Apuntes de Análisis Matemático I

Licenciatura en Ciencias y Técnicas Estadísticas

María Dolores Acosta Vigil

Departamento de Análisis Matemático

Universidad de Granada

Tema 1. Espacios vectoriales.

Análisis Matemático I
1^o Licenciatura de Estadística

Espacios vectoriales

Definición de espacio vectorial

Sea V un conjunto (no vacío). Diremos que V es un **espacio vectorial** si hay definidas dos operaciones (una suma y un producto por escalares) que verifican las siguientes propiedades:

Propiedades de la suma

- 1. Asociativa: $(x+y)+z = x+(y+z), \ \ \forall x, y, z \in V$
- 2. Conmutativa: $x+y = y+x, \ \ \forall x, y \in V$
- 3. Elemento neutro: Existe un elemento $0 \in V$ que verifica

$$x + 0 = x, \quad \forall x \in V.$$

- 4. Elemento opuesto:

$$\forall x \in V, \exists -x \in V : x + (-x) = 0$$

conmutativo.

Espacios vectoriales

Además, existe un producto que asocia a cada par de elementos de $\mathbb{K} \times V$ un elemento de V ($\mathbb{K}$ es $\mathbb{R}$ ó $\mathbb{C}$) y se verifica además:

Propiedades del producto (y suma)

- 1. $t(x+y) = tx + ty, \quad \forall t \in \mathbb{K}, x, y \in V$
- 2. $(s+t)x = sx + tx, \quad \forall s, t \in \mathbb{K}, x \in V$
- 3. $(st)x = s(tx), \quad \forall s, t \in \mathbb{K}, x \in V.$
- 4. $1x = x, \quad \forall x \in V$

Al conjunto V dotado de una suma interna y de un producto por elementos de $\mathbb{K}$ se llama un espacio vectorial si verifica las (ocho) propiedades anteriores.

En ese caso, a los elementos de V se le suelen llamar vectores y a los de $\mathbb{K}$ escalares.

Ejemplos de espacios vectoriales

Ejemplos

- $\mathbb{R}^n$ es un espacio vectorial real con las operaciones dadas por

$$(x_1, x_2, \cdots, x_n) + (y_1, y_2, \cdots, y_n) := (x_1 + y_1, x_2 + y_2, \cdots, x_n + y_n),$$

$$\forall (x_1, x_2, \cdots, x_n), (y_1, y_2, \cdots, y_n) \in \mathbb{R}^n$$

$$t(x_1, x_2, \cdots, x_n) := (tx_1, tx_2, \cdots, tx_n), \quad \forall t \in \mathbb{R}, (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n$$

- El espacio $\mathcal{M}_{m\times n}$ de las matrices con coeficientes reales de orden $m \times n$ es un espacio vectorial con las operaciones dadas por

$$A + B := (a_{ij} + b_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \qquad (A, B \in \mathcal{M}_{m\times n})$$

$$tA := (ta_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \qquad (t \in \mathbb{R}, A \in \mathcal{M}_{m\times n})$$

Ejemplos de espacios vectoriales

Ejemplos

- El espacio $\mathcal{P}(\mathbb{R})$ de los polinomios en una variable con coeficientes reales es un espacio vectorial con la suma y el producto por números reales definidos de la manera usual (término a término).
- El conjunto de las funciones con valores reales definidas en un intervalo real I es un espacio vectorial definiendo la suma y el producto por escalares de manera puntual, esto es,

$$(f+g)(x) := f(x) + g(x) \quad (x \in I)$$

$$(tf)(x) = tf(x) \quad (t \in \mathbb{R}, x \in I)$$

- Si I es un intervalo real, el conjunto $\mathcal{C}(I)$ de las funciones reales definidas en I que son continuas es un espacio vectorial con las operaciones definidas igual que en el apartado anterior.

Ejemplos de espacios vectoriales

Ejemplos

- El espacio $\mathcal{L}_1[0,1]$ de las funciones reales definidas en $[0,1]$ que son integrables (en el sentido de Lebesgue) con la suma y el producto por reales definido de manera puntual.
- El espacio de las sucesiones de números reales dotado de la suma y el producto por reales dado por:

$$\{x_n\} + \{y_n\} := \{x_n + y_n\}$$

$$\alpha\{x_n\} := \{\alpha x_n\}$$

- El espacio ℓ_1 de las sucesiones reales $\{x_n\}$ tales que la serie $\sum x_n$ es absolutamente convergente:

$$\sum_n x_n + \sum_n y_n := \sum_n (x_n + y_n)$$

$$\alpha \sum_n x_n := \sum_n \alpha x_n \,.$$

Dependencia e independencia lineal

Definición

Dado un espacio vectorial V y vectores $v_1, \ldots, v_n$ en V, una combinación lineal de esos vectores es cualquier otro vector que pueda escribirse de la forma $\sum_{i=1}^{n} t_i v_i$ para convenientes escalares $t_1, \ldots, t_n$.
Diremos que los vectores $\{v_1, \ldots, v_n\}$ son linealmente dependientes si el vector 0 se puede escribir como una combinación lineal de los vectores $v_1, \ldots, v_n$ con no todos los escalares nulos. Equivalentemente, si alguno de lo vectores v_i es combinación lineal de los demás vectores.
Los vectores $\{v_1, \ldots, v_n\}$ son linealmente independientes si no son linealmente dependientes, es decir, si se verifica la siguiente condición:

$$t_i \in \mathbb{R}, \forall 1 \leq i \leq n, \; \sum_{i=1}^{n} t_i v_i = 0 \quad \Rightarrow \quad t_i = 0, \forall i \, .$$

Dependencia e independencia lineal

Ejemplos

- Estúdiese si el siguiente conjunto de vectores de $\mathbb{R}^3$ es linealmente independiente o dependiente:

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

Para ello intentamos encontrar una solución del sistema

$$\begin{aligned} 1 &= a+b+c \\ 2 &= b+c \\ 1 &= a+c \end{aligned}$$

De las ecuaciones primera y tercera tenemos $b=0$, de la segunda se obtiene ahora $c=2$, y usando la tercera se obtiene $a=-1$. Así pues, tenemos que

$$(1,2,1)=a(1,0,1)+b(1,1,0)+c(1,1,1)$$

para $(a,b,c)=(-1,0,2)$, por tanto, los vectores anteriores son linealmente dependientes.

Dependencia e independencia lineal

Ejemplos

- En el espacio vectorial $\mathcal{P}_2(\mathbb{R})$ de los polinomios en una variable con coeficientes reales, de grado menor o igual que 2, estúdiese si los polinomios p, q y r son linealmente independientes:

$$p(x) := x^2 + x + 1, \quad q(x) := 2x + 1, \quad r(x) := x^2 + 1$$

Para ello estudiamos las soluciones del sistema

$$ap(x) + bq(x) + cr(x) = 0$$

esto es,

$$0 = a(x^2 + x + 1) + b(2x + 1) + c(x^2 + 1) =$$

$$(a + c)x^2 + (a + 2b)x + (a + b + c)$$

Por tanto, ha de cumplirse

$$a + c = 0, \quad a + 2b = 0, \quad a + b + c = 0$$

Dependencia e independencia lineal

Ejemplos

$$a + c = 0, \quad a + 2b = 0, \quad a + b + c = 0$$

La única solución del sistema anterior es $a = b = c = 0$, luego los polinomios p, q, r son linealmente independientes.

Proposición

Sea V un espacio vectorial y $\{v_1, \ldots, v_{n+r}\}$ vectores de V, entonces se verifica:

- Si $0 \in \{v_1, \ldots, v_n\}$, entonces $\{v_1, \ldots, v_n\}$ son linealmente dependientes.
- $\{v_1\}$ es linealmente independiente $\Leftrightarrow v_1 \neq 0$
- Si $\{v_1, \ldots, v_n\}$ es linealmente dependiente, entonces $\{v_1, \ldots, v_n, v_{n+1}, \ldots, v_{n+r}\}$ es linealmente dependiente.
- Si $\{v_1, \ldots, v_{n+r}\}$ es linealmente independiente, entonces $\{v_1, \ldots, v_n\}$ es linealmente independiente.

Dependencia e independencia lineal

Ejercicio

- Si un conjunto de vectores de un espacio vectorial son linealmente independientes, y se añade otro más, ¿es el conjunto obtenido linealmente independiente?
- ¿Qué ocurre si se suprimen vectores de un conjunto que es linealmente dependiente?

Sistemas de generadores

Definición

Un subconjunto S de un espacio vectorial V es un sistema de generadores de V si todo vector de V se escribe como combinación lineal de elementos de S.

Ejemplo

El conjunto S de $\mathbb{R}^2$ dado por

$$S := \{(1,1),(0,1),(2,1)\}$$

es un conjunto de generadores de $\mathbb{R}^2$, ya que cada elemento $(a,b) \in \mathbb{R}^2$ se puede escribir de la forma

$$(a,b) = x(1,1) + y(0,1) + z(2,1)\ .$$

Basta tomar $z = 0, x = a, y = b - a$. Pero el sistema de generadores anteriores no es mínimo, ya que suprimiendo uno cualquiera de los tres vectores, también se obtienen sistemas de generadores.

Sistemas de generadores

Un par de sencillas observaciones sobre sistemas de generadores:

Lema

Sea V un espacio vectorial y $S \subset V$ un sistema de generadores.

- Si S es linealmente dependiente, entonces S contiene a un subconjunto propio que también es un sistema de generadores.
- Si S es linealmente independiente, entonces ningún subconjunto propio de S es un sistema de generadores.

Sistemas de generadores

Proposición

Sea X un espacio vectorial, $V := \{v_1, \ldots, v_n\}$ un sistema de generadores de X y $U = \{u_1, \ldots, u_m\}$ un subconjunto linealmente independiente de X, entonces $m \leq n$.

Idea de la demostración.

El procedimiento consiste en ir sustituyendo uno a uno los vectores de V por vectores de U, manteniendo la condición de que el conjunto resultante sea un sistema de generadores.

Bases

Ejemplo

El conjunto $S := \{(1,0),(1,1)\}$ es un sistema de generadores de $\mathbb{R}^2$ y además es linealmente independiente.

Definición

Sea V un espacio vectorial y $B := \{v_1, \cdots, v_n\}$ un subconjunto de V. Diremos que B es una base (de Hamel) de V si B es un sistema de generadores de V que es además linealmente independiente.

Ejemplos

- El conjunto $S := \{(1,0),(1,1)\}$ es una base de $\mathbb{R}^2$. Lo mismo ocurre con el conjunto $B = \{(1,0),(0,1)\}$.
- El conjunto $\{1, x, x^2, x^3\}$ es una base de $\mathcal{P}_3(\mathbb{R})$ (espacio de los polinomios con coef. reales de grado menor o igual que 3).

Bases

Ejemplos

- El conjunto $\{x^n : n \in \mathbb{N} \cup \{0\}\}$ es una base de $\mathcal{P}(\mathbb{R})$ (espacio de los polinomios con coef. reales).
  Este espacio vectorial es infinito-dimensional, esto es, tiene una cantidad infinita de vectores linealmente independientes.
- Acto de fe:
  Todo espacio vectorial tiene una base de Hamel. De hecho, todo subconjunto de un espacio vectorial formado por vectores linealmente independientes está contenido en alguna base del espacio vectorial.
  La demostración de la afirmación anterior usa el Lema de Zorn.

Definición

Diremos que un espacio vectorial es finito-dimensional si tiene una base finita. En otro caso, diremos que es infinito-dimensional.

De la colección de ejemplos de espacios vectoriales que tenemos, $\mathbb{R}^n$, $\mathcal{M}_{m\times n}$ y $P_n(\mathbb{R})$ son finito-dimensionales.
Los espacios $\mathcal{C}[0,1]$, $P(\mathbb{R})$, $\mathcal{L}_1[0,1]$, ℓ_1 son infinito-dimensionales.

Bases

Definición

Si V es un espacio vectorial y B un subconjunto (infinito) de V, diremos que B es una base (de Hamel) de V si todo elemento de V es una combinación lineal finita de elementos de B y cada subconjunto finito no vacío de B es linealmente independiente.

Bases

Consecuencias de la def. de base

Si V es un espacio vectorial y $B := \{v_1, \ldots, v_n\}$ una base de V, entonces, para cada elemento x de V han de existir escalares $t_1, \ldots, t_n$ tal que

$$x = \sum_{i=1}^{n} t_i v_i$$

Además los escalares t_i anteriores son únicos, ya que si

$$x = \sum_{i=1}^{n} s_i v_i,$$

tendríamos

$$\sum_{i=1}^{n} t_i v_i = x = \sum_{i=1}^{n} s_i v_i.$$

Bases

Consecuencias de la def. de base

Sabemos que

$$\sum_{i=1}^{n} t_i v_i = x = \sum_{i=1}^{n} s_i v_i,$$

luego

$$\sum_{i=1}^{n} (t_i - s_i) v_i = 0 \ .$$

Como B es una base, los vectores $v_1, \ldots, v_n$ son linealmente independientes, luego han de ser $t_i - s_i = 0$ para todo i, y los escalares $t_1, \ldots, t_n$ son únicos. A los escalares anteriores se llaman coordenadas del vector x respecto de la base $\{v_1, \ldots, v_n\}$.

Bases

Teorema

Todas las bases de un mismo espacio vectorial V que tenga un sistema de generadores finito tienen el mismo número de elementos. A éste número de elementos se llama dimensión del espacio vectorial, que se suele notar por $\dim V$.

Demostración. Sean $U := \{u_1, \ldots, u_n\}$ una base de un espacio vectorial V y $\mathcal{B}$ otra base de V. Como $\mathcal{B}$ está formado por vectores lin. independientes y U es un sistema de generadores de V, en vista de la Proposición enunciada en p. 14, tenemos que el número de elementos de $\mathcal{B}$ es menor o igual que n. Aplicando de nuevo el mismo resultado, obtenemos también la otra desigualdad. $\square$

Cambios de base

Proposición

Sea V un espacio vectorial de dimensión finita y B y B' dos bases diferentes. La ecuación del cambio de base viene dada por la expresión:

$$X = AX',$$

donde X representa las coordenadas de un vector en términos de B, X' las coordenadas del mismo vector en términos de B' y A es la matriz cuyas columnas están formadas por las coordenadas de cada vector de B' en términos de B. La expresión del cambio inverso viene dado por $X' = A^{-1}X$.

