

# Capítulo 1

## Matrices

### Resumen y motivación.

*Este tema es prácticamente un repaso de temas previamente estudiados en bachillerato. Se tratarán con rigor distintos aspectos relacionados con las matrices tales como el rango o los determinantes y se repasará el método de Gauss.*

### Objetivos.

*Al final del tema será imprescindible que los alumnos controlen con soltura el cálculo de matrices, puesto que junto con el tema 2 constituirá la base fundamental del resto de los temas. En particular, deben recordar los conceptos de matriz y determinante, aprender a resolver determinantes, y aprender a aplicar el método de Gauss y ser conscientes de su utilidad para calcular inversas, determinantes y el rango de una matriz.*

### Estructura del tema.

• Conceptos básicos y ejemplos • Operaciones básicas con matrices • Concepto de matriz regular y propiedades • Determinante asociado a una matriz cuadrada • Rango de una matriz

## 1.1. Conceptos básicos y ejemplos

**Definición 1.1.1.** Dados dos números naturales  $m$  y  $n$ , una **matriz** de **orden** o **dimensión**  $m \times n$  es una tabla numérica rectangular con  $m$  filas y  $n$  columnas.

Los elementos de las matrices podrán ser tanto números reales como complejos. La teoría no varía de unos a otros, pero para simplificar, en la mayoría de los ejemplos y de los ejercicios, consideraremos números reales.

Las matrices se suelen denotar por letras mayúsculas:  $A, B, X, \dots$  y los elementos de una matriz se denotan con la misma letra que la matriz pero en minúscula, y con dos subíndices representando la fila y la columna en la que se encuentra el elemento. Cuando escribimos una matriz por medio de sus elementos, escribimos éstos en forma de tabla y delimitados por paréntesis. Por ejemplo, escribimos una matriz genérica de orden  $m \times n$  como

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}.$$

Otras formas abreviadas de denotar la misma matriz son

$$A = (a_{ij})_{j=1, \dots, n}^{i=1, \dots, m}, \quad A = (a_{ij})_{m,n}, \quad \text{o simplemente} \quad A = (a_{ij}).$$

Se define la relación de igualdad entre matrices de la siguiente manera.

**Definición 1.1.2.** Dos matrices  $A$  y  $B$  son **iguales** si tienen la misma dimensión  $m \times n$  y si  $a_{ij} = b_{ij}$  para cualesquiera  $1 \leq i \leq m$  y  $1 \leq j \leq n$ , es decir, si coinciden elemento a elemento.

Algunos casos especiales de matrices según su orden son los siguientes.

**Definición 1.1.3.** Se dice que una matriz  $A$  es una **matriz cuadrada** si  $m = n$ . En este caso podemos decir simplemente que la matriz es una matriz cuadrada de orden  $n$ . Se dice que una matriz es una **matriz fila** o **vector fila** si  $m = 1$ , es decir, si tiene una única fila. Análogamente, una matriz es una **matriz columna** o **vector columna** si  $n = 1$ , es decir, si tiene una única columna. Se sobreentiende que una matriz fila de orden  $n$  es una matriz de orden  $1 \times n$ , y una matriz columna de orden  $m$  es una matriz de orden  $m \times 1$ .

En el caso particular de las matrices cuadradas, podemos definir su diagonal y su traza.

**Definición 1.1.4.** Dada una matriz cuadrada, se define su **diagonal principal**, o simplemente su **diagonal**, como el conjunto de todos los elementos de  $A$  del tipo  $a_{ii}$ . La suma de los elementos de la diagonal principal de una matriz cuadrada se llama **traza**:

$$\text{tr}(A) = \sum_{i=1}^n a_{ii}.$$

Por último, introduciremos el concepto de submatriz.

**Definición 1.1.5.** Dada una matriz  $A$ , llamaremos **submatriz** de  $A$  a cada matriz que se obtenga de ella suprimiendo algunas de sus filas y columnas.

### 1.1.1. Tipos particulares de matrices

Veamos ahora algunos casos particulares de matrices cuadradas de especial relevancia. Sea  $A$  una matriz cuadrada de orden  $n$ . Entonces, se dice que:

- $A$  es una **matriz simétrica** si para cualesquiera  $i, j$  con  $1 \leq i, j \leq n$  se tiene que  $a_{ij} = a_{ji}$ .
- $A$  es una **matriz antisimétrica** si para cualesquiera  $i, j$  con  $1 \leq i, j \leq n$  se tiene que  $a_{ij} = -a_{ji}$ . En particular, los elementos de la diagonal de una matriz antisimétrica valen cero.
- $A$  es una **matriz diagonal** si todos los elementos que no pertenezcan a la diagonal valen cero. En el caso particular en que todos los elementos de la diagonal valgan 1, la matriz se llama **matriz identidad** o **matriz unidad**, y se denota por  $I_n$  donde  $n$  es el orden de la matriz. Cuando no haya lugar a confusión con respecto al orden de la matriz, denotaremos la matriz identidad simplemente por  $I$ .
- $A$  es una **matriz triangular superior (inferior, respectivamente)** si todos los elementos por debajo (por encima, respectivamente) de la diagonal son cero.

Por otro lado, supongamos que  $A$  es una matriz, no necesariamente cuadrada, de orden  $m \times n$ . Llamaremos **pivote** de una fila de  $A$  al primer elemento no nulo de dicha fila (si es que existe). Entonces se dice que:

- $A$  es una matriz **escalonada por filas** si verifica las siguientes condiciones:
  1. Si  $A$  tiene filas compuestas enteramente por ceros (filas nulas), éstas están agrupadas en la parte inferior de la matriz.
  2. El pivote de cada fila no nula es 1.
  3. El pivote de cada fila no nula está a la derecha del de la fila anterior.

Análogamente, se podrían definir las matrices escalonadas por columnas.

Por último, como un caso particular de matriz, definimos la **matriz nula** de orden  $m \times n$ ,  $O_{m \times n}$ , como la matriz de orden  $m \times n$  formada enteramente por ceros. Si la matriz es cuadrada

de orden  $n$  la denotaremos por  $O_n$ , y cuando no haya confusión posible con respecto al orden la denotaremos simplemente por  $O$ .

## 1.2. Operaciones básicas con matrices

En esta sección vamos a repasar las operaciones básicas que se pueden realizar con matrices: traspuesta de una matriz, suma de matrices, producto de una matriz por un escalar y producto de matrices. Veremos también las principales propiedades de estas operaciones.

### 1.2.1. Traspuesta de una matriz

**Definición 1.2.1.** Dada una matriz  $A$  de orden  $m \times n$ , se define la **matriz traspuesta** de  $A$  como la matriz de orden  $n \times m$  que obtenemos cuando intercambiamos las filas y las columnas de  $A$ . Denotamos la matriz traspuesta por  $A^t$ .

De la definición se deduce que:

- Dada una matriz cualquiera  $A$ , la traspuesta de su traspuesta es ella misma:  $(A^t)^t = A$ .
- Si  $A$  es una matriz cuadrada, entonces  $A^t$  también es una matriz cuadrada del mismo orden.

Por otro lado, es posible caracterizar las matrices simétricas y antisimétricas con ayuda de la matriz traspuesta:

- Una matriz cuadrada  $A$  es una matriz simétrica si y sólo si  $A = A^t$ .
- Una matriz cuadrada  $A$  es una matriz antisimétrica si y sólo si  $A = -A^t$ .

### 1.2.2. Suma de matrices

Para sumar dos matrices, las matrices tienen que ser del mismo orden.

**Definición 1.2.2.** Dadas dos matrices  $A$  y  $B$  de orden  $m \times n$ , se define la **suma** de ambas matrices como la matriz del mismo orden  $m \times n$ ,  $A + B$ , que se obtiene al sumar elemento a elemento:

$$c_{ij} = a_{ij} + b_{ij}.$$

La suma de matrices de orden  $m \times n$  tiene las siguientes propiedades:

1. Es una operación *asociativa*: dadas  $A$ ,  $B$  y  $C$  se tiene que  $(A + B) + C = A + (B + C)$ .
2. Es una operación *conmutativa*: dadas  $A$  y  $B$  se tiene que  $A + B = B + A$ .
3. Existe un *elemento neutro*: es decir, existe una única matriz de orden  $m \times n$ , conocida como elemento neutro, de modo que al sumarla con cualquier matriz  $A$  del mismo orden la deje invariante. Esta matriz es la matriz nula  $O_{m \times n}$  ya que dada cualquier matriz  $A_{m \times n}$  se tiene que  $A + O = A$ .
4. Existe el *elemento opuesto*: es decir, dada una matriz  $A$  existe otra matriz del mismo orden, conocida como elemento opuesto de  $A$ , de modo que al sumarla con  $A$  nos da el elemento neutro. Esta matriz se denota por  $-A$  y es la matriz que obtenemos cuando cambiamos cada elemento de  $A$  por su opuesto:  $A + (-A) = O$ .

### 1.2.3. Producto de una matriz por un escalar

Con el término **escalar** nos referiremos a lo largo de la asignatura a un número real o a un número complejo. Denotaremos por  $K$  al conjunto formado por todos los escalares, es decir,  $K = \mathbb{R}$  o  $K = \mathbb{C}$  según corresponda.

**Definición 1.2.3.** Dada una matriz  $A$  de orden  $m \times n$  y un número real  $k \in K$ , se define el **producto de  $A$  por  $k$**  como la matriz de orden  $m \times n$ ,  $kA$ , que se obtiene al multiplicar cada elemento de  $A$  por  $k$ :

$$c_{ij} = ka_{ij}.$$

El producto de una matriz de orden  $m \times n$  por un escalar tiene las siguientes propiedades:

1. Es una operación *asociativa*: dada una matriz  $A$  y dados dos escalares  $k$  y  $l$ , se tiene que  $k(lA) = (kl)A$ .
2. Es una operación *distributiva con respecto a la suma de escalares*: dada una matriz  $A$  y dos escalares  $k$  y  $l$ , se tiene que  $(k + l)A = kA + lA$ .
3. Es una operación *distributiva con respecto a la suma de matrices*: dadas dos matrices  $A$  y  $B$  y un escalar  $k$ , se tiene que  $k(A + B) = kA + kB$ .
4. Existe un *elemento unidad*: es decir, existe un único escalar de modo que al multiplicarlo por cualquier matriz  $A$  la deja invariante. El elemento unidad es el número 1.

Las dos operaciones anteriores junto con las propiedades que hemos enumerado servirán de motivación para estudiar en el tema 3 los espacios vectoriales.

### 1.2.4. Producto de matrices

Para multiplicar dos matrices habremos de tener mucho cuidado con sus dimensiones.

**Definición 1.2.4.** Sea  $A$  una matriz de orden  $m \times p$  y  $B$  una matriz de orden  $p \times n$ , se define su **producto** como la matriz de orden  $m \times n$ ,  $AB$ , cuyos elementos vienen dados por:

$$c_{ij} = \sum_{k=1}^p a_{ik}b_{kj}.$$

El producto de matrices verifica las siguientes propiedades:

1. *No es una operación conmutativa.* De hecho, dadas dos matrices  $A$  y  $B$  puede que sólo tenga sentido realizar uno de los dos productos  $AB$  o  $BA$  (o ninguno de ellos). Aún en el caso de poder realizarse los dos productos, éstos no tienen porqué coincidir.
2. Es una operación *asociativa*:  $A(BC) = (AB)C$  siempre que los productos tengan sentido.
3. Dada una matriz de orden  $m \times n$  se tiene que  $I_m A = A$  y  $A I_n = A$ .
4. Se verifica la propiedad *distributiva con respecto a la suma de matrices*:  $A(B + C) = AB + AC$  y  $(A + B)C = AC + BC$  siempre que los productos tengan sentido.

El comportamiento de la trasposición de matrices con respecto a las operaciones anteriores es:

- Dadas  $A$  y  $B$  dos matrices del mismo orden, se tiene que  $(A + B)^t = A^t + B^t$ .
- Dada  $A$  una matriz cualquiera y dado un escalar  $k$ , se tiene que  $(kA)^t = kA^t$ .
- Dadas dos matrices  $A$  de orden  $m \times p$  y  $B$  de orden  $p \times n$  se tiene que  $(AB)^t = B^t A^t$ .

## 1.3. Concepto de matriz regular y propiedades

**Definición 1.3.1.** Dadas dos matrices  $A$  y  $B$ , se dice que  $B$  es la **matriz inversa** de  $A$  si  $AB = BA = I$ . En el caso en que  $A$  admita una matriz inversa se dice que  $A$  es una matriz **regular**, y su matriz inversa se denota por  $A^{-1}$ . Una matriz se dice que es **singular** si no es regular.

Observemos que una condición necesaria para que  $A$  sea una matriz regular es que sea una matriz cuadrada.

Una caracterización de las matrices regulares viene dada por la siguiente proposición.

**Proposición 1.3.2.** Si  $A$  y  $B$  son dos matrices cuadradas de orden  $n$  tales que  $AB = I$ , entonces  $A$  es regular y  $B = A^{-1}$ .

Las matrices regulares verifican las siguientes propiedades:

- La matriz inversa de una matriz dada, si existe, es única.
- Si  $A$  es una matriz regular entonces  $A^{-1}$  también es regular y  $(A^{-1})^{-1} = A$ .
- Si  $A$  y  $B$  son dos matrices regulares del mismo orden, entonces  $AB$  también es regular y  $(AB)^{-1} = B^{-1}A^{-1}$ .
- Si  $A$  es una matriz regular entonces  $A^t$  también es regular y  $(A^t)^{-1} = (A^{-1})^t$ .

Hay un tipo de matrices para las que su inversa tiene una expresión muy sencilla, las llamadas matrices ortogonales.

**Definición 1.3.3.** Se dice que una matriz cuadrada regular es una **matriz ortogonal** si verifica

$$P^{-1} = P^t.$$

El interés de estas matrices se estudiará en temas posteriores.

### 1.3.1. Método de Gauss

La matriz inversa a una matriz dada se puede obtener directamente a partir de la definición, pero esto implica resolver  $n$  sistemas de  $n$  ecuaciones cada uno, siendo  $n$  el orden de la matriz. Por tanto, este sistema no es muy útil para matrices de orden grande. Hay métodos con un coste computacional menor para calcular la inversa de una matriz. En particular, veamos como calcular la inversa de una matriz utilizando el conocido método de Gauss. Para ello necesitamos conocer antes el siguiente concepto.

