MATEMÁTICAS I

Licenciatura de Administración y Dirección de Empresas

Fernando Casas, María Vicenta Ferrer, Pura Vindel

Departament de Matemàtiques

Universitat Jaume I

Estas notas constituyen el material básico empleado en el curso de Matemáticas I de la Licenciatura de Administración y Dirección de Empresas en la Universitat Jaume I. El contenido es una introducción al álgebra lineal y matricial, e incluye también algunas aplicaciones a la economía. Estas notas también incorporan los enunciados de los problemas habitualmente planteados y resueltos durante el curso.

Índice general

Capítulo 1

MATRICES Y DETERMINANTES

En la mayoría de los modelos matemáticos usados en Economía aparecen sistemas de ecuaciones algebraicas que es preciso resolver. Si las ecuaciones son lineales el estudio de estos sistemas pertenece a la rama de las matemáticas denominada álgebra lineal.

La parte más importante de la economía que usa los sistemas de ecuaciones lineales es el análisis input-output que trabaja con modelos de Leontief. Para comprender y manejar estos modelos es conveniente familiarizarse con una serie de conceptos como son los de matrices, determinantes y vectores. No obstante, nos gustaría recalcar que la utilidad del álgebra lineal va más allá de su capacidad para resolver sistemas de ecuaciones lineales, ya que se usa extensamente en muchas otras materias relacionadas con la ciencia y la técnica, además de la economía. En particular, muchos de los métodos del álgebra lineal son útiles en la teoría de optimización, la estadística y la econometría.

1.1. Matrices

1.1.1. Introducción

Vivimos en un mundo complejo de recursos finitos, demandas mutuamente competitivas y flujos de información que han de ser analizados para asignar los recursos de la mejor manera posible para satisfacer nuestras necesidades. Cualquier herramienta que haga más fácil entender y usar tal información es muy conveniente.

Considérese, por ejemplo, un inventario de camisetas en una sección de un gran almacén. Se tienen camisetas de tres diferentes tamaÒos y cinco colores, y cada noche el supervisor de la sección prepara un inventario de las existencias para la gestión. Un párrafo de dicho inventario podría tener la forma siguiente:

“**Camisetas...**. Nueve amarillas de talla S y cinco amarillas de talla M; ocho S de color verde y seis M verdes; las de tamaño L casi se han agotado pues sólo quedan tres rojas, una rosa y dos negras; también tenemos tres M rosas, cinco M rojas, una M negra y siete S negras...”.

Este informe no es muy fácil de analizar. En particular, se ha de leer el párrafo entero para determinar el número de camisetas pequeñas de color rojo que hay actualmente en stock. En cambio, la tabla rectangular de datos presentada a continuación resume la información mucho mejor:

	Rosa	Amarillo	Verde	Rojo	Negro
S	0	9	8	0	7
M	3	5	6	5	1
L	1	0	0	3	2

Así, de un vistazo, nos percatamos de que no hay camisetas de talla S rojas en stock.

1.1.2. Definición

Llamaremos **matriz** de orden $p \times n$ a un tablero rectangular de elementos de un determinado conjunto $\mathbb{K}$ organizados en p filas y n columnas, considerado como una entidad y delimitado por paréntesis o corchetes.

Las matrices se suelen denotar por letras mayúsculas y las entradas que las componen, denominadas **elementos** o **coeficientes**, por letras minúsculas con subíndices que indican el lugar que ocupan en la matriz. El primer subíndice especifica la posición en las filas y el segundo subíndice su posición en las columnas. Así, a_{12} denota el elemento en la primera fila y segunda columna de la matriz A. En general, una matriz $p \times n$ tiene la forma

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & . & . & . & a_{1n} \\ a_{21} & a_{22} & a_{23} & . & . & . & a_{2n} \\ a_{31} & a_{32} & a_{33} & . & . & . & a_{3n} \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ a_{p1} & a_{p2} & a_{p3} & . & . & . & a_{pn} \end{pmatrix}$$

De manera abreviada se suele escribir

$$A = (a_{ij}), \qquad i : \text{índice de fila}, \qquad j : \text{índice de columna}.$$

Se dice que dos matrices son $i$guales si tienen el mismo tamaño y los elementos correspondientes son iguales.

El conjunto de todas las matrices de orden $p \times n$ se denota por $\mathcal{M}_{p\times n}(\mathbb{K})$, y cuando no hay ambig¸edad en cuanto al conjunto $\mathbb{K}$, simplemente por $\mathcal{M}_{p\times n}$.

Las matrices se usan en muchos contextos, en particular para organizar tableros de datos, para la resolución de ecuaciones lineales, etc. Como curiosidad histórica, digamos que el matemático inglés J. Sylvester (1814-1897) fue el primero que usó el término "matriz" en 1850, para distinguir las matrices de los determinantes (los cuales se estudiarán más adelante). De hecho, la intención era que el término "matriz" tuviera el significado de 'madre de los determinantes'.

1.1.3. Ejemplos

(a) $A = \begin{pmatrix} 2 & 1 & -1 \\ 3 & 7 & 0 \end{pmatrix}$ es una matriz de *tamaño* u *orden* 2×3, siendo $a_{11} = 2$, $a_{12} = 1$, etc.

(b) La matriz $A \in \mathcal{M}_{4\times 3}(\mathbb{R})$ tal que $a_{ij} = 2i - j$ se escribe explícitamente como

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 2 & 1 \\ 5 & 4 & 3 \\ 7 & 6 & 5 \end{pmatrix}.$$

(c) Organiza los siguientes datos en forma de matriz: una tienda tiene dos almacenes como proveedores de electrodomésticos; el primer almacén tiene dos lavadoras, dos cocinas y tres neveras en stock; el segundo tiene cuatro cocinas, tres lavadoras y una nevera.

Solución. La matriz sería en este caso

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

si disponemos los dos almacenes en las filas y los diferentes artículos en las columnas.

(d) Consideremos la matriz $A \in \mathcal{M}_{2\times 3}(\mathbb{R}_2[x])$ dada por

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

Obsérvese que los coeficientes de una matriz no tienen por qué ser necesariamente números reales. En este ejemplo los elementos son polinomios de grado a lo sumo 2.

A partir de ahora trabajaremos, a menos que se especifique lo contrario, con $\mathbb{K} = \mathbb{R}$, es decir, con matrices cuyos elementos son números reales.

1.1.4. Tipos de matrices

Ciertos tipos de matrices aparecen tan frecuentemente que es preferible darles una denominación especial y tratarlas separadamente.

(a) Una *matriz fila* es una matriz con sólo una fila. Una *matriz columna* sólo posee una columna. Por ejemplo, $A = \begin{pmatrix} 1 & 2 & -1 \end{pmatrix}$ es una matriz fila, mientras que $B = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ es una matriz columna.

(b) Una *submatriz* de una matriz A es una matriz que se obtiene de A quitando cualquier número de filas y/o columnas de A. En particular, si

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} \tag{1.1}$$

entonces

$$B = \begin{pmatrix} 10 & 12 \\ 14 & 16 \end{pmatrix}, \qquad C = \begin{pmatrix} 2 & 3 & 4 \\ 10 & 11 & 12 \end{pmatrix}$$

son submatrices de A. En el primer caso se han quitado la primera y segunda filas, y la primera y tercera columnas. En el segundo se han quitado la segunda y cuarta filas y la primera columna.

(c) Una matriz se dice que está *particionada* si está dividida en submatrices por medio de líneas horizontales y verticales entre filas y columnas. Variando la posición de estas líneas horizontales y verticales es posible particionar una matriz de diferentes formas. Así,

$$\left(\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ \hline 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{array}\right) \quad \text{y} \quad \left(\begin{array}{c|cc|c} 1 & 2 & 3 & 4 \\ \hline 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{array}\right)$$

son ejemplos de dos diferentes particiones de la matriz A dada en (1.1).

(d) Una matriz es *cuadrada* si tiene el mismo número de filas que de columnas. En general, una matriz cuadrada de orden n tiene la forma

$$\begin{pmatrix} a_{11} & a_{12} & . & . & . & a_{1n} \\ a_{21} & a_{22} & . & . & . & a_{2n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ a_{n1} & a_{n2} & . & . & . & a_{nn} \end{pmatrix}$$

con los elementos $a_{11}, a_{22}, \ldots, a_{nn}$ formando la *diagonal principal*.

(e) Una *matriz diagonal* es una matriz cuadrada cuyos elementos fuera de la diagonal principal son todos cero. Una matriz identidad, denotada como I, es una matriz diagonal que tiene todos sus elementos de la diagonal principal iguales a 1. Así, la matriz identidad 2×2 es

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

(f) Se define la *matriz cero* O como la matriz cuyos elementos son todos cero.

(g) Una matriz $A = (a_{ij})$ es *triangular superior* si $a_{ij} = 0$ para $i > j$, esto es, todos sus elementos por debajo de la diagonal principal son cero. Si $a_{ij} = 0$ para $i < j$, esto es, si todos los elementos por encima de la diagonal principal son cero, entonces A es *triangular inferior*. Ejemplos de matrices triangular superior y triangular inferior son, respectivamente,

$$\begin{pmatrix} -1 & 2 & 4 & 1 \\ 0 & 1 & 3 & 8 \\ 0 & 0 & 2 & 12 \\ 0 & 0 & 0 & -5 \end{pmatrix} \quad \text{y} \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 5 & 6 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 13 & 14 & 15 & 16 \end{pmatrix}$$

(h) Se define la *traspuesta* de una matriz A, denotada por A^T o bien A', como aquella matriz que se obtiene intercambiando las filas y las columnas de A (preservando su ordenación). Formalmente, si $A = (a_{ij})$ es una matriz $n \times p$, entonces la traspuesta de A, $A^T = (a_{ij}^T)$ es una matriz $p \times n$, donde $a_{ij}^T = a_{ji}$. Por ejemplo, si

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 3 & 2 & 1 \\ 5 & 4 & 3 \\ 7 & 6 & 5 \end{pmatrix} \quad \text{entonces} \quad A^T = \begin{pmatrix} 1 & 3 & 5 & 7 \\ 0 & 2 & 4 & 6 \\ -1 & 1 & 3 & 5 \end{pmatrix}.$$

De acuerdo con esto, es evidente que $(A^T)^T = A$.

(i) Una matriz cuadrada $A = (a_{ij})$ es simétrica si es igual a su propia traspuesta, es decir, $A = A^T$; entonces se verifica que $a_{ij} = a_{ji}$. Una matriz es antisimétrica si $A^T = -A$, donde $-A = (-a_{ij})$. Por ejemplo, $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{pmatrix}$ es una matriz simétrica, mientras que $B = \begin{pmatrix} 0 & 2 & -3 \\ -2 & 0 & 1 \\ 3 & -1 & 0 \end{pmatrix}$ es antisimétrica.

(j) Una matriz se dice que está en forma *escalonada* si satisface las cuatro condiciones siguientes:

1. Todas las filas de ceros aparecen por debajo de filas no nulas cuando ambos tipos están presentes en la matriz.
2. El primer elemento no nulo en cualquier fila no nula es 1.
3. Todos los elementos que están en la misma columna pero en las filas situadas por debajo del primer elemento no nulo de una fila no nula son cero.
4. El primer elemento no nulo de cualquier fila no nula aparece en una columna posterior (hacia la derecha) que el primer elemento no nulo en cualquier fila precedente.

Las matrices en forma escalonada son extrordinariamente útiles a la hora de resolver sistemas de ecuaciones lineales. Usaremos estas matrices ampliamente en las secciones siguientes, pero de momento veamos varios ejemplos.

$$A = \begin{pmatrix} 1 & 1 & -2 & 4 & 7 \\ 0 & 0 & -6 & 5 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

no está en forma escalonada, debido a que el primer elemento no nulo en la segunda fila no es 1. Si a_{23} fuera 1 en vez de -6, entonces la matriz estaría en forma escalonada.

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

no está en forma escalonada debido a que la segunda fila es una fila de ceros y aparece antes que una fila no nula; si se intercambiaran las filas 2 y 3, sí estaría en forma escalonada.

$$C = \begin{pmatrix} 1 & 2 & -2 & 4 \\ 0 & 0 & 1 & 7 \\ 0 & 1 & 0 & 5 \end{pmatrix}$$

no está en forma escalonada, porque el primer elemento no nulo en la fila 2 aparece en una columna posterior (columna 3) que el primer elemento no nulo en la fila 3. Si se intercambiaran las filas 2 y 3 la matriz resultante sí estaría en forma escalonada.

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

no está en forma escalonada, pues el primer elemento no nulo en la fila 2 aparece en la tercera columna, y todos los elementos situados debajo de él no son nulos. Si d_{33} fuera cero en vez de 1, entonces sí estaría en forma escalonada.

1.1.5. Operaciones con matrices

Para muchas de las aplicaciones de las matrices es interesante conocer qué operaciones se pueden realizar con ellas, es decir, conocer el álgebra de las matrices.

Suma de matrices Dadas dos matrices del mismo tamaño $A = (a_{ij})$, $B = (b_{ij}) \in \mathcal{M}_{p\times n}$, se *define* la suma de A y B como la matriz

$$A + B = (a_{ij} + b_{ij}) \in \mathcal{M}_{p\times n}.$$

Ejemplo. Si en el ejemplo (c) de la sección 1.1.3 llega por la noche un envío a los almacenes dado por la matriz de transporte

$$B = \begin{pmatrix} 9 & 0 & 0 \\ 3 & 3 & 3 \end{pmatrix},$$

la nueva matriz de existencias vendría dada por

$$A + B = \begin{pmatrix} 11 & 2 & 3 \\ 6 & 7 & 4 \end{pmatrix}.$$

La suma de matrices cumple las siguientes propiedades, todas ellas fácilmente demostrables a partir de la definición:

1. La suma está bien definida (es consistente). Esto significa que para todas las matrices $A, B \in \mathcal{M}_{p\times n}$ existe una única matriz $A + B \in \mathcal{M}_{p\times n}$.

2. **Asociativa.** $A + (B + C) = (A + B) + C$

3. Existencia de **elemento neutro** O: $A + 0 = 0 + A = A$, donde 0 representa una matriz del mismo tamaño que A con todos sus elementos nulos.

4. Existencia de **elemento opuesto:** $A + (-A) = (-A) + A = 0.$

5. **Conmutativa:** $A + B = B + A$

El conjunto $\mathcal{M}_{p\times n}$ de todas las matrices $p \times n$ con la operación $+$ se dice que tiene entonces estructura de *grupo abeliano o conmutativo*. Además, es fácil comprobar que

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

Nota. Se suele denotar por $A - B = A + (-B)$.

Producto de una matriz por un escalar Dada una matriz $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$ y un escalar $\alpha \in \mathbb{R}$, definimos el producto del número α por la matriz A a la nueva matriz

$$\alpha A = (\alpha a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R}).$$

Ejemplo. $A = \begin{pmatrix} 2 & 1 \\ 3 & -7 \end{pmatrix}$, $\alpha = -2 \quad \Rightarrow \quad \alpha A = \begin{pmatrix} -4 & -2 \\ -6 & 14 \end{pmatrix}$.

Se tienen las siguientes propiedades, siendo α y β escalares arbitrarios:

1. Este producto de matriz por escalar está bien definido, es decir, dada cualquier matriz $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$ y cualquier escalar $\alpha \in \mathbb{R}$, existe una única matriz $\alpha A =\in \mathcal{M}_{p\times n}(\mathbb{R})$.

2. **Distributiva respecto de los escalares:** $(\alpha + \beta)\, A = \alpha A + \beta A$.

3. **Distributiva respecto de las matrices:** $\alpha\,(A + B) = \alpha A + \alpha B$.

4. **Asociativa mixta:** $(\alpha\beta)A = \alpha(\beta A)$.

5. $1.A = A$.

Como veremos en el tema siguiente, el conjunto $\mathcal{M}_{p\times n}(\mathbb{R})$ satisface los axiomas de un espacio vectorial. Además se cumple que

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

Multiplicación de matrices La multiplicación de matrices es la primera operación donde nuestra intuición falla: en primer lugar, dos matrices *no* se multiplican elemento a elemento; en segundo lugar, no siempre es posible multiplicar dos matrices. Nuestro propósito al introducir un nuevo concepto, como el de matriz, es usarlo para intentar ampliar nuestro conocimiento de los fenómenos del mundo real y resolver problemas que previamente eran difíciles de abordar. Una matriz vista solamente como un tablero de números no constituye en verdad ningún concepto novedoso. Las operaciones sobre esos tableros, tales como la suma de matrices o el producto de una matriz por un número son nuevas, pero de hecho no son más que extensiones naturales de operaciones análogas sobre los números reales. Si esperamos usar matrices para analizar problemas de una manera diferente, hemos de introducir algo realmente nuevo, y ese aspecto nuevo es la forma en que se multiplican las matrices.

Una posible motivación para la multiplicación de matrices puede venir del deseo de resolver sistemas de ecuaciones lineales con la misma facilidad como se hace con una ecuación en una variable. Una ecuación lineal en una variable tiene la forma general

$$\text{(constante).(incógnita)} = \text{constante}$$

y resolvemos simplemente dividiendo la ecuación completa por la constante multiplicativa de la izquierda. Queremos de alguna forma imitar este proceso cuando se tienen varias ecuaciones con varias incógnitas. Idealmente nos gustaría tener una sola ecuación de la forma

$$\text{(paquete de constantes).(paquete de incógnitas)} = \text{paquete de constantes}$$

Por ejemplo, para el sistema de dos ecuaciones con dos incógnitas

$$\begin{array}{rcl} 2x+3y & = & 10 \\ 4x+5y & = & 20 \end{array} \tag{1.2}$$

combinando todos los coeficientes de las incógnitas en una *matriz de coeficientes*, todas las incógnitas en una matriz columna y las constantes de la derecha de cada ecuación en otra matriz columna, generamos el sistema matricial

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \end{pmatrix} \tag{1.3}$$

La idea es entonces definir la multiplicación de matrices de manera que el sistema (1.3) sea equivalente al sistema (1.2); esto es, queremos definir la multiplicación de forma que

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x+3y \\ 4x+5y \end{pmatrix} \tag{1.4}$$

de forma que el sistema (1.3) sea

$$\begin{pmatrix} 2x+3y \\ 4x+5y \end{pmatrix} = \begin{pmatrix} 10 \\ 20 \end{pmatrix}. \tag{1.5}$$

Definiremos el producto AB de dos matrices A y B cuando el número de columnas de A es igual al número de filas de B. El resultado será una matriz que tiene el mismo número de filas que A y el mismo número de columnas que B. Así, si A y B son

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

entonces el producto AB está definido, y será matriz 2×4. Por contra, el producto BA no está definido, dado que el número de columnas en B es diferente del número de filas en A.

Si tenemos en general una matriz A $n \times r$ y una matriz B de orden $r \times p$, tomamos como motivación la multiplicación en (1.3) y definimos el elemento ij del producto

$AB = C = (c_{ij})$ multiplicando los elementos de la fila i de A por los correspondientes elementos en la columna j de B y sumando los resultados. Esto es,

$$AB = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} \\ a_{21} & a_{22} & \dots & a_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nr} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1p} \\ b_{21} & b_{22} & \dots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r1} & b_{r2} & \dots & b_{rp} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{np} \end{pmatrix} = C$$

donde

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ir}b_{rj} = \sum_{k=1}^{r} a_{ik}b_{kj}.$$

En particular, c_{11} se obtiene multiplicando los elementos en la primera fila de A por los correspondientes elementos en la primera columna de B y sumándolos; por consiguiente,

$$c_{11} = a_{11}b_{11} + a_{12}b_{21} + \cdots + a_{1r}b_{r1};$$

el elemento c_{12} se obtiene multiplicando los elementos en la primera fila de A por los correspondientes elementos en la segunda columna de B y sumándolos,

$$c_{12} = a_{11}b_{12} + a_{12}b_{22} + \cdots + a_{1r}b_{r2},$$

y así sucesivamente.

Ejemplo. Si en el ejemplo (c) de la sección 1.1.3 el precio de venta al público de una lavadora es de 300 euros, el de una cocina es de 120 euros y el de una nevera de 450 euros, entonces la matriz

$$AX = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 4 & 1 \end{pmatrix} \begin{pmatrix} 120 \\ 300 \\ 450 \end{pmatrix} = \begin{pmatrix} 2310 \\ 2010 \end{pmatrix},$$

donde X es la matriz columna con el precio de venta de cada tipo de electrodoméstico, nos propociona los ingresos de cada almacén si vendieran todos los electrodomésticos que tienen en stock.

La multiplicación de matrices tiene las siguientes propiedades:

1. El producto de matrices es consistente, es decir, dadas cualesquiera matrices $A = (a_{ij}) \in \mathcal{M}_{n\times r}(\mathbb{R})$ y $B = (b_{ij}) \in \mathcal{M}_{r\times p}(\mathbb{R})$, existe una única matriz $AB \in \mathcal{M}_{n\times p}(\mathbb{R})$.

2. **Asociativa:** $A\,(BC) = (AB)\,C$.

3. **Distributiva a derecha:** $(A + B)\,C = AC + BC$.

4. **Distributiva a izquierda:** $C\,(A + B) = CA + CB$.

5. $I_pA = A$ y $AI_n = A$, siendo A una matriz $p \times n$ e I_j la matriz identidad de orden $j \times j$.

6. $(AB)^T = B^TA^T$.

Es importante recalcar, a tenor de los comentarios precedentes, que *no* se verifica la propiedad conmutativa, es decir, en general $AB \neq BA$.
Ejemplo. Calcular AB y BA para

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}, \qquad B = \begin{pmatrix} -7 & -8 \\ 9 & 10 \\ 0 & -11 \end{pmatrix}.$$

Solución.

$$AB = \begin{pmatrix} 11 & -21 \\ 17 & -48 \end{pmatrix}, \qquad BA = \begin{pmatrix} -39 & -54 & -69 \\ 49 & 68 & 87 \\ -44 & -55 & -66 \end{pmatrix}.$$

La multiplicación de matrices también carece de otras propiedades familiares, además de la conmutatividad. Sabemos que con números reales, si el producto $ab = 0$, entonces o bien $a = 0$ o bien $b = 0$ o bien los dos son cero. Esto no es cierto, en general, para las matrices: existen matrices para las cuales $AB = O$ sin que A ni B sean cero. Por ejemplo,

$$A = \begin{pmatrix} 2 & 6 \\ 3 & 9 \end{pmatrix}, \qquad B = \begin{pmatrix} 3 & -6 \\ -1 & 2 \end{pmatrix}.$$

También, en general, la ecuación $AB = AC$ *no* implica que $B = C$. Por ejemplo,

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

1.2. Determinantes

1.2.1. Introducción

Toda matriz cuadrada tiene asociado un escalar, llamado su *determinante*. Hasta hace poco los determinantes jugaban un papel destacado en el estudio del álgebra lineal. Los determinantes eran usados para calcular inversas de matrices, para resolver sistemas de ecuaciones lineales, etc. Hoy día, en su lugar se utilizan otras técnicas, a menudo basadas en operaciones elementales sobre las filas de la matriz, las cuales son más eficientes y mejor adaptadas al cálculo computacional.

Los determinantes se definen en términos de permutaciones sobre enteros positivos. La teoría es complicada, pero una vez completada, da lugar a métodos más simples para calcular determinantes. Nosotros no desarrollaremos la teoría, sino que nos limitaremos a estudiar algunos métodos de cálculo.

Los determinantes se definen *sólo* para matrices cuadradas. Dada una matriz cuadrada A, usaremos $\det(A)$ o bien $|A|$ para denotar del determinante de A. Si la matriz puede ser escrita explícitamente, designaremos su determinante remplazando los paréntesis por líneas verticales. Por ejemplo, si

$$A = \begin{pmatrix} 2 & 5 & 9 \\ 4 & 0 & 8 \\ 1 & 3 & 7 \end{pmatrix} \tag{1.6}$$

entonces

$$\det(A) = \begin{vmatrix} 2 & 5 & 9 \\ 4 & 0 & 8 \\ 1 & 3 & 7 \end{vmatrix} \tag{1.7}$$

Repetimos una vez más que (1.6) y (1.7) representan estructuras completamente diferentes. La expresión (1.6) es una matriz, un tablero rectangular de elementos, mientras que (1.7) representa un escalar, un número asociado a la matriz (1.6).

1.2.2. Definición

En realidad, como hemos dicho, la definición de determinante es un poco ardua, de modo que aquí procederemos de un modo más pragmático. Así, el determinante de una matriz 1×1 (a_{11}) se define como el escalar a_{11}. Si A es una matriz 2×2,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \text{entonces} \quad \det(A) = a_{11}a_{22} - a_{12}a_{21}$$

mientras que si A es 3×3, entonces

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

Nota. Por completitud, el determinante de una matriz $A = (a_{ij})$ de tamaño $n \times n$, es igual a cualquiera de las dos expresiones siguientes:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathcal{S}_n} \varepsilon(\sigma)\, a_{1\sigma(1)}\, a_{2\sigma(2)} \cdots a_{n\sigma(n)} \\ \det(A) &= \sum_{\rho \in \mathcal{S}_n} \varepsilon(\rho)\, a_{\rho(1)1}\, a_{\rho(2)2} \cdots a_{\rho(n)n} \end{aligned}$$

donde $\mathcal{S}_n$ es el conjunto de todas las sustituciones de n elementos, σ y ρ son sustituciones y $\varepsilon(\sigma)$ es $+1$ o -1 dependiendo de si la sustitución tiene un número par o impar de trasposiciones, respectivamente.

Para calcular el determinante de matrices de orden mayor o igual que 3 no recurriremos a la definición, sino que desarrollaremos un método basado en menores y cofactores y reduciremos el determinante de la matriz a una suma de determinantes de orden dos o tres.

Llamaremos *menor* de una matriz A al determinante de cualquier submatriz cuadrada de A. A partir de una matriz cuadrada A se forma un menor quitando el mismo número de filas y columnas de A y calculando el determinante de la submatriz resultante. Así, si

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

entonces $\begin{vmatrix} 2 & 5 \\ 1 & 3 \end{vmatrix}$ y $\begin{vmatrix} 0 & 8 \\ 3 & 7 \end{vmatrix}$ son ambos menores, ya que las matrices $\begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix}$ y $\begin{pmatrix} 0 & 8 \\ 3 & 7 \end{pmatrix}$ son ambas submatrices de A. Por contra, $\begin{vmatrix} 2 & 5 \\ 3 & 7 \end{vmatrix}$ no es un menor, puesto que $\begin{pmatrix} 2 & 5 \\ 3 & 7 \end{pmatrix}$ no es una submatriz de A.

Llamaremos *cofactor* α_{ij} del elemento a_{ij} de una matriz A al escalar obtenido al multiplicar $(-1)^{i+j}$ por el menor obtenido de A quitando la fila i y la columna j. En otras palabras, para calcular el cofactor (también llamado *adjunto*) de un elemento a_{ij} de una matriz A, formamos primero una submatriz de A quitando de A la fila y la columna en la que el elemento está situado, calculamos después el determinante de dicha submatriz y finalmente multiplicamos el determinante por el número $(-1)^{i+j}$.

Ejemplo. Para calcular el cofactor del elemento 4 en la matriz

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

nos fijamos en que el 4 aparece en la segunda fila y primera columna, de modo que $i = 2$, $j = 1$ y $(-1)^{i+j} = (-1)^{2+1} = (-1)^3 = -1$. La submatriz de A obtenida al quitar la segunda fila y primera columna es $\begin{pmatrix} 2 & 3 \\ 8 & 9 \end{pmatrix}$, que tiene por determinante -6. Así pues, el cofactor de 4 es $\alpha_{21} = (-1)(-6) = 6$.

1.2.3. Cálculo de determinantes

Con este concepto podemos calcular el determinante de cualquier matriz cuadrada A $n\times n$. Más concretamente, el determinante de A se puede calcular multiplicando los elementos de una fila (o una columna) por sus cofactores y sumando los productos resultantes. Es decir,

$$\begin{aligned} \det(A) &= a_{i1}\alpha_{i1} + a_{i2}\alpha_{i2} + \cdots + a_{in}\alpha_{in} = \sum_{l=1}^{n} a_{il}\alpha_{il} \\ \det(A) &= a_{1j}\alpha_{1j} + a_{2j}\alpha_{2j} + \cdots + a_{nj}\alpha_{nj} = \sum_{l=1}^{n} a_{lj}\alpha_{lj} \end{aligned}$$

El algoritmo resultante se podría resumir en los siguientes pasos:

Desarrollo de un determinante por los cofactores de una línea

1. Elegir cualquier fila o cualquier columna de la matriz (a gusto del consumidor...)

2. Calcular el cofactor de cada elemento en la fila o columna elegida.

3. Multiplicar cada elemento en la fila o columna elegida por su cofactor y sumar los resultados.

Ejemplo. Calcular $\det(A)$, siendo

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

Desarrollamos arbitrariamente por la segunda columna. Así,

$$\begin{aligned} \det(A) &= 5(-1)^{1+2}\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} + 2(-1)^{2+2}\begin{vmatrix} 3 & 0 \\ 3 & 4 \end{vmatrix} + (-6)(-1)^{3+2}\begin{vmatrix} 3 & 0 \\ -1 & 1 \end{vmatrix} = \\ &= 5(-1)(-4-3) + 2(1)(12-0) + (-6)(-1)(3-0) \\ &= (-5)(-7) + 2(12) + 6(3) = 77. \end{aligned}$$

Si desarrollamos por la primera fila,

$$\begin{aligned} \det(A) &= 3(-1)^{1+1}\begin{vmatrix} 2 & 1 \\ -6 & 4 \end{vmatrix} + 5(-1)^{1+2}\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} + 0(\text{cofactor de } 0) = \\ &= 3(1)(8+6)) + 5(-1)(-4-3) + 0 = 3(14) + (-5)(-7) = 77, \end{aligned}$$

y si se aplica directamente la definición,

$$\det(A) = 24 + 15 + 0 - (0 - 18 - 20) = 39 + 38 = 77.$$

Este ejemplo ilustra dos propiedades importantes del desarrollo de un determinante por cofactores. En primer lugar, el valor de un determinante es el mismo independientemente de qué fila o columna se elige, y en segundo lugar, eligiendo para desarrollar una fila o columna que tenga ceros reduce significativamente el volumen de cálculo requerido.

Ejemplo. Calcular $\det(A)$, siendo

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

En este caso la línea que contiene la mayor cantidad de ceros es la segunda columna, de modo que desarrollamos por ella.

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

Ahora podemos usar el desarrollo por cofactores de cada uno de estos dos determinantes de orden 3 o directamente aplicar la definición. En el primer caso caso se llega a

$$\begin{vmatrix} 1 & 5 & 2 \\ 3 & 4 & 1 \\ -2 & 1 & 3 \end{vmatrix} = 1\begin{vmatrix} 4 & 1 \\ 1 & 3 \end{vmatrix} + 5(-1)^{1+2}\begin{vmatrix} 3 & 1 \\ -2 & 3 \end{vmatrix} + 2(-1)^{1+3}\begin{vmatrix} 3 & 4 \\ -2 & 1 \end{vmatrix} =$$
$$= 11 - 55 + 22 = -22, \quad \text{desarrollando por la primera fila}$$

y

$$\begin{vmatrix} 1 & 5 & 2 \\ -1 & 1 & 0 \\ 3 & 4 & 1 \end{vmatrix} = 2(-1)^{1+3}\begin{vmatrix} -1 & 1 \\ 3 & 4 \end{vmatrix} + 0 + 1(-1)^{3+3}\begin{vmatrix} 1 & 5 \\ -1 & 1 \end{vmatrix} =$$
$$= -14 + 6 = -8 \quad \text{desarrollando por la tercera columna}$$

En consecuencia, $|A| = 4(-22) + 1(-8) = -96$.

Si no tiene elementos nulos, el determinante de una matriz 3×3 requiere $3\cdot 2 = 3!$ multiplicaciones, el de una matriz 4×4 requiere $4 \cdot 3 \cdot 2 = 4!$ multiplicaciones y, en general, una matriz $n \times n$ require $n!$ multiplicaciones, de manera que el número de productos necesarios es prohibitamente elevado conforme el orden de la matriz aumenta. Así, será necesario usar propiedades de los determinantes para, en lo posible, tratar de corregir esta situación.

1.2.4. Propiedades de los determinantes

Las matrices triangulares contienen muchos ceros, de manera que su determinante es particularmente sencillo de calcular.

1. El determinante de una matriz triangular (superior o inferior) es el producto de los elementos de la diagonal principal.

   Las matrices diagonales son a la vez triangulares superiores e inferiores, de modo que la siguiente propiedad es inmediata.

2. El determinante de una matriz diagonal es el producto de los elementos de su diagonal principal.

   El desarrollo por una fila o una columna que tiene muchos ceros simplifica mucho el cálculo de un determinante; el desarrollo por una fila o una columna de ceros, si existe, hace el proceso trivial: al multiplicar cada elemento cero por su cofactor da cero, que al sumar con los demás también da cero, de manera que

3. Si una matriz cuadrada tiene una fila o una columna de ceros su determinante es cero.

4. Si la matriz B se obtiene a partir de una matriz cuadrada A intercambiando la posición de dos filas en A, entonces $\det(B) = -\det(A)$.

   Como consecuencia inmediata de esta propiedad se tiene la siguiente:

5. Si una matriz tiene dos filas idénticas, entonces su determinante es nulo.

6. Si la matriz B se obtiene a partir de una matriz cuadrada A multiplicando todos los elementos de una fila de A por un escalar λ entonces $\det(B) = \lambda \det(A)$.

   Como una extensión inmediata se tiene

7. Si A es una matriz $n \times n$, entonces $\det(\lambda A) = \lambda^n \det(A)$.

8. Si la matriz B se obtiene a partir de una matriz cuadrada A sumando a una fila de A otra fila de A multiplicada por un escalar λ, entonces $\det(B) = \det(A)$.

9. En general $\det(A + B) \neq \det(A) + \det(B)$. No obstante, si las matrices A y B tienen todas las filas iguales salvo la número i, entonces $\det(A) + \det(B)$ será el determinante de una nueva matriz C que tiene todas sus filas, salvo la número i iguales a las de A y B, y cuya fila i es la suma de la fila i de A y la fila i de B.

10. Para cualquier matriz cuadrada A, se tiene que $\det(A) = \det(A^T)$.

    Por consiguiente, todo lo indicado anteriormente para filas se verifica también para columnas.

11. Si A y B son matrices del mismo orden, entonces $\det(AB) = \det(A)\det(B)$.

Las propiedades 4, 6 y 8 muestran el efecto que sobre el determinante de una matriz cuadrada tiene el someter a ésta a las llamadas *operaciones elementales sobre filas*. Más específicamente, a lo largo del curso consideraremos en múltiples contextos las tres operaciones elementales sobre las filas de una matriz siguientes:

(F1) Intercambiar la posición de dos filas cualesquiera de la matriz.

(F2) Multiplicar cualquier fila de la matriz por un escalar no nulo.

(F3) Sumar a una fila de la matriz otra fila de la misma matriz multiplicada por un escalar no nulo.

Como veremos posteriormente, mediante estas sencillas operaciones es posible resolver, en particular, cualquier sistema de ecuaciones lineales.

Un método elegante para reducir sustancialmente el número de operaciones aritméticas necesarias para evaluar determinantes de matrices está basado precisamente en estas operaciones elementales. Se trata, básicamente, de transformar la matriz de partida en otra en una de cuyas columnas haya el mayor número posible de ceros; a continuación, si se desea, el procedimiento se puede repetir para obtener finalmente una matriz escalonada.

En otras palabras, es posible diseñar un algoritmo eficiente para calcular el determinante de una matriz cuyos elementos son todos constantes de la siguiente forma. Se usan operaciones elementales sobre filas para transformar una matriz a la forma escalonada, ya que la matriz resultante es triangular y su determinante es fácil de evaluar por medio de la propiedad 1. Es preciso llevar un registro de los cambios hechos al determinante en el proceso de convertir la matriz en escalonada. El producto de estos cambios por el determinante de la matriz escalonada es entonces el determinante de la matriz original. Ilustremos el procedimiento con un par de ejemplos.

Ejemplo 1. Dada $A = \begin{pmatrix} 0 & -1 & 2 \\ 4 & 5 & 6 \\ -2 & 1 & 4 \end{pmatrix}$, calcular su determinante reduciéndolo al de una matriz escalonada.

Solución. $\det(A) = \begin{vmatrix} 0 & -1 & 2 \\ 4 & 5 & 6 \\ -2 & 1 & 4 \end{vmatrix} \underset{F_2 \longleftrightarrow F_1}{=} - \begin{vmatrix} 4 & 5 & 6 \\ 0 & -1 & 2 \\ -2 & 1 & 4 \end{vmatrix} =$

$$\underset{F_3 \to F_3 + \frac{1}{2}F_1}{=} - \begin{vmatrix} 4 & 5 & 6 \\ 0 & -1 & 2 \\ 0 & \frac{7}{2} & 7 \end{vmatrix} \underset{F_3 \to F_3 + \frac{7}{2}F_2}{=} - \begin{vmatrix} 4 & 5 & 6 \\ 0 & -1 & 2 \\ 0 & 0 & 14 \end{vmatrix} = 56$$

Ejemplo 2. Evaluar $\det \begin{pmatrix} 1 & 2 & 3 \\ -2 & 3 & 2 \\ 3 & -1 & 1 \end{pmatrix}$.