Demostración. Supongamos que $B := \{e_1, e_2, \dots, e_n\}, B' := \{e_1', e_2', \dots, e_n'\}$. Escribimos los vectores de la segunda base en términos de la primera, esto es, han de existir escalares $\{a_{ij} : 1 \leq i, j \leq n\}$ tales que

Cambios de base

$$\begin{array}{rcl} e'_1 & = & a_{11}e_1 + a_{21}e_2 + \ldots a_{n1}e_n \\ e'_2 & = & a_{12}e_1 + a_{22}e_2 + \ldots a_{n2}e_n \\ \vdots & & \\ e'_n & = & a_{1n}e_1 + a_{2n}e_2 + \ldots a_{nn}e_n \end{array}$$

Sea X un vector cuyas coordenadas respecto de B son $(x(1), x(2), \ldots, x(n))$ y respecto de la base B' son $(y(1), y(2), \ldots, y(n))$, esto es,

$$X = \sum_{i=1}^{n} y(i)e'_i =$$

$$\begin{array}{c} = y(1)\Big(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n\Big) + \\ y(2)\Big(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n\Big) + \\ \vdots \\ y(n)\Big(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n\Big) = \end{array}$$

$$
\begin{gathered}
= y(1)\Big(a_{11}e_1 + a_{21}e_2 + \ldots + a_{n1}e_n\Big) + \\
y(2)\Big(a_{12}e_1 + a_{22}e_2 + \ldots + a_{n2}e_n\Big) + \\
\vdots \\
y(n)\Big(a_{1n}e_1 + a_{2n}e_2 + \ldots + a_{nn}e_n\Big) = \\
= \Big(y(1)a_{11} + y(2)a_{12} + \ldots + y(n)a_{1n}\Big)e_1 + \\
\Big(y(1)a_{21} + y(2)a_{22} + \ldots + y(n)a_{2n}\Big)e_2 + \\
\vdots \\
\Big(y(1)a_{n1} + y(2)a_{n2} + \ldots + y(n)a_{nn}\Big)e_n \,.
\end{gathered}
$$

Cambios de base

$$
\begin{aligned}
=&\Big(y(1)a_{11}+y(2)a_{12}+\ldots+y(n)a_{1n}\Big)e_1+\\
&\Big(y(1)a_{21}+y(2)a_{22}+\ldots+y(n)a_{2n}\Big)e_2+\\
&\qquad\vdots\\
&\Big(y(1)a_{n1}+y(2)a_{n2}+\ldots+y(n)a_{nn}\Big)e_n\,.
\end{aligned}
$$

Esto es,

$$
(x(1),x(2),\ldots,x(n))=\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n}\\ a_{21} & a_{22} & \ldots & a_{2n}\\ \vdots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn}\end{pmatrix}\begin{pmatrix} y(1)\\ y(2)\\ \vdots\\ y(n)\end{pmatrix}
$$

Abreviadamente, $X=AX'$, donde cada columna de A está formada por las coordenadas de cada uno de los vectores de la base B' en términos de B.

Tenemos $X = AX'$ y supongamos que otra matriz C da el cambio inverso, esto es, $X' = CX$. Por tanto, obtenemos que para cada vector X se tiene

$$X = AX' = ACX,$$

$$X' = CX = CAX' .$$

Deducimos que $AC = CA = I$, luego $C = A^{-1}$.

Ejemplo

Se consideran en $\mathbb{R}^3$ la base canónica $\{e_1, e_2, e_3\}$ y la base formada por los siguientes vectores $u_1 = (1, 1, 0), u_2 = (0, 1, 1), u_3 = (1, 0, 1)$. Calcular la matriz del cambio de base.
Si X representa a un vector de $\mathbb{R}^3$ usando sus coordenadas respecto de la base canónica e Y las coordenadas respecto de la base $\{u_i : i = 1, 2, 3\}$, la matriz A dada por

$$A := \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

da el cambio de coordenadas, esto es, $X = AY$.

Subespacios vectoriales

Definición

Sea V un espacio vectorial (sobre $\mathbb{K}$) y $W \subset V$. Diremos que W es un subespacio vectorial de V si se verifica:

$$u, w \in W, t \in \mathbb{K} \quad \Rightarrow \quad u + tw \in W \, .$$

Si V es un espacio vectorial, W_1, W_2 son subespacios vectoriales de V, entonces los conjuntos

$$W_1 + W_2 := \{w_1 + w_2 : w_i \in W_i, i = 1, 2\}$$

y $W_1 \cap W_2$ son subespacios vectoriales de V.
Si B es un conjunto de V, al mínimo subespacio vectorial de V que contiene a B se llama subespacio vectorial generado por B.

Escribiremos $V = W_1 \oplus W_2$ cuando $V = W_1 + W_2$ y además $W_1 \cap W_2 = \{0\}$. En tal caso, cada elemento de V se expresa de forma única como suma de un elemento de W_1 y otro de W_2. En este caso diremos que V es la suma directa de W_1 y W_2.

Ejemplos

- El subconjunto de $\mathbb{R}^3$ dado por

$$W := \{(x, y, z) \in \mathbb{R}^3 : x = y\}$$

  es un subespacio vectorial de dimensión 2 y los vectores $\{(0, 0, 1), (1, 1, 0)\}$ son una base de W_1.

- Si añadimos cualquier vector de $\mathbb{R}^3$ que sea linealmente independiente de los anteriores (en este caso, cualquier vector que no pertenezca a W), por ejemplo $(1, 0, 0)$, obtenemos que el subespacio Z generado por $(1, 0, 0)$ verifica que

$$\mathbb{R}^3 = W \oplus Z$$

Es fácil comprobar los siguientes hechos (especialmente para espacios de dimensión finita):

Proposición

Sea X un espacio vectorial. Entonces:

- Todo subconjunto lin. independiente de X se puede prolongar hasta una base de X.
- Si U y V son subespacios vectoriales de X finito-dimensionales se verifica que

$$\dim(U+V)=\dim U+\dim V-\dim(U\cap V)$$

Idea de la demostración:

- Todo subconjunto lin. independiente de X se puede prolongar hasta una base de X.

  Sea $I \subset X$ un conjunto linealmente independiente de X. Si B es un sistema de generadores, ya es una base de X. Si no lo es, existe un vector v_1 que no se puede escribir como combinación lineal de elementos de B. Entonces $B \cup \{v_1\}$ es un conjunto de vectores linealmente independientes. Si X tiene dimensión finita, entonces repitiendo el procedimiento un número finito de veces, conseguimos una base de X.

Demostración de la segunda igualdad:

Si U y V son subespacios vectoriales de X finito-dimensionales se verifica que

$$\dim(U+V) = \dim U + \dim V - \dim(U \cap V)$$

Llamamos $m := \dim U \cap V, r := \dim U, s = \dim V$. Partimos de una base $\{z_1, \ldots, z_m\}$ de $U \cap V$. Podemos completar la base anterior a bases de U y de V, que son, respectivamente

$$\{z_1, \ldots, z_m, u_{m+1}, \ldots, u_r\} \quad \text{base de } U$$

$$\{z_1, \ldots, z_m, v_{m+1}, \ldots, v_s\} \quad \text{base de } V$$

Comprobaremos ahora que el conjunto

$$B := \{z_1, \ldots, z_m, u_{m+1}, \ldots, u_r, v_{m+1}, \ldots, v_s\}$$

es una base de $U+V$.

Demostración de la segunda igualdad:

Comprobaremos ahora que el conjunto

$$B := \{z_1, \ldots, z_m, u_{m+1}, \ldots, u_r, v_{m+1}, \ldots, v_s\}$$

es una base de $U + V$, con lo que

$$\dim(U + V) = r + (s - m) = \dim U + \dim V - \dim(U \cap V).$$

Como el conjunto $\{z_1, \ldots, z_m, u_{m+1}, \ldots, u_r\}$ es una base de U y $\{z_1, \ldots, z_m, v_{m+1}, \ldots, v_s\}$ es base de V, entonces B es un sistema de generadores de $U + V$.
Comprobamos que B es linealmente independiente. Consideramos entonces 0 escrito como combinación lineal de elementos de B

$$0 = \sum_{i=1}^{m} a_i z_i + \sum_{i=m+1}^{r} b_i u_i + \sum_{i=m+1}^{s} c_i v_i,$$

Demostración de la segunda igualdad:

esto es, el vector

$$v = \sum_{i=1}^{m} a_i z_i + \sum_{i=m+1}^{r} b_i u_i = - \sum_{i=m+1}^{s} c_i v_i,$$

está en $U \cap V$, luego se escribe como una única combinación lineal de $\{z_1, \ldots, z_m\}$, esto es,

$$v = - \sum_{i=m+1}^{s} c_i v_i = \sum_{i=1}^{m} d_i z_i$$

Como los vectores $\{z_1, \ldots, z_m, v_{m+1}, \ldots, v_s\}$ son una base de V y

$$\sum_{i=1}^{m} d_i z_i + \sum_{i=m+1}^{s} c_i v_i = v - v = 0$$

entonces han de ser $d_1 = d_2 = \ldots = d_m = c_{m+1} = \ldots c_s = 0.$

Demostración de la segunda igualdad:

Sustituyendo en la igualdad

$$0 = \sum_{i=1}^{m} a_i z_i + \sum_{i=m+1}^{r} b_i u_i + \sum_{i=m+1}^{s} c_i v_i,$$

obtenemos

$$0 = \sum_{i=1}^{m} a_i z_i + \sum_{i=m+1}^{r} b_i u_i,$$

y usando ahora que los vectores $\{z_1, \ldots, z_m, u_{m+1}, \ldots, u_r\}$ son una base de U, entonces han de ser todos los coeficientes cero, esto es, $a_i = b_j = 0, \forall 1 \leq i \leq m, m+1 \leq j \leq r$, con lo que hemos probado que B es un conjunto linealmente independiente de vectores de X.

Tema 2. Matrices y aplicaciones lineales. Determinantes.

Análisis Matemático I
1^o Licenciatura de Estadística

Matrices

Matriz de números reales

Una matriz de números reales y orden $m \times n$ es una aplicación $f : I \times J \longrightarrow \mathbb{R}$, donde $I = \{1, 2, \cdots, m\}, J = \{1, 2, \cdots, n\}$. En lugar de usar la notación anterior se suelen escribir de la forma

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

donde $a_{ij} = f(i, j)$ para $1 \leq i \leq m, 1 \leq j \leq n$.
En el caso anterior, diremos que la matriz tiene m filas y n columnas.
Notaremos por $\mathcal{M}_{m \times n}$ al espacio de las matrices reales de orden $m \times n$. Cuando $m = n$ escribiremos simplemente $\mathcal{M}_n$ y a los elementos de ese conjunto los llamaremos matrices cuadradas.

Rango de una matriz

Para una matriz A de orden $m \times n$, se define su rango como el número máximo de vectores columna de A que son linealmente independientes. Notaremos por $\mathrm{rg}\, A$ el rango de la matriz A.

Ejemplo

Calcúlese el rango de la matriz A, siendo

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

Como la última columna es el vector 0, es claro que una lista de vectores que lo contenga será siempre linealmente dependiente. Así que comprobamos si los vectores que aparecen en las tres primeras columnas son linealmente dependientes.

Rango de una matriz

Ejemplo

$A = \begin{pmatrix} 2 & 2 & 2 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 2 & 0 & 0 \end{pmatrix}$ Calculamos el determinante de la matriz cuadrada

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

que vale $2(-2) + 0 + 1(2 + 2) = -4 + 4 = 0$. Por tanto, el rango de A no puede ser 3. Como las columnas primera y segunda son vectores lin. independientes, $\mathrm{rg}\ A = 2$.

Ejemplo

Si se considera una matriz cuadrada de orden $n \times n$ diagonal, es decir, que tenga la forma

$$\begin{pmatrix} d_1 & 0 & 0 & \cdots & 0 \\ 0 & d_2 & 0 & \cdots & 0 \\ 0 & 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & d_n \end{pmatrix}$$

y suponemos que $d_i \neq 0$, para todo $1 \leq i \leq n$, entonces tiene rango n.

Producto de matrices

Definición

Sea $A \in \mathcal{M}_{n\times m}, B \in \mathcal{M}_{p\times n}$, entonces la matriz producto $BA \in \mathcal{M}_{p\times m}$ viene dada por

$$\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_p \end{pmatrix} \begin{pmatrix} c_1 & c_2 & \cdots & c_m \end{pmatrix} = \begin{pmatrix} < f_1, c_1 > & < f_1, c_2 > & \cdots & < f_1, c_m > \\ < f_2, c_1 > & < f_2, c_2 > & \cdots & < f_2, c_m > \\ \vdots & \vdots & & \vdots \\ < f_p, c_1 > & < f_p, c_2 > & \cdots & < f_p, c_m > \end{pmatrix}$$

donde f_i es la fila i-ésima de B para $1 \leq i \leq p$ y c_j es la columna j-ésima de A para $1 \leq j \leq m$ y $< , >$ es el producto escalar en $\mathbb{R}^n$.

Producto de matrices

Propiedades del producto de matrices

- No es conmutativo en general.
- Asociativa: $(AB)C = A(BC), \quad \forall A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}, C \in \mathcal{M}_{p\times r}$.
- Existen matrices $I_m \in \mathcal{M}_{m\times m}, I_n \in \mathcal{M}_{n\times n}$ tales que

$$I_m A = A, \quad A I_n = A, \quad \forall A \in \mathcal{M}_{m\times n} \ .$$

- Distributiva del producto respecto de la suma de matrices:

$$A(B + C) = AB + AC, \quad \forall A \in \mathcal{M}_{m\times n}, B, C \in \mathcal{M}_{n\times p} \ .$$

Propiedades de la inversión de matrices

Sean $A, B \in \mathcal{M}_n$ matrices inversibles.

