Universidade Vigo

Álgebra Lineal

E.S.E. Informática

Curso 2016-2017

María Begoña Cid Iglesias

Índice General

El nuevo concepto de enseñanza y aprendizaje que se introduce con las nuevas titulaciones de Grado hace que cambie la forma tradicional de impartir la clase por el profesor y de recibirla por el alumnado. No se trata sólo de presentar unos conocimientos, sino que hay que trabajarlos para que estos se incorporen al saber y saber hacer de los alumnos.

Estas notas tan sólo pretenden facilitar este proceso para que los alumnos tengan a mano los distintos ámbitos de conocimiento que se pretenden abordar en esta asignatura y surgen como una recopilación de las notas elaboradas para impartir la asignatura de Álgebra Lineal en la Escuela de Ingeniería Informática de la Universidad de Vigo.

Quiero agradecer la inestimable colaboración de mi colega Eduardo Liz.

Espero que estas notas sean útiles a todos sus posibles lectores.

Tema 1

Estructuras y números complejos

1.1 Operaciones internas y estructura de cuerpo

Definición 1.1 *Una* **operación interna** $*$ *en un conjunto* A *es una correspondencia que asigna a cada par de elementos* $a, b \in A$ *un elemento* $c = a * b \in A$.

Consideraremos dos tipos de operaciones internas, que denotaremos por suma $(+)$ y producto $(\cdot)$. Si A es un conjunto con una o dos operaciones internas, A puede tener distintas estructuras según las propiedades que cumplan estas operaciones. Consideraremos las siguientes propiedades:

1. Propiedad **asociativa**: $(a * b) * c = a * (b * c), \forall a, b, c \in A$. Esta propiedad permite operar más de dos elementos. En este caso escribiremos simplemente $a * b * c$.

2. **Elemento neutro**: Se dice que $(A, *)$ tiene elemento neutro si existe $e \in A$ tal que
$$a * e = e * a = a, \forall a \in A.$$
En la suma, el elemento neutro se llama cero (0) y, en general, en el producto se llama uno (1). El elemento neutro, si existe, es único.

3. **Elemento simétrico**: Se dice que $a \in A$ tiene elemento simétrico si existe $a' \in A$ tal que $a * a' = a' * a = e$. En el caso de la suma, el elemento simétrico se llama elemento **opuesto** y se denota por $-a$, $(a + (-a) = (-a) + a = 0)$. En el caso del producto, se llama **inverso** y se denota por a^{-1}, $(a \cdot a^{-1} = a^{-1} \cdot a = 1)$.

4. Propiedad **conmutativa**: $a * b = b * a, \forall a, b \in A$. Si en una operación producto se cumple la propiedad conmutativa entonces el elemento inverso se suele denotar por $1/a$.

5. Propiedad **distributiva**. Si A tiene definida una suma y un producto, se dice que el producto es distributivo con respecto a la suma si
$$a \cdot (b + c) = a \cdot b + a \cdot c$$
$$(a + b) \cdot c = a \cdot c + b \cdot c,$$
para todo $a, b, c \in A$.

Definición 1.2 *Se dice que un conjunto con una operación interna* $(A, *)$ *es un* **grupo conmutativo** *si cumple las propiedades asociativa y conmutativa, tiene elemento neutro y todo elemento tiene simétrico.*

Ejemplos de grupos conmutativos son $(\mathbb{R}, +), (\mathbb{C}, +), (\mathbb{R} \setminus \{0\}, \cdot)$ y $(\mathbb{C} \setminus \{0\}, \cdot)$.

Observación 1.1 Si B es un subconjunto de A, se denota $A \setminus B = \{x \in A/x \notin B\}$. En particular, si $a \in A, A \setminus a = \{x \in A/x \neq a\}$.

Definición 1.3 *Se dice que un conjunto con dos operaciones internas* $(A, +, \cdot)$ *es un* **cuerpo conmutativo** *si* $(A, +)$ *y* $(A \setminus \{0\}, \cdot)$ *son grupos conmutativos y se cumple la propiedad distributiva del producto respecto a la suma.*

Los conjuntos de números reales y números complejos $(\mathbb{R}, +, \cdot), (\mathbb{C}, +, \cdot)$ son cuerpos conmutativos.

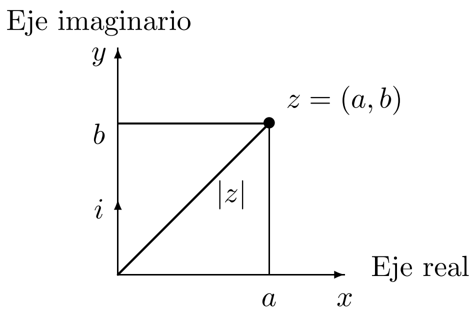
1.2 Números complejos

Definición 1.4 *Un* **número complejo** *es un par ordenado de números reales* $z = (a, b)$. *El número real* a *se llama* **parte real** *de* z *y* b *se llama* **parte imaginaria**.

$$Re(z) = a, \qquad Im(z) = b.$$

Si denotamos $1 = (1, 0), i = (0, 1)$, se escribe $z = (a, b) = a(1, 0) + b(0, 1) = a + bi$ (es la llamadada **forma binómica**). El número complejo $i = (0, 1)$ (en ocasiones se utiliza j), se llama **unidad imaginaria**. Los números complejos de la forma $(0, b)$ son llamados **imaginarios puros**. Así, denotaremos el conjunto de los números complejos como $\mathbb{C} = \{a+bi/a, b \in \mathbb{R}\}$.

De manera natural se identifica $\mathbb{C}$ con el plano $\mathbb{R}^2 : z = a + bi = (a, b)$. Se denominan las **coordenadas cartesianas** de z.



Operaciones en $\mathbb{C}$

- Suma: Sean $z_1 = a_1 + b_1 i, z_2 = a_2 + b_2 i$ dos números complejos. Se define su suma como

$$(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) = (a_1 + a_2) + (b_1 + b_2)i$$

Proposición 1.1 La suma de números complejos tiene las siguientes propiedades:

1. Propiedad asociativa: $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$
2. Elemento neutro de la suma: $(0, 0) = 0, z + 0 = 0 + z = z$
3. Elemento opuesto de la suma: $\forall z = (a, b) \Rightarrow -z = (-a, -b),\ z + (-z) = -z + z = 0$
4. Propiedad conmutativa: $z_1 + z_2 = z_2 + z_1$
5. Resta: $z_1 - z_2 = z_1 + (-z_2)$

- Producto: El producto de números complejos se realiza en forma binómica, teniendo en cuenta que $i^2 = -1$, es decir,

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2)$$

Proposición 1.2 El producto de números complejos tiene las siguientes propiedades:

1. Propiedad asociativa: $(z_1 \cdot z_2) \cdot z_3 = z_1 \cdot (z_2 \cdot z_3)$
2. Elemento neutro del producto: $(1, 0) = 1,\ z \cdot 1 = 1 \cdot z = z$
3. Elemento inverso del producto: $\forall z = (a, b) \neq 0 \Rightarrow z^{-1} = \dfrac{1}{z} = \left(\dfrac{a}{a^2 + b^2}, -\dfrac{b}{a^2 + b^2}\right)$
4. Propiedad conmutativa: $z_1 \cdot z_2 = z_2 \cdot z_1$
5. Cociente: $\dfrac{z_1}{z_2} = z_1 \cdot z_2^{-1},\ \forall z_2 \neq 0$

- $\mathbb{C}$ con las operaciones suma y producto tiene estructura de cuerpo conmutativo.
- En $\mathbb{C}$ no es posible definir una relación de orden compatible con la estructura de cuerpo.

Definición 1.5 *Se define el* **conjugado** *de un número complejo* $z = a + bi$ *como el número complejo* $\bar{z} = a - bi$.

Obsérvese que $z \cdot \bar{z} = a^2 + b^2$ y por tanto

$$z^{-1} = \frac{1}{z} = \frac{\bar{z}}{a^2 + b^2} = \frac{a - bi}{a^2 + b^2}$$

que está bien definido para $z \neq 0$.

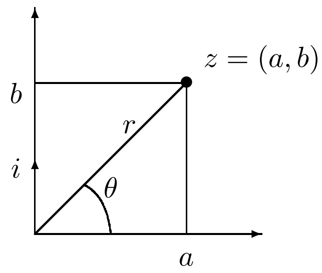
Proposición 1.3 Propiedades del conjugado

1. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \qquad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}.$
2. $z + \bar{z} = 2\ \mathrm{Re}(z), \qquad z - \bar{z} = 2i\ \mathrm{Im}(z).$
3. $\bar{\bar{z}} = z.$

Definición 1.6 *Se define el* **módulo** *de un número complejo* $z = (a, b)$ *como el número real no negativo* $|z| = \sqrt{a^2 + b^2}$. *Esta magnitud representa la distancia en* $\mathbb{R}^2$ *entre el punto* (a, b) *y el origen* $(0, 0)$.

Coordenadas polares

Sean (r, θ) las **coordenadas polares** del punto del plano (a, b) correspondiente al número complejo $z = a + bi \neq 0$.



$$a = r\,cos(\theta), \qquad b = r\,sen(\theta).$$

Entonces:
r es el **módulo** de z, pues:

$$r = \sqrt{x^2 + y^2} = \mid z \mid .$$

θ es el **argumento** de z: es el ángulo en $[0, 2\pi)$ que forma z con el eje real positivo. Se verifica:

$$arg(z) = \theta \text{ tal que } \begin{cases} \cos(\theta) = \dfrac{a}{r} \\ \operatorname{sen}(\theta) = \dfrac{b}{r} \end{cases}$$

Así, la **forma polar** del complejo se denota por $[r]_\theta$. Además, $|z|\ \cos(\theta) = a$, $|z|\ \operatorname{sen}(\theta) = b$. De este modo

$$z = a + bi = |z|(\cos(\theta) + i\ \operatorname{sen}(\theta))$$

es la llamada **forma trigonométrica** de z.

Utilizando las fórmulas trigonométricas para el seno y el coseno de la suma, se obtiene que si $z_1 = |z_1|(\cos(\theta_1) + \operatorname{sen}(\theta_1)i)$ y $z_2 = |z_2|(\cos(\theta_2) + \operatorname{sen}(\theta_2)i)$ son dos números complejos entonces

$$z_1 \cdot z_2 = |z_1||z_2|(\cos(\theta_1 + \theta_2) + \operatorname{sen}(\theta_1 + \theta_2)i),$$

es decir el módulo del producto es el producto de los módulos y el argumento del producto es la suma de los argumentos. De este modo, si $z = |z|(\cos(\theta) + \operatorname{sen}(\theta)i)$, entonces

$$z^n = |z|^n(\cos(n\theta) + \operatorname{sen}(n\theta)i), \forall n \in \mathbb{N}.$$

Con frecuencia se utiliza la notación siguiente, conocida como **fórmula de Euler**:

$$e^{i\theta} = cos(\theta) + isen(\theta),$$

por tanto, puede escribirse:

$$z = |z|\, e^{i\theta}.$$

es la llamada **forma exponencial** de z.

Las fórmulas para el producto y las potencias de números complejos resultan más sencillas cuando se utiliza la forma exponencial:

$$z_1 \cdot z_2 = |z_1|\, e^{i\theta_1} \cdot |z_2|\, e^{i\theta_2} = |z_1||z_2|\, e^{i(\theta_1+\theta_2)}$$

$$z^n = (|z|\, e^{i\theta})^n = |z|^n\, e^{i(n\theta)}$$

Tema 2

Matrices y determinantes

En este capítulo se introducen los conceptos básicos de la teoría de matrices, con especial atención a las operaciones elementales, que serán de mucha utilidad a lo largo del curso. Sus primeras aplicaciones (incluidas en este tema) son el cálculo del rango, la matriz inversa y el determinante.

2.1 Definición y tipos de matrices

Definición 2.1 *Se llama* **matriz** *real de p filas y n columnas a cualquier agrupación de la forma*

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix}$$

donde $a_{ij} \in \mathbb{R}, \forall i = 1, 2, \ldots, p, j = 1, 2, \ldots, n$. *También diremos que A es una matriz de tamaño* $p \times n$ *o de* **orden** $p \times n$.

Denotaremos por $\mathcal{M}_{p\times n}(\mathbb{R})$ el conjunto de todas las matrices de p filas y n columnas con elementos en $\mathbb{R}$. En notación reducida, escribiremos $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$.

Tipos de matrices

- Matriz **fila**: $A = (1, 2, 3, 4) \in \mathcal{M}_{1\times 4}(\mathbb{R})$
- Matriz **columna**: $A = \begin{pmatrix} 1 \\ -2 \\ 5 \end{pmatrix} \in \mathcal{M}_{3\times 1}(\mathbb{R})$
- Matriz **rectangular**: $A \in \mathcal{M}_{p\times n}(\mathbb{R})$
- Matriz **cuadrada**: $A \in \mathcal{M}_{n\times n}(\mathbb{R})$

- Matriz **cero** o **nula**: $\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathcal{M}_{p\times n}(\mathbb{R})$

- Matriz **identidad** o **unidad**: $I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{R})$

- Una matriz cuadrada $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es **diagonal** si los elementos de fuera de la diagonal son todos ceros, es decir, $a_{ij} = 0$ para todo $i \neq j$. Son de la forma

$$A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$$

- Una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ es **triangular superior** si $a_{ij} = 0$ para todo $i > j$, es decir, si los elementos que están por debajo de la diagonal son todos cero. Por ejemplo,

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

- Una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ es **triangular inferior** si $a_{ij} = 0$ para todo $i < j$, es decir, si los elementos que están por encima de la diagonal son todos cero.

Si $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ es una matriz cuadrada, se llama **diagonal** de A al vector de $\mathbb{R}^n$ que contiene los elementos a_{ij} con $i = j$, es decir,

$$\mathrm{diag}(A) = (a_{11}, a_{22}, \ldots, a_{nn}).$$

La suma de los elementos diagonales de A se llama **traza** de A y se denota por $\mathrm{tr}(A)$. Es decir,

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + a_{nn}.$$

Sea $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Se define su **traspuesta** y se denota A^t como la matriz cuyas columnas son las filas de A. En general, cuando hagamos operaciones con matrices que incluyan vectores, éstos se representarán en forma de columna. Si $v \in \mathbb{R}^n$ es un vector columna, el correspondiente vector fila es v^t:

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{R}) \Longrightarrow v^t = (v_1, v_2, \ldots, v_n) \in \mathcal{M}_{1\times n}(\mathbb{R})$$

2.2 Operaciones con matrices

Suma de matrices

La suma es una operación interna en $\mathcal{M}_{p\times n}(\mathbb{R})$. Dadas dos matrices $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$, $B = (b_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$, se define su suma como la matriz

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

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{p1} & b_{p2} & \cdots & b_{pn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1}+b_{p1} & a_{p2}+b_{p2} & \cdots & a_{pn}+b_{pn} \end{pmatrix}$$

Con las propiedades:

1. Asociativa: $A + (B + C) = (A + B) + C$.
2. Elemento neutro: la matriz cero. $A + 0 = 0 + A = A$.
3. Elemento opuesto: Para una matriz A su opuesto es $-A$. $A + (-A) = -A + A = 0$.
4. Conmutativa: $A + B = B + A$.

Con estas cuatro propiedades $(\mathcal{M}_{p\times n}(\mathbb{R}), +)$ tiene estructura de grupo conmutativo.

Producto de una matriz por un escalar

Dada una matriz $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$ y un escalar $\lambda \in \mathbb{R}$, se define

$$\lambda\, A = \lambda\,(a_{ij}) = (\lambda\, a_{ij})$$

$$\lambda \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} = \begin{pmatrix} \lambda\, a_{11} & \lambda\, a_{12} & \cdots & \lambda\, a_{1n} \\ \lambda\, a_{21} & \lambda\, a_{22} & \cdots & \lambda\, a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda\, a_{p1} & \lambda\, a_{p2} & \cdots & \lambda\, a_{pn} \end{pmatrix}$$

Esta es una operación externa que tiene las propiedades:

1. $\lambda(A + B) = \lambda A + \lambda B,\ \forall A, B \in \mathcal{M}_{p\times n}(\mathbb{R}), \forall\lambda \in \mathbb{R}$.
2. $(\lambda + \mu)A = \lambda\, A + \mu\, A,\ \forall A \in \mathcal{M}_{p\times n}(\mathbb{R}), \forall\lambda, \mu \in \mathbb{R}$.
3. $(\lambda\mu)A = \lambda(\mu A),\ \forall A \in \mathcal{M}_{p\times n}(\mathbb{R}), \forall\lambda, \mu \in \mathbb{R}$.
4. $1 \cdot A = A$.

$(\mathcal{M}_{p\times n}(\mathbb{R}), +, \cdot)$ es un espacio vectorial sobre el cuerpo de los números reales.

Producto de matrices

Dadas dos matrices $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$, $B = (b_{ij}) \in \mathcal{M}_{n\times q}(\mathbb{R})$, se define su producto como la matriz $AB = (c_{ij}) \in \mathcal{M}_{p\times q}(\mathbb{R})$ dada por

$$c_{ij} = \sum_{k=1}^{n} a_{ik}\, b_{kj},\ \forall i = 1, 2, \ldots, p,\ \forall j = 1, 2, \ldots, q.$$

Observación 2.1 Para poder realizar el producto AB es necesario que el número de columnas de A coincida con el número de filas de B. Un caso especialmente interesante se presenta cuando ambas matrices son vectores de $\mathbb{R}^n$. Sean

$$u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{R}), \qquad v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{R})$$

Entonces:

$$u^t\, v = (u_1, u_2, \ldots, u_n) \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = u_1v_1 + u_2v_2 + \cdots + u_nv_n \in \mathbb{R}$$

coincide con el producto escalar de ambos vectores, $u \cdot v$, mientras que

$$u\, v^t = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} (v_1, v_2, \ldots, v_n) = \begin{pmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_2v_1 & u_2v_2 & \cdots & u_2v_n \\ \vdots & \vdots & \ddots & \vdots \\ u_nv_1 & u_nv_2 & \cdots & u_nv_n \end{pmatrix} \in \mathcal{M}_{n\times n}(\mathbb{R})$$

Esta operación tiene las siguientes propiedades:

1. Si A, B, C se pueden multiplicar, entonces se cumple la propiedad asociativa:

$$A(BC) = (AB)C$$

2. Elemento neutro: la matriz identidad. Si $I \in \mathcal{M}_{n\times n}(\mathbb{R})$, entonces

$$AI = A, \forall A \in \mathcal{M}_{p\times n}(\mathbb{R}),\ \ IB = B, \forall B \in \mathcal{M}_{n\times q}(\mathbb{R})$$

3. El producto de matrices verifica la propiedad distributiva respecto a la suma, es decir, si $A, B \in \mathcal{M}_{p\times n}(\mathbb{R}), C, D \in \mathcal{M}_{n\times q}(\mathbb{R})$, entonces

$$A(C + D) = AC + AD, \qquad (A + B)C = AC + BC$$

4. El producto de matrices NO es conmutativo. Es decir, si $A, B \in \mathcal{M}_{n\times n}(\mathbb{R}), AB \neq BA$, en general.
   Ejemplo:
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} \neq \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

5. Si $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$, en general $AB = 0 \nRightarrow A = 0$ o $B = 0$. Ejemplo:
$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

Matriz inversa y potencia de una matriz

Para matrices cuadradas tiene sentido definir el concepto de matriz inversa y el de potencia de una matriz.