Solución. $\begin{vmatrix} 1 & 2 & 3 \\ -2 & 3 & 2 \\ 3 & -1 & 1 \end{vmatrix} \underset{F_2 \to F_2 + 2F_1}{=} \begin{vmatrix} 1 & 2 & 3 \\ 0 & 7 & 8 \\ 3 & -1 & 1 \end{vmatrix} =$

$$\underset{F_3 \to F_3 - 3F_1}{=} \begin{vmatrix} 1 & 2 & 3 \\ 0 & 7 & 8 \\ 0 & -7 & -8 \end{vmatrix} \underset{F_2 \to \frac{1}{7}F_2}{=} 7 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 8/7 \\ 0 & -7 & -8 \end{vmatrix} =$$

$$\underset{F_3 \to F_3 + 7F_2}{=} 7 \begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 8/7 \\ 0 & 0 & 0 \end{vmatrix} = 7(0) = 0.$$

1.2.5. Rango de menores de una matriz

Ya introdujimos anteriormente el concepto de *menor* de orden p de una matriz A $m \times n$. Éste no era más que el determinante de cualquier submatriz $p \times p$ de A (con $p \leq m$, $p \leq n$). Introducimos ahora la noción de *rango* de menores de una matriz arbitraria A de tamaño $m \times n$.

Dada una matriz $A \in \mathcal{M}_{m \times n}$, consideremos todos sus menores: los de orden 1, los de orden 2,... (hasta llegar al más pequeño de los m y n). De todos ellos, nos van a interesar los que son no nulos, que serán en general de distintos órdenes. Pues bien, al mayor de estos órdenes (de menores no nulos) se le llamará rango de menores de la matriz A. En otras palabras, este rango es p si hay alguna submatriz de tamaño $p \times p$ con determinante diferente de cero y todas las submatrices cuadradas de mayor tamaño que p tienen determinante cero.

Definición. Se dice que p es el *rango de menores* (o simplemente *rango*) de una matriz A, de tamaño $m \times n$ cualquiera, si A tiene algún menor de orden p que no es nulo y todos los menores de A de orden mayor que p son nulos; o sea, p es el mayor de los órdenes de los menores no nulos de A.

Procedimiento de cálculo. Para hallar el rango de menores de A, se toma un menor M_2 de orden 2 no nulo y se le orla con una fila fija, la i, y con sucesivas columnas; si todos los menores de orden 3 que así se obtienen son nulos, entonces se prescinde de la fila i y se repite el proceso con otra o con otras filas hasta: (1) encontrar un menor M_3 de orden 3 no nulo, en cuyo caso el rango es al menos 3; o (2) descubrir que todos los menores de orden 3 son nulos, en cuyo caso el rango es 2. Si hay un menor M_3 no nulo, se le orla con una fila y con sucesivas columnas, siguiendo el mismo proceso que con M_2, lo que nos lleva o bien a que el rango es 3 (si todos los menores de orden 4 son nulos) o bien a que el rango es al menos 4 (en cuanto se encuentre un menor de orden 4 no nulo). Siguiendo así, se llega a un menor no nulo del mayor tamaño posible; este tamaño es el rango de menores buscado.

Ejemplo. Calcular el rango de menores de la matriz

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

Solución. El menor de orden dos $\begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix}$ es diferente de cero (su valor es -1). Por tanto, el rango de menores de A es al menos 2. A continuación, lo orlamos con la fila 3 y la columna 3. Si el menor resultante de orden 3 es diferente de cero, el rango de A será 3, dado que ya no podremos formar una submatriz de A de tamaño 4×4. Si, por el contrario, el menor de orden 3 así formado a partir del de orden 2 no nulo anterior fuera cero, lo orlaríamos con la tercera fila y la cuarta columna. Si éste también fuera cero, el rango de A sería 2; en caso contrario, el rango sería 3.

Así pues, formamos el menor de orden tres $\begin{vmatrix} 1 & 2 & 3 \\ 2 & 3 & 5 \\ 1 & 3 & 4 \end{vmatrix}$ (orlando con tercera fila y tercera columna), que vale cero. Por tanto, formamos a continuación el menor $\begin{vmatrix} 1 & 2 & 2 \\ 2 & 3 & 1 \\ 1 & 3 & 5 \end{vmatrix}$ (se orlan con tercera fila y cuarta columna), que también es cero. Por consiguiente, el rango de menores de A es 2.

En el tema 2 volveremos sobre el concepto de rango desde un punto de vista diferente, más intuitivo, y proporcionaremos un método de cálculo mucho más sencillo del mismo.

1.3. Inversa de una matriz

La división por números reales es equivalente a la multiplicación por el recíproco. Podemos resolver la ecuación lineal $5x = 20$ para la variable x o bien dividiendo la ecuación por 5 o bien multiplicando la ecuación por 0.2, el recíproco de 5. Un número real b es el recíproco de a si y sólo si $ab = 1$, en cuyo caso escribimos $b = a^{-1}$. El concepto de recíproco se puede extender a matrices. La matriz que hace el papel del número 1 es la matriz identidad I, y se usa la palabra *inversa* de una matriz A en lugar de recíproca, aun cuando se mantiene la notación A^{-1}.

1.3.1. Definición

Diremos que una matriz $B \in \mathcal{M}_{n\times n}$ es la inversa de una matriz $A \in \mathcal{M}_{n\times n}$ si y sólo si

$$A\,B = B\,A = I$$

en cuyo caso escribimos $B = A^{-1}$.

Observemos que el requerimiento de que una matriz conmute con su inversa implica que ambas matrices son cuadradas y del mismo orden. Así, las inversas sólo están definidas para matrices cuadradas. Si una matriz cuadrada A tiene inversa, entonces se dice que A es invertible o *regular*. Si A no tiene inversa, se dice que A es *singular*.

Ejemplo. La matriz $B = \begin{pmatrix} -2 & 1 \\ 3/2 & -1/2 \end{pmatrix}$ es la inversa de $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ ya que

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

1.3.2. Propiedades

1. La inversa de una matriz es única.
2. Si A es una matriz regular, entonces
   $\left(A^{-1}\right)^{-1} = A, \quad \left(A^{-1}\right)^{T} = \left(A^{T}\right)^{-1}, \quad (\lambda A)^{-1} = \frac{1}{\lambda}\,(A)^{-1},$ siendo $\lambda \neq 0$.
3. Si A y B son matrices regulares del mismo orden se cumple que
   $(AB)^{-1} = B^{-1}A^{-1}$.
4. La inversa de una matriz simétrica regular es simétrica.
5. La inversa de una matriz triangular regular también es triangular.

La demostración de estas propiedades es elemental. Como ilustración, consideremos la primera de ellas. Así, supongamos que B y C son ambas inversas de la matriz A. Entonces

$$AB = I, \quad BA = I, \quad AC = I, \quad CA = I$$

De aquí se deduce que

$$C = CI = C(AB) = (CA)B = IB = B$$

Esto es, si B y C son ambas inversas de A, entonces han de ser iguales; de aquí, la inversa es única. Además, se tiene el importante resultado siguiente:
Teorema. Una matriz cuadrada A tiene inversa si y sólo si su determinante $\det(A)$ es diferente de cero. En ese caso $\det(A^{-1}) = 1/\det(A)$.

Aunque demostrar la primera parte de este teorema excede el propósito de este curso, es elemental demostrar la segunda parte. En efecto, si A es invertible, entonces $\det(A) \neq 0$ y $AA^{-1} = I$. Por consiguiente,

$$\det(AA^{-1}) = \det(I) = 1$$

y de aquí

$$\det(A) \cdot \det(A^{-1}) = 1$$

se donde se tiene el resultado buscado.

1.3.3. Procedimientos para el cálculo de la matriz inversa

Cálculo de la inversa mediante la matriz de los adjuntos Dada $A = (a_{ij})$ una matriz cuadrada de tamaño $n \times n$, se llama *matriz adjunta* (o de cofactores) de A a la matriz $\mathrm{adj}(A) = (\alpha_{ij})$ de tamaño $n \times n$ cuyos elementos son los adjuntos α_{ij} de a_{ij}. Se puede demostrar entonces sin excesivas dificultades que si A es regular, entonces su inversa A^{-1} está dada por

$$A^{-1} = \frac{1}{\det(A)} (\mathrm{adj}(A))^T = \frac{1}{\det(A)} \begin{pmatrix} \alpha_{11} & \alpha_{21} & \dots & \alpha_{n1} \\ \alpha_{12} & \alpha_{22} & \dots & \alpha_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n} & \alpha_{2n} & \dots & \alpha_{nn} \end{pmatrix}$$

Como vemos, esta expresión involucra el cálculo del determinante de una matriz $n \times n$ y, en general, el de n^2 determinantes de orden $(n-1) \times (n-1)$, con lo cual, incluso para valores moderados de n ($n \geq 4$), este procedimiento es prohibitivo en la práctica, de manera que conviene introducir otro, más factible computacionalmente.

Cálculo de la inversa mediante operaciones elementales sobre filas Antes de introducir el procedimiento propiamente dicho, conviene hacer notar que cada operación elemental efectuada sobre las filas de una matriz está generada de algún modo por una cierta matriz. Así, llamaremos *matriz elemental* E a una matriz cuadrada que genera una operación elemental sobre una fila de una matriz A (la cual no tiene que ser necesariamente cuadrada) bajo la multiplicación EA. Las matrices elementales se construyen aplicando la operación elemental de fila deseada sobre una matriz identidad del orden apropiado, concretamente el orden de una matriz cuadrada que tiene tantas columnas como filas tiene A, de manera que la multiplicación EA esté definida. En particular,

(i) Para construir una matriz elemental que intercambie la fila i con la fila j, partimos de la matriz identidad I. Primero intercambiamos el 1 en la posición i-i con el cero en la posición j-i y después intercambiamos el 1 en la posición j-j con el cero en la posición i-j.

(ii) Para construir una matriz elemental que multiplica la fila i de una matriz por el escalar no nulo k, se parte de la matriz identidad I y se sustituye el 1 en la posición i-i por k.

(iii) Para construir una matriz elemental que añade a la fila j de una matriz el producto de un escalar k por la fila i de la matriz, se parte de la matriz I y se sustituye el 0 en la posición j-i por k.

Ejemplo. Construir matrices elementales tales que cuando se multiplican por cualquier matriz A 3×5 (a) intercambie la primera y la segunda fila de A; (b) multiplique la tercera fila de A por -2, y (c) añada a la tercera fila de A 4 veces su segunda fila.

Solución.

$$\text{(a)}\ \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};\quad \text{(b)}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix};\quad \text{(c)}\ \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 4 & 1 \end{pmatrix}.$$

Las operaciones elementales se pueden usar para construir un método de cálculo de la matriz inversa bastante extendido. En esencia, una matriz cuadrada A tiene inversa si y sólo si A puede ser transformada en la matriz identidad mediante operaciones elementales sobre filas. En efecto, como cada operación elemental sobre las filas se puede representar por una matriz elemental, concluimos que una matriz A tiene inversa si y sólo si existe una secuencia de matrices elementales $E_1, E_2, \ldots, E_k$ tales que

$$E_k E_{k-1} \cdots E_2 E_1 A = I.$$

Denotando el producto de estas matrices elementales como B, tenemos entonces que $BA = I$, lo cual implica que $B = A^{-1}$. Así pues, para calcular la inversa de una matriz A sólo necesitamos llevar un registro de las operaciones elementales sobre filas usadas para transformar A en la identidad I. Esto se logra simplemente aplicando las mismas operaciones elementales sobre filas a A y a I simultáneamente. El procedimiento se puede resumir en los pasos siguientes:

1. Crear una matriz ampliada (A, I), donde A es la matriz $n \times n$ cuya inversa se desea calcular e I es la matriz identidad $n \times n$.

2. Usar operaciones elementales sobre filas en (A, I) para transformar A a la forma escalonada, aplicando cada operación sobre la matriz ampliada.

3. Si la matriz transformada de A en la forma escalonada tiene ceros en la diagonal principal, parar (A no tiene inversa). Si no, continuar.

4. Usar operaciones elementales sobre filas en la matriz escalonada ampliada para transformar la parte izquierda en la matriz identidad $n \times n$, aplicando cada operación sobre toda la matriz ampliada.

5. La parte derecha de la matriz ampliada final es la inversa de A.

Ejemplo. Calcular la matriz inversa de $A = \begin{pmatrix} 1 & -2 & 1 \\ 2 & 2 & 2 \\ 3 & -4 & -3 \end{pmatrix}$

Solución. $\left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 1 & 0 \\ 3 & -4 & -3 & 0 & 0 & 1 \end{array}\right) \underset{\substack{F_2 \to F_2 - 2F_1 \\ F_3 \to F_3 - 3F_1}}{\sim} \left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 6 & 0 & -2 & 1 & 0 \\ 0 & 2 & -6 & -3 & 0 & 1 \end{array}\right)$

$$\underset{F_2 \to \frac{1}{6}F_2}{\sim} \left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 2 & -6 & -3 & 0 & 1 \end{array}\right) \underset{F_3 \to F_3 - 2F_2}{\sim} \left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 0 & -6 & -\frac{7}{3} & -\frac{1}{3} & 1 \end{array}\right)$$

$$\underset{F_3\to-\frac{1}{6}F_3}{\sim}\left(\begin{array}{ccc|ccc} 1 & -2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{2}{6} & \frac{1}{6} & 0 \\ 0 & 0 & 1 & \frac{7}{18} & \frac{1}{18} & -\frac{1}{6} \end{array}\right)\underset{F_1\to F_1+2F_2-F_3}{\sim}\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{18} & \frac{5}{18} & \frac{1}{6} \\ 0 & 1 & 0 & -\frac{1}{3} & \frac{1}{6} & 0 \\ 0 & 0 & 1 & \frac{7}{18} & \frac{1}{18} & -\frac{1}{6} \end{array}\right)$$

luego

$$A^{-1}=\begin{pmatrix} -\frac{1}{18} & \frac{5}{18} & \frac{1}{6} \\ -\frac{1}{3} & \frac{1}{6} & 0 \\ \frac{7}{18} & \frac{1}{18} & -\frac{1}{6} \end{pmatrix} \text{ y se comprueba que}$$

$$AA^{-1}=A^{-1}A=\begin{pmatrix} 1 & -2 & 1 \\ 2 & 2 & 2 \\ 3 & -4 & -3 \end{pmatrix}\begin{pmatrix} -\frac{1}{18} & \frac{5}{18} & \frac{1}{6} \\ -\frac{1}{3} & \frac{1}{6} & 0 \\ \frac{7}{18} & \frac{1}{18} & -\frac{1}{6} \end{pmatrix}=\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}=I$$

1.4. Sistemas de ecuaciones lineales.

Los sistemas de ecuaciones lineales simultáneas aparecen frecuentemente en multitud de problemas. La necesidad de contar con métodos eficientes para resolver tales sistemas fue históricamente una de las razones que impulsaron la introducción de las matrices, y esa necesidad todavía se da hoy día, especialmente cuando se trata de resolver sistemas grandes, con cientos de ecuaciones y variables.

1.4.1. Definición

Llamaremos sistema lineal de m ecuaciones con n incógnitas a un conjunto de m igualdades del tipo

$$\begin{array}{rcl} a_{11}x_1+a_{12}x_2+a_{13}x_3+\cdots+a_{1n}x_n & = & b_1 \\ a_{21}x_1+a_{22}x_2+a_{23}x_3+\cdots+a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1+a_{m2}x_2+a_{m3}x_3+\cdots+a_{mn}x_n & = & b_m \end{array} \tag{1.8}$$

Los números reales (conocidos) a_{ij}, $1\leq i\leq m$, $1\leq j\leq n$ se llaman *coeficientes* del sistema, los números reales b_i, $1\leq i\leq m$ son los *términos independientes* y las $x_1,x_2,\ldots,x_n$ son las *incógnitas.* Las incógnitas en una ecuación lineal sólo aparecen elevadas a la potencia unidad y están multiplicadas únicamente por escalares conocidos. Las ecuaciones *no* involucran productos de incógnitas, incógnitas elevadas a potencias diferentes de uno o incógnitas apareciendo como argumentos de funciones trascendentes.

Para sistemas que contienen unas cuantas incógnitas es usual denotar a las variables por letras distintas tales como x, y y z. Esta notación claramente es poco práctica para sistemas que involucran cientos de incógnitas, de modo que en esos casos se utiliza una sola letra que identifica todas las variables con diferentes subíndices, tales como $x_1,x_2,\ldots,x_n$.

Una n-tupla $(\alpha_1,\alpha_2,\ldots,\alpha_n)$ se dice que es una *solución* del sistema si al sustituir $x_1\to\alpha_1$, $x_2\to\alpha_2$, ..., $x_m\to\alpha_m$ en *cada ecuación del sistema, se verifican todas y cada una de ellas.*

El conjunto de todas las soluciones del sistema se llama *solución general*; una solución cualquiera se conoce como *solución particular*.

Los sistemas se clasifican así:

$$\begin{cases} \text{Compatible (tiene solución)} \begin{cases} \text{Determinado (Solución única)} \\ \text{Indeterminado (Infinitas soluciones)} \end{cases} \\ \text{Incompatible (no tiene solución)} \end{cases}$$

Si $b_1 = b_2 = \cdots = b_m = 0$ se dice que el sistema es *homogéneo*. Si algún $b_i \neq 0$ el sistema es *inhomogéneo*. Dos sistemas de ecuaciones lineales se dice que son *equivalentes* si tienen la misma solución general.

Veamos ahora unos cuantos ejemplos.

Ejemplo 1. El sistema

$$\begin{aligned} 2x + 3xy &= 25 \\ 4\sqrt{x} + \sin y &= 50 \end{aligned}$$

no es un sistema lineal por varios motivos: contiene un producto xy de las variables; contiene la incógnita x elevada a la potencia $1/2$ y contiene la variable y como el argumento de una función seno.

Ejemplo 2. El sistema de ecuaciones

$$\begin{array}{rcrcrcr} x & + & 2y & + & z & = & 1 \\ 3x & - & y & + & 2z & = & b \\ x & + & y & + & z & = & 3 \\ x & + & y & - & z & = & -1 \end{array}$$

es lineal y resulta ser compatible determinado sólo cuando $b = 15$, siendo entonces la solución $x = 3$, $y = -2$, $z = 2$. Si $b \neq 15$, entonces el sistema es incompatible. Para llegar a esta conclusión, basta con realizar las siguientes manipulaciones: de la tercera ecuación se tiene que $x + z = 3 - y$, que, sustituida en la primera ecuación, lleva a $3 - y + 2y = 1$, y de aquí $y = -2$. Ahora, de la tercera ecuación, $z = 5 - x$, que sustituida en la cuarta, permite concluir que $x - 2 - (5 - x) = -1$, o bien $x = 3$ y de aquí $z = 2$. Por último, sustituyendo estos valores en la segunda ecuación, vemos que ésta sólo se verifica si $b = 15$.

Ejemplo 3. Dado el sistema lineal

$$\begin{array}{rcrcrcr} x & + & y & + & z & = & 1 \\ -x & + & y & - & 2z & = & 2 \\ 2x & & & + & 3z & = & -1 \end{array}$$

al sumar las dos primeras ecuaciones se tiene la igualdad $2y - z = 3$ o bien $z = 2y - 3$, que, sustituida en la primera ecuación, da $x = 1 - y - (2y - 3) = 4 - 3y$. Al sustituir ahora en la tercera ecuación, tenemos $2(4 - 3y) + 3(2y - 3) = -1$, o sea, que ésta se verifica idénticamente, con lo cual el sistema anterior es equivalente a

$$\begin{array}{rcrcrcr} x & + & y & + & z & = & 1 \\ -x & + & y & - & 2z & = & 2 \end{array}$$

el cual tiene como solución general $x = 4-3\lambda$, $y = \lambda$, $z = 2\lambda-3$ con $\lambda \in \mathbb{R}$ cualquier número real. Se trata, pues, de un sistema compatible indeterminado (tiene infinitas soluciones, una para cada valor de λ).

Como ya se vio anteriormente, un sistema de ecuaciones lineales de la forma (1.8) se puede escribir matricialmente como

$$AX = B,$$

con

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \qquad B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

Se dice que la matriz A $m \times n$ es la *matriz de coeficientes*, la matriz columna $n \times 1$ X es la *matriz de incógnitas* y la matriz B $m \times 1$ es la *matriz del término independiente*. En muchas ocasiones es bastante útil la noción de *matriz ampliada* (A, B) del sistema. Ésta no es más que la matriz de coeficientes a la que se añade como columna adicional los términos independientes, es decir,

$$(A, B) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array}\right)$$

1.4.2. Existencia de soluciones para un sistema lineal

Se tiene el siguiente resultado concerniente a los sistemas de ecuaciones lineales:
Teorema. Si X_1 y X_2 son dos soluciones diferentes del sistema $AX = B$, entonces $Z = \alpha X_1 + \beta X_2$ también es una solución del sistema para cualesquiera números reales α y β tales que $\alpha + \beta = 1$.

La demostración de este teorema es elemental. En efecto, si X_1 y X_2 son soluciones de $AX = B$, entonces $AX_1 = B$ y $AX_2 = B$. En consecuencia,

$$AZ = A(\alpha X_1 + \beta X_2) = \alpha(AX_1) + \beta(AX_2) = \alpha B + \beta B = (\alpha + \beta)B = B$$

y así Z también es solución.

Como hay infinitas maneras de formar $\alpha + \beta = 1$ (fijando α como un número cualquiera ponemos $\beta = 1-\alpha$), de este teorema se sigue que una vez hemos identificado dos soluciones, podemos combinarlas para formar infinitas soluciones. Así pues, el número de posibles soluciones de un sistema de ecuaciones lineales es cero, uno o infinito.

Estudiando el rango de la matriz de coeficientes A de un sistema y el de su matriz ampliada (A, B) se puede demostrar el siguiente teorema de existencia de soluciones:
Teorema. El sistema $AX = B$ es compatible si y sólo si el rango de A es igual al rango de (A, B).

La demostración de este resultado se llevará a cabo en el transcurso del tema siguiente.
Ejemplo. Determinar si el sistema siguiente es compatible o no:

$$\begin{aligned} 2x - 3y + z &= -1 \\ x - y + 2z &= 2 \\ 2x + y - 3z &= 3 \end{aligned}$$

Solución. En este caso

$$A = \begin{pmatrix} 2 & -3 & 1 \\ 1 & -1 & 2 \\ 2 & 1 & -3 \end{pmatrix} \quad \text{y} \quad (A, B) = \left(\begin{array}{ccc|c} 2 & -3 & 1 & -1 \\ 1 & -1 & 2 & 2 \\ 2 & 1 & -3 & 3 \end{array}\right)$$

Es fácil demostrar que tanto el rango de A como el rango de (A, B) valen 3. Por tanto, el sistema tiene solución. Ahora bien, ¿cuántas soluciones tiene? Para contestar a esta pregunta, haremos uso del teorema siguiente.

Teorema. Si el sistema $AX = B$ es compatible y si el rango de A, $r(A) = k$, entonces las soluciones del sistema se pueden expresar en términos de $n - k$ parámetros arbitrarios, siendo n el número total de incógnitas en el sistema.

Así, en el ejemplo anterior, como $r(A) = r(A, B) = 3$, las soluciones se pueden expresar en términos de $3 - 3 = 0$ parámetros arbitrarios. Así pues, la solución es única. Se puede comprobar que esta solución es $x = y = 2$, $z = 1$.

Los teoremas anteriores siguen siendo válidos para sistemas homogéneos $AX = O$. Como la estructura de dichos sistemas es bastante más simple, podemos obtener conclusiones para ellos que no son válidas en el caso de sistemas no homogéneos. En particular, un sistema homogéneo es compatible, ya que la solución trivial $X = O$ es siempre una solución de $AX = O$. Además, si el rango de A es igual al número de incógnitas, entonces la solución es única y la solución trivial es la única solución. Por otro lado, si el rango de A es menor que el número de incógnitas, podremos expresar algunas de ellas en términos de otras arbitrarias, de manera que habrá infinitas soluciones. En resumen,
Teorema. Un sistema homogéneo $AX = O$ con n incógnitas admite soluciones no triviales si y sólo si $r(A) < n$.

1.4.3. Resolución de un sistema de ecuaciones lineales

Si bien los teoremas anteriores permiten averiguar si un sistema lineal tiene o no soluciones, no proporcionan, sin embargo, ningún procedimiento para calcular explícitamente las soluciones del mismo. Vamos ahora a revisar brevemente dos de ellos.

Aplicación de la matriz inversa a la resolución de sistemas Si la matriz de coeficientes A del sistema $AX = B$ es regular, entonces podemos multiplicar por la izquierda para obtener

$$A^{-1}(AX) = A^{-1}B \quad \Rightarrow \quad (A^{-1}A)X = A^{-1}B \quad \Rightarrow \quad X = A^{-1}B.$$

Con esta fórmula podemos hallar todas las variables del sistema al mismo tiempo. Esta es la conocida *regla de Cramer*. Alternativamente,

$$x_j = \frac{\det \tilde{A}}{\det(A)}$$

siendo $\tilde{A}$ la matriz obtenida a partir de A remplazando la columna j por B.
Ejemplo. El sistema de ecuaciones

$$\begin{aligned} 5x_1 + 8x_2 + x_3 &= 2 \\ 2x_2 + x_3 &= -1 \\ 4x_1 + 3x_2 - x_3 &= 3 \end{aligned}$$

se puede escribir en forma matricial como $AX = B$, siendo

$$A = \begin{pmatrix} 5 & 8 & 1 \\ 0 & 2 & 1 \\ 4 & 3 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix}, \qquad B = \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix}$$

Aplicando los procedimientos de la sección 3, se tiene que la matriz de coeficientes A es regular y por tanto

$$\begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix} = X = A^{-1}B = \begin{pmatrix} 5 & -11 & -6 \\ -4 & 9 & 5 \\ 8 & -17 & -10 \end{pmatrix} \begin{pmatrix} 2 \\ -1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ -2 \\ 3 \end{pmatrix}$$

y por tanto $x_1 = 3$, $x_2 = -2$, $x_3 = 3$.

La invertibilidad de la matriz de coeficientes A no sólo proporciona una solución del sistema $AX = B$, sino que también proporciona los medios para mostrar que esta solución es la *única* solución del sistema.
Teorema. Si A es invertible, entonces el sistema de ecuaciones lineales definido por $AX = B$ tiene solución única.

¿Y qué ocurre si A no es regular o, más en general, si A no es cuadrada? En ese caso se trata de localizar en A un menor diferente de cero de orden igual al rango de la matriz. Después pasaremos a trabajar con un sistema equivalente sin las filas que no aparecen en el menor y tal que al otro lado de las igualdades estén también las columnas que tampoco estén en el menor. Este sistema equivalente se resolverá mediante la regla de Cramer.
Ejemplo. Para el sistema de ecuaciones

$$\begin{aligned} x + y + z - 2t - 4u &= 6 \\ 2x + y - z + t - 3u &= 1 \\ x - 2y + z - 5t + 2u &= 0 \\ 4x + 3y + z - 3t - 11u &= 13 \end{aligned}$$

la matriz de coeficientes tiene rango 3 y la matriz ampliada tiene también rango 3; el sistema es, pues, compatible, y las soluciones se pueden expresar en términos de

$5-3=2$ parámetros arbitrarios. Dado que el menor formado por las tres primeras filas y las tres primeras columnas de la matriz de coeficientes

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

es distinto de cero, el sistema dado es equivalente a

$$\begin{array}{rcl} x+y+z & = & 6+2t+4u \\ 2x+y-z & = & 1-t+3u \\ x-2y+z & = & 0+5t-2u \end{array}$$

Ahora las variables t y u se toman como parámetros arbitrarios, $t=\lambda$, $u=\mu$, y se resuelve el sistema en las variables x, y, z por medio de la regla de Cramer, pues la correspondiente matriz de coeficientes es invertible (su determinante es el menor previamente identificado, el cual es distinto de cero). Procediendo así, se llega a que la solución general es

$$x=1+\lambda+\mu, \qquad y=2-\lambda+2\mu, \qquad z=3+2\lambda+\mu, \qquad t=\lambda, \qquad u=\mu$$

con $\lambda,\mu\in\mathbb{R}$.

Resolución de un sistema por eliminación gaussiana El método tradicional de resolver un sistema de ecuaciones lineales es manipular las ecuaciones de manera que las ecuaciones resultantes sean más fáciles de resolver y tengan las *mismas* soluciones que las ecuaciones originales, es decir, transformar el sistema en uno equivalente que sea más fácil de resolver. Tres operaciones que alteran las ecuaciones pero no cambian sus soluciones son:

(i) intercambiar las posiciones de cualesquiera dos ecuaciones;

(ii) multiplicar una ecuación por un escalar no nulo;

(iii) sumar a una ecuación el producto de un escalar por otra ecuación.

Si adaptamos estas operaciones a la matriz ampliada del sistema (esto es, la matriz formada por la matriz de coeficientes junto con la del término independiente), vemos que no son más que las operaciones elementales sobre sus filas.

La eliminación gaussiana es un método matricial en cuatro pasos, centrado en operaciones elementales sobre filas, para resolver ecuaciones lineales. Los pasos son:

1. Construir una matriz ampliada para el sistema de ecuaciones dado.

2. Usar operaciones elementales sobre filas para transformar la matriz ampliada en una matriz en forma escalonada.

3. Escribir las ecuaciones asociadas a la matriz ampliada resultante.

4. Resolver el nuevo conjunto de ecuaciones mediante la técnica de la 'sustitución hacia atrás'.

Dicho de otra forma, el método consiste en aplicar operaciones elementales sobre las ecuaciones para encontrar un sistema equivalente al dado de forma que la incógnita x_1 aparezca sólo en la primera ecuación, la x_2 en la segunda, etc. Es decir, se van eliminando incógnitas en las ecuaciones.

Ejemplo 1. Vamos a analizar y resolver el sistema de ecuaciones lineales

$$\begin{aligned} x_1 + 3x_2 + x_3 - x_4 &= 6 \\ 2x_1 + 7x_2 + 3x_3 - 4x_4 &= 15 \\ x_1 + x_2 + 2x_3 + x_4 &= 1 \end{aligned}$$

Realizando operaciones elementales en las filas de la matriz ampliada se obtiene sucesivamente

$$\left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 6 \\ 2 & 7 & 3 & -4 & 15 \\ 1 & 1 & 2 & 1 & 1 \end{array}\right), \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 6 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & -2 & 1 & 2 & -5 \end{array}\right), \left(\begin{array}{cccc|c} 1 & 3 & 1 & -1 & 6 \\ 0 & 1 & 1 & -2 & 3 \\ 0 & 0 & 3 & -2 & 1 \end{array}\right);$$

por tanto, el sistema dado es equivalente al

$$\begin{aligned} x_1 + 3x_2 + x_3 - x_4 &= 6 \\ x_2 + x_3 - x_4 &= 3 \\ 3x_3 - 2x_4 &= 1 \end{aligned}$$

el cual es compatible indeterminado (los rangos de su matriz de coeficientes y de su matriz ampliada son ambos iguales a 3). Haciendo $x_4 = 3\lambda+1$, con $\lambda \in \mathbb{R}$ parámetro, se tiene la solución general:

$$x_1 = -11\lambda - 6, \quad x_2 = 4\lambda + 4, \quad x_3 = 2\lambda + 1, \quad x_4 = 3\lambda + 1$$

Ejemplo 2. Resuélvase el siguiente sistema de ecuaciones lineales:

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ -3x_1 - 5x_2 + 4x_3 &= 1 \\ 4x_1 + 7x_2 - 5x_3 &= 2 \\ 2x_1 + 4x_2 - 3x_3 &= 4 \end{aligned}$$

Realizando operaciones elementales en las filas de la matriz ampliada se obtiene sucesivamente

$$\left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ -3 & -5 & 4 & 1 \\ 4 & 7 & -5 & 2 \\ 2 & 4 & -3 & 4 \end{array}\right), \left(\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & -2 & 4 \\ 0 & -1 & 3 & -2 \\ 0 & 0 & 1 & 2 \end{array}\right),$$

$$\left(\begin{array}{rrr|r} 1 & 2 & -2 & 1 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 2 \end{array}\right), \left(\begin{array}{rrr|r} 1 & 2 & -2 & 1 \\ 0 & 1 & -2 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right);$$

por tanto, el sistema dado es equivalente al

$$\begin{aligned} x_1 + 2x_2 - 2x_3 &= 1 \\ x_2 - 2x_3 &= 4 \\ x_3 &= 2 \end{aligned}$$

que tiene solución única: $x_3 = 2$, $x_2 = 8$, $x_1 = -11$.

1.5. Aplicaciones económicas

1.5.1. Análisis de entrada–salida (input–output)

La macroeconomía es una rama de la economía que trata de los aspectos amplios y generales de un sistema económico, por ejemplo, las relaciones entre los ingresos, las inversiones y los gastos de un país en su totalidad. Se han desarrollado muchas técnicas para tratar estos problemas en la macroeconomía. Discutiremos ahora uno de los más importantes.

Para introducir el modelo suponemos que el parlamento de un cierto país de economía muy desarrollada (por ejemplo, los Estados Unidos) ha aprobado una gran disminución en los gastos para construcción de carreteras. Si no se diera además un aumento en otras inversiones, se esperaría una disminución en los ingresos y en el empleo. Por otra parte, supóngase que el gobierno aumentara sus gastos militares en una cantidad equivalente a la disminución en la construcción de carreteras. ¿Cuál sería el cambio (si es que alguno tiene lugar) en los ingresos y el empleo?

La respuesta es compleja por el hecho de que la construcción de carreteras y los proyectos militares usan el dinero de maneras distintas. Por ello, aunque podría darse un aumento en los ingresos y en el nivel de ocupación entre los trabajadores de industrias como las fabricantes de aviones y barcos, eso podría no compensar las pérdidas y el desempleo en el sector de las obras públicas y la construcción (al menos a corto plazo). El problema radica en que en la economía de un país como Estados Unidos se producen muchos bienes y servicios altamente relacionados entre sí. Los aumentos o recortes en una industria se manifiestan frecuentemente también en otras industrias.

Un modelo para analizar estos efectos fue desarrollado por el economista Wassily W. Leontief en 1936. Tal modelo se llama *análisis de entradas y salidas* o *input–output*. Antes de describir en detalle este modelo presentamos un ejemplo sencillo.

Ejemplo 1. Considérese un modelo muy simplificado de una economía en la que se producen dos artículos: automóviles (incluyendo camiones) y acero. Cada año se da una *demanda externa* de 360000 toneladas de acero y de 110000 automóviles. Aquí la palabra *externa* significa que la demanda proviene de fuera de la economía.

Por ejemplo, si fuera un modelo de una porción de la economía de España, la demanda podría venir de otros países (de tal manera que el acero y los automóviles se exportarían), de otras industrias en España y de empresas privadas.

Pero la demanda externa no es la única que se da en las dos industrias consideradas. Se requiere acero para producir automóviles. También se requieren automóviles para producir automóviles, porque las plantas manufactureras de esos vehículos requieren coches y camiones para transportar los materiales y los empleados. De igual manera, la industria del acero requiere acero (para su maquinaria) y automóviles (para el transporte del producto y de los trabajadores) en su operación. Así que cada una de las dos industrias en el sistema impone demandas a sí misma y a la otra industria. Estas acciones se llaman *demandas internas.*

En nuestro modelo simplificado, se puede suponer que la industria del acero requiere $\frac{1}{4}$ de tonelada de acero y $\frac{1}{12}$ de automóvil (o camión) para producir 1 tonelada de acero (es decir, se usa un automóvil o camión en la producción de 12 toneladas de acero). También la industria automotriz requiere de $\frac{1}{2}$ tonelada de acero y $\frac{1}{9}$ de vehículo para producir un automóvil. La pregunta planteada por el modelo de Leontief de entradas y salidas es entonces la siguiente: ¿cuántas toneladas de acero y cuántos automóviles se deben producir cada año para que la disponibilidad de cada uno sea igual a la demanda total?

Solución. Sean x e y el número total de toneladas de acero y el número de automóviles, respectivamente, en cierto año. Esto constituye la oferta (o lo disponible). Si, por ejemplo, se requiere $\frac{1}{4}$ de tonelada de acero para producir una tonelada de este metal, se necesita entonces $\frac{1}{4}x$ toneladas de acero para producir x toneladas de acero. Similarmente, se requiere $\frac{1}{2}y$ toneladas de acero para producir y automóviles. Por consiguiente, el total de la demanda interna en la industria productora de acero es $\frac{1}{4}x+\frac{1}{2}y$, y la demanda total (sumando la demanda externa) es $\frac{1}{4}x+\frac{1}{2}y+360000$. De manera análoga, la demanda total en la industria automotriz es $\frac{1}{12}x+\frac{1}{9}y+110000$. Igualando la oferta con la demanda se obtiene el sistema

$$\begin{aligned} x &= \frac{1}{4}x+\frac{1}{2}y+360000 \\ y &= \frac{1}{12}x+\frac{1}{9}y+110000 \end{aligned}$$

o equivalentemente

$$\begin{aligned} \frac{3}{4}x-\frac{1}{2}y &= 360000 \\ -\frac{1}{12}x+\frac{8}{9}y &= 110000 \end{aligned}$$

Resolviendo dicho sistema, llegamos a que $x=600000$, $y=180000$. Esto es, para que la oferta sea exactamente igual a la demanda, se deben producir 600000 toneladas de acero y 180000 automóviles. □

Ahora describimos el modelo general de Leontief de entradas y salidas. Supóngase que un sistema económico tiene n industrias. De nuevo, hay dos clases de demandas en cada industria. Primero está la demanda *externa* de fuera del sistema. Si el sistema es un país, por ejemplo, la demanda externa podría ser de otro país. En segundo lugar está la demanda de una industria sobre otra, dentro del mismo sistema.