- La matriz A^{-1} tiene inversa y ésta es A.
- La matriz AB tiene inversa y $(AB)^{-1} = B^{-1}A^{-1}$.
- Si $C \in \mathcal{M}_n$ y $AC = I_n$, entonces $C = A^{-1}$. De forma análoga, si $CA = I_n$, entonces $C = A^{-1}$.

Cálculo de la matriz inversa (Método de Gauss-Jordan)

La idea general consiste en multiplicar una matriz inversible A de orden n por un producto finito de matrices tal que $P_1 P_2 \cdots P_r A = I_n$. En este caso el efecto que produce el producto de cada una de las matrices $P_i A$ es alguno de los siguientes:

- intercambian dos filas de A (la matriz P de orden n que tiene las coordenadas de e_j en la fila i-ésima, y las de e_i en la fila j-ésima y en la fila k-ésima las coordenadas de e_k verifica que al multiplicar PA intercambia las filas i-ésima y j-ésima de la matriz A).
- multiplica una fila por un escalar no nulo (basta con usar como P la matriz que coincide con la identidad en cada fila y que en la fila i-ésima tiene las coordenadas de te_i, al multiplicar por A multiplica por t la fila i-ésima de la matriz A).
- cambiar una fila por la suma de ésta con una combinación lineal otras filas. Por ejemplo si P tiene en la fila i-ésima las coordenadas de $e_i + te_j$ y en el resto de las filas las que tenga la matriz identidad, entonces PA tiene en la fila i-ésima $f_i + tf_j$, donde f_i es la fila i-ésima de A, y el resto de las filas son las de A.

Por supuesto, multiplicando a la derecha la matriz A por las traspuestas de las matrices anteriores, se consiguen los siguientes efectos:

- intercambiar dos columnas de A,
- multiplicar una columna de A por un escalar no nulo,
- sustituir una columna de A por la suma de ésta y de una combinación lineal de otras columnas.

Matrices

Ejemplo

Partimos de la siguiente matriz $A = \begin{pmatrix} 4 & 2 \\ 3 & -1 \end{pmatrix}$, que es inversible.

- Multiplicamos la primera fila por $\frac{1}{4}$, obteniendo $\begin{pmatrix} 1 & \frac{1}{2} \\ 3 & -1 \end{pmatrix}$.
- Ahora sumamos a la segunda fila la primera multiplicada por -3, obteniendo $\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & \frac{-5}{2} \end{pmatrix}$.
- Multiplicamos la segunda por $\frac{-2}{5}$, con lo que obtenemos $\begin{pmatrix} 1 & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$.
- Por último, sumamos a la primera fila la segunda multiplicada por $\frac{-1}{2}$, y queda $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Matrices

Por supuesto, el producto de las matrices por las que hemos multiplicado A por la izquierda es la inversa de A^{-1}. Más adelante, mejoraremos el procedimiento, de manera que al terminar, tengamos la matriz inversa de A.

Ejemplo

Partimos de la siguiente matriz $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}$, que es inversible.

- Cambiamos las filas segunda y tercera por combinaciones lineales de éstas con la primera ($f_2 \mapsto f_2 - 2f_1,\ f_3 \mapsto f_3 - 3f_1$), obteniendo $\begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -5 \\ 0 & -2 & -5 \end{pmatrix}$
- ($f_1 \mapsto f_1 + f_2, f_2 \mapsto -f_2,\ f_3 \mapsto f_3 - 2f_2$), obteniendo $\begin{pmatrix} 1 & 0 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 5 \end{pmatrix}$
- ($f_1 \mapsto f_1 + \frac{3}{5}f_3, f_2 \mapsto f_2 - f_3,\ f_3 \mapsto \frac{1}{5}f_3$), obteniendo $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

Ahora bien, nos interesa que el procedimiento nos permita calcular fácilmente la matriz inversa de la dada.

Nótese que si, por ejemplo, $P_2P_1A = I_n$, donde P_1, P_2 son matrices inversibles tales que P_iB producen uno de los tres tipos de cambios permitidos en B, para una matriz cuadrada B de orden n, entonces sabemos que $P_2P_1 = A^{-1}$, esto es, $P_2P_1I_n = A^{-1}$.

Supongamos, por ejemplo, que P_1B intercambia las filas 1 y 2 de cada matriz B, y que P_2B únicamente cambia la primera fila matriz B por la suma de las filas primera y la segunda.

Entonces P_1I_n intercambia las filas 1 y 2 de la matriz identidad. Después $P_2(P_1I_n)$ cambia las filas 1 de P_1I_n por la suma de la primera y la segunda fila de esta matriz.

En resumen, la inversa de A (que es P_2P_1 en este caso), se obtiene a partir de la matriz identidad, aplicando las mismas transformaciones “elementales” que en cada paso hemos de aplicar para transformar A en la matriz I_n.

Ejemplo

Partimos de la misma matriz A de orden 3 que ya usamos antes. Esto es, $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}$, que es inversible. Ahora aplicaremos las transformaciones que usamos a la matriz

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right).$$

Ejemplo

- $f_2 \mapsto f_2 - 2f_1,\ f_3 \mapsto f_3 - 3f_1$, obteniendo
$$\left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -5 & -2 & 1 & 0 \\ 0 & -2 & -5 & -3 & 0 & 1 \end{array}\right)$$
- $f_1 \mapsto f_1 + f_2, f_2 \mapsto -f_2,\ f_3 \mapsto f_3 - 2f_2$, obteniendo
$$\left(\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & 1 & 0 \\ 0 & 1 & 5 & 2 & -1 & 0 \\ 0 & 0 & 5 & 1 & -2 & 1 \end{array}\right)$$
- Aplicamos ahora los siguientes cambios $f_1 \mapsto f_1 + \frac{3}{5}f_3, f_2 \mapsto f_2 - f_3,\ f_3 \mapsto \frac{1}{5}f_3$
$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{3}{5} \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{array}\right)$$

Matrices

La matriz $\begin{pmatrix} -\frac{2}{5} & -\frac{1}{5} & \frac{3}{5} \\ 1 & 1 & -1 \\ \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{pmatrix}$ es la inversa de A.

Aplicaciones lineales

Aplicación lineal

Sean X e Y espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$. Una aplicación $f : X \longrightarrow Y$ es lineal si verifica

$$f(\alpha u + \beta v) = \alpha f(u) + \beta f(v), \quad \forall \alpha, \beta \in \mathbb{K}, u, v \in X \ .$$

Esto es, f es lineal si conserva las combinaciones lineales y los coeficientes de éstas.

Ejemplos

- La aplicación $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $f(x, y) = (x, x + y)$ para cada vector $(x, y) \in \mathbb{R}^2$ es lineal.
- La aplicación $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $f(x, y) = (x + 2, e^x)$ para cada vector $(x, y) \in \mathbb{R}^2$ no es lineal.

Aplicaciones lineales

Ejemplos

- En el espacio $\mathcal{C}[a,b]$ (funciones reales continuas en el intervalo [a,b]), la aplicación $I : \mathcal{C}[a,b] \longrightarrow \mathbb{R}$ dada por $I(f) = \int_a^b f(t)\,dt$ para cada función $f \in \mathcal{C}[a,b]$ es lineal.
- La aplicación identidad en un espacio vectorial es lineal.

Proposición

Sean X, Y y Z espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$.

- Si $f, g : X \longrightarrow Y$ son aplicaciones lineales, entonces $\alpha f + \beta g$ es una aplicación lineal de X en Y, para cualesquiera escalares α y β.
- Si $f : X \longrightarrow Y$ y $g : Y \longrightarrow Z$ son aplicaciones lineales, entonces la composición $g \circ f$ es una aplicación lineal de X en Z.

La comprobación es totalmente rutinaria.

Inversa

Dados dos conjuntos X e Y y una aplicación $f : X \longrightarrow Y$, diremos que f tiene inversa si existe una aplicación $g : Y \longrightarrow X$ tal que $f \circ g = I_Y$ y $g \circ f = I_X$, donde I_X es la aplicación identidad en X. En tal caso, la aplicación g es única, se llama inversa de f y se suele notar $g = f^{-1}$.

Proposición

Sean X e Y espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$. Si $f : X \longrightarrow Y$ es lineal y tiene inversa, entonces f^{-1} es lineal.

Aplicaciones lineales

Aplicaciones lineales

Si X e Y son espacios vectoriales sobre el cuerpo $\mathbb{K}$, entonces el conjunto $L(X,Y)$ de las aplicaciones lineales de X en Y es un espacio vectorial con las operaciones dadas por:

$$(f+g)(x) := f(x)+g(x), \quad (\alpha f)(x) = \alpha f(x), \quad \forall f,g \in L(X,Y), \alpha \in \mathbb{K}, x \in X\ .$$

En este caso, el elemento neutro de la suma es la aplicación que lleva cada elemento de X en 0.

Propiedades de las aplicaciones lineales

Sean X e Y espacios vectoriales sobre el mismo cuerpo $\mathbb{K}$ y $f : X \longrightarrow Y$ una aplicación lineal. Se verifican las siguientes afirmaciones:

- Si D es un subconjunto de vectores de X linealmente dependiente, entonces $f(D)$ es un subconjunto de Y linealmente dependiente.
- Si $V \subset X$ es un subespacio vectorial, entonces $f(V)$ también lo es.
- Si $W \subset Y$ es un subespacio vectorial, entonces $f^{-1}(W) := \{x \in X : f(x) \in W\}$ también lo es.

Ejemplo

Consideramos la aplicación lineal $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por

$$f(x, y) = (x + 2y, 2x + 4y) \quad ((x, y) \in \mathbb{R}^2) \ .$$

La aplicación anterior lleva la base canónica en los vectores $\{(1, 2), (2, 4)\}$ que son linealmente dependientes, ya que $(2, 4) = 2(1, 2)$, luego una aplicación lineal no tiene porqué conservar la independencia lineal ni tampoco conserva los sistemas de generadores.

Subespacios destacados asociados a una aplicación lineal

Si X e Y son espacios vectoriales y $f : X \longrightarrow Y$ es lineal, entonces se llaman:

- núcleo de f al subespacio de X dado por $f^{-1}(0)$, esto es,

  $$\{x \in X : f(x) = 0\} \,,$$

  subespacio que se nota por $\mathrm{Ker}\, f$.
- La imagen de f es el subespacio vectorial $\mathrm{Im}\,(f)$ dado por

  $$f(X) = \{f(x) : x \in X\} \,.$$

Proposición

Si X e Y son espacios vectoriales y $f : X \longrightarrow Y$ es lineal, entonces se verifican:

- $$\mathrm{Ker}\, f = 0 \quad \Leftrightarrow \quad f \text{ es inyectiva} .$$
- $$\dim \mathrm{Im}\, f \leq \dim X .$$

Ejemplos

- Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ la aplicación dada por

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

En este caso, el núcleo de f viene dado por

$$\mathrm{Ker}\ f = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) = (0,0)\} = \{(x,y,z) \in \mathbb{R}^3 : x+y=0\}$$

que es un subespacio de $\mathbb{R}^3$ de dimensión 2 (generado por e_3 y $e_1 - e_2$, por ejemplo).
Es fácil comprobar que la imagen de f es el conjunto

$$\mathrm{Im}\ f = \{(x,0) : x \in \mathbb{R}\}\ ,$$

que es un subespacio 1-dimensional de $\mathbb{R}^2$.
Nótese que se verifica que

$$\dim \mathbb{R}^3 = \dim \mathrm{Ker}\ f + \dim \mathrm{Im}\ f \quad (3 = 2+1)\ .$$

- Sea $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ la aplicación dada por

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

En este caso, el núcleo de f viene dado por

$$\mathrm{Ker}\ f = \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = (0, 0)\} = \{(0, 0, 0)\} \ .$$

Calculamos ahora la imagen. Si suponemos que $(r, s, t) = f(x, y, z)$, tendremos que

$$(r, s, t) = (x, 2y, x + z) \ .$$

Igualando las coordenadas y resolviendo el sistema obtenemos

$$x = r, y = \frac{s}{2}, z = t - x = t - r \ .$$

Esto es, $f(r, \frac{s}{2}, t - r) = (r, s, t)$ para cualquier elemento $(r, s, t) \in \mathbb{R}^3$, luego $\mathrm{Im}\ f = \mathbb{R}^3$ y también se verifica
$\dim \mathbb{R}^3$(espacio de partida) $= \dim \mathrm{Ker}\ f + \dim \mathrm{Im}\ f \quad (3 = 0 + 3)$.

Teorema (de la dimensión)

Sea $f : X \longrightarrow Y$ una aplicación lineal entre dos espacios vectoriales sobre el cuerpo $\mathbb{K}$ tal que X es finito-dimensional, entonces se verifica que

$$\dim X = \dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f\ .$$

Esquema de la demostración.
Supongamos que $n = \dim X$, $r = \dim \mathrm{Ker}\, f$, $s = \dim \mathrm{Im}\, f$.
Sea $\{y_1, \ldots, y_s\}$ una base de $\mathrm{Im}\, f$ y $\{x_1, \cdots, x_s\}$ vectores de X tales que $f(x_j) = y_j$ para $1 \leq j \leq s$.
Consideramos ahora una base $\{u_1, \cdots, u_r\}$ del subespacio vectorial $\mathrm{Ker}\, f$.
No es difícil probar que

$$\{u_1, \cdots, u_r, x_1, \ldots, x_s\}$$

es una base de X, con lo que

$$\dim X = n = r + s = \dim \mathrm{Ker}\, f + \dim \mathrm{Im}\, f\ .$$

Proposición

Sean dos espacios vectoriales X e Y sobre $\mathbb{K}$ de dimensiones n y m, respectivamente, y B_X y B_Y bases de los espacios X e Y. Para cada aplicación lineal $T : X \longrightarrow Y$, existe una única matriz A de orden $m \times n$ y coeficientes en $\mathbb{K}$ tales que

$$T(x) = Ax^t, \forall x \in X \ .$$

De hecho, si $B_X = \{x_1, x_2, \cdots, x_n\}, B_Y = \{y_1, \cdots, y_m\}$, la columna i-ésima de A está formada por las coordenadas de $T(x_i)$ respecto de la base B_Y.