**Definición 1.3.4.** Llamamos **transformaciones elementales de filas (de columnas, respectivamente)** de una matriz  $A$  a cada uno de los siguientes tipos de transformaciones que se pueden realizar sobre  $A$ :

1. Intercambiar la posición de dos filas (dos columnas, respectivamente). Lo denotamos como  $F_i \leftrightarrow F_j$  ( $C_i \leftrightarrow C_j$ , respectivamente).
2. Multiplicar todos los elementos de una fila (columna, respectivamente) por un escalar no nulo. Lo denotamos como  $F_i \rightarrow aF_i$  ( $C_i \rightarrow aC_i$ , respectivamente).

3. Sumar a una fila (columna, respectivamente) otra multiplicada por un escalar. Lo denotamos como  $F_i \rightarrow F_i + aF_j$  ( $C_i \rightarrow C_i + aC_j$ , respectivamente).

Teniendo en cuenta la definición anterior, podemos establecer la siguiente relación entre dos matrices.

**Definición 1.3.5.** Sean  $A$  y  $B$  dos matrices del mismo orden (no necesariamente cuadradas), entonces se dice que son **equivalentes por filas** (**por columnas**, respectivamente) si podemos obtener una a partir de transformaciones elementales de filas (de columnas, respectivamente) de la otra. En este caso se denota  $A \sim_f B$  ( $A \sim_c B$ , respectivamente).

Se puede probar que una matriz cuadrada  $A$  es regular si y sólo si es equivalente por filas a la matriz identidad, es decir, si existe una serie de transformaciones elementales por filas que transforman  $A$  en la matriz identidad. Si ahora aplicamos las mismas transformaciones elementales por filas a la matriz identidad, obtendremos la matriz inversa de  $A$ ,  $A^{-1}$ . Este procedimiento es el que se conoce como método de Gauss para el cálculo de la matriz inversa.

NOTA: En general cualquier cálculo con matrices que involucre el uso de transformaciones elementales para obtener matrices equivalentes por filas o por columnas más sencillas se conoce como método de Gauss. Este método no es sólo útil para el cálculo de matrices inversas, sino también para el cálculo de determinantes, de rangos o para la resolución de sistemas de ecuaciones lineales.

## 1.4. Determinante asociado a una matriz cuadrada

**Definición 1.4.1.** A una matriz cuadrada  $A$  se le puede asociar un número real, llamado **determinante**, que nos proporciona cierta información sobre la matriz. El concepto de determinante de una matriz cuadrada puede ser definido inductivamente de la siguiente forma.

Para una matriz cuadrada de orden 1,  $A = (a_{11})$ , se define su determinante como  $\det(A) = a_{11}$ .

Supongamos ahora que conocemos el valor del determinante de una matriz cuadrada de orden  $n - 1$  y consideremos una matriz cuadrada de orden  $n$

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



Se llama *ij-ésimo menor adjunto* de  $A$  a  $\alpha_{ij} = (-1)^{i+j} \det(A_{ij})$ , donde  $A_{ij}$  es la submatriz que se obtiene de  $A$  eliminando la fila  $i$ -ésima y la columna  $j$ -ésima, y se define el determinante de  $A$  como

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{11}\alpha_{11} + a_{12}\alpha_{12} + \dots + a_{1n}\alpha_{1n}.$$

Utilizando la definición anterior para matrices cuadradas de orden 2 tenemos

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}\alpha_{11} + a_{12}\alpha_{12} = a_{11}a_{22} - a_{12}a_{21}.$$

Y para matrices de orden 3

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\alpha_{11} + a_{12}\alpha_{12} + a_{13}\alpha_{13} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}. \end{aligned}$$

Así es como se define la conocida **regla de Sarrus** para el cálculo del determinante de una matriz cuadrada de orden 3.

### 1.4.1. Propiedades de los determinantes

Los determinantes verifican las siguientes propiedades:

1. Si tres matrices cuadradas  $A$ ,  $A'$  y  $A''$  son idénticas salvo en que la  $i$ -ésima fila (o columna) de  $A$  es la suma de las filas (o columnas) correspondientes de  $A'$  y  $A''$  entonces

$$\det(A) = \det(A') + \det(A'').$$

2. Si una matriz cuadrada  $A$  tiene dos filas (o dos columnas) iguales entonces su determinante es cero,  $\det(A) = 0$ .

3. Si se intercambian dos filas (o dos columnas) de una matriz  $A$  entonces su determinante cambia de signo.
4. Si se multiplican los elementos de una fila (o columna) de una matriz  $A$  por un escalar  $k$ , entonces su determinante queda multiplicado por  $k$ .
5. Si a una fila (o columna) de una matriz cuadrada  $A$  se le suma otra multiplicada por un escalar  $k$  entonces su determinante no varía.
6. Una matriz cuadrada  $A$  es regular si y sólo si  $\det(A) \neq 0$ .
7. El determinante del producto de matrices cuadradas es igual al producto de sus determinantes,  $\det(AB) = \det(A)\det(B)$ .
8. El determinante de una matriz es igual al de su traspuesta,  $\det(A) = \det(A^t)$ .
9. El determinante de una matriz puede obtenerse desarrollando por cualquiera de sus filas (o columnas),

$$\begin{aligned}\det(A) &= a_{i1}\alpha_{i1} + a_{i2}\alpha_{i2} + \dots + a_{in}\alpha_{in} \\ &= a_{1j}\alpha_{1j} + a_{2j}\alpha_{2j} + \dots + a_{nj}\alpha_{nj}.\end{aligned}$$

para cualquier  $1 \leq i \leq n$  o  $1 \leq j \leq n$ .

Las propiedades anteriores nos facilitan el cálculo de determinantes. De hecho, podemos usarlas para transformar la matriz de modo que obtengamos algunos ceros en una de las filas (o columnas) con el objetivo de desarrollar posteriormente por dicha fila (o columna). Los determinantes los transformaremos usando las transformaciones elementales descritas en la definición 1.3.4, tanto por filas como por columnas. No obstante, hay que tener en cuenta que tras cada transformación el determinante no se mantiene invariante, sino que varía de acuerdo con las propiedades 3, 4 y 5 antes descritas.

Los determinantes nos proporcionan un método alternativo para el cálculo de la matriz inversa.

**Definición 1.4.2.** Dada una matriz cuadrada  $A$  de orden  $n$ , llamaremos **matriz adjunta** de  $A$  a la matriz  $\text{Adj}(A) = (\alpha_{ij})_{m,n}$ , es decir, la matriz que en cada posición tiene el correspondiente menor adjunto de  $A$ .

**Proposición 1.4.3.** Si  $A$  es una matriz regular de orden  $n$  entonces

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

## 1.5. Rango de una matriz

Aplicando transformaciones elementales por filas (o por columnas) a una matriz  $A$  arbitraria de dimensión  $m \times n$ , siempre se puede encontrar otra matriz equivalente por filas (o por columnas) a  $A$  que sea escalonada por filas (o por columnas). Se define el **rango de  $A$** ,  $\text{rg}(A)$ , como el número de filas (o de columnas) no nulas de la matriz equivalente por filas (o por columnas) a  $A$  escalonada por filas (o por columnas) que obtenemos. Es inmediato comprobar que el rango de  $A$  coincide con el número de pivotes que tenga la matriz escalonada correspondiente.

Si  $A$  es de orden  $m \times n$  siempre se tiene que  $\text{rg}(A) \leq \min(m, n)$ .

El rango de una matriz también se puede obtener mediante el cálculo de determinantes.

**Proposición 1.5.1.** *Sea  $A$  una matriz de orden  $m \times n$ , entonces el rango de  $A$  coincide con el mayor orden de una submatriz cuadrada regular de  $A$ .*

Por tanto, se tiene la siguiente caracterización de las matrices regulares.

**Proposición 1.5.2.** *Una matriz cuadrada  $A$  de orden  $n$  es regular si y sólo si  $\text{rg}(A) = n$ .*



## Capítulo 2

---

# Sistemas de ecuaciones lineales

---

### Resumen y motivación.

*De nuevo este tema es un repaso de temas previamente estudiados en bachillerato. La resolución de sistemas es fundamental a lo largo de toda la asignatura, y en general es imprescindible prácticamente en cualquier asignatura de carácter científico que los estudiantes tengan que cursar para superar sus estudios. La discusión de un sistema también es importante, ya que a menudo los alumnos no necesitarán encontrar explícitamente las soluciones, pero sí tener conocimiento sobre la existencia y/o unicidad de éstas.*

---

### Objetivos.

*Al finalizar el tema los estudiantes deben poder entender los conceptos de solución particular y general de un sistema de ecuaciones, saber clasificarlos según el número de soluciones obtenidas, aprender a discutir y a resolver sistemas de ecuaciones lineales y aprender a modelar problemas sencillos a través de sistemas de ecuaciones lineales.*

---

### Estructura del tema.

• Definiciones básicas • Clasificación de los sistemas según el número de soluciones. Teorema de Rouché-Frobenius • Métodos de resolución: métodos de Cramer y de Gauss • Aplicaciones: método de los mínimos cuadrados

## 2.1. Definiciones básicas

Una **ecuación lineal** es una ecuación polinómica de grado 1 en una o varias incógnitas. Es decir, es una expresión de la forma

$$a_1x_1 + \dots + a_nx_n = b,$$

donde los términos  $a_1, \dots, a_n$  son escalares conocidos que se llaman **coeficientes**; el término  $b$  es también otro escalar conocido que se llama **término independiente**, y por último los símbolos  $x_1, \dots, x_n$  se conocen como **incógnitas** y son, a priori, desconocidas. Para un número pequeño de incógnitas será usual también denotarlas por las letras  $x, y, z, t, \dots$

Una **solución** de una ecuación es una asignación de valores de las incógnitas para la cual se verifica la igualdad.

**Definición 2.1.1.** Se llama **sistema de  $m$  ecuaciones lineales con  $n$  incógnitas** a un conjunto de  $m$  ecuaciones lineales en las mismas  $n$  incógnitas:

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

Llamaremos **solución** del sistema a cada asignación de valores de las incógnitas  $x_1 = k_1, \dots, x_n = k_n$  que sea solución común a todas las ecuaciones del sistema, es decir, que verifique todas las igualdades simultáneamente. Se dice también que  $(k_1, \dots, k_n)$  es solución del sistema. Se llama **solución general** del sistema al conjunto de todas las soluciones del sistema. Resolver un sistema es hallar su solución general. Dos sistemas se dice que son **sistemas equivalentes** si tienen la misma solución general, es decir, si tienen exactamente las mismas soluciones.

Para transformar un sistema en otro sistema equivalente podemos realizar las siguientes operaciones elementales:

- Intercambiar dos ecuaciones.
- Multiplicar una ecuación por un escalar no nulo.
- Sumar a una ecuación un múltiplo de otra.
- Eliminar las ecuaciones triviales del tipo  $0 = 0$ , las ecuaciones repetidas o las proporcionales.

### 2.1.1. Forma matricial de un sistema de ecuaciones lineales

Cualquier sistema de ecuaciones lineales se puede expresar de forma matricial. Por ejemplo, (2.1) se puede escribir como

$$AX = \mathbf{b},$$

donde

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

A la matriz  $A$  se le llama **matriz del sistema o de coeficientes**. El vector  $X$  es el **vector de incógnitas** y el vector  $\mathbf{b}$  es el **vector de términos independientes**. Por último llamamos **matriz ampliada** a la matriz que se forma cuando añadimos a la matriz del sistema el vector de términos independientes:

$$A^* = (A|\mathbf{b}) = \left( \begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{array} \right).$$

Como caso particular de sistemas de ecuaciones lineales cabe destacar los sistemas homogéneos.

**Definición 2.1.2.** Un sistema de ecuaciones lineales se dice que es **homogéneo** si su vector de términos independientes es el vector nulo.

Observemos que un sistema homogéneo siempre admite la solución trivial  $\{x_1 = 0, \dots, x_n = 0\}$ .

## 2.2. Clasificación de los sistemas según el número de soluciones. Teorema de Rouché-Frobenius

Atendiendo a la existencia o no de soluciones de un sistema, y al número de éstas, se da la siguiente clasificación.

**Definición 2.2.1.** Un sistema de ecuaciones lineales se dice que es un **sistema incompatible** si no tiene solución. Por el contrario se dice que es un **sistema compatible** si tiene alguna solución. En este último caso sólo caben dos posibilidades: o bien el sistema tiene una única solución, y en este caso se dice que es un **sistema compatible determinado**, o bien tiene infinitas soluciones, llamándose un **sistema compatible indeterminado**.

Como ejemplo, véase que los sistemas homogéneos siempre son compatibles.

En consecuencia, la solución general de un sistema compatible determinado consistirá en la única solución posible del sistema, mientras que la solución general de un sistema compatible indeterminado vendrá expresada en función de uno o más parámetros. En este último caso, para cada asignación de valores que demos a los parámetros obtendremos una solución concreta del sistema.

Al hecho de decidir si un sistema es incompatible, compatible determinado o compatible indeterminado se le llama comúnmente discutir el sistema. Para discutir un sistema no es necesario resolverlo. Nos basta con estudiar el rango de la matriz del sistema y el de la matriz ampliada y aplicar el siguiente resultado.

**Teorema 2.2.2. Teorema de Rouché-Frobenius.** Sea  $AX = \mathbf{b}$  la representación matricial de un sistema de ecuaciones lineales y sea  $A^* = (A|\mathbf{b})$  la matriz ampliada del sistema. Entonces, el sistema es compatible si y sólo si  $\text{rg}(A) = \text{rg}(A^*)$ . En este caso el sistema es compatible determinado si  $\text{rg}(A)$  coincide con el número de incógnitas, y es compatible indeterminado si  $\text{rg}(A)$  es menor que el número de incógnitas.

Algunas consecuencias inmediatas del teorema 2.2.2 son las siguientes.

1. Un sistema es incompatible si y sólo si  $\text{rg}(A) < \text{rg}(A^*)$  ya que, por ser  $A$  una submatriz de  $A^*$ , la otra desigualdad estricta no es posible.
2. Como el número de ecuaciones del sistema coincide con el número de filas de  $A$  y el número de incógnitas con el número de columnas,  $\text{rg}(A)$  siempre es menor o igual que el número de incógnitas del sistema y que el número de ecuaciones.