Definición 2.2 *Una matriz cuadrada $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ se dice* **inversible** *si existe una matriz, que llamaremos inversa de A y denotaremos por A^{-1} , tal que $AA^{-1} = A^{-1}A = I$, donde I es la matriz identidad.*

Proposición 2.1 *Sean $A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$. Si A y B son inversibles entonces AB también lo es y además $(AB)^{-1} = B^{-1}A^{-1}$.*

Definición 2.3 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y $k \in \mathbb{N}$. La* **potencia k-ésima** *de A es la matriz que resulta de multiplicar A por sí misma k veces. Se denota por A^k. Es decir,*
$$A^k = \underbrace{A \cdot A \cdots A}_{k}.$$

Por convenio, $A^0 = I, A^1 = A$.

En general es difícil encontrar la expresión general de A^k en función de k. Sin embargo, es sencillo para matrices diagonales:

Proposición 2.2 *Si A es diagonal entonces A^k también es diagonal. Además,*
$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}^k = \begin{pmatrix} a_{11}^k & 0 & \cdots & 0 \\ 0 & a_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn}^k \end{pmatrix}$$

2.3 Trasposición de matrices

Definición 2.4 *Dada una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ se define su* **matriz traspuesta** *$A^t \in \mathcal{M}_{n\times p}(\mathbb{R})$ como aquella matriz cuyas columnas son las filas de A.*

Proposición 2.3 Se cumplen las siguientes propiedades:

1. $(A^t)^t = A, \forall A \in \mathcal{M}_{p\times n}(\mathbb{R})$
2. $(A+B)^t = A^t + B^t, \forall A, B \in \mathcal{M}_{p\times n}(\mathbb{R})$
3. $(\lambda A)^t = \lambda A^t, \forall A \in \mathcal{M}_{p\times n}(\mathbb{R}), \forall \lambda \in \mathbb{R}$
4. $(AB)^t = B^t A^t, \forall A \in \mathcal{M}_{p\times n}(\mathbb{R}), \forall B \in \mathcal{M}_{n\times q}(\mathbb{R})$
5. Si A es inversible entonces $(A^t)^{-1} = (A^{-1})^t$

En relación con la trasposición de matrices tenemos las siguientes matrices especiales:

Definición 2.5 *Una matriz $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ es* **simétrica** *si $A^t = A$, es decir, si*

$$a_{ij} = a_{ji}, \forall i, j = 1, 2, \ldots, n.$$

La siguiente propiedad permite construir una matriz simétrica a partir de cualquier matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ y será importante en temas posteriores.

Proposición 2.4 *Si $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ entonces $A^t A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es simétrica.*

Definición 2.6 *Una matriz $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ es* **antisimétrica** *si $A^t = -A$, es decir, si*

$$a_{ij} = -a_{ji}, \forall i, j = 1, 2, \ldots, n.$$

Definición 2.7 *Una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es* **ortogonal** *si $AA^t = A^t A = I$, es decir, si A es inversible y $A^t = A^{-1}$.*

Ejemplo 2.1

1. La matriz $A = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 2 & 3 \\ 1 & 3 & 1 \end{pmatrix}$ es simétrica.
2. La matriz $A = \begin{pmatrix} 0 & -2 & 4 \\ 2 & 0 & 2 \\ -4 & -2 & 1 \end{pmatrix}$ es antisimétrica.
3. Si α es cualquier número real, la matriz de rotación de ángulo α

$$\begin{pmatrix} \cos(\alpha) & \operatorname{sen}(\alpha) \\ -\operatorname{sen}(\alpha) & \cos(\alpha) \end{pmatrix}$$

es ortogonal.

2.4 Matrices elementales

Definición 2.8 *Sea* $A \in \mathcal{M}_{p\times n}(\mathbb{R})$*. Se llaman* **operaciones elementales** *(o transformaciones elementales) sobre las filas o columnas de* A *a cualquiera de las siguientes transformaciones:*

1. *Permutar dos filas o dos columnas de* A*.*
2. *Sumar a una fila (o columna) de* A *un múltiplo de otra fila (o columna) de* A*.*
3. *Multiplicar una fila o columna de* A *por un escalar no nulo.*

Las operaciones elementales no afectan a la independencia lineal. Si una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ tiene k filas linealmente independientes y se realizan operaciones elementales por filas en A entonces la matriz resultante también tiene k filas linealmente independientes. Además, el subespacio de $\mathbb{R}^n$ que generan es el mismo.

Definición 2.9 *Una matriz* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *es una* **matriz elemental** *si se obtiene como resultado de efectuar una operación elemental sobre las filas o columnas de la matriz identidad.*

Tipos de matrices elementales

Distinguiremos seis tipos de matrices elementales según los tipos de operaciones elementales definidos arriba y dependiendo de si la operación se realiza sobre las filas o sobre las columnas de la matriz identidad.

1. F_{ij} es la matriz obtenida al permutar las filas i y j en I.
   Intercambia las filas i y j de una matriz A.
2. $F_i(\lambda)$ es la matriz obtenida al multiplicar la fila i de I por un escalar $\lambda \neq 0$.
   Multiplica la fila i de una matriz A por un número $\lambda \neq 0$.
3. $F_{ij}(\lambda)$ es la matriz obtenida al sumar a la fila i de I la fila j multiplicada por el escalar λ.
   Suma a la fila i de una matriz A su fila j multiplicada por $\lambda \neq 0$.
4. K_{ij} es la matriz obtenida al permutar las columnas i y j en I.
   Intercambia las columnas i y j de una matriz A.
5. $K_i(\lambda)$ es la matriz obtenida al multiplicar la columna i de I por un escalar $\lambda \neq 0$.
   Multiplica la columna i de una matriz A por un número $\lambda \neq 0$.
6. $K_{ij}(\lambda)$ es la matriz obtenida al sumar a la columna i de I la columna j multiplicada por el escalar λ.
   Suma a la columna i de una matriz A su fila j multiplicada por $\lambda \neq 0$.

Ejemplo 2.2 Tomando $I \in \mathcal{M}_{3\times 3}(\mathbb{R})$, tenemos

$$F_{23} = K_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad F_2(3) = K_2(3) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$F_{13}(2) = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad K_{13}(2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

Efectos de las matrices elementales

Las operaciones elementales sobre las filas y columnas de una matriz A pueden obtenerse como resultado de multiplicar por una matriz elemental:

1. Realizar una operación elemental sobre las filas de $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ es equivalente a multiplicar A por la izquierda por la correspondiente matriz elemental de filas $F \in \mathcal{M}_{p\times p}(\mathbb{R})$.
2. Realizar una operación elemental sobre las columnas de $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ equivale a multiplicar A por la derecha por la correspondiente matriz elemental de columnas $K \in \mathcal{M}_{n\times n}(\mathbb{R})$.

Ejemplo 2.3 Sea $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

1. Restar a la fila 2 de A la fila 1 multiplicada por 3 es equivalente a multiplicar A por la izquierda por $F_{21}(-3)$:
$$F_{21}(-3)\,A = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & -1 & -3 \end{pmatrix}$$
2. Permutar las columnas 1 y 3 de A es equivalente a multiplicar A por la derecha por K_{13}:
$$A\,K_{13} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 6 & 5 & 4 \end{pmatrix}$$

Inversas de las matrices elementales

Es muy sencillo comprobar que todas las matrices elementales son inversibles y además su inversa es la matriz elemental equivalente a la "transformación inversa". Así,

1. Por filas:
$$(F_{ij})^{-1} = F_{ij}, \quad (F_i(\lambda))^{-1} = F_i(1/\lambda), \quad (F_{ij}(\lambda))^{-1} = F_{ij}(-\lambda).$$
2. Por columnas:
$$(K_{ij})^{-1} = K_{ij}, \quad (K_i(\lambda))^{-1} = K_i(1/\lambda), \quad (K_{ij}(\lambda))^{-1} = K_{ij}(-\lambda).$$

2.5 Forma escalonada y rango de una matriz

Definición 2.10 *Sea $A = (a_{ij}) \in \mathcal{M}_{p\times n}(\mathbb{R})$. Supongamos que la fila i de A no tiene todos los elementos iguales a cero. Se llama* **entrada principal** *de la fila i al primer elemento de dicha fila distinto de cero, es decir, al elemento a_{ij} tal que $a_{ij} \neq 0$ y $a_{ik} = 0$ para todo $k < j$.*

Definición 2.11 *Se dice que la matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ está en* **forma escalonada** *si cumple las dos siguientes condiciones:*

1. *Si hay alguna fila de ceros, está al final.*

2. *Si hay varias filas distintas de cero, entonces la entrada principal de cada fila no nula está más a la izquierda que la de la siguiente fila.*

Definición 2.12 *Se dice que la matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ está en* **forma escalonada reducida** *si cumple las siguientes condiciones:*

1. *Está en forma escalonada.*
2. *Todas las entradas principales son iguales a 1.*
3. *En cada columna donde hay una entrada principal, el resto de los elementos son ceros.*

Ejemplo 2.4 La matriz

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

está en forma escalonada reducida. Se han resaltado sus entradas principales.

El siguiente resultado es clave para las aplicaciones de las operaciones elementales:

Teorema 2.1 *(Reducción de Gauss-Jordan)*
Dada una matriz cualquiera $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ existen matrices F y U tales que $FA = U$ siendo U una matriz escalonada.

La matriz F no es más que el producto de las matrices de las transformaciones elementales filas realizadas para pasar de A a U.

Teorema 2.2 *Toda matriz se puede transformar en una matriz en forma escalonada reducida mediante operaciones elementales por filas.*

Para cada matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, la matriz obtenida mediante el teorema anterior es única y recibe el nombre de **forma escalonada reducida** de A. La denotaremos por $\operatorname{rref}(A)$.

Ejemplo 2.5 Hallar la forma escalonada reducida de

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

$$A = \begin{pmatrix} -1 & -1 & 0 & 3 & -2 \\ 3 & 3 & 2 & -1 & 0 \\ -3 & -3 & -2 & 1 & 0 \\ 2 & 2 & 3 & 0 & -2 \end{pmatrix} \xrightarrow[F_{31}(-3),\, F_{41}(2)]{F_{21}(3)} \begin{pmatrix} -1 & -1 & 0 & 3 & -2 \\ 0 & 0 & 2 & 8 & -6 \\ 0 & 0 & -2 & -8 & 6 \\ 0 & 0 & 3 & 6 & -6 \end{pmatrix}$$

$$\xrightarrow[F_{42}(-3/2)]{F_{32}(1)} \begin{pmatrix} -1 & -1 & 0 & 3 & -2 \\ 0 & 0 & 2 & 8 & -6 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -6 & 3 \end{pmatrix} \xrightarrow{F_{34}} \begin{pmatrix} -1 & -1 & 0 & 3 & -2 \\ 0 & 0 & 2 & 8 & -6 \\ 0 & 0 & 0 & -6 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Esta es la forma escalonada de la matriz. Si queremos obtener la forma escalonada reducida, tenemos que tener un 1 en cada entrada principal y seguir haciendo ceros en la columna de cada entrada principal.

$$\xrightarrow[F_2(1/2),\, F_3(-1/6)]{F_1(-1)} \begin{pmatrix} 1 & 1 & 0 & -3 & 2 \\ 0 & 0 & 1 & 4 & -3 \\ 0 & 0 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[F_{13}(3)]{F_{23}(-4)} \begin{pmatrix} 1 & 1 & 0 & 0 & 1/2 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & -1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Por tanto,

$$\mathrm{rref}(A) = \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & 1/2 \\ 0 & 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & -1/2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Definición 2.13 *Sea $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Se define el* **rango** *de A como el número de filas no nulas de la forma escalonada reducida de A. Se denota rg(A).*

En el ejemplo anterior, $\mathrm{rg}(A) = 3$.

Observación 2.2 En la práctica no es preciso calcular la forma escalonada reducida de A. El rango de filas de A coincide con el número de filas no nulas de cualquier matriz escalonada obtenida realizando operaciones elementales sobre las filas de A. De hecho, para calcular el rango de A se pueden combinar operaciones elementales por filas y por columnas hasta obtener una matriz en forma escalonada.

Proposición 2.5 *El rango de una matriz A coincide con el número de filas linealmente independientes de A.*

Demostración. Es consecuencia de que la independencia lineal de un conjunto de vectores no varía por operaciones elementales y el conjunto de filas no nulas de una matriz escalonada es linealmente independiente. □

El rango de A también coincide con el número de columnas linealmente independientes de A. Esto es equivalente a decir que $\mathrm{rg}(A) = \mathrm{rg}(A^t)$.

La siguiente propiedad proporciona un método para determinar si una matriz tiene inversa usando operaciones elementales.

Proposición 2.6 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz cuadrada. Las siguientes afirmaciones son equivalentes:*

1. *A es inversible.*

2. *$\mathrm{rref}(A) = I$.*

3. *$\mathrm{rg}(A) = n$.*

Demostración. Recordemos que rref(A) se obtiene haciendo operaciones elementales sobre las filas de A. Por tanto, rref$(A) = FA$, donde F es una matriz que resulta de multiplicar matrices elementales. En particular, F es inversible. Veamos que se cumplen las equivalencias:

(1) $\Longrightarrow$ (2): Como A es inversible, rref$(A) = FA$ también es inversible y por tanto no tiene filas de ceros. Necesariamente rref$(A) = I$.

(2) $\Longrightarrow$ (3): Como rref$(A) = I$, rref(A) tiene n filas no nulas y por tanto rg$(A) = n$.

(3) $\Longrightarrow$ (1): Como rg$(A) = n$, rref(A) tiene n filas no nulas y por tanto rref$(A) = I$. Esto quiere decir que existe una matriz F tal que $FA = \text{rref}(A) = I$. Por definición, A es inversible y $F = A^{-1}$. □

2.6 Cálculo de la inversa

Como consecuencia de que la forma escalonada reducida de las matrices inversibles es la identidad, se tiene el siguiente resultado:

Proposición 2.7 *Toda matriz inversible $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ se puede transformar en la matriz identidad mediante operaciones elementales por filas.*

Esta proposición permite calcular la inversa de A utilizando operaciones elementales del siguiente modo: sean $F_1, F_2, \ldots, F_k$ las matrices elementales de filas por las que debemos multiplicar A para llegar a la identidad, es decir, $F_k \ldots F_2F_1A = I$. Entonces $A^{-1} = F_k \ldots F_2F_1$.

En la práctica, se procede del siguiente modo: si escribimos la matriz ampliada $(A|I)$, el resultado de aplicar $F_1, F_2, \ldots, F_k$ sobre esta matriz es $(I|A^{-1})$:

Ejemplo 2.6 Para calcular la inversa de

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

realizamos las siguientes operaciones elementales:

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 0 & 0 & 1 & 0 \\ 1 & 0 & 3 & 0 & 0 & 1 \end{array}\right) \xrightarrow{F_{21}(-1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 1 & 0 & 3 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow{F_{31}(-1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 1 \end{array}\right) \xrightarrow{F_{32}(1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{F_{23}(1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -3 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right) \xrightarrow{F_{13}(-1)} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 3 & -1 & -1 \\ 0 & 1 & 0 & -3 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array}\right)$$

$$\xrightarrow{F_{12}(-1)} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 6 & -3 & -2 \\ 0 & 1 & 0 & -3 & 2 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 \end{array} \right) = (I|A^{-1})$$

Por tanto,

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

Observación 2.3

1. Nunca se pueden combinar operaciones elementales de filas y columnas para calcular la inversa.

2. El número de operaciones necesarias para calcular la inversa de una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ mediante la expresión

$$A^{-1} = \frac{1}{|A|} Adj(A^t)$$

es del orden de $(n+1)!$. Mientras que si se realiza mediante operaciones elementales es de n^3.

2.7 Determinantes

Las operaciones elementales también se usan como un método eficaz para calcular el determinante de una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, teniendo en cuenta las siguientes propiedades:

Proposición 2.8 *Propiedades de los determinantes*

1. $|AB| = |A|\,|B|, \forall A, B \in \mathcal{M}_{n\times n}(\mathbb{R})$

2. $|A^t| = |A|, \forall A \in \mathcal{M}_{n\times n}(\mathbb{R})$

3. *Sumar a una fila o columna de una matriz un múltiplo de otra fila o columna no varía el valor del determinante.*

4. *Si se intercambian dos filas o dos columnas de A, su determinante cambia de signo.*

5. *Si A es una matriz triangular entonces su determinante es el producto de los elementos de la diagonal.*

6. *Si A tiene una fila o una columna de ceros, entonces su determinante es nulo.*

7. *Si A tiene dos líneas (fila o columna) paralelas iguales o proporcionales, su determinante es nulo.*

8. *Si todos los elementos de una línea se multiplican por un número λ, todo el determinante queda multiplicado por dicho número. Si $\lambda \in \mathbb{R}$, entonces*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ \lambda a_{i1} & \lambda a_{i2} & \cdots & \lambda a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

La misma propiedad es válida si una columna está multiplicada por el escalar λ.

9. *Si descomponemos una línea en suma de dos, podemos descomponer el determinante en suma de dos determinantes.*

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{in}+b_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{in} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

10. *El determinante de una matriz no varía si a una línea se le suma una combinación lineal de líneas paralelas.*

11. *Si una línea de la matriz A es combinación lineal de otras paralelas, su determinante es nulo.*

12. $|\lambda A| = \lambda^n\,|A|, \forall A \in \mathcal{M}_{n\times n}(\mathbb{R}), \forall \lambda \in \mathbb{R}$. *En particular,* $|-A| = (-1)^n\,|A|$.

13. *Si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, entonces A es* **inversible** *si y sólo si $|A| \neq 0$. Además, en ese caso, $|A^{-1}| = 1/|A|$.*

De este modo, realizando operaciones elementales en A obtenemos una matriz en forma triangular cuyo determinante se calcula haciendo uso de la quinta propiedad.

Ejemplo 2.7 Vamos a recordar tres métodos para el cálculo del determinante de una matriz cuadrada:

1. Mediante la regla de Laplace.

   Sea $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$. Sea $\tilde{A}_{ij}$ la matriz que se obtiene suprimiendo en A la fila i y la columna j. Entonces, para cada fila i de A, se tiene:

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\, det(\tilde{A}_{ij}).$$

   Esta fórmula permite expresar el determinante de una matriz de orden n en función del determinante de n matrices de orden $(n-1)$. También se verifica una fórmula análoga para cada columna de A.

