo convierte en muy particular. Si calculamos los grados de cada uno de los nodos, nos damos cuenta que la mayoría de ellos tiene un grado comprendido entre dos y cinco. esto supone una distribución de grados muy uniforme y en la que apenas existen diferencias significativas. Esta característica de este tipo de grafos puede llevarnos a resultados no muy precisos cuando intentemos analizar el grafo desde el punto de vista de su centralidad, por ejemplo.

Si pensamos en otro tipo de redes, como por ejemplo las redes sociales, esta característica desaparece. Más bien obtenemos el efecto contrario, puede existir una gran disparidad en los valores del grado de los vértices. Pensemos en personas famosas o conocidas, cuyos contactos o seguidores en las redes sociales se miden por millones, mientras que una gran parte de la población *anónima* se puede mover en niveles de decenas o centenas. Esto provoca un rango en los grados de los nodos que le proporciona al grafo una mayor riqueza, desde el punto de vista de su distribución estadística del grado.

Sucede exactamente lo mismo si consideramos redes urbanas más extensas. Podemos apreciar estas características en la imagen que aparece en la figura 5.17, en la que ahora hemos extendido la red urbana al conjunto del centro de Murcia. Hemos construido el grafo primario correspondiente a esta red; dicho grafo consta de prácticamente 2000 nodos.

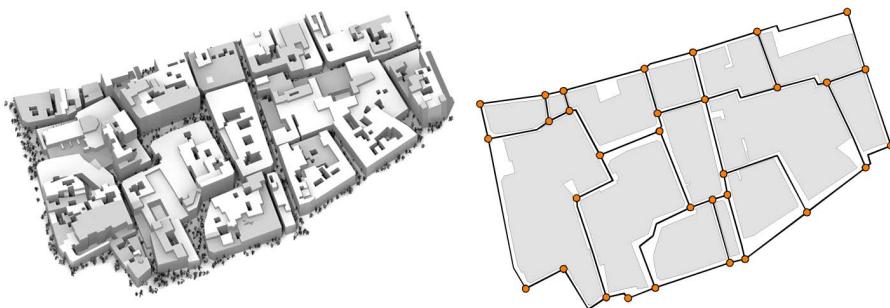


**Figura 5.17** Grafo primario del centro histórico de la ciudad de Murcia..

## 5.7 El grafo dual de un grafo $G$

Pensemos en una red urbana de calles como las que hemos visto en la sección anterior. Cuando nos planteamos representar un trazado urbano mediante una estructura geométrica, inmediatamente pensamos en los grafos como una estructura adecuada para expresar la topología de ese trazado.

Sin embargo, cuando decidimos representar un grafo con nodos y aristas a partir del mapa de una ciudad disponemos de varias posibilidades. La primera consiste en representar el grafo primario. Véase la figura 5.18 para ver un ejemplo.



**Figura 5.18** Grafo primario que representa un trozo de una ciudad..

El grafo que vemos en la figura [figura 5.18](#) marca nodos en las intersecciones y aristas siguiendo el trazado de las calles. Un grafo de este tipo ya sabemos que recibe el nombre de **grafo primario**.

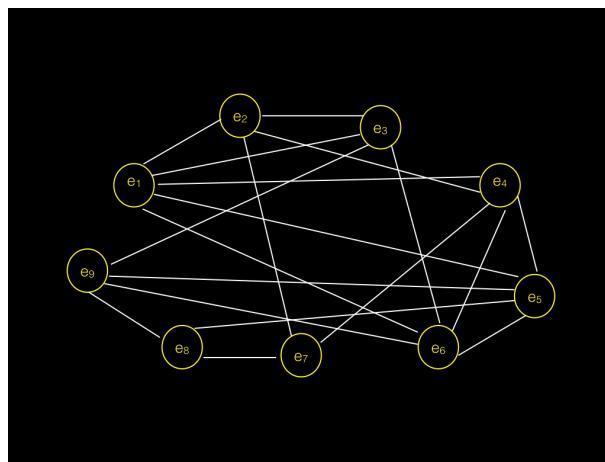
Sin embargo, podemos pensar en otro tipo de grafo donde cada calle represente un nodo del grafo y dibujemos una arista entre los nodos cuando las calles intersecten entre sí. Este concepto de representación gráfica de la ciudad es completamente distinto y su grafo asociado recibe el nombre de **grafo dual** (*dual graph*).

Podemos formalizar el concepto mediante la siguiente definición.

**Definición 5.7.1 (Grafo dual de un grafo  $G$ ) .**

Dado un grafo  $G(V(G),E(G))$  con  $n$  nodos o vértices y  $l$  es el número de aristas, definimos su grafo lineal o dual  $G_L(l,K)$  como el grafo formado por el conjunto de aristas  $l$  como nodos y con  $k$  aristas de forma que existe una arista en  $K$  si y solo si las dos aristas que une tienen un nodo común en  $G$ .

Según la definición de grafo dual, dos nodos son adyacentes en  $G_L$  (grafo dual) si y solo si hay exactamente un nodo de  $G$  en común. Siguiendo con el ejemplo anterior, construimos el grafo dual del grafo de la figura [figura 5.17](#), que es el que aparece en la figura [figura 5.19](#).



**Figura 5.19** Grafo dual del grafo de la figura [figura 5.17](#)..

Observemos el grafo dual dibujado en la figura [figura 5.19](#) para determinar las diferencias entre ambos grafos y ver el proceso de construcción, completamente distinto al del grafo primario. Ahora en el grafo dual aparecen tantos vértices como aristas teníamos en el grafo original, es decir, 9 nodos. Cada uno de estos nodos representa una arista del grafo de la figura [figura 5.17](#). Sabemos que dos nodos son adyacentes en  $G_L$  si tienen un nodo común de  $G$ . Comenzamos preguntándonos si entre los nodos  $e_1$  y  $e_2$  debemos dibujar una arista. Para determinar esta cuestión analizamos el grafo original y vemos que ambas aristas tienen un nodo común, el 1, por lo que dibujamos la arista correspondiente. Observando el grafo

original, vemos que la arista  $e_1$  está compartiendo un nodo con las aristas  $e_2, e_3, e_4, e_5, e_6$ , por lo que procedemos a dibujar aristas entre  $e_1$  y todos estos nodos en el grafo dual.

Seguiríamos un proceso análogo al descrito anteriormente para el resto de vértices, dibujando al final el grafo que aparece en la figura [figura 5.19](#) y que constituye el grafo dual del grafo original. Notemos que mientras que el grafo original tiene  $n = 6$  nodos y  $l = 9$  aristas, el nuevo grafo tiene 9 nodos y 19 aristas.

Del mismo modo que de un grafo obtenemos su matriz de adyacencia, podemos determinar la matriz de adyacencia del grafo dual, que nos proporciona la conectividad del mismo. Denotamos a la matriz de adyacencia del grafo dual como  $A_{G_L}$ .

En nuestro ejemplo,

$$A_{G_L} = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Ahora nos preguntamos si existe una relación entre los grafos  $G$  y  $G_L$  de forma que podamos obtener, de una manera sencilla, la matriz de adyacencia del grafo dual, es decir,  $A_{G_L}$  a partir de las matrices asociadas al grafo  $G$ .

Supongamos que tenemos el grafo  $G$  compuesto de  $n$  vértices y  $l$  aristas. Asociado a este grafo hemos visto que tenemos

- La matriz de adyacencia  $A$ .
- La matriz de incidencia,  $B$ .
- La matriz Laplaciana,  $L$ .

Definimos una matriz  $R$  como

$$R = |B|$$

es decir, la matriz de incidencia en módulo. Con esta nueva matriz, tenemos el siguiente resultado.

**Teorema 5.7.1** Consideremos un grafo  $G$  y su matriz de adyacencia  $A$ , de incidencia  $B$  y Laplaciana  $L$ . Definimos una matriz  $R$  como  $R = |B|$ . Entonces, la matriz de adyacencia  $A$  del grafo  $G$  puede escribirse

$$A = RR^T - \Delta.$$

La matriz de adyacencia del grafo dual  $A_{G_L}$  puede obtenerse mediante la expresión

$$A_{G_L} = R^T R - 2 \cdot I_l.$$

El siguiente resultado resume las caracetrísticas esenciales del grafo dual, caracterizando el número de nodos y de aristas del mismo.

**Teorema 5.7.2** Consideremos un grafo  $G$  y su grafo dual  $A_{G_L}$ . Entonces

- El número de nodos de  $A_{G_L}$  es  $l$  (número de aristas de  $G$ ).
- El número de aristas  $K$  viene dado por la expresión

$$K_{G_L} = \frac{1}{2} \sum_{i=1}^n d_i^2 - l.$$

1. Jérôme Kunegis. KONECT – The Koblenz Network Collection. In Proc. Int. Conf. on World Wide Web Companion, pages 1343–1350, 2013.
2. Hayes, Brian. 2006. "Connecting the dots." American Scientist 94 (5):400-404.



# 6 Breve introducción a la teoría de redes

---

## 6.1 Introducción

Comenzamos analizando las diferencias que existen entre las redes y los grafos. En muchas ocasiones tendemos a pensar que es lo mismo o hablamos de redes y grafos indistintamente. En la actualidad las redes complejas se estudian por su relación con muchos campos de la ciencia. Sin duda, muchos sistemas en la naturaleza se pueden describir por medio de redes complejas, que topológicamente son grafos a los que se agregan algunas características dinámicas que hacen necesario estudiarlos desde varios puntos de vista.

Los grafos descritos en la primera sección del capítulo anterior constituyen ejemplos claros de redes, ya que los nodos y las aristas de los mismos tenían un significado, representaban alguna idea o propiedad asociada a los nodos.

Los ejemplos de la presencia de las redes complejas en la vida real son numerosos; por ejemplo, Internet es una red de routers o dominios conectados por líneas físicas, la World Wide Web es una red de sitios web conectados por hiperenlaces, el cerebro es una red de neuronas conectados por medio de las sinapsis, una organización es una red de personas con diversos tipos de conexiones entre ellas, la economía mundial es una red formada por las economías nacionales, que a su vez son redes de mercados, y éstos son redes de productores y consumidores que interactúan, las redes alimentarias y las rutas metabólicas pueden ser representadas por redes, así como las relaciones (semánticas o sintácticas) entre las palabras de un idioma, los temas en una conversación, e incluso las estrategias para resolver un problema matemático o las redes culturales en las que se mueven los objetos e ideas generadas por el hombre.

Hasta aquí nos quedaríamos con la representación de estos sistemas por medio de la estructura de grafo matemático, pero si consideramos por ejemplo las enfermedades que se transmiten a través de redes sociales, los virus informáticos que en ocasiones se extienden a través de la Internet, o la energía que se distribuye a través de las redes de transporte, nos damos cuenta de que, junto a esa estructura organizativa en forma de grafo es importante

conocer la dinámica que se produce en el flujo de información a través de ella, o incluso la dinámica que se puede producir en la evolución temporal de dicha estructura, que a menudo no es fija ni en el conjunto de nodos que interviene ni en las conexiones que se producen entre ellos.

### Los seis grados de separación.

Seguro que en alguna ocasión nos hemos encontrado algún conocido estando de viaje o a alguien a quien no veíamos en años en el supermercado. O quizás hemos coincidido en el trabajo, en el cine, en un restaurante con amigos de algún amigo. Inmediatamente hemos pensado ¡qué coincidencia! Realmente no es tan coincidencia como nos pensamos. Existe una teoría muy interesante avalada por experimentos llevados a cabo de forma muy precisa que pretende dar respuesta al número de conexiones entre las personas.

Se le llama **seis grados de separación** a la hipótesis que intenta probar que cualquiera en la Tierra puede estar conectado a cualquier otra persona del planeta a través de una cadena de conocidos que no tiene más de cinco intermediarios (conectando a ambas personas con sólo seis enlaces). La teoría fue inicialmente propuesta en 1930 por el escritor húngaro Frigyes Karinthy en un cuento. Recogida también en el libro *Six Degrees: The Science of a Connected Age* del sociólogo Duncan Watts, y que asegura que es posible acceder a cualquier persona del planeta en tan sólo seis *saltos*.

En 1967, el psicólogo estadounidense Stanley Milgram ideó una nueva manera de probar la teoría, que denominó *el problema del pequeño mundo*. El experimento del mundo pequeño de Milgram consistió en la selección al azar de varias personas del medio oeste estadounidense, para que enviaran tarjetas postales a un extraño situado en Massachusetts, situado a varios miles de millas de distancia. Los remitentes conocían el nombre del destinatario, su ocupación y la localización aproximada. Se les indicó que enviaran el paquete a una persona que ellos conocieran directamente y que pensaran que fuera la que más probabilidades tendría, de todos sus amigos, de conocer directamente al destinatario. Esta persona tendría que hacer lo mismo y así sucesivamente hasta que el paquete fuera entregado personalmente a su destinatario final. Aunque los participantes esperaban que la cadena incluyera al menos cientos de intermediarios, la entrega de cada paquete solamente llevó, como promedio, entre cinco y siete intermediarios. Los descubrimientos de Milgram fueron publicados en *Psychology Today* e inspiraron la frase **seis grados de separación**.

En 2011 la empresa Facebook realizó un estudio con todos los usuarios activos de su página en ese momento 721.000.000 miembros (alrededor del 10% de la población mundial) y se analizó el conjunto de amigos en común, para sacar el promedio de cuántos eslabones hay entre cualquier usuario y otro cualquiera. De esta prueba se excluyó a celebridades y famosos. Los resultados mostraron que el 99,6% de pares de usuarios estuvieron conectados por 5 grados de separación. Esta es la prueba más cercana de la teoría a la fecha de hoy y da un resultado aproximado de 4,75 eslabones.