En el caso de un sistema compatible indeterminado sabemos que su solución general depende de parámetros. El número de parámetros necesarios es igual a la diferencia entre el número de incógnitas y el rango de la matriz del sistema:

$$\text{n}^\circ \text{ de parámetros} = \text{n}^\circ \text{ de incógnitas} - \text{rg}(A).$$



## 2.3. Métodos de resolución

### 2.3.1. Método de Cramer

A partir del teorema 2.2.2 y de las observaciones que le suceden se sigue que, si tenemos un sistema de ecuaciones lineales con  $n$  incógnitas, el menor número de ecuaciones para que sea un sistema compatible determinado es también  $n$ . En este caso la matriz del sistema  $A$  es cuadrada y, para que se verifique  $\text{rg}(A) = n$ , ha de tener determinante no nulo. Es decir,  $A$  tiene que ser una matriz regular. Los sistemas que cumplen estas características se denominan sistemas de Cramer.

**Definición 2.3.1.** Un sistema de ecuaciones lineales se dice que es un **sistema de Cramer** si es un sistema compatible determinado de  $n$  ecuaciones con  $n$  incógnitas. O equivalentemente, si la matriz del sistema es una matriz regular.

Es posible dar la solución de un sistema de Cramer mediante determinantes. De hecho, un sistema de Cramer genérico viene dado por

$$\left. \begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ \dots & & \dots & & \dots & & \dots & & \dots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & \dots & + & a_{nn}x_n & = & b_n \end{array} \right\},$$

y se puede probar que su solución es

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} & \dots & a_{1n} \\ b_2 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ b_n & a_{n2} & \dots & a_{nn} \end{vmatrix}}{|A|}, \quad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 & \dots & a_{1n} \\ a_{21} & b_2 & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & b_n & \dots & a_{nn} \end{vmatrix}}{|A|}, \quad \dots, \quad x_n = \frac{\begin{vmatrix} a_{11} & a_{12} & \dots & b_1 \\ a_{21} & a_{22} & \dots & b_2 \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & b_n \end{vmatrix}}{|A|}.$$

Obtener la solución de un sistema compatible determinado mediante determinantes es lo que se conoce como método de Cramer. Los principales inconvenientes del método de Cramer son: tiene un alto coste computacional, y solamente se puede aplicar a sistemas de Cramer. Sin embargo, siguiendo la filosofía del método de Gauss que hemos expuesto en el tema 1, podemos discutir y resolver simultáneamente cualquier sistema de ecuaciones lineales.

### 2.3.2. Método de Gauss

Hay una serie de sistemas de ecuaciones lineales muy fáciles de discutir y de resolver, los sistemas escalonados.

**Definición 2.3.2.** Un sistema de ecuaciones lineales es un **sistema escalonado** si la matriz ampliada es una matriz escalonada por filas tal y como se definió en el tema 1.

Si la matriz ampliada del sistema es una matriz escalonada por filas la matriz del sistema también lo será, y por tanto será inmediato hallar y comparar los rangos de ambas matrices. En consecuencia, gracias al teorema de Rouché-Frobenius la discusión del sistema es inmediata.

Con respecto a la resolución:

- Si el sistema es incompatible no hay nada que resolver.
- Si el sistema es compatible determinado se podrá resolver de forma regresiva: en la última ecuación se despeja la última incógnita y se sustituye en la penúltima ecuación. En esta penúltima ecuación se despeja la penúltima incógnita y se sustituye en la anterior, y así sucesivamente hasta llegar a la primera ecuación dónde despejaremos la primera incógnita.
- Si el sistema es compatible indeterminado lo resolveremos de modo similar, pero nos quedarán incógnitas “libres” que no podremos despejar. Estas incógnitas se pueden considerar como los parámetros de la solución general del sistema.

El método de Gauss para la resolución de sistemas consiste en transformar un sistema de ecuaciones lineales en otro equivalente que sea escalonado, con el fin de discutirlo y resolverlo como se acaba de explicar. Para ello, basta con escribir el sistema en forma matricial, y aplicarle transformaciones elementales de filas a la matriz ampliada hasta obtener una matriz escalonada por filas. También podemos eliminar las filas nulas, o las que sean iguales o proporcionales. La matriz que obtengamos será la matriz ampliada correspondiente a un sistema escalonado equivalente al inicial.

Si se prefiere se puede hacer el mismo proceso sin escribir el sistema en forma matricial, manipulando directamente las ecuaciones.

## 2.4. Aplicaciones: método de los mínimos cuadrados

Por definición un sistema incompatible no tiene solución. No obstante, en ocasiones conviene buscar una aproximación, es decir, hallar una o varias asignaciones de valores para las incógnitas que, aunque no satisfagan todas las ecuaciones del sistema simultáneamente, el “error” cometido con ellas sea de algún modo pequeño.

Este planteamiento es interesante porque en múltiples ocasiones los sistemas de ecuaciones lineales surgen como modelos matemáticos sencillos de problemáticas tangibles y reales. Los modelos más realistas suelen tener un alto número de ecuaciones, y la mayoría de las veces obtenemos sistemas incompatibles (por errores de medición, o simplemente porque no hay solución). Algunos de los problemas más clásicos son: problemas de transporte, de abastecimiento, de flujos, de regresión lineal...

Para buscar las “mejores” asignaciones tenemos que especificar qué tipo de error vamos a considerar. Una vez fijado el error sólo tendremos que minimizarlo. El error que vamos a considerar es el mínimo cuadrático. Supongamos que tenemos el sistema (2.1) y que consideramos una asignación de valores  $(x_1, \dots, x_n) = (k_1, \dots, k_n)$ . En un primer intento de definir el error podríamos pensar que el error cometido en cada una de las ecuaciones vendría dado por

$$|a_{11}k_1 + a_{12}k_2 + \dots + a_{1n}k_n - b_1|, \dots, |a_{m1}k_1 + a_{m2}k_2 + \dots + a_{mn}k_n - b_m|.$$

Por tanto, el error total sería su suma:

$$|a_{11}k_1 + a_{12}k_2 + \dots + a_{1n}k_n - b_1| + \dots + |a_{m1}k_1 + a_{m2}k_2 + \dots + a_{mn}k_n - b_m|.$$

Esta función no tiene buenas propiedades analíticas ya que no es derivable, pero tampoco podemos eliminar sin más los valores absolutos porque los errores individuales deben ser positivos para evitar la posibilidad de que puedan compensarse.

Para solventar este problema de derivabilidad, la función que usualmente se considera es la suma de los cuadrados de los errores individuales, es decir:

$$E(x_1, \dots, x_n) = (a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n - b_1)^2 + \dots + (a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n - b_m)^2.$$

Esta función cuadrática es fácil de minimizar siguiendo los métodos estudiados en la asignatura *Matemáticas I*, y su mínimo será la aproximación que buscábamos.

El método que acabamos de describir es el que se conoce como método de los mínimos cuadrados para aproximar un sistema incompatible. Es interesante observar que el método también podría aplicarse al caso de sistemas compatibles, pero entonces obtendríamos como

mínimos de la función error las propias soluciones del sistema. Obsérvese también que la función error es siempre no negativa y se anula en un punto si y sólo si el punto es solución del sistema.

# Capítulo 3

## Espacios vectoriales

### Resumen y motivación.

*Este es el tema central de la parte de álgebra lineal de la asignatura, y el más extenso. Trataremos el tema desde un punto de vista abstracto para empezar a formar la mente de los alumnos en el razonamiento matemático. No obstante, ilustraremos todos los conceptos con ejemplos más sencillos y familiares como son las matrices estudiadas en el tema 1. Presentaremos los conceptos de base, dimensión y coordenadas como conceptos esenciales del tema. Haremos hincapié en el cambio de base, ya que será fundamental en los temas 4 y 5. También profundizaremos en el concepto de subespacio vectorial, estudiando las distintas formas en las que se pueden presentar los subespacios vectoriales y aprendiendo a identificarlos. Presentaremos los espacios vectoriales euclídeos, y haremos ver a los estudiantes hasta que punto suponen un enriquecimiento de la teoría de espacios vectoriales. En particular, introduciremos los conceptos de base ortogonal y ortonormal.*

### Objetivos.

*Al finalizar el tema los alumnos han de entender los espacios vectoriales de dimensión finita y los espacios vectoriales euclídeos como estructuras que generalizan los vectores libres estudiados en bachillerato, comprender que las matrices son un ejemplo particular de espacio vectorial y entender los conceptos de cambio de base y de subespacio vectorial.*

### Estructura del tema.

• Definición y propiedades. Ejemplos • Dependencia e independencia lineal • Conceptos de base y dimensión. Coordenadas y cambio de base • Subespacios vectoriales • Espacios vectoriales euclídeos

### 3.1. Definición y propiedades

El conjunto de todas las matrices de un mismo orden dotado de las operaciones suma y producto por un escalar, tal cual se ha definido en el tema 1, es un ejemplo particular de la estructura de espacio vectorial que definimos a continuación.

**Definición 3.1.1.** Sea  $K$  el conjunto de los números reales  $\mathbb{R}$  o el conjunto de los números complejos  $\mathbb{C}$ , y sea  $V$  un conjunto no vacío. Supongamos que en  $V$  hay definida una operación interna, que denotaremos por  $+$ , y una operación externa, que denotaremos por  $\cdot$ . Diremos que  $(V, +, \cdot)$  es un **espacio vectorial** sobre  $K$  si se verifican las siguientes propiedades:

1. Propiedad asociativa (op. interna):  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ ,  $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ .
2. Propiedad conmutativa:  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ ,  $\forall \mathbf{u}, \mathbf{v} \in V$ .
3. Existencia de elemento neutro:  $\exists \mathbf{0} \in V \mid \mathbf{0} + \mathbf{v} = \mathbf{v}$ ,  $\forall \mathbf{v} \in V$ .
4. Existencia de elemento opuesto:  $\forall \mathbf{v} \in V \exists -\mathbf{v} \in V \mid \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ .
5. Propiedad distributiva I:  $a \cdot (\mathbf{u} + \mathbf{v}) = a \cdot \mathbf{u} + a \cdot \mathbf{v}$ ,  $\forall a \in K$ ,  $\forall \mathbf{u}, \mathbf{v} \in V$ .
6. Propiedad distributiva II:  $(a + b) \cdot \mathbf{v} = a \cdot \mathbf{v} + b \cdot \mathbf{v}$ ,  $\forall a, b \in K$ ,  $\forall \mathbf{v} \in V$ .
7. Propiedad asociativa (op. externa):  $a \cdot (b \cdot \mathbf{v}) = (ab) \cdot \mathbf{v}$ ,  $\forall a, b \in K$ ,  $\forall \mathbf{v} \in V$ .
8. Elemento unidad:  $1 \cdot \mathbf{v} = \mathbf{v}$ ,  $\forall \mathbf{v} \in V$ , donde 1 se corresponde con el número 1 en  $K$ .

Los elementos de  $V$  suelen denominarse **vectores** y los elementos de  $K$ , tal y como hemos indicado en el tema 1, **escalares**. En el caso en que  $K = \mathbb{R}$  diremos que  $(V, +, \cdot)$  es un **espacio vectorial real**, mientras que en el caso en que  $K = \mathbb{C}$  diremos que  $(V, +, \cdot)$  es un **espacio vectorial complejo**.

NOTA: En realidad, en la definición anterior podríamos considerar como  $K$  un conjunto más general, lo que se conoce en álgebra como un cuerpo. Sin embargo, para los propósitos de esta asignatura nos bastará con considerar espacios vectoriales reales o complejos.

Cabe aclarar algunos detalles de la definición anterior:

- A la operación interna  $+$  se le conoce normalmente como **suma**.
- A la operación externa  $\cdot$  se le conoce normalmente como **producto por un escalar**. A menudo omitiremos el símbolo  $\cdot$ .
- Cuando no haya lugar a confusión con respecto a las operaciones consideradas denotaremos el espacio vectorial simplemente por  $V$ .

- Para evitar confusiones escribiremos los vectores en negrita. En particular, observemos que tenemos dos elementos neutros: el vector  $\mathbf{0}$  y el escalar  $0$ , y no debemos confundirlos.

Todo espacio vectorial satisface las siguientes propiedades.

**Proposición 3.1.2.** *Sea  $V$  un espacio vectorial. Entonces se tiene*

1.  $a \cdot \mathbf{0} = \mathbf{0}, \quad \forall a \in K.$
2.  $0 \cdot \mathbf{v} = \mathbf{0}, \quad \forall \mathbf{v} \in V.$
3. Si  $a \cdot \mathbf{v} = \mathbf{0}$  para  $a \in K$  y  $\mathbf{v} \in V$  entonces, o bien  $a = 0$ , o  $\mathbf{v} = \mathbf{0}$ .
4.  $(-1) \cdot \mathbf{v} = -\mathbf{v}, \quad \forall \mathbf{v} \in V.$

### 3.1.1. Ejemplos clásicos

Algunos ejemplos clásicos de espacios vectoriales reales son:

- Dado cualquier  $n \in \mathbb{N}$ , el conjunto

$$\mathbb{R}^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}$$

dotado de las operaciones suma y producto por un escalar usuales.

- El conjunto  $\mathcal{M}_{m \times n}(\mathbb{R})$  de las matrices con coeficientes reales de orden  $m \times n$  dotados de la suma y el producto por un escalar usuales.
- El conjunto  $\mathcal{P}(\mathbb{R})$  de los polinomios de variable real con coeficientes reales con la suma y el producto por un escalar usuales.
- El conjunto  $\mathcal{P}_n(\mathbb{R})$  de los polinomios de grado a lo sumo  $n$  con coeficientes reales con la suma y el producto por un escalar usuales.
- El conjunto  $\mathcal{C}([0, 1])$  de las funciones continuas en el intervalo cerrado  $[0, 1]$  con la suma y el producto por un escalar usuales.

Por otro lado, podríamos considerar  $\mathbb{C}^n = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{C}\}$ , y todos los ejemplos anteriores con coeficientes complejos o en variable compleja según corresponda. Todos estos casos serían ejemplos de espacios vectoriales tanto reales como complejos (dependiendo del conjunto  $K$  del que consideremos los escalares).

Siempre que sea posible la teoría se expondrá en general para espacios vectoriales tanto reales como complejos. No obstante, como ya hemos dicho en el tema 1, para facilitar que los alumnos asimilen la materia, los ejemplos en clase y los ejercicios que se les propongan versarán en su mayoría sobre espacios vectoriales reales. Sólo trabajaremos con espacios vectoriales complejos cuando queramos poner de manifiesto alguna diferencia significativa.