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 2 \end{vmatrix} = -1 \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} + 1 \begin{vmatrix} 1 & 2 \\ 2 & 2 \end{vmatrix} = -1\cdot 0 + 1\cdot(-2) = -2$$

2. Mediante la regla de Sarrus (SOLO si la matriz es de orden 3)

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 2 \end{vmatrix} = 1\cdot 1\cdot 2 + 1\cdot 1\cdot 2 + 1\cdot 0\cdot 2 - 2\cdot 1\cdot 2 - 1\cdot 1\cdot 0 - 1\cdot 1\cdot 2 = -2$$

3. Mediante operaciones elementales

$$\begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 1 & 2 \end{vmatrix} \underset{F_{31}(-2)}{\overset{F_{21}(-1)}{=}} \begin{vmatrix} 1 & 1 & 2 \\ 0 & 0 & -2 \\ 0 & -1 & -2 \end{vmatrix} \overset{F_{23}}{=} - \begin{vmatrix} 1 & 1 & 2 \\ 0 & -1 & -2 \\ 0 & 0 & -2 \end{vmatrix} = -2$$

Observación 2.4 Para poner de manifiesto la propiedad 8, veamos que

$$\begin{vmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{vmatrix} \underset{F_{31}(-1/2)}{\overset{F_{21}(1)}{=}} \begin{vmatrix} 2 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1/2 \end{vmatrix} = 1.$$

Este es el valor del determinante. Sin embargo,

$$\begin{vmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{vmatrix} \underset{2F_3-F_1}{\overset{F_{21}(1)}{=}} \begin{vmatrix} 2 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 2,$$

con lo cual se obtiene un resultado erróneo.

Tema 3

Sistemas de ecuaciones lineales

Este capítulo está dedicado a la resolución de sistemas de ecuaciones lineales, lo que incluye el estudio de la compatibilidad del sistema (existencia de soluciones), la determinación del conjunto de soluciones y la interpretación geométrica de dicho conjunto. El método principal de resolución es el método de Gauss, basado en operaciones elementales sobre las filas de la matriz ampliada del sistema.

3.1 Expresión matricial

Definición 3.1 *Un* **sistema** *de p ecuaciones lineales con n incógnitas en $\mathbb{R}$ es un conjunto de expresiones:*

$$\begin{array}{rl} a_{11}\,x_1 + a_{12}\,x_2 + \cdots + a_{1n}\,x_n & = b_1 \\ a_{21}\,x_1 + a_{22}\,x_2 + \cdots + a_{2n}\,x_n & = b_2 \\ \cdots & = \cdots \\ a_{p1}\,x_1 + a_{p2}\,x_2 + \cdots + a_{pn}\,x_n & = b_p \end{array}$$

donde los elementos $a_{ij} \in \mathbb{R}$ se llaman **coeficientes del sistema**, *$b_i \in \mathbb{R}$ se llaman* **términos independientes** *y x_i se llaman* **incógnitas**.

El sistema es **homogéneo** *si $b_i = 0, \forall i = 1, 2, \ldots, p$. En otro caso diremos que es* **no homogéneo**.

El sistema se puede expresar en la forma matricial $Ax = b$, donde

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} \end{pmatrix} \in \mathcal{M}_{p\times n}(\mathbb{R}),\ b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{pmatrix} \in \mathbb{R}^p,\ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

La matriz A se llama **matriz de coeficientes del sistema** y b es el **término independiente**. La matriz

$$(A|b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pn} & b_p \end{array}\right) \in \mathcal{M}_{p\times(n+1)}(\mathbb{R})$$

se llama **matriz ampliada del sistema**. Cada una de las ecuaciones se puede identificar con la correspondiente fila de la matriz $(A|b)$. Obsérvese que el número de columnas de A coincide con el número de incógnitas del sistema.

3.2 Existencia de soluciones

Definición 3.2 *Un vector* $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ *es una* **solución** *del sistema si* $Av = b$.

Resolver el sistema es determinar el conjunto de sus soluciones (que es un subconjunto de $\mathbb{R}^n$). Si no existe ninguna solución, el sistema es **incompatible**. Si existe alguna solución, diremos que el sistema es **compatible determinado** si la solución es única y **compatible indeterminado** si existe más de una solución.

Eliminación gaussiana

La siguiente propiedad permitirá estudiar con facilidad si un sistema es compatible y calcular el conjunto de sus soluciones.

Teorema 3.1 *Sea* $Ax = b$ *un sistema de* p *ecuaciones lineales con* n *incógnitas. Si efectuamos operaciones elementales sobre las filas de la matriz ampliada* $(A|b)$ *hasta obtener una nueva matriz* $(A'|b')$ *entonces los sistemas* $Ax = b$ *y* $A'x = b'$ *son equivalentes, es decir, tienen el mismo conjunto de soluciones.*

Utilizando esta proposición, para resolver un sistema se realizan operaciones elementales sobre las filas de $(A|b)$ hasta obtener su forma escalonada reducida $(A'|b')$. Sea $r = \text{rg}(A|b) = \text{rg}(A'|b')$. El sistema $A'x = b'$ se resuelve de forma inmediata, despejando las r incógnitas correspondientes a las entradas principales en función de las $(n-r)$ restantes (incógnitas libres).

Teorema 3.2 Toerema de Rouché-Frobenius
Sea $Ax = b$ *un sistema de* p *ecuaciones lineales con* n *incógnitas.*

- *Si* $rg(A) \neq rg(A|b)$ *entonces el sistema es incompatible.*
- *Si* $rg(A) = rg(A|b) = n$ *(*$n =$ *número de incógnitas =número de columnas de* A*) entonces el sistema es compatible determinado.*
- *Si* $rg(A) = rg(A|b) < n$ *entonces el sistema es compatible indeterminado y el conjunto de soluciones se puede escribir en función de las* $(n-r)$ *incógnitas libres.*

3.3 Conjuntos de soluciones

Una de las características especiales de los sistemas de ecuaciones lineales es que aunque el conjunto de soluciones puede ser infinito, siempre queda determinado por un conjunto finito de vectores de $\mathbb{R}^n$. Comenzamos analizando el caso de sistemas homogéneos.

Sistemas homogéneos

Consideremos un sistema homogéneo $Ax = 0$, donde $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. En primer lugar, observemos que un sistema homogéneo siempre es compatible, ya que $x = 0$ es solución. El conjunto de soluciones se denomina **núcleo** de A y se denota por $\mathrm{Ker}(A)$, es decir,

$$\mathrm{Ker}(A) = \{x \in \mathbb{R}^n / Ax = 0\}.$$

Por tanto sólo hay dos posibilidades:

- Si $\mathrm{rg}(A) = n$ entonces el sistema es compatible determinado y su única solución es el vector cero ($\mathrm{Ker}(A) = \{0\}$).
- Si $\mathrm{rg}(A) = r < n$ entonces el sistema es compatible indeterminado y el núcleo de A es el conjunto de todas las combinaciones lineales de $k = n - r$ vectores de $\mathbb{R}^n$ $u_1, u_2, \dots, u_k$, es decir,

  $$\mathrm{Ker}(A) = \{\lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_k u_k / \lambda_i \in \mathbb{R}, i = 1, \dots, k\}.$$

  Estos vectores se determinan despejando las incógnitas correspondientes a las entradas principales de la forma escalonada reducida de A en función del resto. En otras palabras, el núcleo de A es el subespacio de $\mathbb{R}^n$ generado por los vectores $u_1, u_2, \dots, u_k$:

  $$\mathrm{Ker}(A) =< \{u_1, u_2, \dots, u_k\} >$$

  y $\dim(\mathrm{Ker}(A)) = n - \mathrm{rg}(A)$.

  Resolver el sistema homogéneo $Ax = 0$ en el caso compatible indeterminado equivale a calcular una base del núcleo de A.

Ejemplo 3.1 Consideremos el sistema

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

Realizando operaciones elementales sobre las filas de la matriz A, tenemos:

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 2 \end{pmatrix} \xrightarrow[F_{31}(-1)]{F_{21}(-1)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & -1 & 1 & 1 \end{pmatrix} \xrightarrow{F_{32}(1)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{F_{12}(-1)} \begin{pmatrix} \boxed{1} & 0 & 2 & 2 \\ 0 & \boxed{1} & -1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = A' = \mathrm{rref}(A)$$

Como $\mathrm{rg}(A) = \mathrm{rg}(A') = 2 < 4 =$ número de incógnitas, el sistema es compatible indeterminado. Además, el conjunto de soluciones de $Ax = 0$ coincide con el conjunto de soluciones del sistema equivalente $A'x = 0$, es decir, del sistema

$$\begin{cases} \boxed{x} + 2z + 2t = 0 \\ \boxed{y} - z - t = 0 \end{cases}$$

Despejando las incógnitas x e y correspondientes a las entradas principales en función de las incógnitas libres z y t, tenemos que el conjunto de soluciones es:

$$\begin{aligned}\mathrm{Ker}(A) &= \{(x,y,z,t)\in\mathbb{R}^4/x=-2z-2t, y=z+t\} = \{(-2z-2t, z+t, z, t)/z,t\in\mathbb{R}\} = \\ &= \{z(-2,1,1,0)+t(-2,1,0,1)/z,t\in\mathbb{R}\} =< \{(-2,1,1,0),(-2,1,0,1)\} > .\end{aligned}$$

El conjunto de soluciones es el subespacio de $\mathbb{R}^4$ de dimensión 2 formado por todas las combinaciones lineales de $u_1=(-2,1,1,0)$ y $u_2=(-2,1,0,1)$.

Sistemas no homogéneos

Consideremos ahora un sistema no homogéneo $Ax=b$, con $A\in\mathcal{M}_{p\times n}(\mathbb{R}), b\in\mathbb{R}^p$. El sistema es compatible indeterminado si $\mathrm{rg}(A)=r=\mathrm{rg}(A|b)<n$. En este caso el conjunto de soluciones está determinado por los $k=n-r$ generadores del núcleo de A y un vector p llamado solución particular. En concreto, se tiene el siguiente resultado:

Teorema 3.3 *Si $rg(A)=r=rg(A|b)<n$, el conjunto de soluciones del sistema $Ax=b$ es*

$$S=\{p+\lambda_1u_1+\lambda_2u_2+\cdots+\lambda_ku_k/\lambda_i\in\mathbb{R}, i=1,\dots,k\} := p+ < \{u_1,u_2,\dots,u_k\} >,$$

donde p es una solución de $Ax=b$ (es decir, $Ap=b$) y $<\{u_1,u_2,\dots,u_k\}>= Ker(A)$. En notación abreviada, escribiremos el conjunto de soluciones en la forma $S=p+Ker(A)$.

Ejemplo 3.2 Consideremos el sistema

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

Realizando operaciones elementales sobre las filas de la matriz ampliada $(A|b)$, tenemos:

$$(A|b)=\left(\begin{array}{ccc|c}1&1&1&1\\1&2&0&1\\1&0&2&1\end{array}\right)\xrightarrow[F_{31}(-1)]{F_{21}(-1)}\left(\begin{array}{ccc|c}1&1&1&1\\0&1&-1&0\\0&-1&1&0\end{array}\right)\xrightarrow{F_{32}(1)}\left(\begin{array}{ccc|c}1&1&1&1\\0&1&-1&0\\0&0&0&0\end{array}\right)$$

$$\xrightarrow{F_{12}(-1)}\left(\begin{array}{ccc|c}1&0&2&1\\0&1&-1&0\\0&0&0&0\end{array}\right)=(A'|b')$$

En primer lugar, $\mathrm{rg}(A|b)=\mathrm{rg}(A'|b')=2<3=$ número de incógnitas, y por tanto el sistema es compatible indeterminado. Además, el conjunto de soluciones de $Ax=b$ coincide con el conjunto de soluciones de $A'x=b'$, es decir, del sistema

$$x+2z=1$$

$$y-z=0.$$

Despejando $x=1-2z, y=z$, tenemos que el conjunto de soluciones es

$$\begin{aligned}S &= \{(x,y,z)\in\mathbb{R}^3/y=z, x=1-2z\} = \{(1-2z,z,z)/z\in\mathbb{R}\} = \\ &= \{(1,0,0)+z(-2,1,1)/z\in\mathbb{R}\} = (1,0,0)+<\{(-2,1,1)\}>= p+\mathrm{Ker}(A).\end{aligned}$$

Ejemplo 3.3 Queremos resolver el sistema

$$\begin{cases} x-2y+2z &= 3 \\ 3x-5y+4z &= 1 \\ 2x-3y+2z &= 1 \end{cases}$$

Buscamos la forma escalonada de la matriz ampliada $(A|b)$ del sistema:

$$(A|b) = \left(\begin{array}{ccc|c} 1 & -2 & 2 & 3 \\ 3 & -5 & 4 & 1 \\ 2 & -3 & 2 & 1 \end{array}\right) \xrightarrow[F_{31}(-2)]{F_{21}(-3)} \left(\begin{array}{ccc|c} 1 & -2 & 2 & 3 \\ 0 & 1 & -2 & -8 \\ 0 & 1 & -2 & -5 \end{array}\right) \xrightarrow{F_{32}(-1)} \left(\begin{array}{ccc|c} 1 & -2 & 2 & 3 \\ 0 & 1 & -2 & -8 \\ 0 & 0 & 0 & 3 \end{array}\right)$$

La última fila representa la ecuación $0x+0y+0z=3$ lo que produce un sistema incompatible pues $0 \neq 3$.

3.4 Matrices cuadradas y uso de la factorización LU

Cuando A es una matriz cuadrada, es más sencillo determinar si el sistema $Ax=b$ es compatible determinado:

Proposición 3.1 *Sean $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y $b \in \mathbb{R}^n$. El sistema $Ax=b$ tiene solución única si y sólo si rg$(A)=n$.*

Obsérvese que en este caso la única solución del sistema homogéneo asociado $Ax=0$ es la solución trivial, es decir, $\mathrm{Ker}(A)=\{0\}$.

Proposición 3.2 *Las siguientes propiedades son equivalentes para una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$:*

1. *El sistema $Ax=b$ es compatible determinado para cada $b \in \mathbb{R}^n$.*
2. *Ker$(A)=\{0\}$.*
3. *rg$(A)=n$.*
4. *A es inversible.*
5. *$|A| \neq 0$.*

Observación 3.1 Si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es inversible, entonces la única solución del sistema $Ax=b$ se puede escribir en la forma $x=A^{-1}b$. Sin embargo, en la práctica no se suele calcular la inversa de A para resolver el sistema.

Factorización *LU*

La factorización LU consiste en descomponer una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ en el producto $A=LU$, donde $L \in \mathcal{M}_{n\times n}(\mathbb{R})$ es una matriz triangular inferior con todos los elementos diagonales iguales a 1, y $U \in \mathcal{M}_{n\times n}(\mathbb{R})$ es una matriz triangular superior. Diremos que A admite factorización LU si es posible encontrar estas dos matrices.

El método de cálculo de L y U se basa en la eliminación gaussiana. Para poder obtener L y U por este procedimiento será necesario pedir condiciones adicionales a la matriz A.

Proposición 3.3 *Si todos los menores principales de A son distintos de cero entonces A admite factorización LU. Además, en este caso, dicha factorización es única.*

Cálculo de la factorización LU

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz en las condiciones de la proposición anterior. Entonces es posible transformar la matriz A en una matriz triangular superior U mediante operaciones elementales sobre las filas de A del tipo $F_{ij}(\lambda)$, con $i > j$, es decir, sin efectuar permutaciones de filas y utilizando sólo las filas superiores para modificar las inferiores.

Sean $F_1, F_2, \ldots, F_k$ las correspondientes matrices elementales de filas tales que $F_k \ldots F_2F_1A = U$. Entonces $L = (F_k \ldots F_2F_1)^{-1} = F_1^{-1}F_2^{-1} \ldots F_k^{-1}$ es triangular inferior, sus elementos diagonales son iguales a 1 y además $A = LU$.

Ejemplo 3.4 Consideremos la matriz:

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

Veamos que A admite factorización LU. Los menores principales de la matriz A son:

$$\Delta_1 = 2 \neq 0$$

$$\Delta_2 = \begin{vmatrix} 2 & -1 \\ 4 & -4 \end{vmatrix} = -4 \neq 0$$

$$\Delta_3 = \begin{vmatrix} 2 & -1 & 0 \\ 4 & -4 & 1 \\ -2 & 1 & -1 \end{vmatrix} = 4 \neq 0$$

$$\Delta_4 = \begin{vmatrix} 2 & -1 & 0 & 1 \\ 4 & -4 & 1 & 5 \\ -2 & 1 & -1 & 0 \\ -2 & 5 & -4 & -1 \end{vmatrix} = 16 \neq 0$$

Todos los menores principales de A son no nulos y por tanto admite factorización LU. Para calcular dicha factorización, en primer lugar determinaremos la matriz triangular superior U mediante operaciones elementales sobre las filas de la matriz A del tipo $F_{ij}(\lambda)$, con $i > j$. Así,

$$\begin{pmatrix} 2 & -1 & 0 & 1 \\ 4 & -4 & 1 & 5 \\ -2 & 1 & -1 & 0 \\ -2 & 5 & -4 & -1 \end{pmatrix} \xrightarrow[F_{41}(1)]{F_{21}(-2),\, F_{31}(1)} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 4 & -4 & 0 \end{pmatrix}$$

$$\xrightarrow{F_{42}(2)} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & -2 & 6 \end{pmatrix} \xrightarrow{F_{43}(-2)} \begin{pmatrix} 2 & -1 & 0 & 1 \\ 0 & -2 & 1 & 3 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix} = U.$$

De esto se deduce que

$$[F_{43}(-2)\ F_{42}(2)\ F_{41}(1)\ F_{31}(1)\ F_{21}(-2)]A = U$$

y entonces

$$\begin{aligned} L &= [F_{43}(-2)\ F_{42}(2)\ F_{41}(1)\ F_{31}(1)\ F_{21}(-2)]^{-1} = \\ &= F_{21}(2)\ F_{31}(-1)\ F_{41}(-1)\ F_{42}(-2)\ F_{43}(2). \end{aligned}$$

Calcular el producto de estas matrices elementales es equivalente a realizar las correspondientes operaciones elementales a la matriz identidad:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[F_{41}(-1)]{F_{43}(2),\ F_{42}(-2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & -2 & 2 & 1 \end{pmatrix}$$

$$\xrightarrow{F_{31}(-1)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -2 & 2 & 1 \end{pmatrix} \xrightarrow{F_{21}(2)} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & -2 & 2 & 1 \end{pmatrix} = L.$$

Observación 3.2 En la práctica no es necesario comprobar previamente que todos los menores principales de A son no nulos. Esto es equivalente a que se pueda obtener la matriz U mediante operaciones elementales sobre las filas de A del tipo $F_{ij}(\lambda)$, con $i > j$, y además los elementos diagonales de U sean distintos de cero.