Demostración.
Sea $u \in X$; escribimos u en términos de la base B_X. Por tanto, existen escalares $u(1), \cdots, u(n)$ tales que $u = \sum_{i=1}^{n} u(i) x_i$. Como T es lineal, tendremos que

$$T(u) = \sum_{i=1}^{n} u(i) T(x_i) \ .$$

Matrices y aplicaciones lineales

Por ser B_Y una base de Y, podremos escribir cada elemento $T(x_i)$ como combinación lineal de los elementos de B_Y. Por tanto, existen escalares (a_{ij}) tales que

$$\begin{array}{rcl} T(x_1) & = & a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m = \sum_{j=1}^{m} a_{j1}y_j \\ T(x_2) & = & a_{12}y_1 + a_{22}y_2 + \cdots + a_{m2}y_m = \sum_{j=1}^{m} a_{j2}y_j \\ \vdots & & \\ T(x_n) & = & a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{mn}y_m = \sum_{j=1}^{m} a_{jn}y_j \end{array}$$

Sustituyendo en la expresión $T(u) = \sum_{i=1}^{n} u(i)T(x_i)$ los vectores $T(x_i)$ por las expresiones anteriores obtenemos

Matrices y aplicaciones lineales

$$
\begin{aligned}
T(u) = \quad & u(1)\Big(a_{11}y_1 + a_{21}y_2 + \cdots + a_{m1}y_m\Big) + \\
& u(2)\Big(a_{12}y_1 + a_{22}y_2 + \cdots + a_{m2}y_m\Big) + \\
& \qquad\qquad \vdots \\
& u(n)\Big(a_{1n}y_1 + a_{2n}y_2 + \cdots + a_{mn}y_m\Big) =
\end{aligned}
$$

$$
\begin{aligned}
& \Big(u(1)a_{11} + u(2)a_{12} + \cdots + u(n)a_{1n}\Big)y_1 + \\
& \Big(u(1)a_{21} + u(2)a_{22} + \cdots + u(n)a_{2n}\Big)y_2 + \\
& \qquad\qquad \vdots \\
& \Big(u(1)a_{m1} + u(2)a_{m2} + \cdots + u(n)a_{mn}\Big)y_m
\end{aligned}
$$

$$\begin{aligned}&\Big(u(1)a_{11} + u(2)a_{12} + \cdots + u(n)a_{1n}\Big)y_1 +\\ &\Big(u(1)a_{21} + u(2)a_{22} + \cdots + u(n)a_{2n}\Big)y_2 +\\ &\qquad\qquad\vdots\\ &\Big(u(1)a_{m1} + u(2)a_{m2} + \cdots + u(n)a_{mn}\Big)y_m\end{aligned}$$

Por tanto, las coordenadas de $T(u)$ respecto de la base B_Y vienen dadas por el producto de las matrices

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} u(1) \\ u(2) \\ \vdots \\ u(n) \end{pmatrix} = Au^t$$

Matrices y aplicaciones lineales

Por supuesto, la matriz depende de las bases que se fijen en ambos espacios vectoriales.
Cuando se de la matriz asociada a la aplicación lineal sin especificar la base, se entenderá que las bases fijadas son las canónicas de ambos espacios.

Ejemplos

Se considera la aplicación $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por $T(x, y, z) = (2x, x + z), \forall (x, y, z) \in \mathbb{R}^3$.
Para las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ la matriz asociada a T es

$$\begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

ya que las imágenes de e_1, e_2 y e_3 son, respectivamente, los vectores $(2, 1)$, $(0, 0)$ y $(0, 1)$, que forman las tres columnas de la matriz. El hecho de que la matriz asociada a T sea la anterior significa simplemente que

$$T(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (2x, x + z), \quad \forall (x, y, z) \in \mathbb{R}^3 .$$

Ejemplos

Se considera la aplicación $T : \mathbb{R}^3 \longrightarrow \mathbb{R}^2$ dada por $T(x, y, z) = (2x, x + z), \forall (x, y, z) \in \mathbb{R}^3$.

En $\mathbb{R}^3$ consideramos la base $B = \{(1, 1, 0), (1, 0, 1), (0, 1, 1)\}$, y en $\mathbb{R}^2$ consideramos la base canónica. Como se verifica que

$$T(1, 1, 0) = (2, 1), \quad T(1, 0, 1) = (2, 2), \quad T(0, 1, 1) = (0, 1) \ ,$$

la matriz asociada a T para estas nuevas bases viene dada por

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

Teorema

Sean X e Y espacios vectoriales sobre el cuerpo $\mathbb{K}$ de dimensiones m y n, respectivamente. Si fijamos una base B_X de X y otra base B_Y de Y, entonces la aplicación $\Phi : \mathcal{M}_{n\times m}(\mathbb{K}) \longrightarrow L(X, Y)$ dada por

$$\phi(A)(x) = Ax^t \quad \forall A \in \mathcal{M}_{n\times m}(\mathbb{K}), x \in X$$

es un isomorfismo lineal entre espacios vectoriales, esto es, Φ es biyectiva y lineal.

Además, la identificación anterior lleva producto de matrices en composición de aplicaciones, esto es, si X, Y y Z son espacios vectoriales sobre $\mathbb{K}$ de dimensiones m, n y p, respectivamente, se tiene que

$$\Phi(BA) = \Phi(B) \circ \Phi(A), \quad \forall A \in \mathcal{M}_{n\times m}(\mathbb{K}), B \in \mathcal{M}_{p\times n}(\mathbb{K})$$

Es inmediato comprobar la primera parte del enunciado.
Probaremos que la identificación Φ transforma el producto de matrices en composición de aplicaciones. Tengamos en cuenta que las aplicaciones $\Phi(BA)$ y $\Phi(B) \circ \Phi(A)$ son aplicaciones lineales de X en Z. Dado un vector $x \in X$ se verifica que

$$\Phi(BA)(x) = (BA)x^t = B(Ax^t) = \Phi(B)((Ax^t)^t) =$$

$$\Phi(B)(\Phi(A)(x)) = \big(\Phi(B) \circ \Phi(A)\big)(x).$$

Como x es un vector arbitrario de X hemos probado que
$\Phi(BA) = \Phi(B) \circ \Phi(A)$.

Proposición

Sean X e Y dos espacios vectoriales sobre el cuerpo $\mathbb{K}$ y $T : X \longrightarrow Y$ una aplicación lineal.

- Si $B := \{x_1, \cdots, x_n\}$ es una base de X, entonces $\{T(x_1), \cdots, T(x_n)\}$ es un sistema de generadores de $\mathrm{Im}\ T$.
- Si consideramos en Y una base $\{y_1, \cdots, y_m\}$ y A es la matriz asociada a T en términos de las bases fijadas, entonces los vectores que aparecen en cada una de las columnas de A generan la imagen de T.

Demostración:
Probamos el primer apartado; el segundo es consecuencia del primer apartado y del hecho de que los vectores que aparecen en las columnas de A son $\{T(x_1), \cdots, T(x_n)\}$.
Sea $v \in \mathrm{Im}\ T$; entonces existe $u \in X$ tal que $v = T(u)$. Como B es una base de X, entonces u es una combinación lineal de B, esto es, existen escalares $\{t_1, \cdots, t_n\}$ tales que $u = \sum_{i=1}^{n} t_i x_i$, entonces

$$v = T(u) = T\left(\sum_{i=1}^{n} t_i x_i\right) = \sum_{i=1}^{n} t_i T(x_i),$$

y v es una combinación lineal de los vectores $\{T(x_1), \cdots, T(x_n)\}$, luego estos vectores generan la imagen de T.

Ejemplo

Si consideramos la aplicación lineal $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3$ asociada a la matriz A, siendo $A = \begin{pmatrix} 2 & 2 & 2 & 0 \\ 0 & -1 & 1 & 0 \\ 1 & 2 & 0 & 0 \end{pmatrix}$ entonces sabemos que

$$\dim \mathrm{Im}\ f = 2(= \mathrm{rg}\ (A)),$$

por tanto, el núcleo de f tiene dimensión 2, ya que

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

Además, el subespacio $\mathrm{Im}\ f$ está generado por $\{(2,0,1),(2,-1,2)\}$ mientras que

$$\mathrm{Ker}\ f = \{(x,y,u,v) \in \mathbb{R}^4 : f(x,y,u,v) = 0\}\ .$$

Ejemplo

Como

$$f(x,y,u,v) = A\begin{pmatrix} x \\ y \\ u \\ v \end{pmatrix} = (2x+2y+2u, -y+u, x+2y)$$

entonces

$$\mathrm{Ker}\, f = \{(x,y,u,v) \in \mathbb{R}^4 : -y+u=0, x+2y=0\}\ .$$

Algunas propiedades más de las matrices obtenidas usando operadores

Otra propiedad más del producto de matrices

Se verifica que

$$\mathrm{rg}\ (AB) \leq \min\{\mathrm{rg}\ (A), \mathrm{rg}\ (B)\}, \quad \forall A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}\ .$$

Demostración. Llamamos S y T a las aplicaciones lineales asociadas a las matrices A y B. Entonces sabemos que $S \in L(\mathbb{R}^n, \mathbb{R}^m)$, $T \in L(\mathbb{R}^p, \mathbb{R}^n)$. Tenemos que

$$\mathrm{rg}\ (AB) = \dim \mathrm{Im}\ (S \circ T) \leq \dim \mathrm{Im}\ S = \mathrm{rg}\ A\ ,$$

donde hemos usado que $\mathrm{Im}\ (S \circ T) \subset \mathrm{Im}\ S$, luego $\dim \mathrm{Im}\ (S \circ T) \leq \dim \mathrm{Im}\ S$. Razonando de forma parecida, obtenemos

$$\mathrm{rg}\ (AB) = \dim \mathrm{Im}\ (S \circ T) \leq \dim \mathrm{Im}\ T = \mathrm{rg}\ B\ ,$$

ya que si B es una base de $\mathrm{Im}\ T$, entonces $S(B)$ es un sistema de generadores de $\mathrm{Im}\ (S \circ T)$, luego $\dim \mathrm{Im}\ (S \circ T) \leq \dim \mathrm{Im}\ T$.

Proposición

Una matriz A cuadrada de orden n tiene inversa si, y sólo si, $\mathrm{rg}\ A = n$.

Demostración.
Si A tiene inversa, tenemos $AA^{-1} = I_n = A^{-1}A$. Si llamamos S y T a las aplicaciones lineales en $\mathbb{R}^n$ asociadas a S y T tendremos $S \circ T = I = T \circ S$, en particular, S es inyectivo. Por tanto $\mathrm{Ker}\ S = \{0\}$, luego $\mathrm{rg}\ A = \dim \mathrm{Im}\ S = n$.
Recíprocamente, si $\mathrm{rg}\ (A) = n$, entonces $\dim \mathrm{Im}\ S = n$, luego $\dim \mathrm{Ker}\ S = \{0\}$, luego S es inyectivo y sobreyectivo, entonces S tiene inversa, que sabemos que es lineal. Usando que la identificación de matrices y aplicaciones lineales transforma productos en composición de aplicaciones, la matriz asociada a la inversa de S es la inversa de A.

Traza de una matriz

Definición

Dada una matriz cuadrada A de orden n, se define la traza de A, que notaremos $\mathrm{tr}\,(A)$, como la suma de los elementos de la diagonal principal de A, esto es,

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

Ejemplo

Si $A = \begin{pmatrix} 1 & -1 & 0 \\ 2 & 4 & -6 \\ 3 & 1 & 7 \end{pmatrix}$, entonces $\mathrm{tr}\, A = 1 + 4 + 7 = 12$

Traza de una matriz

Propiedades de la aplicación traza

- La traza es una aplicación lineal de $\mathcal{M}_n(\mathbb{K})$ en $\mathbb{K}$ ($\mathbb{K}=\mathbb{R}$ ó $\mathbb{K}=\mathbb{C}$), esto es,
$$\operatorname{tr}(sA+B)=s\operatorname{tr}(A)+\operatorname{tr}(B), \quad \forall A,B\in\mathcal{M}_n(\mathbb{K}), s\in\mathbb{K}\,.$$
- Se verifica que
$$\operatorname{tr}(AB)=\operatorname{tr}(BA), \forall A\in\mathcal{M}_{m\times n}, B\in\mathcal{M}_{n\times m}\,.$$

Demostración.
La primera propiedad es trivial. Comprobamos la segunda. Llamamos $C=AB$, $D=BA$, entonces

$$\operatorname{tr}(C)=\sum_{i=1}^{m}c_{ii}=\sum_{i=1}^{m}\sum_{k=1}^{n}a_{ik}b_{ki}=\sum_{k=1}^{n}\sum_{i=1}^{m}b_{ki}a_{ik}=\operatorname{tr}(D)$$

Traza de una matriz

En general, la aplicación traza no lleva producto de matrices en producto, esto es, si $A, B \in \mathcal{M}_n$, en general se tiene que

$$\mathrm{tr}\,(AB) \neq \mathrm{tr}\,(A)\mathrm{tr}\,(B)\ .$$

Por ejemplo, si $A = B = I_n$, entonces $AB = I_n$, con lo que, si $n > 1$

$$\mathrm{tr}\,(AB) = n \neq n^2 = \mathrm{tr}\,(A)\mathrm{tr}\,(B)\ .$$

Por la misma razón, no llevan inversos en inversos.

Permutaciones

Definición

Una permutación del conjunto $\{1, 2, \cdots, n\}$ es una aplicación biyectiva $\sigma : \{1, 2, \cdots, n\} \longrightarrow \{1, 2, \cdots, n\}$. Una trasposición en $\{1, 2, \cdots, n\}$ es una aplicación $\sigma : \{1, 2, \cdots, n\} \longrightarrow \{1, 2, \cdots, n\}$ tales que existen $i, j \in \{1, 2, \cdots, n\}$ de forma que

$$\sigma(i) = j, \ \sigma(j) = i, \ \sigma(k) = k, \forall k \in \{1, 2, \cdots, n\} \backslash \{i, j\} \ .$$

Notaremos por Π_n al conjunto de todas las permutaciones de $\{1, 2, \cdots, n\}$.

Se verifica que toda permutación se puede expresar como composición de trasposiciones.

Si $\sigma \in \Pi_n$, notaremos por s_σ al número de trasposiciones necesarias para expresar σ como composición de trasposiciones (para alguna descomposición). También coincide con el número de trasposiciones necesarias para que al componer con σ se obtenga la identidad.