Trabajaremos a menudo con los ejemplos que acabamos de introducir, por lo que mantendremos la notación siempre que sea posible. Mientras no se especifiquen las operaciones a considerar, asumiremos que son la suma y el producto por un escalar usuales en cada uno de los conjuntos que estemos tratando.

Como un último ejemplo de espacio vectorial consideremos un conjunto formado por un único elemento. Se puede dotar a este conjunto de estructura de espacio vectorial (real o complejo). Para ello, el único elemento del conjunto tiene que ser el elemento neutro:  $V = \{\mathbf{0}\}$  y se definen las operaciones suma y producto por un escalar del único modo posible:  $\mathbf{0} + \mathbf{0} = \mathbf{0}$  y  $a \cdot \mathbf{0} = \mathbf{0}$ ,  $\forall a \in K$ . Este ejemplo se conoce como **espacio vectorial trivial**.

## 3.2. Dependencia e independencia lineal

**Definición 3.2.1.** Dado un conjunto de vectores  $\mathbf{v}_1, \dots, \mathbf{v}_k$  de un espacio vectorial  $V$ , llamaremos **combinación lineal** de ellos a cualquier vector de la forma

$$a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k,$$

donde los coeficientes  $a_i$ ,  $i = 1, \dots, k$ , son escalares de  $K$ .

El concepto de combinación lineal nos permite dar la definición de dependencia e independencia lineal que será fundamental a lo largo de toda la asignatura.

**Definición 3.2.2.** Se dice que un conjunto finito de vectores  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  es **linealmente dependiente** si el elemento neutro se puede escribir como combinación lineal de ellos con no todos los coeficientes nulos. Es decir, si existen escalares  $a_1, \dots, a_k$  no todos nulos tales que

$$a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k = \mathbf{0}.$$

Por el contrario, si  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  no es linealmente dependiente, se dice que es un conjunto **linealmente independiente**. Es decir,

$$\text{si } a_1 \mathbf{v}_1 + \dots + a_k \mathbf{v}_k = \mathbf{0} \Rightarrow a_1 = \dots = a_k = 0.$$



Como una consecuencia inmediata de esta definición, observemos que cualquier conjunto de vectores que contenga al elemento neutro es linealmente dependiente.

De la definición también se deduce fácilmente el siguiente resultado.

**Proposición 3.2.3.** 1. Si  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  es un conjunto de vectores linealmente dependiente, entonces cualquier otro conjunto de vectores  $S'$  que lo contenga ( $S \subseteq S'$ ) también es linealmente dependiente.

2. Si  $S = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$  es un conjunto de vectores linealmente independiente, entonces cualquier subconjunto de  $S$  también es linealmente independiente.

Estudiar la independencia o dependencia lineal de un conjunto finito de vectores es equivalente a estudiar si un determinado sistema homogéneo es compatible determinado o indeterminado, respectivamente.

Una caracterización equivalente de la dependencia lineal es la que sigue.

**Proposición 3.2.4.** Un conjunto de vectores  $\{\mathbf{v}_1, \dots, \mathbf{v}_k\}$  es linealmente dependiente si y sólo si podemos expresar uno de los vectores del conjunto como combinación lineal del resto.

NOTA: Que podamos expresar uno de los vectores como combinación lineal del resto no significa que podamos expresar cualquier vector como combinación lineal del resto.

Es interesante también remarcar que en la definición de combinación lineal no sólo intervienen los vectores, sino también los escalares. Si consideramos un mismo conjunto como espacio vectorial real y como espacio vectorial complejo el concepto de dependencia e independencia lineal puede variar.

### 3.3. Conceptos de base y dimensión

**Definición 3.3.1.** Se dice que  $V$  es un **espacio vectorial de dimensión finita** si existe un conjunto finito de vectores  $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\} \subseteq V$  de modo que cualquier otro vector  $\mathbf{v} \in V$  se pueda expresar como una combinación lineal de los elementos de  $S$ . En este caso, se dice que el conjunto  $S$  es un **sistema de generadores** o un **sistema generador** de  $V$ .

De ahora en adelante, y salvo que se indique lo contrario, trabajaremos con espacios vectoriales de dimensión finita. La teoría se puede extender a espacios vectoriales que no admitan un conjunto finito de generadores, pero esto excede de los contenidos de la asignatura. Diremos que los espacios para los que se da esta situación son **espacios vectoriales de dimensión infinita**.

Un sistema generador  $S$  de  $V$  no tiene porqué ser un conjunto linealmente independiente. No obstante, siempre podemos extraer un subconjunto de  $S$  que sea linealmente independiente. Basta con quitar, uno a uno, los vectores que se puedan expresar como combinación lineal del resto. Además, puesto que los vectores que eliminamos de esta manera se pueden expresar como combinación lineal del resto, es fácil razonar que el subconjunto linealmente independiente con el que nos vamos a quedar también será un sistema generador de  $V$ .

Por otro lado, un conjunto de vectores linealmente independiente de  $V$  no es necesariamente un sistema generador de  $V$ . Sin embargo, podemos completar el conjunto añadiendo uno a uno vectores que sean linealmente independientes con los anteriores, hasta obtener un sistema generador de  $V$ . Son estos conjuntos los que nos van a interesar.

**Definición 3.3.2.** Dado un espacio vectorial  $V$ , un subconjunto de vectores  $B \subseteq V$  se dice que es una **base** de  $V$  si es simultáneamente un conjunto linealmente independiente y un sistema generador de  $V$ .

Las bases de un espacio vectorial no son únicas.

**Teorema 3.3.3. Teorema de la base.** Si un espacio vectorial  $V$  tiene una base formada por un número finito de vectores, entonces todas las bases de  $V$  son finitas y tienen el mismo número de vectores. En este caso llamaremos **dimensión** de  $V$ , y lo denotaremos por  $\dim(V)$ , al número de vectores de cualquiera de sus bases.

Vamos a describir ahora las bases *más sencillas* de algunos de los espacios vectoriales clásicos. Cada una de estas bases se conoce como **base estándar** o **base canónica**:

- Para el espacio vectorial real  $\mathbb{R}^n$ ,  $B = \{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}$ .
- Para el espacio vectorial real  $\mathbb{C}^n$ ,  
 $B = \{(1, 0, \dots, 0), (i, 0, \dots, 0), (0, 1, \dots, 0), (0, i, \dots, 0), \dots, (0, 0, \dots, 1), (0, 0, \dots, i)\}$ .
- Para el espacio vectorial complejo  $\mathbb{C}^n$ ,  $B = \{(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1)\}$ .
- Para el espacio vectorial real  $\mathcal{P}_n(\mathbb{R})$ ,  $B = \{1, x, x^2, \dots, x^n\}$ .
- Para el espacio vectorial real  $\mathcal{M}_{m \times n}(\mathbb{R})$ ,

$$B = \left\{ \begin{pmatrix} 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}_{m \times n}, \begin{pmatrix} 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{pmatrix}_{m \times n}, \dots, \begin{pmatrix} 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}_{m \times n} \right\}.$$

Como consecuencia  $\dim(\mathbb{R}^n) = n$ ,  $\dim(\mathcal{P}_n(\mathbb{R})) = n + 1$  y  $\dim(\mathcal{M}_{m \times n}(\mathbb{R})) = m \times n$ . Y, en cuanto a  $\mathbb{C}^n$ , tendremos que su dimensión real es  $2n$ ,  $\dim_{\mathbb{R}}(\mathbb{C}^n) = 2n$ , y su dimensión compleja es  $n$ ,  $\dim_{\mathbb{C}}(\mathbb{C}^n) = n$ .

Por otro lado, es fácil probar que  $\mathcal{P}(\mathbb{R})$  y  $\mathcal{C}([0, 1])$  son espacios vectoriales reales de dimensión infinita.

Consideremos, por último, el espacio vectorial trivial  $V = \{\mathbf{0}\}$ . Ya hemos visto que cualquier conjunto de vectores que contenga al elemento neutro es linealmente dependiente. Por tanto,  $V$  no tiene ningún subconjunto linealmente independiente. En consecuencia  $\dim(V) = 0$  y no tiene sentido hablar de base, no existe.

Recordemos que hemos dicho que, dado un espacio vectorial  $V$ , cualquier subconjunto  $S \subset V$  de vectores linealmente independientes se puede extender a una base de  $V$ . Y, por otro lado, cualquier sistema generador se puede reducir también a una base de  $V$ . Además, por el teorema de la base todas las bases de  $V$  tienen el mismo número de elementos.

**Proposición 3.3.4.** *Sea  $V$  un espacio vectorial de dimensión  $n$ . Entonces:*

1. *Todo conjunto de  $n$  vectores linealmente independientes de  $V$  es una base de  $V$ .*
2. *Todo sistema generador de  $V$  de  $n$  elementos es una base de  $V$ .*

### 3.3.1. Coordenadas y cambio de base

Dado un sistema generador  $S$  de  $V$  y un vector cualquiera  $\mathbf{v} \in V$ , en general no hay una única manera de expresar  $\mathbf{v}$  como combinación lineal de los elementos de  $S$ . Sin embargo, esta combinación lineal es única en el caso en que  $S$  sea una base de  $V$ .

**Proposición 3.3.5.** *Sea  $V$  un espacio vectorial. Si  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  es una base de  $V$ , entonces cualquier vector  $\mathbf{v} \in V$  se puede escribir de forma única como una combinación lineal de los vectores de  $B$ . Es decir, existe una única elección de escalares  $a_1, \dots, a_n \in K$  de modo que*

$$\mathbf{v} = a_1 \mathbf{v}_1 + \dots + a_n \mathbf{v}_n.$$

*A los coeficientes de esta combinación lineal se les llama **coordenadas** de  $\mathbf{v}$  con respecto a la base  $B$ , y se denota como*

$$\mathbf{v} = (a_1, \dots, a_n)_B.$$

De la definición de coordenadas dada en la proposición 3.3.5 se observa que, si  $V$  es un espacio vectorial de dimensión  $n$  y  $B$  es una base cualquiera de  $V$ , entonces cualquier vector  $\mathbf{v} \in V$  se puede identificar mediante sus coordenadas con respecto a  $B$  con un vector de  $K^n$  (es decir,  $\mathbb{R}^n$  o  $\mathbb{C}^n$  según corresponda). Esto nos va a permitir, una vez expresados todos los

vectores en función de una misma base, trabajar en cualquier espacio vectorial de dimensión finita como si lo estuviéramos haciendo en  $K^n$ .

Las coordenadas simplifican las operaciones con vectores. De hecho, sea  $V$  un espacio vectorial de dimensión  $n$ , sea  $B$  una base de  $V$ , y sean  $\mathbf{u}$  y  $\mathbf{v}$  dos vectores de  $V$ . Entonces, si  $\mathbf{u} = (a_1, \dots, a_n)_B$  y  $\mathbf{v} = (b_1, \dots, b_n)_B$  se tiene que:

- $\mathbf{u} + \mathbf{v} = (a_1 + b_1, \dots, a_n + b_n)_B$ .
- $k\mathbf{u} = (ka_1, \dots, ka_n)_B, \forall k \in K$ .

Las coordenadas también nos permiten decidir sobre la dependencia o independencia lineal de un conjunto de vectores.

**Teorema 3.3.6.** *Sea  $V$  un espacio vectorial y sea  $B$  una base cualquiera de  $V$ . Un conjunto de vectores  $\{\mathbf{u}_1, \dots, \mathbf{u}_r\}$  es linealmente independiente si y sólo si la matriz cuyas columnas (o filas) son las coordenadas de cada uno de estos vectores con respecto a la base  $B$  tiene rango  $r$ .*

Nótese que, por lo general, las expresiones en coordenadas de un vector con respecto a dos bases distintas serán también distintas. Veamos como se relacionan ambas coordenadas. Para ello, sea  $V$  un espacio vectorial de dimensión  $n$  y sean  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  y  $B' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_n\}$  dos bases de  $V$ . Dado un vector  $\mathbf{v} \in V$  podemos considerar sus coordenadas con respecto a ambas bases:

$$\mathbf{v} = (a_1, \dots, a_n)_B = (a'_1, \dots, a'_n)_{B'}.$$

Por otro lado, por ser  $B$  una base, podemos expresar cada uno de los vectores de  $B'$  en coordenadas con respecto a  $B$ ,  $\mathbf{v}'_j = (p_{1j}, \dots, p_{nj})_B, \forall 1 \leq j \leq n$ . De aquí se llega fácilmente a que

$$\mathbf{v} = \sum_{j=1}^n a'_j \mathbf{v}'_j = \sum_{j=1}^n a'_j \left( \sum_{i=1}^n p_{ij} \mathbf{v}_i \right) = \sum_{i=1}^n \left( \sum_{j=1}^n p_{ij} a'_j \right) \mathbf{v}_i.$$

Como conclusión, por la unicidad de coordenadas con respecto a la misma base tenemos que

$$a_i = \sum_{j=1}^n p_{ij} a'_j, \forall 1 \leq i \leq n. \text{ Escrito en forma matricial tenemos}$$

$$\begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}_B = \begin{pmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \dots & \dots & \dots & \dots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{pmatrix} \begin{pmatrix} a'_1 \\ a'_2 \\ \dots \\ a'_n \end{pmatrix}_{B'}.$$

La matriz  $P_{B'}^B = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}$  se conoce como matriz cambio de base de  $B'$  a  $B$ .

Observemos que se trata de la matriz cuyas columnas son las coordenadas de los vectores de la base  $B'$  en función de la base  $B$ .

Análogamente se puede deducir el cambio de base de  $B$  a  $B'$ . Así, de modo similar obtendríamos que  $\begin{pmatrix} a'_1 \\ a'_2 \\ \cdots \\ a'_n \end{pmatrix}_{B'} = P_B^{B'} \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}_B$  donde  $P_B^{B'}$  es la matriz cambio de base de  $B$  a  $B'$ . Es decir, la matriz cuyas columnas son las coordenadas de los vectores de la base  $B$  en función de la base  $B'$ .

Se puede comprobar que las matrices cambio de base siempre son regulares, y que además se verifica  $P_B^{B'} = (P_{B'}^B)^{-1}$ .

### 3.4. Subespacios vectoriales

**Definición 3.4.1.** Sea  $V$  un espacio vectorial y sea  $U$  un subconjunto no vacío de  $V$ . Decimos que  $U$  es un **subespacio vectorial** de  $V$  si  $U$  es en sí mismo un espacio vectorial con las operaciones inducidas de  $V$ . Lo denotaremos como  $U \leq V$ .

Una caracterización equivalente de los subespacios vectoriales es la siguiente.