Uso de la factorización LU

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz cuadrada de rango n. Supongamos que A admite factorización LU. Entonces resolver el sistema de ecuaciones lineales $Ax = b$ es equivalente a resolver consecutivamente los sistemas $Lz = b, Ux = z$. (En efecto, $Ax = LUx = Lz = b$).

Ejemplo 3.5 Sean

$$A = \begin{pmatrix} 2 & -1 & 0 & 1 \\ 4 & -4 & 1 & 5 \\ -2 & 1 & -1 & 0 \\ -2 & 5 & -4 & -1 \end{pmatrix},\ b = \begin{pmatrix} -5 \\ -14 \\ 1 \\ 1 \end{pmatrix}$$

Vamos a resolver el sistema $Ax = b$ usando la factorización LU. En el ejemplo anterior hemos calculado la factorización LU de la matriz A:

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

Como $A = LU$, la resolución del sistema $Ax = b$ es equivalente a la resolución sucesiva de dos sistemas triangulares:

$$Ax = b \Longleftrightarrow L\underbrace{U\,x}_{z} = b \Longleftrightarrow \begin{cases} Lz = b \\ Ux = z \end{cases}$$

La solución $z = (z_1, z_2, z_3, z_4)^t$ del sistema $Lz = b$ se calcula resolviendo por **descenso**, y viene dada por

$$\begin{aligned} z_1 &= -5 \\ 2z_1 + z_2 = -14 \Longrightarrow z_2 &= -4 \\ -z_1 + z_3 = 1 \Longrightarrow z_3 &= -4 \\ -z_1 - 2z_2 + 2z_3 + z_4 = 1 \Longrightarrow z_4 &= -4 \end{aligned}$$

Calculamos ahora la solución del sistema $Ux = z$ por **remonte**:

$$\begin{aligned} 4x_4 = -4 &\Longrightarrow x_4 = -1 \\ -x_3 + x_4 = -4 &\Longrightarrow x_3 = 3 \\ -2x_2 + x_3 + 3x_4 = -4 &\Longrightarrow x_2 = 2 \\ 2x_1 - x_2 + x_4 = -5 &\Longrightarrow x_1 = -1 \end{aligned}$$

Se puede comprobar que $x = (x_1, x_2, x_3, x_4) = (-1, 2, 3, -1)$ es la solución del sistema original $Ax = b$.

3.5 Mínimos cuadrados. Ajuste

Consideremos un sistema de ecuaciones lineales $Ax = b$, donde $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ y $b \in \mathbb{R}^p$.

Definición 3.3 *Se define la* **imagen** *de A, y se denota por $Im(A)$, como el subespacio generado por las columnas de A.*

La compatibilidad del sistema $Ax = b$ se caracteriza en términos de la imagen de A de forma sencilla.

Proposición 3.4 *El sistema $Ax = b$ es compatible si y sólo si $b \in Im(A)$.*

La proposición anterior dice que $Ax = b$ es compatible sólo cuando b es combinación lineal de las columnas de A. En el caso de que el sistema sea incompatible, se puede buscar una "solución aproximada". Una posibilidad es determinar el vector $b' \in \mathrm{Im}(A)$ cuya distancia al término independiente b sea la menor posible. Los vectores $x \in \mathbb{R}^n$ tales que $Ax = b'$ serán lo que llamaremos soluciones del sistema $Ax = b$ en el sentido de mínimos cuadrados.

Definición 3.4 *Sean $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ y $b \in \mathbb{R}^p$. Se dice que $x_0 \in \mathbb{R}^n$ es una solución* **en el sentido de mínimos cuadrados** *del sistema $Ax = b$ si se cumple la siguiente igualdad:*

$$||Ax_0 - b|| = \min\{||Ax - b||/x \in \mathbb{R}^n\}.$$

La distancia mínima de b a la imagen de A es la distancia de b a la proyección ortogonal de b sobre $\mathrm{Im}(A)$, es decir, el único vector $b' \in \mathrm{Im}(A)$ tal que $b - b'$ es ortogonal a todos los vectores de la imagen de A. Por tanto x_0 es una solución de $Ax = b$ en el sentido de mínimos cuadrados si y sólo si $v = Ax_0 - b$ es ortogonal a las columnas de A. Esto es equivalente a la relación

$$A^t(Ax_0 - b) = 0.$$

Por lo tanto, se cumple el siguiente resultado:

Teorema 3.4 *Sean $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ y $b \in R^p$. Un vector x_0 es una solución en el sentido de mínimos cuadrados de $Ax = b$ si y sólo si*

$$A^tAx_0 = A^tb.$$

El siguiente resultado es una consecuencia de que en $\mathbb{R}^p$ siempre es posible calcular la proyección ortogonal de un vector b sobre un subespacio U. Además, si $b \in U$ entonces la proyección ortogonal es el propio b.

Teorema 3.5 *Sean $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ y $b \in R^p$. El sistema de ecuaciones lineales $A^tAx = A^tb$ es un sistema compatible. Además:*

1. *Si $Ax = b$ es compatible entonces el conjunto de soluciones de $A^tAx = A^tb$ coincide con el conjunto de soluciones de $Ax = b$.*
2. *Si $Ax = b$ es incompatible entonces el conjunto de soluciones de $A^tAx = A^tb$ coincide con el conjunto de soluciones de $Ax = b$ en el sentido de mínimos cuadrados.*
3. *El sistema $A^tAx = A^tb$ tiene solución única si y sólo si $rg(A) = n$.*

Ejemplo 3.6 Calcula la solución en el sentido de mínimos cuadrados del sistema $Mx = b$ si

$$M = \begin{pmatrix} 2 & -1 & 2 \\ 1 & 1 & -1 \\ 2 & 1 & 2 \\ 1 & -1 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

Se puede comprobar que el sistema anterior es incompatible. Por tanto vamos a resolverlo en el sentido de mínimos cuadrados. Tenemos que buscar la solución de $M^tMx = M^tb$

$$\begin{pmatrix} 10 & 0 & 6 \\ 0 & 4 & 0 \\ 6 & 0 & 10 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 5 \\ 1 \\ 3 \end{pmatrix}$$

De modo que la solución es

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/4 \\ 0 \end{pmatrix}$$

Ajuste polinómico de datos mediante mínimos cuadrados

Supongamos que se calcula experimentalmente el valor de una cierta cantidad y que se supone que es función polinómica de otra cantidad x:

$$y = p(x) = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n.$$

Si se realizan k experimentos en los que se obtienen las mediciones $y_1, y_2, \ldots, y_k$ para los datos de entrada respectivos $x_1, x_2, \ldots, x_k$, los coeficientes del polinomio $p(x)$ vendrían dados por las soluciones del sistema de ecuaciones lineales

$$\begin{aligned}
y_1 &= a_0 + a_1 x_1 + a_2 x_1^2 + \ldots + a_n x_1^n \\
y_2 &= a_0 + a_1 x_2 + a_2 x_2^2 + \ldots + a_n x_n^2 \\
&\vdots \\
y_k &= a_0 + a_1 x_k + a_2 x_k^2 + \ldots + a_n x_k^n,
\end{aligned}$$

o, en forma matricial, $Ax = b$, donde

$$\underbrace{\begin{pmatrix} 1 & x_1 & x_1^2 & \ldots & x_1^n \\ 1 & x_2 & x_2^2 & \ldots & x_2^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_k & x_k^2 & \ldots & x_k^n \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}}_{x} = \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_k \end{pmatrix}}_{b}$$

Si el sistema $Ax = b$ es compatible entonces la gráfica del polinomio cuyos coeficientes son la solución del sistema pasa por todos los puntos $(x_1, y_1), (x_2, y_2), \ldots, (x_k, y_k)$. Si no es compatible, la solución del sistema de ecuaciones normales $A^t A x = A^t b$ proporciona los coeficientes del polinomio de grado n que mejor ajusta los datos en el sentido de mínimos cuadrados.

Observación 3.3 Si el polinomio $p(x)$ que buscamos es de grado 1 se dice que el ajuste es **lineal**. Si $p(x)$ es de grado 2, se dice que el ajuste es **cuadrático**.

Ejemplo 3.7 Encontrar la recta y la parábola de ajuste en el sentido de mínimos cuadrados para los siguientes datos:

x	-2	-1	1	2
y	3	1	1	5

La recta tiene la forma $y = a_0 + a_1 x$, de modo que buscamos la solución de mínimos cuadrados del sistema $A_1 x_1 = b$

$$\begin{pmatrix} 1 & -2 \\ 1 & -1 \\ 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 5 \end{pmatrix}$$

El sistema de mínimos cuadrados $A_1^t A_1 x_1 = A_1^t b$ es

$$\begin{pmatrix} 4 & 0 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \end{pmatrix}$$

Por tanto, $a_0 = 5/2, a_1 = 2/5$ y la recta es

$$y = 5/2 + 2/5x.$$

Si ahora buscamos la parábola $y = a_0 + a_1 x + a_2 x^2$ que ajusta mejor estos datos en el sentido de mínimos cuadrados, planteamos el sistema $A_2 x_2 = b$

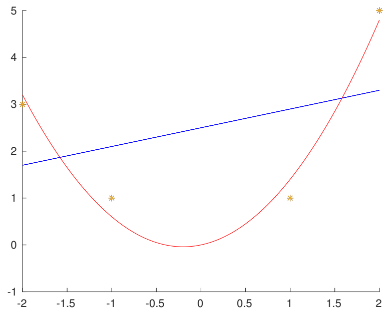
$$\begin{pmatrix} 1 & -2 & 4 \\ 1 & -1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 1 \\ 5 \end{pmatrix}$$

El sistema de ecuaciones normales es $A_2^t A_2 x_2 = A_2^t b$

$$\begin{pmatrix} 4 & 0 & 10 \\ 0 & 10 & 0 \\ 10 & 0 & 34 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} 10 \\ 4 \\ 34 \end{pmatrix}$$

y tiene como solución $(a_0, a_1, a_2) = (0, 2/5, 1)$. En consecuencia, la ecuación de la parábola de ajuste es

$$y = a_0 + a_1 x + a_2 x^2 = \frac{2}{5} x + x^2.$$



En la figura anterior se representan los puntos que corresponden a los datos y las aproximaciones lineal (azul) y cuadrática (rojo). Se observa que ésta última es mucho más precisa.

Tema 4

Espacios vectoriales

En este capítulo introduciremos la definición de espacio vectorial y los principales conceptos relacionados, como la independencia lineal, generadores, base y dimensión. También veremos los espacios vectoriales euclídeos.

4.1 Espacios y subespacios vectoriales

Definición 4.1 *Se llama* **espacio vectorial** *sobre* $\mathbb{R}$ *o espacio vectorial real a un conjunto* V *dotado de dos operaciones,* $(V, +, \cdot)$*:*

- *Una operación interna (***suma***), de tal forma que (V, +) es un grupo conmutativo con las propiedades:*

  1. *Asociativa:* $u + (v + w) = (u + v) + w, \forall u, v, w \in V$.
  2. *Elemento neutro: El vector 0,* $v + 0 = v + 0 = v, \forall v \in V$.
  3. *Elemento simétrico:* $v + (-v) = -v + v = 0, \forall v \in V$.
  4. *Conmutativa:* $u + v = v + u, \forall u, v \in V$.

- *Una operación externa (***producto por escalares***) que asigna a cada escalar* $\lambda \in \mathbb{R}$ *y a cada elemento* $v \in V$ *un nuevo elemento* $\lambda v \in V$*, de tal forma que se cumplen las siguientes propiedades:*

  1. $\lambda(v + w) = \lambda v + \lambda w, \forall \lambda \in \mathbb{R}, \forall v, w \in V$.
  2. $(\lambda + \mu)v = \lambda v + \mu v, \forall \lambda, \mu \in \mathbb{R}, \forall v \in V$.
  3. $(\lambda\mu)v = \lambda(\mu v), \forall \lambda, \mu \in \mathbb{R}, \forall v \in V$.
  4. $1v = v, \forall v \in V$*, donde 1 es el elemento neutro del producto en* $\mathbb{R}$.

A los elementos de V *los llamaremos* **vectores** *y a los elementos de* $\mathbb{R}$ *los llamaremos* **escalares**. *Generalmente denotaremos a estos últimos con letras del alfabeto griego.*

Ejemplo 4.1

1. $\mathbb{R}^n$ es un espacio vectorial real con las operaciones usuales de suma y producto por escalares.

2. El conjunto $\mathcal{M}_{p\times n}(\mathbb{R})$ de las matrices reales de p filas y n columnas es un espacio vectorial sobre $\mathbb{R}$ con las operaciones suma de matrices y producto de un escalar por una matriz.

3. El conjunto $\Pi_n(\mathbb{R})$ de los polinomios en una variable de grado menor o igual que n y con coeficientes en $\mathbb{R}$ es un espacio vectorial real con las operaciones habituales de suma de polinomios y producto de un escalar por un polinomio.
$$\Pi_n(\mathbb{R}) = \{a_0 + a_1x + \cdots + a_nx_n/a_0, a_1, \dots, a_n \in \mathbb{R}\}.$$

4. El conjunto $\mathcal{C}_0(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R}/f \text{ es continua}\}$ es un espacio vectorial real con las operaciones habituales de suma de funciones y producto de un escalar por una función.

Muchos de los conceptos definidos para $\mathbb{R}^n$ se extienden a otros espacios vectoriales. A continuación repasamos algunos.

Subespacios vectoriales

Definición 4.2 *Sea V un espacio vectorial. Un subconjunto U de V es un subespacio vectorial de V si cumple las siguientes propiedades:*

1. $0 \in U$.
2. $u_1 + u_2 \in U, \forall u_1, u_2 \in U$.
3. $\lambda u \in U, \forall \lambda \in \mathbb{R}, \forall u \in U$.

Ejemplo 4.2

1. Si $A \in \mathcal{M}_{p\times n}(\mathbb{R})$, entonces $\text{Ker}(A) = \{v \in \mathbb{R}^n/Av = 0\}$ es un subespacio vectorial de $\mathbb{R}^n$.

2. El conjunto $U = \{A \in \mathcal{M}_{n\times n}(\mathbb{R})/A^t = A\}$ es un subespacio vectorial de $\mathcal{M}_{n\times n}(\mathbb{R})$.

3. El conjunto $W = \{A \in \mathcal{M}_{2\times 2}(\mathbb{R})/det(A) = 0\}$ no es un subespacio vectorial de $\mathcal{M}_{n\times n}(\mathbb{R})$. Aunque $0 \in W$, veamos que no se cumple la propiedad (2); para ello basta tomar
$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \; A_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$
Es claro que A_1 y A_2 pertenecen a W ya que $det(A_1) = det(A_2) = 0$. Sin embargo,
$$|A_1 + A_2| = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \neq 0 \Longrightarrow A_1 + A_2 \notin W.$$

Definición 4.3 *Si $v_1, v_2, \dots, v_n$ son n vectores de un espacio vectorial V y $\lambda_1, \dots, \lambda_n$ son números reales, entonces cualquier vector de la forma*
$$v = \lambda_1 v_1 + \lambda_2 v_2 + \dots \lambda_n v_n$$
se llama **combinación lineal** *de $v_1, v_2, \dots, v_n$.*

Proposición 4.1 *Un subconjunto no vacío U de un espacio vectorial V es un subespacio vectorial si y sólo si todas las combinaciones lineales de vectores de U pertenecen a U. Es decir, dados $u_1, u_2 \in U, \lambda, \mu \in \mathbb{R}$, entonces $\lambda u_1 + \mu u_2 \in \mathbb{R}$.*

Definición 4.4 *Sea U un subespacio vectorial de un espacio vectorial V. Se dice que un subconjunto S de U es un* **conjunto de generadores** *de U si todo vector de U es combinación lineal de vectores de S. Si S es un conjunto de generadores de U, diremos que U es el* **subespacio generado** *por S.*

En muchas ocasiones la forma más sencilla de probar que un subconjunto U de un espacio vectorial V es un subespacio consiste en encontrar un conjunto de generadores.

Ejemplo 4.3 Sea $U = \{p(x) \in \Pi_2(\mathbb{R})/p(1) = 0\}$.
Consideremos un polinomio arbitrario $p(x) = a + bx + cx^2 \in \Pi_2(\mathbb{R})$. Entonces:

$$p(x) \in U \Longleftrightarrow p(1) = 0 \Longleftrightarrow a + b + c = 0.$$

Podemos reescribir U como:

$$\begin{aligned} U &= \{a + bx + cx^2 \in \Pi_2(\mathbb{R})/a + b + c = 0\} = \{a + bx + cx^2 \in \Pi_2(\mathbb{R})/c = -a - b\} = \\ &= \{a + bx + (-a - b)x^2/a, b \in \mathbb{R}\} = \{a(1 - x^2) + b(x - x^2)/a, b \in \mathbb{R}\} = \\ &= < \{1 - x^2, x - x^2\} > . \end{aligned}$$

Por tanto, U es el subespacio vectorial de $\Pi_2(\mathbb{R})$ generado por $1 - x^2$ y $x - x^2$.

4.2 Independencia lineal

Los conceptos de dependencia e independencia lineal se extienden de manera natural a cualquier espacio vectorial.

Definición 4.5 *Sea V un espacio vectorial y $S = \{v_1, v_2, \ldots, v_n\}$ un subconjunto de V. Se dice que S* **es linealmente dependiente** *(o que forma un sistema* **ligado***) si existen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, no todos nulos, tales que*

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0.$$

Un conjunto de vectores es **linealmente independiente** *(o forma un sistema* **libre***) si de cualquier combinación lineal de ellos igualada a cero*

$$\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n = 0$$

se deduce que $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Se llama **rango** *de un conjunto de vectores al número de vectores linealmente independientes que contiene. Por tanto, un conjunto de n vectores es linealmente independiente si y sólo si su rango es n.*

Si $V=\mathbb{R}^n$, estudiar si un conjunto de vectores S es **libre** se reduce a calcular el rango de la matriz que tiene como filas los vectores de S: un conjunto $S=\{v_1,v_2,\dots,v_p\}$ de vectores de $\mathbb{R}^n$ es libre si y sólo si

$$\operatorname{rg}\begin{pmatrix} v_1^t \\ v_2^t \\ \vdots \\ v_p^t \end{pmatrix} = p.$$

Ejemplo 4.4

1. Sea $S=\{(1,0),(1,-1)\}$. Entonces:

$$\operatorname{rg}(S)=\operatorname{rg}\begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \stackrel{F_{21}(-1)}{=} \operatorname{rg}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 2$$

   Por tanto, S es libre. Visto de otro modo

$$\alpha(1,0)+\beta(1,-1)=(0,0)$$

   sólo admite la solución trivial.