Así, cuando hablamos de redes nos referimos a una amplia variedad de estructuras cuyas entidades están representadas por los nodos de la red y las relaciones o interacciones entre esos elementos se representan en la red mediante enlaces o aristas entre entidades. En consecuencia, las aristas admiten muy diversas representaciones. En el esquema siguiente tenemos algunas redes físicas o virtuales y cómo pueden ser representadas mediante grafos, especificando sus nodos y aristas.

Red compleja	Nodos	Aristas
Bioquímica	Moléculas	Reacciones químicas
Neuronales	Neuronas	Conexiones sinápticas
Epidemiológicas	Individuos infectados	Contactos infecciosos
world wide web	páginas web	Hiperenlaces
Redes de potencia	Generadores, usuarios	Líneas de potencia, transformadores
Colaborativas	Científicos, ingenieros	Colaboraciones entre individuos
Sociales	Gente, individuos	Amistad, relación social
internet	Ordenadores	Routers, cables ethernet

Estos son solo algunos ejemplos de los muchos que existen puesto que vivimos en un **mundo conectado**, donde las redes son protagonistas en nuestra vida cotidiana y no nos referimos únicamente a las redes sociales.

En el capítulo anterior estudiamos algunos grafos en la sección inicial que en realidad representaban grafos asociados a redes. Concretamente hablamos de una red de tráfico aéreo con conexiones entre distintos aeropuertos, el grafo de un equipo de fútbol donde las conexiones estaban relacionadas con las asociaciones entre los distintos futbolistas, así como una red de caminos y carreteras donde las conexiones eran físicas entre los nodos.

A finales de 1950, dos matemáticos, Erdős y Rényi (ER), hicieron un gran avance en la teoría matemática clásica de grafos que revolucionaría la forma en que se pueden modelar estos problemas describiendo una red con topología compleja por medio de un grafo aleatorio, estableciendo los fundamentos de la teoría de redes aleatorias. Aunque la intuición indica claramente que muchas redes complejas de la vida real no son ni totalmente regulares ni completamente aleatorias, el modelo ER fue el único enfoque sensato y riguroso que dominó el pensamiento de los científicos acerca de las redes complejas durante la segunda mitad del siglo XX.

Durante las últimas décadas y de forma paralela al auge de *la ciencia de los datos*, se ha desarrollado enormemente lo que hoy día conocemos como ciencia de las redes o *nework science*. En 1998, con el fin de describir la transición de una red regular en una red aleatoria, Watts y Strogatz (WS) introdujeron el concepto de red de mundo pequeño. Debe notarse que el fenómeno de mundo pequeño es de hecho muy común, y nada alejado de nuestras experiencias diarias. A menudo, poco después de conocer a un extraño, uno se sorprende al descubrir que tenemos un amigo común con él, de modo que es habitual la expresión: *¡Qué pequeño es el mundo!*

## 6.2 Algunos conceptos y medidas básicas en redes

Hemos visto entonces que existen redes de muy diferentes tipos, características y significados, lo que nos lleva al problema de intentar clasificarlas o establecer semejanzas y

diferencias entre ellas, creando modelos o tipos que tengan unas características comunes.

Ha de indicarse que, debido a la alta complejidad de las redes reales, no se han encontrado todavía caracterizaciones completas de las mismas, es decir, en la actualidad todavía no disponemos de un conjunto de medibles que caractericen por completo a cada red (algo así como un código genético que nos permita con absoluta precisión establecer comparaciones entre ellas).

Sin embargo, aunque no dispongamos de medidas absolutas que caracterizan las redes, es necesario establecer una serie de medidas o conceptos que cumplan un papel importante en su estudio. Dichos conceptos y medidas cuantitativas básicas en las redes son:

- Longitud media de los caminos.
- Coeficiente de comunidades.
- Distribución de grado.

### 6.2.1 Longitud media de caminos en una red

El tema de la longitud de caminos en la red es similar al del estudio de los caminos de mínima distancia entre los nodos de un grafo. Así, sabemos por la teoría de grafos que la distancia  $d_{ij}$  entre dos nodos  $i$  y  $j$  se define como el número de enlaces del camino más corto que los conecta.

El diámetro de una red es la máxima distancia entre cualesquiera par de nodos dentro de la red.

Por último, la longitud promedio  $L$  de los caminos de una red la podemos definir como la media de las distancias entre todos los pares de nodos.

A medida que se tenían más datos de redes complejas reales se constató que la longitud promedio de los caminos de la mayoría de ellas era relativamente pequeña, incluso en los casos en que estos tipos de redes tuvieran muchos menos enlaces de los posibles que se podrían dar (a menor cantidad de conexiones en el mundo, parece claro que haya que hacer recorridos más largos para poder llegar de un nodo a otro). Esta característica es lo que se llamó efecto de mundo pequeño, y de ahí el nombre de redes de mundo pequeño que intentan modelarlo. Vuelve a aparecer la idea de un mundo conectado por un número pequeño de enlaces.

### 6.2.2 Coeficiente de comunidades o agrupamiento

En el estudio de redes complejas, se dice que una red tiene una estructura de comunidad si los nodos de la red se pueden agrupar fácilmente en conjuntos de nodos (potencialmente superpuestos) de modo que cada conjunto de nodos esté conectado densamente internamente. Si los conjuntos de nodos muy conectados entre sí son no superpuestos, esto implica que la red se divide naturalmente en grupos de nodos con conexiones densas internamente y conexiones más dispersas entre grupos. Pero las comunidades superpuestas también están permitidas.

La definición más general se basa en el principio de que los pares de nodos tienen más probabilidades de estar conectados si ambos son miembros de la misma comunidad, siendo menos probable que se conecten si no comparten comunidades. Un problema relacionado

pero diferente es la búsqueda en la comunidad, aquí el objetivo es encontrar una comunidad a la que pertenezca un cierto vértice.

En las redes sociales es muy fácil que se produzca esta estructura de comunidades donde unos individuos se relacionan más frecuentemente entre sí.

El coeficiente de agrupamiento (mencionado en la literatura también como *clustering coefficient*) de un vértice en un grafo o red cuantifica qué tanto está de agrupado (o interconectado) con sus vecinos. Se puede decir que si el vértice está agrupado dentro de un grafo completo su valor es máximo, mientras que un valor pequeño indica un vértice poco agrupado en la red. Duncan J. Watts y Steven Strogatz fueron los primeros en idear este coeficiente, en 1998 para determinar si un grafo es una red de mundo pequeño. Se suele representar formalmente como  $C_i$ .

Recordemos que el grado del nodo  $i$  lo denotamos por  $k_i$ .

Así, el coeficiente de comunidades o agrupamiento  $C_i$  de un nodo  $i$  viene dado por la proporción entre los enlaces o conexiones entre vértices que son sus vecinos dividido entre el número de pares de vecinos de  $i$ .

Lo escribimos más formalmente. Si  $k_i$  es el grado de un vértice  $i$  y  $t_i$  es el número de aristas o conexiones entre sus vecinos, su **coeficiente de agrupamiento o comunidades** viene dado por la expresión

$$C_i = \frac{t_i}{k_i(k_i - 1)/2}.$$

Debemos establecer dos casos particulares:

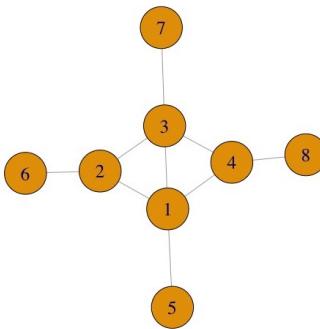
- Cuando  $k_i = 0$ , es decir, tenemos un nodo aislado, entonces su coeficiente de agrupamiento es, evidentemente, igual a 0.
- Cuando  $k_i = 1$ , es decir, tenemos un nodo con grado igual a 1, entonces su coeficiente de agrupamiento es, evidentemente, igual a 1.

En general, el coeficiente de agrupamiento de la red  $C$  será la media de los coeficientes de agrupamiento de todos los nodos de la red. Evidentemente,  $C \leq 1$  y solamente será igual a 1 cuando la red sea completa.

Notemos que hemos tomado la decisión de asignar el coeficiente igual a 1 en los nodos con grado 1; podríamos perfectamente haber tomado la decisión de que cuando el grado del nodo  $i$  sea igual a 1, entonces  $C_i = 0$ . En cualquier caso hablamos de nodos finales en una red, también llamados *pelos de la red*.

---

**Ejemplo 6.2.1** Consideramos el grafo que aparece en la figura 6.1 con 7 nodos y 9 aristas. Queremos calcular el coeficiente de agrupamiento de cada uno de los nodos.



**Figura 6.1** Red de ejemplo..

Si observamos el grafo, antes de comenzar a realizar los cálculos oportunos, nos damos cuenta que los nodos 1 y 3 son equivalentes, ya que ambos forman parte de dos triángulos y tienen grado 4. Así, comenzamos calculando su coeficiente. Notemos que para los nodos 1 y 3, se tiene que  $k_i = 4$ . Entonces,

$$C_1 = C_3 = \frac{t_1}{k_i(k_i - 1)/2} = \frac{2}{6} = \frac{1}{3}.$$

Lo mismo sucede con los vértices 2 y 4, aunque ahora tienen grado 3, es decir,  $k_i = 3$ .

$$C_2 = C_4 = \frac{t_2}{k_i(k_i - 1)/2} = \frac{1}{3}.$$

El resto de vértices que tenemos en el grafo tienen grado igual a 1, por lo que podemos afirmar que:

$$C_5 = C_6 = C_7 = C_8 = 1.$$

Si queremos determinar el coeficiente de agrupamiento de la red, tendremos que hacer

$$C = 1/n \sum_i C_i = \frac{1/3 + 1/3 + 1/3 + 1/3 + 1 + 1 + 1 + 1}{8} = \frac{2}{3}.$$

El objetivo principal del libro *Linked: The New Science of Networks* de Albert-László Barabási es presentar la ciencia de las redes a la audiencia general. Este libro fue escrito por Barabási en 2002 y supuso una revolución en el mundo de la ciencia de las redes, ya que significó el primer intento de acercar esta teoría a una audiencia y a un público no especialista en ciencias.

Proporciona una introducción fácil y legible a los principales modelos y propiedades de las redes y sus aplicaciones en muchas áreas de la vida real, como la propagación de epidemias, la lucha contra el terrorismo, el manejo de crisis económicas o la solución de problemas sociales de la sociedad. Como dice el subtítulo del libro, *Linked revela Cómo todo está conectado a todo lo demás y lo que significa para los negocios, la ciencia y la vida cotidiana.*

### La regla del 80/20.

Se conoce como la regla 80/20 o el principio de Pareto y constituye uno de los capítulos del interesante libro de Barabási. Establece que en muchas situaciones o sucesos, el 80% de los efectos viene de un 20% de causas. Esencialmente, Pareto demostró que aproximadamente el 80% de la tierra en Italia era propiedad del 20% de la población; Pareto desarrolló el principio al observar que alrededor del 20% de las vainas de los guisantes en su jardín contenían el 80% de los guisantes. Existe un buen número de situaciones donde se aplica la regla 80/20 en todos los ámbitos de la sociedad.

En el Informe del Programa de las Naciones Unidas para el Desarrollo de 1992 mostraba que la distribución de los ingresos mundiales es muy desigual. Su conclusión fue categórica e irrefutable: el 20% más rico de la población mundial controla el 82.7% de la riqueza mundial.

En informática, Microsoft se dio cuenta que arreglando el 20% de los bugs más repetidos en su sistema operativo, se solucionaba el 80% de los errores y fallos de un sistema concreto.

El sistema sanitario americano ha estimado que el 20% de los pacientes son los responsables del 80% de los recursos sanitarios.

Los profesionales de la salud y la seguridad en el trabajo usan el principio de Pareto para subrayar la importancia de la priorización del peligro. Suponiendo que el 20% de los peligros representan el 80% de las lesiones y, clasificando los peligros, los profesionales de seguridad pueden enfocarse en el 20% de los peligros que causan el 80% de las lesiones o accidentes.

### 6.2.3 Distribución de grados

La idea que subyace el concepto de distribución de grados de una red es la de encontrar la mejor aproximación para la distribución de probabilidades de los grados de los nodos de una red. El objetivo en esta sección es que nos demos cuenta que redes de tipos distintos presentan distribuciones de grado distintas; por esta razón el estudio estadístico de los grados de los nodos nos proporciona una información muy valiosa sobre la topología de la red. Vamos a ver algunos ejemplos de redes y el análisis de su distribución de grados.

Antes de entrar con los ejemplos, es necesario definir lo que entendemos por distribución de grados en una red o grafo.

**Definición 6.2.1** Dado un grafo  $G$  con  $n$  nodos y  $l$  aristas, la distribución del grado de los nodos se define como

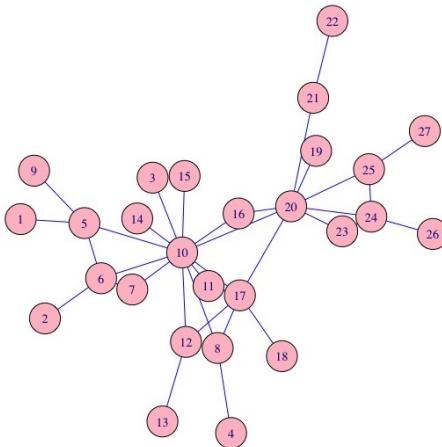
$$P(k) = \frac{n_k}{n}$$

donde  $n_k$  es el número de nodos con grado  $k$  y  $n$  es el número total de nodos de la red.