**Proposición 3.4.2.** Sea  $V$  un espacio vectorial y sea  $U$  un subconjunto no vacío de  $V$ . Entonces, son equivalentes:

- $U$  es un subespacio vectorial de  $V$ .
- Dados dos vectores cualesquiera  $\mathbf{u}, \mathbf{v} \in U$  y dos escalares  $a, b \in K$  se tiene que  $a\mathbf{u} + b\mathbf{v} \in U$ .

Todo espacio vectorial no trivial  $V$  tiene al menos dos subespacios vectoriales distintos:  $U = V$  y  $U = \{\mathbf{0}\}$ . Estos dos subespacios se conocen como **subespacios impropios** mientras que el resto de subespacios intermedios se llaman **subespacios propios**.

Uno de los modos más habituales y sencillos de obtener subespacios vectoriales es a partir de un subconjunto cualquiera de  $V$ .

$$L(S) = \{a_1 \mathbf{s}_1 + \dots + a_n \mathbf{s}_n; n \in \mathbb{N}, a_i \in K, \mathbf{s}_i \in S, i = 1, \dots, n\}.$$

Se puede demostrar que  $L(S)$  es el menor subespacio vectorial de  $V$  que contiene a  $S$ .  
 do cualquier subespacio vectorial  $U \neq \emptyset$ , podemos encontrar un conjunto  $S \subset V$  de modo  
 e  $U = L(S)$ . Diremos en este caso que  $S$  es un **conjunto de generadores** de  $U$ .

### 3.4.1. Ecuaciones de un subespacio

$$\left. \begin{array}{l} x_1 = a_{11} \lambda_1 + a_{12} \lambda_2 + \dots + a_{1r} \lambda_r \\ x_2 = a_{21} \lambda_1 + a_{22} \lambda_2 + \dots + a_{2r} \lambda_r \\ \dots \quad \dots \quad \dots \quad \dots \\ x_n = a_{n1} \lambda_1 + a_{n2} \lambda_2 + \dots + a_{nr} \lambda_r \end{array} \right\} \quad (3.1)$$

El sistema de ecuaciones (3.1) se conoce como **ecuaciones paramétricas** de  $U$  respecto a la base  $B$ . A partir de estas ecuaciones, eliminando los parámetros, vemos que también podemos obtener todos los elementos de  $U$ ,  $\mathbf{x} = (x_1, \dots, x_n)_B$ , como el conjunto de soluciones de un sistema homogéneo de  $n$  incógnitas y  $n - r$  ecuaciones. A este sistema homogéneo se

le conoce como **ecuaciones cartesianas** de  $U$  respecto a  $B$ . Tanto unas ecuaciones como las otras nos proporcionan las condiciones necesarias para que un determinado vector de  $V$  pertenezca al subespacio  $U$ . Hay que tener en cuenta que las ecuaciones cartesianas no son únicas.

### 3.4.2. Operaciones entre dos subespacios

Sea  $V$  un espacio vectorial y sean  $U$  y  $W$  dos subespacios vectoriales de  $V$ . Es posible realizar ciertas operaciones entre  $U$  y  $W$  para obtener nuevos subespacios.

**Proposición 3.4.4.** *La **intersección**  $U \cap W$  es un subespacio vectorial de  $V$ .*

Sin embargo, es fácil ver que la unión  $U \cup W$  no es, por lo general, un subespacio vectorial. No obstante, podemos considerar el menor subespacio vectorial que contenga a  $U \cup W$ .

**Definición 3.4.5.** El menor subespacio que contiene a  $U \cup W$  se denomina **suma** de los subespacios  $U$  y  $W$  y se denota por  $U + W$ .

En el caso particular en que  $U \cap W = \{0\}$  se dice que la suma es una **suma directa** y se denota por  $U \oplus W$ .

Se puede comprobar que  $U + W = \{\mathbf{u} + \mathbf{w} \mid \mathbf{u} \in U, \mathbf{w} \in W\}$ .

Por lo general la descomposición de un vector  $\mathbf{v} \in U + W$  como suma de un vector de  $U$  y otro de  $W$  no es única, pero sí lo es en el caso en que la suma sea directa.

La suma directa nos permite definir el concepto de subespacio complementario (o suplementario).

**Proposición 3.4.6.** *Sea  $V$  un espacio vectorial y sea  $U$  un subespacio vectorial de  $V$ . Entonces, siempre existe otro subespacio vectorial de  $V$ ,  $W$ , de modo que  $V = U \oplus W$ . En este caso se dice que  $W$  es un **subespacio complementario (o suplementario)** de  $U$ .*

El subespacio complementario (o suplementario) a uno dado no es único.

Por último, cabe destacar la siguiente relación entre las dimensiones de los subespacios  $U$ ,  $W$ ,  $U + W$  y  $U \cap W$ .

**Proposición 3.4.7.** *Sea  $V$  un espacio vectorial y sean  $U$  y  $W$  dos subespacios vectoriales de  $V$ , entonces:*

$$\dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W).$$

NOTA: De modo similar se podrían definir las operaciones entre un número arbitrario de subespacios vectoriales.

### 3.5. Espacios vectoriales euclídeos

A lo largo de esta sección,  $V$  va a representar un espacio vectorial real. En esta sección queremos ver cómo podemos enriquecer la estructura de  $V$ . Para ello introduciremos el concepto de producto escalar y de espacio vectorial euclídeo.

**Definición 3.5.1.** Sea  $V$  un espacio vectorial real. Un **producto escalar** en  $V$  es una aplicación

$$\langle, \rangle : V \times V \rightarrow \mathbb{R}$$

verificando las siguientes propiedades:

1. Simetría:  $\langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ ,  $\forall \mathbf{u}, \mathbf{v} \in V$ .
2. Bilinealidad:  $\langle a\mathbf{u} + b\mathbf{v}, \mathbf{w} \rangle = a\langle \mathbf{u}, \mathbf{w} \rangle + b\langle \mathbf{v}, \mathbf{w} \rangle$ ,  $\forall a, b \in \mathbb{R}$ ,  $\forall \mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ .
3. Definido positivo:  $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ ,  $\forall \mathbf{v} \in V$ , y  $\langle \mathbf{v}, \mathbf{v} \rangle = 0$  si y sólo si  $\mathbf{v} = \mathbf{0}$ .

**Definición 3.5.2.** Un **espacio vectorial euclídeo** es un par  $(V, \langle, \rangle)$  formado por un espacio vectorial real  $V$  y un producto escalar definido en él.

Hay que tener en cuenta que, si para un mismo espacio vectorial real consideramos distintos productos escalares, obtendremos distintos espacios vectoriales euclídeos.

El producto escalar de un espacio euclídeo de dimensión finita se puede expresar matricialmente como sigue. Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo de dimensión finita, sea  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  una base de  $V$ , y consideremos  $c_{ij} = \langle \mathbf{v}_i, \mathbf{v}_j \rangle$ ,  $1 \leq i, j \leq n$ . Se llama **matriz de Gram** de  $(V, \langle, \rangle)$  con respecto de la base  $B$  a la matriz

$$G = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \dots & c_{nn} \end{pmatrix}.$$

Consideremos ahora  $\mathbf{u}, \mathbf{v} \in V$ , e identifiquemos estos vectores con las matrices filas que forman sus coordenadas con respecto a  $B$ , es decir:  $\mathbf{u} = (a_1, \dots, a_n)_B$  y  $\mathbf{v} = (b_1, \dots, b_n)_B$ . Entonces el producto escalar se puede expresar como el siguiente producto de matrices:

$$\langle \mathbf{u}, \mathbf{v} \rangle = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} G \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

A esta expresión es a la que se le conoce como **forma o expresión matricial del producto escalar**.



### 3.5.1. Norma de un vector

**Definición 3.5.3.** Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y sea  $\mathbf{v} \in V$ . Se define la **norma** o **módulo** de  $\mathbf{v}$  como

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

De las definiciones de norma y de producto escalar se deducen las siguientes propiedades para la norma de un vector.

**Proposición 3.5.4.** Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y sean  $\mathbf{v} \in V$  y  $a \in \mathbb{R}$ . Entonces,

1.  $\|\mathbf{v}\| \geq 0$  y  $\|\mathbf{v}\| = 0$  si y sólo si  $\mathbf{v} = 0$ .
2.  $\|a\mathbf{v}\| = |a|\|\mathbf{v}\|$ .

Se dice que un vector  $\mathbf{v}$  es **unitario** si tiene norma 1. Podemos obtener un vector unitario proporcional a uno dado dividiéndolo por su norma.

Las siguientes desigualdades son dos propiedades no tan inmediatas pero muy importantes.

**Proposición 3.5.5.** Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y sean  $\mathbf{u}, \mathbf{v} \in V$ . Entonces,

- Desigualdad de Cauchy-Schwarz:  $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$ .
- Desigualdad triangular:  $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ .

### 3.5.2. Ángulo entre dos vectores. Vectores ortogonales

La desigualdad de Cauchy-Schwarz nos permite definir el ángulo entre dos vectores.

**Definición 3.5.6.** Dado  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y dados  $\mathbf{u}, \mathbf{v} \in V$ , se llama **ángulo** entre los vectores  $\mathbf{u}$  y  $\mathbf{v}$  al único número real  $0 \leq \theta \leq \pi$  de forma que

$$\cos \theta = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{u}\|\|\mathbf{v}\|}.$$

Dos vectores  $\mathbf{u}, \mathbf{v} \in V$  se dice que son **ortogonales**, y se denota por  $\mathbf{u} \perp \mathbf{v}$ , si el ángulo que forman es  $\frac{\pi}{2}$ , o equivalentemente si  $\langle \mathbf{u}, \mathbf{v} \rangle = 0$ . Esto nos permite definir unas bases con características especiales que las hacen más operativas.

**Definición 3.5.7.** Dado un espacio vectorial euclídeo  $(V, \langle, \rangle)$ , una base  $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  se dice que es

- una **base ortogonal** si los vectores que la forman son ortogonales dos a dos.
- una **base ortonormal** si es ortogonal y además todos los vectores que la forman son unitarios.

Estas bases nos permiten, por ejemplo, obtener una expresión sencilla de las coordenadas de un vector con respecto a una de ellas.

**Proposición 3.5.8.** Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y sea  $B = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$  una base ortogonal. Entonces, las coordenadas de un vector  $\mathbf{v} \in V$  con respecto a  $B$ ,  $\mathbf{v} = (v_1, \dots, v_n)_B$ , vienen dadas por

$$v_i = \frac{\langle \mathbf{v}, \mathbf{e}_i \rangle}{\|\mathbf{e}_i\|^2}.$$

Siempre podemos construir una base ortonormal a partir de una base dada. Para ello hay que aplicar el proceso conocido como **método de Gram-Schmidt** descrito a continuación.

**Teorema 3.5.9.** Sea  $(V, \langle, \rangle)$  un espacio vectorial euclídeo y sea  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  una base cualquiera de  $V$ . Entonces existe una base ortonormal de  $V$ ,  $B' = \{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ , de forma que  $L(\mathbf{v}_1, \dots, \mathbf{v}_k) = L(\mathbf{e}_1, \dots, \mathbf{e}_k)$  para cualquier  $1 \leq k \leq n$ .

Los vectores de  $B'$  se construyen según el siguiente algoritmo. En primer lugar construimos una base ortogonal de  $V$ ,  $\{\mathbf{e}'_1, \dots, \mathbf{e}'_n\}$ , mediante las siguientes expresiones:

$$\begin{aligned} \mathbf{e}'_1 &= \mathbf{v}_1 \\ \mathbf{e}'_2 &= \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{e}'_1 \rangle}{\|\mathbf{e}'_1\|^2} \mathbf{e}'_1 \\ &\dots \\ \mathbf{e}'_n &= \mathbf{v}_n - \frac{\langle \mathbf{v}_n, \mathbf{e}'_1 \rangle}{\|\mathbf{e}'_1\|^2} \mathbf{e}'_1 - \dots - \frac{\langle \mathbf{v}_n, \mathbf{e}'_{n-1} \rangle}{\|\mathbf{e}'_{n-1}\|^2} \mathbf{e}'_{n-1}. \end{aligned}$$

Basta ahora con dividir cada uno de estos vectores por su módulo para obtener los vectores de la base ortonormal,  $\mathbf{e}_i = \frac{\mathbf{e}'_i}{\|\mathbf{e}'_i\|}$ ,  $\forall 1 \leq i \leq n$ .

### 3.5.3. Complemento ortogonal

Se dice que un vector  $\mathbf{v} \in V$  es ortogonal a un subespacio vectorial  $U \leq V$ , y se denota  $\mathbf{v} \perp U$ , si es ortogonal a todos los vectores de  $U$ , es decir, si  $\mathbf{v} \perp \mathbf{u}$ ,  $\forall \mathbf{u} \in U$ . Si consideramos el conjunto de todos los vectores ortogonales a  $U$  obtendremos un nuevo subespacio vectorial.

**Proposición 3.5.10.** Sea  $(V, \langle \cdot, \cdot \rangle)$  un espacio vectorial euclídeo de dimensión finita y sea  $U$  un subespacio de  $V$ . Entonces, el conjunto

$$U^\perp = \{\mathbf{v} \in V \mid \mathbf{v} \perp U\}$$

es un subespacio vectorial de  $V$ . Además,  $V = U \oplus U^\perp$  y, en particular,

$$\dim(U^\perp) = \dim(V) - \dim(U).$$

Llamamos a  $U^\perp$  el **complemento ortogonal** de  $U$ .

Obsérvese que  $(U^\perp)^\perp = U$ .

El complemento ortogonal nos permite definir la proyección ortogonal de un vector sobre un cierto subespacio. Como  $V = U \oplus U^\perp$ , cualquier vector  $\mathbf{v} \in V$  se puede escribir de forma única como  $\mathbf{v} = \mathbf{u} + \mathbf{w}$  con  $\mathbf{u} \in U$  y  $\mathbf{w} \in U^\perp$ . En esta situación diremos que  $\mathbf{u}$  es la **proyección ortogonal** de  $\mathbf{v}$  sobre  $U$  y lo denotaremos por  $\mathbf{u} = p_U(\mathbf{v})$ . Análogamente, el vector  $\mathbf{w}$  será la proyección ortogonal de  $\mathbf{v}$  sobre  $U^\perp$ ,  $\mathbf{w} = p_{U^\perp}(\mathbf{v})$ .