2. Sea $S=\{(2,-4),(-3,6)\}$. Entonces:

$$\operatorname{rg}(S)=\operatorname{rg}\begin{pmatrix} 2 & -4 \\ -3 & 6 \end{pmatrix} = \operatorname{rg}\begin{pmatrix} 2 & -4 \\ 0 & 0 \end{pmatrix} = 1$$

   Por tanto, S no es libre. Esto quiere decir que

$$\alpha(2,-4)+\beta(-3,6)=(0,0)$$

   admite más soluciones que la trivial. Por ejemplo $3(2,-4)+2(-3,6)=(0,0)$.

3. Sea $S=\{(1,2,1,1),(-1,1,0,0),(1,5,2,2)\}$. Entonces:

$$\operatorname{rg}(S)=\operatorname{rg}\begin{pmatrix} 1 & 2 & 1 & 1 \\ -1 & 1 & 0 & 0 \\ 1 & 5 & 2 & 2 \end{pmatrix} \underset{F_{31}(-1)}{\stackrel{F_{21}(1)}{=}} \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 3 & 1 & 1 \\ 0 & 3 & 1 & 1 \end{pmatrix} \stackrel{F_{32}(-1)}{=} \operatorname{rg}\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 2$$

   Por tanto, S no es libre.

Observación 4.1 Si sólo se realizan operaciones elementales por filas en A para determinar una matriz escalonada A' y obtener el rango de S entonces el subespacio generado por S coincide con el subespacio generado por las filas no nulas de A'. Esta propiedad no es cierta si se combinan operaciones de filas y columnas para calcular el rango.

En el ejemplo anterior,

$$U=<S>=<\{(1,2,1,1),(-1,1,0,0),(1,5,2,2)\}>=<\{(1,2,1,1),(0,3,1,1)\}>.$$

Para otros espacios vectoriales, resulta útil la siguiente caracterización de la independencia lineal:

Proposición 4.2 *Un conjunto $S = \{v_1, v_2, \ldots, v_n\}$ de vectores es linealmente independiente si y sólo si se cumple la siguiente propiedad:*

Si $\lambda_1, \ldots, \lambda_n$ son números reales tales que $\lambda_1 v_1 + \lambda_2 v_2 + \cdots + \lambda_n v_n = 0$ entonces necesariamente $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$.

Por ejemplo, el conjunto

$$\left\{ \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 4 & 0 \end{pmatrix} \right\}$$

es libre porque

$$\alpha \begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 1 \\ 4 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Longleftrightarrow \begin{pmatrix} 2 & 0 & 1 \\ 1 & 1 & 1 \\ 3 & 2 & 4 \\ 1 & 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

y la única solución del sistema es $(\alpha, \beta, \gamma) = (0, 0, 0)$ porque el rango de la matriz de coeficientes coincide con el número de incógnitas.

4.3 Bases y dimensión

Definición 4.6 *Un conjunto linealmente independiente de generadores de un espacio vectorial V se llama* **base** *de V.*

El conjunto de generadores de un espacio vectorial no es único.

Teorema 4.1 *Todo espacio vectorial V finito y no nulo posee, al menos, una base.*

Ejemplo 4.5

1. El conjunto $\mathcal{B} = \{1, x, x^2, \ldots, x^n\}$ es una base del espacio de polinomios $\Pi_n(\mathbb{R})$.
2. El conjunto
$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$
es una base de $\mathcal{M}_{2\times 2}(\mathbb{R})$.

Teorema 4.2 *Sea V un espacio vectorial generado por n vectores. Entonces*

1. *Todas las bases de V tienen n elementos.*
2. *Todo conjunto linealmente independiente de n elementos es una base de V.*
3. *Todo conjunto generador de V de n elementos es una base.*

Dimensión

Dado que todas las bases de un espacio vectorial V tienen el mismo número de vectores, el número de vectores de cualquier base de V se llama **dimensión** de V y se denota por $\dim(V)$.

Para los espacios vectoriales que hemos mencionado anteriormente, se tiene:

$$\dim(\Pi_n(\mathbb{R})) = n + 1, \quad \dim \mathcal{M}_{2\times 2}(\mathbb{R}) = 4.$$

En general, $\dim(\mathcal{M}_{p\times n}(\mathbb{R})) = p \times n$.

Observación 4.2 Si $V = \{0\}$ entonces no existe ninguna base de V y, por convenio, definiremos $\dim(V) = 0$.

Cálculo de la dimensión de un subespacio vectorial

En primer lugar, si $V =< \{v_1, v_2, \dots, v_p\} >$ entonces $\dim(V) = \mathrm{rg}(\{v_1, v_2, \dots, v_p\})$.

Por ejemplo: Sea $U =< \{(1, 2, 1, 1), (0, 1, -1, -1), (0, 0, 0, 1)\} >$. Entonces

$$\dim(U) = \mathrm{rg}\begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 3.$$

Ya sabemos que si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, entonces $\mathrm{Ker}(A)$ es un subespacio de $\mathbb{R}^n$ y

$$\dim(\mathrm{Ker}(A)) = n - \mathrm{rg}(A).$$

Esta propiedad se puede extender a cualquier espacio vectorial de dimensión finita V: Si U es un subespacio de V entonces la dimensión de U es igual a la dimensión de V menos el número de ecuaciones linealmente independientes que definen a U.

Por ejemplo, si $U = \{A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R}) / a_{ii} = 0, \forall i = 1, 2, \dots, n\}$ entonces

$$\dim(U) = \dim(\mathcal{M}_{n\times n}(\mathbb{R})) - n = n^2 - n.$$

4.4 Cambio de base en $\mathbb{R}^n$

La siguiente propiedad es una consecuencia inmediata de la definición de base y permite introducir el concepto de vector de coordenadas:

Proposición 4.3 *Sea $\mathcal{B} = \{u_1, u_2, \dots, u_n\}$ una base de $\mathbb{R}^n$. Cada $x \in \mathbb{R}^n$ se puede escribir de modo único como*

$$x = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n.$$

El vector $(\lambda_1, \lambda_2, \dots, \lambda_n)$ se llama **vector de coordenadas** *de x respecto de la base $\mathcal{B}$ y se suele denotar $x = (\lambda_1, \lambda_2, \dots, \lambda_n)_{\mathcal{B}}$. Las coordenadas de un vector respecto de una base son únicas.*

Definición 4.7 *De entre todas las bases del espacio vectorial* $(\mathbb{R}^n, +, \cdot)$ *hay una que recibe el nombre de* **base canónica** *y suele denotarse por*

$$\mathcal{C} = \{e_1, e_2, \ldots, e_n\} \text{ siendo } \begin{cases} e_1 = (1, 0, 0, \ldots, 0) \\ e_2 = (0, 1, 0, \ldots, 0) \\ \vdots \\ e_n = (0, 0, 0, \ldots, 1) \end{cases}$$

Las coordenadas de un vector respecto de la base canónica coinciden con los n valores que componen el vector.

Dado un espacio vectorial V de dimensión finita, $\dim(V) = n$, una vez elegida una base $\mathcal{B}$ para dicho espacio, podemos establecer una relación de uno a uno entre los vectores del espacio V y los del espacio $\mathbb{R}^n$. Es decir, podemos asociar (identificar) cada vector de V con un único elemento de $\mathbb{R}^n$ que representa sus coordenadas respecto de la base elegida. Con esta idea, podemos prescindir de trabajar con los vectores originales (matrices, polinomios, funciones, etc.) y trabajar con sus coordenadas.

Ejemplo 4.6 En $\mathbb{R}^3$ se considera la base $\mathcal{B} = \{(1, 1, 1), (1, 2, 0), (0, 0, 1)\}$.
Calculamos las coordenadas de $x = (1, 0, 0)$ respecto de $\mathcal{B}$:
Si $(1, 0, 0) = (\alpha, \beta, \gamma)_{\mathcal{B}}$ entonces:

$$(1, 0, 0) = \alpha(1, 1, 1) + \beta(1, 2, 0) + \gamma(0, 0, 1) = (\alpha + \beta, \alpha + 2\beta, \alpha + \gamma) \Longleftrightarrow$$

$$\Longleftrightarrow \begin{cases} \alpha + \beta = 1 \\ \alpha + 2\beta = 0 \\ \alpha + \gamma = 0 \end{cases} \Longleftrightarrow \begin{cases} \alpha = 2 \\ \beta = -1 \\ \gamma = -2 \end{cases}$$

Por tanto, $(1, 0, 0) = (2, -1, -2)_{\mathcal{B}}$.

Si $\mathcal{B}$ es una base de $\mathbb{R}^n$ y $x = (\lambda_1, \lambda_2, \ldots, \lambda_n)_{\mathcal{B}}$ entonces denotaremos

$$x_{\mathcal{B}} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{R})$$

Observemos que si consideramos la base canónica $\mathcal{C}$, entonces las coordenadas de un vector $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ respecto de $\mathcal{C}$ son precisamente $(x_1, x_2, \ldots, x_n)$, es decir,

$$x_{\mathcal{C}} = x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathcal{M}_{n\times 1}(\mathbb{R})$$

A continuación veremos cómo cambian las coordenadas de un vector x al cambiar de base.

Definición 4.8 *Sea $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ una base de $\mathbb{R}^n$. Se llama* **matriz de cambio de base** *de $\mathcal{B}$ a la base canónica $\mathcal{C}$ a la matriz $P_{\mathcal{BC}} \in \mathcal{M}_{n\times n}(\mathbb{R})$ cuyas columnas son los vectores de $\mathcal{B}$, es decir,*

$$P_{\mathcal{BC}} = (u_1|u_2|\ldots|u_n).$$

Ejemplo 4.7 Sea $B = \{(1,1,1), (1,2,0), (0,0,1)\}$. La matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ es

$$P_{\mathcal{BC}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

La propiedad que caracteriza a la matriz de cambio de base es la siguiente:

Proposición 4.4 *Si $P_{\mathcal{BC}}$ es la matriz de cambio de base de $\mathcal{B}$ a $\mathcal{C}$ entonces*

$$P_{\mathcal{BC}}\, x_{\mathcal{B}} = x_{\mathcal{C}}, \forall x \in \mathbb{R}^n.$$

<u>Demostración</u>. Sea $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ y $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ su vector de coordenadas respecto de $\mathcal{B}$. Entonces:

$$x = x_{\mathcal{C}} = \lambda_1 u_1 + \lambda_2 u_2 + \ldots + \lambda_n u_n = (u_1|u_2|\ldots|u_n) \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = P_{\mathcal{BC}} x_{\mathcal{B}}$$

□

De modo análogo, si $\mathcal{B}$ y $\mathcal{B}'$ son dos bases de $\mathbb{R}^n$ se define la matriz de cambio de base $P_{\mathcal{B}'\mathcal{B}}$ de $\mathcal{B}'$ a $\mathcal{B}$ como la que tiene la siguiente propiedad:

$$P_{\mathcal{B}'\mathcal{B}}\, x_{\mathcal{B}'} = x_{\mathcal{B}}, \forall x \in \mathbb{R}^n.$$

El cambio de base de $\mathcal{B}'$ a $\mathcal{B}$ se puede hacer utilizando las siguientes propiedades:

Proposición 4.5 *Sean $\mathcal{B}$ y $\mathcal{B}'$ dos bases de $\mathbb{R}^n$. Entonces:*

1. *$P_{\mathcal{BC}}$ es inversible y $(P_{\mathcal{BC}})^{-1} = P_{\mathcal{CB}}$.*
2. *$P_{\mathcal{B}'\mathcal{B}} = P_{\mathcal{CB}} P_{\mathcal{B}'\mathcal{C}} = (P_{\mathcal{BC}})^{-1} P_{\mathcal{B}'\mathcal{C}}$.*

Ejemplo 4.8 La matriz de cambio de base de $\mathcal{C} = \{(1,0,0), (0,1,0), (0,0,1)\}$ a $\mathcal{B} = \{(1,1,1), (1,2,0), (0,0,1)\}$ es

$$P_{\mathcal{CB}} = (P_{\mathcal{BC}})^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 1 & 0 \\ -2 & 1 & 1 \end{pmatrix}$$

4.5 Ecuaciones de los subespacios

Definición 4.9 *Se denominan* **ecuaciones paramétricas** *de un subespacio U de V a las relaciones que ligan las coordenadas de un vector cualquiera $x \in U$ respecto de las bases $\mathcal{B}'$ de U y $\mathcal{B}$ de V.*

$$\forall x \in U, x = \sum_{i=1}^{r} \lambda_i u_i, \qquad \forall x \in U \subseteq V, x = \sum_{j=1}^{n} x_j v_j$$

$$u_1, u_2, \ldots, u_n \in V \Longrightarrow u_i = a_{1i}v_1 + a_{2i}v_2 + \cdots + a_{ni}v_n,\ i = 1, 2, \ldots, r$$

$$x = \lambda_1 u_1 + \lambda_2 u_2 + \cdots + \lambda_r u_r = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \Longrightarrow$$

$$\lambda_1(a_{11}v_1 + \cdots + a_{n1}v_n) + \cdots + \lambda_r(a_{1r}v_1 + \cdots + a_{nr}v_n) = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n \Longrightarrow$$

$$(\lambda_1 a_{11} + \cdots + \lambda_r a_{1r})v_1 + \cdots + (\lambda_1 a_{n1} + \cdots + \lambda_r a_{nr})v_n = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

y al ser únicas las coordenadas de un vector respecto a una base, se tiene:

$$\left.\begin{array}{l} x_1 = \lambda_1 a_{11} + \cdots + \lambda_r a_{1r} \\ \vdots \\ x_n = \lambda_1 a_{n1} + \cdots + \lambda_r a_{nr} \end{array}\right\} \text{ecuaciones paramétricas de } S$$

Se trata de un sistema de n ecuaciones con r incógnitas siendo $r < n$.

Definición 4.10 *Si en el sistema anterior de eliminan los parámetros $\lambda_1, \lambda_2, \ldots, \lambda_n$ se obtiene las* **ecuaciones cartesianas** *o* **implícitas** *del subespacio U.*

Ejemplo 4.9 Hallar las ecuaciones paramétricas e implícitas del subespacio U de $\mathbb{R}^5$ generado por los vectores

$$U =< \{(1, 2, 1, 0, 0), (0, -1, 1, 1, 0), (1, 0, -1, 0, 1), (1, 1, 2, 1, 0)\} >$$

Determinamos, en primer lugar una base de U

$$\begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & 0 & 1 \\ 1 & 1 & 2 & 1 & 0 \end{pmatrix} \xrightarrow[F_{41}(-1)]{F_{31}(-1)} \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & -2 & -2 & 0 & 1 \\ 0 & -1 & 1 & 1 & 0 \end{pmatrix}$$

$$\xrightarrow[F_{42}(-1)]{F_{32}(-2)} \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & -4 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

por tanto, una base de U es $\mathcal{B} = \{(1, 2, 1, 0, 0), (0, -1, 1, 1, 0), (1, 0, -1, 0, 1)\}$ y $\dim(U) = 3$. Por tanto, cualquier vector $x \in U$ puede expresarse de la forma:

$$x = (x_1, x_2, x_3, x_4, x_5) = \lambda_1(1, 2, 1, 0, 0) + \lambda_2(0, -1, 1, 1, 0) + \lambda_3(1, 0, -1, 0, 1)$$

de donde

$$\left.\begin{array}{l} x_1 = \lambda_1 + \lambda_3 \\ x_2 = 2\lambda_1 - \lambda_2 \\ x_3 = \lambda_1 + \lambda_2 - \lambda_3 \\ x_4 = \lambda_2 \\ x_5 = \lambda_3 \end{array}\right\} \text{ ecuaciones paramétricas de } U.$$

Obsérvese que las ecuaciones paramétricas no son únicas, dependen de las bases elegidas. Por ejemplo, otra base de U está formada por las filas no nulas y finales de la matriz escalonada resultante:

$$\mathcal{B}' = \{(1,2,1,0,0),(0,-1,1,1,0),(0,0,-4,-2,1)\},$$

por lo que podemos elegir libremente la base que mejor nos convenga. Vamos a hallar ahora unas ecuaciones implícitas a partir de las anteriores ecuaciones paramétricas:

$$\begin{pmatrix} 1 & 0 & 1 & x_1 \\ 2 & -1 & 0 & x_2 \\ 1 & 1 & -1 & x_3 \\ 0 & 1 & 0 & x_4 \\ 0 & 0 & 1 & x_5 \end{pmatrix} \xrightarrow[F_{31}(-1)]{F_{21}(-2)} \begin{pmatrix} 1 & 0 & 1 & x_1 \\ 0 & -1 & -2 & x_2 - 2x_1 \\ 0 & 1 & -2 & x_3 - x_1 \\ 0 & 1 & 0 & x_4 \\ 0 & 0 & 1 & x_5 \end{pmatrix} \xrightarrow[F_{42}(1)]{F_{32}(1)}$$

$$\begin{pmatrix} 1 & 0 & 1 & x_1 \\ 0 & -1 & -2 & x_2 - 2x_1 \\ 0 & 0 & -4 & -3x_1 + x_2 + x_3 \\ 0 & 0 & 2 & -2x_1 + x_2 + x_4 \\ 0 & 0 & 1 & x_5 \end{pmatrix} \xrightarrow[4F_5+F_3]{2F_4+F_3} \begin{pmatrix} 1 & 0 & 1 & x_1 \\ 0 & -1 & -2 & x_2 - 2x_1 \\ 0 & 0 & -4 & -3x_1 + x_2 + x_3 \\ 0 & 0 & 0 & -7x_1 + 3x_2 + x_3 + 2x_4 \\ 0 & 0 & 0 & -3x_1 + x_2 + x_3 + 4x_5 \end{pmatrix}$$

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

Estas son unas ecuaciones implícitas de U.

Observación 4.3 Se puede pasar de las ecuaciones cartesianas a las ecuaciones paramétricas, simplemente resolviendo el sistema de ecuaciones.

4.6 Operaciones con subespacios vectoriales

Veamos a continuación cómo se comporta la noción de subespacio vectorial con las operaciones entre conjuntos. Antes, definimos una nueva operación suma del siguiente modo. Sean W_1 y W_2 subespacios vectoriales de V y definimos la **suma** de ambos como

$$W_1 + W_2 = \{u + v / u \in W_1, v \in W_2\}.$$

Proposición 4.6 *Sean W_1 y W_2 subespacios vectoriales de V. Entonces*

1. *$W_1 \cap W_2$ es un subespacio vectorial de V.*
2. *$W_1 \cup W_2$ no es un subespacio vectorial de V en general.*
3. *$W_1 + W_2$ es un subespacio vectorial de V.*

Definición 4.11 *Se dirá que la suma de dos subespacios es* **directa** *si su intersección es el vector 0, esto es,* $W_1 \cap W_2 = \{0\}$*. Se denota* $W_1 \oplus W_2$*.*

La intersección y la suma de subespacios vectoriales son operaciones que permiten construir nuevos subespacios vectoriales más pequeños o más grandes.