Sea e_i la demanda externa sobre la industria i. Sea a_{ij} la demanda interna de la industria j sobre la industria i. De un modo más preciso, a_{ij} representa el número de unidades de producto de la industria i necesarias para producir 1 unidad de producto de la industria j. Sea x_i la producción de la industria i. Ahora suponemos que la producción de cada industria es igual a su demanda (es decir, no hay sobreproducción). La demanda total es igual a la suma de las demandas interna y externa. Para calcular la demanda interna en la industria 2, por ejemplo, notamos que $a_{21}x_1$ es la demanda sobre la industria 2 por parte de la industria 1. Así, el total de la demanda interna sobre la industria 2 es

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n$$

De esta forma llegamos al siguiente sistema de ecuaciones, el cual se obtiene al igualar la demanda total con la producción de cada industria:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + e_1 &= x_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + e_2 &= x_2 \\
\vdots \quad \vdots \qquad\quad \vdots \qquad &\quad \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + e_n &= x_n
\end{aligned}$$

o, escribiéndolo de otra manera,

$$\begin{aligned}
(1-a_{11})x_1 - a_{12}x_2 - \cdots - a_{1n}x_n &= e_1 \\
-a_{21}x_1 + (1-a_{22})x_2 - \cdots - a_{2n}x_n &= e_2 \\
\vdots \qquad \vdots \qquad &\quad \vdots \\
-a_{n1}x_1 - a_{n2}x_2 - \cdots + (1-a_{nn})x_n &= e_n
\end{aligned}$$

Muchas veces conviene escribir los números a_{ij} en forma de matriz A, llamada *matriz de tecnología* o también *matriz de requerimientos directos*. Se tiene así

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

la cual es una matriz cuadrada. Obsérvese entonces que el sistema de ecuaciones anterior se puede escribir de una manera mucho más simple como

$$(I - A)X = E$$

donde X es la matriz columna de las producciones x_i y E es la matriz columna de las demandas externas x_i. La matriz $I - A$ en este modelo se llama *matriz de Leontief.*

Suponiendo que la matriz de Leontief es invertible, la matriz de producción X puede expresarse como

$$X = (I - A)^{-1}E$$

Esta forma de escribir la matriz de producción tiene algunas ventajas. La matriz de tecnología A es la matriz de las demandas internas, las cuales, en periodos relativamente largos, permanecen fijas. Sin embargo, la matriz E de la demanda externa puede cambiar con cierta frecuencia. Normalmente se requiere mucho cálculo para obtener $(I-A)^{-1}$, pero una vez obtenida esta matriz se puede encontrar la matriz de producción X correspondiente a cualquier matriz E de demanda mediante una simple multiplicación de matrices. Si no se halla $(I-A)^{-1}$, habría que resolver un sistema de ecuaciones para cada matriz E.

Ejemplo 2. Supóngase que en un sistema económico con tres industrias las demandas externas son, respectivamente, de 10, 25 y 20. Considérese que $a_{11}=0{,}2$, $a_{12}=0{,}5$, $a_{13}=0{,}15$, $a_{21}=0{,}4$, $a_{22}=0{,}1$, $a_{23}=0{,}3$, $a_{31}=0{,}25$, $a_{32}=0{,}5$ y $a_{33}=0{,}15$. Encontrar la producción en cada industria para equilibrar con exactitud la oferta con la demanda. Repetir el problema para unas demandas externas de 15, 20, 40.
Solución. En este caso $n=3$ y la matriz A de tecnología está dada por

$$A=\begin{pmatrix} 0{,}2 & 0{,}5 & 0{,}15 \\ 0{,}4 & 0{,}1 & 0{,}3 \\ 0{,}25 & 0{,}5 & 0{,}15 \end{pmatrix}$$

con lo cual la matriz de Leontief es

$$I-A=\begin{pmatrix} 0{,}8 & -0{,}5 & -0{,}15 \\ -0{,}4 & 0{,}9 & -0{,}3 \\ -0{,}25 & -0{,}5 & 0{,}85 \end{pmatrix}$$

Calculando su inversa, tenemos

$$(I-A)^{-1}=\begin{pmatrix} 2{,}78594 & 2{,}26497 & 1{,}29103 \\ 1{,}87992 & 2{,}91048 & 1{,}35896 \\ 1{,}92521 & 2{,}37818 & 2{,}35555 \end{pmatrix}$$

y así, tenemos en el primer caso

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}=\begin{pmatrix} 2{,}78594 & 2{,}26497 & 1{,}29103 \\ 1{,}87992 & 2{,}91048 & 1{,}35896 \\ 1{,}92521 & 2{,}37818 & 2{,}35555 \end{pmatrix}\begin{pmatrix} 10 \\ 25 \\ 20 \end{pmatrix}=\begin{pmatrix} 110{,}30 \\ 118{,}74 \\ 125{,}82 \end{pmatrix}$$

mientras que en el segundo caso, $x_1 \simeq 139$, $x_2 \simeq 141$ y $x_3 \simeq 171$. □

1.5.2. Modelización lineal

Supongamos que la práctica sugiere que una variable dependiente y es una función lineal de $k-1$ variables independientes $x_2, x_3, \ldots, x_k$. Por ejemplo, la cantidad demandada de un cierto bien del mercado y puede considerarse como una función lineal del precio de este bien, x_2, del precio de otros bienes que influyan sobre la cantidad demandada de éste, $x_3, \ldots, x_{k-2}$ y de la renta de los consumidores, x_{k-1}. Así, podríamos escribir

$$y=\beta_1+\beta_2x_2+\beta_3x_3+\cdots+\beta_kx_k+u \tag{1.9}$$

siendo u una variable que trata de dar cuenta del error existente en el modelo.

El problema que surge en la modelización económica consiste esencialmente en determinar el valor de los parámetros desconocidos $\beta_1, \ldots, \beta_k$ de tal manera que los errores en el modelo sean mínimos, a partir de un número determinado de observaciones muestrales del comportamiento en la realidad de las variables consideradas.

Supongamos que disponemos de n observaciones muestrales para cada una de las variables que intervienen en el modelo, recogidas en n instantes del tiempo diferentes, esto es, en términos matriciales disponemos de las siguientes matrices columna de datos:

$$Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}, \quad X_2 = \begin{pmatrix} x_{21} \\ \vdots \\ x_{2n} \end{pmatrix}, \ldots, X_k = \begin{pmatrix} x_{k1} \\ \vdots \\ x_{kn} \end{pmatrix}.$$

Entonces, para que se verifique la ec. (1.9), se deben satisfacer las n igualdades

$$y_i = \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + \cdots + \beta_k x_{ki} + u_i, \qquad i = 1, \ldots, n,$$

igualdades que pueden reescribirse matricialmente como

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{21} & \ldots & x_{k1} \\ \vdots & \vdots & & \vdots \\ 1 & x_{2n} & \ldots & x_{kn} \end{pmatrix} \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \tag{1.10}$$

o simbólicamente,

$$Y = XB + U \tag{1.11}$$

donde Y y X son datos del problema, mientras que B y U son matrices, en principio, desconocidas.

Una vez formulado el modelo observacional en la forma matricial (1.10) o (1.11), es necesario introducir una serie de hipótesis adicionales sobre el comportamiento de los distintos elementos del mismo (como por ejemplo que la matriz X tiene rango k, para lo cual se debe suponer que $n \geq k$) para poder abordar el problema de la estimación de los valores de los parámetros β_i. Si admitimos estas hipótesis subyacentes en el proceso de estimación denominado de *mínimos cuadrados*, se puede comprobar que la "mejor estimación" de la matriz B (en el sentido de mínimos cuadrados) viene dada por la expresión

$$B = (X^T X)^{-1} X^T Y.$$

Sin profundizar más en el modelo empírico aquí expuesto, nuestra intención con la inclusión del mismo es dejar patente las múltiples aplicaciones que encuentran las matrices en la Economía y, en este caso particular, en su construcción empírica.

Ejercicio. Supongamos que la demanda de un cierto bien del mercado y viene dada como una función lineal del precio de este bien, x, según la ecuación

$$y = \beta_1 + \beta_2 x + u$$

Además, supongamos que conocemos los datos muestrales

$$Y = \begin{pmatrix} 10 \\ 20 \\ 15 \\ 25 \\ 30 \end{pmatrix}, \qquad X = \begin{pmatrix} 6 \\ 3 \\ 4 \\ 2 \\ 1 \end{pmatrix}.$$

Estimar los valores mínimos cuadráticos de β_1 y β_2, al igual que los errores generados por esta estimación. Finalmente, predecir la demanda de este bien que se tendrá para un precio del mismo de $x = 7{,}5$ (suponiendo el error que cometemos nulo).

1.6. Problemas

1.6.1. Operaciones con matrices

1. Dadas las matrices:

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

calcula $3A + 4B - 2C$.

2. Sean $A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ -3 & 4 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 & -5 \\ 3 & 4 & 0 \end{pmatrix}$. Encuentra AB, BA, BB^T, B^TB.

3. Dada la matriz $A = \begin{pmatrix} 1 & 3 \\ 4 & -3 \end{pmatrix}$, encuentra una matriz columna U diferente de cero de manera que $AU = 3U$.

Matrices cuadradas. Determinantes. Inversas.

1. Halla la inversa de $C = \begin{pmatrix} -1 & 2 & -3 \\ 2 & 1 & 0 \\ 4 & -2 & 5 \end{pmatrix}$, $D = \begin{pmatrix} 2 & 1 & -1 \\ 0 & 2 & 1 \\ 5 & 2 & -3 \end{pmatrix}$

2. Eval˙a el determinante de

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

3. Calcula la potencia n-ésima de la matriz $A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$

4. Demuestra que la matriz $C = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$ es ortogonal ($C^{-1} = C^T$)

5. Sean A y B dos matrices cuadradas, con A regular. Sabiendo que $AB = BA$, demuestra que $A^{-1}B = BA^{-1}$.

6. Demuestra que toda matriz cuadrada se puede expresar como una suma de una matriz simétrica más una matriz antisimétrica

7. Demuestra que la matriz A es idempotente ($A^2 = A$), siendo $A = \begin{pmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{pmatrix}$

8. Calcula los valores de m y n para que la matriz A verifique la ecuación $A^2 + mA + nI = 0$, siendo $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

9. Calcula las matrices X e Y que verifiquen el sistema
$$2X - 5Y = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \quad -X + 3Y = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$$

10. Calcula la matriz X en la ecuación
$$X \cdot \begin{pmatrix} 1 & 1 & -1 \\ 2 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 3 \\ 4 & 3 & 2 \\ 1 & -2 & 5 \end{pmatrix}$$

11. Calcula A^p, siendo $A = \begin{pmatrix} 1 & 1/p & 1/p \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

12. a) Construye la matriz $\mathbf{A}$ 2×2 teniendo $a_{ij} = (-1)^{i+j}$. b) Construye la matriz $\mathbf{A}$ 3×3 con $a_{ij} = i/j$. c) Construye la matriz $\mathbf{D}$ 3×4 teniendo
$$d_{ij} = \begin{cases} i+j & i > j \\ 0 & i = j \\ i-j & i < j \end{cases}$$

13. Sean A y S dos matrices cuadradas, siendo S simétrica. Analiza si se verifica que:

    a) A^TA es simétrica.

    b) A^TSA es simétrica.

    c) A antisimétrica $\Rightarrow A^2$ simétrica.

14. Sean A y B dos matrices cuadradas y simétricas. Demuestra que AB es simétrica si y sólo si $AB = BA$.

15. Dadas las matrices $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ y $B = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}$ calcula las matrices P y Q de manera que se cumpla: $PQ = A$ y $P + Q^{-1} = B$.

16. Dadas las matrices $A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 2 & 1 \\ 3 & 2 & 1 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ calcula las matrices R y S de manera que se cumpla: $RS = A$ y $R + S^{-1} = B$.

17. Dadas las matrices A y B tales que $B = P^{-1}AP$, siendo A simétrica y P ortogonal, demuestra que B es simétrica.

18. Dadas las matrices A y B tales que $B = PAP^{-1}$, siendo A antisimétrica y P ortogonal, demuestra que B es antisimétrica.

1.6.2. Aplicaciones de las matrices a la Economía

1. Una empresa, además de pagar a sus ejecutivos un salario extraordinario, a manera de gratificación anual, les da acciones de la compañía. Durante el año 2000 el presidente recibió 8000 euros y 50 acciones, cada uno de los tres vicepresidentes 4500 euros y 20 acciones y el tesorero 4000 euros y 10 acciones. Se pide.

   a) Expresar estos datos en términos de una matriz A.

   b) Expresar el número de ejecutivos de cada rango mediante un vector columna $\vec{x}$.

   c) ¿Qué representa $A\vec{x}$?

2. La empresa IMAGE DEVELOPMENT COMPANY (IDC) fabrica en su planta de Zaragoza tres tipos de televisores: de 14, 21 y 25 pulgadas. Los almacenes principales se encuentran en Madrid, Barcelona, Vigo y Sevilla. Las ventas durante el año 2005 del almacén de Madrid se cifraron en 400,100 y 500 televisores de 14, 21 y 25 pulgadas respectivamente; las del almacén de Barcelona en 300, 150 y 400; las del almacén de Vigo en 100, 100 y 200; y las de Sevilla en 200, 150 y 300. Los precios de venta de los televisores en 2005 fueron de 120 euros los de 14 pulgadas, de 300 los de 21 y de 500 los de 25 pulgadas. Se pide:

   a) Expresar estos datos en términos de una matriz A.

   b) Expresar el precio de cada tipo de televisor mediante un vector columna $\vec{x}$.

   c) ¿Qué representa $A\vec{x}$?

3. El almacén 1 de una cadena de 3 tiendas de electrodomésticos tiene 3 neveras, 5 hornos, 3 lavadoras y 4 secadoras en stock. El almacén 2 no tiene ninguna nevera, pero sí tiene 2 hornos, 9 lavadoras y 5 secadoras, mientras que el almacén 3 tiene en stock 4 neveras, 2 hornos y ninguna lavadora ni secadora. Presentar el inventario de toda la cadena en forma de matriz.

4. El número de artículos estropeados servidos por la Compañía de Colchones S.A. desde todas sus plantas durante el año pasado está dado por la matriz de

daños

$$\begin{pmatrix} 80 & 12 & 16 \\ 50 & 40 & 16 \\ 90 & 10 & 50 \end{pmatrix}$$

Las filas pertenecen a sus tres plantas en Madrid, Tarragona y Zaragoza; las columnas pertenecen a sus modelos regular, firme y extra-firme, respectivamente. El objetivo de la compañía para el año siguiente es reducir en un 10 % el número de colchones regulares dañados servidos por cada planta, reducir en un 20 % el número de colchones firmes suministrados por su planta de Tarragona, reducir en un 30 % el número de colchones extrafirmes dañados servidos por su planta de Zaragoza y mantener todas las demás entradas como el último año. ¿Cuál será la matriz de daños si todos los objetivos se cumplen?

5. La tabla de precios para el vuelo Chicago-Los Angeles está dado por $\mathbf{P} = (200, 350, 500)$, donde los elementos de la matriz fila pertenecen, respectivamente, a clase turista, clase de negocios y primera clase. El número de billetes vendidos en cada clase para un vuelo particular está dado por la matriz columna
$$\mathbf{N} = \begin{pmatrix} 130 \\ 20 \\ 10 \end{pmatrix}$$
Calcular los productos $\mathbf{PN}$ y $\mathbf{NP}$, y determinar el significado de cada uno.

6. Los tiempos necesarios para que una compañía produzca tres productos están contenidos en la siguiente matriz
$$\mathbf{T} = \begin{pmatrix} 0{,}2 & 0{,}5 & 0{,}4 \\ 1{,}2 & 2{,}3 & 1{,}7 \\ 0{,}8 & 3{,}1 & 1{,}2 \end{pmatrix}$$
donde las filas pertenecen a bases de lámparas, armarios y mesas, respectivamente. Las columnas indican las horas de trabajo requeridas para cortar la madera, pegar las piezas y pintar. El salario (por hora) de un carpintero para cortar la madera, de un artesano para ensamblar las piezas y de un decorador para pintar el producto está dado por las entradas de la siguiente matriz columna $\mathbf{W} = \begin{pmatrix} 10{,}50 \\ 14 \\ 12{,}25 \end{pmatrix}$

    a) Calcular el producto $\mathbf{TW}$ y explicar su significado.

    b) Si el número de artículos a fabricar de bases de lámparas, armarios y mesas está dado por la matriz fila $\mathbf{Q} = (1000 \quad 100 \quad 200)$, calcular el producto $\mathbf{QTW}$ y explicar su significado.

1.6.3. Sistemas de ecuaciones lineales

1. Usa el método de eliminación gaussiana para resolver los sistemas

a) $\left.\begin{array}{rcrcrcr} x & + & 4y & - & z & = & 0 \\ -4x & - & 8y & + & 2z & = & 0 \\ -2x & - & 4y & + & z & = & -1 \end{array}\right\}$

b) $\left.\begin{array}{rcrcrcr} -3x & + & 6y & - & 3z & = & 0 \\ x & - & 2y & + & z & = & 0 \\ x & - & 2y & + & z & = & 0 \end{array}\right\}$

2. Discute y resuelve en su caso los siguientes sistemas:

a) $\left.\begin{array}{rcrcrcr} 2x & + & 3y & - & z & = & 0 \\ 4x & + & 6y & - & z & = & 0 \\ 8x & + & 12y & - & 3z & = & 0 \end{array}\right\}$

b) $\left.\begin{array}{rcrcrcrcr} & & y & - & z & - & 2t & = & 1 \\ x & & & - & z & - & t & = & -2 \\ x & + & y & & & - & 3t & = & -1 \end{array}\right\}$

c) $\left.\begin{array}{rcrcrcr} x & + & 4y & + & z & = & b \\ 3x & - & y & + & 2z & = & 1 \\ 2x & - & 5y & + & az & = & -2 \end{array}\right\}$

d) $\left\{\begin{array}{rcrcrcr} 4x & + & 2y & + & z & = & ax \\ 2x & + & 4y & + & 2z & = & ay \\ 2x & + & 4y & + & 8z & = & az \end{array}\right.$

e) $\left\{\begin{array}{rcrcrcrcr} x_1 & + & 3x_2 & + & x_3 & - & x_4 & = & 6 \\ 2x_1 & + & 7x_2 & + & 3x_3 & - & 4x_4 & = & 15 \\ x_1 & + & x_2 & + & 2x_3 & + & x_4 & = & 1 \end{array}\right.$

Aplicaciones a la Economía

1. Un fabricante produce escritorios y estanterías. Los escritorios d requieren un tiempo de 5 horas para cortar y 10 horas para ensamblar las piezas. Las estanterías b requieren 15 min. para cortar y 1 hora para ensamblar. Cada día el fabricante tiene disponibles 200 horas de tiempo de personal para cortar y 500 horas para ensamblar. ¿Cuántos escritorios y estanterías se pueden fabricar cada día usando toda la potencia de trabajo disponible?

2. Una compañía minera tiene un contrato para suministrar 70000 Tm de mineral de bajo grado, 181000 Tm de mineral de grado medio y 41000 Tm de mineral de alto grado. La compañía tiene 3 minas en explotación. La mina A produce 8000 Tm de mineral de bajo grado, 5000 Tm de mineral de grado medio y 1000 Tm de mineral de alto grado en cada día de explotación. La mina B produce 3000 Tm de mineral de bajo grado, 12000 Tm de mineral de grado medio y 3000 Tm de mineral de alto grado por cada día que está en explotación. Los números para la mina C son 1000, 10000 y 2000, respectivamente. ¿Cuántos días ha de estar cada mina operando para satisfacer las demandas contractuales sin producir un excedente?

3. En una fábrica se produce queso y mantequilla. Para fabricar una unidad de queso se precisan 10 unidades de leche y 6 horas de mano de obra. Para la mantequilla se necesitan 5 unidades de leche y 8 horas de mano de obra por unidad. Sabiendo que tenemos disponibles 100 unidades de leche y 110 de mano de obra, calcular la producción posible de queso y mantequilla, considerando que utilizamos todo lo disponible.

4. Supongamos un proceso generador de dos productos, los cuales deben pasar por dos máquinas A y B, consumiendo el primer producto 2 unidades de tiempo (u.t.) de la primera y 5 u.t. de la segunda, y el segundo producto 4 u.t. de A y 3 de B. Si el tiempo disponible es de 35 u.t. en A y 40 u.t. en B, determinar si existe alguna producción posible y en su caso obtenerla.

5. El Ministerio de Economía en su departamento de transacciones exteriores ha obtenido las relaciones que ligan el saldo de la balanza de mercancías (exportaciones menos importaciones de bienes), el de la balanza de servicios (ídem de servicios) y el de la de capital (entradas menos salidas de capital a largo plazo) con tres variables básicas como son la tasa de inflación (π), el tipo de interés (r) y el tipo de cambio euro/dólar (e), relaciones que se resumen en las ecuaciones
$$\begin{array}{rcrcrcr} SBM & = & & - & 300\pi & + & 10e \\ SBS & = & & - & 80\pi & + & 6e \\ SBC & = & 60r & - & 50\pi & + & 5e \end{array}$$
¿Cuáles tendrían que ser los valores de las variables r, π y e si se desea que este trimestre la balanza de mercancías tenga un déficit de 1000, la de servicios un superávit de 200 y la de capital un superávit de 1200?

1.6.4. Modelos de Leontief

1. Consideremos una economía dividida en un sector agrícola (A) y un sector industrial (I). Para producir una unidad de A se necesitan 1/6 de unidades de A y 1/4 de unidades de I. Para producir una unidad de I se necesitan 1/4 de unidades de A y 1/4 de unidades de I. Supongamos que las demandas finales en cada uno de los sectores son de 60 unidades.

   a) Escribir el sistema de Leontief de esta economía.

   b) Hallar el número de unidades que hay que producir en cada sector para cubrir las demandas finales.

2. Pablo, Jaime y María deciden ayudarse mutuamente para construir casas. Pablo empleará la mitad de su tiempo en su propia casa y una cuarta parte de su tiempo en cada una de las casas de Jaime y María. Jaime empleará un tercio de su tiempo en cada una de las casas en construcción. María empleará un sexto de su tiempo en la casa de Pablo, un tercio en la casa de Jaime y la mitad de su tiempo en su propia casa. Por razones de impuestos, cada uno ha de fijar un precio a su trabajo, pero ellos lo quieren hacer de manera que nadie gane ni pierda dinero. Mostrad que el proceso de determinar los salarios de

acuerdo con esas premisas es un modelo cerrado de Leontief conteniendo tres ecuaciones homogéneas y encontrad después los salarios de cada persona.

3. Consideremos cuatro países del Tercer Mundo y supongamos que cada uno de ellos produce un tipo de fruta diferente destinada a exportación, y que cada uno usa el dinero obtenido con la venta para pagar la importación de las frutas de los otros países. El país A exporta el 20 % de su fruta al país B, el 30 % al país C, el 35 % al país D y usa el resto de su fruta para consumo interno. El país B exporta el 10 % de su fruta al país A, el 15 % al país C, el 35 % al país D y retiene el resto para sus propios ciudadanos. El país C no exporta al país A; divide su cosecha por igual entre los países B y D y su propia población. El país D no consume su propia fruta, sino que la exporta toda, con un 15 % yendo al país A, un 40 % al país B y un 45 % al país C. Mostrad que el problema de determinar los precios de las cosechas anuales de fruta de modo que cada país no salga beneficiado ni perjudicado es equivalente a resolver cuatro ecuaciones homogéneas con cuatro incógnitas y después hallad los precios.

4. Consideremos un modelo input-output con tres sectores. El sector 1 es industria pesada, el 2 es industria ligera y el 3 es agricultura. Supongamos que los requerimientos de input están dados por la siguiente tabla:

	Ind. pesada	Ind. ligera	Agricultura
Unid. bienes ind. pesada	0.1	0.2	0.1
Unid. bienes ind. ligera	0.3	0.2	0.2
Unid. bienes agric.	0.2	0.2	0.1

y que las demandas finales de los tres bienes sean de 85, 95 y 20 unidades, respectivamente. Escribe el modelo de Leontief del problema y encontrar una solución.

5. Una versión muy simplificada de la tabla de entradas y salidas para la economía de Israel en 1958 divide la economía en tres sectores, agricultura, manufacturas y energía, con el siguiente resultado:

	Agricultura	Manufactura	Energía
Agricultura	0.293	0	0
Manufactura	0.014	0.207	0.017
Energía	0.044	0.010	0.216

Las exportaciones (en miles de libras israelíes) en 1958 fueron las siguientes:

Agricultura: 138213; Manufacturas: 17597; Energía: 1786

a) ¿Cuántas unidades de producción agrícola se requieren para generar una unidad de producción agrícola?

b) Determina el total de libras israelíes en productos agrícolas, productos manufacturados y en energía requeridos para activar este modelo de la economía israelí y exportar los valores indicados de productos.

1.6.5. Matrices de transición

1. Uno de los temas que más ha preocupado a los economistas es el estudio de la distribución de la riqueza. Supongamos que en un país imaginario la población está dividida en cuatro clases, A, B, C y D, de acuerdo con su riqueza (de mayor a menor y según algún criterio dado); una persona que se encuentra en una determinada posición en un momento dado puede ascender, mantenerse o descender en el siguiente con probabilidades dadas por la matriz adjunta,

	A	B	C	D
A	0.7	0.2	0.1	0
B	0.2	0.4	0.1	0.3
C	0.1	0.3	0.4	0.2
D	0	0.1	0.4	0.5

   donde el elemento a_{ij} la probabilidad de que un individuo que en un momento dado pertenece a la clase j en el siguiente período pertenezca a la clase i. Se pide:

   a) Si en el año 2007 el 17 % de la población pertenece a la clase A, el 24 % a la B, el 30 % a la C y el 29 % a la D, ¿cuál será la distribución en 2008? ¿y en 2009?

   b) ¿Cuál fue la distribución en 2006?

2. En la comunidad de Murcia el 15 % de las rentas familiares anuales en 2006 son inferiores a 12 000 euros, el 80 % está comprendido entre 12 000 y 24 000 euros y el 5 % restante supera esta cifra. Se sabe que año tras año, el 50 % de las familias con renta baja permanece en dicho tramo, mientras que un 30 % pasan a renta media y un 20 % a renta alta. De las familias con renta media, un 60 % permanecen en este tramo, un 20 % pasan a renta baja y otro 20 % a renta alta. Por último, el 90 % de la familias con renta alta siguen siéndolo mientras que el 10 % restante pasan a renta media.

   a) ¿Cúal es el porcentaje de familias en cada uno de los tramos en 2007? ¿Y en 2008?

3. Los obreros, dentro de la población activa de un país, se clasifican en dos categorías profesionales: obreros especializados (x) y obreros no especializados (y). Se sabe que:

   - Cada trabajador activo tiene sólo un hijo.
   - Los hijos de los obreros especializados se reparten en las dos categorías según los porcentajes 60 % y 40 %.
   - Para los hijos de los obreros no especializados estos porcentajes son 40 %, 60 %.

   Se pide:

a) Escribir el vector que representa las categorías profesionales de los obreros del país y plantear el modelo que representa la distribución de esta fuerza del trabajo del país de generación en generación.

b) Si en 2000 hay 3 millones de obreros especializados y 6 de obreros no especializados ¿cuántos habrá en la siguiente generación?

Capítulo 2

ESPACIOS VECTORIALES

2.1. Introducción. Vectores en el plano.

Supongamos que una tienda vende n bienes distintos. Cada mes anota el número de unidades que hay de cada bien: $v_1, v_2, v_3, \ldots, v_n$. Conviene representar estas existencias por una fila $(v_1, v_2, v_3, \ldots, v_n)$ o una columna $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$.

Un conjunto ordenado de números que se distingue no sólo por los elementos que contiene, sino también por el orden en que están dispuestos los elementos se conoce como **vector**, o n**-upla**. Los vectores anteriores se llaman vector fila y vector columna, respectivamente, y también pueden ser considerados como matrices fila $1 \times n$ o matrices columna $n \times 1$.

Los vectores se escriben en negrita o con flechas. Los números $v_1, v_2, v_3, \ldots, v_n$ se llaman **componentes** o coordenadas del vector. El número de componentes da la **dimensión** del vector.

El término *coordenadas* proviene de la representación de los vectores respecto de los ejes coordenados. Así, los vectores de dimensión 2 se pueden representar en el plano. Su primera componente coincide con la proyección del vector sobre el eje X y se conoce también como coordenada x; su segunda componente o coordenada y coincide con la proyección del vector sobre el eje Y.

2.2. Operaciones con vectores.

Dos vectores de la misma dimensión se dice que son iguales si sus componentes son todas iguales.

Dados los vectores $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$ y $\mathbf{b} = (b_1, b_2, b_3, \ldots, b_n)$, se definen las siguientes operaciones:
suma: $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots, a_n + b_n)$.
multiplicación por un escalar: $\alpha\mathbf{a} = (\alpha a_1, \alpha a_2, \alpha a_3, \ldots, \alpha a_n)$.
diferencia: $\mathbf{a} - \mathbf{b} = (a_1 - b_1, a_2 - b_2, a_3 - b_3, \ldots, a_n - b_n)$.

El vector nulo es aquel que tiene 0 en todas sus componentes: $\mathbf{0} = (0, 0, 0, \ldots, 0)$.

Si **a** y **b** tienen la misma dimensión y α y β son números reales, el vector $\alpha\mathbf{a}+\beta\mathbf{b}$ se conoce como **combinación lineal** de los vectores **a** y **b**.

Ejemplo 1 *Una tienda dispone inicialmente de 2 lavadoras, 3 frigoríficos, 5 estufas y 6 microondas. Al mes siguiente recibe 1 frigorífico, 2 estufas, 3 microondas y 5 lavadoras. ¿Cuántos tendrá en total?*

Solución. **a**=(lavadoras, frigoríficos, estufas, microondas)
$\mathbf{a}_0 = (2,3,5,6)$, $\mathbf{a}_1 = (5,1,2,3)$, $\mathbf{a}_0 + \mathbf{a}_1 = (7,4,7,9)$.

Propiedades

Si **a**, **b** y **c** son vectores arbitrarios con la misma dimensión y α, β son números reales se cumple:

1. $(\mathbf{a}+\mathbf{b})+\mathbf{c} = \mathbf{a}+(\mathbf{b}+\mathbf{c})$;
2. $\mathbf{a}+\mathbf{b} = \mathbf{b}+\mathbf{a}$;
3. $\mathbf{a}+\mathbf{0} = \mathbf{0}+\mathbf{a} = \mathbf{a}$;
4. $\mathbf{a}+(-\mathbf{a}) = \mathbf{0}$;
5. $(\alpha+\beta)\,\mathbf{a} = \alpha\mathbf{a}+\beta\mathbf{a}$;
6. $\alpha\,(\mathbf{a}+\mathbf{b}) = \alpha\mathbf{a}+\alpha\mathbf{b}$;
7. $(\alpha\beta)\,\mathbf{a} = \alpha\,(\beta\mathbf{a})$;
8. $1\mathbf{a} = \mathbf{a}$.

Ejercicio 1 *Si* $3\,(x,y,z)+5\,(-1,2,3) = (4,1,3)$*, calcular el vector* (x,y,z)*.*

2.3. Interpretación geométrica de los vectores.

Es habitual representar un vector (o una matriz fila) (a,b) como una flecha (o segmento dirigido) con punto inicial el origen de coordenadas y punto final el punto (a,b) de $\mathbb{R}^2$. La flecha queda determinada por su dirección y longitud. Dos vectores se dice que son iguales o equivalentes si tienen la misma longitud y están sobre rectas paralelas con el mismo sentido.

Ejercicio 2 *Representar el vector que une los puntos* $(2,3)$ *y* $(-1,1)$*, un vector equivalente a él cuyo origen sea el punto* $(1,1)$ *y un vector equivalente a ambos cuyo origen sea el origen de coordenadas.*

Las componentes de estos vectores coinciden con las proyecciones sobre el origen de coordenadas del vector $\mathbf{a} = (-3,-2)$.

Se define la **norma** de un vector como su longitud:

$$\mathbf{a} = (a_1,a_2,a_3,\ldots,a_n) \Rightarrow \|\mathbf{a}\| = \sqrt{a_1^2+a_2^2+a_3^2+\cdots+a_n^2}.$$

Así, por ejemplo, la norma del vector $\mathbf{a} = (-3, -2)$ es $\|\mathbf{a}\| = \sqrt{(-3)^2 + (-2)^2} = \sqrt{13}$.

En el caso de vectores de dimensión 2 la suma, diferencia y multiplicación por un escalar pueden representarse también gráficamente de manera sencilla.

También pueden dibujarse los vectores de tres componentes en el espacio tridimensional, cada una de las componentes correspondiendo en este caso a la proyección del vector sobre cada uno de los ejes coordenados.

Ejercicio 3 *Dados los vectores* $\mathbf{a} = (2, 1, 1)$ *y* $\mathbf{b} = (-1, 1, 2)$*, calcular y dibujar:* $\mathbf{a} + \mathbf{b}$*,* $\mathbf{a} - \mathbf{b}$*,* $2\mathbf{a}$*,* $2\mathbf{b}$*.*

2.4. Producto escalar. Propiedades. Vectores ortogonales.

Ejercicio 4 *Considérese una tienda que dispone de 2 estufas, 2 cocinas y 3 neveras. El precio de una estufa es 20 000 ptas, las cocinas valen 35 000 y las neveras 60 000 ptas. Si la tienda vende todos los electrodomésticos, ¿cuánto dinero obtendrá?*

Definición 1 *Se define el* ***producto escalar*** *de dos vectores* $\mathbf{a} = (a_1, a_2, a_3, \ldots, a_n)$ *y* $\mathbf{b} = (b_1, b_2, b_3, \ldots, b_n)$ *como el escalar que resulta de hacer las operaciones*

$$\mathbf{a} \cdot \mathbf{b} = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

Propiedades

1. $\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a}$
2. $\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c}$
3. $\mathbf{a} \cdot (\alpha \mathbf{b}) = (\alpha \mathbf{a}) \cdot \mathbf{b} = \alpha (\mathbf{a} \cdot \mathbf{b})$
4. $\mathbf{a} \cdot \mathbf{a} \geq 0$, $\mathbf{a} \cdot \mathbf{a} = 0$ si y sólo si $\mathbf{a} = \mathbf{0}$

Ejercicio 5 *Pepe va a una frutería y compra 3 kg de cerezas, 4 de peras, 3 de naranjas y 5 de manzanas. Luis compra 2 de cerezas, 2 de naranjas y 3 de manzanas. Se sabe que 1 kg de cerezas vale 120 ptas, 1 de peras 190 ptas, 1 de naranjas 80 ptas y uno de manzanas 135 ptas. Se pide:*

1. *¿Cuál es el vector de compras de Pepe? ¿Y el de Luis?*
2. *¿Cuál es el vector de precios?*
3. *¿Cuánto gasta cada uno de ellos?*
4. *Comprobar que se satisfacen las propiedades anteriores.*

Teniendo en cuenta que al calcular el producto escalar de un vector consigo mismo se obtiene el cuadrado de su norma, se observa que el producto escalar de dos vectores está relacionado con sus normas y el ángulo que forman. De hecho, el ángulo que forman dos vectores puede calcularse como:

$$\cos\theta = \frac{\mathbf{a}\cdot\mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$$

Dos vectores son **ortogonales** si su producto escalar es 0. En el plano y en el espacio de dimensión 3 es fácil comprobar que esto equivale a que los vectores son perpendiculares entre sí.

$$\mathbf{a}\cdot\mathbf{b} \Leftrightarrow \cos\theta = 0 \Leftrightarrow \theta = \frac{\pi}{2}$$

Ejercicio 6 *Estudiar qué pares de vectores son ortogonales:*

1. $\mathbf{a} = (1, 2)$ *y* $\mathbf{b} = (-2, -1)$.
2. $\mathbf{a} = (1, -1, 2)$ *y* $\mathbf{b} = (-2, 0, 1)$.

Ejercicio 7 *Calcular el valor de x para que los siguientes vectores sean ortogonales:*

$$\mathbf{a} = (x, x, -x-2, x) \quad y \quad \mathbf{b} = (x, 1, 1, x)$$

2.5. Espacios vectoriales. Definición.

Una de las tareas básicas del análisis matemático consiste en identificar estructuras fundamentales que aparecen con alguna regularidad en diferentes situaciones, desarrollarlas en abstracto y aplicar después los resultados obtenidos a las diferentes situaciones particulares. De este modo se pueden entender simultáneamente muchas situaciones diferentes investigando las propiedades que rigen todas ellas. Las matrices parecen tener poco que ver con los polinomios y éstos a su vez con los segmentos orientados, pero resulta que todos ellos comparten características fundamentales que, una vez desarrolladas, proporcionan un mejor entendimiento.

¿Cuáles son algunas de las propiedades fundamentales de las matrices, los segmentos orientados, las n-uplas e incluso los polinomios? En primer lugar, *se pueden sumar*; así tenemos el concepto de *clausura bajo suma*: se tienen objetos definidos en un conjunto particular y se establece una operación de suma sobre esos objetos de manera que el resultado es un nuevo objeto del *mismo* conjunto. En segundo lugar, también se tiene el concepto de *clausura bajo multiplicación por un escalar*. Además, estas operaciones tienen ciertas propiedades que ya hemos visto en el caso de n-uplas y segmentos orientados.

Así pues, hemos identificado rápidamente una serie de características comunes. ¿Hay otras? Y lo que es más interesante, ¿cuál es el número más pequeño de características que necesitamos identificar de modo que todas las demás se sigan a partir de éstas?

Para empezar, creamos una nueva etiqueta, que aplicamos a cualquier conjunto de objetos que tengan estas características: *espacio vectorial*, y nos referimos a los objetos de este conjunto como *vectores*. Después podemos mostrar que las matrices, los segmentos orientados, las n-uplas, etc., son ejemplos particulares de espacios vectoriales, de manera que los podemos denominar con el término más general de vectores.