Podemos representar la distribución de grados dibujando  $P(k)$  frente a  $k$ .

### Ejemplo 1. Redes libres de escala.

Ahora estudiaremos mediante algunos ejemplos sencillos la distribución del grado en las redes. El primer ejemplo que estudiamos es el de la red que aparece en la [figura 6.2](#).

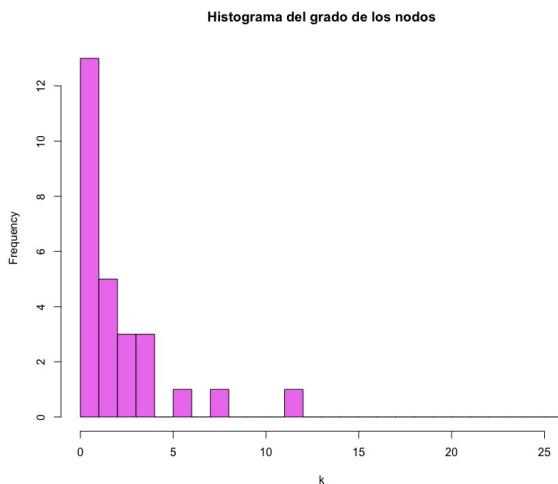


**Figura 6.2** Red de ejemplo..

Se trata de una red que está formada por 27 nodos y 35 aristas. Si calculamos para esta red la distribución de grados, obtenemos la siguiente tabla:

$k$	1	2	3	4	5	6	7	8
$P(k)$	0.4814	0.1851	0.1111	0.1111	0.0000	0.0370	0.0000	0.0370
$k$	9	10	11	12	0.0000	0.0370	0.0000	0.0370
$P(k)$	0.0000	0.0000	0.0000	0.0370				

Vemos claramente la distribución de los nodos si representamos el histograma de las frecuencias de los grados. Lo vemos en la [figura 6.3](#).



**Figura 6.3** Histograma de la red de la [figura 6.2..](#)

Este histograma presenta una forma muy curiosa. Se observa que la mayoría de los nodos de la red tienen unos valores del grado muy bajo, aunque hay algunos pocos nodos con grado máximo o cercano al máximo. Estas son un tipo muy frecuente de redes que llamamos **redes libres de escala**. Se caracterizan porque siguen una distribución del grado de los nodos del tipo exponencial negativa:

$$P(k) = c \cdot k^{-\gamma}.$$

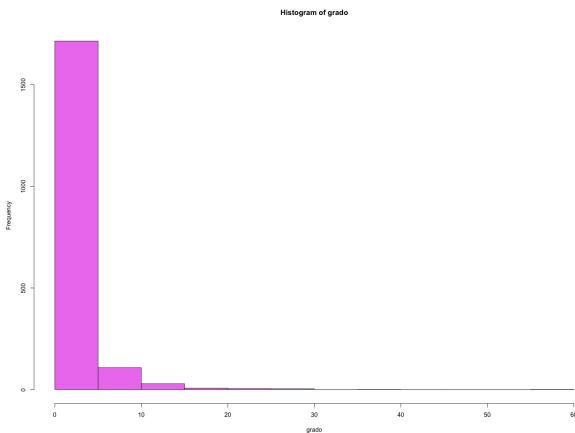
En estadística, la prueba de Kolmogórov-Smirnov (también prueba K-S) es una prueba no paramétrica que determina la bondad de ajuste de dos distribuciones de probabilidad entre sí. Aplicamos el test de Kolmogorov-Smirnov sobre el ajuste de la topología de una red a la propiedad libre de escala. Esta función devuelve un objeto donde el valor  $KS_p$  es el  $p$ -valor correspondiente a rechazar la hipótesis nula que en este caso aserta que la red estudiada es libre de escala. Por lo tanto, un valor alto de  $KS_p$  indica la ausencia de evidencia para afirmar que la red estudiada no es libre de escala.

Para la red del ejemplo mostrada en la [figura 6.2](#), el test K-S nos proporciona un valor de 0.91, lo que significa que se trata de una red claramente *libre de escala*. En nuestro red ejemplo podemos observar que los dos nodos que se comportan principalmente como hubs son el 10 y 20.

En la [sección 5.1.4](#) mostramos un grafo de la base de datos de redes **konect** relacionada con la interacción de proteínas de la levadura. Recordemos que dicho grafo está constituido por 1870 proteínas y se han contabilizado 2270 interacciones entre ellas. Se trata, como se deduce fácilmente de estos datos, de un grafo muy poco denso (de hecho, su coeficiente de densidad es de 0.0013) y el diámetro total del grafo es de 19.

Si estudiamos la distribución del grado de los nodos de este grafo, nos encontramos con

un grafo claramente del tipo libre de escala. Si representamos gráficamente el histograma de la densidad de grados, tenemos la imagen de la [figura 6.4](#).



**Figura 6.4** Histograma de la red de proteínas..

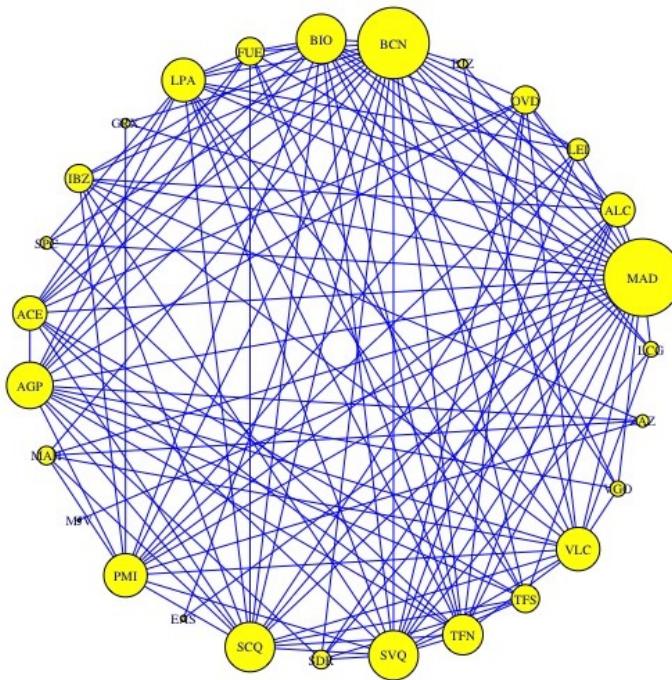
Notemos en la imagen de la [figura 6.4](#) que más del 80 % de los grados de los nodos de la red están entre 1 y 5. También se observa el descenso absoluto que se produce a partir de grado 5. Estamos, como se aprecia claramente, ante una distribución del tipo exponencial negativa, por lo que estamos ante un ejemplo real claro de una red del tipo libre de escala.

Un cierto número de sistemas complejos comparten una propiedad y es que algunos nodos tienen un gran número de conexiones con otros nodos, mientras que el resto de nodos tienen una conexión muy pobre con el resto. Los nodos populares, llamados hubs, pueden tener miles o millones de conexiones y la red parece que no tiene ninguna escala.

El conocimiento sobre este tipo de redes sin escala y con esta propiedad tiene implicaciones importantes en la comprensión de la expansión de los virus de ordenador o en la transmisión de las enfermedades y las epidemias. Existen teorías de difusión que intentan predecir un valor crítico para la propagación de un contagio a través de una población. Cualquier virus que sobrepase el valor umbral se multiplicará exponencialmente por el sistema.

### Ejemplo del grafo de las rutas aéreas.

Nosotros estudiamos ahora el ejemplo de la red de las rutas aéreas de la [sección 5.1.1](#). Vemos una representación de esta red en forma de círculo en la [figura 6.5](#).



**Figura 6.5** Red de rutas de vuelos nacionales..

Ya hemos dicho anteriormente que se trata de una red formada por 27 nodos (aeropuertos nacionales) y 135 aristas (las rutas comerciales entre estos aeropuertos).

Vamos a estudiar las características de esta red y a comprobar si cumple con las condiciones para ser una red libre de escala.

Si calculamos el coeficiente de agrupamiento de cada uno de los nodos  $C_i$  y calculamos la media de todos ellos en la red, obtenemos el coeficiente de agrupamiento  $C$  global de la red. En este caso

$$C = 0.5674,$$

lo que representa un valor intermedio.

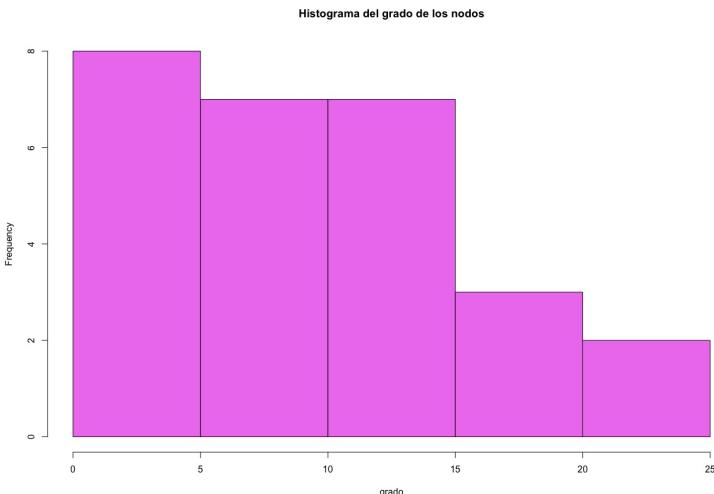
Ahora estudiaremos también el valor de la distancia media entre los nodos, lo que nos proporcionará una información muy valiosa sobre la topología de la red.

Cabe decir que la densidad de esta red es de 0.3846, mientras que su diámetro es 3, lo que significa que la distancia máxima entre dos nodos de la red es de 3. Es decir, para conectar cualesquiera dos aeropuertos de la red nacional, por pequeños y poco activos que sean, necesitaremos como máximo dos escalas. Por ejemplo, si queremos ir desde Las Palmas de Gran Canaria hasta Zaragoza, debemos hacer la ruta siguiente:

$$LCG \longrightarrow MAD \longrightarrow IBZ \longrightarrow ZAZ.$$

Esto ya nos da que pensar sobre el hecho de que la distancia media entre dos aeropuertos de la red es pequeña. Concretamente, la distancia media calculada a partir de la matriz de las distancias, nos proporciona un valor de 1.6296, es decir, entre un camino de longitud 1 y 2. Esto representa un valor muy bajo de la distancia media entre los nodos de la red.

Podemos representar el histograma del grado de los nodos de la red para ver el tipo de distribución que sigue. Lo vemos en la [figura 6.6](#).



**Figura 6.6** Histograma del grafo de la red de vuelos nacionales..

Si observamos la distribución del grado de los nodos, podemos comparar con los histogramas dibujados en los casos anteriores de redes libre de escala vemos que no sigue un dibujo exactamente igual que las anteriores.

Notamos claramente que, aunque esta red tiene propiedades de las redes de libre escala, también tiene algunas propiedades de las redes del mundo pequeño, especialmente en su distancia media entre dos nodos de la red, como veremos más adelante cuando hablaremos de este tipo de redes.

Puesto que tenemos la creencia de que se trata de una red libre de escala pero con algunas propiedades que nos hacen dudar, podemos utilizar el índice de Kolmogorov-Smirnov para obtener una respuesta definitiva. Si calculamos para esta red el índice K-S obtendremos el valor

$$KS.p = 0.7658.$$

Este valor del índice nos indica lo que sospechábamos; se trata de una red libre de escala, aunque el valor de su índice no es tan alto como el de las redes estudiadas en este apartado, que muestran un carácter más puro de las redes libres de escala.

### Ejemplo 2. Redes aleatorias.

Paul Erdős, nacido en Budapest (Hungría) en marzo de 1913, fue un matemático húngaro inmensamente prolífico y famoso por su excentricidad que, con cientos de colaboradores, trabajó en problemas sobre combinatoria, teoría de grafos, teoría de números, análisis clásico, teoría de aproximación, teoría de conjuntos y probabilidad. Erdős fue uno de los publicadores de artículos matemáticos más prolíficos de todos los tiempos, únicamente superado por Leonhard Euler (Erdős publicó más artículos, pero Euler publicó más páginas). Escribió aproximadamente 1500 artículos en el transcurso de su vida, colaborando con alrededor de 500 coautores. Él creía firmemente en las matemáticas como una actividad social.

Debido a sus numerosos aportes, colaboradores y amigos inventaron el número de Erdős como un homenaje con tintes de humor matemático. Erdős tiene asignado el número 0, todos aquellos que colaboraron en algún artículo con él tienen el 1, alguien que haya colaborado con alguno de sus colaboradores tiene el 2, y así sucesivamente.... Sencillas estimaciones comprueban que el 90% de los matemáticos activos tienen un número de Erdős menor que 8 (parece sorprendente si uno no conoce la teoría de Seis grados de separación).

El estudio de redes aleatorias alcanzó prominencia gracias al trabajo fundamental de Pál Erdős y Alfréd Rényi. Escribieron diversos artículos entre 1959 y 1968 en los que combinaban la teoría de la probabilidad con la teoría de grafos. Esto dio origen a lo que llamamos **teoría de grafos o redes aleatoria**.