Ejemplo 4.10 Dado $\mathbb{R}^3$ y los subespacios $W_1 = \{(x, y, z) \in \mathbb{R}^3 / x + y = 0, x + z = 0\}$ y $W_2 = \{(x, y, z) \in \mathbb{R}^3 / x = 0, y = 0\}$, calculemos su suma e intersección. Para ello, démonos cuenta que $(x, y, z) \in W_1 \cap W_2$ si satisface a la vez las ecuaciones que definen ambos subespacios, esto es

$$\begin{cases} x + y = 0, \\ x + z = 0, \\ x = 0, \\ y = 0 \end{cases}$$

de donde obtenemos resolviendo el sistema que $x = y = z = 0$, o lo que es lo mismo

$$W_1 \cap W_2 = \{(0, 0, 0)\}.$$

Por otro lado, las ecuaciones paramétricas de W_1 y W_2 son

$$\begin{cases} x = -\lambda, \\ y = \lambda, \\ z = \lambda \end{cases}, \lambda \in \mathbb{R} \qquad \begin{cases} x = 0, \\ y = 0, \\ z = \lambda \end{cases}, \lambda \in \mathbb{R}$$

respectivamente. Entonces es fácil ver que $W_1 =< \{(-1, 1, 1)\} >$ y $W_2 =< \{(0, 0, 1)\} >$. Así $(x, y, z) \in W_1 + W_2$ si y sólo si

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

de donde tenemos el sistema

$$\begin{cases} x = -\alpha, \\ y = \alpha, \\ z = \alpha + \beta \end{cases}$$

que es compatible. Al calcular los rangos de las matrices asociadas

$$\left(\begin{array}{rr|c} -1 & 0 & x \\ 1 & 0 & y \\ 1 & 1 & z \end{array} \right) \xrightarrow[F_{31}(1)]{F_{21}(1)} \left(\begin{array}{rr|c} -1 & 0 & x \\ 0 & 0 & y + x \\ 0 & 1 & z + y \end{array} \right)$$

tenemos que ambos son iguales a dos si $y + x = 0$, por lo que

$$W_1 + W_2 = \{(x, y, z) \in \mathbb{R}^3 / x + y = 0\}.$$

Proposición 4.7 *Sean* W_1 *y* W_2 *subespacios vectoriales de* V *de dimensión finita. Entonces*

$$dim(W_1 + W_2) = dim(W_1) + dim(W_2) - dim(W_1 \cap W_2).$$

En particular, si la suma es directa

$$dim(W_1 \oplus W_2) = dim(W_1) + dim(W_2).$$

Ejemplo 4.11 El resultado anterior tiene su aplicación a ejemplos como el siguiente. Sean

$$W_1 = \{(x,y,z) \in \mathbb{R}^3 / x+y+z=0\}, \quad W_2 = \{(x,y,z) \in \mathbb{R}^3 / x+y=0, y+z=0\}.$$

Sus dimensiones son 2 y 1, respectivamente. Por otra parte, $(x,y,z) \in W_1 \cap W_2$ si se satisfacen

$$\begin{cases} x+y+z=0, \\ x+y=0, \\ y+z=0. \end{cases}$$

Si resolvemos este sistema obtenemos que $x=y=z=0$, por lo que $W_1 \cap W_2 = \{(0,0,0)\}$. Entonces

$$\dim(W_1 \oplus W_2) = \dim(W_1) + \dim(W_2) = 2+1 = 3,$$

y dado que $W_1 \oplus W_2 \subseteq \mathbb{R}^3$, se verifica que $W_1 \oplus W_2 = \mathbb{R}^3$.

4.7 Espacio vectorial euclídeo

Definición 4.12 *Sea V un espacio vectorial sobre $\mathbb{R}$. Se denomina* **producto escalar** *a una aplicación*

$$<,>: V \times V \longrightarrow \mathbb{R}$$

que a cada par de vectores $u, v \in V$ les asocia un número real $< u, v > \in \mathbb{R}$ que verifica las siguientes propiedades

1. $< u, u > \geq 0, \forall u \in V$. *Además,* $< u, u >= 0 \Longleftrightarrow u = 0$.
2. $< \alpha u + \beta v, w >= \alpha < u, w > + \beta < v, w >, \forall \alpha, \beta \in \mathbb{R},\ \forall u, v, w \in V$.
3. $< u, v >=< v, u >,\ \forall u, v \in V$.

Observación 4.4 En ocasiones el producto escalar de dos vectores $u, v \in V$ se escribe como $u \cdot v$.

Definición 4.13 *Un espacio vectorial real sobre el que se ha definido un producto escalar se dice que es un* **espacio vectorial euclídeo** *y se representa por el par $(V, <,>)$.*

El producto escalar no es único pues pueden definirse diferentes aplicaciones sobre un mismo espacio vectorial.

Definición 4.14 *Se define el* **producto escalar usual** *de dos vectores $x = (x_1, x_2, \ldots, x_n)$ e $y = (y_1, y_2, \ldots, y_n)$ de $\mathbb{R}^n$ como*

$$< x, y >= x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n = \sum_{i=1}^{n} x_i y_i.$$

Definición 4.15 *Sea $(V, <,>)$ un espacio vectorial euclídeo. Se denomina* **norma** *del vector $v \in V$ al número real positivo*

$$||v|| = +\sqrt{< v, v >}$$

que tiene sentido ya que $< v, v > \geq 0\ \forall v \in V$.

Proposición 4.8 *Propiedades de la norma*

- $||x|| \geq 0, \forall x \in V$. *Además* $||x|| = 0 \Longleftrightarrow x = 0$.
- $||\lambda x|| = |\lambda| \, ||x||, \ \forall x \in V$ *y* $\forall \lambda \in \mathbb{R}$
- *Ley del paralelogramo:*
$$||x+y||^2 + ||x-y||^2 = 2(||x||^2 + ||y||^2), \ \forall x, y \in V$$
- *Desigualdad de Cauchy-Schwartz:*
$$| < x, y > | \leq ||x|| \, ||y||, \ \forall x, y \in V$$
- *Desigualdad de Minkowski (o desigualdad triangular):*
$$||x+y|| \leq ||x|| + ||y||, \ \forall x, y \in V$$
- $||x|| - ||y|| \leq ||x - y||, \ \forall x, y \in V$

Como consecuencia de la desigualdad de Cauchy-Schwartz, dados dos vectores $x, y \in V$ se tiene
$$-1 \leq \frac{< x, y >}{||x||||y||} \leq 1.$$
De forma que se puede hacer la siguiente definición

Definición 4.16 *Sean $x, y \in V$ no nulos. Se define el* **ángulo** *que forman los vectores x e y como el número real α tal que*
$$\cos(\alpha) = \frac{< x, y >}{||x||||y||}$$
De modo que el producto escalar verifica la fórmula
$$< x, y >= ||x||||y|| \cos(\alpha)$$

Ejemplo 4.12 El ángulo α generado por los vectores $(1,1)$ y $(1,0)$ se calcula mediante la expresión
$$\cos(\alpha) = \frac{< (1,1), (1,0) >}{||(1,1)||||(1,0)||} = \frac{1}{\sqrt{2}}$$

4.8 Ortogonalidad

Definición 4.17 *Sea $(V, <,>)$ un espacio vectorial euclídeo. Se dice que los vectores $x, y \in V$, no nulos, son* **ortogonales** *respecto del producto escalar $<,>$ si se verifica que $< x, y >= 0$.*

Teorema 4.3 *Teorema de Pitágoras*
Sea $(V, <,>)$ un espacio vectorial euclídeo. Dos vectores $x, y \in V$ son ortogonales si, y sólo si,
$$||x+y||^2 = ||x||^2 + ||y||^2$$

Definición 4.18 *Sea $(V, <,>)$ un espacio vectorial euclídeo. Se dice que los vectores $x, y \in V$, no nulos, son* **ortonormales** *respecto del producto escalar $<,>$ si se verifica que $< x, y >= 0$ y $||x|| = ||y|| = 1$.*

4.9 Bases ortonormales

Definición 4.19 *Una base $\mathcal{B} = \{u_1, u_2, \ldots, u_p\}$ de un subespacio vectorial U de $\mathbb{R}^n$ es una* **base ortonormal** *si todos los vectores son unitarios y ortogonales entre sí, es decir, $< u_i, u_j >= 0$ si $i \neq j$ y $||u_i|| = 1$ para todo $i = 1, 2, \ldots, p$.*

El **procedimiento de ortonormalización de Gram-Schmidt** permite calcular una base ortonormal a partir de una base de U. Sea $\mathcal{B} = \{v_1, v_2, \ldots, v_p\}$ una base de un subespacio vectorial U de $\mathbb{R}^n$. Es posible construir una base ortonormal $\mathcal{B}' = \{u_1, u_2, \ldots, u_p\}$ de U a partir de $\mathcal{B}$ del siguiente modo:

1. Se construye u_1 dividiendo v_1 por su norma:
$$u_1 = \frac{v_1}{||v_1||}.$$
2. Para cada $i \geq 2$ se construye u_i en dos etapas:
   (a) Se calcula un vector $\tilde{u}_i$ dado por:
   $$\tilde{u}_i = vi - \sum_{j=i}^{i-1}(v_i^t u_j)u_j = v_i - (v_i^t u_1)u_1 - \cdots - (v_i^t u_{i-1})u_{i-1}.$$
   (b) Se normaliza el vector $\tilde{u}_i$:
   $$u_i = \frac{\tilde{u}_i}{||\tilde{u}_i||}.$$

Ejemplo 4.13 Vamos a calcular una base ortonormal del subespacio $U =< \{(1, 0, 1), (1, 1, 1)\} >$. Denotemos por $v_1 = (1, 0, 1), v_2 = (1, 1, 1)$. Entonces:

$$u_1 = \frac{v_1}{||v_1||} = \frac{1}{\sqrt{2}}(1, 0, 1) = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$$

$$\tilde{u}_2 = v_2 - < v_2, u_1 > u_1 = (1, 1, 1) - \frac{2}{\sqrt{2}}\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right) = (1, 1, 1) - (1, 0, 1) = (0, 1, 0)$$

$$u_2 = \frac{\tilde{u}_2}{||\tilde{u}_2||} = (0, 1, 0)$$

El conjunto $\mathcal{B}' = \{u_1, u_2\} = \left\{\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right), (0, 1, 0)\right\}$ es una base ortonormal de U.

El siguiente resultado relaciona las bases ortonormales con las matrices ortogonales y será de utilidad en el Capítulo 6.

Proposición 4.9 *Una matriz $P \in \mathcal{M}_{n\times n}(\mathbb{R})$ es ortogonal si y sólo si sus columnas son una base ortonormal de $\mathbb{R}^n$. En particular, si $\mathcal{B}$ es una base ortonormal, la matriz de cambio de coordenadas $P_{\mathcal{BC}}$ es una matriz ortogonal.*

Tema 5

Aplicaciones lineales

En este capítulo se interpretan las matrices como aplicaciones lineales.

5.1 Aplicación lineal y matriz asociada

Vamos a recordar algunas cuestiones previas sobre aplicaciones.

Definición 5.1 *Una aplicación f entre dos conjuntos X e Y es cualquier regla que nos permite asignar a cada elemento x del conjunto inicial X un único elemento y del conjunto final Y. Se denota por $f : X \longrightarrow Y$.*

- $\forall x \in X, \exists\, y \in Y$ *tal que* $f(x) = y$.
- $x = y \Longrightarrow f(x) = f(y)$

Definición 5.2 *Una aplicación $f : X \longrightarrow Y$ diremos que es*

- **Inyectiva** *si*

$$x \neq y \Longrightarrow f(x) \neq f(y) \text{ o lo que es lo mismo } f(x) = f(y) \Longrightarrow x = y.$$

- **Sobreyectiva** *si* $\forall y \in Y, \exists\, x \in X$ *tal que* $f(x) = y$.
- **Biyectiva** *si es simultáneamente inyectiva y sobreyectiva.*

Definición 5.3 *Sean V y W dos espacios vectoriales. Una aplicación $L : V \to W$ es* **lineal** *si cumple las siguientes propiedades:*

1. $L(x + y) = L(x) + L(y), \forall x, y \in V$.
2. $L(\lambda x) = \lambda L(x), \forall \lambda \in R, \forall x \in V$.

De estas propiedades se obtiene por inducción que

$$L(\lambda_1 v_1 + \lambda_2 v_2 + \ldots + \lambda_n v_n) = \lambda_1 L(v_1) + \lambda_2 L(v_2) + \ldots + \lambda_n L(v_n),$$

donde $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R}, v_1, v_2, \ldots, v_n \in V$. En otras palabras, si $L : \mathbb{R}^n \to \mathbb{R}^p$ es una aplicación lineal entonces la imagen de la combinación lineal de n vectores de V es igual a la combinación lineal de sus imágenes. En particular $L(0) = 0$.

Matriz asociada a una aplicación lineal

Una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ define una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^p$ dada por $L(x) = Ax$, donde $x \in \mathbb{R}^n$ es un vector columna. Recíprocamente, el siguiente resultado prueba que una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^p$ siempre se puede escribir en la forma $L(x) = Ax$ para una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$.

Proposición 5.1 *Dada una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^p$, existe una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ tal que $L(x) = Ax, \forall x \in \mathbb{R}^n$.*

Demostración. Denotemos por $\mathcal{C} = \{e_1, e_2, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$.
Sea $x = (x_1, x_2, \ldots, x_n) = x_1e_1 + x_2e_2 + \ldots + x_ne_n \in \mathbb{R}^n$. Como L es una aplicación lineal:

$$L(x) = L(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = x_1L(e_1) + x_2L(e_2) + \ldots + x_nL(e_n) =$$

$$(L(e_1)|L(e_2)|\ldots|L(e_n)) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = Ax.$$

□

La matriz A se llama matriz asociada a L y **sus columnas son las imágenes de los vectores de la base canónica**. En la práctica, la matriz asociada a una aplicación lineal se puede obtener directamente.

Ejemplo 5.1 Sea $L : \mathbb{R}^3 \to \mathbb{R}^2$ definida por $L(x, y, z) = (x + 2y - z, y + 4z)$. Entonces:

$$L(x, y, z) = \begin{pmatrix} x + 2y - z \\ y + 4z \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

La matriz asociada a L es

$$A = \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & 4 \end{pmatrix} = (L(e_1)|L(e_2)|L(e_3)) \in \mathcal{M}_{2\times 3}(\mathbb{R}),$$

donde

$$L(1, 0, 0) = (1, 0),\ L(0, 1, 0) = (2, 1),\ L(0, 0, 1) = (-1, 4).$$

5.2 Núcleo e imagen de una aplicación lineal

Definición 5.4 *Sea $L : V \to W$ una aplicación lineal. Se define el* **núcleo** *de L como*

$$Ker(L) = \{v \in V/L(v) = 0\}.$$

La **imagen** *de L se define como el subespacio formado por todos los vectores de W que son imagen de algún vector de V por la aplicación L:*

$$Im(L) = \{w \in W/w = L(v), v \in V\}.$$

Proposición 5.2 *Sea $L : V \to W$ una aplicación lineal. Entonces Ker(L) es un subespacio vectorial de V.*

Demostración. Para ello sean $\alpha, \beta \in \mathbb{R}$ y $u, v \in \mathrm{Ker}(L)$ y comprobemos que $\alpha u + \beta v \in \mathrm{Ker}(L)$ comprobando que su imagen es nula.

$$L(\alpha u + \beta v) = \alpha L(u) + \beta L(v) = \alpha 0 + \beta 0 = 0.$$

□

Proposición 5.3 *Sea $L : V \to W$ una aplicación lineal. Entonces Im(L) es un subespacio vectorial de W.*

Demostración. Para ello sean $\alpha, \beta \in \mathbb{R}$ y $u, v \in \mathrm{Im}(L)$ y tenemos que ver si $\alpha u + \beta v \in \mathrm{Im}(L)$. Dado que $u, v \in \mathrm{Im}(L)$, entonces existen $u_1, v_1 \in V$ tales que $L(u_1) = u$ y $L(v_1) = v$. De forma que por la linealidad de L tenemos

$$\alpha u + \beta v = \alpha L(u_1) + \beta L(v_1) = L(\alpha u_1 + \beta v_1) \in W.$$

□

Proposición 5.4 *Sea $L : V \to W$ una aplicación lineal entonces*

$$\mathit{dim}(\mathit{Ker}(L)) + \mathit{dim}(\mathit{Im}(L)) = \mathit{dim}(V).$$

Proposición 5.5 *Una aplicación lineal $L : V \to W$ queda definida cuando conocemos la imagen $L(v_i)$ de una base $\mathcal{B} = \{v_1, \ldots, v_n\}$ de V.*

Proposición 5.6 *Sea $L : \mathbb{R}^n \to \mathbb{R}^p$ una aplicación lineal y $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ su matriz asociada. Entonces*

$$\mathit{Ker}(L) = \mathit{Ker}(A) = \{x \in \mathbb{R}^n / Ax = 0\}.$$

$$\mathit{Im}(L) = \mathit{Im}(A) = \{Ax, x \in \mathbb{R}^n\}.$$

es decir, la imagen de L está generada por las columnas de A.