Definición 2 *Un conjunto de objetos* $V = \{\mathbf{a}, \ \ \mathbf{b}, \ \mathbf{c}, \ldots, \}$ *y de escalares* $\{\alpha, \beta, \gamma, \ldots\}$ *junto con una operación de suma vectorial* $+$ *sobre los objetos y un producto por un escalar* $\cdot$ *se dice que es un* **espacio vectorial** *si posee las siguientes diez propiedades:*
Suma de vectores

1. *ley interna:* $\forall\, \mathbf{a}, \mathbf{b} \in V \Rightarrow \mathbf{a} + \mathbf{b} \in V$
2. *conmutativa:* $\mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}$
3. *asociativa:* $\mathbf{a} + (\mathbf{b} + \mathbf{c}) = (\mathbf{a} + \mathbf{b}) + \mathbf{c}$
4. *elemento neutro: existe un vector* $\mathbf{0} \in V$ *tal que* $\mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}$
5. *elemento opuesto: para todo vector* $\mathbf{a} \in V$ *existe un vector* $-\mathbf{a}$*, llamado el opuesto de* $\mathbf{a}$ *tal que* $\mathbf{a} + (-\mathbf{a}) = (-\mathbf{a}) + \mathbf{a} = \mathbf{0}$

   Producto por un escalar
6. *ley externa:* $\forall\, \mathbf{a} \in V, \forall \alpha \in \mathbb{R} \Rightarrow \alpha\mathbf{a} \in V$
7. $\alpha\,(\beta\mathbf{a}) = (\alpha\beta)\,\mathbf{a} = (\alpha\mathbf{a})\,\beta$
8. *elemento unidad:* $1 \cdot \mathbf{a} = \mathbf{a}$
9. *distributiva respecto de la suma de escalares:* $(\alpha + \beta)\,\mathbf{a} = \alpha\mathbf{a} + \beta\mathbf{a}$
10. *distributiva respecto de la suma de vectores:* $\alpha\,(\mathbf{a} + \mathbf{b}) = \alpha\mathbf{a} + \alpha\mathbf{b}$

Si los escalares son números reales, entonces V es llamado un espacio vectorial real.

Ejercicio 8 *Sea* $V = \left\{(x, y, z) \in \mathbb{R}^3 \,/\, x = 0\right\}$*. Comprobar que tiene estructura de espacio vectorial con la suma de vectores y el producto por un escalar de* $\mathbb{R}^3$*.*

Ejercicio 9 *Demostrar que el conjunto de matrices* $p \times n$ *es un espacio vectorial con la suma de matrices y el producto por un escalar. Se denota por* $\mathcal{M}_{p\times n}$*.*

Trabajamos en primer lugar con el vector cero $\mathbf{0}$. ¿Tiene las propiedades que uno asocia normalmente con la palabra 'cero'? Si multiplicamos el vector cero por un escalar no nulo, ¿se obtendrá de nuevo el vector cero? Si multiplicamos cualquier vector por el número 0, ¿el resultado es el vector cero? La respuesta en ambos casos es afirmativa, pero hay que demostrarlo a partir de la definición de espacio vectorial.

- Para cualquier vector $\mathbf{a} \in V$, $0 \cdot \mathbf{a} = \mathbf{0}$.
- Para cualquier escalar α, $\alpha \cdot \mathbf{0} = \mathbf{0}$.
- El opuesto de cualquier vector $\mathbf{a} \in V$ es único.
- El vector cero en V es único.
- Para cualquier vector $\mathbf{a} \in V$, $-1 \cdot \mathbf{a} = -\mathbf{a}$.
- Para cualquier vector $\mathbf{a} \in V$, $-(-\mathbf{a}) = \mathbf{a}$.
- Si $\alpha \cdot \mathbf{a} = \mathbf{0}$, entonces o bien $\alpha = 0$ o bien $\mathbf{a} = \mathbf{0}$.

2.6. Subespacios vectoriales.

Para mostrar que un conjunto de objetos es un espacio vectorial, hemos de ver que se cumplen las diez propiedades del espacio vectorial. Este proceso se puede hacer más corto si el conjunto de objetos *es un subconjunto de un espacio vectorial conocido* V. Entonces, en vez de las 10 propiedades sólo hemos de verificar dos: las de clausura, porque las otras ocho se siguen inmediatamente de estas dos.

Definimos un conjunto no vacío S de un espacio vectorial V como un *subespacio* de V si S es un espacio vectorial bajo las *mismas* operaciones de suma vectorial y producto por escalar definidas en V.

Teorema 2.6.1 *Sea V un espacio vectorial y $S \subset V$; S es un subespacio vectorial si y sólo si se verifican las dos propiedades siguientes (de clausura):*

(i) Clausura bajo suma: si $\mathbf{a} \in S$ y $\mathbf{b} \in S$, entonces $\mathbf{a} + \mathbf{b} \in S$.

(ii) Clausura bajo producto por escalar: si $\mathbf{a} \in S$ y α es cualquier escalar, entonces $\alpha \cdot \mathbf{a} \in S$.

Ejemplo 2 *Demostrar que el conjunto de las matrices matrices (2×2) con ceros en la diagonal principal es un subespacio de $\mathcal{M}_{2\times 2}$.*

Ejemplo 3 *El conjunto $S = \{(x, y, z) \in \mathbb{R}^3 \ / \ y = 0\}$ es un subespacio vectorial de $\mathbb{R}^3$.*

Ejemplo 4 *El conjunto $T = \{(x, y, z) \in \mathbb{R}^3 \ / \ x = 3\}$ NO es un subespacio vectorial de $\mathbb{R}^3$. (Fijémonos que el conjunto T no contiene al vector cero, por lo que no puede ser espacio vectorial.)*

En general, si un subconjunto de un espacio vectorial no incluye al vector $\mathbf{0}$ no puede ser subespacio.

Un subespacio sencillo asociado con cualquier espacio vectorial es el siguiente:

Teorema 2.6.2 *El subconjunto que contiene únicamente al vector cero es un subespacio de todo espacio vectorial V.*

Las dos condiciones especificadas en el teorema 2.6.1 se pueden agrupar en una sola.

Teorema 2.6.3 *$S \subset V$ es un subespacio vectorial de V si y sólo si para cualesquiera dos elementos $\mathbf{a}, \mathbf{b} \in S$ y para cualesquiera dos escalares α, β*

$$\alpha \mathbf{a} + \beta \mathbf{b}$$

también está en S.

La expresión $\alpha \mathbf{a} + \beta \mathbf{b}$ del teorema anterior es un caso especial de una combinación lineal. Decimos que un vector $\mathbf{v}$ es una **combinación lineal** de los vectores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n\}$ si existen escalares $\alpha_1, \alpha_2, \dots, \alpha_n$ tales que

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \cdots + \alpha_n \mathbf{v}_n$$

con algún $\alpha_i \neq 0$.

Ejemplo 5 *Determinar si $\mathbf{u} = (1, 2, 3)$ es una combinación lineal de $\mathbf{v}_1 = (1, 1, 1)$, $\mathbf{v}_2 = (2, 4, 0)$, $\mathbf{v}_3 = (0, 0, 1)$.*

2.6.1. Variedad lineal.

Dado un conjunto de vectores $S = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$, al conjunto de *todas* las combinaciones lineales de estos vectores se le llama **variedad lineal** engendrada por S. Se denota por $L(S)$ o también por $\langle \mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \rangle$.

La variedad lineal engendrada por un conjunto finito de vectores tiene estructura de espacio vectorial. Por tanto, si se tiene un espacio vectorial y se toma un conjunto finito de vectores de este espacio, la variedad lineal engendrada es un subespacio vectorial del espacio inicial. Así, podemos crear subespacios simplemente formando combinaciones lineales con unos cuantos vectores.

El conjunto de vectores que engendra una variedad lineal se conoce como **sistema generador**.

Ejemplo 6 *Considérese el espacio* $\mathbb{R}^3$. *Un sistema generador de este espacio viene dado por el conjunto* $\{(1,0,0), (0,1,0), (0,0,1)\}$ *porque cualquier vector* (x, y, z) *lo podemos escribir como una combinación lineal de ellos:*

$$\mathbf{a} = (x, y, z) = x(1,0,0) + y(0,1,0) + z(0,0,1)$$

Si se toman dos vectores cualesquiera, por ejemplo $(1,2,3)$ *y* $(0,1,-1)$*, el conjunto de todas sus combinaciones lineales*

$$\alpha(1,2,3) + \beta(0,1,-1)$$

forma una variedad lineal, que es un subespacio vectorial de $\mathbb{R}^3$. *Esta variedad lineal determina un plano cuyo sistema generador son los vectores dados, y se denota por*

$$\langle (1,2,3), (0,1,-1) \rangle$$

Si se considera un único vector $\langle (1,2,-1) \rangle$*, la variedad lineal que engendra es una recta que pasa por el origen.*

2.6.2. Intersección de subespacios.

Dado un espacio vectorial V y dos subespacios $U, W \subset V$, el conjunto de los vectores que pertenecen tanto a U y a W se conoce como el subespacio intersección $U \cap W$. Se puede demostrar que el conjunto $U \cap W$ es un subespacio de V.

$$U \cap W = \{\mathbf{v} \in V \,/\, \mathbf{v} \in U \quad \text{y} \quad \mathbf{v} \in W\}.$$

Se dice que dos espacios son **disjuntos** si su intersección es únicamente el vector cero $U \cap W = \{\mathbf{0}\}$.

Ejemplo 7 *Calcular la intersección de los subespacios* $U = \langle (1,2,0), (0,-1,1) \rangle$, $W = \langle (2,1,-1) \rangle$.

2.6.3. Suma y suma directa de subespacios.

Dado un espacio vectorial V y dos subespacios $U, W \subset V$, se define el **subespacio suma** $U + W$ como

$$U + W = \{\mathbf{v} \in V \,/\, \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2, \quad \mathbf{v}_1 \in U \ \text{ y } \ \mathbf{v}_2 \in W\},$$

esto es, como el conjunto de todos los vectores de V que pueden expresarse como suma de un vector de U y un vector de W. Se verfica que $U + W$ es un subespacio vectorial de V.

Normalmente, si $\mathbf{v} \in U + W$, ocurre que $\mathbf{v}$ se puede expresar como suma de varios vectores diferentes, es decir, podemos escribir

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{w}_1 \qquad \text{y} \qquad \mathbf{v} = \mathbf{u}_2 + \mathbf{w}_2$$

con $\mathbf{u}_1, \mathbf{u}_2 \in U$, $\mathbf{w}_1, \mathbf{w}_2 \in W$ y $\mathbf{u}_1 \neq \mathbf{u}_2$, $\mathbf{w}_1 \neq \mathbf{w}_2$.

Diremos que el subespacio $U + W$ es *suma directa* de U y W, y escribiremos $U \oplus W$ si todo vector de $U + W$ *se expresa de forma única* como suma de un vector de U y un vector de W. Los subespacios se llaman entonces *independientes.*

Se puede demostrar que dos subespacios U, W son independientes, esto es, se tiene $U \oplus W$ si y sólo si $U \cap W = \{\mathbf{0}\}$.

Ejemplo 8 *Calcular el subespacio suma de*

$$\begin{aligned} U &= \left\{(x, y, z, t) \in \mathbb{R}^4 \,/\, y = t = 0\right\} \\ W &= \left\{(x, y, z, t) \in \mathbb{R}^4 \,/\, y = t, z = 0\right\}. \end{aligned}$$

Comprobar si es suma directa.

Ejemplo 9

$$\begin{aligned} U &= \left\{(x, y, z) \in \mathbb{R}^3 \,/\, y = z\right\} \\ W &= \left\{(x, y, z) \in \mathbb{R}^3 \,/\, x = y = z\right\}. \end{aligned}$$

Comprobar si es suma directa.

Subespacio suplementario

Dado un espacio vectorial V y dos subespacios U y W, se dice que son suplementarios si $V = U \oplus W$, es decir, los dos subespacios U y W han de ser disjuntos e independientes.

2.6.4. Dependencia e independencia lineal.

La mayoría de los espacios vectoriales contienen infinitos vectores. En particular, si $\mathbf{u} \in V$, entonces $\alpha \cdot \mathbf{u} \in V$ para *cualquier* número real α. Por consiguiente, es útil determinar si un espacio vectorial puede ser caracterizado completamente sólo con unos cuantos representantes. Si es así, podemos describir el espacio vectorial mediante esos representantes, y en lugar de escribir todos los vectores del espacio vectorial (que a menudo son infinitos) sólo escribimos los representantes. Entonces usamos esos representantes para estudiar el espacio vectorial entero.

Uno de los principales objetivos del álgebra lineal es la caracterización *eficiente* de un espacio vectorial por medio de sus representantes, donde por "eficiente" entendemos escribir tan pocos representantes como podamos. Nos dedicamos ahora a determinar las propiedades que deben poseer esos conjuntos de representantes.

Un conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \ldots, \mathbf{v}_n\}$ de un espacio vectorial V se dice que es *linealmente dependiente* si existen escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ no todos nulos tales que

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n = \mathbf{0}. \tag{2.1}$$

Los vectores son *linealmente independientes* si el único conjunto de escalares que satisface la ecuación (2.1) es el conjunto

$$\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0.$$

Para ver si un conjunto de vectores dado es linealmente independiente, escribimos primero la ecuación (2.1) y preguntamos: ¿para qué valores de los α_i se satisface la ecuación? Claramente $\alpha_1 = \alpha_2 = \cdots \alpha_n = 0$ es un conjunto apropiado. Si éste es el *único* conjunto de valores que satisface (2.1) entonces los vectores son linealmente independientes. Si existe un conjunto de valores *no todos nulos*, entonces los vectores son linealmente dependientes.

Ejemplo 10 *¿El conjunto $\{(1,2),(3,4)\} \subset \mathbb{R}^2$ es linealmente independiente?*

Solución: Sí.

Las ecuaciones que definen las combinaciones lineales y la dependencia lineal son similares, por lo que no es extraño que los dos conceptos estén relacionados:

Teorema 2.6.4 *Un conjunto finito de vectores es linealmente dependiente si y sólo si uno de los vectores es una combinación lineal de los demás.*

No es preciso que cualquier vector en un conjunto dado sea combinación lineal de los demás si ese conjunto es linealmente dependiente, sino sólo que *al menos uno* tenga esa propiedad.

Teorema 2.6.5 *Si un conjunto de vectores S en un espacio vectorial V es linealmente independiente, entonces cualquier subconjunto de S también es linealmente independiente.*

Teorema 2.6.6 *Si un conjunto de vectores $S \subset V$ es linealmente dependiente, entonces cualquier conjunto $W \subset V$ que contenga a S, $S \subset W$ también es linealmente dependiente.*

2.7. Base y dimensión de un espacio vectorial.

Queremos caracterizar un espacio vectorial sólo con unos cuantos de sus representantes. Una propiedad que sería interesante tener es la posibilidad de expresar *cada vector* del espacio vectorial a partir de sus representantes, esto es, de combinar representantes para generar todos los otros vectores. Los únicos medios que tenemos para combinar vectores es la suma de vectores y el producto por un escalar, de modo que las únicas combinaciones posibles son las combinaciones lineales.

Definimos un conjunto de vectores $S \subset V$ como un sistema generador de V si cualquier vector en V se puede escribir como combinación lineal de los vectores de S, es decir, $V = L(S)$.

Si S es un sistema generador de V, S representa V completamente debido a que todo vector de V se puede obtener a partir de los vectores de S. Si requerimos también que S sea un conjunto linealmente independiente, tenemos garantizado que ningún vector de S se puede escribir como combinación lineal de los otros vectores de S. Así, la independencia lineal asegura que el conjunto S no contiene vectores superfluos.

Definición 3 *Una* base *de un espacio vectorial V es un conjunto de vectores que es linealmente independiente y a la vez genera V.*

Así pues, una base genera un espacio vectorial y tiene el menor número de vectores necesario para generarlo, con lo que satisface todos nuestros requerimientos para representar de manera eficiente un espacio vectorial dado.

Ejemplo 11 *Base de* $\mathbb{R}^2 = \{(1,0)\,,(0,1)\}$.
Base de $\mathbb{R}^3 = \{(1,0,0)\,,(0,1,0)\,,(0,0,1)\}$.
Base de $\mathbb{R}^4 = \{(1,0,0,0)\,,(0,1,0,0)\,,(0,0,1,0)\,,(0,0,0,1)\}$, *etc.*

En general, el conjunto de n-uplas de la forma

$$\{\mathbf{e}_1 = (1,0,0,\ldots,0)\,, \quad \mathbf{e}_2 = (0,1,0,\ldots,0)\,, \quad \ldots, \quad \mathbf{e}_n = (0,0,0,\ldots,0,1)\}$$

es una base de $\mathbb{R}^n$, llamada *base canónica o standard.*

Ejemplo 12 *Encontrar una base del espacio vectorial de las matrices* 2×2.

Ejemplo 13 *Encontrar una base de los polinomios de segundo grado.*

Ejemplo 14 *Demostrar que el conjunto formado por los vectores*
$\left\{ \begin{pmatrix} 3 & 6 \\ 3 & -6 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -8 \\ -12 & -4 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix} \right\}$ *es una base de* $\mathcal{M}_{2\times 2}$.

Un espacio vectorial V es de dimensión finita si tiene una base conteniendo un número finito de vectores. Un espacio vectorial de dimensión finita puede tener diferentes bases. El hecho de que todas las bases hayan de tener el mismo número de vectores es consecuencia del resultado siguiente:

Teorema 2.7.1 *Si $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de un espacio vectorial V, entonces cualquier conjunto conteniendo más de n vectores ha de ser linealmente dependiente. Por consiguiente, cualquier conjunto linealmente independiente de vectores en V ha de contener a lo sumo n vectores.*

Éste es uno de los principios fundamentales del álgebra lineal:

Teorema 2.7.2 *Cualquier base de un espacio vectorial de dimensión finita ha de contener el mismo número de vectores.*

Como el número de vectores en una base para un espacio vectorial de dimensión finita V es siempre el mismo, le podemos dar un nombre: *dimensión de* V.

Definimos la dimensión del espacio vectorial que sólo contiene al vector cero como 0.

Ejemplo 15 *La dimensión de* $\mathbb{R}^n$ *es* n.

Como consecuencia inmediata del teorema 2.7.2 se obtiene uno de los más importantes resultados del álgebra lineal:

Teorema 2.7.3 *En un espacio vectorial* n*-dimensional, cualquier conjunto de* $n+1$ *o más vectores es linealmente dependiente.*

Ejemplo 16 *El conjunto* $\{(1,5),(2,-4),(-3,-4)\}$ *es un conjunto de 3 vectores en* $\mathbb{R}^2$*; por consiguiente, es linealmente dependiente.*

2.8. Componentes de un vector en una base. Cambio de base.

Si $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de un espacio vectorial V, entonces B genera V, por lo que cualquier vector de V será combinación lineal de los vectores de la base, es decir,

$$\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n \qquad \forall \mathbf{v} \in V.$$

Se puede demostrar que los escalares $\alpha_1, \alpha_2, \ldots, \alpha_n$ son únicos para cada vector $\mathbf{v}$ y se llaman coordenadas de $\mathbf{v}$ en la base B. Así pues, se tiene

$$\mathbf{v} \longleftrightarrow \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}_B \longleftrightarrow (\alpha_1, \alpha_2, \ldots, \alpha_n)_B$$

y el vector $\mathbf{v}$ se representa por la n-upla $(\alpha_1, \alpha_2, \ldots, \alpha_n)_B$. (A veces, cuando está claro en qué base se trabaja, se suele prescindir del subíndice B).

Ejemplo 17 *Obtener las coordenadas del vector* $\mathbf{v} = (7,2) \in \mathbb{R}^2$ *respecto a*

(a) la base canónica $C = \{(1,0),(0,1)\}$

(b) la base $D = \{(1,1),(1,-1)\}$

Solución. (a) $(7,2) = 7(1,0) + 2(0,1)$, por lo que $(7,2) = (7,2)_C$.

(b) $(7,2) = \frac{9}{2}(1,1) + \frac{5}{2}(1,-1)$, de manera que $(7,2) = \left(\frac{9}{2}, \frac{5}{2}\right)_D$. $\square$

Observamos que las coordenadas del vector $\mathbf{v}$ con respecto a la base canónica es el vector mismo. Éste es siempre el caso para vectores en $\mathbb{R}^n$ con respecto a la base canónica.

Si se toma otra base diferente B' de V, las coordenadas de los vectores cambian, ya que son los coeficientes de la combinación lineal de los vectores de la nueva base.

Ejemplo 18 *Dado el vector* $\mathbf{v} = (7,3) \in \mathbb{R}^2$*, encontrar sus coordenadas en la base canónica y en las bases* $B = \{(1,1), (1,-1)\}$ *y* $B' = \{(2,1), (-1,-1)\}$.

Solución. En la base canónica las coordenadas del vector coinciden con sus componentes.

$$(7,3) = \alpha(1,1) + \beta(1,-1) \Rightarrow \left\{ \begin{array}{c} 7 = \alpha + \beta \\ 3 = \alpha - \beta \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \alpha = 5 \\ \beta = 2 \end{array} \right. \Rightarrow (7,3) = (5,2)_B$$

$$(7,3) = \alpha(2,1) + \beta(-1,-1) \Rightarrow \left\{ \begin{array}{c} 7 = 2\alpha - \beta \\ 3 = \alpha - \beta \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \alpha = 4 \\ \beta = 1 \end{array} \right. \Rightarrow (7,3) = (4,1)_{B'}.$$

$\square$

Ejemplo 19 *Encontrar una base para el subespacio de* $\mathbb{R}^4$ *generado por los vectores*

$$\{(1,1,-4,3), (2,0,2,2), (2,-1,3,2)\}$$

Solución. Se pueden escribir estos vectores como las filas de una matriz y estudiar cuáles son linealmente independientes aplicando operaciones elementales sobre las filas:

$$\begin{pmatrix} 1 & 1 & -4 & -3 \\ 2 & 0 & 2 & -2 \\ 2 & -1 & 3 & 2 \end{pmatrix} \underset{\substack{F_2 \to F_2 - 2F_1 \\ F_3 \to F_3 - 2F_1}}{\sim} \begin{pmatrix} 1 & 1 & -4 & -3 \\ 0 & -2 & 10 & 4 \\ 0 & -3 & 15 & 8 \end{pmatrix} \underset{F_2 \to \frac{-1}{2} F_2}{\sim} \begin{pmatrix} 1 & 1 & -4 & -3 \\ 0 & 1 & -5 & -2 \\ 0 & -3 & 15 & 8 \end{pmatrix}$$

$$\underset{F_3 \to F_3 + 3F_1}{\sim} \begin{pmatrix} 1 & 1 & -4 & -3 \\ 0 & 1 & -5 & -2 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Por tanto, $\{(1,1,-4,3), (2,0,2,2), (2,-1,3,2)\}$ es una base y la dimensión del subespacio es 3. $\square$

Ejercicio 10 *Encontrar una base para el subesapcio de* $\mathbb{R}^4$ *generado por los vectores*

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

Ejercicio 11 *Encontrar las coordenadas del vector* $\mathbf{v} = (2,-1,3) \in \mathbb{R}^3$ *en las bases*

- $B = \{(1,0,0), (2,2,0), (3,3,3)\}$
- $B' = \{(1,0,0), (0,-1,1), (0,0,2)\}$.

2.9. Aplicaciones económicas

En modelos estáticos de equilibrio competitivo, el comportamiento de los agentes económicos individuales puede ser descrito por un elemento del espacio vectorial $\mathbb{R}^n$.

Para verlo, tengamos primeramente en cuenta que cualquier economía individual que intercambia con otras n productos diferentes puede ser descrita por n cantidades reales

$$x_1, x_2, \ldots, x_n$$

donde cada x_i representa la cantidad intercambiada del bien i, con signo positivo si es el resultado de su actividad productiva, o con signo $-$ si esta cantidad corresponde a bienes adquiridos por este agente como factor utilizado en la producción o para su consumo final.

Si estuviéramos estudiando el comportamiento de un consumidor, estas cantidades se podrían interpretar como aquellas cantidades de bienes consumidos por el mismo (si tienen signo menos) o como las cantidades de diferentes trabajos que oferta (si tienen signo más).

Por consiguiente, en estos casos *el conjunto* $\mathbb{R}^n$ *representará el conjunto de distribuciones posibles de intercambios para esta economía.* Además, podemos definir dos operaciones importantes sobre el conjunto de distribuciones de bienes:

(a) <u>Suma</u>: para todo $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $(y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, definimos su suma mediante la igualdad

$$(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n) \in \mathbb{R}^n$$

Nótese que, si todas las cantidades involucradas son positivas, esta operación describe la unión de ambas distribuciones de intercambios (bienes). Si alguna cantidad de las que intervienen es negativa, entonces esta operación nos indicará el *resultado neto* de dos procesos de intercambio, obtenidos posiblemente mediante el uso de algunos de los productos de un proceso como factores de producción en el otro o viceversa.

(b) <u>Producto por un número real</u>: para todo $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, definimos esta operación por

$$\lambda \cdot (x_1, x_2, \ldots, x_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n) \in \mathbb{R}^n.$$

Nótese que esta operación multiplica todas las cantidades de una distribución de bienes intercambiados dada por un factor de proporcionalidad común λ.

Como ya hemos visto anteriormente, el conjunto de intercambios de esta economía, $\mathbb{R}^n$, con estas dos operaciones definidas, verifica una serie de propiedades que le confieren la estructura de espacio vectorial real.

Ejemplo. Supongamos una situación simplista en la que un individuo fabrica mantequilla, para lo cual necesita leche, consumiendo a su vez ropa y otros comestibles agregados en un solo bien. Esta economía individual podría ser descrita por el espacio vectorial $\mathbb{R}^4$, esto es, cada combinación de intercambio para estos bienes podría venir determinada por

$$(x_1, x_2, x_3, x_4) \in \mathbb{R}^4,$$

representando x_1 mantequilla, x_2 leche, x_3 ropa y x_4 otros comestibles, medidos en unidades apropiadas. Posibles distribuciones de estos bienes podrían ser $(10, -20, 5, 50)$ o $(20, 10, 0, 15) \in \mathbb{R}^4$. En este caso, el resultado neto de ambos intercambios vendría dado por la operación suma:

$$(10, -20, 5, 50) + (20, 10, 0, 15) = (30, -10, 5, 65).$$

Por otro lado, si la intención de este individuo es aumentar en un 20 % las cantidades intercambiadas en el primer proceso, los nuevos intercambios pueden venir dados por la operación multiplicación por un número real:

$$1{,}2 \cdot (10, -20, 5, 50) = (12, -24, 6, 60) \in \mathbb{R}^4$$

2.10. Problemas

1. Determina si los conjuntos siguientes son espacios vectoriales bajo las operaciones que se indican:

   a) El conjunto de las matrices cuadradas 2×2 con la suma de matrices y la multiplicación por un escalar.

   b) El conjunto de las matrices reales triangulares inferiores 3×3 de la forma

   $$\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{pmatrix}$$

   con la suma matricial y la multiplicación por un escalar.

   c) $\{(a, b) \in \mathbf{R}^2 / a + b = 2\}$ con la suma y la multiplicación por un escalar estándar.

2. Muestra que el conjunto de soluciones de la ecuación matricial $AX = \mathrm{O}$, donde A es una matriz $p \times n$, es un subespacio de $\mathbb{R}^n$. Mostrar también que el conjunto de soluciones de la ecuación $AX = B$ no es un subespacio de $\mathbb{R}^n$ cuando $B \neq \mathrm{O}$.

3. Consideremos los vectores de $\mathbb{R}^3$ $\vec{u} = (1, -3, 2)$ y $\vec{v} = (2, -1, 1)$. a) Escribe $(1, 7, -4)$ como combinación lineal de $\vec{u}$ y $\vec{v}$. b) Escribe $(2, -5, 4)$ como combinación lineal de $\vec{u}$ y $\vec{v}$. c) ¿Para qué valores de k es $(1, k, 5)$ una combinación lineal de $\vec{u}$ y $\vec{v}$?

4. Estudia si el conjunto de vectores $S = \{(2, 1, 0), (3, 1, 1), (3, 2, -1)\}$ forma un sistema generador de $\mathbb{R}^3$. Si no es así, determina el subespacio que genera.

5. Enccuentra un sistema generador del subespacio de $\mathbb{R}^4$ definido por $\{(x, y, z, t) \in \mathbf{R}^4 : x + 2z - 3t + 2y = 0\}$.

6. Determina si los siguientes conjuntos son bases de $\mathbb{R}^3$:

   a) $\{(2,1,1),\ (2,2,1),\ (2,2,-1)\}$
   b) $\{(1,2,1),\ (1,3,1),\ (1,4,1),\ (1,5,1)\}$

7. Encuentra una base para $L(S)$, donde S viene dado por:

   a) $S = \{(0,1,1,1),\ (1,0,0,1),\ (-1,1,1,0),\ (1,1,0,1)\}$
   b) $S = \{(1,0,-1,1),\ (3,1,0,1),\ (1,1,2,-1),\ (3,2,3,-1),\ (2,1,0,0)\}$

8. Halla una base del subespacio $S = \{(x, y, z) \in \mathbb{R}^3\ :\ x + y = z\}$.

9. Dado el conjunto de vectores $\{(1,1,1), (-2,1,0), (-1,0,1)\}$, ¿son linealmente independientes? ¿Son base de $\mathbb{R}^3$?

10. Dados los subespacios $S = \langle(1,0,1,0), (1,1,1,0), (0,1,0,0)\rangle$ y $T = \langle(1,0,0,0), (1,0,1,0), (0,0,1,0)\rangle$, caracteriza los subespacios $S+T$ y $S\cap T$.

11. Sean $S = \langle(2,0,2), (0,2,0), (-1,0,-1)\rangle$ y $T = \left\{(x, y, z) \in \mathbb{R}^3 /\quad z - y = 0\right\}$ variedades lineales en $\mathbb{R}^3$. Se pide:

    a) Encuentra las ecuaciones de la variedad S, una base y su dimensión. Encuentra también una base y la dimensión de T.
    b) ¿Pertenece el vector $\vec{v} = (4, -2, 4,)$ al subespacio S? Si es así, da sus coordenadas en función de la base encontrada anteriormente.
    c) Calcula $S \cap T$ y $S + T$ dando una base de cada uno de ellos.

12. Calcula:

    a) $L_1 \cap L_2$, siendo $L_1 = \{(x, y, z, t) \in \mathbb{R}^4\ :\ x = y = z = t\}$ y $L_2 = \{(x, y, z, t) \in \mathbb{R}^4\ :\ x = z,\ y = t\}$.
    b) $L_2 \cap L_3$, siendo $L_3 = \{(x, y, z, t) \in \mathbb{R}^4\ :\ y = 2x,\ t = 2z\}$.
    c) Complementario de L_3.
    d) Complementario de L_4, siendo $L_4 = \{(x, y, z, t) \in \mathbb{R}^4\ :\ 2y + 3z = 0\}$.
    e) Indica si la suma es directa en cada uno de los casos siguientes: $L_1 + L_2$, $L_1 + L_3$, $L_2 + L_3$, $L_2 + L_4$.

13. Dados los subespacios vectoriales

$$L_1 = \left\{\vec{x} \in \mathbb{R}^4 /\ x_1 + x_3 = x_2 = -x_4\right\}, \qquad L_2 = \langle(1,0,0,-1)\rangle$$

    determina:

    a) $L_1 \cap L_2$ ¿Son suplementarios?

  b) $L_1 + L_2$
  c) $\dim L_1$, $\dim L_2$
  d) Comprueba que L_1 es un subespacio vectorial de $\mathbb{R}^4$

14. Dados los subespacios vectoriales del espacio vectorial de las matrices cuadradas 2×2 con coeficientes reales

$$H_1 = \left\{ \left(\begin{array}{cc} a & b \\ -b & a \end{array} \right) \Big/ a, b \in \mathbb{R} \right\}, \qquad H_2 = \left\{ \left(\begin{array}{cc} c & d \\ e & -c \end{array} \right) \Big/ c, d, e \in \mathbb{R} \right\},$$

calcula la dimensión y una base de H_1, H_2, $H_1 + H_2, H_1 \cap H_2$.

15. Sea $S = \langle (1, 1, -1), (2, 0, 1), (1, -1, 2) \rangle$ una variedad lineal en $\mathbb{R}^3$. Se pide:

  a) Calcula una base y la dimensión de S.
  b) Encuentra las coordenadas del vector $\vec{v} = (5, 3, -2)$ respecto de dicha base.
  c) Sea $H = \{(\alpha, \beta, \alpha + \beta),\ \alpha, \beta \in \mathbb{R}\}$; encuentra una base y la dimensión de H.
  d) Calcula $S \cap H$, dando una base y su dimensión. Calcula $S + H$. ¿Es suma directa? ¿Por qué?

16. Dados los espacios vectoriales $F = \left\{(x, y, z) \in \mathbb{R}^3 / x = y = z\right\}$ y $G = \langle (1, 0, -1), (0, 1, -1), (1, -1, 0) \rangle :$

  a) Prueba que son subespacios de $\mathbb{R}^3$.
  b) Prueba que $F \oplus G = \mathbb{R}^3$.

17. Calcula las coordenadas de los vectores de $\mathbb{R}^3$ $(1, 2, 4)$ y $(2, 3, -1)$ respecto de la base $\{(1, 0, 0), (0, 2, 1), (0, 0, -1)\}$

  a) ¿Son linealmente independientes los vectores $(1, 1, 4)$, $(0, 2, 1)$ y $(3, 1, 9)$?
  b) Dado el vector con coordenadas $(3, 2, 1)$ en la base $\{(1, 0, 2), (0, 2, 1), (1, 2, 0)\}$, determina sus coordenadas en la base $\{(3, 0, 4), (0, 3, 4), (3, 4, 0)\}$.

18. Sea $S = \langle (1, 1, 1, 1), (1, 0, 1, 0), (1, -1, 1, -1), (0, 1, 0, 1) \rangle$ una variedad lineal en $\mathbb{R}^4$. Se pide:

  a) Encuentra las ecuaciones de dicha variedad, una base y su dimensión.
  b) ¿Pertenece el vector $\vec{v} = (4, -2, 4, -2)$ a dicho subespacio? Si es así, da sus coordenadas en función de los vectores dados.

19. Sean $S = \left\{(x, y, z) \in \mathbb{R}^3 : x = z\right\}$ y $T = \langle (1, 2, 3), (3, 2, 1), (1, 0, -1) \rangle$. Se pide:

  a) Demuestra que S es subespacio vectorial de $\mathbb{R}^3$.

b) Da una base y la dimensión de S y T.

c) ¿Pertenece el vector $(1,1,1)$ a S?

20. En $\mathbb{R}^3$ se consideran los subespacios vectoriales

$$\begin{aligned} H_1 &= \{(x,y,z) \in \mathbb{R}^3 \,:\, x+y=0,\, x+z=0\} \\ H_2 &= \langle (1,1,0),\, (0,1,1) \rangle \end{aligned}$$

a) Justifica que H_2 es, efectivamente, un subespacio vectorial de $\mathbb{R}^3$. Calcula el valor de a para que el vector $(2,a,2)$ pertenezca a H_2.

b) Calcula el subespacio vectorial $H_1 \cap H_2$. ¿Cuál es la dimensión del subespacio vectorial $H_1 + H_2$? ¿Es suma directa? ¿Se verifica que $H_1 \oplus H_2 = \mathbb{R}^3$?

2.11. Aplicaciones de los vectores a la Economía

1. Una compañía petrolera puede convertir un barril de crudo en tres clases distintas de combustible. Sin aditivos de plomo sus producciones de las tres clases de combustible a partir de un barril de crudo vienen dadas por el vector $(2,2,4)$. Con el máximo de aditivos de plomo permitidos legalmente las producciones son $(5,0,3)$. Supongamos que los efectos de los aditivos de plomo son proporcionales, es decir, que al usar una fracción θ del máximo permitido $(0 \leq \theta \leq 1)$ se tiene la producción $(1-\theta)(2,2,4) + \theta(5,0,3)$.

a) ¿Es posible que la compañía produzca los siguientes vectores: a) $\left(\frac{5}{2}, 1, \frac{7}{2}\right)$, b) $\left(\frac{9}{2}, \frac{1}{3}, \frac{19}{6}\right)$, c) $(1,6,9)$?

b) En caso afirmativo, ¿qué proporción de aditivos de plomo legalmente permitido hay que usar en cada caso?

2. Una compañía constructora almacena tres mezclas básicas A, B y C. Las cantidades se miden en gramos y cada "unidad" de la mezcla pesa 60 gr. Pueden hacerse mezclas especiales de argamasa efectuando combinaciones de las tres mezclas básicas. Por ello, las mezclas especiales posibles pertenecen al espacio generado por los tres vectores que representan las tres mezclas básicas. La composición de éstas es:

	A	B	C
Cemento	20	18	12
Agua	10	10	10
Arena	20	25	15
Grava	10	5	15
Tobas	0	2	8

a) ¿Es posible hacer una mezcla que contenga 1000 gr. de cemento, 200 de agua, 1000 de arena, 500 de grava y 300 de tobas? ¿Por qué? Si se puede, ¿cuántas unidades de cada mezcla básica se necesitan para esta fórmula?

b) Supóngase que se desean hacer 5400 gr. de argamasa de manera que contenga 1350 gr de cemento, 1675 de arena y 1025 de grava. Si la razón de agua con cemento es de 2 a 3, ¿qué cantidad de tobas debe utilizarse para obtener los 5400 gr de argamasa? ¿Se puede formular ésta como una mezcla especial? Si es así, ¿cuántas unidades de A, B y C se necesitan para formular esta mezcla especial?