Una red aleatoria consiste en una red de  $n$  nodos donde cada pareja de nodos se conecta con una probabilidad  $p$ . Existen básicamente dos definiciones basándonos en el modelo utilizado. Así

- Modelo  $G(n,L)$ . La red consta de  $n$  nodos y existen  $L$  aristas aleatoriamente situadas. Este es el modelo que desarrollaron Pál Erdős y Alfréd Rényi.
- Modelo  $G(n,p)$ . Aquí tenemos  $n$  nodos conectados entre ellos de acuerdo a una probabilidad  $p$ .

En una red aleatoria siguiendo el modelo  $G(n,p)$ , la probabilidad de que un nodo  $i$  tenga exactamente  $k$  conexiones es el producto

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k},$$

donde

$$\binom{n-1}{k}$$

representa el número de formas en que podemos seleccionar  $k$  conexiones de  $n-1$  posibles,  $p^k$  representa la probabilidad de tener  $k$  links desde  $i$  y  $(1-p)^{n-1-k}$  representa la probabilidad de que las restantes  $(n-1)-k$  conexiones no se produzcan.

Consecuentemente, la distribución de grados que presenta una red del tipo aleatoria es la distribución binomial dada por

$$P(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}.$$

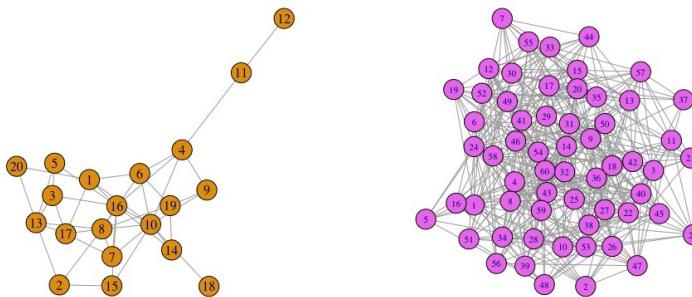
La forma de esta distribución de grados depende del tamaño de la red  $n$  y de la probabilidad  $p$ .

Sin embargo, para valores de  $n$  mucho mayores que  $k$ , esta distribución binomial se aproxima mediante una distribución de Poisson, dada por la expresión

$$P(k) = \exp^{-\bar{k}} \frac{\bar{k}^k}{k!}$$

donde  $\bar{k}$  representa el grado promedio de la red.

A continuación mostramos dos ejemplos de redes creadas aleatoriamente, definidas como  $G(20,1/4)$  y  $G(60,1/4)$ , respectivamente, que se muestran en la figura 6.7. En ambas se ha tomado la probabilidad de establecer conexiones entre dos nodos como  $1/4$ . Esto lo hacemos para comparar ahora la distribución del grado en ambas redes.



**Figura 6.7** Redes generadas aleatoriamente de 20 nodos (izquierda) y 60 nodos (derecha)..

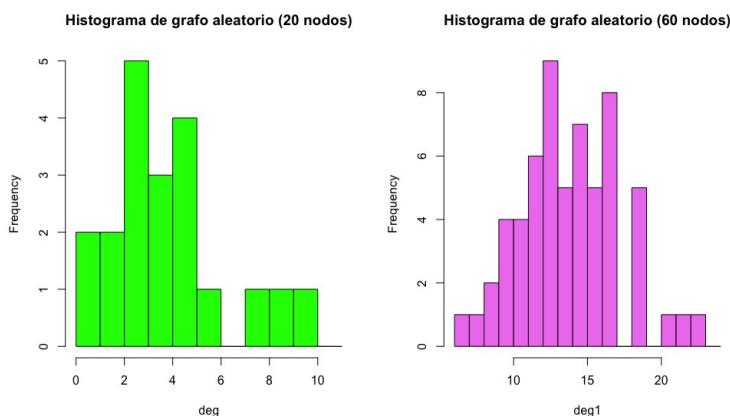
Bertrand Russel sostenía que de un enunciado falso se podía deducir todo. Alguien le preguntó que demostraría que si  $2 + 2 = 5$  entonces él era el Papa. Contestó que si suponemos que  $2 + 2 = 5$ , entonces estaría de acuerdo que si restamos 2 de cada lado obtenemos  $2 = 3$ . Invirtiendo la igualdad y restando 1 de cada lado conseguimos  $2 = 1$ . Como el papa y yo somos dos personas y  $2 = 1$ , entonces el papa y yo somos uno, luego yo soy el Papa.

Blaise Pascal, matemático francés del siglo 17, quien generalizó el triángulo aritmético para el desarrollo de un binomio, de Tartaglia y contribuyó al cálculo de probabilidades. Pascal construyó la primera máquina de sumar y restar que se conoce, a la edad de 18 años, para su padre que tenía problemas con estas operaciones matemáticas. Su padre era cobrador de impuestos.

Einstein en una ocasión se encontró en una fiesta con el gran actor de película Charles Chaplin. Después de las presentaciones de rigor, el autor de la teoría de la relatividad comentó:

- ¡Que envidiable popularidad la suya, señor Chaplin! ¡Su humor es universal y todo el mundo lo entiende!
- Más admirable es su caso, profesor- repuso el actor-, pues todo el mundo lo admira y nadie lo entiende.

Ahora la idea es comparar los histogramas de las distribuciones de grado de ambas redes, lo que se muestra en la [figura 6.8](#).

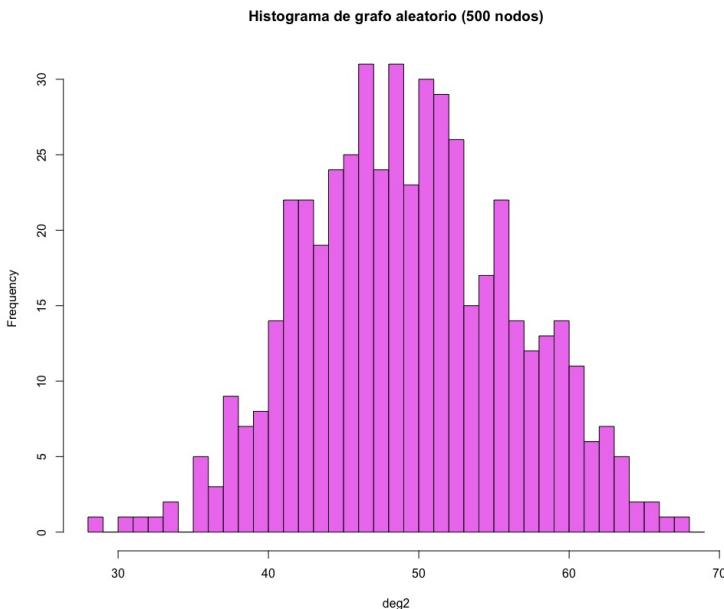


**Figura 6.8** Redes generadas aleatoriamente de 20 nodos (izquierda) y 60 nodos (derecha)..

Revisando ambos gráficos se aprecia la característica de una distribución de grados siguiendo una binomial en ambos casos. Se aprecia una binomial mucho más claramente cuando aumentamos el número de nodos de la red. Para acabar de visualizar mejor este efecto cuando la red va creciendo en tamaño, se ha generado una red de forma aleatoria siguiendo el parámetro  $G(500,1/10)$ , es decir, una red aleatoria con 500 nodos y una probabilidad de conexión entre dos nodos de  $1/10$ .

Con la instrucción  $G(500,1/10)$  hemos generado en R una red formada por 500 nodos

y 12451 aristas. Nos interesa el histograma del grado de los nodos, que también lo hemos calculado ayudándonos del mismo software. Dicho histograma se muestra en la figura 6.9.



**Figura 6.9** Histograma de grado de una red generada aleatoriamente de 500 nodos..

### Ejemplo 3. Redes del mundo pequeño (Small worlds networks).

Las redes del mundo pequeño están relacionadas con el fenómeno que hemos descrito al comienzo del capítulo de los seis grados de separación. La característica fundamental de estas redes es que todo nodo puede ser alcanzado desde un nodo cualquiera en un pequeño número de pasos. La distancia típica  $L$  entre dos nodos elegidos aleatoriamente crece proporcionalmente al logaritmo del número de nodos de la red, es decir,

$$L \propto \log N,$$

teniendo en cuenta que el coeficiente de agrupamiento no es pequeño.

Algunos grafos que cumplen estas características son los grafos que representan redes sociales, las citas de trabajos científicos, internet, some webs like Wikipedia, redes de interacción de proteínas, etc.

Las redes de mundo pequeño tienden a contener *camarillas* y *camarillas cercanas*, es decir, subredes que tienen conexiones entre casi dos nodos dentro de ellas. Pensemos en

una red con nuestros amigos, en la que podemos tener distintas camarillas o subgrupos como los amigos del trabajo, los amigos de la universidad, los amigos del colegio y la infancia, los padres y madres de nuestros hijos, etc. Entre estos grupos las asociaciones y conexiones son muy altas y se forman pequeños hubs que caracterizan la red. Esto se desprende de la propiedad definitoria de un alto coeficiente de agrupamiento.

En segundo lugar, hay que destacar que la mayoría de los pares de nodos estarán conectados por al menos una ruta corta. Esto se deduce de la propiedad definitoria de que la longitud del camino medio más corto sea pequeña. Varias otras propiedades a menudo se asocian con redes de mundo pequeño. Por lo general, hay una gran cantidad de concentradores: nodos en la red con un alto número de conexiones (conocidos como nodos de alto grado). Estos centros sirven como las conexiones comunes que median las longitudes de ruta corta entre otros bordes.

Existe un índice para determinar si una red se encuentra dentro del grupo de las small world. Dicho coeficiente  $\sigma$  se calcula comparando el coeficiente de agrupamiento  $C$  y la longitud de los caminos de una red dada respecto de su equivalente red aleatoria con el mismo grado y longitud media de caminos, es decir,

$$\sigma = \frac{C/C_r}{L/L_r}.$$

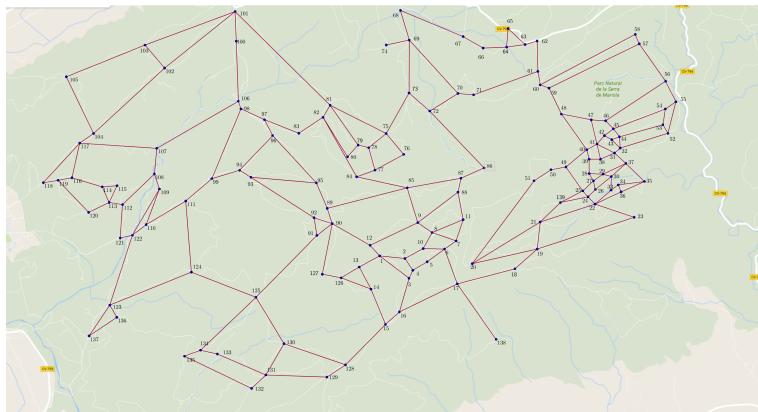
Si  $\sigma > 1$ , entonces la red es de mundo pequeño.

#### Ejemplo 4. Redes urbanas.

En la [sección 5.6](#) analizamos el modo en que podíamos transformar un mapa urbano con su trazado de calles, plazas y espacios públicos en una estructura matemática como es un grafo. Llegamos a la conclusión de que un grafo primario, donde los nodos son las intersecciones de calles y las aristas son las propias calles, podría resultar una solución adecuada al problema de obtener un grafo de un mapa.

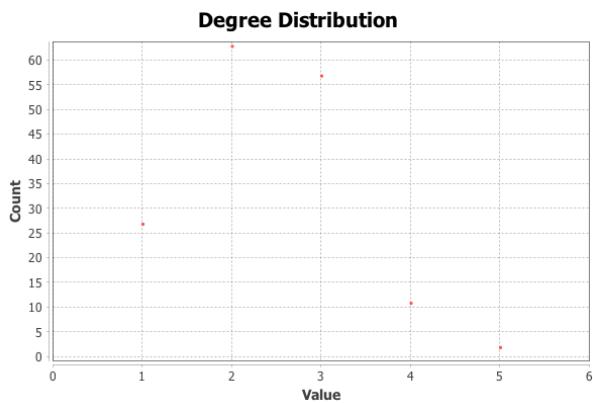
Los grafos asociados a las redes urbanas tienen una particularidad muy destacable respecto a la distribución del grado de los nodos. Dicha característica es que la mayoría de los nodos de la red tienen un grado que se encuentra entre 2 y 5, aunque la red esté formada por miles de nodos.

Analizamos el ejemplo de la [sección 5.1.3](#), donde construimos una red de carreteras y caminos en un entorno natural. Dicha red está compuesta por 160 nodos y 189 aristas, lo que representa una densidad muy baja como podemos suponer. Vemos el grafo que construimos con los nodos numerados en la [figura 6.10](#). Concretamente, la densidad es de 0.015, lo que supone un valor muy bajo de la densidad.



**Figura 6.10** Grafo de la red de caminos de Mariola..

Si calculamos los grados de todos los nodos que la integran, obtenemos un valor de grado medio de 2.36, mientras que el diámetro total de la red es 25. Igualmente, la distancia media entre cualesquiera par de nodos es de 10.96. Con los datos recopilados hasta ahora, podemos decir que no se trata de una red libre de escala, ni tampoco aleatoria. Por tanto, no se ajusta a ninguno de los tipos de redes que hemos estudiado hasta ahora.



**Figura 6.11** Histograma de grados del grafo de la red de caminos de Mariola..

En la [figura 6.11](#) se muestran los valores de la distribución de grados del grafo de los caminos y carreteras de la sierra de Mariola. Podemos construir la [sección 3.1](#) siguiente con las distribuciones de los grados.