### 3.5.4. El producto vectorial en $\mathbb{R}^3$

Para finalizar el tema, consideremos el espacio euclídeo ordinario, o espacio euclídeo estándar,  $\mathbb{R}^3$ . Es decir, consideremos el producto escalar sobre  $\mathbb{R}^3$  que hace que el conjunto  $B = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$  sea una base ortonormal, es decir, el producto escalar usual.

En este contexto vamos a dar la definición y las principales propiedades del producto vectorial de dos vectores de  $\mathbb{R}^3$ .

**Definición 3.5.11.** Dados dos vectores  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ , se define el producto vectorial de  $\mathbf{u}$  y  $\mathbf{v}$  como el único vector  $\mathbf{u} \wedge \mathbf{v} \in \mathbb{R}^3$  tal que

$$\langle \mathbf{u} \wedge \mathbf{v}, \mathbf{w} \rangle = \det(\mathbf{u}, \mathbf{v}, \mathbf{w}), \quad \forall \mathbf{w} \in \mathbb{R}^3.$$

En concreto, si  $\mathbf{u} = (u_1, u_2, u_3)_B$  y  $\mathbf{v} = (v_1, v_2, v_3)_B$  se tiene que

$$\mathbf{u} \wedge \mathbf{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \mathbf{e}_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \mathbf{e}_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \mathbf{e}_3$$

siendo  $\mathbf{e}_1 = (1, 0, 0)$ ,  $\mathbf{e}_2 = (0, 1, 0)$  y  $\mathbf{e}_3 = (0, 0, 1)$ .

El producto vectorial verifica las siguientes propiedades. Dados tres vectores  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$  y dos escalares  $a, b \in \mathbb{R}$  se verifican las siguientes propiedades:

1. Antisimetría:  $\mathbf{u} \wedge \mathbf{v} = -\mathbf{v} \wedge \mathbf{u}$ .
2. Bilinealidad:  $(a\mathbf{u} + b\mathbf{w}) \wedge \mathbf{v} = a\mathbf{u} \wedge \mathbf{v} + b\mathbf{w} \wedge \mathbf{v}$ .
3. No-asociatividad:  $(\mathbf{u} \wedge \mathbf{v}) \wedge \mathbf{w} = \langle \mathbf{u}, \mathbf{w} \rangle \mathbf{v} - \langle \mathbf{v}, \mathbf{w} \rangle \mathbf{u} \neq \mathbf{u} \wedge (\mathbf{v} \wedge \mathbf{w})$ .
4.  $\mathbf{u} \wedge \mathbf{v} = 0$  si y sólo si  $\mathbf{u}$  y  $\mathbf{v}$  son linealmente dependientes.
5.  $\mathbf{u} \wedge \mathbf{v}$  es ortogonal a  $\mathbf{u}$  y a  $\mathbf{v}$ .
6.  $\|\mathbf{u} \wedge \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \sin^2 \theta$ , siendo  $\theta$  el ángulo formado por los vectores  $\mathbf{u}$  y  $\mathbf{v}$ .

# Capítulo 4

## Aplicaciones lineales

### Resumen y motivación.

*Estudiamos en este tema el concepto de aplicación lineal entre espacios vectoriales como una generalización del concepto de función lineal ya estudiado en la asignatura Matemáticas I. Veremos la correspondencia que existe entre las aplicaciones lineales y las matrices. Definiremos los conceptos de núcleo e imagen de una aplicación lineal, y estudiaremos la caracterización de los monomorfismos y epimorfismos (es decir, aplicaciones lineales inyectivas y sobreyectivas respectivamente) a partir del núcleo y de la imagen. Estudiaremos el cambio de base para aplicaciones lineales. Una vez comprendido el concepto de cambio de base para espacios vectoriales no debería resultar complicado entenderlo en este contexto.*

### Objetivos.

*Tras estudiar esta unidad didáctica, los alumnos han de comprender el concepto de aplicación lineal como una generalización de las funciones lineales, han de entender la relación existente entre las aplicaciones lineales y las matrices, han de conocer los conceptos de imagen y núcleo de una aplicación lineal y han de saber obtener dichos subconjuntos.*

### Estructura del tema.

• Definición y propiedades. Ejemplos • Expresión matricial de una aplicación lineal • Núcleo e imagen de una aplicación lineal • Cambio de base para aplicaciones lineales

Tal y como hemos dicho en el tema 3, supondremos en este tema, y en el resto de la asignatura, que todos los espacios vectoriales son de dimensión finita.

## 4.1. Definición y propiedades. Ejemplos

Una aplicación lineal es una función o aplicación entre espacios vectoriales que preserva la estructura de espacio vectorial, es decir, que se comporta bien con respecto a las operaciones suma y producto por un escalar.

**Definición 4.1.1.** Dados  $V$  y  $V'$  dos espacios vectoriales, una aplicación  $f : V \rightarrow V'$  se dice que es una **aplicación lineal** si para cualesquiera  $\mathbf{u}, \mathbf{v} \in V$  y  $a, b \in K$  se tiene

- $f(\mathbf{u} + \mathbf{v}) = f(\mathbf{u}) + f(\mathbf{v})$ .
- $f(a\mathbf{v}) = af(\mathbf{v})$ .

De la definición de aplicación lineal se siguen las siguientes propiedades inmediatas.

**Proposición 4.1.2.** Sea  $f : V \rightarrow V'$  una aplicación lineal entre dos espacios vectoriales  $V$  y  $V'$ . Entonces,

1.  $f(\mathbf{0}) = \mathbf{0}$ .
2.  $f(-\mathbf{v}) = -f(\mathbf{v})$ ,  $\forall \mathbf{v} \in V$ .
3.  $f(a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n) = a_1f(\mathbf{v}_1) + \dots + a_nf(\mathbf{v}_n)$ ,  $\forall \mathbf{v}_1, \dots, \mathbf{v}_n \in V$ ,  $\forall a_1, \dots, a_n \in K$ .

Los dos ejemplos más sencillos de aplicaciones lineales son:

- La **aplicación identidad** de un espacio vectorial  $V$  en sí mismo:

$$\begin{aligned} I : V &\rightarrow V \\ I(\mathbf{v}) &= \mathbf{v}. \end{aligned}$$

- La **aplicación nula, cero o trivial** entre dos espacios vectoriales  $V$  y  $V'$ :

$$\begin{aligned} 0 : V &\rightarrow V' \\ 0(\mathbf{v}) &= \mathbf{0}. \end{aligned}$$

### 4.1.1. Monomorfismos, epimorfismos e isomorfismos

Las aplicaciones lineales reciben denominaciones especiales según las propiedades que tengan.

**Definición 4.1.3.** Sea  $f : V \rightarrow V'$  una aplicación lineal entre dos espacios vectoriales  $V$  y  $V'$ :

- Si  $V = V'$  se dice que  $f$  es un **endomorfismo**.
- Si  $f$  es una aplicación inyectiva se dice que es un **monomorfismo**.
- Si  $f$  es una aplicación sobreyectiva se dice que es un **epimorfismo**.
- Si  $f$  es una aplicación biyectiva se dice que es un **isomorfismo**.
- Si  $f$  es simultáneamente un endomorfismo y un isomorfismo se dice que es un **automorfismo**.

Estas aplicaciones se pueden caracterizar según el comportamiento de conjuntos de vectores linealmente independientes, de sistemas generadores, o de bases.

**Proposición 4.1.4.** *Dada  $f : V \rightarrow V'$  una aplicación lineal entre dos espacios vectoriales  $V$  y  $V'$  se verifica que:*

1.  *$f$  es un monomorfismo si y sólo si para cualquier conjunto linealmente independiente de vectores de  $V$   $\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ , el conjunto  $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_r)\}$  es linealmente independiente en  $V'$ .*
2.  *$f$  es un epimorfismo si y sólo si para cualquier sistema generador de  $V$   $\{\mathbf{v}_1, \dots, \mathbf{v}_s\}$ , el conjunto  $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_s)\}$  es un sistema generador de  $V'$ .*
3.  *$f$  es un isomorfismo si y sólo si para cualquier base de  $V$   $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ ,  $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)\}$  es una base de  $V'$ .*

Por otro lado, se dice que dos espacios vectoriales son **isomorfos** si existe un isomorfismo entre ellos. Por la proposición anterior dos espacios isomorfos tienen la misma dimensión, pero es más, comparten la misma estructura y las mismas propiedades. De hecho, es fácil probar que cualquier espacio vectorial de dimensión finita  $n$  es isomorfo al espacio vectorial  $K^n$  sobre  $K$ . Por tanto, podemos suponer sin pérdida de generalidad que, siempre que consideremos un espacio de dimensión finita, estaremos trabajando en  $K^n$  (bien sea  $\mathbb{R}^n$  o  $\mathbb{C}^n$ ).

### 4.1.2. Operaciones entre aplicaciones lineales

Podemos realizar varias operaciones entre aplicaciones lineales de modo que el resultado sea una nueva aplicación lineal.

**Proposición 4.1.5.** *Sean  $f, g : V \rightarrow V'$  dos aplicaciones lineales entre dos espacios vectoriales  $V$  y  $V'$  y sea  $a$  un escalar. Entonces, las operaciones suma y producto por un escalar,*

$$\begin{array}{lcl} f + g : V \rightarrow V' & & af : V \rightarrow V' \\ (f + g)(\mathbf{v}) = f(\mathbf{v}) + g(\mathbf{v}) & \text{y} & (af)(\mathbf{v}) = af(\mathbf{v}) \end{array},$$

son también aplicaciones lineales.

Lo mismo ocurre con la composición de aplicaciones lineales (cuando ésta es posible).

**Proposición 4.1.6.** Sean  $V$ ,  $V'$  y  $V''$  tres espacios vectoriales, y sean  $f : V \rightarrow V'$  y  $g : V' \rightarrow V''$  dos aplicaciones lineales entre ellos. Entonces la composición  $g \circ f : V \rightarrow V''$  también es una aplicación lineal.

Y, por último, con la aplicación inversa.

**Proposición 4.1.7.** Sea  $f : V \rightarrow V'$  un isomorfismo entre espacios vectoriales, entonces su aplicación inversa  $f^{-1} : V' \rightarrow V$  también es un isomorfismo.

## 4.2. Expresión matricial de una aplicación lineal

Como ya hemos indicado en la sección anterior, cualquier espacio vectorial de dimensión finita es isomorfo a un cierto  $K^n$ . Supongamos, por tanto, que estamos trabajando únicamente en estos espacios.

Sea  $f : K^n \rightarrow K^m$  una aplicación lineal y sea  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  una base de  $K^n$ . Entonces  $f$  queda completamente determinada por las imágenes de los vectores de la base  $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)$ . En efecto, consideremos  $\mathbf{v} \in K^n$  y supongamos que sus coordenadas con respecto a  $B$  son  $\mathbf{v} = (a_1, \dots, a_n)_B$ . Gracias a las propiedades de las aplicaciones lineales se tiene que

$$f(\mathbf{v}) = a_1 f(\mathbf{v}_1) + \dots + a_n f(\mathbf{v}_n).$$

Sea ahora  $B' = \{\mathbf{v}'_1, \dots, \mathbf{v}'_m\}$  una base de  $K^m$ . Podemos expresar los vectores  $f(\mathbf{v}_1), \dots, f(\mathbf{v}_n) \in K^m$  en coordenadas con respecto a  $B'$ :

$$f(\mathbf{v}_i) = (c_{1i}, \dots, c_{mi})_{B'}, \quad \forall 1 \leq i \leq n.$$

Y, en consecuencia,

$$f(\mathbf{v}) = (c_{11}a_1 + c_{12}a_2 + \dots + c_{1n}a_n, \dots, c_{m1}a_1 + c_{m2}a_2 + \dots + c_{mn}a_n)_{B'}.$$

Escrito en forma matricial obtenemos que las coordenadas de  $f(\mathbf{v})$  con respecto a la base  $B'$  vienen dadas por

$$f(\mathbf{v}) = \begin{pmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \dots & c_{mn} \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}_B.$$



A la matriz  $Mf_B^{B'} = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{m1} & \cdots & c_{mn} \end{pmatrix}$  le llamaremos matriz asociada a la aplicación lineal  $f$  con respecto a las bases  $B$  y  $B'$  de  $K^n$  y  $K^m$  respectivamente. Esta matriz nos permite obtener las coordenadas con respecto a la base  $B'$  de la imagen de un vector de  $K^n$  dado en coordenadas con respecto a la base  $B$ . Las columnas de la matriz  $Mf_B^{B'}$  son las coordenadas de las imágenes de cada uno de los vectores de la base  $B$  con respecto a la base  $B'$ .

Las operaciones entre aplicaciones lineales se traducen en operaciones análogas en sus expresiones matriciales.

**Proposición 4.2.1.** Sean  $f, g : K^n \rightarrow K^m$  dos aplicaciones lineales y sea  $a \in K$ . Consideremos  $B$  y  $B'$  bases de  $K^n$  y  $K^m$  respectivamente. Entonces:

- Si  $h = f + g : K^n \rightarrow K^m$ , entonces  $Mh_B^{B'} = Mf_B^{B'} + Mg_B^{B'}$ .
- Si  $h = af : K^n \rightarrow K^m$ , entonces  $Mh_B^{B'} = aMf_B^{B'}$ .

**Proposición 4.2.2.** Sean  $f : K^n \rightarrow K^m$  y  $g : K^m \rightarrow K^l$  dos aplicaciones lineales y sean  $B, B'$  y  $B''$  bases de  $K^n, K^m$  y  $K^l$  respectivamente. Sea  $h = g \circ f : K^n \rightarrow K^l$ , entonces

$$Mh_B^{B''} = Mg_{B'}^{B''} \cdot Mf_B^{B'}.$$

**Proposición 4.2.3.** Sea  $f : K^n \rightarrow K^n$  un isomorfismo y sean  $B$  y  $\bar{B}$  dos bases cualesquiera de  $K^n$ . Entonces, si  $h = f^{-1} : K^n \rightarrow K^n$

$$Mh_B^{\bar{B}} = (Mf_B^{\bar{B}})^{-1}.$$

## 4.3. Núcleo e imagen de una aplicación lineal

A toda aplicación lineal se le puede asociar dos conjuntos relevantes: su núcleo y su imagen.

**Definición 4.3.1.** Sea  $f : K^n \rightarrow K^m$  una aplicación lineal. Entonces se definen su **núcleo**,  $Ker(f)$ , y su **imagen**,  $Im(f)$ , como los subconjuntos

$$Ker(f) = \{\mathbf{v} \in K^n : f(\mathbf{v}) = \mathbf{0}\} \subseteq K^n \quad \text{e} \quad Im(f) = \{f(\mathbf{v}) : \mathbf{v} \in K^n\} \subseteq K^m,$$

respectivamente.