Permutaciones

Ejemplo

Si $n = 3$ y $\sigma = \{3, 2, 1\}$, esto es, $\sigma(1) = 3, \sigma(2) = 2, \sigma(3) = 1$. Si ahora intercambiamos 1 y 3, obtenemos la identidad. Como podemos usar una trasposición para componer con σ y conseguir la identidad, en este caso tenemos $s_\sigma = 1$.

Determinante de una matriz

Definición

Dada una matriz cuadrada A de orden n, se define el determinante de A, que notaremos por $\det(A)$ ó $|A|$, por la suma siguiente

$$\det(A) = \sum_{\sigma \in \Pi_n} (-1)^{s_\sigma} a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \ .$$

Cálculo del determinante

Supongamos que $n = 2$, entonces hay dos permutaciones del conjunto $\{1, 2\}$, que son $I = \{1, 2\}$ y $\tau = \{2, 1\}$. Como la segunda es una trasposición tenemos $s_I = 0$, $s_\tau = 1$, por lo que si $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, tenemos

$$\det(A) = (-1)^{s_I} a_{11} a_{22} + (-1)^{s_\tau} a_{12} a_{21} = a_{11} a_{12} - a_{12} a_{21}$$

Cálculo del determinante

Si $n = 3$, entonces hay seis permutaciones del conjunto $\{1, 2, 3\}$, que listamos a continuación:

$$\begin{array}{ll} I = \{1,2,3\} & s_I = 0 \\ \sigma_2 = \{2,1,3\} & s_{\sigma_2} = 1 \\ \sigma_3 = \{2,3,1\} & s_{\sigma_3} = 2 \\ \sigma_4 = \{3,2,1\} & s_{\sigma_4} = 1 \\ \sigma_5 = \{3,1,2\} & s_{\sigma_5} = 2 \\ \sigma_6 = \{1,3,2\} & s_{\sigma_6} = 1 \end{array}$$

Por tanto, si $A \in \mathcal{M}_3$, obtenemos

$$\det(A) = a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} +$$

$$-a_{13}a_{22}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32}$$

Determinante

Reescribiendo la expresión anterior obtenemos

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

que es el desarrollo del determinante usando la primera fila de la matriz.

De forma análoga, se puede desarrollar por cada fila y por cada columna.

Propiedades del determinante

Sean $A, B, C \in \mathcal{M}_n$ y $s \in \mathbb{R}$. Se verifica que:

- Si todas las filas de A, B y C coinciden, salvo la j-ésima, y éstas verifican que la fila j-ésima de C es la suma de las filas j-ésimas de A y B, entonces
$$\det(C) = \det(A) + \det(B)\ .$$
- Si se multiplica una fila de A por s, entonces el valor del determinante también se multiplica por s.
- $\det(sA) = s^n \det(A)$.
- Si alguna de las filas de A es nula, entonces $det(A) = 0$.
- Si en una matriz se intercambian dos filas, el determinante de la nueva matriz cambia de signo.
- Si dos filas son proporcionales en una matriz, entonces su determinante es cero.

Determinante

Propiedades del determinante

Sean $A, B, C \in \mathcal{M}_n$ y $s \in \mathbb{R}$. Se verifica que:

- Si en una matriz se cambia una fila por la suma de ésta con una combinación lineal de las restantes, el determinante no cambia.
- $\det(AB) = \det(A)\det(B)$, como consecuencia $\det A^{-1} = \frac{1}{\det(A)}$ si A tiene inversa.

Proposición

Si $A \in \mathcal{M}_n$, se tiene que $\mathrm{rg}\,(A) < n$ si, y sólo si, $\det(A) = 0$.

Demostración.
Si $\mathrm{rg}\,(A) < n$, entonces los n vectores que aparecen en las columnas de A son linealmente dependientes. Por las propiedades del determinante, se tiene $\det(A) = 0$.
Supongamos que $\det(A) = 0$ y que $\mathrm{rg}\,(A) = n$. Usando la identificación de matrices y operadores y el Teorema de la dimensión, se obtiene que A es inversible. Por tanto $AA^{-1} = I_n$. Aplicando el determinante y el hecho de que conserva productos, se obtiene $1 = \det I_n = \det A \det A^{-1}$, luego $\det A \neq 0$, y obtenemos una contradicción. Por tanto, si $\det(A) = 0$, entonces $\mathrm{rg}\,(A) < n$.

Proposición

Si $A \in \mathcal{M}_{m\times n}$. Si $\mathrm{rg}\,(A) = r$, entonces A tiene r vectores fila linealmente independientes y éste es el número máximo de vectores fila que verifican esa condición.

Demostración.
Si A tiene rango r, sabemos que tiene r columnas linealmente independientes y éste es el número máximo de columnas que verifica esa condición. Sea s el máximo número de filas de A linealmente independientes. Probaremos que $s = r$.
Podemos suponer que las s primeras filas son linealmente independientes. Así, pues, para el resto de las filas, es decir, si $i = s+1, \cdots, m$ existen escalares $\{t_{ij}\}$ tales que

$$a_{ij} = t_{i1}a_{1j} + \cdots + t_{is}a_{sj}, \quad \forall j \in \{1, \cdots, m\}\ .$$

$$a_{ij} = t_{i1}a_{1j} + \cdots + t_{is}a_{sj}, \quad \forall j \in \{1, \cdots, m\}, i = s+1, \cdots, m.$$

Por esto, la columna j-ésima de A ($1 \leq j \leq n$) tendrá la forma

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{sj} \\ \sum_{k=1}^{s} t_{s+1\,k}a_{kj} \\ \vdots \\ \sum_{k=1}^{s} t_{mk}a_{kj} \end{pmatrix}.$$

Rango de una matriz

La columna j-ésima de A ($1 \leq j \leq n$) es

$$\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{sj} \\ \sum_{k=1}^{s} t_{s+1k} a_{kj} \\ \vdots \\ \sum_{k=1}^{s} t_{mk} a_{kj} \end{pmatrix},$$

luego la podemos escribir como la siguiente combinación lineal:

$$a_{1j} \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \\ t_{s+11} \\ \vdots \\ t_{m1} \end{pmatrix} + a_{2j} \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \\ t_{s+12} \\ \vdots \\ t_{m2} \end{pmatrix} + \cdots + a_{sj} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ t_{s+1s} \\ \vdots \\ t_{ms} \end{pmatrix}$$

Rango de una matriz

$$a_{1j}\begin{pmatrix}1\\0\\\vdots\\0\\t_{s+11}\\\vdots\\t_{m1}\end{pmatrix}+a_{2j}\begin{pmatrix}0\\1\\\vdots\\0\\t_{s+12}\\\vdots\\t_{m2}\end{pmatrix}+\cdots+a_{sj}\begin{pmatrix}0\\\vdots\\1\\t_{s+1s}\\\vdots\\t_{ms}\end{pmatrix}$$

Si llamamos $w_1, w_2, \cdots, w_s$ a las matrices columna que aparecen antes, hemos comprobado que el subespacio vectorial generado por $\{w_1, \cdots, w_s\}$ genera todas las columnas de la matriz, luego como máximo, hay s linealmente independientes. Por tanto $\mathrm{rg}\,(A) = r \leq s$.
Cambiando el papel de filas y columnas de la matriz, se comprueba que $s \leq r$.

Rango de una matriz no cuadrada

Proposición

Si $A \in \mathcal{M}_{m\times n}$ y $r := \mathrm{rg}\,(A) \leq \min\{m, n\}$, si y sólo si, existe una submatriz cuadrada A_r de A de orden r tal que $\det A_r \neq 0$ y r es el mayor natural que verifica esa condición.

Definición

Si $A \in \mathcal{M}_{m\times n}$, una submatriz de A es una matriz que se obtiene a partir de A suprimiendo filas y columnas.

Demostración.

Si $\mathrm{rg}\,(A) = r$, por definición, sabemos que ha de haber r columnas de A que son linealmente independientes. Podemos suponer sin pérdida de generalidad que son las r primeras columnas. Esto es, las columnas de la submatriz B son linealmente independientes

$$B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mr} \end{pmatrix}$$

Rango de una matriz no cuadrada

Por el resultado que hemos probado antes, B ha de tener también r filas linealmente independientes; de nuevo podemos suponer que son las r primeras, con lo que la matriz cuadrada C tiene rango r, siendo

$$C = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1r} \\ a_{21} & a_{22} & \cdots & a_{2r} \\ \vdots & \vdots & & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rr} \end{pmatrix} .$$

Por el resultado ya probado para matrices cuadradas , si C tiene rango máximo, entonces $\det C \neq 0$ y es claro que C es una submatriz de A. Basta tomar $A_r = C$. Como r es el número máximo de columnas lin. independientes, y para matrices cuadradas hay equivalencia entre tener rango máximo y determinante no nulo, entonces todas las submatrices cuadradas de A de orden mayor que r tienen determinante no nulo.
Para probar la otra implicación, basta usar que la hipótesis implica que la submatriz A_r tiene rango r, luego las r columnas de A_r son lin. independientes. Por ser A_r una submatriz de A, hay (al menos) r columnas de A linealmente independientes, luego $\mathrm{rg}\,(A) \geq r$.

Si $\mathrm{rg}\,(A) > r$, entonces, por lo probado en la primera parte, habría una submatriz de A cuadrada e inversible de orden mayor que r, cosa que no ocurre.

Ejemplo

La matriz A dada por

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

tiene rango 3 ya que $\det A = 1 + (3-2) = 2 \neq 0$.

Ejemplo

La matriz A dada por

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

tiene rango 2 ya que $\det A_2 = \det \begin{pmatrix} 2 & 1 \\ 3 & 4 \end{pmatrix} = 8 - 3 = 5 \neq 0$ y las únicas 4 submatrices de A de orden 3 no son inversibles, ya que son las siguientes

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

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

y todas tienen determinante cero.

Tipos de matrices

Definición

Sea $A \in \mathcal{M}_n(\mathbb{K})$

- La matriz A es triangular superior si todos los elementos situados debajo de la diagonal principal son nulos, esto es, si

$$a_{ij} = 0, \quad \forall i > j \, .$$

- La matriz A es triangular inferior si todos los elementos situados encima de la diagonal principal son nulos, esto es, si

$$a_{ij} = 0, \quad \forall i < j \, .$$

- Si $B \in \mathcal{M}_n(\mathbb{K})$, diremos que B es la matriz traspuesta de A si las filas de B son las columnas de A, es decir, si

$$b_{ij} = a_{ji}, \quad \forall i, j \, .$$

  Notaremos por A^t a la matriz traspuesta de A.

Tipos de matrices

Definición

Sea $A \in \mathcal{M}_n(\mathbb{K})$

- A es simétrica si coincide con su traspuesta, esto es, si
$$a_{ij} = a_{ji}, \quad \forall i, j \ .$$
- A es antisimétrica si $A^t = -A$, esto es, si
$$a_{ij} = -a_{ji}, \quad \forall i, j \ .$$
- A es hermitiana si $\overline{A} = A^t$, donde $\overline{A}$ es la matriz conjugada de A, esto es, si
$$\overline{a_{ij}} = a_{ji}, \quad \forall i, j \ .$$
En caso de que $\mathbb{K} = \mathbb{R}$ hermitiana significa simplemente simétrica.

Definición

Sea $A \in \mathcal{M}_n(\mathbb{K})$

- A es semipositiva si se verifica

$$a_{ij} \geq 0, \quad \forall i, j \ .$$

- A es positiva si se verifica

$$a_{ij} \geq 0, \quad \forall i, j \ \text{ y } \ \text{existen} \quad i_0, j_0 : a_{i_0 j_0} > 0 \ .$$

- A es estrictamente positiva si

$$a_{ij} > 0, \quad \forall i, j \ .$$

- A es estocástica si es semipositiva y la suma de cada una de las filas es uno, esto es, si

$$a_{ij} \geq 0, \quad \forall i, j, \quad \sum_{j=1}^{n} a_{ij} = 1, \forall i \ .$$

Tipos de matrices

Definición

Sea $A \in \mathcal{M}_n(\mathbb{K})$

- Una matriz estocástica A es doblemente estocástica si la suma de los elementos de cada una de sus columnas es uno, esto es, si

$$\sum_{i=1}^{n} a_{ij} = 1, \forall j \, .$$

- A es ortogonal si

$$AA^t = I_n \, .$$

  Esto ocurre cuando las filas son ortogonales dos a dos y tienen norma euclídea 1. Las matrices ortogonales en $\mathbb{R}^n$ corresponden a las aplicaciones lineales que son isometrías en $\mathbb{R}^n$.

- A es idempotente si

$$A^2 = A \, .$$

  Las matrices idempotentes corresponden a las proyecciones.

Propiedades de las matrices triangulares superiores

- El conjunto de las matrices cuadradas de orden n que son triangulares superiores es un subespacio vectorial del espacio $\mathcal{M}_n(\mathbb{K})$ que es estable por productos e inversas.
- Si $A \in \mathcal{M}_n(\mathbb{K})$ es triangular superior, entonces $\mathrm{rg}\,(A)$ es mayor o igual que el máximo número de elementos no nulos de la diagonal principal.
- Si $A \in \mathcal{M}_n(\mathbb{K})$ es triangular superior, entonces $\det(A) = \prod_{i=1}^{n} a_{ii}$

Demostración.

- Es inmediato comprobar que el conjunto de las matrices cuadradas de orden n que son triangulares superiores es un subespacio vectorial. Si $A, B \in \mathcal{M}_n(\mathbb{K})$ son triangulares superiores e $i > j$, entonces

$$1 \leq k \leq n \quad \Rightarrow \quad k < i \ \text{ó}\ k \geq i > j \quad \Rightarrow \quad a_{ik} = 0 \quad \text{ó} \quad b_{kj} = 0\,.$$

Por tanto, si $C = AB$, entonces $c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj} = 0$ para $i > j$, luego AB es triangular superior.

Propiedades de las matrices triangulares superiores

Demostración.