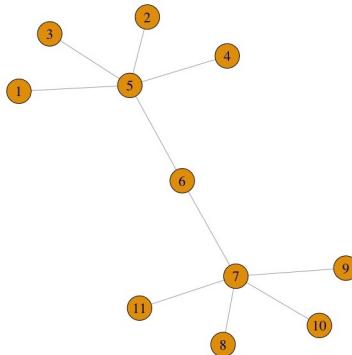
$k$	1	2	3	4	5
$P(k)$	0.1687	0.3937	0.3562	0.0687	0.0187

Se observa que los grados toman valores entre 1 y 5. Destaca que existe un número elevado de nodos con rango 2 y 3, de la misma forma que sucede en las redes urbanas.

### 6.3 La centralidad en redes complejas

La idea de centralidad constituye uno de los conceptos más importantes dentro del moderno análisis de redes complejas. La centralidad de un nodo mide la importancia del nodo dentro de la red, existiendo un buen número de medidas de centralidad atendiendo a las diferentes interpretaciones y objetivos que persigamos en la aplicación concreta que estamos estudiando. Es decir, la idea de importancia es muy relativa, ya que de acuerdo con la aplicación o red analizada, podemos considerar que un nodo es muy importante si está conectado con otros muchos nodos dentro de esa red. Sin embargo, también podemos considerar que la importancia de un nodo dentro de una red está relacionada con el hecho de que sea un nodo de paso en muchos caminos que conectan otros nodos. O también podríamos considerar la importancia de un nodo en función de si está conectado con otros nodos importantes.

Antes de entrar en una descripción más detallada de distintas medidas de centralidad clásicas basadas en la topología de la red, intentamos motivar con un ejemplo la idea de centralidad. Consideraremos el grafo de la figura 6.12.



**Figura 6.12** Red de ejemplo..

Nos preguntamos, después de observar la figura 6.12, ¿cuál es el nodo más importante del grafo? En cuanto al tema de la importancia de los nodos, podemos decir que dividiríamos los nodos del grafo en tres grupos: el primer grupo estaría formado por el nodo 6. Este nodo es fundamental puesto que es el que conecta las dos partes del grafo. El segundo grupo estaría formado por los nodos 5 y 7, que son nodos que tienen un grado elevado, están conectados con muchos nodos del grafo. El tercer grupo lo forman el resto de los nodos.

La cuestión es que dependiendo del contexto, uno de los vértices puede ser más importante que los otros. Supongamos ahora que eliminamos el nodo 6. Ahora el grafo está claramente dividido en dos componentes. Sin embargo, si eliminamos el nodo 5, el grafo es completamente distinto puesto que se divide en varias componentes, quedando los nodos 1,2,3,4 aislados.

Hasta aquí hemos tenido en cuenta la idea de importancia de un nodo fijándonos en los caminos dentro del mismo, buscando aquéllos elementos que son esenciales en los caminos que se establecen entre cualesquiera pares de nodos. Si nos fijamos de nuevo en el grafo de la figura 6.12 y pensamos en la importancia de un nodo como aquél que tiene más conexiones, no podemos decir que el nodo 6 sea el más importante. Los nodos 5 y 7 son los más importantes desde el punto de vista de la distribución de grado.

En consecuencia, los ejemplos anteriores nos hacen pensar sobre la posible diversidad de las medidas de centralidad existentes en la moderna teoría de redes complejas, así como en sus diversos ámbitos de aplicación. No es nuestro objetivo desarrollar un exhaustivo estudio de las medidas de centralidad existentes, sino describir brevemente aquéllas que podemos considerar *clásicas* o más utilizadas desde hace varias décadas.

En principio, introducimos brevemente, cuatro medias de centralidad fundamentales para determinar la importancia de los nodos de una red, basándonos en sus características topológicas. Son

**degree** La centralidad de grado.

**closeness** Centralidad de cercanía.

**betweenness** Centralidad de intermediación.

**eigenvector** Centralidad de vector propio.

### La centralidad de grado

Consideremos un grafo  $V$  no dirigido con  $n$  vértices y sea  $A$  su matriz de adyacencia. La centralidad de grado está relacionada con el número de enlaces o conexiones que tiene un vértice  $j$ . Más formalmente, podemos decir que la centralidad de grado  $C_d$  del vértice  $j$  viene dada por

$$C_d(j) = \text{grado}(j) = \sum_i a_{ij}$$

siendo  $a_{ij}$  elementos de la matriz de adyacencia del grafo  $V$ .

Para grafos dirigidos debemos definir dos medidas de grado diferentes, una para el grado de entrada y otra para el de salida. En términos de relaciones sociales, la medida de grado de entrada puede darnos una idea de popularidad del nodo, mientras que el grado de salida puede relacionarse con el grado de sociabilidad.

Siguiendo con el grafo de la figura 6.12 calculamos la centralidad de grado de cada uno de los nodos.

nodo	1	2	3	4	5	6	7	8	9	10	11
grado	1	1	1	1	5	2	5	1	1	1	1

Los nodos con una mayor centralidad de grado son el 5 y el 7.

### La centralidad de cercanía

La medida de cercanía,  $C_c$ , definida por el matemático Murray Beauchamp en 1965 y luego popularizada por Freeman en 1979 es la más conocida y utilizada de las medidas radiales de longitud. Se basa en calcular la suma o bien el promedio de las distancias más cortas desde un nodo hacia todos los demás.

Puede determinarse cuantitativamente mediante la expresión

$$C_c(j) = \frac{1}{\sum_{j=i}^n(S)_{ij}}, \quad (6.1)$$

donde  $S$  representa la matriz de distancias de la red, es decir, aquella matriz cuyo elemento genérico  $S(i,j)$  representa la distancia más corta desde el nodo  $i$  hasta el nodo  $j$ .

Cuanto menor sea el valor anterior, se puede decir que el nodo está más próximo al centro de la red. La idea relacionada con esta medida es su interpretación como en una red con flujos, el tiempo de llegada a destino de algo que está fluyendo por la misma. de alguna forma estamos midiendo la accesibilidad de un nodo en la red.

Hay que notar que en esta medida, cuando hablamos de caminos entre nodos, nos estamos refiriendo a caminos euclídeos dentro de la red, es decir, caminos más cortos desde el punto de vista topológico. esto es ciertamente discutible en algunos tipos de red, lo que ha dado a diversas variantes de esta medida, considerando caminos aleatorios dentro de la red, por ejemplo.

Siguiendo con el ejemplo que estamos utilizando para el cálculo de las centralidades, veamos la centralidad de cercanía de los nodos del grafo de la [figura 6.12](#). Para ello, de acuerdo con la definición dada por la expresión [expresión \(6.1\)](#), debemos calcular la matriz de distancias. Calculamos la distancia entre los caminos más cortos entre los nodos  $i$  y  $j$ . La matriz de distancias es:

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

Así, tenemos que

$$C_c(1) = \frac{1}{\sum_{j=1}^{11}(S)_{1j}} = \frac{1}{S_{11} + S_{12} + \dots + S_{111}} = \frac{1}{28}$$

$$C_c(2) = \frac{1}{\sum_{j=1}^{11}(S)_{2j}} = \frac{1}{S_{21} + S_{22} + \dots + S_{211}} = \frac{1}{28}$$

$$C_c(3) = \frac{1}{\sum_{j=i}^{11} (S)_{3j}} = \frac{1}{S_{31} + S_{32} + \dots + S_{311}} = \frac{1}{28}$$

Análogamente,

$$C_c(4) = \frac{1}{28}, C_c(5) = \frac{1}{19}, C_c(6) = \frac{1}{18}, C_c(7) = \frac{1}{19},$$

$$C_c(8) = \frac{1}{28}, C_c(9) = \frac{1}{28}, C_c(10) = \frac{1}{28}, C_c(11) = \frac{1}{28},$$

Como era de esperar, el nodo con mejor centralidad de cercanía es el nodo 6, seguido de los nodos 5 y 7.

### La centralidad de intermediación

La centralidad de intermediación  $C_b$  parte de unas premisas bastante diferentes a las vistas en los casos anteriores de centralidad; ahora resulta fundamental el concepto de intermediación de un nodo. Podemos decir que la intermediación de un nodo cuantifica la frecuencia o el número de veces que un nodo actúa como un *punto de interrupción* a lo largo del camino más corto entre otros dos nodos de la red.

La medida fue introducida por Linton Freeman en 1977, como forma de cuantificar el control de un humano en la comunicación existente con otros humanos en una red social. La idea intuitiva es que si se eligen dos nodos al azar, y luego también al azar uno de los eventuales posibles caminos más cortos entre ellos, entonces los nodos con mayor intermediación serán aquellos que aparezcan con mayor probabilidad dentro de este camino.

Si suponemos que  $g_i^{(st)}$  es el número de caminos geodésicos desde el vértice  $s$  al vértice  $t$  que pasan por  $i$  y suponemos que  $n_{st}$  es el número total de caminos geodésicos que unen  $s$  con  $t$ , entonces podemos escribir la medida betweenness de  $i$  como

$$C_b(i) = \frac{\sum_{s < t} g_i^{(st)} / n_{st}}{\frac{1}{2}n(n-1)}$$

donde  $n$  es el número total de vértices o nodos de la red.

Los nodos con una alta intermediación, al igual que las aristas con una alta intermediación en los lazos interpersonales (tema que no tiene que ver directamente con centralidad y que no abordaremos aquí), suelen jugar un rol crítico en la estructura de la red, cuando hay grandes flujos de información que son transportados por nodos pertenecientes a grupos compactos. En una red social están relacionados con aquellos nodos de los que depende la integración de algunas componentes de la red. Los nodos que poseen una posición de intermediarios de alguna manera son también controladores o reguladores del flujo de información.

Continuamos con el mismo ejemplo de la figura 6.12, donde el grafo tenía 11 nodos y 10 aristas. Como ejemplo, nos proponemos calcular la medida de centralidad betweenness del nodo 6, a modo de ejemplo.

La clave del cálculo de esta centralidad se basa en el cálculo de los caminos más cortos entre todos los pares de vértices. Si observamos detenidamente el grafo, encontramos

una peculiaridad muy importante para este caso. Podemos decir que entre dos pares de nodos cualesquiera existe únicamente un camino de longitud mínima. Esta característica se traduce en que

$$n_{st} = 1,$$

para cualesquiera  $s$  y  $t$ , con  $s \neq t$ . Esto se debe a la particular configuración topológica del grafo. de esta forma, el cálculo de la centralidad se reduce a encontrar en dichos caminos más cortos, el número de caminos en los que se pasa por el vértice 6.

Así, por ejemplo, si consideramos los caminos  $g_6^{(1t)}$ , para  $t = 2, 3, \dots, 11$ , vemos que

$$\frac{g_6^{(12)}}{n_{12}} + \frac{g_6^{(13)}}{n_{13}} + \frac{g_6^{(14)}}{n_{14}} + \frac{g_6^{(15)}}{n_{15}} + \frac{g_6^{(16)}}{n_{16}} + \frac{g_6^{(17)}}{n_{17}} + \frac{g_6^{(12)}}{n_{18}} + \frac{g_6^{(19)}}{n_{19}} + \frac{g_6^{(110)}}{n_{110}} + \frac{g_6^{(111)}}{n_{111}} = 5.$$

Ahora, para  $s = 2$ , hacemos lo mismo y obtenemos

$$\frac{g_6^{(23)}}{n_{23}} + \frac{g_6^{(24)}}{n_{24}} + \frac{g_6^{(25)}}{n_{25}} + \frac{g_6^{(26)}}{n_{26}} + \frac{g_6^{(27)}}{n_{27}} + \frac{g_6^{(28)}}{n_{28}} + \frac{g_6^{(29)}}{n_{29}} + \frac{g_6^{(210)}}{n_{210}} + \frac{g_6^{(211)}}{n_{211}} = 5.$$

De la misma forma hacemos para  $s = 3, 4, \dots, 11$ . Obtenemos

$$C_b(6) = \frac{\sum_{s < t} g_6^{(st)} / n_{st}}{\frac{1}{2} 11(10)} = \frac{25}{55}.$$

Si calculamos el resto de centralidades del resto de nodos, obtenemos el vector de centralidades

nodo	1	2	3	4	5	6	7	8	9	10	11
$C_b(i)$	0	0	0	0	$\frac{30}{55}$	$\frac{25}{55}$	$\frac{30}{55}$	0	0	0	0

Notemos la cantidad de nodos que tienen una centralidad de intermediación nula. Eso es consecuencia de la topología y las pocas aristas del grafo.

### La centralidad de vector propio

La centralidad de vector propio mide la influencia de un nodo en una red. Fue propuesta por Phillip Bonacich en 1972 y corresponde al principal vector propio de la matriz de adyacencia del grafo analizado.

Intuitivamente, los nodos que poseen un valor alto de esta medida de centralidad están conectados a muchos nodos que a su vez están bien conectados, también en este sentido; por lo tanto, son buenos candidatos para difundir información, divulgar rumores o enfermedades, etc. Los nodos más centrales en este sentido corresponden a centros de grandes grupos cohesivos. Mientras que en el caso de la centralidad de grado, cada nodo pesa lo mismo dentro de la red, en este caso la conexión de los nodos pesa de forma diferente.