Se puede probar que ambos conjuntos tienen estructura de espacio vectorial.

**Proposición 4.3.2.** *Dada una aplicación lineal  $f : K^n \rightarrow K^m$ ,  $\text{Ker}(f)$  e  $\text{Im}(f)$  son subespacios vectoriales de  $K^n$  y  $K^m$ , respectivamente.*

Para obtener el núcleo de una aplicación lineal basta con resolver un sistema homogéneo de ecuaciones. En cuanto a la imagen, el siguiente resultado es muy útil para su cálculo.

**Lema 4.3.3.** *Sea  $f : K^n \rightarrow K^m$  una aplicación lineal. Si  $\{\mathbf{v}_1, \dots, \mathbf{v}_p\}$  es un sistema generador de  $K^n$ , entonces  $\{f(\mathbf{v}_1), \dots, f(\mathbf{v}_p)\}$  es un sistema generador de  $\text{Im}(f)$ .*

El núcleo y la imagen también nos proporcionan una nueva caracterización para los monomorfismos y los epimorfismos.

**Proposición 4.3.4.** *Sea  $f : K^n \rightarrow K^m$  una aplicación lineal. Entonces,*

- *$f$  es un monomorfismo si y sólo si  $\text{Ker}(f) = \{\mathbf{0}\}$ .*
- *$f$  es un epimorfismo si y sólo si  $\text{Im}(f) = K^m$ .*

La matriz asociada a una aplicación lineal facilita sobremanera el cálculo de su núcleo y de su imagen. De hecho, dada  $f : K^n \rightarrow K^m$  una aplicación lineal y dadas  $B$  y  $B'$  dos bases de  $K^n$  y  $K^m$  respectivamente, es inmediato comprobar que las columnas de la matriz asociada  $Mf_B^{B'}$  se corresponden con las coordenadas de las imágenes de los  $n$  vectores de la base  $B$  mediante  $f$  con respecto a la base  $B'$ , y que estos vectores son un sistema generador de  $\text{Im}(f)$ . En particular, se deduce que

$$\dim(\text{Im}(f)) = \text{rg}(Mf_B^{B'}).$$

Por otro lado, para calcular el núcleo de la aplicación lineal  $f$  podemos considerar el sistema homogéneo

$$Mf_B^{B'} \cdot \mathbf{x} = \mathbf{0},$$

donde  $\mathbf{x} = (x_1, \dots, x_n)$  es el vector de incógnitas. Las soluciones de este sistema serán los elementos del núcleo escritos en coordenadas con respecto a la base  $B$ . Recordando del tema 2 el modo de obtener los parámetros de los que dependía un sistema compatible indeterminado, deducimos que

$$\dim(\text{Ker}(f)) = n - \text{rg}(Mf_B^{B'}).$$

Como consecuencia, destacamos la siguiente relación entre las dimensiones de  $K^n$ ,  $\text{Im}(f)$  y  $\text{Ker}(f)$ .

**Proposición 4.3.5.** *Sea  $f : K^n \rightarrow K^m$  una aplicación lineal, entonces*

$$n = \dim(K^n) = \dim(\text{Im}(f)) + \dim(\text{Ker}(f)).$$

## 4.4. Cambio de base para aplicaciones lineales

Las matrices asociadas a una misma aplicación lineal con respecto a bases diferentes no son iguales pero comparten varias propiedades interesantes. Por ejemplo, de la sección anterior deducimos que tienen el mismo rango. Aún es más, están relacionadas del modo que veremos ahora.

Sea  $f : K^n \rightarrow K^m$  una aplicación lineal y sean  $B$  y  $\bar{B}$  dos bases de  $K^n$  y  $B'$  y  $\bar{B}'$  dos bases de  $K^m$ . Asociadas a estas bases y a la aplicación lineal podemos hallar  $Mf_B^{B'}$  y  $Mf_{\bar{B}}^{\bar{B}'}$ . Por otro lado, en el tema 3 estudiamos la matriz cambio de base en un espacio vectorial. Podemos obtener por tanto  $P_B^{\bar{B}}$  y  $P_{B'}^{\bar{B}'}$ . Es fácil razonar que estas cuatro matrices están relacionadas mediante la expresión

$$Mf_{\bar{B}}^{\bar{B}'} = P_{B'}^{\bar{B}'} Mf_B^{B'} P_B^{\bar{B}}.$$

De hecho, si hacemos un diagrama de esta situación con la aplicación lineal y las cuatro matrices implicadas, vemos que la relación anterior se deriva de una composición de aplicaciones (véase la Figura 4.1).

$$\begin{array}{ccc}
 f : K^n & \longrightarrow & K^m \\
 & \searrow Mf_B^{B'} & \nearrow \\
 B & \xrightarrow{\quad} & B' \\
 P_B^{\bar{B}} \uparrow & & \uparrow P_{B'}^{\bar{B}'} \\
 \bar{B} & \xrightarrow{\quad} & \bar{B}'
 \end{array}$$

Figura 4.1: Cambio de base para una aplicación lineal.

Recordemos del tema anterior que las matrices cambio de base son matrices regulares. En general, si dadas dos matrices  $A$  y  $C$  existen dos matrices regulares  $P$  y  $Q$  tales que  $C = PAQ$ , se dice que  $A$  y  $C$  son dos **matrices equivalentes**. Por tanto, todas las matrices asociadas a una misma aplicación lineal son matrices equivalentes.

NOTA: Volviendo a lo estudiado en el tema 1, se puede probar que dos matrices son equivalentes si y sólo si una se puede obtener a partir de transformaciones por filas y/o por columnas de la otra. En particular, las matrices equivalentes por filas o por columnas son matrices equivalentes.

Supongamos ahora que realizamos el mismo razonamiento a partir de un endomorfismo  $f : K^n \rightarrow K^n$ . En este caso podríamos considerar las mismas bases en el espacio de partida

y en el de llegada. El diagrama quedaría como la Figura 4.2.

$$\begin{array}{ccc}
 f : K^n & \longrightarrow & K^n \\
 & & \\
 B & \xrightarrow{Mf_B^B} & B \\
 \uparrow P_{\bar{B}}^B & & \uparrow P_B^{\bar{B}} \\
 \bar{B} & \xrightarrow{Mf_{\bar{B}}^{\bar{B}}} & \bar{B}
 \end{array}$$

Figura 4.2: Cambio de base para un endomorfismo.

y tendríamos que

$$Mf_{\bar{B}}^{\bar{B}} = P_B^{\bar{B}} Mf_B^B P_{\bar{B}}^B = (P_B^{\bar{B}})^{-1} Mf_B^B P_{\bar{B}}^B.$$

En este caso la relación es algo más fuerte. Dadas dos matrices  $A$  y  $C$ , si existe una matriz regular  $P$  tal que  $C = P^{-1}AP$ , se dice que  $A$  y  $C$  son **matrices semejantes**. Esto es lo que ocurre con las matrices  $Mf_{\bar{B}}^{\bar{B}}$  y  $Mf_B^B$ .

En general, dado un endomorfismo  $f : K^n \rightarrow K^n$ , nos referiremos a  $Mf_B^B$  como la matriz asociada a  $f$  con respecto a la base  $B$  (daremos por hecho que la base es la misma para el espacio de partida que para el de llegada).

## Capítulo 5

---

# Diagonalización de matrices y endomorfismos

---

### Resumen y motivación.

*En este tema planteamos la diagonalización de endomorfismos, entendiendo ésta como la búsqueda del cambio de base que proporcione la expresión más “sencilla” para el endomorfismo. Veremos que es equivalente a la diagonalización de matrices. Introduciremos todos los elementos necesarios para diagonalizar una matriz, y obtendremos una caracterización de las matrices diagonalizables. Por último, estudiaremos la diagonalización por congruencia para el caso de las matrices simétricas. Los conceptos estudiados en este tema serán de utilidad para la resolución de sistemas de ecuaciones diferenciales que se estudiará en la asignatura Matemáticas III.*

---

### Objetivos.

*Al finalizar el tema es imprescindible que los alumnos entiendan que es equivalente hablar de diagonalización de matrices que de diagonalización de endomorfismos y comprendan su utilidad, que aprendan a calcular los valores y vectores propios de una matriz cuadrada, que sepan decidir cuándo una matriz es diagonalizable y en ese caso sepan calcular la matriz diagonal semejante y una matriz de paso, y que comprendan el concepto de cadena de Markov y sepan trabajar con ellas.*

---

### Estructura del tema.

• Introducción • Vectores y valores propios. Definición y propiedades • Polinomio y ecuación característica • Matrices diagonalizables. Caracterización • Diagonalización de matrices simétricas reales • Aplicaciones. Ecuaciones en diferencias, cadenas de Markov

## 5.1. Introducción

En el tema 4 estudiamos las aplicaciones lineales y, en particular, los endomorfismos. Vimos que la matriz asociada a un endomorfismo es una matriz cuadrada, y que las matrices asociadas a un mismo endomorfismo son matrices semejantes. Nuestro propósito ahora es: dado un endomorfismo  $f : V \rightarrow V$  pretendemos hallar una base para la cual la expresión de  $f$  sea lo más “sencilla” posible. O, dicho de otra manera, dada una matriz cuadrada queremos encontrar, de entre todas sus matrices semejantes, la más “sencilla”. Falta explicitar que queremos decir con “sencilla”. Con ello nos referiremos a que el endomorfismo (o la matriz) tenga forma diagonal. En concreto, podemos dar las siguientes definiciones.

**Definición 5.1.1.** Se dice que un endomorfismo  $f : V \rightarrow V$  es **diagonalizable** si existe una base  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\} \subset V$  de modo que  $f(\mathbf{v}_i) = \lambda_i \mathbf{v}_i$ ,  $\forall 1 \leq i \leq n$ , y para ciertos escalares  $\lambda_i$ .

**Definición 5.1.2.** Se dice que una matriz cuadrada  $A$  es **diagonalizable** si es semejante a una matriz diagonal  $D$ . Es decir, si existe una matriz regular  $P$  de modo que  $D = P^{-1}AP$ . Esta matriz  $P$  se conoce como **matriz de paso**.

El principal objetivo de este tema es determinar cuando un determinado endomorfismo (o una matriz cuadrada) es diagonalizable. Además, si la respuesta es afirmativa buscaremos un método para hallar la base que diagonalice el endomorfismo (o la matriz de paso que diagonalice la matriz).

Se prueba fácilmente que estudiar la diagonalización de un endomorfismo es equivalente a estudiar la diagonalización de cualquiera de sus matrices asociadas. Para simplificar, estudiaremos la diagonalización de matrices cuadradas.

De nuevo trabajaremos en este tema con espacios vectoriales del tipo  $K^n$ .

## 5.2. Vectores y valores propios. Definición y propiedades

Dada una matriz cuadrada  $A$  de orden  $n$  con coeficientes en  $K$ , pueden existir vectores  $\mathbf{v} \in K^n$  no nulos y escalares  $\lambda \in K$  tales que  $A\mathbf{v} = \lambda\mathbf{v}$ . Estos elementos, los vectores y los escalares, serán fundamentales en el desarrollo del tema.

**Definición 5.2.1.** Sea  $A$  una matriz cuadrada de orden  $n$  con coeficientes en  $K$ . Se dice que un escalar  $\lambda \in K$  es un **valor propio** de  $A$  (o **autovalor**) si existe un vector  $\mathbf{v} \in K^n$

no nulo de modo que  $A\mathbf{v} = \lambda\mathbf{v}$ . Todo vector no nulo que satisfaga esta relación se denomina **vector propio** de  $A$  (o **autovector**) asociado al valor propio  $\lambda$ .

Obsérvese que la condición  $\mathbf{v} \neq \mathbf{0}$  es necesaria para evitar el caso trivial: cualquier número real  $\lambda$  verifica la condición  $A\mathbf{0} = \lambda\mathbf{0}$ . Por tanto  $\mathbf{v} = \mathbf{0}$  no se considera vector propio. Por el contrario, los valores propios sí pueden tomar el valor 0.

Los valores y vectores propios tienen las siguientes propiedades básicas:

**Propiedades:**

1. Todo vector propio está asociado a un único valor propio.
2. Un valor propio tiene asociados infinitos vectores propios. De hecho, se puede demostrar que el conjunto de todos los vectores propios asociados a un mismo valor propio  $\lambda$  junto con el vector nulo  $\mathbf{0}$  tiene estructura de subespacio vectorial de  $K^n$ . En particular, si  $\mathbf{v}$  es un vector propio asociado a  $\lambda$ , cualquier múltiplo suyo  $a\mathbf{v}$  ( $a \neq 0$ ,  $a \in K$ ) también lo será.
3. Si  $A$  es una matriz de orden  $n$ , existen a lo sumo  $n$  valores propios distintos de  $A$ .

Teniendo en cuenta la segunda propiedad, podemos dar la siguiente definición.

**Definición 5.2.2.** Sea  $A$  una matriz cuadrada de orden  $n$  y sea  $\lambda$  un valor propio de  $A$ . El conjunto de todos los vectores propios  $\mathbf{v}$  asociados a un mismo valor propio  $\lambda$  junto con el vector nulo  $\mathbf{v} = \mathbf{0}$  se denomina **subespacio propio** de  $A$  asociado a  $\lambda$  y se denota como  $E_\lambda$ , es decir:

$$E_\lambda = \{\mathbf{v} \in K^n : A\mathbf{v} = \lambda\mathbf{v}\}.$$

De las propiedades 1 y 2, y de las definiciones de valores y vectores propios, se deducen los siguientes resultados.

**Proposición 5.2.3.** Cada subespacio propio  $E_\lambda$  tiene estructura de subespacio vectorial de  $K^n$ , y además  $\dim(E_\lambda) = n - \text{rg}(A - \lambda I)$ .

**Proposición 5.2.4.** Si  $\lambda_1$  y  $\lambda_2$  son dos autovalores distintos de  $A$ , entonces  $E_{\lambda_1} \cap E_{\lambda_2} = \{\mathbf{0}\}$ . En particular, vectores propios asociados a valores propios distintos son siempre linealmente independientes.

Obsérvese que de este último resultado se sigue la propiedad 3 antes descrita.

### 5.3. Polinomio y ecuación característica

Pretendemos ahora aprender a calcular los valores y los vectores propios de una matriz cuadrada.