- Para comprobar la estabilidad por inversas basta observar que en este caso, aplicando el método de Gauss-Jordan, se obtiene una matriz triangular superior.
- Usando la definición de rango, se puede comprobar que para matrices triangulares superiores, las r columnas cuyos elementos de la diagonal son no nulos son linealmente independientes, luego el rango es mayor o igual que el máximo número de elementos de la diagonal principal no nulos.
- Basta desarrollar el determinante usando la última columna de la matriz y usar inducción sobre el orden de la matriz, ya que suprimiendo en una matriz triangular superior la última fila y la última columna se obtiene otra matriz del mismo tipo y de orden uno menos que la original.

Tipos de matrices

Ejemplo

La matriz $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ tiene rango 3 y sólo dos elementos en la diagonal no nulos.

Proposición

Sea $A \in \mathcal{M}_n$ tal que $\det A_i \neq 0$ para $1 \leq i \leq n-1$, donde

$$A_i = \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ a_{21} & \cdots & a_{2i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix}.$$ Entonces A se puede expresar de la forma $A = LU$ donde $L, U \in \mathcal{M}_n$, L es triangular inferior y los elementos de la diagonal principal son 1 y U es triangular superior.

Demostración.

Basta aplicar el método de Gauss-Jordan a A para encontrar la matriz L, que es la matriz que se obtiene multiplicando A por las transformaciones necesarias para transformar a A en una matriz diagonal superior. En este caso, basta multiplicar filas de A por escalares y sumar a una fila de A una combinación lineal de otras filas.

Ejemplo

Dada la matriz $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 1 \\ 0 & 5 & 1 \end{pmatrix}$, la descomponemos de la forma que aparece en la proposición anterior. En este caso las matrices L y U han de tener la siguiente forma

$$L = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}.$$

Imponemos que $A = LU$ y resolvemos el sistema que aparece.

Tipos de matrices

Ejemplo

$$LU = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix} =$$

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} \\ l_{21}u_{11} & l_{21}u_{12} + u_{22} & l_{21}u_{13} + u_{23} \\ l_{31}u_{11} & l_{31}u_{12} + l_{32}u_{22} & l_{31}u_{13} + l_{32}u_{23} + u_{33} \end{pmatrix} = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 1 \\ 0 & 5 & 1 \end{pmatrix}.$$

Nótese que igualando los elementos y resolviendo el sistema se puede averiguar la primera fila de U, la primera columna de L, segunda fila de U, segunda columna de L, etc.

Ejemplo

En este caso, obtenemos

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 5 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

Propiedades de la trasposición de matrices

Sean $A \in \mathcal{M}_{m\times n}, C \in \mathcal{M}_{n\times p}$. Se verifica que

- La trasposición de matrices es una aplicación lineal.
- $(AC)^t = C^t A^t$.
- $(A^t)^t = A$.
- $\mathrm{rg}\,(A^t) = \mathrm{rg}\,(A)$.
- Si $A \in \mathcal{M}_n$, entonces $\mathrm{tr}\,(A^t) = \mathrm{tr}\,(A)$ y $\det A = \det A^t$.
- Si $A \in \mathcal{M}_n$ y $\det A \neq 0$, entonces $(A^t)^{-1} = (A^{-1})^t$
- Si $A \in \mathcal{M}_n$, entonces $\mathrm{rg}\,(A) = \mathrm{rg}\,(A^t A) = \mathrm{rg}\,(AA^t)$.

Propiedades de las matrices simétricas

- El conjunto de las matrices simétricas de orden n es un espacio vectorial.
- Si $A, B \in \mathcal{M}_n$ son matrices simétricas y $AB = BA$, entonces AB es simétrica.
- Si A es inversible y simétrica, entonces A^{-1} también es simétrica.
- Si $C \in \mathcal{M}_{m\times n}$, entonces C^tC es simétrica.

Tipos de matrices

Propiedades de las matrices antisimétricas

- El conjunto de las matrices antisimétricas de orden n es un espacio vectorial.
- Si $A, B \in \mathcal{M}_n$ son matrices antisimétricas y $AB = BA$, entonces AB es antisimétrica.
- Si A es inversible y antisimétrica, entonces A^{-1} también es antisimétrica.
- Si $A \in \mathcal{M}_n$ es una matriz antisimétrica, entonces $\det A = (-1)^n \det(A)$, por tanto, si n es impar, entonces $\det(A) = 0$.

Propiedades de las matrices hermitianas

- El conjunto de las matrices hermitianas de orden n es un espacio vectorial real.
- Si $A, B \in \mathcal{M}_n$ son matrices hermitianas y $AB = BA$, entonces AB es hermitiana.
- Si A es hermitiana, entonces A^t y $\overline{A}$ son también hermitianas.
- Si $A \in \mathcal{M}_n$ es una matriz hermitiana e inversible, entonces A^{-1} es también hermitiana.
- Si $A \in \mathcal{M}_n$ es una matriz hermitiana, entonces $\det A \in \mathbb{R}$.

Tipos de matrices

Propiedades de las matrices estocásticas

- Si $A \in \mathcal{M}_{m\times n}, B \in \mathcal{M}_{n\times p}$ son matrices estocásticas, entonces AB es estocástica.
- Si $A \in \mathcal{M}_n$ es una matriz estocástica, entonces A^k también es estocástica, para todo natural k.
- Si $A \in \mathcal{M}_{m\times n}$ es una matriz doblemente estocástica, entonces A^t también es lo es.

Propiedades de las matrices ortogonales

Sean $A, B \in \mathcal{M}_n$ matrices ortogonales, entonces

- AB es ortogonal.
- $\det(A) = 1$ ó $\det(A) = -1$.
- $\mathrm{rg}\,(A) = n$.
- A^{-1} es ortogonal y $A^{-1} = A^t$.

Propiedades de las matrices idempotentes

Sean $A, B \in \mathcal{M}_n$ matrices idempotentes.

- Si $AB = BA$, entonces AB es idempotente.
- Si $C = C^t C$, entonces C es simétrica e idempotente.
- $\det(A) \in \{0, 1\}$.
- Si $\det(A) = 1$, entonces $A = I_n$.
- $I_n - A$ es idempotente.

Matriz traspuesta

Base dual de una dada

Sea X un espacio vectorial sobre $\mathbb{K}$ y $B := \{e_1, \cdots, e_n\}$ una base de X. Entonces el conjunto $\{f_1, \cdots, f_n\}$, donde

$$f_j\Big(\sum_{k=1}^{n} x(k)e_k\Big) = x(j), \quad \forall (x(1), x(2), \cdots, x(n)) \in \mathbb{K}^n, 1 \leq j \leq n$$

es una base del espacio $X^* = L(X, \mathbb{K})$ que se llama base dual de B.

Operador traspuesto de uno dado

Sean X e Y espacios vectoriales sobre $\mathbb{K}$ y $T \in L(X, Y)$, entonces se define el operador traspuesto de T, que notaremos por T^t como sigue

$$T^t(f) = f \circ T, \quad \forall f \in Y^* .$$

Se verifica que $T^t : Y^* \longrightarrow X^*$ es una aplicación lineal.

Proposición

Sean X e Y espacios vectoriales sobre $\mathbb{K}$ y $T \in L(X, Y)$. Si $B_X = \{x_1, \cdots, x_n\}$ es una base de X, $B_Y = \{y_1, \cdots, y_m\}$ una base de Y y $A \in \mathcal{M}_{m\times n}$ la matriz asociada a T es función de las bases fijadas, entonces A^t es la matriz asociada a T^t en función de las bases duales de B_Y y de B_X.

Demostración.
Tenemos que probar que $T^t(f) = A^t f^t, \forall f \in Y^*$, usando en Y^* y en X^* las coordenadas asociadas a las bases duales de B_Y y de B_X. Como las aplicaciones $f \mapsto T^t(f), f \mapsto A^t f^t$ son lineales, basta comprobar que coinciden sobre una base de Y^*. Llamamos $\{f_1, \cdots, f_n\}$ y $\{g_1, \cdots, g_m\}$ a las bases duales de B_X y B_Y, respectivamente, por tanto, sabemos que

$$f_j(x_i) = \delta_{ij}, \forall 1 \leq i, j \leq n, \quad g_k(y_s) = \delta_{sk}, \forall 1 \leq k, s \leq m\,.$$

Matriz traspuesta

Sabemos que

$$T^t(g_j)(x_i) = (g_j \circ T)(x_i) = g_j(T(x_i)) \; .$$

Ahora bien

$$T(x_i) = Ax_i^t = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2i} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mi} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1 \; i) \\ \vdots \\ 0 \end{pmatrix} =$$

$$(a_{1i}, a_{2i}, \cdots, a_{mi}) = \sum_{k=1}^{m} a_{ki} y_k \; .$$

Como consecuencia $g_j(T(x_i)) = a_{ji}$.

Matriz traspuesta

Hemos comprobado que $T^t(g_j)(x_i) = a_{ji}$. Ahora hemos de comprobar que a_{ji} es la coordenada i-ésima de $A^t {g_j}^t$.

$$A^t {g_j}^t = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{j1} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{j2} & \cdots & a_{m2} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{jn} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 1\ j) \\ \vdots \\ 0 \end{pmatrix} =$$

$$(a_{j1}, a_{j2}, \cdots, a_{jn}) = \sum_{k=1}^{m} a_{jk} f_k \ .$$

Por tanto, la coordenada i-ésima de $\sum_{k=1}^m a_{jk} f_k$ es $(\sum_{k=1}^m a_{jk} f_k)(x_i) = a_{ji}$. Hemos comprobado que $T^t(g_j)(x_i) = a_{ji} = A^t g_j^t(i)$, para todo $1 \leq i \leq n$. Como dos funcionales lineales en X coinciden sobre B_X, entonces $T^t(g_j) = A^t g_j^t$ para cada $1 \leq j \leq m$. Usando de nuevo la linealidad de $g \mapsto T^t(g)$ y de $g \mapsto A^t g^t$, obtenemos que T^t está representada por la matriz A^t.

Tema 3. Diagonalización de matrices.

Análisis Matemático I
1^o Licenciatura de Estadística

Algunos ejemplos

Ejemplos de matrices asociadas a operadores sencillos

Consideramos algunos operadores lineales en $\mathbb{R}^2$ y sus matrices asociadas

- Homotecia de razón k $\quad T(x,y) = (kx, ky)$

$$A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}$$

- Giro de ángulo α $\quad T(x,y) = (x\cos\alpha - y \operatorname{sen}\alpha, x\operatorname{sen}\alpha + y\cos\alpha)$

$$A = \begin{pmatrix} \cos\alpha & -\operatorname{sen}\alpha \\ \operatorname{sen}\alpha & \cos\alpha \end{pmatrix}$$

- Simetría respecto del origen $\quad T(x,y) = (-x,-y)$

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

Algunos ejemplos

Ejemplos de matrices asociadas a operadores sencillos

- Simetría respecto de la recta $y = x$ $\quad T(x, y) = (y, x)$

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

En el primer ejemplo, cualquier subespacio vectorial de $\mathbb{R}^2$ es invariante por el operador y k es el único valor propio.
En el segundo ejemplo, no hay subespacios vectoriales no triviales invariantes por el operador asociado a la matriz, ni valores propios (salvo que α sea múltiplo de 2π). En el tercero, -1 es el único valor propio y el operador fija cualquier subespacio vectorial.
En el último, el subespacio vectorial $\{(x, x) : x \in \mathbb{R}\}$ es invariante por T, también lo es $\{(x, -x) : x \in \mathbb{R}\}$ y 1 y -1 son los únicos valores propios.

Vectores propios y valores propios

Definición

Sea X un espacio vectorial sobre $\mathbb{K}$ y $T \in L(X)$. Diremos que un escalar λ es un valor propio (ó autovalor) de T si existe un vector v no nulo en X tal que $T(v) = \lambda v$. En tal caso, el vector v es un vector propio (ó autovector) asociado al valor propio λ.

Algunas observaciones.

- El vector propio asociado a un valor propio no es único. De hecho, si $v \neq 0$ verifica que $T(v) = \lambda v$, por ser T lineal, tenemos que $T(\alpha v) = \alpha T(v) = \alpha\lambda v = \lambda(\alpha v)$, luego αv también es un vector propio asociado a λ, para cualquier escalar no nulo α.
- Un mismo vector no puede ser vector propio asociado a dos valores propios diferentes, ya que si λ, μ son escalares y se verifica que

$$T(v) = \lambda v, \quad T(v) = \mu v \ ,$$

entonces $\lambda v = \mu v$ y como v es no nulo, ha de ser $\lambda = \mu$.

Vectores propios y valores propios

Ejemplos

- Consideramos en $\mathbb{R}^2$ la simetría respecto del origen $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por
$$T(x,y) = (-x,-y), \quad \forall (x,y) \in \mathbb{R}^2 .$$
En este caso -1 es el único valor propio y cualquier vector no nulo de $\mathbb{R}^2$ es un vector propio asociado a -1.
- Si consideramos la aplicación lineal $T : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ dada por $T(x,y) = (y,x)$, los únicos valores propios de T son 1 y -1 y los conjuntos de vectores propios asociados a los valores propios son, respectivamente,
$$\{(x,x) : x \in \mathbb{R}\} \quad \text{y} \quad \{(x,-x) : x \in \mathbb{R}\} .$$

Vectores propios y valores propios

Ejemplos

- En $\mathcal{P}_2$ (polinomios en una variable de grado menor o igual que 2) definimos el siguiente operador lineal

  $$T(a_0 + a_1 x + a_2 x^2) = 2a_0 + 2(a_0 + a_1) + 2a_2 x^2, \quad \forall a_0, a_1, a_2 \in \mathbb{R}$$

  Si imponemos la ecuación $T(p(x)) = \lambda p(x)$ para algún polinomio $p(x)$ no nulo de grado menor o igual que 2, obtenemos que $\lambda = 2$ y además el polinomio $p(x)$ tiene la forma $p(x) = a_1 x + a_2 x^2$. Es decir, 2 es el único valor propio de T y el conjunto de los vectores propios asociados a T viene dado por los polinomios no nulos de grado menor o igual que 2 sin término independiente.

Aunque en los ejemplos anteriores es muy sencillo calcular los valores propios asociados a los operadores, en general, puede no serlo tanto. Usaremos la identificación con matrices para facilitar el cálculo.