### Un ejemplo de cálculo de la centralidad

Una vez explicados brevemente estos modelos de centralidad basados en la estructura topológica de la red, nos centramos en un ejemplo concreto donde podamos comprender la utilidad de este concepto en redes reales. En la sección 5.1.5, presentamos una red

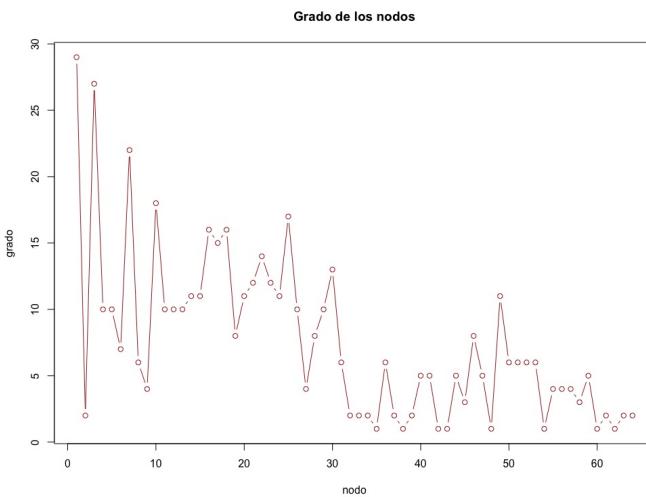
relacionada con la organización criminal asociada a los atentados terroristas de los trenes de Madrid en 2004. En dicho grafo se muestran las relaciones entre los distintos terroristas, de manera que constituyen un grafo no dirigido compuesto por 64 nodos y 243 relaciones o aristas.

Hemos determinado algunas de sus características como son:

- Presenta una densidad de grado de 0.12, lo que significa que no es muy densa.
- Su diámetro es de 6.
- Su distancia media entre pares de nodos es de 2.69.

Si calculamos el grado de los nodos del grafo, tendremos los grados del vector

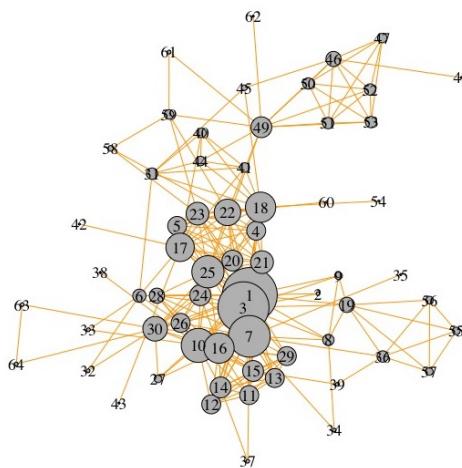
$$dg = (29, 2, 27, 10, 10, 7, 22, 6, 4, 18, 10, 10, 10, 11, 11, 16, 15, 16, 8, 11, 12, 14, 12, 11, 17, 10, 4, 8, 10, 13, 6, 2, 2, 1, 6, 2, 1, 2, 5, 5, 1, 1, 5, 3, 8, 5, 1, 11, 6, 6, 6, 6, 1, 4, 4, 4, 3, 5, 1, 2, 1, 2, 2)$$



**Figura 6.13** Grados de los distintos nodos..

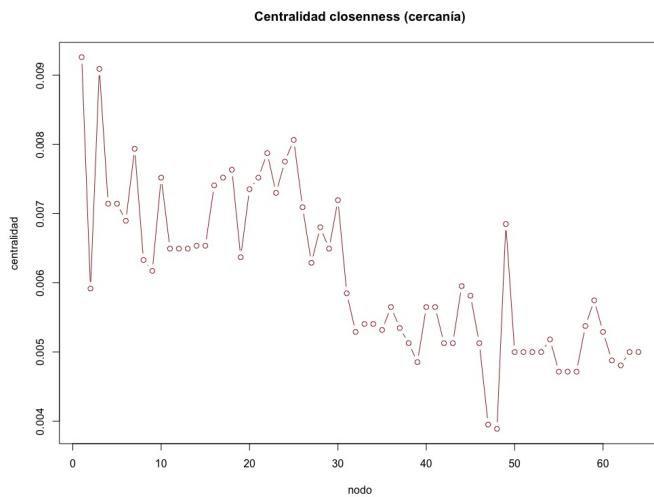
Como podemos observar en la figura 6.13 en la que hemos representado los grados de los 64 nodos del grafo, los nodos 1,3,7 son los nodos que presentan un mayor grado, es decir, son los nodos que tienen un número mayor de conexiones. Desde el punto de vista de las relaciones o contactos entre ellos, podríamos decir que los nodos más importantes son éstos.

Se ha representado el grafo haciendo los tamaños de los nodos en función de su importancia cuando calculamos la centralidad de grado. Esta representación aparece en la figura 6.14.

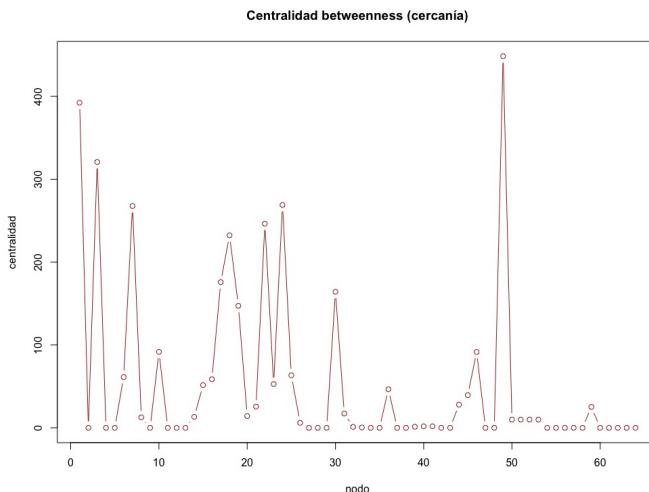


**Figura 6.14** Representación de la centralidad de grado..

La centralidad de grado se basa en la idea de las conexiones individuales de cada uno de los nodos de la red. Pero esto es únicamente un criterio; en la red estudiada significa que calculamos los nodos con más conexiones. Esto no significa que los nodos con más conexiones sean los más importantes siguiendo otros criterios y otras medidas de centralidad que analizaremos a continuación.



**Figura 6.15** Representación de la centralidad de cercanía..



**Figura 6.16** Representación de la centralidad betweenness o intermediación..

La siguiente centralidad que podemos calcular es la centralidad de cercanía, es decir, la centralidad que tiene en cuenta los nodos más centrales desde el punto de vista de su distancia al resto, entendiendo por distancia como los nodos de distancia en los caminos entre el nodo en cuestión y el resto.

Se ha calculado la centralidad de cercanía de cada nodo y su representación gráfica aparece en la [figura 6.15](#).

Si comparamos la centralidad de grado con la centralidad de cercanía, notamos que aunque los resultados son parecidos, no podemos decir que son absolutamente iguales. Los grados 1 y 3 siguen siendo los más importantes en ambas centralidades, pero el nodo que ocupa el tercer lugar ya cambia. El nodo 25 ahora toma un papel importante en la centralidad de cercanía, ligeramente por delante del 7. Como vemos, no son resultados muy distintos, existe bastante similitud en los mismos, lo que nos da una idea aproximada ya de qué nodos son fundamentales en la red.

Seguimos con una de las centralidades más importantes por su significado en esta red real concreta. Se trata de la centralidad de intermediación, que nos proporciona una visión distinta de las medidas anteriores. Ahora la importancia de un nodo la basamos en su importancia como nodo de paso en los caminos más cortos que conectan el resto de los nodos. Es decir, estamos asignando una importancia muy significativa a aquéllos nodos que son de paso obligado dentro de la estructura topológica de la red.

Esto cobra un gran sentido en el ejemplo que estamos estudiando. Buscamos nodos que sean fundamentales a la hora de tomar decisiones dentro de la trama de la red, esto significa nodos que se encuentran en el camino de las trayectorias entre el resto de nodos. Hemos calculado esta centralidad para este ejemplo de forma análoga a las centralidades anteriores y el resultado se muestra en [figura 6.16](#).

Ahora, si observamos la figura 6.16 la cuestión cambia un poco y el nodo más importante según las centralidades anteriores, ya no es el nodo más central de la medida betweenness. Se trata del nodo 1; en la centralidad betweenness el nodo más importante ahora es el 49. Esto significa que el nodo 49 es un nodo que no es el mejor relacionado con el resto ni el que se encuentra más cerca del resto, pero sí que es fundamental en la toma de decisiones puesto que cualquier trayectoria en el grafo pasa por él, es decir, podemos interpretar que es fundamental en la estructura global. Los nodos 1 y 3 siguen siendo muy importantes en la red.

## 6.4 El concepto PageRank

Realizar una búsqueda por internet actualmente se ha convertido en una rutina o tarea cotidiana. No le damos la importancia que merece a la historia de los buscadores y no nos preocupamos sobre el fundamento que hay detrás de una búsqueda cualquiera en uno de los muchos buscadores existentes. La gran revolución en los buscadores vino de la mano de dos estudiantes de Stanford, cuya gran y genial idea fue implementar un algoritmo de clasificación de las páginas web basado no solo en el número de entradas de cada página, sino en la calidad o importancia de éstas. Lo explicamos detenidamente en esta sección.

Fue en 1993 cuando la información digital vivió su gran revolución, con la aparición del primer buscador propiamente dicho. Se llamó *Wandex* (*World Wide Web Wanderer*) y fue desarrollado por un científico del MIT, Mathew Gray, aunque en origen no pretendía ser un buscador ni mucho menos.

El germen de *Wandex* estaba en una herramienta científica que pretendía medir el tamaño de Internet con fines académicos, cuantificando el número total de informaciones que existían en la Red. No obstante, una vez desarrollado el motor de esta aplicación, el equipo del MIT se dio cuenta de que ese robot también podía ?con algunas modificaciones- leer URLs e indexarlas para encontrarlas fácilmente. Así nacieron las búsquedas en Internet.

Posteriormente surgieron otros motores de búsqueda, entre los que destacaríamos **Altavista**. *AltaVista* fue otro de los buscadores más conocidos y apareció en 1995. La web fue muy innovadora porque tenía dos características que ningún otro buscador poseía en ese momento: usaba un rastreador rápido que manejaba varios sub-procesos que le permitían buscar más páginas web de las que se creía que existían y hacia una búsqueda muy eficiente gracias a un hardware avanzado en los servidores. *AltaVista* también tenía un diseño minimalista y limitaba los resultados de un dominio para evitar recibir varios resultados de la misma fuente.

Con tantas funciones interesantes, AltaVista se convirtió en un éxito después de estar dos años en línea, recibiendo más de 80 millones de visitas cada día. En 1998 con la llegada y la popularidad de Google, AltaVista comenzó a perder usuarios, hasta que finalmente cerró el 8 de julio de 2013.

El buscador de *Google* (en inglés **Google Search**) es un motor de búsqueda en la web propiedad de *Alphabet Inc.*, es el motor de búsqueda más utilizado en la Web. Recibe cientos de millones de consultas cada día a través de sus diferentes servicios. El objetivo principal del buscador de *Google* es buscar texto en las páginas web, en lugar de otro tipo

de datos, fue desarrollado originalmente por Larry Page y Sergey Brin en 1997, en la Universidad de Stanford mientras estudiaban el posgrado en Ciencias de la Computación.

El orden de los resultados de búsqueda en las páginas de resultados de *Google* se basa, en parte, en un rango de prioridad llamado *PageRank*. El Buscador de *Google* proporciona muchas opciones para la búsqueda personalizada. Por poner un ejemplo, en 2013, *Google* indexó más de 60 trillones de páginas web, 154.300 consultas por segundo, 400 millones de consultas por día, 1,3 mil millones de imágenes y más de mil millones de mensajes de Usenet.

*PageRank* es una marca registrada y patentada por *Google* el 9 de enero de 1999 que ampara una familia de algoritmos utilizados para asignar de forma numérica la relevancia de los documentos (o páginas web) indexados por un motor de búsqueda.

*PageRank* ha tomado su modelo del *Science Citation Index* (SCI) elaborado por Eugene Garfield para el *Instituto para la Información Científica* (ISI) en los Estados Unidos durante la década del 50. El SCI pretende resolver la asignación objetiva de méritos científicos suponiendo que los investigadores cuyo factor de impacto (número de publicaciones y/o referencias bibliográficas en otros trabajos científicos) es más alto, colaboran en mayor medida con el desarrollo de su área de investigación. El índice de citación es un elemento determinante para seleccionar qué investigadores reciben becas y recursos de investigación.

*Google* ordena los resultados de la búsqueda utilizando su propio algoritmo *PageRank*. A cada página web se le asigna un número en función del número de enlaces de otras páginas que la apuntan, el valor de esas páginas y otros criterios no públicos. *Google* interpreta un enlace de una página A a una página B como un voto, de la página A, para la página B. Pero *Google* mira más allá del volumen de votos, o enlaces que una página recibe; también analiza la página que emite el voto. Los votos emitidos por las páginas consideradas "importantes", es decir con un *PageRank* elevado, valen más, y ayudan a hacer a otras páginas "importantes". Por lo tanto, el *PageRank* de una página refleja la importancia de la misma en Internet.

#### 6.4.1 El *PageRank*

Supongamos que tenemos un conjunto de  $n$  páginas  $\{1, 2, \dots, n\}$ . Si la página  $i$  tiene un enlace a la página  $j$  es porque la página  $i$  considera relevante la información que proporciona la página  $j$ . Si un buen número de páginas tienen enlaces a la página  $j$ , entonces debemos pensar que dicha página es ciertamente relevante en ese campo. Sin embargo, también debemos tener en cuenta la importancia de los enlaces, es decir, la relevancia o importancia de la página que enlaza con la que se evalúa.