3. El señor Oriol cuenta, entre las muchas empresas que integran su holding, con una que se dedica a la fabricación de televisores, Oriol TV S.A., en la cual se producen cuatro tipos de artículos:

   Gama alta: TV 14", TV 18"

   Gama baja: TV 14", TV 18".

   Con el fin de realizar una previsión de ventas para el año próximo, ha encargado a la consultoría PRIVASA un estudio de mercado. Los resultados del análisis fueron los siguientes:

   Gama alta: Ventas de TV 14" $(A_1) = 100N + 218E$

   Gama alta: Ventas de TV 18" $(A_2) = 90N + 153E$

   Gama baja: Ventas de TV 14" $(B_1) = 300E$

   Gama baja: Ventas de TV 18" $(B_2) = 500E$,

   siendo N el número de familias con renta anual superior a los 24000 euros y E el número de grandes almacenes en los que se comercializan los productos. Se pide:

   a) Construye el vector de nivel de ventas de los cuatro artículos y comprueba que el conjunto de esos vectores constituye un subespacio vectorial de $\mathbb{R}^4$.

   b) Calcula la base del subespacio de $\mathbb{R}^4$ generado por el vector formado por el nivel de ventas de los cuatro artículos.

   c) Interpreta la dimensión de dicho subespacio.

   d) ¿En qué condiciones las ventas de los cuatro productos serán nulas?

4. Supongamos un país en el que el gobierno fija políticamente los precios del pan, de la carne y de la gasolina, estando estos precios relacionados por:

$$P_c = 100P_p, \quad P_g = 10P_p$$

   siendo P_p, P_c y P_g respectivamente las variaciones de precios del pan, de la carne y de la gasolina con respecto al período anterior. Comprueba que el conjunto de vectores de variaciones de precios administrados es un subespacio vectorial de $\mathbb{R}^3$. ObtÈn una base y la dimensión. ¿Qué interpretación económica se le puede dar a la dimensión?

5. Sabemos que en un país existen multitud de tipos de interés y también que en los libros de economía se habla sólo del tipo de interés a secas. Consideremos tipos de interés reales (nominales menos la tasa de inflación) y la economÌa de

la eurozona; para no hacer interminable la lista, tomemos cuatro de los más conocidos: el tipo de interés del mercado interbancario a un mes, el rendimiento interno de obligaciones privadas, el de los pagarés del tesoro a un año y el de las obligaciones, y denotémoslos por IM, ROB, RP y ROE, respectivamente. Esto nos permite definir un vector de tipos de interés reales dado por

$$R = (IM, ROB, RP, ROE).$$

De una cuidadosa observación de los distintos tipos vemos que están relacionados por:

$$ROB = 1{,}6IM, \quad RP = 0{,}75ROB, \quad ROE = 1{,}4IM$$

Se pide:

a) ¿Podemos decir que los vectores de tipos de interés reales forman un subespacio vectorial de $\mathbb{R}^4$?
b) En caso afirmativo, calcúlese su dimensión y una base.
c) ¿Consideras que la abstracción de los libros de economía es excesiva o que por el contrario no se pierde ninguna información trabajando con un único tipo de interés? Si podemos trabajar con uno solo, ¿cambian algo los resultados al tomar diferentes tipos?

6. Sea el subespacio vectorial

$$V = \{(w, t_d, t_i, P, GP)/P = w, \ t_d + t_i = GP\}$$

donde w es el crecimiento de los salarios nominales, t_d es el crecimiento de los impuestos directos, t_i es el crecimiento de los impuestos indirectos, P es el crecimiento de los precios y GP es el crecimiento del gasto público. Se pide:

a) Demuestra que V es efectivamente un subespacio vectorial.
b) Calcula una base y la dimensión. Dale una interpretación económica.

7. En una economía hay cinco variables básicas: variación de precios (p), variación de la cantidad de dinero (m), variación de renta (y), variación de impuestos (t) y variación de gasto público (g), componentes del vector en que se fijan los distintos agentes económicos:

$$(p, m, y, t, g) \in \mathbb{R}^5$$

Los agentes de la economía se dividen en dos grandes grupos: el sector privado, cuyas creencias sobre la evolución de las cinco variables clave de la economía se resumen en

$$(0{,}2x + z, 0{,}2x + z, 0, 0{,}8x, x), \quad x, z \in \mathbb{R}$$

mientras el otro gran grupo, el sector público, cree que estas variables están ligadas por

$$(\beta, \gamma + \beta, \gamma, \alpha, \alpha), \quad \alpha, \beta, \gamma \in \mathbb{R},$$

significando las cinco coordenadas de cada subespacio las arriba indicadas. Se pide:

a) Demostrar que las creencias de cada grupo constituyen un subespacio vectorial de $\mathbb{R}^5$.

b) Si la evolución de la economía viene dada por la suma de los dos tipos de creencias, obtÈnla.

c) ¿Tienen en común ambos grupos alguna creencia?

8. Una empresa de fabricación de ordenadores personales ofrece en el mercado dos productos, el 'súper' y el 'extra'. Los costes totales en que incurre la empresa vienen dados por la expresión

$$CT = 2q_s + 2q_e,$$

siendo q_s y q_e las cantidades producidas de ordenadores 'súper' y 'extra', respectivamente. Por otra parte, en la fabricación se emplean dos clases de tecnología: nacional y extranjera. La empresa sabe que para producir un ordenador súper necesita emplear tres componentes nacionales y dos extranjeros, mientras que para producir un ordenador extra necesita tres nacionales y un extranjero. Se pide obtener los costes totales de la empresa en función de las dos clases de tecnología empleadas. Suponed que los componentes nacionales son un vector de la forma $(CN, 0)$, mientras que los extranjeros vienen dados por $(0, CE)$.

Capítulo 3

APLICACIONES LINEALES

3.1. Funciones.

Relaciones entre objetos suceden a menudo en las interacciones de la vida diaria, y si las matemáticas tienen como objeto modelar o explicar esas interacciones, entonces han de tratar las *relaciones*. En el comercio hay relaciones entre trabajo y producción, entre producción y beneficio, y entre beneficio e inversión. En Física hay relaciones entre fuerza y aceleración, y entre masa y energía. Necesitamos, por consiguiente, estructuras matemáticas para representar relaciones. Una de esas estructuras es la *función*.

Diremos que una *función* es una regla de correspondencia entre dos conjuntos, llamados habitualmente conjunto inicial o *dominio* y conjunto final, que asigna a cada elemento en el dominio exactamente un elemento (no necesariamente diferente) en el conjunto final.

La *imagen* de una función es el conjunto de todos los elementos del conjunto final que están relacionados con elementos del dominio mediante la regla de correspondencia.

Un elemento y del conjunto final está en la imagen sólo si existe un elemento x en el dominio tal que a x se le asigna el valor y mediante la regla de correspondencia.

El dominio y el conjunto final de una función pueden ser cualquier tipo de conjunto, mientras que la regla de correspondencia puede ser especificada mediante flechas, tablas, fórmulas o palabras. Si nos restringimos a conjuntos de números reales y reglas de correspondencia dadas por ecuaciones, tenemos las funciones comúnmente estudiadas en cálculo.

Siempre que tengamos dos conjuntos numéricos y una función f relacionando el elementos arbitrario x en el dominio con el elemento y en la imagen mediante una ecuación, decimos que y es una función de x y escribimos $y = f(x)$. En general, la ecuación $S = f(N)$ es una forma simbólica de escribir la siguiente frase: 'tenemos una función que consiste en dos conjuntos de números y una ecuación, donde N y S denotan elementos en el dominio y en el conjunto final, respectivamente'. Si el dominio no se especifica, se supone que son todos los números reales para los cuales la regla de correspondencia tiene sentido.

Cuando se tiene una regla de correspondencia definida mediante la fórmula $f(x)$, entonces se obtiene el elemento del conjunto final asociado a un valor particular de x simplemente sustituyendo x por el valor particular en la fórmula.

Ejemplo 20 $\left.\begin{array}{cccc} f: & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^2 \end{array}\right\} \quad f(x) = x^2$

Ejemplo 21 $\left.\begin{array}{cccc} f: & \mathbb{Q} & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \sqrt{x^2+1} \end{array}\right\} \quad f(x) = \sqrt{x^2+1}$

3.2. Aplicaciones lineales.

Dos sinónimos de la palabra función usados frecuentemente son *aplicación* y *transformación.* En cálculo elemental, el dominio y el conjunto final están restringidos a subconjuntos de números reales y se usa exclusivamente la palabra función. En álgebra lineal, el dominio y el conjunto final son espacios vectoriales y se prefiere usar las palabras transformación y aplicación.

Una *transformación* T es una regla de correspondencia entre dos espacios vectoriales V y W, que asigna a cada elemento en V exactamente un elemento (pero no necesariamente diferente) en W. Tal transformación se denota mediante $T : V \longrightarrow W$. Escribimos $\mathbf{w} = T(\mathbf{v})$ si el vector $\mathbf{w} \in W$ está relacionado con el vector $\mathbf{v} \in V$ mediante la regla de correspondencia asociada con T.

La imagen de T es el conjunto de todos los vectores de W que están relacionados con vectores de V bajo la regla de correspondencia. Así, $\mathbf{w}$ está en la imagen de T si y sólo si existe un vector $\mathbf{v} \in V$ tal que $\mathbf{w} = T(\mathbf{v})$.

Definición 4 *Una aplicación $T : V \longrightarrow W$ es lineal si para cualesquiera dos escalares α, β y cualesquiera dos vectores $\mathbf{u}, \mathbf{v} \in V$ se verifica la siguiente igualdad*

$$T(\alpha\,\mathbf{u} + \beta\,\mathbf{v}) = \alpha\,T(\mathbf{u}) + \beta\,T(\mathbf{v}) \tag{3.1}$$

Si $\alpha = \beta = 1$, entonces $T(\mathbf{u}+\mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v})$; si $\beta = 0$, entonces $T(\alpha\,\mathbf{u}) = \alpha\,T(\mathbf{u})$.

La parte izquierda de (3.1) es la transformación de la combinación lineal $\alpha\mathbf{u}+\beta\mathbf{v}$ del espacio vectorial V en el espacio vectorial W. Si T es lineal, entonces el resultado de transformar $\alpha\mathbf{u}+\beta\mathbf{v}$ en W es el mismo que transformar por separado $\mathbf{u}$ y $\mathbf{v}$ en W ($T(\mathbf{u})$, $T(\mathbf{v})$) y después formar *idéntica* combinación lineal con $T(\mathbf{u})$ y $T(\mathbf{v})$ en W a la que teníamos en V con $\mathbf{u}$ y $\mathbf{v}$. Las combinaciones lineales son fundamentales en los espacios vectoriales, ya que involucran las únicas operaciones (suma y producto por escalar) cuya existencia está garantizada en los expacios vectoriales. *De todas las posibles transformaciones, las aplicaciones lineales son aquéllas que preservan las combinaciones lineales.*

Ejemplo 22 *Determinar si la transformación $T : V \longrightarrow V$ definida por $T(\mathbf{v}) = k\mathbf{v}$ es lineal, siendo k un escalar.*

Ejemplo 23 *Sea $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $T(x, y) = (x, x+y, x-y)$. Demostrar que es una aplicación lineal.*

Solución:
$\mathbf{u} = (x_1, y_1) \Rightarrow T(\mathbf{u}) = (x_1, x_1 + y_1, x_1 - y_1)$
$\mathbf{v} = (x_2, y_2) \Rightarrow T(\mathbf{v}) = (x_2, x_2 + y_2, x_2 - y_2)$
$\mathbf{u}+\mathbf{v} = (x_1 + x_2, y_1 + y_2) \Rightarrow T(\mathbf{u}+\mathbf{v}) = (x_1 + x_2, x_1 + x_2 + y_1 + y_2, x_1 + x_2 - y_1 - y_2)$

- $T(\mathbf{u}) + T(\mathbf{v}) = (x_1, x_1 + y_1, x_1 - y_1) + (x_2, x_2 + y_2, x_2 - y_2) =$
  $= (x_1 + x_2, x_1 + x_2 + y_1 + y_2, x_1 + x_2 - y_1 - y_2) = T(\mathbf{u}+\mathbf{v})$
- $\alpha\mathbf{u} = (\alpha x_1, \alpha y_1) \Rightarrow T(\alpha\mathbf{u}) = (\alpha x_1, \alpha x_1 + \alpha y_1, \alpha x_1 - \alpha y_1)$
  $\alpha T(\mathbf{u}) = \alpha(x_1, x_1 + y_1, x_1 - y_1) = (\alpha x_1, \alpha(x_1 + y_1), \alpha(x_1 - y_1)) = T(\alpha\mathbf{u})$.
  $\square$

Ejercicio 12 *Dada* $\left.\begin{array}{rccc} T: & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x,y,z) & \longmapsto & (2x+y, 3y-4z) \end{array}\right\}$, *demostrar si es o no una aplicación lineal.*

Ejercicio 13 *Sea* $\left.\begin{array}{rccc} T: & \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ & (x,y) & \longmapsto & (x^2, y) \end{array}\right\}$; *¿es una aplicación lineal?*

Ejercicio 14 *Dada* $\left.\begin{array}{rccc} T: & \mathcal{M}_{2\times 2} & \longrightarrow & \mathbb{R} \\ & A & \longmapsto & \det A \end{array}\right\}$, *demostrar si es o no una aplicación lineal.*

Ejercicio 15 *Dada* $\left.\begin{array}{rccc} T: & \mathbb{M}_{2\times 2} & \longrightarrow & \mathbb{R} \\ & A & \longmapsto & \mathrm{tr}A \end{array}\right\}$, *demostrar si es o no una aplicación lineal, siendo* $\mathrm{tr}A = \sum a_{ii}$, *la suma de los elementos de la diagonal principal de la matriz* A.

Ejemplo 24 *La transformación* R *definida por*

$$R\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix}$$

donde a *y* b *denotan números reales arbitrarios y* θ *es constante, es una aplicación lineal.*

De hecho, la solución de este último ejemplo se generaliza fácilmente a cualquier aplicación definida por una multiplicación de matrices por n-tuplas. Consecuentemente, *cualquier matriz define una aplicación lineal.*

Teorema 3.2.1 *Si* $L: \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *está definida mediante* $L(\mathbf{u}) = A\mathbf{u}$ *para cualquier matriz* A *de tamaño* $m \times n$, *entonces* L *es una aplicación lineal.*

Ejemplo 25 *Sea* $T: \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ *una aplicación definida por* $T\binom{x}{y} = A\binom{x}{y}$, *siendo* $A = \begin{pmatrix} 2 & -1 \\ -8 & 4 \end{pmatrix}$. *Comprobar que es una aplicación lineal.*

Solución. Dados $\mathbf{v}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$, $\mathbf{v}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \in \mathbb{R}^2$,

$T(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = A(\alpha\mathbf{v}_1 + \beta\mathbf{v}_2) = A(\alpha\mathbf{v}_1) + A(\beta\mathbf{v}_2) = \alpha A(\mathbf{v}_1) + \beta A(\mathbf{v}_2) = \alpha T(\mathbf{v}_1) + \beta T(\mathbf{v}_2)$, luego T es una aplicación lineal. $\square$

3.3. Representaciones matriciales

En el tema 2 mostramos que cualquier vector en un espacio vectorial de dimensión finita puede ser representado como una n-tupla con respecto a una base dada. En consecuencia, podemos estudiar espacios vectoriales de dimensión finita simplemente analizando n-tuplas. Ahora mostraremos que cualquier transformación lineal de un espacio vectorial n-dimensional en un espacio vectorial m-dimensional se puede representar por una matriz $m \times n$. Así, podemos reducir el estudio de las aplicaciones lineales sobre espacios vectoriales de dimensión finita al estudio de matrices.

Más en detalle, vimos que sólo hay una forma de expresar $\mathbf{v}$ como combinación lineal de un conjunto de vectores de una base dada. Si $\mathbf{v}$ es cualquier vector en un espacio vectorial V de dimensión finita, y si $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ es una base de V, entonces existe un único conjunto de escalares $c_1, c_2, \dots, c_n$ tal que

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n \tag{3.2}$$

y escribimos

$$\mathbf{v} \longleftrightarrow (c_1, c_2, \dots, c_n)_B \longleftrightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \tag{3.3}$$

para indicar que la n-tupla es una representación de coordenadas del vector $\mathbf{v}$.

Si $T : V \longrightarrow W$ es una aplicación lineal y $\mathbf{v}$ es cualquier vector en V expresado en la forma (3.2), entonces

$$T(\mathbf{v}) = T(c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n) = c_1T(\mathbf{v}_1) + c_2T(\mathbf{v}_2) + \cdots + c_nT(\mathbf{v}_n).$$

Así pues, T *es descrita completamente mediante su acción sobre una base*. Una vez se conoce cómo T transforma los vectores de la base se puede determinar cómo afecta T a cualquier vector $\mathbf{v}$ en V.

Ejemplo 26 *Sea* $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ *una aplicación lineal tal que* $T(1,0) = (1,2,0)$, $T(0,1) = (0,3,4)$. *Entonces, para cualquier vector* $\mathbf{v} = (a,b) \in \mathbb{R}^2$ *se tiene, como* $\mathbf{v} = a(1,0) + b(0,1)$, *que*

$$T(\mathbf{v}) = aT(1,0) + bT(0,1) = a(1,2,0) + b(0,3,4) = (a, 2a+3b, 4b)$$

Con estos dos conceptos tenemos las herramientas necesarias para mostrar que cualquier aplicación lineal de un espacio vectorial de dimensión finita en otro se puede representar mediante una matriz.

En efecto, sea T una aplicación de un espacio vectorial V de dimensión n en un espacio vectorial W m-dimensional, y sea $B = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ una base de V y $C = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m\}$ una base de W. Entonces $T(\mathbf{v}_1), T(\mathbf{v}_2), \dots, T(\mathbf{v}_n)$ son todos vectores de W, que se pueden expresar como una combinación lineal de los vectores de la base C. En particular,

$$\begin{aligned} T(\mathbf{v}_1) &= a_{11}\mathbf{w}_1 + a_{21}\mathbf{w}_2 + \cdots + a_{m1}\mathbf{w}_m \\ T(\mathbf{v}_2) &= a_{12}\mathbf{w}_1 + a_{22}\mathbf{w}_2 + \cdots + a_{m2}\mathbf{w}_m \\ \vdots \quad & \quad \vdots \\ T(\mathbf{v}_n) &= a_{1n}\mathbf{w}_1 + a_{2n}\mathbf{w}_2 + \cdots + a_{mn}\mathbf{w}_m \end{aligned}$$

para algunos escalares $a_{i,j}$. La representación en coordenadas de estos vectores son

$$T(\mathbf{v}_1) \longleftrightarrow \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}_C, \qquad T(\mathbf{v}_2) \longleftrightarrow \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}_C, \ldots, \quad T(\mathbf{v}_n) \longleftrightarrow \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}_C$$

Si usamos estas n-tuplas como las columnas de una matriz A, entonces, como veremos, *A es la representación matricial de la aplicación lineal T*. Como esta matriz depende de la base (de hecho, depende de la base B de V y C de W elegidas), podemos escribir A_B^C para explicitar esa dependencia. Así, *A_B^C denota la representación matricial* (también llamada *matriz asociada*) *de la aplicación lineal T con respecto a la base B en el dominio V y la base C en el espacio W*. A menudo, el subíndice B o el superíndice C no se pone cuando corresponde a la base canónica en $\mathbb{R}^n$ y $\mathbb{R}^m$, respectivamente.

Ejemplo 27 *Obtener la representación matricial con respecto a la base canónica en $\mathbb{R}^2$ y la base canónica en $\mathbb{R}^3$ para la aplicación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $T(a, b) = (a + b, a - b, 2b)$.*

Solución. Tenemos que

$$T(1,0) = (1,1,0) \longleftrightarrow \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad T(0,1) = (1,-1,2) \longleftrightarrow \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix},$$

de manera que

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

□

Ejemplo 28 *Rehacer el ejemplo precedente con la base del dominio cambiada a $B = \{(1,1), (1,-1)\}$.*

Solución. Ahora

$$T(1,1) = (2,0,2) \longleftrightarrow \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \qquad T(1,-1) = (0,2,-2) \longleftrightarrow \begin{pmatrix} 0 \\ 2 \\ -2 \end{pmatrix},$$

de manera que

$$A_B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ 2 & -2 \end{pmatrix}_B$$

$\square$

Para probar que A_B^C es una representación matricial de una aplicación lineal T, empezamos con un vector $\mathbf{v}$ en el dominio V. Si $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V, entonces

$$\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n = \sum_{j=1}^{n} c_j\mathbf{v}_j \longleftrightarrow \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}_B$$

Si $\mathbf{w} = T(\mathbf{v})$, entonces

$$\begin{aligned} \mathbf{w} = T(\mathbf{v}) &= T\left(\sum_{j=1}^{n} c_j\mathbf{v}_j\right) = \sum_{j=1}^{n} c_j T(\mathbf{v}_j) \\ &= \sum_{j=1}^{n} c_j(a_{ij}\mathbf{w}_1 + a_{2j}\mathbf{w}_2 + \cdots + a_{mj}\mathbf{w}_m) = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{m} a_{ij}\mathbf{w}_i\right) \\ &= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}c_j\right) \mathbf{w}_i \end{aligned}$$

con lo que $\mathbf{w}$ está escrita como combinación lineal de los vectores de la base $C = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$. Entonces la representación en coordenadas de $\mathbf{w}$ respecto a la base C es

$$T(\mathbf{v}) = \mathbf{w} \longleftrightarrow \begin{pmatrix} \sum_{j=1}^{n} a_{1j}c_j \\ \sum_{j=1}^{n} a_{2j}c_j \\ \vdots \\ \sum_{j=1}^{n} a_{mj}c_j \end{pmatrix}_C$$

Este vector no es más que el producto de matrices

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}$$

Así,

$$T(\mathbf{v}) = \mathbf{w} \longleftrightarrow A_B^C V_B. \tag{3.4}$$

Podemos calcular $T(\mathbf{v})$ de dos formas: primero, la *forma directa*, usando la parte izquierda de (3.4), evaluando directamente cómo T transforma $\mathbf{v}$; segundo, la *forma indirecta*, usando la parte derecha de (3.4), multiplicando la matriz asociada a T por la representación en coordenadas de $\mathbf{v}$ para obtener $A_B^C V_B$, la representación en coordenadas de $\mathbf{w}$ (m-tupla), de donde es fácil obtener $\mathbf{w}$.

3.4. Propiedades de las aplicaciones lineales

Como es posible representar una transformación lineal entre espacios vectoriales n-dimensionales mediante una matriz, podemos usar todo lo que sabemos acerca de las matrices para obtener información sobre las aplicaciones lineales. Al revés, como las matrices son aplicaciones lineales, podemos formular propiedades de las matrices a partir de las aplicaciones lineales. Algunas veces será más fácil descubrir propiedades trabajando con matrices, debido a que su estructura es muy concreta. Otras veces será más fácil trabajar directamente con aplicaciones lineales en abstracto, ya que su estructura es muy simple.

Teorema 3.4.1 *Si $T : V \longrightarrow W$ es una aplicación lineal, entonces $T(\mathbf{0}) = \mathbf{0}$.*

En términos de matrices, este teorema establece simplemente que el producto de una matriz por una matriz columna cero es de nuevo una matriz columna cero.

El teorema 3.4.1 afirma que una transformación lineal siempre transforma el vector cero del dominio en el vector cero de W. Puede ocurrir, sin embargo, que éste no sea el único vector que se transforma en el vector cero. Así, por ejemplo, la aplicación lineal $L : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ definida como $L(a, b) = (a, 0)$ proporciona $L(0, 1) = (0, 0)$ y en general $L(0, k) = (0, 0)$ para todo k. Aquí, L transforma infinitos vectores en el vector cero. Por contra, la aplicación identidad $I(\mathbf{v}) = \mathbf{v}$ sólo transforma el vector cero en el vector cero.

Definimos el *núcleo* (o espacio nulo) de una aplicación lineal $T : V \longrightarrow W$, denotado por $Ker(T)$ o $N(T)$, como el conjunto de todos los vectores $\mathbf{v} \in V$ que se transforman mediante T en el vector cero de W; esto es, *todos los $\mathbf{v} \in V$ para los cuales $T(\mathbf{v}) = \mathbf{0}$.*

Teorema 3.4.2 *El núcleo de una aplicación lineal es un subespacio vectorial del dominio.*

En términos de una matriz A concreta, el núcleo es el conjunto de vectores columna X que satisfacen la ecuación matricial $AX = 0$. Esto es, $N(A)$ es el conjunto de todas las soluciones del sistema de ecuaciones homogéneo $AX = 0$.

Ejemplo 29 *Determinar el núcleo de la matriz $A = \begin{pmatrix} 1 & 1 & 5 \\ 2 & -1 & 1 \end{pmatrix}$.*

<u>Solución</u>. Los vectores del núcleo son de la forma $\mathbf{x} = (x, y, z)$ y verifican

$$\begin{pmatrix} 1 & 1 & 5 \\ 2 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

es decir,

$$\begin{aligned} x + y + 5z &= 0 \\ 2x - y + z &= 0 \end{aligned}$$

con solución $x = -2z$, $y = -3z$. Esto es, es el subespacio vectorial de $\mathbb{R}^3$ generado por el vector $(-2, -3, 1)$. $\square$

La *imagen* de una aplicación $T : V \longrightarrow W$, denotada $Im(T)$, es el conjunto de los vectores de W que están relacionados con al menos un vector de V; esto es, $\mathbf{w}$ está en la imagen de T si y sólo si existe al menos un vector $\mathbf{v}$ en el dominio para el que $T(\mathbf{v}) = \mathbf{w}$. Si T es lineal, $Im(T)$ siempre contiene al vector $\mathbf{0} \in W$, ya que $T(\mathbf{0}) = \mathbf{0}$. Más aún,

Teorema 3.4.3 *La imagen de una aplicación lineal $T : V \longrightarrow W$ es un subespacio vectorial de W.*

Es importante darse cuenta de que el núcleo y la imagen de una aplicación lineal $T : V \longrightarrow W$ conceptualmente son subespacios diferentes: el núcleo es un subespacio del dominio V, mientras que la imagen es un subespacio de W. Se tiene además el importante resultado siguiente.

Teorema 3.4.4 *Si $T : V \longrightarrow W$ es una aplicación lineal, V es un espacio vectorial de dimensión finita y $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ es una base de V, entonces*

- *$\{T(\mathbf{v}_1), T(\mathbf{v}_2), \ldots, T(\mathbf{v}_n)\}$ es un sistema generador de $Im(T)$.*
- *$\dim(N(T)) + \dim(Im(T)) = \dim(V) = n$.*

Fijémonos que este resultado no depende de la dimensión de W.

Una aplicación lineal $T : V \longrightarrow W$ es uno a uno (también llamada *inyectiva*) si la igualdad $T(\mathbf{u}) = T(\mathbf{v})$ implica que $\mathbf{u} = \mathbf{v}$. Una transformación lineal uno a uno transforma vectores diferentes de V en vectores diferentes de W. A menudo el método más directo para ver si una aplicación lineal es uno a uno es usar el siguiente

Teorema 3.4.5 *Una aplicación lineal $T : V \longrightarrow W$ es uno a uno si y sólo si el núcleo de T contiene únicamente al vector cero, esto es, si $\dim(N(T)) = 0$.*

Ejemplo 30 *Sea $B = \{\mathbf{v}_1 = (1,1,1), \mathbf{v}_2 = (1,1,0), \mathbf{v}_3 = (1,0,0)\}$ una base de $\mathbb{R}^3$ y sea $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ una aplicación lineal tal que $T(\mathbf{v}_1) = (1,0)$, $T(\mathbf{v}_2) = (2,-1)$, $T(\mathbf{v}_3) = (4,3)$. Calcular:*

1. *$Ker(T)$, $Im(T)$ y sus dimensiones.*
2. *La imagen del vector $(2, -3, 5)$*

<u>Solución</u>.

1. Por definición, $Ker(T) = \{\mathbf{v} \in \mathbb{R}^3 / T(\mathbf{v}) = \mathbf{0}\}$ e $Im(T) = \{\mathbf{w} \in \mathbb{R}^2 / T(\mathbf{v}) = \mathbf{w}, \quad \forall \mathbf{v} \in \mathbb{R}^3\}$.

   Un vector $\mathbf{v} \in \mathbb{R}^3$ lo podemos expresar en la base B como

   $\mathbf{v} = (x,y,z)_B = x(1,1,1) + y(1,1,0) + z(1,0,0)$ y por tanto $T(\mathbf{v}) = xT(1,1,1) + yT(1,1,0) + zT(1,0,0) = x(1,0) + y(2,-1) + z(4,3) = (0,0) \Rightarrow$

   $$\left.\begin{array}{r} x + 2y + 4z = 0 \\ -y + 3z = 0 \end{array}\right\} \Rightarrow \left.\begin{array}{c} x = -10z \\ y = 3z \end{array}\right\} \Rightarrow Ker(T) = \langle (-10, 3, 1)_B \rangle \Rightarrow$$

$\dim(Ker(T)) = 1$.

Por otra parte, $Im(T) = \langle(1,0),(2,-1),(4,3)\rangle$ y se puede comprobar que el tercero es combinación lineal de los otros dos, por tanto una base de $Im(T)$ es $\{(1,0),(2,-1)\} \Rightarrow \dim(Im(T)) = 2$.

2. Hemos de expresar el vector en la base B,

$(2,-3,5) = x(1,1,1) + y(1,1,0) + z(1,0,0) \Rightarrow$

$$\left.\begin{array}{r} x+y+z=2 \\ x+y=-3 \\ x=5 \end{array}\right\} \Rightarrow \left.\begin{array}{c} x=5 \\ y=-8 \\ z=5 \end{array}\right\} \Rightarrow$$

$(2,-3,5) = 5\,(1,1,1) - 8\,(1,1,0) + 5\,(1,0,0) = (5,-8,5)_B$

$\Rightarrow T\,(2,-3,5) = 5T\,(1,1,1) - 8T\,(1,1,0) + 5T\,(1,0,0) = 5\,(1,0) - 8\,(2,-1) + 5\,(4,3) = (9,23)$

□

3.5. Cambio de base

Las representaciones en coordenadas de los vectores en un espacio vectorial n-dimensional son dependientes de la base, de forma que *diferentes bases generalmente resultan en diferentes representaciones de n-tuplas para el mismo vector*. Es natural preguntarse, por consiguiente, si las *diferentes* representaciones en coordenadas de un *mismo* vector están relacionadas.

Sean $C = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ y $D = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dos bases de un espacio vectorial V. Si $\mathbf{v} \in V$, entonces $\mathbf{v}$ se puede expresar como una única combinación lineal de los vectores de la base C:

$$\mathbf{v} = c_1\mathbf{u}_1 + c_2\mathbf{u}_2 + \cdots + c_n\mathbf{u}_n = \sum_{j=1}^{n} c_j\mathbf{u}_j$$

De forma similar, si consideramos la base D, tendremos

$$\mathbf{v} = d_1\mathbf{v}_1 + d_2\mathbf{v}_2 + \cdots + d_n\mathbf{v}_n = \sum_{i=1}^{n} d_i\mathbf{v}_i.$$

Las representaciones en coordenadas de $\mathbf{v}$ con respecto a C y D son, respectivamente,

$$V_C = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}_C \qquad \text{y} \qquad V_D = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}_D$$

Como cada vector de la base C también es un vector de V, se puede expresar como combinación lineal de los vectores de la base D. En particular,

$$\begin{aligned}
\mathbf{u}_1 &= p_{11}\mathbf{v}_1 + p_{21}\mathbf{v}_2 + \cdots + p_{n1}\mathbf{v}_n \\
\mathbf{u}_2 &= p_{12}\mathbf{v}_1 + p_{22}\mathbf{v}_2 + \cdots + p_{n2}\mathbf{v}_n \\
\vdots \;\; & \;\;\; \vdots \\
\mathbf{u}_n &= p_{1n}\mathbf{v}_1 + p_{2n}\mathbf{v}_2 + \cdots + p_{nn}\mathbf{v}_n
\end{aligned}$$

para algunos escalares $p_{i,j}$, con lo cual se puede escribir

$$\mathbf{u}_1 \longleftrightarrow \begin{pmatrix} p_{11} \\ p_{21} \\ \vdots \\ p_{n1} \end{pmatrix}_D , \quad \mathbf{u}_2 \longleftrightarrow \begin{pmatrix} p_{12} \\ p_{22} \\ \vdots \\ p_{n2} \end{pmatrix}_D , \ldots , \quad \mathbf{u}_n \longleftrightarrow \begin{pmatrix} p_{1n} \\ p_{2n} \\ \vdots \\ p_{nn} \end{pmatrix}_D .$$

Si usamos estas n-tuplas como las columnas de una matriz P, tenemos

$$P_C^D = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$$

llamada *matriz de transición desde la base C a la base D* o también *matriz de cambio de base.*

Resulta entonces

$$\mathbf{v} = \sum_{j=1}^{n} c_j \mathbf{u}_j = \sum_{j=1}^{n} c_j \left(\sum_{i=1}^{n} p_{ij}\mathbf{v}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} p_{ij} c_j \right) \mathbf{v}_i$$

pero también

$$\mathbf{v} = \sum_{i=1}^{n} d_i \mathbf{v}_i$$

y como la representación ha de ser única,

$$d_i = \sum_{j=1}^{n} p_{ij} c_j, \qquad i = 1, \ldots, n$$

que se puede escribir como producto de matrices:

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix}_D = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix}_C$$

o $V_D = P_C^D\, V_C$. Tenemos así probado el siguiente

Teorema 3.5.1 *Si V_C y V_D son las representaciones en coordenadas (n-tuplas) de un vector* $\mathbf{v}$ *con respecto a las bases C y D, respectivamente, y si P_j es la representación en coordenadas (n-tupla) del vector j-ésimo de la base C $(j = 1, \ldots, n)$ con respecto a la base D, entonces $V_D = P_C^D V_C$, donde la j-ésima columna de P_C^D es P_j.*

Ejemplo 31 *Obtener la matriz de transición entre las bases $C = \{(1,0),(0,1)\}$ y $D = \{(1,1),(1,-1)\}$ en $\mathbb{R}^2$ y verificar el teorema para el vector* $\mathbf{v} = (7,2)$.

Solución.

$$\begin{aligned}(1,0) &= \frac{1}{2}(1,1) + \frac{1}{2}(1,-1) \longleftrightarrow \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}_D \\ (0,1) &= \frac{1}{2}(1,1) - \frac{1}{2}(1,-1) \longleftrightarrow \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{pmatrix}_D\end{aligned}$$

con lo cual la matriz de cambio de base es

$$P_C^D = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}$$

de manera que, como

$$V_C = \begin{pmatrix} 7 \\ 2 \end{pmatrix}_C, \qquad \Longrightarrow V_D = P_C^D V_C = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 7 \\ 2 \end{pmatrix} = \begin{pmatrix} \frac{9}{2} \\ \frac{5}{2} \end{pmatrix}$$

□

Aunque el teorema tiene que ver con la matriz de transición de C a D, es igualmente válido en la dirección inversa para la matriz de transición de D a C. Si P_D^C representa esta matriz, entonces

$$V_C = P_D^C V_D$$

Ejemplo 32 *Verificar esta igualdad para el vector* $\mathbf{v}$ *y las bases del ejemplo 31.*

Solución. Recordemos que $C = \{(1,0),(0,1)\}$ y $D = \{(1,1),(1,-1)\}$. Ahora

$$\begin{aligned}(1,1) &\longleftrightarrow \begin{pmatrix} 1 \\ 1 \end{pmatrix}_C \\ (1,-1) &\longleftrightarrow \begin{pmatrix} 1 \\ -1 \end{pmatrix}_C\end{aligned}$$

con lo cual la matriz de cambio de base ahora es

$$P_D^C = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

y

$$P_D^C V_D = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \frac{9}{2} \\ \frac{5}{2} \end{pmatrix} = \begin{pmatrix} 7 \\ 2 \end{pmatrix} = V_C$$

□

Obsérvese que la matriz de transición P_D^C encontrada en este último ejemplo es la inversa de la matriz de transición P_C^D encontrada en el ejemplo 31. Esto no es una coincidencia.

Teorema 3.5.2 *La matriz de transición desde C a D, donde C y D son bases del mismo espacio vectorial de dimensión finita, es invertible, y su inversa es la matriz de transición desde D hasta C.*

Mostramos anteriormente que una aplicación entre espacios vectoriales de dimensión finita se puede representar mediante una matriz. Esa matriz, sin embargo, depende de la base que se considere. *En general, una aplicación lineal tiene muchas representaciones matriciales, una matriz diferente para cada pareja de bases en el dominio y en el conjunto final.*

Es natural preguntarse entonces si las *diferentes* matrices que representan la *misma* aplicación lineal están relacionadas. Aquí nos limitaremos a transformaciones lineales de un espacio vectorial en sí mismo, esto es, aplicaciones lineales de la forma $T : V \longrightarrow V$ (también llamadas *endomorfismos*), por lo que las diferentes matrices asociadas serán cuadradas. El resultado se puede generalizar fácilmente al caso $T : V \longrightarrow W$, con V y W diferentes.