Demostración. Denotemos por $\mathcal{C} = \{e_1, e_2, \ldots, e_n\}$ la base canónica de $\mathbb{R}^n$. Teniendo en cuenta que $A = (L(e_1)|L(e_2)| \ldots |L(e_n))$ y que L es una aplicación lineal:

$$\begin{aligned} b \in \mathrm{Im}(L) \quad & \Longleftrightarrow \exists\, x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n / b = L(x) \Longleftrightarrow \\ & \Longleftrightarrow b = L(x_1e_1 + x_2e_2 + \ldots + x_ne_n) = x_1L(e_1) + x_2L(e_2) + \ldots + x_nL(e_n) \Longleftrightarrow \\ & \Longleftrightarrow b \in < \{L(e_1), L(e_2), \ldots, L(e_n)\} \Longleftrightarrow b \in \mathrm{Im}(A). \end{aligned}$$

□

La fórmula de las dimensiones $\dim(\mathrm{Ker}(A)) + rg(A) = n$ para una matriz $A \in \mathcal{M}_{p\times n}(\mathbb{R})$ se reescribe para una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^p$ como

$$\dim(\mathrm{Ker}(L)) + \dim(\mathrm{Im}(L)) = n = \dim(\mathbb{R}^n).$$

Ejemplo 5.2 Se considera la aplicación lineal $L : \mathbb{R}^4 \to \mathbb{R}^3$ definida por

$$L(x, y, z, t) = (x + y + z, y - 2z + t, 2x + y + 4z - t).$$

Vamos a calcular una base de $\mathrm{Ker}(L)$ y otra de $\mathrm{Im}(L)$. La matriz asociada es

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

Por tanto, $\mathrm{Ker}(L) = \mathrm{Ker}(A) = \{x \in \mathbb{R}^4 / Ax = 0\}$. Para resolver el sistema, hacemos operaciones elementales sobre las filas de la matriz de coeficientes:

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 2 & 1 & 4 & -1 \end{pmatrix} \xrightarrow{F_{31}(-2)} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & -1 & 2 & -1 \end{pmatrix} \xrightarrow{F_{32}(1)} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{F_{12}(-1)} \begin{pmatrix} 1 & 0 & 3 & -1 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Así,

$$\begin{aligned} \mathrm{Ker}(L) &= \{(x, y, z, t) \in \mathbb{R} / x = -3z + t, y = 2z - t\} = \{(-3z + t, 2z - t, z, t)/z, t \in \mathbb{R}\} = \\ &= \{z(-3, 2, 1, 0) + t(1, -1, 0, 1)/z, t \in \mathbb{R}\} =< \{(-3, 2, 1, 0), (1, -1, 0, 1)\} > . \end{aligned}$$

Por tanto, $\dim(\mathrm{Ker}(L)) = 2$ y una base de $\mathrm{Ker}(L)$ es

$$\mathcal{B}_1 = \{(-3, 2, 1, 0), (1, -1, 0, 1)\}.$$

Por otra parte, la imagen de L está generada por las columnas de A:

$$\mathrm{Im}(L) = \mathrm{Im}(A) =< \{(1, 0, 2), (1, 1, 1), (1, -2, 4), (0, 1, -1)\} > .$$

Para calcular una base de la imagen de L hacemos operaciones elementales para eliminar los vectores linealmente dependientes:

$$\begin{pmatrix} 1 & 0 & 2 \\ 1 & 1 & 1 \\ 1 & -2 & 4 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow[F_{31}(-1)]{F_{21}(-1)} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & -2 & 2 \\ 0 & 1 & -1 \end{pmatrix} \xrightarrow[F_{42}(-1)]{F_{32}(2)} \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Por tanto, $\dim(\mathrm{Im}(L)) = 2$ y una base de $\mathrm{Im}(L)$ es

$$\mathcal{B}_2 = \{(1, 0, 2), (0, 1, -1)\}.$$

Se puede comprobar que se cumple la relación general

$$\dim(\mathrm{Ker}(L)) + \dim(\mathrm{Im}(L)) = 2 + 2 = 4 = \dim(\mathbb{R}^4).$$

5.3 Clasificación de las aplicaciones lineales

Definición 5.5

- *Si* $L : V \to W$ *se denomina* **homomorfismo**.
- *Si* $L : V \to V$ *se denomina* **endomorfismo**.
- $L : V \to W$ *se dice que es* **inyectiva** *o que es un* **monomorfismo** *si y sólo si* $Ker(L) = \{0\} \Longleftrightarrow dim(Ker(L)) = 0$
- $L : V \to W$ *se dice que es* **sobreyectiva** *o que es un* **epimorfismo** *si y sólo si* $dim(Im(L)) = dim(W)$
- $L : V \to W$ *se dice que es* **biyectiva** *o que es un* **isomorfismo** $\Longleftrightarrow V \cong W$
- $L : V \to W$ *se dice que es un* **automorfismo** *si y sólo si es un endomorfismo y un isomorfismo.*

Ejemplo 5.3

1. La aplicación lineal del ejemplo anterior no es inyectiva ($\dim(\mathrm{Ker}(L)) \neq 0$) ni sobreyectiva ($\dim(\mathrm{Im}(L)) = 2 \neq 3 = \dim(\mathbb{R}^3)$).
2. Si $\dim(V) > \dim(W)$, entonces la aplicación nunca puede ser inyectiva.
3. Si $\dim(V) < \dim(W)$, entonces la aplicación nunca puede ser sobreyectiva.

5.4 Inversas de aplicaciones lineales

El siguiente resultado muestra qué aplicaciones lineales son inversibles y cómo calcular la aplicación inversa.

Proposición 5.7 *Sean* V *y* W *dos espacios vectoriales y sea* $L : V \longrightarrow W$ *un isomorfismo (aplicación lineal biyectiva). Entonces:*

1. *La aplicación* $L^{-1} : W \longrightarrow V$ *es lineal.*
2. *Si* $\mathcal{B}$ *es una base de* V *y* $\mathcal{B}'$ *es una base de* W *, entonces*

$$P_{\mathcal{B}\mathcal{B}'}(L^{-1}) = (P_{\mathcal{B}'\mathcal{B}}(L))^{-1}.$$

Proposición 5.8 *Sea* $L : \mathbb{R}^n \to \mathbb{R}^n$ *una aplicación lineal y sea* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *su matriz asociada. Entonces* L *es inversible si y sólo si* A *es inversible. Además, la matriz asociada a* L^{-1} *es* A^{-1}.

Ejemplo 5.4 Consideremos la aplicación lineal $L : \mathbb{R}^2 \to \mathbb{R}^2$ dada por $L(x, y) = (x+y, 2x+y)$. Su matriz asociada es

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

Como $|A| = -1 \neq 0, A$ es inversible y por tanto L es inversible. La matriz asociada a L^{-1} es

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

y en consecuencia la aplicación inversa $L^{-1} : \mathbb{R}^2 \to \mathbb{R}^2$ está definida por

$$L^{-1}(x, y) = A^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -x + y \\ 2x - y \end{pmatrix}$$

5.5 Transformaciones ortogonales

Definición 5.6 *Se dice que una aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^n$ es una* **transformación ortogonal** *si conserva el producto escalar, es decir, si para cada par de vectores x e y de $\mathbb{R}^n$ se cumple que*

$$(L(x)) \cdot L(y) = x \cdot y.$$

Observemos que si A es la matriz asociada a la aplicación L entonces

$$(L(x)) \cdot L(y) = (Ax)^t Ay = x^t A^t Ay.$$

De esta relación se obtiene el siguiente resultado que caracteriza las transformaciones ortogonales:

Proposición 5.9 *Sea $L : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal y sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ su matriz asociada. Entonces L es una transformación ortogonal si y sólo si A es una matriz ortogonal.*

Es fácil probar que las transformaciones ortogonales conservan la norma, la distancia y el ángulo. Por esta razón se suelen llamar **movimientos rígidos**. En $\mathbb{R}^2$ las únicas transformaciones ortogonales son giros alrededor del origen y las simetrías axiales con ejes por el origen. En $\mathbb{R}^3$ las únicas son la rotación alrededor de un eje por el origen, la simetría especular respecto a un plano que contiene al origen y la simetría rotacional.

Definición 5.7 *Un movimiento rígido o* **isometría** *es una aplicación $F : \mathbb{R}^n \to \mathbb{R}^n$ que conserva las distancias.*

Son isometrías las transformaciones ortogonales, las traslaciones y la composición de ambas. Se demuestra que toda isometría es una transformación ortogonal o bien una traslación compuesta con una transformación ortogonal.

Tema 6

Diagonalización

Los conceptos principales de este capítulo son los de autovalor y autovector de una matriz cuadrada. Se introduce el polinomio característico para el cálculo de autovalores y se dan aplicaciones a la diagonalización de matrices. También se introduce el concepto de valor singular y su aplicación en la obtención de la mejor aproximación de rango k de una matriz.

6.1 Autovalores y autovectores

Definición 6.1 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Un vector v es un* **autovector** *de A si $v \neq 0$ y existe un escalar λ tal que $Av = \lambda v$. El escalar λ se llama* **autovalor** *de A asociado al autovector v.*

Aunque en la mayoría de las aplicaciones que veremos este curso trabajaremos con autovalores reales y por tanto el autovector es un vector de $\mathbb{R}^n$, veremos que es posible que el escalar λ sea complejo. En ese caso el autovector asociado será un vector $v \in \mathbb{C}^n$.

Definición 6.2 *El conjunto de todos los autovalores de una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ se llama* **espectro** *de A y se denota $Sp(A)$.*

Ejemplo 6.1

1. Consideremos la matriz
$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{R})$$
Veamos que $\lambda = 3$ es un autovalor de A y $v = (1,1,1)$ es un autovector asociado a dicho autovalor:
$$A\,v = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \lambda\, v$$

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

no tiene autovalores reales. Sin embargo, $\lambda = i \in \mathrm{Sp}(A)$:

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

6.1.1 Cálculo de autovalores: polinomio característico

La forma de calcular los autovalores de una matriz la proporciona el siguiente resultado:

Teorema 6.1 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y sea λ un escalar. Entonces $\lambda \in Sp(A) \Longleftrightarrow det(A-\lambda\, I) = 0$. En consecuencia, $Sp(A) = \{\lambda \in \mathbb{C} / det(A - \lambda\, I) = 0\}$.*

Demostración. Observemos que

$$A\,v = \lambda\,v \Longleftrightarrow Av - \lambda v = 0 \Longleftrightarrow (A - \lambda\, I)v = 0 \Longleftrightarrow v \in \mathrm{Ker}(A - \lambda\, I).$$

Por tanto,

$$\lambda \in \mathrm{Sp}(A) \Longleftrightarrow \mathrm{Ker}(A - \lambda\, I) \neq \{0\} \Longleftrightarrow |A - \lambda\, I| = 0.$$

□

Definición 6.3 *Si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, se llama* **polinomio característico** *de A al polinomio definido por $q_A(x) = det(A - x\, I)$.*

El teorema anterior dice que los autovalores de A son las raíces de su polinomio característico.

Ejemplo 6.2 Sea

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \in \mathcal{M}_{3\times 3}(\mathbb{R})$$

El polinomio característico de A es

$$q_A(x) = |A - x\, I| = \begin{vmatrix} 1-x & 2 \\ 2 & 1-x \end{vmatrix} = x^2 - 2x - 3.$$

Los autovalores de A son las raíces de $q_A(x)$. En este caso,

$$x^2 - 2x - 3 = 0 \Longleftrightarrow x = \frac{2 \pm \sqrt{16}}{2} = \frac{2 \pm 4}{2},$$

los autovalores de A son $\lambda_1 = 3, \lambda_2 = -1$.

Si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ entonces su polinomio característico tiene grado exactamente n y su coeficiente principal es $(-1)^n$. Es decir,

$$q_A(x) = (-1)^n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

Recordamos ahora algunas notas sobre raíces de polinomios necesarias para enunciar otros resultados sobre el polinomio característico.

- Sea $p(x)$ un polinomio de grado n con coeficientes en $\mathbb{R}$. Se dice que λ es una **raíz** de $p(x)$ de **multiplicidad** k si existe un polinomio $p_1(x)$ tal que $p(x) = (x-\lambda)^k p_1(x)$ y $p_1(\lambda) \neq 0$.
- Un polinomio $p(x)$ de grado n con coeficientes reales tiene exactamente n raíces en $\mathbb{C}$ contadas con su multiplicidad, es decir,

$$p(x) = c(x-\lambda_1)^{\alpha_1}(x-\lambda_2)^{\alpha_2} \dots (x-\lambda_r)^{\alpha_r},$$

donde $c \in \mathbb{R}, \lambda_1, \lambda_2, \dots, \lambda_r \in \mathbb{C}, \alpha_1, \alpha_2, \dots, \alpha_r \in \mathbb{N}$ y $\alpha_1 + \alpha_2 + \cdots + \alpha_r = n$.

Definición 6.4 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y sea $\lambda \in Sp(A)$. Se llama* **multiplicidad algebraica** *de λ a la multiplicidad de λ como raíz de $q_A(x)$, es decir al número natural α tal que $q_A(x) = (x-\lambda)^\alpha p(x), p(\lambda) \neq 0$. Se denota $m.a.(\lambda)$.*

Por tanto, una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tiene exactamente n autovalores (contados con su multiplicidad), aunque algunos de ellos pueden no ser reales.

6.1.2 Cálculo de autovectores. Subespacios propios

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y sea $\lambda \in \mathrm{Sp}(A)$. Si $\lambda \in \mathbb{R}$ entonces los autovectores asociados son vectores de $\mathbb{R}^n$.

Definición 6.5 *Se llama* **subespacio propio** *asociado a λ al conjunto*

$$V(\lambda) = \{v \in \mathbb{R}^n / Av = \lambda\, v\} = \{v \in \mathbb{R}^n / (A - \lambda\, I)v = 0\} = Ker(A - \lambda\, I).$$

Se llama **multiplicidad geométrica** *de λ a la dimensión del subespacio propio $V(\lambda)$, es decir,*

$$m.g.(\lambda) = dim(V(\lambda)) = dim(Ker(A - \lambda\, I)).$$

Observación 6.1

1. Recordemos que si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ entonces $\dim(\mathrm{Ker}(A)) = n - \mathrm{rg}(A)$. Por tanto,

$$\mathrm{m.g.}(\lambda) = \dim(\mathrm{Ker}(A - \lambda\, I)) = n - \mathrm{rg}(A - \lambda\, I).$$

2. Los subespacios propios asociados a autovalores diferentes son disjuntos.

$$\lambda_1 \neq \lambda_2 \Longrightarrow V(\lambda_1) \cap V(\lambda_2) = \{0\}.$$

3. Los autovectores asociados a autovalores diferentes son linealmente independientes.

Si $\lambda \in \mathrm{Sp}(A)$, tanto la multiplicidad algebraica como la multiplicidad geométrica de λ son al menos 1. De hecho se tiene el siguiente resultado:

Proposición 6.1 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y sea $\lambda \in Sp(A)$. Entonces*

$$1 \leq m.g.(\lambda) \leq m.a.(\lambda) \leq n.$$

Corolario 6.1 *Si $\lambda \in Sp(A)$ y $m.a.(\lambda) = 1$ entonces $m.g.(\lambda) = m.a.(\lambda) = 1$.*

Ejemplo 6.3 Se considera la matriz

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

Calculamos el polinomio característico de A:

$$|A - x\,I| = \begin{vmatrix} -x & 1 & 1 \\ -1 & 1-x & 0 \\ 1 & 0 & 1-x \end{vmatrix} = -x(1-x)^2 = 0$$

Por tanto, $\text{Sp}(A) = \{0, 1\}$, con $\text{m.a.}(0) = 1, \text{m.a.}(1) = 2$.
Como $\text{m.a.}(0) = 1$, se tiene que $\text{m.g.}(0) = \text{m.a.}(0) = 1$.
A continuación calculamos la multiplicidad geométrica del autovalor $\lambda = 1$:

$$\text{m.g.}(1) = 3 - \text{rg}(A - I) = 3 - rg \begin{pmatrix} -1 & 1 & 1 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} = 3 - 2 = 1.$$

Los subespacios propios asociados a 0 y 1 son:

$$V(0) = \text{Ker}(A) = \{(x, y, z) \in \mathbb{R}^3 / y = x, z = -x\} =< \{(1, 1, -1)\} > .$$

$$V(1) = \text{Ker}(A - I) = \{(x, y, z) \in \mathbb{R}^3 / x = 0, z = -y\} =< \{(0, 1, -1)\} > .$$

Proposición 6.2 *Si $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ y $Sp(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ (cada autovalor aparece tantas veces como indica su multiplicidad algebraica), entonces:*

- *$det(A) = \displaystyle\prod_{i=1}^{n} \lambda_i = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$.*
- *$tr(A) = \displaystyle\sum_{i=1}^{n} \lambda_i = \lambda_1 + \lambda_2 + \cdots + \lambda_n$.*

Esta última propiedad es útil para comprobar si los autovalores se han calculado correctamente, ya que su suma debe coincidir con la traza de la matriz.

6.2 Matrices diagonalizables

Definición 6.6 *Una matriz $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es* **diagonalizable** *si existen dos matrices $P, D \in \mathcal{M}_{n\times n}(\mathbb{R})$ tales que P es inversible, D es diagonal y $A = PDP^{-1}$.*

Denotemos por

$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \quad P = (u_1|u_2|\ldots|u_n).$$

Obsérvese que

$$A = PDP^{-1} \Longleftrightarrow AP = PD \Longleftrightarrow (Au_1|Au_2|\ldots|Au_n) = (\lambda_1 u_1|\lambda_2 u_2|\ldots|\lambda_n u_n).$$

Esto quiere decir que si A es diagonalizable entonces los elementos diagonales de la matriz D son los autovalores de A (contados con su multiplicidad) y las columnas de la matriz P son los correspondientes autovectores asociados (en el mismo orden). Para poder construir D y P es necesario que todos los autovalores de A sean reales y que cada autovalor proporcione tantos autovectores linealmente independientes como indica su multiplicidad algebraica. En resumen, se tiene el siguiente resultado:

Teorema 6.2 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Entonces:*

1. *A es diagonalizable si y sólo si todos los autovalores de A son reales y además*
$$m.a.(\lambda) = m.g.(\lambda), \forall \lambda \in Sp(A).$$
2. *Si A es diagonalizable, las matrices P y D tales que $A = PDP^{-1}$ se construyen del siguiente modo:*
$$D = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix}, \quad P = (u_1|u_2|\ldots|u_n),$$
donde $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los autovalores de A (contados con su multiplicidad) y $u_1, u_2, \ldots, u_n$ son los correspondientes autovectores asociados.

La diagonalización se puede aplicar al cálculo de potencias de matrices.

Proposición 6.3 *Si $A = PDP^{-1}$ entonces $A^k = PD^kP^{-1}, \forall k \geq 1$.*

Ejemplo 6.4 Hallar la expresión de A^k para la matriz

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

En este caso $\mathrm{Sp}(A) = \{0, 3\}$, con $\mathrm{m.a.}(0) = \mathrm{m.g.}(0) = 2$. Además,

$$\mathrm{Ker}(A) =< \{(1,0,-1),(0,1,-1)\} >, \;\; \mathrm{Ker}(A-3I) =< \{(1,1,1)\} > .$$

Por tanto, podemos tomar

$$D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \;\; P = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix}$$

de tal forma que $A = PDP^{-1}$. Por tanto,

$$\begin{aligned} A^k &= PD^kP^{-1} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3^k \end{pmatrix} \begin{pmatrix} 2/3 & -1/3 & -1/3 \\ -1/3 & 2/3 & -1/3 \\ 1/3 & 1/3 & 1/3 \end{pmatrix} \\ &= 3^{k-1} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \end{aligned}$$

6.3 Diagonalización ortogonal

Recordemos que una matriz $P \in \mathcal{M}_{n\times n}(\mathbb{R})$ es ortogonal si $P^{-1} = P^t$, es decir $P^tP = I$.

Definición 6.7 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$. Se dice que A es* **ortogonalmente diagonalizable** *si existen dos matrices $P, D \in \mathcal{M}_{n\times n}(\mathbb{R})$ tales que P es ortogonal, D es diagonal y $A = PDP^t$. En tal caso, se dice que la descomposición $A = PDP^t$ es una* **diagonalización ortogonal** *de A.*

Teorema 6.3 *(***Teorema espectral para matrices simétricas***)*
Una matriz real $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ es ortogonalmente diagonalizable si y sólo si A es simétrica.