Así pues, nos encontramos con el problema de establecer una clasificación de páginas en las que existen relaciones de conectividad entre ellas. Podemos trasladar nuestro problema al lenguaje de las redes o grafos si consideramos que cada página constituye un nodo y unimos dos nodos entre sí si existe un enlace entre esas páginas. De esta manera tendremos un grafo dirigido en el que una arista del nodo  $i$  al nodo  $j$  si la página  $i$  tiene un enlace a la página  $j$ .

También consideramos el número de enlaces que tiene cada página, ya que interpretamos que cada página reparte la importancia en sus conexiones entre todos sus enlaces a partes

iguales. Concretamente, denotaremos por  $c_i$  al número de conexiones de la página  $i$ , para  $i = 1, 2, \dots, n$ , siendo  $n$  el número total de nodos o páginas.

De esta forma, disponemos de un grafo dirigido al que podemos asociar una cierta matriz  $A^*$  de tamaño  $n \times n$ , cuyo elemento  $a_{ij}^*$  podemos definir de la forma siguiente:

$$a_{ij}^* = \frac{1}{c_j}$$

si la página  $j$  tiene un enlace a la página  $i$ . Si la página  $j$  no tiene ningún enlace a la página  $i$ , entonces  $a_{ij}^* = 0$ .

Construida de esta forma, la matriz  $A^*$  representa, por columnas, las conexiones entre las páginas. Podemos llamarla matriz de transición. Notemos su similitud con la matriz de adyacencia de un grafo.

Además, la matriz  $A^*$  tiene las siguientes características:

1.  $A^*$  es de términos positivos.
2.  $A^*$  es estocástica por columnas.
3. El radio espectral de  $A^*$  es 1. El radio espectral es el valor máximo en valor absoluto de los valores propios.
4. La matriz  $A^*$  es irreducible y primitiva.

Como consecuencia de estas características, vemos que se cumplen las hipótesis del teorema de Perron-Frobenius.

**Teorema 6.4.1 (Perron-Frobenius)** *Sea una matriz  $A$  positiva y estocástica. Entonces, se cumple que*

1.  $\lambda = 1$  es un valor propio de  $A$  de multiplicidad 1.
2.  $\lambda = 1$  es, en módulo, el mayor valor propio.
3. El vector propio asociado a  $\lambda = 1$  tiene todos sus valores positivos. En particular, solo uno de ellos cumple que la suma de todas sus componentes es igual a 1.

El teorema de Perron-Frobenius es de vital importancia en este problema, ya que es el que nos asegura que podemos encontrar un vector  $\mathbf{x}$  asociado al valor propio  $\lambda = 1$ , de manera que

$$\mathbf{A}\mathbf{x} = 1 \cdot \mathbf{x}.$$

Con las características que presenta la matriz  $A^*$  podemos deducir que el vector Page-Rank es el vector límite de distribución de probabilidad de una cadena de Markov ergódica, lo que significa que el vector de estado límite existe y coincide con el vector de estado estacionario, que es un vector de probabilidad asociado al valor propio 1. Además, es independiente del vector inicial.

Dicho vector  $\mathbf{x}$  constituye nuestro vector buscado, que llamaremos **pagerank**. Nos proporciona la ordenación de las páginas de acuerdo con el criterio de conectividad entre ellas.

El cálculo de este vector **pagerank** presenta problemas cuando nos encontramos con nodos aislados o nodos que no tienen enlaces a otras páginas.

Concretamente, Page y Brin se dan cuenta de que la estructura de la web da lugar a que la matriz  $A^*$  no sea primitiva, por lo que introducen un nuevo modelo definiendo una nueva matriz, que llamamos  $G$ , como

$$G = (1 - p)A^* + pB, \quad (6.2)$$

donde  $A$  es la matriz de transición definida anteriormente,  $p$  es un factor (*damping factor*) y  $B$  es la matriz que definimos como

$$B = \frac{1}{n}I_n.$$

Podemos escribir la matriz ([expresión \(6.2\)](#)) como

$$M = (1 - p)A^* + p\mathbf{v}\mathbf{e}^T, \quad (6.3)$$

donde  $\mathbf{e} \in \mathbb{R}^{n \times 1}$ , es decir es un vector cuyas todas sus componentes son 1 y  $\mathbf{v}$  es el llamado vector de personalización, que se suele tomar como  $\mathbf{v} = \frac{1}{n}\mathbf{e}$ .

La idea de la definición de la matriz  $B$  consiste en la posibilidad de que un cierto usuario no navegue de una página a otra, a través de un enlace con la primera. El factor  $p$  refleja la probabilidad de que un usuario abandone la página en la que se encuentre y vaya a cualquier otra, sin seguir enlaces directos. Normalmente, un valor típico de  $p$  es 0.15.

Sin embargo, seguimos teniendo problemas en el caso de que tengamos nodos que no tienen enlaces salientes, por lo que tendremos columnas de ceros. Para ello, se define una variante de ([expresión \(6.3\)](#)), que denotamos por  $G$  y que definimos como

$$G = (1 - p)(A^* + \mathbf{v}\mathbf{d}^T) + p\mathbf{v}\mathbf{e}^T, \quad (6.4)$$

donde  $\mathbf{e}$ , y  $\mathbf{v}$  son los mismos que en el caso de la matriz  $M$  y donde  $\mathbf{d}$  es ahora un vector de tamaño  $n \times 1$  cuyo elemento  $d_i = 1$  en el caso de que  $c_i = 0$ , es decir, el vector  $\mathbf{d}$  es de ceros, salvo en aquellos elementos que no tienen enlaces al exterior, es decir  $c_i = 0$ .

A  $G$  se la llama matriz de Google.

Se puede demostrar que la matriz de Google dada por ([expresión \(6.4\)](#)) es una matriz estocástica por columnas y de términos positivos, por lo que tenemos garantizadas las propiedades que se siguen del teorema de Perron-Frobenius.

#### 6.4.2 Cálculo del vector PageRank por sistemas de ecuaciones

El vector PageRank, es decir, el vector estacionario de la cadena de Markov ergódica definida por  $M$  es el vector de distribución de probabilidad que verifica el problema de valores y vectores propios

$$M\mathbf{x} = \mathbf{x}.$$

Esta ecuación puede escribirse de la forma

$$(I - M)\mathbf{x} = 0, \quad (6.5)$$

sujeto a la condición de que  $\mathbf{x}^T \mathbf{e} = 1$ .

Se cumple que  $I - M$  es una matriz estocástica por columnas. El problema consiste en resolver el sistema (expresión (6.5)) siendo  $M$  grande.

Introducimos un nuevo enfoque en el que sustituimos en (expresión (6.5)) el valor de  $M$ , de la forma siguiente:

$$(I - M)\mathbf{x} = 0,$$

$$[I - [(1-p)A - p\mathbf{v}\mathbf{e}^T]]\mathbf{x} = 0,$$

por lo que

$$[I - (1-p)A]\mathbf{x} - p\mathbf{v}\mathbf{e}^T\mathbf{x} = 0,$$

$$[I - (1-p)A]\mathbf{x} = p\mathbf{v}\mathbf{e}^T\mathbf{x},$$

por lo que como  $\mathbf{e}^T \mathbf{x} = 1$ , entonces

$$[I - (1-p)A]\mathbf{x} = p\mathbf{v}.$$

### 6.4.3 Un ejemplo

**Ejemplo 6.4.1** Consideremos un ejemplo concreto con un conjunto de 10 plazas  $P_1, P_2, \dots, P_{10}$  que se encuentran conectadas de la forma que se muestra en la figura 6.17. Queremos establecer un ranking con estas plazas basándonos en su conectividad con el resto.

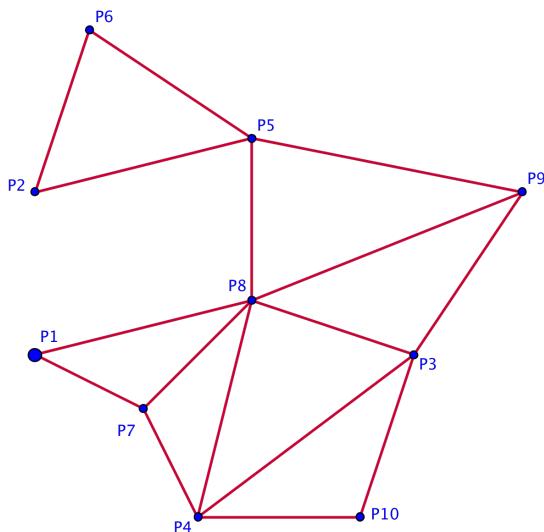


Figura 6.17 Red de diez plazas de ejemplo..

Vemos en la figura el ejemplo con las plazas y su conectividad, que podemos resumir a través de la matriz de transición  $A^*$  dada por

$$A^* = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1/3 & 1/6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/4 & 0 & 0 & 0 & 1/6 & 1/3 & 1/2 \\ 0 & 0 & 1/4 & 0 & 0 & 0 & 1/3 & 1/6 & 0 & 1/2 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 & 1/6 & 1/3 & 0 \\ 0 & 1/2 & 0 & 0 & 1/4 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/4 & 0 & 0 & 0 & 1/6 & 0 & 0 \\ 1/2 & 0 & 1/4 & 1/4 & 1/4 & 0 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 1/4 & 0 & 1/4 & 0 & 0 & 1/6 & 0 & 0 \\ 0 & 0 & 1/4 & 1/4 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Como se trata de una matriz estocástica por columnas con sus valores positivos, sabemos, por el teorema de Perron-Frobenius, que tiene un valor propio  $\lambda = 1$  dominante y cuyo vector propio con la suma de sus componentes uno es único.

Una forma de calcular dicho vector propio es mediante el método de las potencias, en el que se calculan las potencias sucesivas

$$\mathbf{x}_1 = A^* \mathbf{x}_0, \mathbf{x}_2 = A^* \mathbf{x}_1 = (A^*)^2 \mathbf{x}_0, \mathbf{x}_3 = A^* \mathbf{x}_2 = (A^*)^3 \mathbf{x}_0, \dots, \mathbf{x}_k = A^* \mathbf{x}_{k-1} = (A^*)^k \mathbf{x}_0, \dots$$

siendo  $\mathbf{x}_k$  el vector propio buscado.

Notemos que partimos de un vector inicial  $\mathbf{x}_0$ , que en nuestro caso es el vector inicial cuyas componentes son todas iguales a  $1/n$ .

En nuestro ejemplo calculamos

$$\mathbf{x}_1 = A^* \mathbf{x}_0, \mathbf{x}_2 = (A^*)^2 \mathbf{x}_0, \mathbf{x}_3 = (A^*)^3 \mathbf{x}_0, \dots, \mathbf{x}_{50} = (A^*)^{50} \mathbf{x}_0, \dots$$

cuyo resultado es:

$$\mathbf{x}_{50} = \begin{bmatrix} 0.0625 \\ 0.0625 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.1250 \\ 0.0625 \\ 0.0937 \\ 0.1875 \\ 0.0938 \\ 0.0625 \end{bmatrix}. \quad (6.6)$$

Este es el vector que hemos venido llamando vector pagerank.

Si observamos el resultado que nos proporciona el vector  $\mathbf{x}_{50}$ , notamos que disponemos de una clasificación ordenada de los nodos. Concretamente, el nodo que sale ganador, por llamarlo de alguna manera, en este ranking es el nodo 8, que corresponde con la plaza  $P_8$ . Este resultado es absolutamente lógico si observamos en la figura que se trata del nodo que tiene más conexiones con el resto. Después siguen en el ranking los nodos que tienen un mayor número de conexiones.

Construimos ahora para este ejemplo la matriz dada por (6.8),

$$M = (1 - \alpha) * A^* + \alpha \cdot 1/10 \cdot I_{10},$$

que es

$$M = \begin{bmatrix} 0.015 & 0.015 & 0.015 & 0.015 & 0.015 & 0.015 & 0.2983 & 0.1567 & 0.015 & 0.015 \\ 0.015 & 0.015 & 0.015 & 0.015 & 0.2275 & 0.44 & 0.015 & 0.015 & 0.015 & 0.015 \\ 0.015 & 0.015 & 0.015 & 0.2275 & 0.015 & 0.015 & 0.015 & 0.1567 & 0.2983 & 0.44 \\ 0.015 & 0.015 & 0.2275 & 0.015 & 0.015 & 0.015 & 0.2983 & 0.1567 & 0.015 & 0.44 \\ 0.015 & 0.44 & 0.015 & 0.015 & 0.015 & 0.44 & 0.015 & 0.1567 & 0.2983 & 0.015 \\ 0.015 & 0.44 & 0.015 & 0.015 & 0.2275 & 0.015 & 0.015 & 0.015 & 0.015 & 0.015 \\ 0.44 & 0.015 & 0.015 & 0.2275 & 0.015 & 0.015 & 0.015 & 0.1567 & 0.015 & 0.015 \\ 0.44 & 0.015 & 0.2275 & 0.2275 & 0.2275 & 0.015 & 0.2983 & 0.015 & 0.2983 & 0.015 \\ 0.015 & 0.015 & 0.2275 & 0.015 & 0.2275 & 0.015 & 0.015 & 0.1567 & 0.015 & 0.015 \\ 0.015 & 0.015 & 0.2275 & 0.2275 & 0.015 & 0.015 & 0.015 & 0.015 & 0.015 & 0.015 \end{bmatrix}.$$

En esta matriz ya se contempla la idea de establecer un factor para que no haya tantas componentes iguales a cero. Así, se está contemplando la posibilidad de desplazarse de una plaza a otra sin que estén conectadas. Tomamos  $\alpha = 0.15$ , por lo que muchos elementos de esta matriz son ahora  $1/10 \cdot 0.15$ . Notemos que la nueva matriz  $M$  sigue con las características básicas que nos permiten aplicar el teorema de Perron-Frobenius.