Por definición, un escalar  $\lambda$  es un valor propio de  $A$  si y sólo si existe un vector no nulo  $\mathbf{v}$  de modo que

$$A\mathbf{v} = \lambda\mathbf{v}.$$

Es decir, si se verifica

$$A\mathbf{v} - \lambda\mathbf{v} = (A - \lambda I)\mathbf{v} = 0. \quad (5.1)$$

Observemos que (5.1) no es más que un sistema de ecuaciones lineales homogéneo, y estamos diciendo que  $\lambda$  es un valor propio de  $A$  si y sólo si el sistema es compatible indeterminado. A su vez esto ocurrirá si y sólo si  $\text{rg}(A - \lambda I) < n$  y, como se trata de una matriz cuadrada, esto es equivalente a comprobar que  $\det(A - \lambda I) = 0$ .

Resumiendo,  $\lambda$  es un valor propio de  $A$  si y sólo si  $\det(A - \lambda I) = 0$ . Por tanto, el procedimiento a seguir para hallar los valores propios de una matriz cuadrada  $A$  de orden  $n$  es considerar  $\lambda$  como un parámetro arbitrario y calcular

$$\det(A - \lambda I).$$

El determinante anterior es un polinomio de orden  $n$  en la variable  $\lambda$ , y sus raíces en  $K$  serán los valores propios de  $A$  que estamos buscando.

**Definición 5.3.1.** La matriz  $A - \lambda I$  se llama **matriz característica** de  $A$ . El polinomio  $p_A(\lambda) = \det(A - \lambda I)$  se conoce como **polinomio característico** de  $A$ . Si  $\lambda$  es un valor propio de  $A$ , entonces se llama **multiplicidad algebraica de  $\lambda$**  a la multiplicidad de  $\lambda$  como raíz del polinomio característico.

Una vez calculados los valores propios de  $A$  estamos en condiciones de hallar sus subespacios propios, y por tanto sus vectores propios asociados. Sea  $\lambda$  un valor propio de  $A$ , el subespacio propio asociado a  $\lambda$ ,  $E_\lambda$ , es el conjunto de todas las soluciones del sistema compatible indeterminado (5.1). La dimensión de  $E_\lambda$  se conoce como **multiplicidad geométrica de  $\lambda$** .

Es importante destacar que para cualquier valor propio  $\lambda$  siempre se verifica que

$$1 \leq \text{multiplicidad geométrica de } \lambda \leq \text{multiplicidad algebraica de } \lambda. \quad (5.2)$$



## 5.4. Matrices diagonalizables. Caracterización

Una vez presentados todas las herramientas necesarias, estamos en condiciones de presentar el resultado principal del tema.

**Teorema 5.4.1.** *Una matriz  $A$  cuadrada de orden  $n$  es diagonalizable si y sólo si tiene  $n$  vectores propios linealmente independientes. En este caso, la matriz diagonal  $D$  a la que es semejante es aquella cuyos elementos diagonales son los valores propios de  $A$ . Por último, la matriz de paso  $P$  es la matriz cuyas columnas son los  $n$  vectores propios linealmente independientes.*

Como los valores propios son las raíces de un polinomio de orden  $n$  y se verifica la desigualdad (5.2), una caracterización equivalente de las matrices diagonalizables viene dada por la siguiente proposición.

**Proposición 5.4.2.** *Una matriz  $A$  cuadrada de orden  $n$  es diagonalizable si y sólo si la suma de las multiplicidades algebraicas de las raíces de su polinomio característico es  $n$  y para cualquier valor propio coincide la multiplicidad algebraica con la geométrica. En particular, si la matriz tiene  $n$  valores propios distintos, entonces es diagonalizable.*

NOTA: Si estamos trabajando en  $\mathbb{C}$ , la suma de las multiplicidades algebraicas de las raíces del polinomio característico será siempre  $n$ . Sin embargo en  $\mathbb{R}$  podría no ser así.

Si ahora volvemos a pensar en la diagonalización de endomorfismos, tendremos que un endomorfismo es diagonalizable si y sólo si lo es su matriz asociada con respecto a cualquier base. Y, en este caso, la base que diagonaliza el endomorfismo es la formada por los vectores propios.

## 5.5. Diagonalización de matrices simétricas reales

Consideremos ahora un espacio vectorial euclídeo  $(V, \langle, \rangle)$ , aunque lo denotaremos simplemente por  $V$  para simplificar. Por definición  $V$  es un espacio vectorial real, luego supondremos que es  $\mathbb{R}^n$  para algún  $n \in \mathbb{N}$ . En este contexto damos la siguiente definición.

**Definición 5.5.1.** Un endomorfismo  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  se dice que es **simétrico** si verifica

$$\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f(\mathbf{v}) \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n.$$

Es fácil probar que un endomorfismo simétrico admite matrices simétricas asociadas.

**Proposición 5.5.2.** *Sea  $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$  un endomorfismo y sea  $B$  una base ortonormal de  $\mathbb{R}^n$ . Entonces,  $f$  es un endomorfismo simétrico si y sólo si su matriz asociada  $Mf_B^B$  es simétrica.*

Veamos en esta sección que las matrices simétricas reales, y en consecuencia los endomorfismos simétricos, siempre son diagonalizables y, que además, en este caso obtenemos propiedades adicionales especiales.

En primer lugar se verifica el siguiente resultado.

**Proposición 5.5.3.** *Sea  $A$  un matriz simétrica real de orden  $n$ . Entonces todas las raíces de su polinomio característico son reales.*

Ya tenemos asegurada la primera condición para garantizar la diagonalización de  $A$ . Veamos que la segunda también se verifica.

**Proposición 5.5.4.** *Dada una matriz simétrica real  $A$ , la multiplicidad algebraica de cualquier valor propio coincide con su multiplicidad geométrica. Es decir, si  $\lambda_1, \dots, \lambda_r$  son los autovalores de  $A$ , entonces  $\mathbb{R}^n = E_{\lambda_1} \oplus \dots \oplus E_{\lambda_r}$ .*

Por tanto, las matrices simétricas reales son siempre diagonalizables, lo que deriva en el resultado más importante de esta sección.

**Teorema 5.5.5. Teorema espectral.** *Toda matriz simétrica real es diagonalizable en  $\mathbb{R}$ .*

Pero, además, las matrices simétricas se pueden diagonalizar de un modo especial: por congruencia o semejanza ortogonal.

**Definición 5.5.6.** Dos matrices cuadradas  $A$  y  $C$  se dice que son **matrices congruentes** si existe una matriz regular  $P$  de modo que  $C = P^t A P$ .

Veamos que toda matriz simétrica real es congruente a su matriz diagonal semejante. En concreto, veamos que podemos obtener una matriz de paso  $P$  ortogonal. Necesitamos para ello la siguiente caracterización de las matrices ortogonales.

**Proposición 5.5.7.** *Una matriz cuadrada real  $P$  de orden  $n$  es ortogonal si y sólo si sus filas (o columnas) forman una base ortonormal de  $\mathbb{R}^n$ .*

Teniendo esto en cuenta, lo que tenemos que probar es que podemos obtener una base ortonormal de vectores propios. Para ello, es fácil probar primero que los subespacios propios son ortogonales dos a dos.

**Lema 5.5.8.** *Sea  $A$  una matriz simétrica real, y sean  $\lambda_1$  y  $\lambda_2$  dos valores propios distintos de  $A$ . Entonces, los correspondientes subespacios propios  $E_{\lambda_1}$  y  $E_{\lambda_2}$  son ortogonales. Es decir, para cualesquiera  $\mathbf{u} \in E_{\lambda_1}$  y  $\mathbf{v} \in E_{\lambda_2}$ , se tiene  $\mathbf{u} \perp \mathbf{v}$ .*

Una vez probado que los subespacios propios son ortogonales dos a dos, nos basta con encontrar una base ortonormal de cada uno de ellos mediante el procedimiento de Gram-Schmidt expuesto en el tema 3.

En resumen, podemos enunciar ya el siguiente resultado principal de esta sección.

**Teorema 5.5.9.** *Toda matriz simétrica real admite una matriz de paso ortogonal.*

Aún es más, se puede probar que este resultado es, en realidad, una caracterización de las matrices simétricas reales.

**Proposición 5.5.10.** *Una matriz cuadrada real  $A$  es una matriz simétrica si y sólo si es una matriz diagonalizable que admite una matriz de paso ortogonal.*

Por último, las matrices simétricas reales admiten una descomposición como suma de matrices más sencillas con propiedades interesantes. Dicha descomposición se obtiene a partir de sus valores y vectores propios, y se recoge en el siguiente resultado.

**Teorema 5.5.11. Teorema de descomposición espectral.** *Sea  $A$  una matriz simétrica real. Sea  $B = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$  una base ortonormal de  $\mathbb{R}^n$  formada por vectores propios de  $A$  y sean  $\lambda_1, \dots, \lambda_n$  los correspondientes valores propios. Entonces,*

$$A = \lambda_1 A_1 + \dots + \lambda_n A_n,$$

siendo  $A_i = \mathbf{v}_i \mathbf{v}_i^t$ ,  $1 \leq i \leq n$ , (entendiendo cada vector  $\mathbf{v}_i$  como un vector columna).

Obsérvese que, por construcción, las matrices  $A_i$  de la descomposición anterior son matrices simétricas, que además verifican las siguientes propiedades:

$$A^2 = A \text{ (matriz idempotente)}, \quad A_1 + \dots + A_n = I \quad \text{y} \quad A_i A_j = O, \quad \forall i \neq j.$$

## 5.6. Aplicaciones de la diagonalización

### 5.6.1. Cálculo de potencias de una matriz diagonalizable

El producto de dos matrices, estudiado en el tema 1, es una operación sencilla, pero que puede llegar a resultar tediosa por el alto número de sumas y productos por un escalar que

involucra. De hecho, es relativamente común cometer errores en dichos cálculos. Por tanto, realizar una potencia de una matriz directamente puede ser un proceso largo y delicado con un alto coste computacional (especialmente si el exponente es grande).

Sin embargo, hay un caso de matrices para las que el cálculo de las potencias resulta más sencillo: las matrices diagonales. La potencia de una matriz diagonal es otra matriz diagonal del mismo orden obtenida al reemplazar cada término de la diagonal por su respectiva potencia. Esto nos permite simplificar el cálculo de las potencias de una matriz diagonalizable.

**Proposición 5.6.1.** *Sea  $A$  una matriz diagonalizable, entonces*

$$A^n = PD^nP^{-1},$$

*siendo  $D$  una matriz diagonal semejante a  $A$ , y  $P$  una matriz de paso asociada.*

### 5.6.2. Ecuaciones en diferencias. Cadenas de Markov

Una ecuación en diferencias es una ecuación determinada mediante una relación de recurrencia. En concreto,

**Definición 5.6.2.** Dados  $a_1, \dots, a_p \in \mathbb{R}$ ,  $a_p \neq 0$ , una **ecuación en diferencias lineal y homogénea de orden  $p$**  es una relación de recurrencia del tipo

$$z_n = a_1 z_{n-1} + \dots + a_p z_{n-p}, \quad \forall n > p, \quad n \in \mathbb{N}. \quad (5.3)$$

Observemos que esta ecuación en diferencias nos permite, dados  $z_1, \dots, z_p \in \mathbb{R}$ , definir una sucesión de números reales  $\{z_n\}_{n \in \mathbb{N}}$ . El problema que nos planteamos es acerca de la existencia de una expresión explícita para  $z_n$ .

A partir de (5.3) podemos considerar el sistema de  $p$  ecuaciones

$$\left. \begin{array}{rcll} z_n & = & a_1 z_{n-1} + \dots + a_{p-1} z_{n-p+1} + a_p z_{n-p} \\ z_{n-1} & = & z_{n-1} \\ \dots & & \dots & \dots \\ z_{n-p+1} & = & & z_{n-p+1} \end{array} \right\}.$$

Aunque no se trata de un sistema de ecuaciones lineales, sino de un sistema de ecuaciones en diferencias, se parece bastante. De hecho, si lo escribimos en forma matricial tenemos

$$\mathbf{u}_n = A\mathbf{u}_{n-1}, \quad (5.4)$$

siendo

$$\mathbf{u}_n = \begin{pmatrix} z_n \\ z_{n-1} \\ \dots \\ z_{n-p+1} \end{pmatrix} \quad \forall n > p, \quad \text{y} \quad A = \begin{pmatrix} a_1 & \dots & a_{p-1} & a_p \\ 1 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & 0 \end{pmatrix}.$$

Puesto que  $z_n$  es la primera coordenada del vector  $\mathbf{u}_n$ , si obtenemos una expresión explícita de dicho vector ya tendremos lo que buscamos. Aplicando sucesivamente la relación de recurrencia (5.4) se tiene

$$\mathbf{u}_n = A\mathbf{u}_{n-1} = A^2\mathbf{u}_{n-2} = \dots = A^n\mathbf{u}_0,$$

por lo que, conociendo los primeros  $p$  datos, el problema se reduce al cálculo de las potencias de una matriz cuadrada, que ya hemos estudiado en el caso en que la matriz  $A$  sea diagonalizable.

Otra cuestión interesante es si la ecuación en diferencias (5.3) tiene un estado estacionario. Es decir, ¿qué ocurre si hacemos tender  $n$  a infinito? Hay un tipo particular de ecuaciones en diferencias y de matrices asociadas que garantizan la estabilidad.

**Definición 5.6.3.** Se dice que una matriz cuadrada real  $A$  es una **matriz de Markov** o **matriz estocástica** si todos sus elementos son no negativos, y si además la suma de los elementos de cada una de sus columnas es igual a 1.

**Definición 5.6.4.** Una **cadena de Markov** es un sistema de ecuaciones en diferencias de la forma (5.4), para el que la matriz  $A$  es una matriz de Markov.

Las matrices de Markov tienen propiedades interesantes. Sea  $A$  una matriz de Markov, entonces:

1.  $A^n$  también es una matriz de Markov para cualquier  $n \in \mathbb{N}$ .
2. 1 es un valor propio de  $A$ .
3. Todo valor propio  $\lambda$  de  $A$  satisface  $|\lambda| \leq 1$ .

Gracias a estas propiedades se puede demostrar que si tenemos una cadena de Markov el valor de  $z_n$  está siempre acotado. Y además, en el caso en que la matriz  $A$  no admita a  $-1$  como valor propio, existe el límite cuando hacemos tender  $n$  a infinito.