Proposición

Sea X un espacio vectorial de dimensión n sobre el cuerpo $\mathbb{K}$ y B una base de X. Si $T \in L(X)$ y $A \in \mathcal{M}_n(\mathbb{K})$ es la matriz asociada a T en términos de la base B, entonces se verifica

$$T(v) = \lambda v \quad \Leftrightarrow \quad Av^t = \lambda v \; .$$

Por tanto, λ es un valor propio de T si, y sólo si, $\det(A - \lambda I_n) = 0$.

Demostración.
La primera igualdad es consecuencia de la forma que tiene la identificación entre aplicaciones lineales y matrices.
Para la equivalencia, usaremos que, por definición, λ es un valor propio de T si, y sólo si, $\mathrm{Ker}\,(T - \lambda I) \neq \{0\}$, esto es, cuando el operador $T - \lambda I$ no es inyectivo. Equivalentemente, el operador $T - \lambda I$ no es inversible, por tanto, su matriz asociada, que es $A - \lambda I_n$ no tiene inversa, luego $\det(A - \lambda I_n) = 0$.
Nótese que en todas las implicaciones usadas se ha reformulado la condición original de manera equivalente, por lo que $\det(A - \lambda I_n) = 0$ implica que λ es un valor propio de T.

Notas

- En vista del resultado anterior, si se consideran dos bases en un espacio vectorial y las matrices asociadas a un mismo operador en términos de esas dos bases, ambas tendrán los mismos valores propios. Por supuesto, los vectores propios asociados a un mismo autovalor pueden tener coordenadas diferentes dependiendo de la base que se use.

- Convenio: A partir de ahora, cuando digamos que una matriz A es la asociada a un operador lineal en $\mathbb{K}^n$, se sobreentiende que consideramos las bases canónicas en $\mathbb{K}^n$.

Ejemplo

La aplicación lineal en $\mathbb{R}^2$ asociada a la matriz $A = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}$ no tiene valores propios (reales), ya que el determinante de la matriz $A - \lambda I_2 = \begin{pmatrix} 1-\lambda & 2 \\ & \\ -2 & 1-\lambda \end{pmatrix}$ no es nulo para ningún valor de λ.

Definición

Si A es una matriz cuadrada de orden n, se llama polinomio característico de A al polinomio P_n de grado n dado por

$$P_n(\lambda) = \det(A - \lambda I_n) \ ,$$

y la ecuación $P_n(\lambda) = 0$ es la ecuación característica de A.

Vectores propios y valores propios

Definición

Si el polinomio característico P_n de la matriz A se puede expresar de la forma

$$P_n(\lambda) = (\lambda - \lambda_i)^r Q(\lambda), \quad \forall \lambda \in \mathbb{K}$$

donde r es un número natural y Q es un polinomio de grado $n - r$ tal que $Q(\lambda_i) \neq 0$, entonces diremos que el valor propio λ_i tiene multiplicidad r.

Proposición

- Si $A \in \mathcal{M}_n$ y λ es uno de sus valores propios, entonces los vectores propios asociados a λ son las soluciones del sistema $(A - \lambda I_n)x^t = 0$. El conjunto
$$\{x \in X : Ax^t = \lambda x\}$$
es un subespacio vectorial de X.
- Si $A, B \in \mathcal{M}_n$ y B es inversible, entonces las matrices A y $C := BAB^{-1}$ tienen el mismo polinomio característico.

Demostración.
Probamos la segunda afirmación. El polinomio característico de A viene dado por $\det(A - \lambda I_n)$. Escribimos ahora el de C

$$\det(C - \lambda I_n) = \det(BAB^{-1} - \lambda I_n) =$$

$$\det(B(A - \lambda I_n)B^{-1}) = \det(B)\det(A - \lambda I_n)\det(B^{-1}) =$$

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

Ejemplo

La matriz $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ tiene como polinomio característico a

$$\det(A - \lambda I_3) = \det \begin{pmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{pmatrix} =$$

$$(1-\lambda)\Big((2-\lambda)(1-\lambda) - 1\Big) + 1(-1+\lambda) =$$

$$-\lambda^3 + 4\lambda^2 - 3\lambda = -\lambda(\lambda^2 - 4\lambda + 3) =$$

$$-\lambda(\lambda - 1)(\lambda - 3)$$

Por tanto, los valores propios son $0, 1$ y 3. Ahora calculamos los subespacios propios asociados a cada uno.

Vectores propios y valores propios

Ejemplo

Vectores propios asociados a cero

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

$$\begin{array}{l} x - y = 0 \\ -x + 2y - z = 0 \\ -y + z = 0 \end{array} \quad \Rightarrow x = y = z$$

Las soluciones son el subespacio vectorial dado por

$$\{x, x, x) : x \in \mathbb{R}\} .$$

Ejemplo

Vectores propios asociados a 1

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

$$\begin{array}{l} x - y = x \\ -x + 2y - z = y \\ -y + z = z \end{array} \quad \Rightarrow y = 0, z = -x$$

Las soluciones son el subespacio vectorial dado por

$$\{x, 0, -x) : x \in \mathbb{R}\} \ .$$

Ejemplo

Vectores propios asociados a 3

$$\begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = (3x, 3y, 3z)$$

$$\begin{array}{l} x - y = 3x \\ -x + 2y - z = 3y \\ -y + z = 3z \end{array} \quad \Rightarrow y = -2x, z = x$$

Las soluciones son el subespacio vectorial dado por

$$\{x, -2x, x) : x \in \mathbb{R}\} .$$

Propiedades de los autovalores

Proposición

Sea $A \in \mathcal{M}_n$ una matriz cuyos valores propios son $\lambda_1, \cdots, \lambda_n$. Se verifica que

- Los autovalores de A coinciden con los de A^t.
- $\det(A) = \prod_{i=1}^n \lambda_i$
- $\mathrm{tr}\,(A) = \sum_{i=1}^n \lambda_i$
- Los autovalores de αA son $\alpha\lambda_1, \cdots, \alpha\lambda_n$.
- Si $B \in \mathcal{M}_n$, los autovalores de AB y de BA son los mismos.
- Si $\det(A) \neq 0$, entonces los autovalores de A^{-1} son $\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}$.
- λ_i^k es un valor propio de A^k, para cada $1 \leq i \leq n$.
- Si $\lambda_i \neq -1\ \forall i$, entonces $I_n + A$ tiene inversa y sus valores propios son $\frac{1}{1+\lambda_1}, \cdots, \frac{1}{1+\lambda_n}$.

Demostración.

- Los autovalores de A coinciden con los de A^t.
  Es consecuencia de que $\det(B) = \det B^t$ para cada matriz cuadrada B, por tanto, $\det(A - \lambda I_n) = \det(A^t - \lambda I_n)$.
- $\det(A) = \prod_{i=1}^{n} \lambda_i$
  Por ser el polinomio característico de A un polinomio de grado n con coeficiente líder $(-1)^n$ y raíces $\lambda_1, \cdots, \lambda_n$ sabemos que

$$P(\lambda) = \prod_{i=1}^{n} (\lambda_i - \lambda).$$

  Por otra parte, sabemos que

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

  Si igualamos las dos expresiones del polinomio característico y evaluamos en $\lambda = 0$ obtenemos que $\prod_{i=1}^{n} \lambda_i = \det A$.

Propiedades de los autovalores

- $\mathrm{tr}\,(A) = \sum_{i=1}^{n} \lambda_i$
  Igual que en el apartado anterior, basta calcular el coeficiente de λ^{n-1} en las dos expresiones del polinomio característico de A e igualarlas, obteniendo que
  $$(-1)^{n-1} \sum_{i=1}^{n} \lambda_i = (-1)^{n-1} \mathrm{tr}\,(A)\ .$$
- Los autovalores de αA son $\alpha\lambda_1, \cdots, \alpha\lambda_n$.
  Basta usar que $\det(\alpha A - \alpha\lambda I_n) = \alpha^n \det(A - \lambda I_n)$, por tanto, si λ es un valor propio de A, $\alpha\lambda$ es un valor propio de αA.
- Si $B \in \mathcal{M}_n$, los autovalores de AB y de BA son los mismos.
  La demostración en caso de que una de las matrices sea inversible es especialmente sencilla. Por ejemplo, si A es inversible, sabemos que los autovalores de AB coinciden con los de $A^{-1}(AB)A = BA$.

Propiedades de los autovalores

- Si $B \in \mathcal{M}_n$, los autovalores de AB y de BA son los mismos.
  Comprobaremos que $\det(AB - \lambda I_n) = \det(BA - \lambda I_n)$, para todo escalar λ, igualdad que sabemos que es cierta para $\lambda = 0$. Supongamos entonces que $\lambda \neq 0$.
  Para ello construimos las siguientes matrices cuadradas de orden $2n$:

$$H_1 := \begin{pmatrix} I_n & A \\ 0_n & I_n \end{pmatrix}, \quad C_1 := \begin{pmatrix} A & -\lambda I_n \\ -I_n & B \end{pmatrix}.$$

Se verifica que $\det(H_1) = 1$ y

$$C_1^* := H_1 C_1 = \begin{pmatrix} 0_n & AB - \lambda I_n \\ -I_n & B \end{pmatrix}$$

Por tanto $\det(C_1) = \det(C_1^*) = \det(AB - \lambda I_n)$.

Propiedades de los autovalores

Cambiando los papeles de las matrices A y B, definimos

$$H_2 := \begin{pmatrix} I_n & B \\ 0_n & I_n \end{pmatrix}, \quad C_2 := \begin{pmatrix} B & -\lambda I_n \\ -I_n & A \end{pmatrix}.$$

Se verifica que $\det(H_2) = 1$ y

$$C_2^* := H_2 C_2 = \begin{pmatrix} 0_n & BA - \lambda I_n \\ -I_n & A \end{pmatrix}$$

y en este caso sabemos que $\det(C_2) = \det(C_2^*) = \det(BA - \lambda I_n)$.

Basta comprobar entonces que $\det(C_1) = \det(C_2)$ para terminar la demostración.

Propiedades de los autovalores

En efecto,

$$\det(C_2) = \det\begin{pmatrix} B & -\lambda I_n \\ -I_n & A \end{pmatrix} = (-1)^n \det\begin{pmatrix} -\lambda I_n & B \\ A & -I_n \end{pmatrix} =$$

$$(-1)^n(-1)^n \det\begin{pmatrix} A & -I_n \\ -\lambda I_n & B \end{pmatrix} =$$

$$\lambda^n \det\begin{pmatrix} \frac{1}{\lambda}A & -\frac{1}{\lambda}\lambda I_n \\ -I_n & B \end{pmatrix} =$$

$$\lambda^n \frac{1}{\lambda^n} \det\begin{pmatrix} A & -\lambda I_n \\ -I_n & B \end{pmatrix} = \det(C_1)\ .$$

Propiedades de los autovalores

- Si $\det(A) \neq 0$, entonces los autovalores de A^{-1} son $\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_n}$.
  Si λ es un autovalor de A y T es el operador asociado, tenemos que para algún vector x no nulo $Tx = \lambda x$. Si A tiene inversa, T también y entonces $T^{-1}(\lambda x) = x$. Como λ no puede ser cero, entonces el vector $v = \lambda x$ es un vector propio de T^{-1} asociado al valor propio $\frac{1}{\lambda}$, y lo mismo le ocurre a su matriz asociada A^{-1}.
- λ_i^k es un valor propio de A^k, para cada $1 \leq i \leq n$.
  Si T es el operador asociado a A y $Tx = \lambda x$ para algún vector no nulo x y algún escalar λ, entonces $T^2 x = T(Tx) = T(\lambda x) = \lambda^2 x$. Por inducción se prueba que $T^k(x) = \lambda^k x$, luego λ^k es un valor propio de T^k, y también de la matriz que representa a ese operador A^k.

Propiedades de los autovalores

- Si $\lambda_i \neq -1$ $\forall i$, entonces $I_n + A$ tiene inversa y sus valores propios son $\frac{1}{1+\lambda_1}, \dots, \frac{1}{1+\lambda_n}$.
  El hecho de que -1 no sea valor propio de la matriz A significa que $\det(A + I_n) \neq 0$, luego $A + I_n$ es inversible, por tanto 0 no es un valor propio de $A + I_n$.
  Si λ es un valor propio de A y v un vector propio asociado a λ, entonces $(A + I_n)v^t = \lambda v + v = (1 + \lambda)v$, luego $1 + \lambda$ es un valor propio de $A + I_n$.
  De la igualdad anterior se deduce también que todos los valores propios de $A + I_n$ se obtienen sumando uno a un valor propio de A, ya que

  $$(A + I_n)v^t = \lambda v \quad \Rightarrow \quad Av^t = (\lambda - 1)v$$

  luego si λ es un autovalor de $A + I_n$, entonces $\lambda - 1$ es un valor propio de A.
  Finalmente, usando el apartado que nos asegura que los valores propios de la inversa de una matriz son los inversos de los valores propios de la matriz se obtiene que los valores propios de la inversa de $A + I_n$ son $\frac{1}{1+\lambda_1}, \dots, \frac{1}{1+\lambda_n}$.

Propiedades de los vectores propios

Es inmediato comprobar que si v es un vector propio de una matriz cuadrada A de orden n, entonces v es también un vector propio de las matrices αA (α es un escalar), A^k (para k natural) y de $I_n + A$. Enunciamos algunas propiedades más de los vectores propios.

Proposición

Sean $A \in \mathcal{M}_n$, $\lambda_1, \cdots, \lambda_r$ los valores propios distintos de A, con multiplicidades $m_1, \cdots, m_r$ y supongamos que se verifica que $\sum_{i=1}^r m_i = n$. Entonces

- Para cada i, el conjunto

$$V(\lambda_i) := \{v \in \mathbb{K}^n : Av^t = \lambda_i v\}$$

  es un subespacio vectorial de dimensión menor o igual que m_i (llamado el subespacio propio asociado a λ_i).
- Los vectores propios asociados a valores propios distintos son linealmente independientes.
- Si $r = n$, entonces $\mathbb{K}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_n)$.

Demostración.