Sea $T : V \longrightarrow V$ una aplicación lineal sobre un espacio vectorial n-dimensional V, con $\mathbf{w} = T(\mathbf{v})$. Si C es una base de V, entonces la representación en coordenadas de $\mathbf{w}$ con respecto a C se puede obtener como

$$W_C = A_C^C \, V_C$$

siendo A_C^C la matriz asociada a T en la base C. Si usamos una base diferente D, tendremos también

$$W_D = A_D^D \, V_D$$

Como V_C y V_D son representaciones en coordenadas del mismo vector $\mathbf{v}$ con respecto a bases diferentes, existe una matriz de transición P_C^D tal que

$$V_D = P_C^D \, V_C$$

e igual ocurrirá para el vector $\mathbf{w}$

$$W_D = P_C^D \, W_C$$

con lo que

$$W_D = A_D^D \, V_D = A_D^D \, P_C^D \, V_C$$

y también

$$W_D = P_C^D \, W_C = P_C^D \, A_C^C \, V_C$$

de modo que

$$P_C^D \, A_C^C \, V_C = A_D^D \, P_C^D \, V_C$$

Esta igualdad es válida para todas las n-tuplas V_C con respecto a la base C. A partir de aquí es inmediato concluir que

$$P_C^D \, A_C^C = A_D^D \, P_C^D$$

o bien

$$A_C^C = \left(P_C^D\right)^{-1} A_D^D \, P_C^D.$$

A la inversa, el mismo razonamiento muestra que si esta expresión es válida, entonces A_C^C y A_D^D son matrices asociadas a la misma aplicación lineal con respecto a las bases C y D, respectivamente, estando estas bases relacionadas mediante la matriz de transición P_C^D. Así pues,

Teorema 3.5.3 *Dos matrices A, B $n \times n$ representan la misma transformación lineal si y sólo si existe una matriz invertible P tal que $A = P^{-1}BP$.*

En otras palabras, y como veremos en el tema siguiente, A y B representan la misma aplicación lineal si y sólo si son similares.

De todas las matrices similares que pueden representar una aplicación lineal, algunas son más simples que otras. Una matriz especialmente simple es una diagonal. Dado un endomorfismo, ¿es posible saber por adelantado si existe una base del espacio vectorial tal que la matriz asociada al endomorfismo en esa base es diagonal? La respuesta a esta y otras cuestiones se verá en detalle en el siguiente tema.

3.6. Problemas

1. Estudia en cada uno de los casos siguientes si se trata o no de una aplicación lineal:

   a) $\begin{array}{rccc} h_\alpha : & V & \longrightarrow & V \\ & \vec{v} & \longmapsto & \alpha\,\vec{v} \end{array}$, donde V es un espacio vectorial y α es un escalar.

   b) $\begin{array}{rccc} t_a : & V & \longrightarrow & V \\ & \vec{v} & \longmapsto & \vec{v}+\vec{a} \end{array}$, donde V es un espacio vectorial y $\vec{a} \in V$.

   c) $\begin{array}{rccc} t_i : & \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ & (x_1,\ldots,x_n) & \longmapsto & x_i \end{array}$

   d) $\begin{array}{rccc} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ & (x,y,z) & \longmapsto & (2x-z, y+z, x) \end{array}$

   e) $\begin{array}{rccc} g : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^4 \\ & (x,y,z) & \longmapsto & (2x-z, 2x, 4z-y, z) \end{array}$

   f) $\begin{array}{rccc} h : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x,y,z) & \longmapsto & (x-y, x-z) \end{array}$

2. Encuentra una aplicación lineal f de $\mathbb{R}^3$ en $\mathbb{R}^3$ tal que $\operatorname{Ker} f = \langle(0,0,1)\rangle$ e $\operatorname{Im} f = \{(x,y,z) \in \mathbb{R}^3 / z = 0\}$.

3. Consideremos el endomorfismo (aplicación lineal de un espacio en sí mismo) de $\mathbb{R}^3$ dado por la relación $(x,y,z) \mapsto (y-z, -x+4z, y-z)$. Hállese el núcleo y la imagen, así como un sistema generador del núcleo.

4. Obtener $T(\vec{u})$ para una transformación lineal T si se sabe que $T(\vec{u}-\vec{v}) = 4\vec{u}+5\vec{v}$ y $T(\vec{u}+\vec{v}) = \vec{u}-3\vec{v}$.

5. Sean $S : V \longrightarrow W$ y $T : V \longrightarrow W$ dos aplicaciones lineales. Su suma es otra aplicación de V a W definida mediante $(S+T)(\vec{v}) = S(\vec{v}) + T(\vec{v})$ para todos los $\vec{v}$ de V. Probar que la aplicación $S+T$ es lineal.

6. Sea $T : V \longrightarrow W$ una transformación lineal y k un escalar dado. Se define una nueva transformación $k\,T : V \longrightarrow W$ mediante $(k\,T)(\vec{v}) = k(T(\vec{v}))$ para todos los $\vec{v}$ de V. Demostrar que la aplicación kT es lineal.

7. Sean $S : V \longrightarrow W$ y $T : V \longrightarrow V$ dos aplicaciones lineales, y definamos su producto como otra aplicación de V en V tal que $(ST)(\vec{v}) = S(T(\vec{v}))$ para todos los $\vec{v}$ de V. Este producto primero aplica T a un vector y después S al resultado. Demostrar que la aplicación ST es lineal.

8. Dada la aplicación lineal representada por la matriz

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

en la base canónica, halla el vector transformado del $\vec{x} = (1, -1, 2)$. Suponiendo que se cambia la base del espacio vectorial y se sustituye por la formada por los vectores $(1, 0, 3)$, $(3, 2, 1)$ y $(3, -5, 1)$, halla las coordenadas del vector dado y de su transformado en la nueva base. Halla la matriz de la aplicación en la nueva base.

9. Dado el endomorfismo de $\mathbb{R}^3$ cuya matriz en la base canónica es

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

demostrar que para cualquier valor de a la dimensión del subespacio imagen es dos. Hallar el núcleo y la imagen para $a = -2$.

10. Sea f la aplicación lineal de $\mathbb{R}^4$ en $\mathbb{R}^3$ definida por $f(1, 1, 1, 1) = (0, 0, 1)$, $f(1, 0, 1, 0) = (1, 1, -1)$, $f(1, 1, 1, 0) = (0, 0, -1)$, $f(-1, -2, 0, 0) = (1, 1, 1)$. Halla:

    a) La matriz asociada a f respecto de las bases canónicas.

    b) Bases de $\operatorname{Im} f$ y $\operatorname{Ker} f$.

11. Dada la aplicación lineal f definida por: $f(1, 0, 0, 0) = (1, 0, 0)$, $f(0, 1, 0, 0) = (0, 1, 0)$, $f(1, 1, 1, 0) = (1, -1, 0)$, $f(1, 1, 1, 1) = (0, 0, 1)$

    a) Halla la matriz de la aplicación respecto de las bases canónicas.

    b) Halla la matriz de la aplicación respecto de la bases canónica de $\mathbb{R}^4$ y la base $B' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$ de $\mathbb{R}^3$.

    c) Calcula la imagen del vector (x, y, z, t) y obtén las ecuaciones que caracterizan la transformación.

    d) Calcula los subespacios núcleo e imagen de f dando una base y la dimensión de los mismos.

12. Sea $f : \mathbb{R}^3 \to \mathbb{R}^3$ una aplicación lineal tal que

$$f(1,1,0) = (2,\alpha,2), \quad f(0,1,1) = (1+\alpha,\alpha,4), \quad f(0,1,-1) = (1-\alpha,\alpha,0)$$

    a) Determina la matriz asociada a f en la base canónica.
    b) Calcula $f(x,y,z)$ para cualquier $(x,y,z) \in \mathbb{R}^3$. Calcula $f(1,-1,2)$.
    c) Calcula la dimensión del subespacio imagen de f en función del parámetro α. ¿Existe algún valor de α tal que $\operatorname{Im} f = \mathbb{R}^3$? ¿Y alguno para el que el núcleo de f sea sólo el vector nulo? Justifica las respuestas.

13. Se considera la aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^4$ definida por $f(1,1,0) = (3,2,0,1)$, $f(2,3,1) = (1,-2,1,1)$, $f(0,-2,1) = (4,0,1,2)$. Calcula la matriz asociada a f en las bases canónicas y los subespacios núcleo e imagen.

14. Sea la aplicación $f : \mathbb{R}^3 \to \mathbb{R}^2$ dada por $f(-1,1,1) = (1,2)$, $f(0,-1,1) = (-1,0)$. Se pide:

    a) Calcula $f(x,y,z)$ para cualquier vector $(x,y,z) \in \mathbb{R}^3$.
    b) Calcula los subespacios núcleo e imagen de f.
    c) Determina los vectores de $\mathbb{R}^3$ que se transforman en el vector $(0,1)$ mediante f.

15. Dada la aplicación $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida por

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

calcula su matriz asociada respecto de la base $B = \{(1,0,-1),(0,1,-1),(1,1,1)\}$.

16. Dada la aplicación lineal $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por

$$\begin{aligned} f(1,1) &= (1,2,0) \\ f(1,-1) &= (5,0,2) \end{aligned}$$

    a) Halla su matriz asociada, A, en las bases canónicas.
    b) Obtener una base y la dimensión de los subespacios imagen y núcleo de f.
    c) Dado el vector $\vec{v} = (-1,3,-1) \in \mathbb{R}^3$ encontrar el vector $\vec{u}$ tal que $f(\vec{u}) = \vec{v}$. Calcula las coordenadas de $\vec{v}$ en la base del apartado (b).

17. Obtener la representación matricial de la aplicación lineal $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ definida mediante $T(a,b,c) = (2a+3b-c, 4b+5c)$ respecto a las bases $B = \{(1,0,0),(1,1,0),(1,1,1)\}$ y $C = \{(1,1),(0,1)\}$.

18. Dada las aplicación lineal $T : V \longrightarrow V$

$$T(a,b,c) = (3a-b+c, 2a-2c, 3a-3b+c)$$

y las bases de V : $B = \{(1,0,0),(0,1,0),(0,0,1)\}$ y $B' = \{(1,1,0),(0,1,1),(1,0,1)\}$, *a*) Obtener su representación matricial con respecto a dichas bases bases

b) Comprobar que ambas son similares, con una matriz de transición apropiada.

19. Dada las aplicación lineal $T : V \longrightarrow V$

$$T(a, b, c) = (a, 2b, -3c)$$

y las bases de V : $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ y $B' = \{(1, 0, 0), (1, 1, 0), (1, 1, 1)\}$,

a) Obtener su representación matricial con respecto a dichas bases.

b) Comprobar que ambas son similares, con una matriz de transición apropiada.

20. Sean $B = \{(1, 1, 1), (1, 1, 0), (1, 0, 0)\}$ y $V = \{(1, 2), (3, 0)\}$ bases de $\mathbb{R}^3$ y $\mathbb{R}^2$, respectivamente. Calcular la matriz asociada a las aplicaciones:

a) $f : \mathbb{R}^3 \to \mathbb{R}^3$ definida por $f(x, y, z) = (x, x + y + 2z, 2y - 3z)$ respecto a la base B.

b) $f : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $f(x, y, z) = (2x - 3y + z, x + y - z)$ respecto a la bases B y V.

c) $f : \mathbb{R}^2 \to \mathbb{R}^3$ definida por $f(x, y) = (2x + 2y, 0, 3x - y)$ respecto a V y B.

21. Sean U, V dos espacios vectoriales y $f : U \to V$ una aplicación lineal tal que en ciertas bases de U y V tiene asociada la matriz

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

Calcular $\dim U$, $\dim V$, $\dim(\mathrm{Im} f)$, $\dim(\mathrm{Ker} f)$.

22. Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ la aplicación lineal definida mediante

$$f(x_1, x_2, x_3) = (x_2 - x_3,\ x_1 + x_2,\ x_1 - x_3,\ x_1 + x_2 + x_3)$$

Se pide:

a) Matriz asociada a f respecto de las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^4$.

b) Subespacios Núcleo e Imagen de f, facilitando una base de cada uno de ellos.

c) Matriz asociada a f respecto de la base $\{\vec{u}_1 = (0, 1, 1),\ \vec{u}_2 = (1, 1, 0),\ \vec{u}_3 = (1, 0, 1)\}$ de $\mathbb{R}^3$ y de la base canónica de $\mathbb{R}^4$.

Justificar las respuestas.

23. Sean V y W espacios vectoriales reales, $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ base de V y $\{\vec{u}_1, \vec{u}_2, \vec{u}_3, \vec{u}_4\}$ base de W. Sea $f : V \longrightarrow W$ la aplicación lineal definida mediante

$$f(\vec{e}_1) = \lambda\vec{u}_1 + \vec{u}_2 \quad f(\vec{e}_2) = \vec{u}_3 + \vec{u}_4 \quad f(\vec{e}_3) = \vec{u}_1 + \vec{u}_2 + \vec{u}_3 + \vec{u}_4$$

con $\lambda \in \mathbb{R}$ dado. Hállese:

   a) La matriz asociada a f en las bases dadas.
   b) El núcleo y el rango de f, analizando cuándo es inyectiva, esto es, cuando su núcleo tiene dimensión cero.

24. Sea $f : \mathbb{R}^3 \to \mathbb{R}^4$ el homomorfismo definido mediante $f(1,0,0) = (1,0,1,1)$, $f(0,1,0) = (0,1,1,1)$, $f(0,0,1) = (1,-1,0,a)$, donde $a \in \mathbb{R}$ es fijo.

   a) Obténgase la matriz asociada a f respecto de las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^4$.
   b) Dado un vector $\vec{u} \in \mathbb{R}^3$ de componentes (x, y, z) respecto de la base canónica, calcular $f(\vec{u})$.
   c) Localícese el núcleo de f, dando una base del mismo.
   d) Calcúlese el subespacio imagen de f, dando una base del mismo.

25. Sea $a \in \mathbb{R}$ y f una aplicación lineal de $\mathbb{R}^3$ en $\mathbb{R}^3$ definida por $f(x, y, z) = (3x + 4y + z, 2x + 6y + 3z, x + 3y + az)$. Se pide:

   a) Calcular la matriz asociada a f en la base canónica de $\mathbb{R}^3$.
   b) Calcular la matriz asociada a f en la base formada por

$$\vec{v}_1 = (1,0,3), \quad \vec{v}_2 = (3,2,1), \quad \vec{v}_3 = (3,-5,1)$$

   c) Analizar para qué valores de a la dimensión del núcleo de f es cero.

26. Sea f el endomorfismo $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ definido por $f(x, y, z) = \frac{1}{2}(-x - y, x + y, 2z)$.

   a) Calcular una base de Ker f y una base de Im f.
   b) Encontrar una base del subespacio vectorial Ker $f \cap$ Im f.
   c) Sea $B' = \{\vec{b}_1, \vec{b}_2, \vec{b}_3\}$ la base de $\mathbb{R}^3$ formada por los vectores

$$\vec{b}_1 = \frac{1}{\sqrt{2}}(1,1,0), \qquad \vec{b}_2 = \frac{1}{\sqrt{2}}(-1,1,0), \qquad \vec{b}_3 = (0,0,1)$$

   Encontrar la matriz del endomorfismo f en esta base B'.
   d) Sea $\vec{x} = 3\vec{b}_1 + 2\vec{b}_2 - \vec{b}_3$. Encontrar $f(\vec{x})$ y expresarlo en la base B' y en la base canónica de $\mathbb{R}^3$.
   e) ¿Es la imagen de la base canónica una base de $\mathbb{R}^3$? ¿Es la imagen de B' una base de $\mathbb{R}^3$?

27. Dada la aplicación lineal $f:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$ definida por $f(x,y)=(x-2y,y,3x+y)$,

    a) obtener su matriz asociada si en ambos espacios consideramos las bases canónicas;

    b) calcular una base de Ker f, una base de Im f y una base de Ker $f \cap$ Im f;

    c) hallar la matriz asociada a la aplicación lineal si consideramos las bases $B_2=\{(2,1),(0,3)\}$, $B_3=\{(1,0,0),(1,1,0),(0,0,2)\}$;

    d) ¿es la imagen de B_2 una base de $\mathbb{R}^3$? ¿Por qué?

28. Dada la aplicación lineal $f:\mathbb{R}^3 \to \mathbb{R}^3$ definida por:

$$\begin{aligned} f(2,0,2) &= (2,1,3) \\ f(1,-1,0) &= (-1,-2,-2) \\ f(0,0,2) &= (0,2,2) \end{aligned}$$

    Calcular:

    a) La matriz asociada a f en la base canónica de $\mathbb{R}^3$.

    b) El vector transformado mediante f de $\vec{u}=(x,y,z)$, así como la imagen del vector $(3,1,-2)$.

    c) Subespacios núcleo e imagen de f, dando una base de dichos subespacios.

    d) El vector $\vec{u}$ tal que $f(\vec{u})=\vec{v}$, siendo $\vec{v}$ un vector cuyas coordenadas en la base
    $B=\{\vec{e}_1=(0,1,1),\vec{e}_2=(0,2,0),\vec{e}_3=(1,0,-2)\}$ son $(1,2,1)$.

29. Sean las aplicaciones $f:\mathbb{R}^2\to\mathbb{R}^2$ y $g:\mathbb{R}^2\to\mathbb{R}^3$ dadas por $f(x,y)=(x-y,x+y)$ y $g(1,0)=(1,0,1)$, $g(0,1)=(-1,0,1)$.

    a) Hallar las ecuaciones de la aplicación $g\circ f:\mathbb{R}^2\to\mathbb{R}^3$ y calcular el núcleo y la imagen de $g\circ f$. (Recordad que $(g\circ f)(\vec{x})=g(f(\vec{x}))$).

    b) Hallar la matriz de la aplicación $g\circ f$ respecto de la base de $\mathbb{R}^2$ $\{(1,1),(0,2)\}$ y de la canónica de $\mathbb{R}^3$.

    c) Calcular $(g\circ f)(1,1)$ y $(g\circ f)^{-1}(1,1,1)$.

30. Dada la matriz

$$A=\begin{pmatrix} 1 & -1 & 3 \\ 5 & 6 & -4 \\ 7 & 4 & 2 \end{pmatrix}$$

    a) Encontrar una aplicación lineal $f:\mathbb{R}^3\to\mathbb{R}^3$ de manera que A sea la matriz asociada a f en la base canónica de $\mathbb{R}^3$. Hallar la imagen de cualquier vector $\vec{u}=(x,y,z)$.

    b) Hallar el núcleo y la imagen de f, dando una base y su dimensión.

c) Calcular la matriz asociada a f en la base de $\mathbb{R}^3$ formada por los vectores $\{(1,1,1),(1,1,0),(1,0,0)\}$

31. Sea la aplicación $f : \mathbb{R}^2 \to \mathbb{R}^3$ dada por $f(-1,1) = (0,1,2)$, $f(0,-1) = (4,-1,0)$. Se pide:

    a) Calcular $f(x,y)$ para cualquier vector $(x,y) \in \mathbb{R}^2$.

    b) Calcular los subespacios núcleo e imagen de f.

    c) Determinar los vectores de $\mathbb{R}^2$ que se transforman en el vector $(0,1,2)$ mediante f.

32. Dada la aplicación lineal definida por

$$f(1,0,2) = (5,1), \qquad f(0,-1,1) = (3,-3), \qquad f(0,0,1) = (2,0)$$

    a) Determina los espacios inicial y final. Da explícitamente las ecuaciones que caracterizan la aplicación.

    b) Halla la matriz asociada a dicha aplicación respecto de las bases canónicas.

    c) Halla la matriz asociada a f respecto de las bases:

$$\begin{aligned} B_1 &= \{(1,0,-1),(2,1,0),(0,1,1)\} \\ B_2 &= \{(-2,1),(1,-1)\} \end{aligned}$$

    d) Halla los subespacios núcleo e imagen de la aplicación dada.

3.7. Aplicaciones lineales en economía

1. Un fabricante produce 3 artículos diferentes, cada uno de los cuales requiere para su elaboración de 2 materias primas. Los tres productos se denotas por p_1, p_2 y p_3 y las dos materias primas por M_1 y M_2. En la tabla que se indica a continuación se representa el número de unidades de cada materia prima que se requiere para elaborar una unidad de producto

	p_1	p_2	p_3
M_1	1	2	1
M_2	3	1	2

Se pide

    a) Determinar la ley que asocia a cada vector de producción (p_1,p_2,p_3) el vector de materias primas $(M_1\ ,M_2)$ que le corresponde para que dicha producción sea posible.

    b) ¿Es una aplicación lineal la ley determinada en el apartado anterior?

    c) Determinar los subespacios núcleo e imagen e interpretarlos.

2. En la Bolsa de Madrid se cotizan dos grandes grupos de valores: las acciones de las empresas privadas y la deuda pública. Tenemos definido un índice de cotización 'normal' para cada grupo, basado en la actividad normal del mercado. Además se consideran índices 'extraordinarios' que miden distorsiones positivas o negativas alrededor de la cotización normal y que denominamos IEA e $IEDP$, cuyos valores dependerán de la masa monetaria m, del exceso de reservas bancarias e y del crédito del Banco de España al sistema bancario c de acuerdo con las expresiones

$$\begin{aligned} IEA &= -2m - 3e + 2c \\ IEDP &= -m - 4e + c \end{aligned}$$

Se pide:

a) ¿Se puede interpretar este problema como una aplicación lineal? Demostrarlo.

b) Obtener los subespacios núcleo e imagen de la aplicación e interpretarlos económicamente.

3. Estudiemos los movimientos de capital entre España, Francia y Reino Unido. Definamos el saldo de la balanza de capitales de un país como la diferencia entre las entradas de capital en ese país menos las salidas. Se sabe que en cada país la situación depende de los tipos de interés reales vigentes (el nominal menos la tasa de inflación del país), y la interrelación se puede resumir en:

$$\begin{aligned} S_E &= 0{,}4R_E - 0{,}5R_F - 0{,}3R_I \\ S_F &= -0{,}3R_E + 0{,}6R_F - 0{,}4R_I \\ S_I &= -0{,}1R_E - 0{,}1R_F + 0{,}3R_I \end{aligned}$$

siendo S_E, S_F y S_I los saldos de España, Francia y Reino Unido, respectivamente, y R_E, R_F y R_I los tipos de interés reales en esos tres países. Se pide:

a) ¿Podríamos hablar aquí de aplicación lineal?

b) Calcular el núcleo y la imagen. Interpretar económicamente el resultado.

c) Supóngase que la interrelación pasa a ser:

$$\begin{aligned} S_E &= 0{,}4R_E - 0{,}5R_F - 0{,}3R_I \\ S_F &= -0{,}3R_E + 0{,}6R_F - 0{,}4R_I \\ S_I &= -0{,}1R_E - 0{,}1R_F + 0{,}9R_I \end{aligned}$$

Calcular el nuevo núcleo interpretando económicamente el resultado.

4. Consideremos una economía dividida en tres sectores: agrícola, industrial y de servicios. Sean P_A, P_I y P_S el porcentaje de variación de los precios de un año

para otro en los sectores agrícola, industrial y de servicios, respectivamente, información que vamos a representar en un vector de la forma

$$\vec{P} = (P_A, P_I, P_S)$$

Sea w la tasa de variación de salarios de un año a otro, p_m la tasa de variación de precios de importaciones en el mismo periodo y t la tasa de variación de los impuestos, que se pueden agrupar en un vector de la forma

$$\vec{x} = (w, p_m, t)$$

Sabemos que las componentes del vector $\vec{P}$ dependen de las componentes de $\vec{x}$, pero no conocemos la relación concreta que las liga; lo único que sabemos es que un aumento del 1 % en la tasa de variación de salarios, manteniéndose constantes los precios de importaciones y los impuestos, provoca que los precios agrícolas aumenten un 2 %, los industriales un 1,5 % y los del sector servicios otro 2 %, es decir:

$$(1, 0, 0) \longmapsto (2, 1{,}5, 2)$$

Si lo que aumenta son los precios de importaciones

$$(0, 1, 0) \longmapsto (1{,}5, 1{,}5, 1)$$

y si aumentan los impuestos

$$(0, 0, 1) \longmapsto (1, 1, 1{,}5)$$

Se pide:

a) Calcular la relación que liga las componentes del vector $\vec{x}$ con las del vector $\vec{P}$.

b) Se espera que para el próximo año los salarios crezcan un 8 %, los precios de importaciones un 2 % y la presión fiscal un 10 %. ¿En qué sector aumentarán más los precios?

5. CERÁMICAS ORIOL S.A. es una empresa que se dedica a la fabricación de azulejos y goza de gran prestigio gracias a una política de segundas marcas que consiste en lo siguiente. Según la calidad de las piezas, que por razones técnicas no es constante, éstas se distribuyen en tres marcas, de forma que únicamente llevan la marca CERÁMICAS ORIOL las de calidad alta, mientras que las de calidad media y baja se comercializan con otros nombres que nada tienen que ver con el de la empresa. Las ventas V de cada tipo de azulejos dependen del grado de actividad del sector de la construcción I, de su precio y del precio de las otras dos marcas (P_a, P_m, P_b) según las siguientes relaciones:

$$\begin{aligned} V_a &= 2I - 0{,}2P_a + 0{,}4P_m + 0{,}1P_b \\ V_m &= 4I + 0{,}4P_a - 0{,}3P_m + 0{,}5P_b \\ V_b &= 5I + 0{,}1P_a + 0{,}5P_m - 0{,}2P_b \end{aligned}$$

Se pide:

a) Calcular una base de los subespacios Imagen y Núcleo de la aplicación que se desprende del enunciado.

b) ¿En qué condiciones las ventas totales serán nulas?

c) ¿Se modifica la respuesta si las funciones de ventas de cada tipo de azulejos sólo dependieran del grado de actividad del sector de la construcción y de los precios del resto de los productos?

6. El señor Oriol montó hace años una fábrica de juguetes de artesanía que ahora dirige su hijo mayor. El motivo que le llevó a realizar dicha inversión roza lo sentimental, y es que su infancia no estuvo sobrada de juguetes, más bien al contrario, pues su familia era humilde y no podía permitirse esos lujos. Por ello cuando le plantearon el negocio no se lo pensó dos veces.

   Desde el principio la fábrica sólo se ha dedicado a producir dos tipos de juguetes, caballos de cartón y muñecas de trapo. Dado lo peculiar de la producción y tras duras negociaciones, hace dos años consiguieron una subvención del gobierno cuya cuantía S se determina cada año y que repercute en diferente proporción sobre los costes variables según el tipo de producto de que se trate.

   Como es natural, en una fábrica de este tipo la utilización de maquinaria es casi simbólica, por eso la repercusión en los costes variables del factor capital K es mucho menor que la del factor trabajo L, lo cual queda reflejado en las expresiones siguientes:

	Costes variables netos
Caballos de cartón	$2L + 0{,}3K - 1{,}5S$
Muñecas de trapo	$3L + 0{,}1K - 2S$

   siendo los costes variables netos = costes variables - subvención (por unidad de producto).

   No obstante, dado el notable incremento que se ha producido en la demanda de estos productos, el hijo del señor Oriol ha considerado la posibilidad de introducir nueva maquinaria en el proceso productivo, de tal manera que éste pase de ser intensivo en mano de obra a ser intensivo en capital. Si se sabe que con el nuevo equipo la utilización de mano de obra se reduce a la cuarta parte, mientras que la de capital permanece constante y la subvención se duplica, se pide:

   a) Hallar la matriz asociada a la aplicación que, con el equipo existente, transforma el precio de los factores productivos y la cuantía de la subvención en los costes variables netos unitarios.

   b) Hallar la matriz asociada a la misma aplicación pero en el caso de que se incorporara la nueva tecnología.

   c) Calcular el Núcleo en ambos casos e interpretar el resultado.

   d) Si el coste de la mano de obra es de 400, el de la maquinaria es de 250 y la subvención es de 80, calcular la ventaja que tiene el proceder al cambio tecnológico en términos de costes variables netos.

7. El secretario de Estado de Economía ha observado que la evolución de los precios de los bienes duraderos es diferente de la de los bienes de consumo (donde se incluyen los servicios). Bajo esta hipótesis ha definido un índice de precios 'normal' para cada grupo, pero además considera la existencia de índices 'extraordinarios' que miden las variaciones positivas o negativas alrededor de la tendencia del índice de precios normal (IED, IEC), cuyos valores ha contrastado que dependen de la masa monetaria M, de los tipos de interés i y del nivel de producción del país Q según las relaciones siguientes:

$$\begin{aligned} IED &= 2M + 3i - Q \\ IEC &= M + 0{,}5i - 2Q \end{aligned}$$

Antes de comunicarle al ministro de Economía su teoría, fue a pedir consejo a su amigo, profesor de Matemáticas en la universidad, el cual, tras comprobar que las relaciones entre los índices extraordinarios y las variables M, i y Q eran aceptables, pasó a analizar el problema como una aplicación lineal.

Tras un análisis pausado y metódico le dijo: "Estás en lo cierto, estas variaciones existen e influyen notablemente en las evoluciones normales de los precios. No obstante, antes de ir a hablar con el ministro, comprueba cuál es el valor que en la actualidad toman las variables M, i, Q, no fuera a darse el caso de que dicho vector perteneciera al núcleo de la aplicación y nos derrumbara toda la teoría".

Se pide:

a) ¿Hace lo correcto el profesor al analizar el planteamiento como un problema de aplicación lineal?

b) ¿Cuál es el núcleo de la aplicación?

c) ¿Por qué le dice el profesor que si los valores de M, i y Q pertenecen al núcleo la teoría del secretario de Estado se derrumba?

8. El gobierno de un país ha decidido, en vista del elevado incremento del número de accidentes de circulación, establecer medidas sancionadoras con el fin de frenar dicha tendencia. Para ello ha diseñado un sistema de puntos mediante el cual al llegar al máximo se produce la retirada por un año del carnet de conducir.

La medida ha alarmado a las dos empresas afincadas en el país que se dedican a la fabricación de automóviles, Tatra y Mosva, puesto que las ventas de los dos tipos de automóviles que ambas producen (de lujo y utilitario) están en función de la renta disponible de la clase asalariada (Y), del incremento del número de accidentes de circulación producido en el periodo anterior (N) y del incremento neto de carnets de conducir expedidos en el perido actual (carnets expedidos - carnets retirados C), según las relaciones siguientes:

	Tatra	Mosva
de lujo	$T_L = 0{,}2Y + 0{,}1N + 0{,}2C$	$M_L = 0{,}3Y + 0{,}1N + 0{,}1C$
utilitario	$T_U = 0{,}3Y + 0{,}2N + 0{,}1C$	$M_U = 0{,}1Y + 0{,}2N + 0{,}1C$

Sabiendo que en el momento en que va a entrar en vigor la nueva normativa las variables económicas que preocupan a los fabricantes de coches toman los siguientes valores: $Y = 3000$, $N = 50$, $C = 60$, se pide:

a) ¿Cuáles son las ventas de automóviles de las dos marcas nacionales en ese momento?

b) ¿En qué condiciones, al entrar la nueva normativa las ventas de coches nacionales se anularían?

c) ¿Variaría la respuesta en caso de que la normativa no entrara en vigor?

Capítulo 4

DIAGONALIZACIÓN DE MATRICES

4.1. Introducción

Comenzaremos el tema exponiendo la noción de semejanza, así como algunas propiedades de la misma. Como veremos, esta definición juega un papel importante en el concepto de matriz diagonalizable.

Definición 5 *Dadas dos matrices A, B, cuadradas $n \times n$, se dice que A es* **semejante** *a B si existe una matriz cuadrada $n \times n$ no singular P tal que $B = P^{-1}AP$.*

La matriz P se suele denominar *matriz de paso* asociada a esta relación de semejanza.

Éstas son algunas propiedades elementales de la relación de semejanza:

1. Propiedad reflexiva: A es semejante a ella misma.

2. Propiedad simétrica: si A es semejante a B, entonces B es semejante a A.

3. Propiedad transitiva: si A es semejante a B y B es semejante a C, entonces A y C son semejantes.

4. Dos matrices semejantes tienen el mismo rango.

5. Si A y B son semejantes, entonces $\det(A) = \det(B)$.

6. Si A y B son semejantes, entonces A^k y B^k son también semejantes con la misma matriz de paso.

Definición 6 *Una matriz cuadrada $A \in \mathcal{M}_n$ se dice que es* **diagonalizable** *si existe una matriz semejante a ella diagonal. Dicho de otro modo, A es diagonalizable si existe una matriz $P \in \mathcal{M}_n$ regular y una matriz $D \in \mathcal{M}_n$ diagonal tales que*

$$D = P^{-1}AP$$

Las matrices diagonales tienen propiedades atractivas. Se pueden multiplicar fácilmente, sus determinantes se calculan sin dificultad, se puede determinar rápidamente si tales matrices tienen inversa y, en caso afirmativo, sus inversas son fáciles de obtener. En particular, dada la matriz diagonal

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

resulta sencillo calcular su inversa (si $\lambda_i \neq 0$, $i = 1, 2, 3$):

$$D^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & 0 \\ 0 & \frac{1}{\lambda_2} & 0 \\ 0 & 0 & \frac{1}{\lambda_3} \end{pmatrix},$$

así como sus potencias sucesivas:

$$D^n = \begin{pmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{pmatrix}.$$

Si A es una matriz diagonalizable, es decir, si es semejante a una matriz diagonal, entonces también es sencillo calcular cualquier potencia de A, ya que $A = PDP^{-1}$ y así

$$A^n = PDP^{-1}PDP^{-1} \cdots PDP^{-1} = PD^nP^{-1}.$$

En este tema estudiaremos qué condiciones se tienen que verificar para que una matriz sea diagonalizable y describiremos un método para obtener tanto la matriz diagonal semejante a la matriz dada como la matriz de paso. Para ello introducimos primero los conceptos de valor propio y vector propio de una matriz cuadrada y estudiaremos algunas propiedades de los mismos, que nos permitirán conformar la técnica de *diagonalización* mencionada.

Es importante hacer notar que a lo largo de este tema identificaremos vectores de n componentes (i.e., vectores de $\mathbb{R}^n$) con matrices columna $n \times 1$.

4.2. Valores y vectores propios. Propiedades

Definición 7 *Decimos que un vector* <u>*no nulo*</u> $\mathbf{x}$ *es un* **vector propio** *de una matriz cuadrada* A *si existe un escalar* λ*, llamado* **valor propio***, tal que*

$$A\mathbf{x} = \lambda\,\mathbf{x}$$

Un vector propio ha de ser diferente del vector $\mathbf{0}$; los valores propios pueden, en cambio, ser cero. No toda matriz tiene valores propios reales; por ejemplo, la matriz

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

correspondiente a una rotación en el plano de ángulo θ.

Ejemplo 33 *El vector* $\mathbf{x} = (-1, 1)$ *es un vector propio de* $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ *con valor propio* $\lambda = -1$.

Ejemplo 34 *El vector* $\mathbf{x} = (4, 1, -2)$ *es un vector propio de* $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix}$, *ya que*

$$A\mathbf{x} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 4 \\ 1 \\ -2 \end{pmatrix} = 0\mathbf{x}$$

Los valores y vectores propios van en parejas. Si $\mathbf{x}$ es un vector propio de una matriz A entonces ha de existir un valor propio λ tal que $A\mathbf{x} = \lambda\mathbf{x}$, lo cual es equivalente a la ecuación $A\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$, o bien,

$$(A - \lambda I)\mathbf{x} = \mathbf{0}. \tag{4.1}$$

Éste es un sistema de ecuaciones lineales homogéneo para las componentes del vector $\mathbf{x}$. De aquí se sigue que $\mathbf{x}$ es un vector propio de A correspondiente al valor propio λ si y sólo si $(A - \lambda I)$ no tiene inversa. Equivalentemente, si y sólo si

$$\det(A - \lambda I) = 0. \tag{4.2}$$

La ecuación (4.2) es la *ecuación característica* de A. Si A es una matriz $n \times n$, entonces $\det(A - \lambda I) = 0$ es un polinomio en λ de grado n, y la ecuación característica de A tiene exactamente n soluciones (pudiendo estar algunas repetidas), que son los valores propios de A. Una vez se ha localizado un valor propio, los vectores propios correspondientes se obtienen resolviendo la ec. (4.1).

Así pues, para obtener los valores y vectores propios de una matriz A, primero se resuelve la ecuación característica, $\det(A - \lambda I) = 0$, para los valores propios y después, para cada valor propio, se resuelve $(A - \lambda I)\mathbf{x} = \mathbf{0}$ para los correspondientes vectores propios.

Ejemplo 35 *Obtener los valores y vectores propios de*

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

<u>Solución</u>.

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & -1 & 0 \\ 3 & -2-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)(\lambda^2 - 1) \Rightarrow \lambda = 1, 1, -1$$

$$\lambda = 1 \Rightarrow (A - \lambda I)\mathbf{x} = \mathbf{0} \Rightarrow \begin{pmatrix} 1 & -1 & 0 \\ 3 & -3 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$\Rightarrow x = y \Rightarrow \mathbf{x} = \begin{pmatrix} x \\ x \\ z \end{pmatrix} = x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \text{ con } x, z \neq 0.$$

$$\lambda = -1 \Rightarrow (A - \lambda I)\mathbf{x} = \mathbf{0} \Rightarrow \begin{pmatrix} 3 & -1 & 0 \\ 3 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Rightarrow$$

$$\Rightarrow \left\{ \begin{array}{l} 3x - y = 0 \\ z = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{l} y = 3x \\ z = 0 \end{array} \right. \Rightarrow \mathbf{x} = x \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \text{ con } x \neq 0. \qquad \square$$

Las soluciones (también llamadas 'raíces') de una ecuación característica pueden estar repetidas. Si $\lambda_1 = \lambda_2 = \cdots = \lambda_k$, el valor propio se dice que tiene *multiplicidad* k. Así, en este ejemplo, $\lambda = 1$ tiene multiplicidad 2, mientras que $\lambda = -1$ es un valor propio con multiplicidad 1.