Calculamos el vector propio asociado al valor propio  $\lambda = 1$  de  $M$ , siguiendo un procedimiento similar al realizado anteriormente, calculando las sucesivas potencias y estudiando su convergencia. El vector propio resultante es:

$$\mathbf{x}_{50} = \begin{bmatrix} 0.0657 \\ 0.0733 \\ 0.1185 \\ 0.1187 \\ 0.1278 \\ 0.0733 \\ 0.0926 \\ 0.1728 \\ 0.0918 \\ 0.0654 \end{bmatrix}. \quad (6.7)$$

Podemos comparar el vector dado por (expresión (6.7)) con el que obtuvimos en (expresión (6.6)); se advierte claramente que los resultados son muy similares, lo que resulta lógico dadas las características del ejemplo que estamos estudiando.

Notemos que cuando definimos la matriz  $B$  en la ecuación (expresión (6.8)), asignamos la misma probabilidad a que un usuario se desplazara desde una plaza a cualquier otra; no tuvimos en cuenta ningún criterio de prioridad a la hora de establecer preferencias en los posibles caminos. Sin embargo, nos cuestionamos qué ocurriría con la ordenación o pagerank si establecemos algún tipo de distribución de probabilidad no uniforme entre los nodos.

En nuestro caso, podemos considerar que existen algunos nodos (plazas) que tienen más probabilidad de ser visitados que otros, por las razones que sea (importancia estratégica, localización de organismos públicos, centro de cultura u ocio, etc).

En consecuencia, consideraremos que la matriz dada por (expresión (6.8)) es ahora modificada y que la que calculamos es

$$M' = (1 - \alpha)A^* + \alpha P, \quad (6.8)$$

donde  $A$  es la matriz de transición,  $\alpha$  es un factor entre 0 y 1 (*damping factor*) y  $P$  es la matriz en la que el elemento  $p_{ij}$  representa la probabilidad de que un usuario que

se encuentra en la plaza  $P_j$  desee desplazarse a la plaza  $P_i$ . Notemos que ahora esta probabilidad no es la misma en todos los casos.

Supongamos en nuestro ejemplo que las plazas  $P_5$ ,  $P_7$  y  $P_8$  son las más concurridas puesto que en ellas se encuentran monumentos de gran atractivo turístico. Entonces nuestra matriz  $P$  la definimos como:

$$P = \begin{bmatrix} 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 \\ 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 & 0.10 \\ 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 \\ 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 & 0.20 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \\ 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 & 0.05 \end{bmatrix},$$

donde se observa que en las filas correspondientes a las plazas  $P_5$ ,  $P_7$  y  $P_8$  los valores de la probabilidad son mayores que en el resto.

Una vez definida  $P$ , tenemos que

$$M' = (1 - \alpha)A + \alpha P =$$

$$\begin{bmatrix} 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.2908 & 0.1492 & 0.0075 & 0.0075 \\ 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.2200 & 0.4325 & 0.0075 & 0.0075 & 0.0075 & 0.0075 \\ 0.0075 & 0.0075 & 0.0075 & 0.2200 & 0.0075 & 0.0075 & 0.0075 & 0.1492 & 0.2908 & 0.4325 \\ 0.0075 & 0.0075 & 0.2200 & 0.0075 & 0.0075 & 0.0075 & 0.2908 & 0.1492 & 0.0075 & 0.4325 \\ 0.0300 & 0.4550 & 0.0300 & 0.0300 & 0.0300 & 0.4550 & 0.0300 & 0.1717 & 0.3133 & 0.0075 \\ 0.0150 & 0.4400 & 0.0150 & 0.0150 & 0.2275 & 0.0150 & 0.0150 & 0.0150 & 0.0150 & 0.0075 \\ 0.4550 & 0.0300 & 0.0300 & 0.2425 & 0.0300 & 0.0300 & 0.0300 & 0.1717 & 0.0300 & 0.0075 \\ 0.4550 & 0.0300 & 0.2425 & 0.2425 & 0.2425 & 0.0300 & 0.3133 & 0.0300 & 0.3133 & 0.0075 \\ 0.0075 & 0.0075 & 0.2200 & 0.0075 & 0.2200 & 0.0075 & 0.0075 & 0.1492 & 0.0075 & 0.0075 \\ 0.0075 & 0.0075 & 0.2200 & 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.0075 & 0.0075 \end{bmatrix}.$$

Volvemos a calcular el vector propio asociado ahora a la matriz  $M'$  y obtenemos

$$\mathbf{x}_{50} = \begin{bmatrix} 0.0526 \\ 0.0568 \\ 0.0857 \\ 0.0895 \\ 0.1157 \\ 0.0609 \\ 0.0869 \\ 0.1530 \\ 0.0709 \\ 0.0435 \end{bmatrix}.$$

#### 6.4.4 Analizando una red terrorista

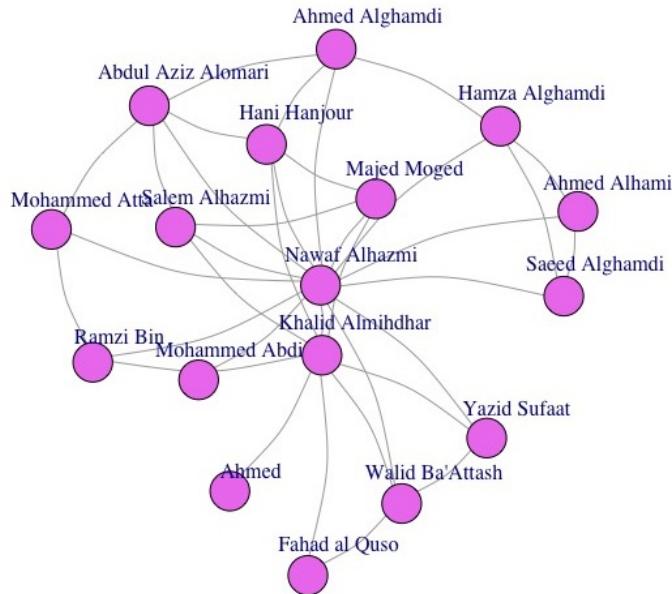
En la mañana del 11 de septiembre de 2001, cuatro aviones comerciales fueron secuestrados y convertidos en proyectiles por terroristas de Al Qaeda. Dos de ellos impactaron contra las dos torres del World Trade Center de Nueva York, un tercero contra el ala oeste del pentágono en Washington DC y el último, cuyo objetivo era la Casa Blanca, fue desviado por los ocupantes del mismo, estrellándose en Pensilvania. Los terroristas que abordaron

los aviones fueron dirigidos por Khalid Sheik Mohammed, de nacionalidad paquistaní, que fue capturado en 2003.

Después de lo ocurrido se inició una investigación exhaustiva participando la Agencia de Seguridad Nacional Americana (NSA) quien ya había advertido anteriormente que disponía de indicios que se estaba preparando un ataque terrorista.

Se determinó la identidad de los 19 terroristas que participaron en los ataques con conexiones entre ellos. Eran los siguientes: Nawaf Alhzami, Khalid Almihdar, Majed Moged, Salem Alhzami, Hani Hanjour, Ramzi Bin, Mohammed Atta, Abdul Aziz Alomari, Ahmed Alghamdi, Hamza Alghamdi, Ahmed Alhami, Saeed Alghamdi, Mohammed Abdi, Yazid Sufaat, Waliid Ba'Attash, Fahad al Quso, Ahmed.

El siguiente paso fue establecer las conexiones y contactos entre ellos para acabar de establecer la red que tenían para comunicarse y proceder con los diferentes operativos. El grafo que muestra las conexiones en esta red lo vemos en la figura 6.18.

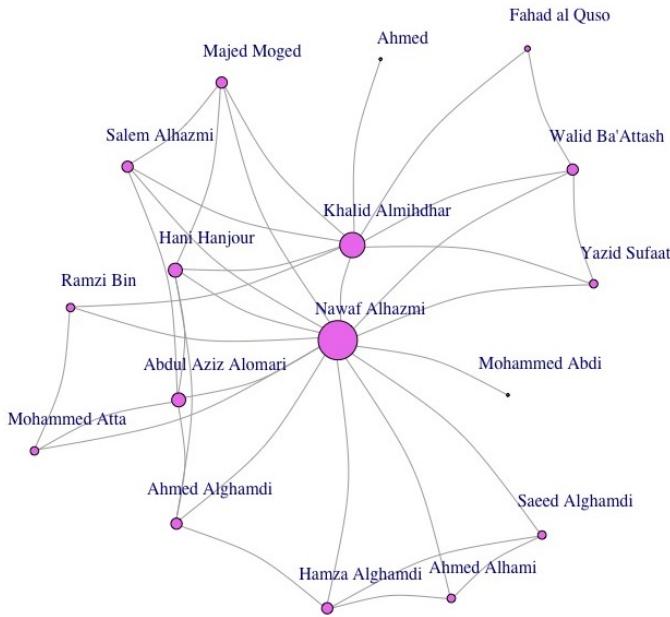


**Figura 6.18** Grafo de la red terrorista del 11 de septiembre..

A la vista del grafo que hemos dibujado en la figura 6.18, se observa que existen dos nodos fundamentales si tenemos en cuenta únicamente los enlaces entre ellos y en esta trama, como son Nawaf Alhzami y Khalid Almihdar. Parecen los verdaderos cabecillas de la organización.

Si representamos el grafo dibujando el tamaño de los nodos en función del grado que poseen, se obtiene una imagen mucho más clarificadora de lo que queremos decir, como

se aprecia en la figura 6.19.



**Figura 6.19** Grafo de la red terrorista del 11 de septiembre en función de los contactos de cada nodo..

Parece claro entonces la importancia de los nodos Nawaf Alhazmi y Khalid Almihdhar en la red. Sin embargo, es conveniente confirmar nuestras sospechas mediante el cálculo de la centralidad de los nodos.

Podemos construir una tabla donde visualicemos todas las medidas que vamos a calcular relacionadas con la centralidad de los nodos.

De la sección 6.4.4 extraemos un ránking de los nodos más importantes, de acuerdo con el criterio de centralidad o medida aplicada. Así, los tres primeros nodos del ranking son los siguientes, para cada una de las centralidades calculadas.

**Grado** 1- Nawaf Alhazmi, 2- Khalid Almihdhar, 3- Hani Hanjour / Abdul Aziz.

**Closeness** 1- Nawaf Alhazmi, 2- Khalid Almihdhar, 3- Hani Hanjour.

**Betweenness** 1- Nawaf Alhazmi, 2- Khalid Almihdhar, 3- Walid Ba'Attash.

**PageRank** 1- Nawaf Alhazmi, 2- Khalid Almihdhar, 3- Abdul Aziz Alomari.

A la vista de estos resultados, la conclusión es muy clara ya que se observa perfectamente que los dos componentes fundamentales de la trama son Nawaf Alhazmi y Khalid

**Tabla 6.1** Datos sobre la centralidad de los nodos de la red terrorista..

Nodo	Grado	Closeness	Betweenness	PageRank
Nawaf Alhazmi	14	0.0555	69.58	0.1800
Khalid Almihdhar	9	0.0434	29.83	0.1217
Majed Moged	4	0.0357	0.25	0.0538
Salem Alhazmi	4	0.0357	1.25	0.0539
Hani Hanjour	5	0.0370	3.08	0.0654
Ramzi Bin	3	0.0345	1.33	0.0435
Mohammed Atta	3	0.0323	0.50	0.0433
Abdul Aziz Alomari	5	0.0345	2.25	0.0661
Ahmed Alghamdi	4	0.0333	1.00	0.0541
Hamza Alghamdi	4	0.0333	1.00	0.0563
Ahmed Alhami	3	0.0323	0.00	0.0443
Saeed Alghamdi	3	0.0323	0.00	0.0443
Mohammed Abdi	1	0.0303	0.00	0.0198
Yazid Sufaat	3	0.0345	0.00	0.0434
Walid BaÁttash	4	0.0357	3.92	0.0574
Fahad al Quso	2	0.0270	0.00	0.0325
Ahmed	1	0.0263	0.00	0.0203

Almihdhar. Después de ellos, hay alguna diferencia entre los elementos que ocupan el tercer lugar en importancia. por ejemplo, si nos fijamos en la centralidad betweenness o de intermediación, Wali BaÁttash pasa por delante de Hani hanjour, aunque la diferencia entre ellos es muy pequeña. Igualmente, si tomamos como referencia la centralidad de PageRank, Abdul Aziz Alomari se encuentra por delante de Hani Hanjour, aunque las diferencias vuelven a ser muy pequeñas.

Podríamos afirmar que después de los dos máximos responsables del operativo, se encuentran los nodos Hani Hanjour, Abdul Aziz Alomari y Walid BaÁttash.

Arne Frick, Andreas Ludwig, Heiko Mehldau: A Fast Adaptive Layout Algorithm for Undirected Graphs, Proc. Graph Drawing 1994, LNCS 894, pp. 388-403, 1995.

Kamada, T. and Kawai, S.: An Algorithm for Drawing General Undirected Graphs. Information Processing Letters, 31/1, 7–15, 1989.

Ron Davidson, David Harel: Drawing Graphs Nicely Using Simulated Annealing. ACM Transactions on Graphics 15(4), pp. 301-331, 1996.

Cox, T. F. and Cox, M. A. A. (2001) Multidimensional Scaling. Second edition. Chapman and Hall.



## **Bibliografía**

---