Diagonalización ortogonal de una matriz simétrica

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. Veamos cómo construir las matrices P y D tales que $A = PDP^t$.

La matriz D se construye en la forma habitual, es decir, es una matriz diagonal cuyos elementos diagonales son los autovalores de A, repetidos un número de veces igual a su multiplicidad algebraica. Una observación importante es que **todos los autovalores de una matriz simétrica son reales**.

Como $A = PDP^t = PDP^{-1}$, las columnas de la matriz P deben ser autovectores de A, pero necesitamos además que P sea ortogonal. En virtud de la propiedad probada en la sección 4.9, las columnas de A deben ser una base ortonormal. La siguiente propiedad hace que sea posible calcular una base ortonormal formada por autovectores:

Lema 6.1 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. Si v_1 y v_2 son autovectores asociados a dos autovalores distintos de A entonces v_1 y v_2 son ortogonales.*

Demostración. Sean $\lambda_1 \neq \lambda_2$ dos autovalores de A y sean $v_1 \in V(\lambda_1), v_2 \in V(\lambda_2)$. Teniendo en cuenta que $A = A^t$ y $\lambda_1, \lambda_2 \in \mathbb{R}$:

$$\lambda_1 v_1^t v_2 = (\lambda_1 v_1)^t v_2 = (Av_1)^t v_2 = v_1^t A^t v_2 = v_1^t A v_2 = v_1^t \lambda_2 v_2 = \lambda_2 v_1^t v_2.$$

Por tanto, $\lambda_1 v_1^t v_2 = \lambda_2 v_1^t v_2$. Como $\lambda_1 \neq \lambda_2$, necesariamente $v_1^t v_2 = 0$. □

Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. Teniendo en cuenta las propiedades anteriores, los pasos para calcular una diagonalización ortogonal $A = PDP^t$ son los siguientes:

1. Se calculan los autovalores de A. Los elementos diagonales de la matriz D son los autovalores de A (repetidos tantas veces como indica su multiplicidad algebraica).

2. Para cada autovalor $\lambda \in \mathrm{Sp}(A)$ se halla una base del subespacio propio asociado $V(\lambda)$ y se le aplica el proceso de ortonormalización de Gram-Schmidt para obtener una base ortonormal de $V(\lambda)$.

3. La matriz P es la que tiene por columnas los elementos de las bases ortonormales de $V(\lambda_1), V(\lambda_2), \ldots, V(\lambda_k)$ (donde $\lambda_1, \lambda_2, \cdots, \lambda_k$ son los autovalores distintos de A) colocadas en el mismo orden que ocupan los correspondientes autovalores en la diagonal de D.

Ejemplo 6.5 Hallar una diagonalización ortogonal de la matriz $A \in \mathcal{M}_{4\times 4}(\mathbb{R})$ dada por

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

Dado que A es una matriz simétrica real, es ortogonalmente diagonalizable, es decir, existen dos matrices $P, D \in \mathcal{M}_{4\times 4}(\mathbb{R})$ tales que P es ortogonal, D es diagonal y $A = PDP^t$. La matriz diagonal D tiene como elementos diagonales los autovalores de A.

El polinomio característico de A es $q_A(x) = (-3-x)^3(3-x)$.

Por tanto los autovalores de A son $\lambda_1 = -3$ y $\lambda_2 = 3$, con $\text{m.a.}(-3) = 3, \text{m.a.}(3) = 1$, y la matriz D es

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

Los vectores columna de la matriz ortogonal $P = (u_1|u_2|u_3|u_4)$ constituyen una base ortonormal de $\mathbb{R}^4$ formada por autovectores de A. Para determinarlos, aplicaremos el procedimiento de ortonormalización de Gram-Schmidt a sendas bases de los subespacios propios asociados a los autovalores $\lambda_1 = -3$ y $\lambda_2 = 3$.

Resolviendo el correspondiente sistema homogéneo, se tiene:

$$V(-3) = \text{Ker}(A + 3I) =< \{(1,1,0,0), (0,2,0,1), (0,0,1,0)\} > .$$

Si denotamos $v_1 = (1,1,0,0), v_2 = (0,2,0,1), v_3 = (0,0,1,0)$ entonces los tres primeros vectores columna u_1, u_2, u_3 de la matriz P se calculan del siguiente modo:

$$u_1 = \frac{v_1}{||v_1||} = \left(1/\sqrt{2}, 1/\sqrt{2}, 0, 0\right)$$

$$\tilde{u}_2 = v_2 - < v_2, u_1 > u_1 = (-1,1,0,1), \quad u_2 = \frac{\tilde{u}_2}{||\tilde{u}_2||} = \left(-1/\sqrt{3}, 1/\sqrt{3}, 0, 1/\sqrt{3}\right)$$

$$\tilde{u}_3 = v_3 - < v_3, u_1 > u_1 - < v_3, u_2 > u_2 = (0,0,1,0), \quad u_3 = \frac{\tilde{u}_3}{||\tilde{u}_3||} = (0,0,1,0)$$

Del mismo modo,

$$V(3) = \text{Ker}(A - 3I) =< \{(1,-1,0,2)\} >=< v_4 >,$$

de modo que el vector columna u_4 de P viene dado por

$$u_4 = \frac{v_4}{||v_4||} = \left(1/\sqrt{6}, -1/\sqrt{6}, 0, 2/\sqrt{6}\right)$$

Así, la matriz ortogonal

$$P = (u_1|u_2|u_3|u_4) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{3}} & 0 & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & 0 & \frac{-1}{\sqrt{6}} \\ 0 & 0 & 1 & 0 \\ 0 & \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix}$$

cumple que $A = PDP^t$.

6.4 Formas bilineales

Definición 6.8 *Sea V un espacio vectorial real. Una* **forma bilineal** *b sobre V es una aplicación $b : V \times V \longrightarrow \mathbb{R}$ tal que:*

1. $b(x + x', y) = b(x, y) + b(x', y), \forall x, x', y \in V$
2. $b(x, y + y') = b(x, y) + b(x, y'), \forall x, y, y' \in V$
3. $b(\lambda x, y) = b(x, \lambda y) = \lambda \cdot b(x, y), \forall x, y \in V, \forall \lambda \in \mathbb{R}$

Se dice que b es **simétrica** *si*

$$b(x, y) = b(y, x),\ \forall x, y \in V$$

Se dice que b es **definida positiva** *si*

$$b(x, x) \geq 0,\ \forall x \in V \ \textit{con}\ b(x, x) = 0 \Longleftrightarrow x = 0$$

Definición 6.9 *Sea $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ una base de V y $b : V \times V \longrightarrow \mathbb{R}$ una forma bilineal. Entonces*

$$x = \sum_{i=1}^{n} x_i u_i, \qquad y = \sum_{j=1}^{n} y_j u_j$$

$$b(x, y) = b\left(\sum_{i=1}^{n} x_i u_i, \sum_{j=1}^{n} y_j u_j\right) = \sum_{i,j=1}^{n} x_i y_j b(u_i, u_j)$$

y en forma matricial

$$b(x, y) = (x_1, x_2, \ldots . x_n) \begin{pmatrix} b(u_1, u_1) & b(u_1, u_2) & \cdots & b(u_1, u_n) \\ b(u_2, u_1) & b(u_2, u_2) & \cdots & b(u_2, u_n) \\ \vdots & \vdots & \ddots & \vdots \\ b(u_n, u_1) & b(u_n, u_2) & \cdots & b(u_n, u_n) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

$$b(x, y) = x^t A y$$

siendo $A = (b(u_i, u_j))_{1 \leq i,j \leq n}$. Esta matriz A se denomina **matriz asociada a la forma bilineal** *b respecto de la base $\mathcal{B}$.*

6.5 Formas cuadráticas

Definición 6.10 *Una* **forma cuadrática** *sobre* $\mathbb{R}^n$ *es una aplicación bilineal simétrica,* $\omega : \mathbb{R}^n \to \mathbb{R}$ *definida por* $\omega(x) = x^t Ax, \forall x \in \mathbb{R}^n$, *donde* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *es una matriz simétrica.*

Si $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$ entonces la forma cuadrática $\omega(x) = x^t Ax$ se expresa como:

$$\omega(x_1, x_2, \ldots, x_n) = (x_1, x_2, \ldots, x_n) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^{n} a_{ij} x_i x_j$$

Recíprocamente, si tenemos una expresión extendida de la forma cuadrática como la anterior, podemos encontrar una única matriz simétrica $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ tal que $\omega(x) = x^t Ax, \forall x \in \mathbb{R}^n$. Esta matriz se llama **matriz asociada** a la forma cuadrática.

Ejemplo 6.6 Sea $\omega(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + x_3^2 - 4x_1x_2 + 2x_1x_3 - 2x_2x_3$. Entonces:

$$\omega(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x^t Ax$$

6.5.1 Formas cuadráticas degeneradas y no degeneradas

Definición 6.11 *Sea* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *una matriz simétrica y sea* $\omega : \mathbb{R}^n \to \mathbb{R}$ *la forma cuadrática definida por* $\omega(x) = x^t Ax, \forall x \in \mathbb{R}^n$. *Se dice que* ω *es* **no degenerada** *si* $rg(A) = n$, *es decir, si* $|A| \neq 0$. *Si el determinante de* A *es cero entonces se dice que la forma cuadrática* ω *es* **degenerada**.

Por ejemplo, la forma cuadrática $\omega(x_1, x_2, x_3) = 2x_1^2 + 3x_2^2 + x_3^2 - 4x_1x_2 + 2x_1x_3 - 2x_2x_3$ es no degenerada porque $\omega(x) = x^t Ax$, con

$$|A| = \begin{vmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{vmatrix} = 1 \neq 0$$

Clasificación de formas cuadráticas no degeneradas

Las formas cuadráticas no degeneradas $\omega : \mathbb{R}^n \to \mathbb{R}$ pueden ser de tres tipos:

1. ω es **definida positiva** si $\omega(x) = x^t Ax > 0, \forall x \neq 0$,
2. ω es **definida negativa** si $\omega(x) = x^t Ax < 0, \forall x \neq 0$,
3. ω es **indefinida** si existen dos vectores $x, y \in \mathbb{R}^n$ tales que $\omega(x) > 0, \omega(y) < 0$.

Definición 6.12 *Una matriz simétrica* $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ *se dice* **definida positiva, definida negativa o indefinida** *según lo sea la forma cuadrática* $\omega_A : \mathbb{R}^n \to \mathbb{R}$ *definida por* $\omega_A(x) = x^t Ax$.

Ejemplo 6.7

1. $\omega(x,y,z) = x^2+y^2+z^2$ es definida positiva ya que $x^2+y^2+z^2 \geq 0, \forall (x,y,z) \in \mathbb{R}^3$ y además $x^2+y^2+z^2 = 0 \Longleftrightarrow x=y=z=0$.

2. $\omega(x,y,z) = x^2+y^2-z^2$ es indefinida ya que, por ejemplo, $\omega(1,0,0) = 1 > 0$ y $\omega(0,0,1) = -1 < 0$. Además es no degenerada ya que

$$\omega(x,y,z) = (x,y,z)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = x^tAx$$

con $|A| = -1 \neq 0$.

Sin embargo, en general es difícil determinar la clasificación de ω si aparecen "términos cruzados". Por ejemplo, la forma cuadrática $\omega(x_1,x_2,x_3) = 2x_1^2+3x_2^2+x_3^2-4x_1x_2+2x_1x_3-2x_2x_3$ es definida positiva, pero no es inmediato deducirlo a simple vista.

Clasificación de formas cuadráticas degeneradas

Las formas cuadráticas degeneradas $\omega : \mathbb{R}^n \to \mathbb{R}$ pueden ser de tres tipos:

1. ω es **semidefinida positiva** si $\omega(x) = x^tAx \geq 0, \forall x \in \mathbb{R}^n$,

2. ω es **semidefinida negativa** si $\omega(x) = x^tAx \leq 0, \forall x \in \mathbb{R}^n$,

3. ω es **indefinida** si existen dos vectores $x, y \in \mathbb{R}^n$ tales que $\omega(x) > 0, \omega(y) < 0$.

6.5.2 Clasificación de formas cuadráticas usando menores principales

Las formas cuadráticas no degeneradas se pueden clasificar analizando el signo de los menores principales de la matriz.

Definición 6.13 *Sea $A = (a_{ij}) \in \mathcal{M}_{n\times n}(\mathbb{R})$. Para cada $k = 1,2,\ldots,n$, se llama* **menor principal** *de orden k de A y se denota Δ_k al siguiente determinante:*

$$\Delta_k = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{vmatrix}$$

Teorema 6.4 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. Entonces A es definida positiva si y sólo si todos los menores principales de A son mayores que cero.*

Ejemplo 6.8 Consideremos la forma cuadrática $\omega : \mathbb{R}^3 \to \mathbb{R}$ definida por $\omega(x) = x^tAx$, donde

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

Los menores principales de A son:

$$\Delta_1 = 2 > 0,\ \Delta_2 = \begin{vmatrix} 2 & -2 \\ -2 & 3 \end{vmatrix} = 2 > 0,\ \Delta_3 = \begin{vmatrix} 2 & -2 & 1 \\ -2 & 3 & -1 \\ 1 & -1 & 1 \end{vmatrix} = 1 > 0$$

Como todos son positivos, A es definida positiva.

El resultado anterior se puede aplicar también a matrices definidas negativas, teniendo en cuenta que A es **definida negativa** si y sólo si $B = -A$ es definida positiva y que si $A_k \in \mathcal{M}_{k\times k}(\mathbb{R})$ entonces $|-A_k| = (-1)^k |A_k|$. De este modo se obtiene el siguiente resultado:

Proposición 6.4 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. A es definida negativa si y sólo si los menores principales de orden impar son menores que cero y los de orden par son mayores que cero, es decir, $\Delta_1 < 0, \Delta_2 > 0, \Delta_3 < 0, \ldots$*

El uso de los menores principales se puede resumir en el siguiente resultado:

Teorema 6.5 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica tal que $|A| \neq 0$. Entonces la forma cuadrática $\omega(x) = x^t A x$ se clasifica en función de los menores principales del siguiente modo:*

1. *Si todos los menores principales de A son positivos entonces ω es definida positiva.*

2. *Si los menores principales de orden impar son negativos y los de orden par son positivos entonces ω es definida negativa.*

3. *En cualquier otro caso, ω es indefinida.*

La clasificación de las formas cuadráticas degeneradas no se puede deducir directamente de los menores principales.

6.5.3 Clasificación de formas cuadráticas usando diagonalización ortogonal

Sea $\omega : \mathbb{R}^n \longrightarrow \mathbb{R}$ una forma cuadrática definida por $\omega(x) = x^t A x$, donde A es una matriz simétrica. La clasificación de la forma cuadrática se puede hacer utilizando la diagonalización ortogonal de A.

Como A es ortogonalmente diagonalizable, existen dos matrices $P, D \in \mathcal{M}_{n\times n}(\mathbb{R})$ tales que D es diagonal, P es ortogonal y $A = PDP^t$.

Sea $x \in \mathbb{R}^n$. Entonces:

$$\omega(x) = x^t A x = x^t P D P^t x = (P^t x)^t D (P^t x).$$

Si denotamos $y = P^t x$ entonces la forma cuadrática se escribe en la nueva variable como

$$\omega(y) = y^t D y = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2 = \sum_{i=1}^{n} \lambda_i y_i^2,$$

donde $y = (y_1, y_2, \ldots, y_n)$ y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los autovalores de A contados con su multiplicidad.

De aquí se deduce el siguiente resultado:

Teorema 6.6 *Sea $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ una matriz simétrica. Entonces:*

1. *A es definida positiva si y sólo si $\lambda > 0, \forall\lambda \in Sp(A)$.*
2. *A es definida negativa si y sólo si $\lambda < 0, \forall\lambda \in Sp(A)$.*
3. *A es semidefinida positiva si y sólo si $\lambda \geq 0, \forall\lambda \in Sp(A)$.*
4. *A es semidefinida negativa si y sólo si $\lambda \leq 0, \forall\lambda \in Sp(A)$.*
5. *En cualquier otro caso, A es indefinida.*

Ejemplo 6.9 *La matriz*

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

es semidefinida negativa ya que $Sp(A) = \{0, -3\}$, con $m.a.(0) = 1, m.a.(-3) = 2$.

6.6 Descomposición en valores singulares

Sea $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Entonces $A^tA \in \mathcal{M}_{n\times n}(\mathbb{R})$ es una matriz simétrica. En particular, todos los autovalores de A^tA son reales. Además son no negativos:

Proposición 6.5 *Todos los autovalores de A^tA son mayores o iguales que cero.*

Demostración. Sea $\lambda \in \mathrm{Sp}(A^tA)$ y v un autovector asociado. Entonces $A^tAv = \lambda v$ y por tanto:

$$||Av||^2 = (Av)^t(Av) = v^tA^tAv = \lambda v^tv = \lambda||v||^2 \Longrightarrow \lambda = \frac{||Av||^2}{||v||^2} \geq 0.$$

Definición 6.14 *Sean $A \in \mathcal{M}_{p\times n}(\mathbb{R})$. Se llaman* **valores singulares** *de A a las raíces cuadradas positivas de los autovalores de A^tA, es decir, si $Sp(A^tA) = \{\lambda_1, \ldots, \lambda_n\}$ entonces los valores singulares de A son*

$$\sigma_1 = \sqrt{\lambda_1}, \ldots, \sigma_n = \sqrt{\lambda_n}.$$

Se ordenan de tal forma que

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0.$$

Ejemplo 6.10 Calcular los valores singulares de

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix} \in \mathcal{M}_{4\times 3}(\mathbb{R}).$$

$$A^tA = \begin{pmatrix} 0 & 1 & -1 & 1 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 2 \\ 1 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & -1 & 0 \\ -1 & 3 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

Los autovalores de $A^t A$ son 2, 4 y 9, de modo que los valores singulares de A son

$$\sigma_1 = \sqrt{9} = 3$$
$$\sigma_2 = \sqrt{4} = 2$$
$$\sigma_3 = \sqrt{2}$$

Una de las principales aplicaciones de los valores singulares es que permiten obtener una descomposición de A como suma de r matrices de rango 1, donde $r = \mathrm{rg}(A)$.

Observación 6.2 El rango de A coincide con el número de valores singulares no nulos de A (contados con su multiplicidad).

Bibliografía

[1] D. C. Lay, *Álgebra Lineal y sus Aplicaciones*, Pearson Educación, 2012.

[2] G. Nakos y D. Joyner, *Álgebra Lineal con aplicaciones*, Thomson, 1999.

[3] D. Poole, *Álgebra Lineal: una introducción moderna*, Thomson, 2007.