Teorema 4.2.1 *Matrices semejantes tienen la misma ecuación característica (y, por consiguiente, los mismos valores propios).*

De este teorema se deduce que si dos matrices *no* tienen la misma ecuación característica, entonces las matrices *no pueden ser* semejantes.

Ejemplo 36 *Determinar si* $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ *es semejante a* $B = \begin{pmatrix} 1 & 4 \\ 3 & 2 \end{pmatrix}$.

Solución. No.

Los vectores propios $\mathbf{x}$ correspondientes al valor propio λ de una matriz A son soluciones no nulas de la ecuación matricial $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Esta ecuación matricial define el *núcleo* de $(A - \lambda I)$, un espacio vectorial conocido como el **subespacio propio** de A correspondiente al valor propio λ. Los vectores no nulos de un subespacio propio son los vectores propios. De esta manera, los vectores propios correspondientes a un valor propio particular se describen de forma más simple dando una base del correspondiente subespacio propio.

Ejemplo 37 *Obtener bases para los subespacios propios de la matriz*

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

Solución. Del ejemplo anterior, deducimos que una base del subespacio propio $S_{\lambda=1}$ correspondiente a $\lambda = 1$ es $\{(1,1,0), (0,0,1)\}$, mientras que para $S_{\lambda=-1}$ podría ser $\{(1,3,0)\}$. $\square$

Ejemplo 38 *Sea*

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

Calcular sus valores y vectores propios, así como una base de cada uno de sus subespacios propios.

Solución.

$$\det(A-\lambda I)=\begin{vmatrix} 3-\lambda & -2 & 0 \\ -2 & 3-\lambda & 0 \\ 0 & 0 & 5-\lambda \end{vmatrix}=(5-\lambda)\,(\lambda-1)\,(\lambda-5)\Rightarrow\lambda=1,5,5$$

$$\lambda=1\Rightarrow(A-\lambda I)\mathbf{x}=\mathbf{0}\Rightarrow\begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Rightarrow$$

$$\Rightarrow\left\{\begin{array}{l} 2x-2y=0 \\ 4z=0 \end{array}\right.\Rightarrow\left\{\begin{array}{l} x=y \\ z=0 \end{array}\right.\Rightarrow\mathbf{x}=x\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}\Rightarrow S_{\lambda=1}=\langle(1,1,0)\rangle$$

$$\lambda=5\Rightarrow(A-\lambda I)\mathbf{x}=\mathbf{0}\Rightarrow\begin{pmatrix} -2 & -2 & 0 \\ -2 & -2 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}\Rightarrow$$

$$\Rightarrow -2x-2y=0\Rightarrow x=-y\Rightarrow\mathbf{x}=x\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}+z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\Rightarrow S_{\lambda=5}=$$

$\langle(1,-1,0),(0,0,1)\rangle$ □

La ecuación característica de una matriz real puede tener raíces complejas, por lo que no son valores propios reales. Los vectores propios correspondientes a esas raíces complejas tienen componentes complejos, de modo que tales 'vectores' no son elementos de un espacio vectorial real. Por consiguiente, no hay vectores en un espacio vectorial real tales que satisfacen $A\mathbf{x}=\lambda\mathbf{x}$ cuando λ es complejo.

4.2.1. Propiedades de los valores y vectores propios

La **traza** de una matriz cuadrada A, denotada $\mathrm{Tr}(A)$, es la suma de los elementos de la diagonal principal de A.

Teorema 4.2.2 *La suma de los valores propios de una matriz es igual a la traza de la matriz.*

Este resultado proporciona un buen test respecto al cálculo de los valores propios. Si la suma de los valores propios calculados no es igual a la traza de la matriz, entonces hay un error.

El determinante de una matriz triangular es el producto de los elementos de la diagonal principal, de modo que se tiene el siguiente

Teorema 4.2.3 *Los valores propios de una matriz triangular son los elementos de la diagonal principal.*

Una vez se conocen los valores propios de una matriz, se puede determinar inmediatamente si la matriz es singular:

Teorema 4.2.4 *Una matriz es singular y sólo si tiene un valor propio cero.*

Veamos algunas otras propiedades:

- Si $\mathbf{x}$ es un vector propio de una matriz regular A, correspondiente al valor propio λ, entonces $\mathbf{x}$ también es un vector propio de A^{-1} correspondiente al valor propio $\frac{1}{\lambda}$.
- El producto de todos los valores propios de una matriz (contando su multiplicidad) es igual al determinante de la matriz.
- Si $\mathbf{x}$ es un vector propio de A correspondiente al valor propio λ, entonces λ^n y $\mathbf{x}$ son una pareja de valor propio y vector propio de A^n, para cualquier n entero positivo.
- Si λ es un valor propio de A, entonces λ también es un valor propio de A^T.

4.3. Diagonalización

Recordemos que, según vimos al principio del tema, una matriz es diagonalizable si es semejante a una matriz diagonal.

Si una matriz A es semejante a una matriz diagonal D, entonces $D = P^{-1}AP$ y la forma de D está determinada. Tanto A como D tienen los mismos valores propios, y los valores propios de una matriz diagonal (que es a la vez triangular superior e inferior) son los elementos de su diagonal principal. Por consiguiente, *la diagonal principal de D ha de tener los valores propios de A.*

Se puede, además, demostrar el importante resultado siguiente:

Teorema 4.3.1 *Una matriz A $n \times n$ es diagonalizable si y sólo si la matriz posee n vectores propios linealmente independientes.*

En el transcurso de la demostración de este teorema se pone de manifiesto que las columnas de la matriz de paso P son precisamente los vectores propios correspondientes a los valores propios de A.

Ejemplo 39 *Determinar si $A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$ es diagonalizable.*

<u>Solución</u>. Es fácil comprobar que los valores propios son $\lambda_1 = -1$ y $\lambda_2 = 5$, mientras que una base de los respectivos subespacios propios es $\mathbf{x}_1 = (-1, 1)$ y $\mathbf{x}_2 = (1, 2)$. Estos dos vectores son linealmente independientes, de modo que A es diagonalizable. Podemos elegir

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

y entonces

$$D = P^{-1}AP = \begin{pmatrix} -1 & 0 \\ 0 & 5 \end{pmatrix}$$

□

En general, ni P ni D son únicas. No obstante, una vez se ha seleccionado P, entonces D está determinada completamente. El elemento de D situado en la fila j, columna j ha de ser el valor propio correspondiente al vector propio en la columna j de P.

Ejemplo 40 *Determinar si* $A = \begin{pmatrix} 2 & -1 & 0 \\ 3 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *es diagonalizable.*

Solución. Sabemos que $\{\mathbf{x}_1 = (1,1,0), \mathbf{x}_2 = (0,0,1)\}$ es una base del subespacio propio correspondiente al valor propio $\lambda = 1$ de multiplicidad 2 y $\mathbf{x}_3 = (1,3,0)$ es una base correspondiente a $S_{\lambda=-1}$. Estos tres vectores son linealmente independientes, por lo que A es diagonalizable. Si elegimos

$$P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix} \quad \Rightarrow \quad P^{-1}AP = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

□

El proceso de determinar si un conjunto dado de vectores propios es linealmente independiente se simplifica con estos dos resultados:

Teorema 4.3.2 *Vectores propios de una matriz correspondientes a valores propios diferentes son linealmente independientes.*

Por consiguiente, cualquier matriz real $n \times n$ que tenga n valores propios de multiplicidad 1 es diagonalizable.

Teorema 4.3.3 *Si λ es un valor propio de multiplicidad k de una matriz A $n \times n$, entonces el número de vectores propios linealmente independientes de A correspondientes a λ es $n - \text{rango}(A - \lambda I)$.*

Ejemplo 41 *Sea*

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

Diagonalizar la matriz A, es decir, calcular una matriz diagonal D y la correspondiente matriz de paso P tal que $D = P^{-1}AP$.

Solución. La matriz de paso estará formada por los vectores propios encontrados,

$$P = \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow$$

$$D = P^{-1}AP = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

□

El teorema 4.3.1 se puede formular de varias maneras, todas ellas equivalentes. Así, la matriz A $n \times n$ es diagonalizable si y sólo si es posible construir una base de $\mathbb{R}^n$ formada por vectores propios de A.

También se puede demostrar que A es diagonalizable si y sólo si la multiplicidad de cada valor propio de A coincide con la dimensión de su subespacio propio correspondiente.

4.4. Ortogonalidad.

Ya se vio en el tema de los espacios vectoriales la noción de vectores ortogonales. Recordemos que dos vectores **u**, **v** son ortogonales si y sólo si su producto escalar es cero, esto es, $\mathbf{u}.\mathbf{v} = 0$. Esta noción se puede extender a un conjunto arbitrario de vectores. Así, un conjunto de vectores se llama *ortogonal* si cada vector en el conjunto es ortogonal a cualquier otro vector del conjunto.

Ejemplo 42 *Los vectores* $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ *de* $\mathbb{R}^3$ *siguientes*

$$\mathbf{x} = (1, 1, 1), \quad \mathbf{y} = (1, 1, -2), \quad \mathbf{z} = (1, -1, 0)$$

forman un conjunto ortogonal de vectores. Por contra, el conjunto $\{\mathbf{a}, \mathbf{b}, \mathbf{c}\}$

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

no es ortogonal, ya que $\mathbf{a}.\mathbf{c} \neq 0$.

Un conjunto ortogonal de vectores *unitarios* se llama conjunto *ortonormal.* Por tanto, el conjunto de vectores $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ es ortonormal si y sólo si $\mathbf{v}_i.\mathbf{v}_j = 0$, $i \neq j$ y $\mathbf{v}_i.\mathbf{v}_i = 1$.

Ejemplo 43 *El conjunto de vectores* $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ *de* $\mathbb{R}^3$ *definido por*

$$\mathbf{u} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), \quad \mathbf{v} = (\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0), \quad \mathbf{w} = (0, 0, 1)$$

es un conjunto ortonormal de vectores.

Cualquier conjunto ortogonal de vectores no nulos se puede transformar en un conjunto ortonormal dividiendo cada vector por su norma.

Teorema 4.4.1 *Un conjunto ortonormal de un número finito de vectores es linealmente independiente.*

Si el conjunto ortonormal de vectores es una base, se llama *base ortonormal.*

Se dice que una matriz cuadrada P es **ortogonal** si su inversa coincide con su traspuesta, es decir,

$$P^{-1} = P^T$$

o, equivalentemente,

$$PP^T = P^T P = I.$$

4.4.1. Diagonalización de matrices reales simétricas.

Vimos antes que no todas las matrices cuadradas $n \times n$ cuyos elementos son números reales son diagonalizables, obteniéndose una condición necesaria y suficiente para que lo sean. Cuando se considera una matriz cuadrada real que además es simétrica, esa simetría da lugar a varias simplificaciones. En particular,

- todas las raíces de la ecuación característica son reales

- la matriz es diagonalizable
- es posible elegir n vectores propios formando un conjunto ortonormal, es decir, existe siempre una base ortonormal de $\mathbb{R}^n$ formada por vectores propios de la matriz.

Teorema 4.4.2 *Todas las soluciones de la ecuación característica de una matriz real simétrica son números reales, por lo que todos sus valores propios son reales.*

Teorema 4.4.3 *Para una matriz real simétrica los vectores propios correspondientes a valores propios distintos son ortogonales.*

Se dice que una matriz A es **diagonalizable ortogonalmente** si existe una matriz ortogonal P tal que la matriz $P^{-1}AP$ es diagonal.

Teorema 4.4.4 *Toda matriz A real y simétrica $n \times n$ es ortogonalmente diagonalizable, es decir, existe una matriz ortogonal P tal que la matriz $D = P^{-1}AP$ es diagonal y real o, lo que es lo mismo, existe una base ortonormal de $\mathbb{R}^n$ formada por vectores propios de A.*

La forma de obtener esta matriz ortogonal es encontrando un conjunto ortonormal de vectores propios de A.

Ejemplo 44 *La matriz real simétrica*

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

es ortogonalmente diagonalizable.

<u>Solución</u>. En efecto, sus valores propios son $\lambda_1 = 5$ (con multiplicidad 1) y $\lambda_2 = 3$ (con multiplicidad 2), mientras que sus subespacios propios son

$$\begin{aligned} S_{\lambda_5} &= \langle (1,1,0) \rangle \\ S_{\lambda_3} &= \langle (1,-1,0), (0,0,1) \rangle \end{aligned}$$

con lo que

$$\{\mathbf{u}_1 = (1,1,0),\ \mathbf{u}_2 = (1,-1,0),\ \mathbf{u}_3 = (0,0,1)\}$$

forman una base de $\mathbb{R}^3$ constituida por vectores propios, que además son ortogonales. A partir de ésta formamos una base ortonormal dividiendo cada vector por su norma:

$$\{\mathbf{v}_1 = \frac{1}{\sqrt{2}}(1,1,0),\ \mathbf{v}_2 = \frac{1}{\sqrt{2}}(1,-1,0),\ \mathbf{v}_3 = (0,0,1)\}$$

y la correspondiente matriz de paso ortogonal

$$P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

de manera que

$$D = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{pmatrix} = P^{-1}AP = P^T AP.$$

□

Cuando todos los valores propios tienen multiplicidad 1 los subespacios propios tienen dimensión 1 y al localizar un vector en cada uno de ellos, éstos son ortogonales, de modo que al dividirlos por sus normas se obtienen vectores propios ortonormales, los cuales constituyen las columnas de P. *Cuando la dimensión de un subespacio propio es mayor o igual que 2 se han de tomar vectores en ese subespacio de manera que sean unitarios y ortogonales*, lo cual puede hacerse de infinitas maneras.

Ejemplo 45 *Diagonaliza ortogonalmente la matriz simétrica*

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

Solución. Ya sabemos que los subespacios propios de esta matriz son:

$S_{\lambda=1} = \langle(1,1,0)\rangle$, $S_{\lambda=5} = \langle(1,-1,0),(0,0,1)\rangle$. Vamos a comprobar si son ortogonales:

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

Por tanto buscaremos los vectores unitarios:

$$\left.\begin{matrix} (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0) \\ (\frac{1}{\sqrt{2}}, \frac{-1}{\sqrt{2}}, 0) \\ (0,0,1) \end{matrix}\right\} \Rightarrow P = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \Rightarrow P^T = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$D = P^T AP = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

□

Ejercicio 16 *Diagonaliza ortogonalmente la matriz*

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

4.5. Introducción a las cadenas de Markov

Las cadenas de Markov constituyen una poderosa herramienta que posibilita el estudio del comportamiento de un sistema económico en una situación dinámica. Más específicamente, a partir del comportamiento pasado del problema planteado podemos describir su comportamiento futuro, siempre que éste se ajuste a ciertas hipótesis de funcionamiento implícitas en toda modelización económica. Por otra parte, su aplicación al estudio del comportamiento de diversos sistemas sociales (movimientos de población, preferencias políticas, fidelidad a una marca determinada, etc.) posibilitan el diseño de estrategias, marketing, análisis de precios, etc., para "influir" de alguna forma en el sistema que interesa.

Desde un punto de vista matemático, las cadenas de Markov constituyen una sencilla aplicación de la teoría de los valores y vectores propios y combinan la teoría de la probabilidad con el cálculo matricial. Una comprensión elemental de las cadenas de Markov sólo requiere nociones elementales de la teoría de la probabilidad; en particular, sólo hace falta saber que una probabilidad cuantifica la posibilidad de que ocurra un cierto suceso, que es un número comprendido entre 0 y 1 y que, si el conjunto de todos los sucesos posibles está limitado a un número finito y además son mutuamente excluyentes, entonces la suma de todas las probabilidades es 1. Para demostrar los teoremas relevantes acerca de las cadenas de Markov se necesita mucho más aparato probabilístico, de modo que aquí nos vamos a limitar a un tratamiento elemental.

Definición 8 *Una cadena de Markov finita es un conjunto de objetos (por ejemplo, personas), un conjunto de periodos temporales consecutivos (por ejemplo, intervalos de cinco años) y un conjunto finito de diferentes estados (por ejemplo, empleados y desempleados) tales que*

(i) durante cualquier periodo temporal dado cada objeto sólo está en un estado (aunque objetos diferentes pueden estar en estados diferentes), y

(ii) la probabilidad de que un objeto pase de un estado a otro (o permanezca en el mismo estado) al cabo de un periodo temporal sólo depende de los estados inicial y final.

Denotamos los estados como estado 1, estado 2, estado 3, ..., estado N y designamos por p_{ij} la probabilidad de pasar en un periodo temporal desde el estado j hasta el estado i $(i, j = 1, 2, \ldots, N)$. La matriz $P = (p_{ij})$ se llama *matriz de transición*. Así pues, una matriz de transición para una cadena de Markov de N estados es una matriz $N \times N$ con elementos comprendidos entre 0 y 1; la suma de los elementos en cada columna es 1.

Ejemplo 46 *Construir una matriz de transición para la siguiente cadena de Markov. Un gestor de control de tráfico clasifica cada día como despejado o nublado. Los datos recopilados en el transcurso del tiempo muestran que la probabilidad de que se tenga un día despejado tras un día nublado es 0.6, mientras que la probabilidad de que se tenga un día despejado tras otro día despejado es 0.9.*

Solución. Aunque podemos imaginar muchas otras clasificaciones, tales como lluvioso, muy nublado, parcialmente soleado, etc., este gestor en particular ha optado por solamente dos, de manera que sólo tenemos dos estados: despejado y nublado, y cada día ha de pertenecer a una y sólo una de estas dos categorías. Fijamos arbitrariamente "despejadoçomo estado 1 y "nubladoçomo estado 2. La unidad natural de tiempo es 1 día. De los datos del problema sabemos que $p_{12} = 0{,}6$, de manera que $p_{22} = 0{,}4$, porque después de un día nublado el siguiente ha de ser o bien despejado o bien nublado y la probabilidad de que ocurra uno u otro de estos dos sucesos es 1. De forma similar, sabemos que $p_{11} = 0{,}9$, de manera que $p_{21} = 0{,}1$. Así pues, la matriz de transición es

$$P = \begin{pmatrix} 0{,}9 & 0{,}6 \\ 0{,}1 & 0{,}4 \end{pmatrix}.$$

□

Las potencias enteras de una matriz de transición tienen las mismas propiedades que una matriz de transición: todos los elementos están comprendidos entre 0 y 1, y la suma de todos los elementos de una columna dada es igual a 1. Además se tiene el siguiente

Teorema 4.5.1 . *Si P es una matriz de transición de una cadena de Markov finita y si $p_{ij}^{(k)}$ denota el elemento ij de P^k (la potencia k-ésima de P), entonces $p_{ij}^{(k)}$ es la probabilidad de moverse al estado i desde el estado j tras k periodos temporales.*

Para la matriz de transición creada en el ejemplo 46, la segunda potencia resulta ser

$$P^2 = \begin{pmatrix} 0{,}87 & 0{,}78 \\ 0{,}13 & 0{,}22 \end{pmatrix}$$

Por consiguiente, $p_{12}^{(2)} = 0{,}78$ es la probabilidad de que se tenga un día despejado dos días después de un día nublado, mientras que $p_{22}^{(2)} = 0{,}22$ es la probabilidad de que se tenga un día nublado. Calculando la décima potencia de esta misma matriz de transición y redondeando a cuatro cifras decimales se tiene

$$P^{10} = \begin{pmatrix} 0{,}8571 & 0{,}8571 \\ 0{,}1429 & 0{,}1429 \end{pmatrix}$$

Como $p_{11}^{(10)} = p_{12}^{(10)} = 0{,}8571$ se deduce que la probabilidad de tener un día despejado 10 días después de un día despejado es igual a la probabilidad de tener un día despejado 10 días después de un día nublado.

Un objeto en una cadena de Markov ha de estar en un solo estado en un tiempo dado, pero ese estado no siempre es conocido con exactitud. A menudo se proporcionan probabilidades para describir la posibilidad de que un objeto esté en uno cualquiera de los posibles estados en un instante dado. Un *vector distribución* $\mathbf{d}$ para una cadena de Markov de N estados en un instante dado es un vector cuyas N componentes (una por cada estado) proporcionan la probabilidad de que un objeto del sistema esté en cada uno de los respectivos estados en ese instante. Dicho de otra forma, la i-ésima componente es la probabilidad de que un objeto esté en el i-ésimo estado en ese instante.

Ejemplo 47 *Obtener el vector distribución para la cadena de Markov descrita en el ejemplo 46 si se sabe que el día de hoy es nublado.*

Solución. Los objetos en el sistema son días, que son clasificados como despejados (estado 1) o nublados (estado 2). Se nos dice que el día de hoy es nublado, de manera que la probabilidad de que el día sea nublado es 1 y la probabilidad de que el día sea despejado es 0. Por consiguiente,

$$\mathbf{d} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

□

Periodos temporales diferentes pueden tener vectores distribución diferentes, de manera que denotaremos por $\mathbf{d}^{(k)}$ un vector distribución *tras k* periodos temporales. En particular, $\mathbf{d}^{(1)}$ es un vector distribución tras un periodo temporal, $\mathbf{d}^{(2)}$ es un vector distribución tras 2 periodos temporales y $\mathbf{d}^{(10)}$ es un vector distribución tras 10 periodos temporales. Un vector distribución inicial para una cadena de Markov se designa por $\mathbf{d}^{(0)}$. Los vectores distribución correspondientes a periodos temporales diferentes están relacionados.

Teorema 4.5.2 . *Si P es una matriz de transición para una cadena de Markov, entonces*

$$\mathbf{d}^{(k)} = P^k\mathbf{d}^{(0)} = P\mathbf{d}^{(k-1)}$$

donde P^k denota la potencia k-ésima de P.

Para el vector distribución y la matriz de transición creados en los ejemplos 46 y 47 calculamos

$$\begin{aligned}
\mathbf{d}^{(1)} &= P\mathbf{d}^{(0)} = \begin{pmatrix} 0{,}9 & 0{,}6 \\ 0{,}1 & 0{,}4 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}6 \\ 0{,}4 \end{pmatrix} \\
\mathbf{d}^{(2)} &= P^2\mathbf{d}^{(0)} = \begin{pmatrix} 0{,}87 & 0{,}78 \\ 0{,}13 & 0{,}22 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}78 \\ 0{,}22 \end{pmatrix} \\
\mathbf{d}^{(10)} &= P^{10}\mathbf{d}^{(0)} = \begin{pmatrix} 0{,}8571 & 0{,}8571 \\ 0{,}1429 & 0{,}1429 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0{,}8571 \\ 0{,}1429 \end{pmatrix}
\end{aligned}$$

Si empezamos con un día nublado, entonces las probabilidades de tener un día nublado tras un periodo temporal, 2 periodos temporales y 10 periodos temporales son, respectivamente, 0.4, 0.22 y 0.1429.

La potencia 10 de la matriz de transición creada en el ejemplo 46 es

$$P^{10} = \begin{pmatrix} 0{,}8571 & 0{,}8571 \\ 0{,}1429 & 0{,}1429 \end{pmatrix}$$

Si seguimos calculando potencias más altas de P veremos que cada una de ellas es idéntica a P^{10} cuando se redondea cada elemento a cuatro cifras decimales. Podemos escribir entonces que

$$\lim_{n\to\infty} P^n = \begin{pmatrix} 0{,}8571 & 0{,}8571 \\ 0{,}1429 & 0{,}1429 \end{pmatrix}$$

No todas las matrices de transición tienen potencias que convergen a una matriz límite L. Una matriz de transición para una cadena de Markov finita es *regular* is ella o una de sus potencias contiene solamente elementos positivos. Las potencias de una matriz regular siempre convergen a una matriz límite L.

La matriz de transición creada en el ejemplo 46 es regular porque todos sus elementos son positivos. Por contra, la matriz de transición

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

no es regular porque cada una de sus potencias es o bien ella misma o bien la matriz identidad 2×2, y ambas contienen ceros.

Por definición, alguna potencia de una matriz regular P, digamos la m-ésima, contiene solamente elementos positivos. Como los elementos de P son no negativos, de la multiplicación de matrices se sigue que cualquier potencia de P mayor que m también ha de tener todos sus elementos positivos. Más aún, si $L = \lim_{k\to\infty} P^k$, entonces también es cierto que $L = \lim_{k\to\infty} P^{k-1}$. Por consiguiente,

$$L = \lim_{k\to\infty} P^k = \lim_{k\to\infty} (PP^{k-1}) = P\left(\lim_{k\to\infty} P^{k-1}\right) = PL \tag{4.3}$$

Denotemos las columnas de L como $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$, respectivamente, de manera que $L = (\mathbf{x}_1\,\mathbf{x}_2 \cdots \mathbf{x}_N)$. Entonces (4.3) se puede escribir como

$$(\mathbf{x}_1\,\mathbf{x}_2 \cdots \mathbf{x}_N) = P(\mathbf{x}_1\,\mathbf{x}_2 \cdots \mathbf{x}_N)$$

de donde $\mathbf{x}_j = P\mathbf{x}_j$ $(j = 1, 2, \ldots, N)$, o bien $P\mathbf{x}_j = 1\mathbf{x}_j$. De esta manera, cada columna de L es un vector propio de P correspondiente al valor propio 1. Hemos probado así una parte del importante resultado siguiente.

Teorema 4.5.3 *Si una matriz de transición P $N \times N$ es regular, entonces las potencias enteras sucesivas de P convergen a una matriz límite L cuyas columnas son vectores propios de P asociados al valor propio $\lambda = 1$. Las componentes de este vector propio son positivas y suman 1.*

Más todavía: si P es regular, entonces su valor propio $\lambda = 1$ tiene multiplicidad 1 y hay solamente un vector propio linealmente independiente asociado a él. Este vector propio está dado en térmios de una constante arbitraria, que se determina de forma única exigiendo que la suma de las componentes sea 1. Así, cada columna de L es *el mismo vector propio*.

Definimos el *vector distribución del estado límite* para una cadena de Markov de N estados como

$$\mathbf{d}^{(\infty)} = \lim_{n\to\infty} \mathbf{d}^{(n)}$$

cuando el límite exista. Esto siempre ocurre cuando P es regular. Entonces las componentes de $\mathbf{d}^{(\infty)}$ son las probabilidades límite de que un objeto en el sistema esté en cada uno de los estados tras un gran número de periodos temporales. Más aún,

$$\mathbf{d}^{(\infty)} = \lim_{n\to\infty} \mathbf{d}^{(n)} = \lim_{n\to\infty} (P^n \mathbf{d}^{(0)}) = \left(\lim_{n\to\infty} P^n\right) \mathbf{d}^{(0)} = L\mathbf{d}^{(0)}$$

Cada columna de L es idéntica a las otras, de manera que cada fila de L contiene un solo número repetido N veces. Combinando esto con el hecho de que $\mathbf{d}^{(0)}$ tiene componentes que suman 1, se sigue que el producto $L\mathbf{d}^{(0)}$ es igual a cada una de las columnas idénticas de L.

Ejemplo 48 *Encontrar el vector distribución del estado límite para la cadena de Markov del ejemplo 46.*

Solución. La matriz de transición es

$$P = \begin{pmatrix} 0{,}9 & 0{,}6 \\ 0{,}1 & 0{,}4 \end{pmatrix}$$

que es regular. Los vectores propios de esta matriz tienen la forma

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}.$$

Los vectores propios correspondientes a $\lambda = 1$ satisfacen la ecuación matricial $(P - 1\mathbf{I})\mathbf{x} = \mathbf{0}$ o, equivalentemente, el conjunto de ecuaciones

$$\begin{aligned} -0{,}1x + 0{,}6y &= 0 \\ 0{,}1x - 0{,}6y &= 0 \end{aligned}$$

Resolviendo el sistema (por ejemplo mediante eliminación gaussiana), se tiene que $x = 6y$, con y arbitraria. Así,

$$\mathbf{x} = \begin{pmatrix} 6y \\ y \end{pmatrix}.$$

Si elegimos y de manera que la suma de las componentes de $\mathbf{x}$ sea 1, se tiene $7y = 1$ o $y = 1/7$. El vector propio resultante es el vector distribución del estado límite, esto es,

$$\mathbf{d}^{(\infty)} = \begin{pmatrix} 6/7 \\ 1/7 \end{pmatrix}$$

y

$$L = \begin{pmatrix} 6/7 & 6/7 \\ 1/7 & 1/7 \end{pmatrix}.$$

A largo plazo seis de cada siete días estará despejado y uno de cada siete estará nublado. □

4.6. Aplicación económica de las cadenas de Markov

Vamos a estudiar ahora desde esta perspectiva el problema de la fidelidad a una marca determinada en el consumo de un producto (por ejemplo, un detergente).

Supongamos que este producto es ofertado por cuatro empresas que abastecen, en un alto porcentaje, su demanda. Las cuatro empresas, que denominaremos X, Y, W y Z, lo comercializan bajo cuatro marcas diferentes, con nombres respectivos A, B, C y D.

Imaginemos que se realiza un análisis sobre 2000 consumidores de ese producto, obteniéndose los siguientes resultados:

- Consumen marca A: 475 personas.
- Consumen marca B: 550 personas.

- Consumen marca C: 485 personas.
- Consumen marca D: 490 personas.

Tras sucesivos muestreos repetidos en diferentes periodos, se observa la siguiente evolución de la demanda de las citadas marcas:

	Aumentos				**Disminuciones**				
	A	B	C	D	A	B	C	D	Total
A	0	10	5	10	0	5	20	30	445
B	5	0	5	5	10	0	5	25	525
C	20	5	0	15	5	5	0	10	505
D	30	25	10	0	10	5	15	0	525

en donde indicamos por filas los aumentos o disminuciones que ha sufrido cada marca y de dónde o hacia dónde han ido las preferencias de los consumidores. Por ejemplo, si leemos la fila de la marca C sabemos que ha sufrido un aumento de 40 consumidores (20 de la marca A, 5 de la B y 15 de la D), mientras que sufre una disminución de 20 usuarios de su marca (5 se han cambiado a la marca A, 5 a la B y 10 a la D).

Si aceptamos que la preferencia de los consumidores se estabiliza en el tiempo, los datos anteriores se pueden resumir en la siguiente tabla:

	A	B	C	D
A	420	10	5	10
B	5	510	5	5
C	20	5	465	15
D	30	25	10	460
Total	475	550	485	490

donde hemos indicado en la diagonal principal los consumidores que se mantienen fieles a su marca (los que tenía al principio menos las disminuciones sufridas).

A partir de estos datos podemos calcular la matriz de transición de un periodo a otro, que nos indicará las probabilidades de retención o de transición de clientes para cada marca, esto es, la fracción de clientes que seguramente siguen consumiendo la misma marca o que se cambian a otras. Esta matriz se puede obtener dividiendo cada elemento de la tabla anterior por la suma total de la columna a la que pertenece. En nuestro caso resulta (redondeando a cuatro cifras decimales)

$$P = \begin{pmatrix} 0{,}8842 & 0{,}0182 & 0{,}0103 & 0{,}0204 \\ 0{,}0105 & 0{,}9273 & 0{,}0103 & 0{,}0102 \\ 0{,}0421 & 0{,}0091 & 0{,}9588 & 0{,}0306 \\ 0{,}0632 & 0{,}0454 & 0{,}0206 & 0{,}9388 \end{pmatrix}$$

y cada elemento p_{ij} representa la probabilidad de que un consumidor de la marca j se pase a la marca i, para $i, j = 1, 2, 3, 4$ (A,B,C,D) y los elementos de la diagonal principal indican la probabilidad correspondiente a la fidelidad de los consumidores a su propia marca.

Con ayuda de esta matriz podemos estudiar el comportamiento de los consumidores, y por tanto definir nuevas estrategias empresariales en su competencia con las otras marcas.

Si denominamos $\mathbf{b}^{(k)}$ al vector de 4 componentes que nos proporciona las distintas situaciones en el periodo k-ésimo (la cantidad de clientes que posee cada empresa), medidos éstos en unidades convenientes (meses, años, etc.), tendremos

$$\begin{aligned} \mathbf{b}^{(1)} &= P\mathbf{b}^{(0)} \\ \mathbf{b}^{(2)} &= P^2\mathbf{b}^{(0)} \\ \mathbf{b}^{(3)} &= P^3\mathbf{b}^{(0)} \end{aligned}$$

y en general

$$\mathbf{b}^{(n)} = P^n\mathbf{b}^{(0)}$$

Si P es diagonalizable, entonces

$$P^n = QD^nQ^{-1}$$

donde D es la matriz diagonal conteniendo los valores propios de P en su diagonal principal y Q es la matriz de cambio de base formada por los vectores propios asociados a esos valores propios. En nuestro caso

$$Q = \begin{pmatrix} -0{,}208258 & 0{,}122896 & -0{,}12674 & 0{,}769566 \\ -0{,}219646 & -0{,}00427674 & 0{,}247561 & -0{,}00421248 \\ -0{,}722763 & -0{,}758539 & 0{,}616107 & -0{,}143024 \\ -0{,}621287 & 0{,}639919 & -0{,}736928 & -0{,}622329 \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0{,}926216 & 0 & 0 \\ 0 & 0 & 0{,}917195 & 0 \\ 0 & 0 & 0 & 0{,}865689 \end{pmatrix}$$

con lo que las potencias sucesivas de P vienen dadas por

$$\begin{aligned} P^2 &= \begin{pmatrix} 0{,}783724 & 0{,}0339892 & 0{,}0195906 & 0{,}03769 \\ 0{,}020099 & 0{,}860633 & 0{,}0197451 & 0{,}0195636 \\ 0{,}0796198 & 0{,}019319 & 0{,}920455 & 0{,}0590182 \\ 0{,}116558 & 0{,}0860586 & 0{,}0402091 & 0{,}883728 \end{pmatrix} \\ P^3 &= \begin{pmatrix} 0{,}696532 & 0{,}0476713 & 0{,}0279823 & 0{,}0523175 \\ 0{,}0288759 & 0{,}799499 & 0{,}0284062 & 0{,}028159 \\ 0{,}113084 & 0{,}0304191 & 0{,}884767 & 0{,}0853935 \\ 0{,}161508 & 0{,}122411 & 0{,}0588443 & 0{,}83413 \end{pmatrix} \\ P^4 &= \begin{pmatrix} 0{,}620859 & 0{,}0595124 & 0{,}0355725 & 0{,}0646674 \\ 0{,}0369023 & 0{,}743438 & 0{,}0363482 & 0{,}0360488 \\ 0{,}142954 & 0{,}042194 & 0{,}851552 & 0{,}109859 \\ 0{,}199285 & 0{,}154856 & 0{,}0765273 & 0{,}789425 \end{pmatrix} \end{aligned}$$

y así

$$\begin{aligned}
\mathbf{b}^{(1)} &\simeq (445, 525, 505, 525)\\
\mathbf{b}^{(2)} &\simeq (419, 502, 524, 555)\\
\mathbf{b}^{(3)} &\simeq (396, 481, 541, 581)\\
\mathbf{b}^{(4)} &\simeq (377, 462, 558, 604)
\end{aligned}$$

dado que el vector distribución inicial está dado por $\mathbf{b}^{(0)} = (475, 550, 485, 490)$ (el número de consumidores de cada marca en el instante inicial).

¿Cuál es la tendencia a largo plazo del sistema? Nos podemos plantear si el sistema llegará a estabilizarse, esto es, si llegará un momento en el que las proporciones de mercado de cada una de las marcas se mantienen constantes.

Para resolver este problema podemos estudiar directamente las potencias de la matriz P (con ayuda de un ordenador) o bien aplicar la teoría precedente relativa a la matriz límite L en una cadena de Markov y cómo calcularla.

Procediendo de la primera manera, observamos que

$$P^{100} = \begin{pmatrix} 0{,}117538 & 0{,}11765 & 0{,}117473 & 0{,}117552\\ 0{,}123936 & 0{,}124107 & 0{,}123937 & 0{,}123935\\ 0{,}407858 & 0{,}407042 & 0{,}408259 & 0{,}407772\\ 0{,}350668 & 0{,}3512 & 0{,}350331 & 0{,}350742 \end{pmatrix}$$

mientras que

$$P^{101} = \begin{pmatrix} 0{,}117537 & 0{,}117642 & 0{,}117477 & 0{,}11755\\ 0{,}123938 & 0{,}124095 & 0{,}123938 & 0{,}123937\\ 0{,}407861 & 0{,}407101 & 0{,}408232 & 0{,}407781\\ 0{,}350664 & 0{,}351162 & 0{,}350352 & 0{,}350732 \end{pmatrix}$$

Por ello podemos considerar que el mercado se estabiliza sobre el valor $\mathbf{b}^{(100)} \simeq (235, 248, 815, 701)$.

Si procedemos de la segunda forma, hemos de calcular el vector propio de P asociado al valor propio $\lambda = 1$. Éste resulta ser

$$\mathbf{x} = y(-0{,}208258, -0{,}219646, -0{,}722763, -0{,}621287)$$

(primera columna de la matriz Q). Elegimos y de manera que la suma de las componentes de $\mathbf{x}$ sea 1: $y = -1{,}77195$. Así pues,

$$\mathbf{d}^{(\infty)} = (0{,}11753, 0{,}123957, 0{,}40789, 0{,}350623)$$

la matriz límite viene dada por

$$L = \begin{pmatrix} 0{,}11753 & 0{,}11753 & 0{,}11753 & 0{,}11753\\ 0{,}123957 & 0{,}123957 & 0{,}123957 & 0{,}123957\\ 0{,}40789 & 0{,}40789 & 0{,}40789 & 0{,}40789\\ 0{,}350623 & 0{,}350623 & 0{,}350623 & 0{,}350623 \end{pmatrix}$$

y el estado final es

$$\mathbf{b}^{(\infty)} = (235, 248, 816, 701)$$

Observemos la similitud de ambos resultados y la mayor sencillez del segundo procedimiento.

Desde un punto de vista económico, si las hipótesis del modelo propuesto permanecen válidas, se puede concluir que las empresas A y B disminuiránn notablemente su clientela (casi en un 50 %), mientras que las empresas C y D aumentarán sus ventas en una proporción similar. Es preciso remarcar que este estado de equilibrio (que como se observa es independiente de los valores iniciales) se produce si no existe ninguna perturbación en el sistema estudiado (promociones, descuentos, publicidad, etc.).

4.7. Problemas

1. Dada la matriz

$$\begin{pmatrix} 1 & 4 & -12 \\ 0 & 3 & -9 \\ 0 & 1 & -3 \end{pmatrix}$$

determinar sus valores propios así como una base del subespacio propio asociado a cada uno de ellos.

2. Calcular los valores propios, los subespacios propios y, si es posible, construir una base de $\mathbb{R}^3$ formada por vectores propios de cada una de las matrices siguientes:

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 5 \\ 0 & 0 & 10 \end{pmatrix}, \quad B = \begin{pmatrix} 5 & 1 & 2 \\ 0 & 3 & 1 \\ 0 & 0 & 5 \end{pmatrix}$$

¿Son diagonalizables?

3. Sea la matriz $A = \begin{pmatrix} 2 & -2 & 6 \\ 0 & a & 4-a \\ 0 & a & -a \end{pmatrix}$, $a \geq 0$.

   a) Estudia su diagonalización en función de los valores de a.

   b) Diagonalizar A (si es posible) cuando $a = 4$, hallando explícitamente la matriz de cambio de base.

4. Estudiar en función del parámetro real a la diagonalizabilidad de las matrices

$$A = \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}, \quad B = \begin{pmatrix} a & -a & a \\ a & -a & a \\ 2 & -1 & 0 \end{pmatrix}$$

5. Estudiar para qué valores de los parámeros reales a y b las matrices siguientes son diagonalizables:

$$\begin{pmatrix} 5 & 0 & 0 \\ 0 & -1 & b \\ 3 & 0 & a \end{pmatrix}, \qquad \begin{pmatrix} a & b & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

6. Dada la matriz

$$\mathbf{A} = \begin{pmatrix} a & b & 0 \\ c & a & b \\ 0 & c & a \end{pmatrix}$$

¿Qué relaciones deben verificarse entre los parámetros a, b y c para que $\mathbf{A}$ tenga una valor propio de multiplicidad 3?

7. Dada la matriz

$$M = \begin{pmatrix} -7 & -6 \\ 12 & 10 \end{pmatrix}$$

calcular su potencia de exponente 275.

8. Sabemos que la matriz

$$M = \begin{pmatrix} a & 1 & p \\ b & 2 & q \\ c & -1 & r \end{pmatrix}$$

admite como vectores propios $(1, 1, 0)$, $(-1, 0, 2)$ y $(0, 1, -1)$. Hallar los elementos de dicha matriz así como sus valores propios.

9. Dada la matriz

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

   a) Obtener $P(A)$, donde $P(x) = 3 + 2x - 5x^2$.

   b) Calcular su inversa.

10. Diagonalizar las matrices $A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix}$, $B = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

    a) ¿Es posible diagonalizarlas ortogonalmente? Si es así, determinar la matriz de cambio de base ortogonal.

    b) Calcular $\varphi(A) = A^7 + 4A^5 - 2A + 3A^2$

11. Dada la matriz $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1+a & 1 \\ 0 & 0 & a \end{pmatrix}$

    a) Estudiar su diagonalización en función de los valores de a

    b) Diagonalizarla para $a = 1$, si es posible, calculando explícitamente la matriz de paso y su inversa. Comprobar la diagonalización.

4.8. Aplicaciones a la economía. Cadenas de Markov

1. Supongamos un mercado duopolista, en el que dos empresas A y B fabrican la totalidad de un cierto producto. Supongamos también que el producto se adquiere mensualmente, y que por medio de estudios de mercado se ha llegado a las siguientes conclusiones sobre la intención de compra de los consumidores.

   - El 50 % de los consumidores que compran durante un mes el producto fabricado por la empresa A volverá a hacerlo así al mes siguiente, y el resto cambiará al fabricado por la empresa B.
   - El 25 % de los consumidores que compran durante un mes el producto fabricado por la empresa B volverá a proceder así al mes siguiente y el resto cambiará al fabricado por la empresa A.

   Sean 50 y 100 las cuotas de mercado de las empresas A y B, respectivamente, durante el mes de octubre de 2006. ¿Cuáles serán las cuotas de mercado de cada una de las empresas al cabo de diez meses, es decir, en el mes de agosto de 2007?

2. Dos marcas de tabaco controlan el mercado repartiéndoselo al 60 % y 40 % respectivamente. Este año en el mercado se mueven 500 millones de euros. Si los consumidores de la marca A son cada año fieles en un 30 % y los de la marca B son fieles en un 40 %. ¿Cómo se repartirán el mercado al cabo de 5 años las dos marcas suponiendo que su volumen se mantiene constante?

3. Consideremos la siguiente cadena de Markov. Cada cuatro años, los votantes de un pueblo de Nueva Inglaterra (EE. UU.) eligen un nuevo alcalde, debido a que una ordenanza municipal prohíbe a los alcaldes sucederse a sí mismos. Los datos de elecciones anteriores indican que un alcalde Demócrata es sucedido por otro Demócrata el 30 % de las veces y por un Republicano el 70 % de las veces. Por otra parte, un alcalde Republicano es sucedido por otro Republicano un 60 % del tiempo y por un Demócrata el 40 % de las veces. Se pide:

   a) Construir la matriz de transición.

   b) Determinar la probabilidad de que haya un alcalde Republicano 8 años después de que haya un Republicano en la alcaldía. Ídem tras 12 años.

   c) Determinar un vector distribución inicial si el pueblo tiene ahora mismo un alcalde Demócrata. Mostrar que las componentes de $\vec{d}^{(1)}$ son las probabilidades de que el siguiente alcalde sea un Republicano y un Demócrata, respectivamente.

   d) Obtener el vector distribución límite, y usarlo para determinar la probabilidad de tener un alcalde Republicano a largo plazo.

4. Ante el buen resultado de sus empresas, el señor Oriol piensa acometer un ambicioso proyecto de inversiones que la comprende la adquisición de la cadena

de supermercados MERCADSA, la cual en la actualidad únicamente disfruta del 25 % de cuota de mercado, estando el 75 % restante ocupado por la multinacional MARKETS CO.

El objeto de dicha actuación es conseguir, con la nueva gestión del equipo ejecutivo de las empresas ORIOL, alcanzar una cuota de mercado predominante, es decir, superior al 50 %.

Para asegurarse la consecución de dicho objetivo, el señor Oriol realiza una consulta a PRIVASA, la cual, tras un intenso estudio sobre cómo va a evolucionar en el tiempo el comportamiento del mercado, llega a las siguientes conclusiones básicas:

- La evolución dinámica en el tiempo está caracterizada por una cadena de Markov.
- Como consecuencia del cambio cualitativo de gestión, a MERCADSA se le supone con capacidad para retener anualmente el 72 % de la cuota de mercado del año anterior, perdiendo únicamente el 28 % restante en favor de MARKETS CO., la cual a su vez retendrá un 58 % de la cuota del año precedente.
- El comportamiento definido por las anteriores probabilidades permanece constante en el tiempo.

Se pide:

a) Construir la matriz de transición que describe el cambio en las cuotas de mercado de los dos supermercados.

b) ¿Qué cuota de mercado alcanzará MERCADSA después de un año de gestión del equipo ejecutivo del señor Oriol? En consecuencia, ¿conseguirá MERCADSA alcanzar su objetivo inicial en ese corto periodo de tiempo?

c) Si se contempla un horizonte temporal a largo plazo, ¿qué cuota de mercado será capaz de abarcar?

5. Estudiemos los mercados de mantequilla y margarina de un determinado país; estos dos mercados estaban en equilibrio, pero una disposición del consejo de ministros sobre las grasas comestibles ha generado considerables distorsiones en los mismos. La distorsión en cualquiera de esos mercados en el periodo $t+1$ es una función de la distorsión en los dos mercados en el periodo anterior, y más concretamente según las relaciones

$$\begin{aligned} x_{t+1} &= 0{,}6x_t + 0{,}5y_t \\ y_{t+1} &= 0{,}8x_t + 0{,}3y_t \end{aligned}$$

siendo x, y las distorsiones en los mercados de mantequilla y margarina respectivamente. ¿Desaparecerán estas distorsiones a largo plazo?

6. En una determinada comunidad autónoma española el 20 % de las rentas familiares anuales en 2006 son inferiores a 20000 euros, el 75 % está comprendido entre 20000 y 40000 euros y el 5 % restante supera esta cifra. A estos tres tramos de renta se les denomina tramos de renta baja, media y alta respectivamente.

   Se sabe que año tras año, el 60 % de las familias con renta baja permanece en dicho tramo, mientras que un 30 % pasan a renta media y un 10 % a renta alta. De las familias con renta media, un 40 % permanecen en este tramo, un 30 % pasan a renta baja y un 10 % a renta alta. Por último, el 60 % de familias conrenta alta siguen siéndolo, mientras que un 10 % pasan a renta baja y un 30 % a renta media.

   a) ¿Cuál será el porcentaje de las familias en cada uno de los tramos en 2007? ¿Y en 2008?

   b) Las autoridades de la comunidad autónoma están preocupadas por el tema de la distribuión futura de la renta y están pensando en aplicar medidas correctoras, ya que creen que la situación actual puede empeorar las cosas en el futuro. Para ello deben conocer lo que sucederá al cabo de 20 años. Calcularlo.

   c) Calcular también el porcentaje de las familias que estará en cada uno de los tramos al cabo de ese tiempo.

7. Los obreros, dentro de la población activa de un país, se clasifican en dos categorías profesionales: obreros especializados (x) y obreros no especializados (y). Se sabe que:

   - Cada trabajador activo tiene sólo un hijo.
   - Los hijos de los obreros especializados se reparten en las dos categorías según los porcentajes 60 % y 40 %.
   - Para los hijos de los obreros no especializados estos porcentajes son 40 %, 60 %. Se pide:

   a) Escribir el vector que representa las categorías profesionales de los obreros del país.

   b) Plantear el modelo que representa la distribución de esta fuerza del trabajo del país de generación en generación.

   c) Si en 2000 hay 3 millones de obreros especializados y 6 de obreros no especializados ¿cuántos habrá en la siguiente generación? ¿Y al cabo de 5 generaciones? ¿Qué sucederá a largo plazo?

8. La Unión Europea, Estados Unidos y Japón están muy preocupadas por las repercusiones que una nueva subida en los precios del petróleo pueden tener sobre sus economías. Cada una de estas comunidades ha definido una renta que consideran 'normal' (en ausencia de efectos exógenos extraños) y está interesada en estudiar los efectos que sobre $y_t = Y_t - Y$ puede tener esta subida del

petróleo, siendo Y_t la renta del periodo t, Y la renta 'normal' e y_t la desviación de aquélla respecto a ésta. El problema se complica debido a la gran interrelación que existe entre estas economías, donde las crisis se propagan de unas a otras con gran rapidez. Estas ideas aparecen expresadas matemáticamente en las relaciones

$$\begin{aligned} y_{E,t+1} &= 0{,}4y_{E,t} + 0{,}2y_{U,t} + 0{,}2y_{J,t} \\ y_{U,t+1} &= 0{,}1y_{E,t} + 0{,}3y_{U,t} + 0{,}2y_{J,t} \\ y_{J,t+1} &= 0{,}2y_{E,t} + 0{,}2y_{U,t} + 0{,}3y_{J,t} \end{aligned}$$

Estudiar si ante distorsiones en el sistema éste tiende a volver al equilibrio a largo plazo.

9. El 1 de enero un país decide cambiar su estructura impositiva y establecer dos únicos impuestos, sobre la renta del trabajo y sobre las ganancias de capital. En función de una serie de factores relacionados con la marcha de la economía, los agentes económicos toman sus decisiones de trabajar e invertir; al no responder cada comportamiento individual de la misma forma ante distintos factores, lo que se observa es un agregado en el que de una semana a otra el 30 % de las rentas de trabajo pasan a ser rentas del capital, y un 10 % de las ganancias de capital se convierten en rentas de trabajo. Sabiendo que el 1 de enero las rentas del trabajo ascendían a 1000 unidades monetarias y las rentas del capital a 500 unidades monetarias, que el impuesto sobre aquéllas es del 10 % y sobre éstas del 20 % y que el impuesto se recauda teniendo en cuenta la última semana del año, ¿cuál será la recaudación del Gobierno?

10. Una agencia naviera tiene sus barcos distribuidos entre los puertos de Barcelona, Málaga y Mallorca. De los barcos que al comienzo de cada mes están en Barcelona, al final de mes sólo vuelve la mitad, un 20 % se va a Málaga y el resto a Mallorca. De los que al principio de mes están en Málaga, al final de mes hay un 20 % en Barcelona, un 40 % en Mallorca y el resto vuelve a Málaga. Finalmente, de los barcos que hay en Mallorca, un 80 % regresa a puerto y el resto va a Barcelona. Suponiendo constante el número de barcos se pide:

    a) Plantear un modelo matricial que represente la distribución de la flota.

    b) Sabiendo que en el instante actual hay 350, 500 y 200 barcos respectivamente en Barcelona, Málaga y Mallorca, determinar el número de barcos que habrá en cada puerto dentro de 3 meses. Determinarlo también para dentro de un año.

    c) ¿Cuál será la flota de barcos en cada puerto a largo plazo?

11. En una población animal, la edad máxima alcanzada por sus individuos es de 12 años, y la población se clasifica en cuatro grupos de edades:

    pequeños: de 0 a 3 años.

    jóvenes: de 3 a 6 años.

    adultos: de 6 a 9 años.

ancianos: de 9 a 12 años.

Además se ha observado que la relación existente entre la población que hay en un peíodo k con respecto a la que había en el período anterior $k-1$ es la que se recoge en la tabla siguiente, expresada en porcentaje, y entendiéndose que un período es un trienio

		período $k-1$			
		pequeños	jóvenes	adultos	ancianos
período k	pequeños	0.25	0.25	0.25	0.25
	jóvenes	0.25	0	0.5	0.25
	adultos	0.25	0.5	0.25	0
	ancianos	0.25	0.25	0	0.5

Se pide:

a) Formular matemáticamente un modelo que represente la evolución temporal de la población.

b) ¿Cuál es la distribución de la población a largo plazo si en la actualidad ($k=0$) la proporción de individuos en cada uno de los cuatro grupos es del 20 %, 20 %, 20 % y 40 %, respectivamente?

Capítulo 5

FORMAS CUADRÁTICAS

5.1. Definición.

En este tema aplicaremos los resultados obtenidos con las matrices simétricas al estudio de las ecuaciones cuadráticas y de las formas cuadráticas. Las formas cuadráticas surgen de una gran variedad de problemas de física, geometría, estadística, etc. En particular, aparecen en los problemas de optimización de dos variables, importantes en muchos campos de la economía.

Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + ey + g = 0, \quad a, b, c, d, e, g \in \mathbb{R}$$

es una **ecuación cuadrática** en x e y. La expresión

$$q(x, y) = ax^2 + 2bxy + cy^2$$

se denomina **forma cuadrática asociada**.

Las gráficas de las ecuaciones cuadráticas en x e y se llaman **cónicas**. Las cónicas más importantes son la circunferencia, la elipse, la hipérbola y la parábola. A éstas se las conoce como *cónicas no degeneradas*. Las cónicas restantes se llaman degeneradas e incluyen puntos aislados y pares de rectas.

La ecuación de una forma cuadrática en x, y puede escribirse matricialmente como

$$q(x, y) = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = X^T A X.$$

La matriz simétrica A se llama **matriz de la forma cuadrática**.

Las ecuaciones cuadráticas generales pueden reducirse mediante traslación y giro de ejes a formas cuadráticas, por lo que estudiaremos éstas con detalle.

En general, una forma cuadrática en $\mathbb{R}^n$ es una función $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ definida por

$$q(\mathbf{x}) = X^T A X$$

donde A es una matriz simétrica de orden n y X es la matriz columna de coordenadas de $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

Ejemplo 49 *Consideremos la forma cuadrática* $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ *dada por*

$$q(x, y, z) = 7x^2 + 5y^2 - 4xy + 26xz + 30yz$$

Podemos escribir dicha forma cuadrática en la forma $q(x, y, z) = (x\ y\ z)A \begin{pmatrix} x \\ y \\ z \end{pmatrix}$, *con*

$$A = \begin{pmatrix} 7 & -2 & 13 \\ -2 & 5 & 15 \\ 13 & 15 & 0 \end{pmatrix}$$

5.2. Clasificación de las formas cuadráticas.

Una forma cuadrática $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ es

- *definida positiva* si $q(\mathbf{x}) > 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.
- *semidefinida positiva* si $q(\mathbf{x}) \geq 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.
- *semidefinida negativa* si $q(\mathbf{x}) \leq 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.
- *definida negativa* si $q(\mathbf{x}) < 0$ para todo $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{x} \neq \mathbf{0}$.
- *indefinida* en cualquier otra situación.

Ejemplo 50 *La forma cuadrática* $q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ *dada por* $q(x, y) = 3x^2 + 7y^2$ *es definida positiva, pues es claro que* $q(x, y) > 0$ *para todo* $(x, y) \in \mathbb{R}^2$, $(x, y) \neq (0, 0)$. *Del mismo modo, se ve claramente que la forma cuadrática* $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ *dada por* $q(x, y, z) = -x^2 - 8y^2 - 10z^2$ *es definida negativa y que la forma cuadrática* $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ *dada por* $q(x, y, z) = 2x^2 + 3z^2$ *es semidefinida positiva.*

Con este ejemplo podemos darnos cuenta de que la propiedad de una forma cuadrática de ser definida positiva o negativa se descubre inmediatamente cuando ésta solamente tiene términos no acoplados. Más en concreto, podemos decir que la forma cuadrática $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

$$q(x_1, x_2, \ldots, x_n) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \cdots + \lambda_n x_n^2$$

será definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si y sólo si todos los coeficientes λ_i son números positivos (negativos, no negativos, no positivos, respectivamente). Nótese que en este caso la matriz simétrica A de la forma cuadrática es en realidad una matriz *diagonal* (cuyos elementos en la diagonal principal son los números λ_i).

Ejemplo 51 *Consideremos la forma cuadrática* $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ *dada por*

$$q(x, y, z) = 6x^2 + 5y^2 + 7z^2 - 4xy + 4xz = X^T AX$$

con

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}.$$

En este caso no es posible afirmar a primera vista que esta forma cuadrática es definida positiva (como de hecho lo es). Sin embargo, si escribiéramos nuestra forma cuadrática como

$$q(x, y, z) = 3\left(\frac{2}{3}x + \frac{2}{3}y - \frac{1}{3}z\right)^2 + 6\left(\frac{1}{3}x - \frac{2}{3}y - \frac{2}{3}z\right)^2 + 9\left(\frac{2}{3}x - \frac{1}{3}y + \frac{2}{3}z\right)^2$$

resultaría evidente que es definida positiva, pues es una suma de cuadrados con coeficientes positivos.

Este ejemplo nos muestra que la propiedad de que una forma cuadrática sea definida positiva o negativa se descubre fácilmente si escribimos la expresión que defina a la forma como una suma de cuadrados. Se plantean así dos preguntas naturales: 1) ¿Es siempre posible escribir una forma cuadrática como una suma de cuadrados? 2) Si la respuesta es afirmativa, ¿cómo hacerlo?

Consideremos la forma cuadrática $q : \mathbb{R}^n \longrightarrow \mathbb{R}$, $q(\mathbf{x}) = X^T AX$. En el tema 4 veíamos que, por ser A simétrica, entonces es diagonalizable ortogonalmente, es decir, existe una matriz ortogonal P tal que $P^T AP$ es una matriz diagonal. Se vio también que las columnas de P forman una base ortonormal de $\mathbb{R}^n$. Sea $B = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ dicha base. Recordemos que los vectores $\mathbf{v}_i$ son vectores propios de A, tienen norma 1 y son ortogonales. Escribamos el vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ en términos de la base B: $\mathbf{x} = x'_1\mathbf{v}_1 + x'_2\mathbf{v}_2 + \cdots + x'_n\mathbf{v}_n$ o, en forma de matriz,

$$X' = \begin{pmatrix} x'_1 \\ x'_2 \\ \vdots \\ x'_n \end{pmatrix}$$

con lo que $X = PX'$. Entonces la forma $q(\mathbf{x})$ queda como

$$q(\mathbf{x}) = X^T AX = (PX^{'})^T A(PX^{'}) = (X^{'})^T(P^T AP)X^{'} = (X^{'})^T DX^{'}$$

donde D es una matriz diagonal cuyos elementos en la diagonal principal son los valores propios de A. Sean $\lambda_1, \lambda_2, \ldots, \lambda_n$ tales valores propios. Escribiendo explícitamente la última expresión tenemos

$$q(\mathbf{x}) = (X^{'})^T DX^{'} = \lambda_1(x'_1)^2 + \lambda_2(x'_2)^2 + \cdots + \lambda_n(x'_n)^2,$$

con lo que hemos logrado escribir $q(\mathbf{x})$ como una suma de cuadrados de las coordenadas x'_i del vector $\mathbf{x}$ respecto a la base ortonormal B. No es difícil ver que tales coordenadas se pueden calcular como $x'_i = \mathbf{x} \cdot \mathbf{v}_i$.

En resumen, la forma cuadrática $q : \mathbb{R}^n \longrightarrow \mathbb{R}$, $q(\mathbf{x}) = X^T AX$, con A matriz simétrica se puede escribir como

$$q(\mathbf{x}) = \lambda_1(\mathbf{x} \cdot \mathbf{v}_1)^2 + \lambda_2(\mathbf{x} \cdot \mathbf{v}_2)^2 + \cdots + \lambda_n(\mathbf{x} \cdot \mathbf{v}_n)^2$$

donde λ_i son los valores propios de A y $\mathbf{v}_i$ son vectores propios unitarios correspondientes a tales valores propios. Se tiene así probado el siguiente

Teorema 5.2.1 *La forma cuadrática* $q : \mathbb{R}^n \longrightarrow \mathbb{R}$, $q(\mathbf{x}) = X^T AX$, *con* A *matriz simétrica, es definida positiva (definida negativa, semidefinida positiva, semidefinida negativa) si y sólo si todos los valores propios de* A *son positivos (negativos, no negativos, no positivos, respectivamente). La forma cuadrática es indefinida si* A *tiene valores propios positivos y negativos.*

Regresando al ejemplo anterior, vemos que la matriz

$$A = \begin{pmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{pmatrix}.$$

tiene por valores propios $\lambda_1 = 3$, $\lambda_2 = 6$, $\lambda_3 = 9$ y que los siguientes vectores propios forman una base ortonormal:

$$\mathbf{v}_1 = \left(\frac{2}{3}, \frac{2}{3}, -\frac{1}{3}\right), \quad \mathbf{v}_2 = \left(\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right), \quad \mathbf{v}_3 = \left(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}\right),$$

de donde se obtiene el resultado ya conocido.

Ejemplo 52 *Clasificar la forma cuadrática* $q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ *dada por*

$$q(x, y, z) = (x\ y\ z) \begin{pmatrix} 2 & 2 & -2 \\ 2 & 5 & -4 \\ -2 & -4 & 5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Solución. Para investigar si esta forma cuadrática es definida positiva, etc., sólo tenemos que obtener los valores propios de la matriz simétrica asociada a la forma q. La ecuación característica de esta matriz es

$$\det(A - \lambda I) = -(\lambda - 1)^2(\lambda - 10) = 0.$$

Así pues, los valores propios de A son $\lambda_1 = 10$, $\lambda_2 = 1$. Como éstos son positivos, concluimos que la forma dada es definida positiva. □

Ejemplo 53 *Clasifica la siguiente forma cuadrática*

$$q\left(x_1, x_2, x_3\right) = 5x_1^2 + x_1x_2 - 3x_1x_3 + 3x_2x_1 + x_2^2 - 2x_2x_3 + 5x_3x_1 + 2x_3x_2 + x_3^2$$

Solución. Se puede resescribir como

$$q\left(x_1, x_2, x_3\right) = 5x_1^2 + 2x_1x_2 + x_1x_3 + 2x_2x_1 + x_2^2 + x_3x_1 + x_3^2$$

cuya matriz simétrica es $A = \begin{pmatrix} 5 & 2 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$, y sus autovalores son $\lambda = 0, 1, 6$, por tanto Q es semidefinida positiva. □

A menudo, dada una matriz simétrica correspondiente a una forma cuadrática, no resulta nada sencillo calcular sus valores propios (al menos sin recurrir a técnicas numéricas aproximadas), con lo cual el criterio de clasificación que hemos dado anteriormente tiene una validez limitada. Por ello damos a continuación otro par de criterios sin incluir su demostración, por caer ésta fuera del presente curso.

Segundo criterio de clasificación. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

una matriz cuadrada real. Llamemos *menor preferente* de A al determinante de cualquier submatriz $A_{i_1 i_2 \ldots i_r}$ de r filas y r columnas, obtenida suprimiendo en A todas las filas y columnas que no tengan sus subíndices en el conjunto $\{i_1, i_2, \ldots, i_r\}$ de $\{1, 2, \ldots, n\}$, con $1 \leq i_1 < i_2 < \cdots < i_r \leq n$, es decir,

$$A_{i_1 i_2 \ldots i_r} = \begin{vmatrix} a_{i_1 i_1} & a_{i_1 i_2} & \dots & a_{i_1 i_r} \\ \vdots & \vdots & & \vdots \\ a_{i_r i_1} & a_{i_r i_2} & \dots & a_{i_r i_r} \end{vmatrix}$$

Así, dada la matriz

$$A = \begin{pmatrix} 4 & 7 & 2 \\ 0 & 3 & 1 \\ 1 & 5 & 6 \end{pmatrix}$$

sus menores preferentes son

$$A_1 = 4;\ A_2 = 3;\ A_3 = 6;\ A_{12} = \begin{vmatrix} 4 & 7 \\ 0 & 3 \end{vmatrix} = 12;\ A_{13} = 22;\ A_{23} = \begin{vmatrix} 3 & 1 \\ 5 & 6 \end{vmatrix} = 13;\ A_{123} = \det(A).$$

Pues bien, entonces se tiene el siguiente criterio:

(2a) q es definida positiva si y sólo si todos los menores preferentes de su matriz simétrica asociada A son estrictamente positivos.

(2b) q es semidefinida positiva si y sólo si todos los menores preferentes de A son no negativos (i.e., mayores o iguales que cero) y $\det(A) = 0$.

(2c) q es definida negativa si y sólo si todos los menores preferentes de A de orden par son estrictamente positivos y todos los de orden impar son estrictamente negativos.

(2d) q es semidefinida negativa si y sólo si todos los menores preferentes de A de orden par son no negativos (i.e., mayores o iguales que cero), los de orden impar son no positivos (i.e., menores o iguales que cero) y $\det(A) = 0$.

(2e) q es indefinida si y sólo si no se verifican los apartados anteriores.

Las dificultades que tienen lugar cuando se lleva a la práctica este criterio (sobre todo porque el número de menores preferentes a analizar aumenta espectacularmente con la dimensión de la matriz) nos lleva a incluir un criterio adicional, más sencillo y eficiente (a veces).

Tercer criterio de clasificación. Sea

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

una matriz cuadrada real. Llamamos *menores principales* de A a los siguientes determinantes:

$$\Delta_1 = a_{11},\ \Delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix};\ \dots,\ \Delta_n = \det(A).$$

Entonces se tiene:

Si $\det(A) \neq 0$,

(3a) q es definida positiva si y sólo si $\Delta_i > 0$ para todo $i = 1, 2, \dots, n$.

(3b) q es definida negativa si y sólo si $(-1)^i \Delta_i > 0$ para todo $i = 1, 2, \dots, n$.

(3c) q es indefinida en cualquier otro caso.

Si $\det(A) = 0$,

(4a) si q es semidefinida positiva entonces $\Delta_i \geq 0$ para todo $i = 1, 2, \dots, n$ (y existe al menos uno no nulo).

(4b) si q es semidefinida negativa entonces $(-1)^i \Delta_i \geq 0$ para todo $i = 1, 2, \dots, n$ (y existe al menos uno no nulo).

(4c) q es indefinida en cualquier otro caso.

Veamos a continuación algunos ejemplos adicionales.

Ejemplo 54 *Clasificar la forma cuadrática que tiene por matriz simétrica asociada la siguiente*

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

<u>Solución</u>. Si calculamos sus menores principales resulta

$$\Delta_1 = 2 > 0,\ \Delta_2 = 1 > 0; \Delta_3 = \det(A) = 2 > 0,$$

por lo que, en virtud de (3a), la forma cuadrática correspondiente a A es definida positiva. Nótese también que sus valores propios son todos estrictamente positivos (primer criterio), así como todos sus menores preferentes (2a). □

Ejemplo 55 *La matriz $-A$ (o su forma cuadrática correspondiente) es definida negativa, como se puede comprobar por cualquiera de los tres criterios, o simplemente, porque A, tal y como se vio en el ejemplo anterior, es definida positiva.*

Ejemplo 56 *Clasificar la forma cuadrática*

$$q(x, y, z) = x^2 + y^2 + z^2 + 2xy + 2xz + 2yz,$$

con matriz simétrica asociada

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

Solución. Esta forma cuadrática no se puede clasificar con el tercer criterio, ya que sus menores principales son

$$\Delta_1 = 1, \ \Delta_2 = 0, \ \Delta_3 = \det(A) = 0$$

y aunque podemos decir con toda seguridad que no es definida positiva ni definida negativa, no podemos afirmar otra cosa basándonos únicamente en este criterio. Sin embargo, sus menores preferentes

$$A_1 = A_2 = A_3 = 1 > 0, \ A_{12} = A_{23} = 0, \ A_{123} = \det(A) = 0$$

con lo cual la forma cuadrática q es semidefinida positiva. También se puede comprobar que los valores propios de la matriz A son todos no negativos y que tiene uno cero. □

Ejercicio 17 *Clasifica la forma cuadrática dada por*

$$q(x, y, z) = x^2 + y^2 + 2xy + 2xz + 2yz.$$

5.3. Problemas

1. Clasificar las siguientes formas cuadráticas:

   a) $f(x, y) = 4x^2 + 5y^2 + 8yx$
   b) $f(x, y) = -x^2 - 3y^2 + yx$
   c) $f(x, y) = \frac{1}{2}x^2 + \frac{1}{4}y^2 - yx$

2. Escribir las siguientes formas cuadráticas en forma matricial, con **A** simétrica, y clasificarlas:

   a) $f(x, y, z) = 3x^2 - 2xy + 3xz + y^2 - 4yz + 3z^2$
   b) $f(x, y, z) = x^2 + 8z^2$
   c) $f(x, y, z) = -3x^2 + 2xy - y^2 + 4yz - 8z^2$

3. Dada la forma cuadrática $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida mediante

   $$f(x, y, z) = x^2 + ay^2 + 2yz + az^2$$

   hállese $a \in \mathbb{R}$ para que f sea semidefinida, indicando si lo es positiva o negativa.

4. Dada la forma cuadrática $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ definida mediante
$$f(x, y, z) = \alpha x^2 + \alpha y^2 + (\alpha - 1)z^2 + 2xy$$
donde $\alpha \in R$ es fijo, clasificar f para los distintos valores de α.

5. Encontrar una base de $\mathbb{R}^3$ en la que la expresión analítica de la forma cuadrática
$$f(x, y, z) = 7x^2 - 2xy + 5y^2 + 5z^2 - 2xz + 2yz$$
sea canónica, esto es, no aparezcan términos cruzados.

6. En la base canónica de $\mathbb{R}^2$ una forma cuadrática $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ adopta la siguiente expresión:
$$f(x, y) = 4x^2 + 8y^2 - 6xy$$
Determinar su nueva expresión en la base formada por los vectores $\vec{v}_1 = (1, 1)$ y $\vec{v}_2 = (-1, 1)$.

7. Dada la forma cuadrática
$$f(x_1, x_2, x_3) = x_1^2 - x_2^2 + 3x_3^2 + 4x_1x_2 + 6x_1x_3 + 8x_2x_3$$
encontrar su expresión canónica y clasificarla.

8. Dada la forma cuadrática $q : \mathbb{R}^4 \rightarrow \mathbb{R}^4$ definida por
$$q(x, y, z) = 6x^2 + 3y^2 + 2z^2 + t^2 + 4xy + 6xz + 4xt + 2yz + \frac{8}{3}yt + 2zt$$
Clasificar q.

9. Determinar si son definidas positivas o negativas las formas cuadráticas siguientes sujetas a la restricción lineal dada:

    a) $f(x, y) = x^2 - 2xy + y^2$ sujeta a $x + y = 0$
    b) $f(x, y) = 2x^2 - 4xy + y^2$ sujeta a $3x + 4y = 0$
    c) $f(x, y) = -x^2 + xy - y^2$ sujeta a $5x - 2y = 0$

10. Clasificar la forma cuadrática
$$f(x, y, z) = x^2 + 2xy + y^2 + 3z^2$$
restringida a:

    a) $x - y + z = 0$
    b) $T = \{(x, y, z) \in R^3 / z = 0\}$

11. Clasificar la forma cuadrática
$$f(\vec{x}) = X^T \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} X$$

12. Dada la forma cuadrática $f : \mathbb{R}^3 \longmapsto \mathbb{R}$ definida mediante $f(x,y,z) = 5x^2 + 5y^2 + 4z^2 + 2xy$, diagonalizar ortogonalmente f y clasificar f.

13. Dada la forma cuadrática $f : \mathbb{R}^3 \longmapsto \mathbb{R}$

$$f(x,y,z) = [x \;\; y \;\; z] \begin{bmatrix} 3 & -1 & 4 \\ -1 & 3 & 10 \\ -4 & -10 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

hállese una matriz simétrica que de lugar a la misma forma cuadrática y diagonalícese ortogonalmente, dando explícitamente la matriz ortogonal que diagonaliza.

5.4. Aplicaciones de las formas cuadráticas

13 La función de costes de una empresa viene representada por

$$C = L^2 + 3LK + 3K^2$$

siendo K y L el trabajo y el capital respectivamente.

El equipo de economistas de dicha empresa recibe una notificación según la cual deben facilitar al consejo directivo de la empresa las cantidades de capital y trabajo que minimizan los costes.

Reunido dicho equipo llegan a la conclusión de que las cantidades pedidas son $L = K = 0$.

Sabiendo que la minimización de la función de costes tiene lugar si se verifica que la forma cuadrática

$$W = (\Delta L, \Delta K) \begin{pmatrix} 2 & 3 \\ 3 & 6 \end{pmatrix} \begin{pmatrix} \Delta L \\ \Delta K \end{pmatrix}$$

es definida positiva. Comprobar si la decisión del equipo es correcta.

14 El objetivo de la política fiscal del gobierno se basa en la reducción del déficit público, para lo cual está estudiando la posibilidad de introducir un nuevo impuesto en función de lo pagado (o devuelto) en concepto del Impuesto sobre la Renta de las Personas Físicas (R) y el Impuesto sobre el Patrimonio (P). El impuesto sería

$$T = 2R^2 + 4P^2 - 4RP$$

El gobierno no está dispuesto a devolver dinero por este nuevo concepto tributario, por lo que te ha encargado el estudio del mismo a fin de comprobar este último punto.

15 Los economistas de una empresa aseguran que la función de producción por la que se sigue la misma es del tipo $P = L^2 + K^2 - 2LK$, siendo L y K el número de trabajadores y el de máquinas respectivamente. Además se sabe que para que funcione una máquina se necesitan dos trabajadores. Comprobar que P es una verdadera función de producción.

16 El nivel de contaminación que produce una empresa viene dado por

$$C = x_1^2 - 3x_2^2 - 2x_3^2 - 2x_1x_2$$

donde x_1, x_2, x_3 son cantidades de materias primas utilizadas en el proceso productivo y C es el nivel de contaminación. Además, en su proceso de producción utiliza la misma cantidad de la materia prima x_1 que la suma de las cantidades de x_2 y x_3. El gobierno no autoriza los procesos productivos que generen niveles de contaminación positivos. Si tú eres el gerente de esta empresa, ¿pedirías la autorización? ¿Por qué?

Bibliografía

[1] J.A. Barrios, C. González, J.C. Moreno. *Álgebra matricial para economistas.* Ed. AC, 1993.

[2] R. Bronson. *Linear Algebra. An introduction.* Academic Press, 1995.

[3] R. E. Caballero Fernández, A.C. González Pareja, F.A. Triguero Ruiz. *Métodos matemáticos para la economia.* McGraw-Hill, 1992.

[4] J.R. Cancelo, J. López Ortega, C. González Conde, J.M. Montero. *Problemas de Álgebra lineal para economistas.* Ed. Tebar-Flores, 1987. (Dos volúmenes).

[5] S.I. Grossman. *Aplicaciones del álgebra lineal.* Grupo Editorial Iberoamérica, 1988.

[6] S.I. Grossman. *Álgebra lineal*, 5ª edición. McGraw-Hill, 1996.

[7] D.C. Lay. *Álgebra lineal y sus aplicaciones*, 2ª edición. Prentice Hall, 2001.

[8] P. Sanz, F.J. Vázquez, P. Ortega. *Álgebra lineal. Cuestiones, ejercicios y tratamiento en Derive.* Prentice Hall, 1998.

[9] K. Sydsaeter, P.J. Hammond. *Matemáticas para el análisis económico.* Prentice Hall, 1996.