- $$V(\lambda_i) := \{v \in \mathbb{K}^n : Av^t = \lambda_i v\}$$

  es un subespacio vectorial de dimensión menor o igual que m_i (llamado el subespacio propio asociado a λ_i).
  Lamamos T al operador asociado a la matriz A cuando usamos la base canónica. Por supuesto, los autovalores y los subconjuntos $V(\lambda_i)$ dependen del operador T y son iguales para todas las matrices que representen a T. Como $V(\lambda_i) = \mathrm{Ker}\,(T - \lambda_i I)$, entonces $V(\lambda_i)$ es un subespacio vectorial. Supongamos que $r := \dim V(\lambda_i)$ $(r \geq 1)$. Entonces podemos elegir una base $\{x_1, \cdots, x_r\}$ del subespacio $V(\lambda_i)$. Ampliamos hasta conseguir una base $\{x_1, \cdots, x_n\}$ de $\mathbb{K}^n$. En términos de esta nueva base, la matriz asociada al operador T tiene el siguiente aspecto

  $$B := \begin{pmatrix} \lambda_i I_r & C \\ 0_{n-r\times r} & D \end{pmatrix}$$

  siendo $C \in \mathcal{M}_{r\times n-r}(\mathbb{K})$ y $D \in \mathcal{M}_{n-r}(\mathbb{K})$.

Propiedades de los vectores propios

Demostración.

$$B := \begin{pmatrix} \lambda_i I_r & C \\ 0_{n-r\times r} & D \end{pmatrix}$$

Por tanto, el polinomio característico de T viene dado por

$$\det(B-\lambda I) = \det \begin{pmatrix} (\lambda_i-\lambda)I_r & C \\ 0_{n-r\times r} & D-\lambda I_{n-r} \end{pmatrix}$$

Desarrollando el determinante por la primera columna (y repitiendo de nuevo este procedimiento si $r \geq 2$) obtenemos

$$\det(B-\lambda I) = (\lambda_i-\lambda)^r \det(D-\lambda I_{n-r}),$$

por lo que la multiplicidad del valor propio λ_i es, como mínimo r, esto es, $m_i \geq r = \dim V(\lambda_i)$.

Demostración.

- Los vectores propios asociados a valores propios distintos son linealmente independientes.
  Sean x e y vectores propios de T (aplicación lineal asociada a A) para distintos valores propios de T λ y μ y supongamos que para ciertos escalares a y b se tiene
  $$0 = ax + by \ ,$$
  Aplicando el operador T, de la igualdad anterior deducimos que
  $$0 = \lambda ax + \lambda by \ , \quad 0 = a\lambda x + b\mu y \ ,$$
  y restando las dos ecuaciones obtenemos que $b(\lambda - \mu)y = 0$. Como $\lambda \neq \mu, y \neq 0$, entonces ha de ser $b = 0$ y volviendo a la igualdad original obtenemos que $a = 0$, luego x e y son linealmente independientes.

  Ahora usaremos el mismo argumento para probar que esto ocurre no sólo para dos vectores propios sino para cualquier conjunto de vectores propios asociados a valores propios diferentes.

Propiedades de los vectores propios

Sean $\mu_1, \cdots \mu_r$ valores propios distintos de T y, para cada $1 \leq i \leq r$, $x_i \in \mathbb{K}^n$ un vector propio de T asociado a μ_i. Supongamos que para algunos escalares $a_1, \cdots, a_r$ se verifica que

$$0 = \sum_{i=1}^{r} a_i x_i \ .$$

Aplicando el operador T obtenemos que

$$0 = \sum_{i=1}^{r} \mu_i a_i x_i \ \text{ y como además } \ 0 = \sum_{i=1}^{r} \mu_1 a_i x_i \ .$$

Restando ambas expresiones obtenemos que

$$0 = \sum_{i=2}^{r} (\mu_i - \mu_1) a_i x_i = 0 \ .$$

Propiedades de los vectores propios

Multiplicando por μ_2 por una parte, y aplicando de nuevo el operador T obtenemos

$$0 = \sum_{i=2}^{r} (\mu_i - \mu_1)\mu_2 a_i x_i \quad 0 = \sum_{i=2}^{r} (\mu_i - \mu_1)\mu_i a_i x_i \ .$$

Restando de nuevo se obtiene

$$0 = \sum_{i=3}^{r} (\mu_i - \mu_1)(\mu_i - \mu_2) a_i x_i \ .$$

Repitiendo el procedimiento obtenemos

$$0 = \sum_{i=k}^{r} (\mu_i - \mu_1)(\mu_i - \mu_2) \cdots (\mu_i - \mu_{k-1}) a_i x_i = 0 \ , \ \ \forall 2 \leq k \leq r \ ,$$

y haciendo $k = r$ tenemos

$$(\mu_r - \mu_1)(\mu_r - \mu_2) \cdots (\mu_r - \mu_{r-1}) a_r x_r = 0 \ ,$$

y como los valores propios son todos distintos y los vectores propios son no nulos, deducimos que $a_r = 0$.

Volviendo a la expresión que habíamos conseguido en el paso para $k = r - 1$, que es

$$0 = (\mu_{r-1} - \mu_1)(\mu_{r-1} - \mu_2)\cdots(\mu_{r-1} - \mu_{r-2})a_{r-1}x_{r-1}+$$

$$(\mu_r - \mu_1)(\mu_r - \mu_2)\cdots(\mu_r - \mu_{r-2})a_r x_r =$$

$$(\mu_{r-1} - \mu_1)(\mu_{r-1} - \mu_2)\cdots(\mu_{r-1} - \mu_{r-2})a_{r-1}x_{r-1} \ ,$$

obtenemos que $a_{r-1} = 0$. Repitiendo el procedimiento y usando la expresión

$$0 = \sum_{i=k}^{r}(\mu_i - \mu_1)(\mu_i - \mu_2)\cdots(\mu_i - \mu_{k-1})a_i x_i = 0 \ , \ \ \forall 2 \leq k \leq r \ ,$$

obtenemos que $a_i = 0$ para todo $1 \leq i \leq r$, luego $x_1, \cdots, x_r$ son linealmente independientes.

Propiedades de los vectores propios

- Si $r = n$, entonces $\mathbb{K}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_n)$.

  Si $r = n$, como $\dim V(\lambda_i) \geq 1$ para dada $1 \leq i \leq n$, del apartado anterior deducimos que la dimensión del subespacio vectorial generado por $\{V(\lambda_i) : 1 \leq i \leq n\}$ es n, por tanto coincide con $\mathbb{K}^n$. Usando, de nuevo el apartado anterior, obtenemos que

$$\mathbb{K}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_n) \, .$$

Ejemplo

La matriz

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

tiene por polinomio característico a

$$P(\lambda) = \det(A - \lambda I_5) = (1-\lambda)(-2-\lambda)(2-\lambda)\Big((4-\lambda)(-5-\lambda)+18\Big) =$$

$$(1-\lambda)(-2-\lambda)(2-\lambda)\Big(\lambda^2+\lambda-2\Big) =$$

$$(1-\lambda)(-2-\lambda)(2-\lambda)(\lambda-1)(\lambda+2) = -(\lambda-1)^2(\lambda+2)^2(\lambda-2)\ .$$

Por tanto, los valores propios son $\lambda_1 = 1, \lambda_2 = 2, \lambda_3 = -2$ con multiplicidades son $m_1 = 2, m_2 = 1$ y $m_3 = 2$, respectivamente, luego estamos en las hipótesis del resultado anterior.

Propiedades de los vectores propios

El subespacio $\mathrm{Ker}\,(T - \lambda I)$ viene dado por las ecuaciones

$$x_1 = \lambda x_1, \;\; -2x_2 = \lambda x_2, \;\; x_1 - 5x_3 + 6x_4 = \lambda x_3, \;\; x_1 - 3x_3 + 4x_4 = \lambda x_4, \;\; 2x_5 = \lambda x_5 \,.$$

Resolviendo los sistemas anteriores para los tres valores propios, obtenemos la siguientes descripciones.

$$V(1) = \mathrm{Ker}\,(T - I) = \{(0, 0, x, x, 0) : x \in \mathbb{R}\}\,.$$

Se verifica que $\dim V(\lambda_1) = \dim V(1) = 1 < 2 = m_1$.

$$V(2) = \mathrm{Ker}\,(T - 2I) = \{(0, 0, 0, 0, x) : x \in \mathbb{R}\}\,.$$

En este caso $\dim V(\lambda_2) = \dim V(2) = 1 = m_2$.

$$V(-2) = \{(0, y, 2x, x, 0) : x, y \in \mathbb{R}\}\,.$$

Para el valor propio -2 se da la igualdad $\dim V(\lambda_3) = \dim V(-2) = 2 = m_3$.

Proposición

Si A es una matriz triangular de orden n, entonces los valores propios de A son los elementos de la diagonal principal.

Demostración.
Si A es triangular superior, entonces $A - \lambda I_n$ también lo es, luego

$$\det(A - \lambda I_n) = \prod_{i=1}^{n} (a_{ii} - \lambda) \; ,$$

luego los valores propios son $a_{11}, a_{22}, \cdots, a_{nn}$.

Producto escalar

Definición

Si $x, z \in \mathbb{C}^n$, definimos el producto escalar de x y z como sigue:

$$(x|z) = \sum_{i=1}^{n} x(i)\overline{z(i)} \,.$$

Se dice que dos vectores x, z son ortogonales si $(x|z) = 0$. Un subconjunto A de $\mathbb{C}^n$ es ortonormal si $\|a\|_2 = 1$ para cada $a \in A$ y cualesquiera dos elementos distintos de A son ortogonales (por definición $\|x\| = \sqrt{(x|x)}$, para un vector x de $\mathbb{C}^n$).

Producto escalar

Proposición

Se verifica que

- El producto escalar en $\mathbb{C}^n$ es lineal en la primera variable y conjugado lineal en la segunda, esto es,

$$(x+\alpha y|z) = (x|z)+\alpha(y|z), \quad (z|x+\alpha y) = (z|x)+\overline{\alpha}(x|y), \quad \forall x, y, z \in \mathbb{C}^n, \alpha \in$$

- La restricción del producto escalar a $\mathbb{R}^n$ es bilineal (lineal en cada variable).
- Se verifica que
$$(y|x) = \overline{(x|y)}, \quad \forall x, y \in \mathbb{C}^n .$$
- Si A es una matriz cuadrada de orden n con coeficientes en $\mathbb{C}$, se verifica que
$$(Ax^t|y) = (x|\overline{A}^t y^t), \quad \forall x, y \in \mathbb{C}^n .$$

Demostración.

Sólo probamos la última afirmación. Para ello, en vista de la primera propiedad, basta comprobar la igualdad para vectores de la base canónica de $\mathbb{C}^n$. En efecto, si $1 \leq i, j \leq n$ tenemos que

$$(Ae_i^t|e_j) = ((a_{1i}, a_{2i}, \cdots, a_{ni})|e_j) = a_{ji} \ ,$$

mientras que

$$(e_i|\overline{A}^t e_j^t) = (e_i|(\overline{a_{j1}}, \overline{a_{j2}}, \cdots, \overline{a_{jn}})) = a_{ji} \ ,$$

por tanto $(Ae_i^t|e_j) = a_{ji} = (e_i|\overline{A}^t e_j^t)$ y, se tiene como consecuencia que $(Ax^t|y) = (x|\overline{A}^t y^t)$ para cualesquiera vectores $x, y \in \mathbb{C}^n$.

Procedimiento de Gram-Schmidt

Proposición

Todo subespacio M de $\mathbb{R}^n$ (ó de $\mathbb{C}^n$) tiene una base ortonormal. De hecho, cualquier subconjunto ortonormal de M se puede ampliar a una base ortonormal.
Además, si $\{u_1, u_2, \cdots, u_r\}$ es una base de M, se puede conseguir una base ortonormal $\{m_1, m_2, \cdots, m_r\}$ de M tal que $L(u_1, \cdots, u_k) = L(m_1, \cdots, m_k)$ para cada $k \leq r$.

Demostración.
Sólo probamos la última afirmación. La primera es una consecuencia directa de esta afirmación y para probar la segunda basta aplicar el mismo procedimiento que usaremos ahora par ampliar un conjunto ortonormal a una base. Probamos la afirmación usando inducción sobre la dimensión del subespacio M. Si $r = 1$ y $\{u_1\}$ es una base de M, entonces $\left\{\frac{u_1}{\|u_1\|_2}\right\}$ es una base ortonormal de M.

Procedimiento de Gram-Schmidt

Supongamos que es cierto para subespacios de dimensión r y supongamos que M es un subespacio de $\mathbb{R}^n$ de dimensión $r+1$. Aplicando la hipótesis de inducción al subespacio generado por $\{u_1, \cdots, u_r\}$, sabemos que éste tiene una base ortonormal $\{m_1, \cdots, m_r\}$ que verifica que $L(u_1, \cdots, u_k) = L(m_1, \cdots, m_k)$ para cada $k \leq r$. Definimos ahora el vector

$$m := u_{r+1} - \sum_{i=1}^{r} (u_{r+1}|m_i) m_i$$

que está en M por estarlo u_{r+1} y por la hipótesis de inducción es no nulo, por ser los vectores $\{u_i : 1 \leq i \leq r+1\}$ una base de M. Comprobamos que m es ortogonal a los vectores m_j $(1 \leq j \leq r)$, ya que

$$(m|m_j) = (u_{r+1}|m_j) - \sum_{i=1}^{r} (u_{r+1}|m_i)(m_i|m_j) =$$

$$(u_{r+1}|m_j) - (u_{r+1}|m_j)(m_j|m_j) = 0 \ .$$

Tomamos $m_{r+1} = \frac{m}{\|m\|_2}$. Es inmediato que $\{m_1, \cdots, m_{r+1}\}$ es una base ortonormal de M que verifica el enunciado.

Proposición

Si A es una matriz simétrica cuyos elementos son números reales. Se verifica que:

- Todas las soluciones de la ecuación característica son números reales.
- Dos valores propios distintos de A tienen asociados vectores propios ortogonales.
- Para cada autovalor de A, la dimensión del subespacio propio asociado coincide con la multiplicidad de ese valor propio.

Demostración.

- Todas las soluciones de la ecuación característica son números reales. Supongamos que $\lambda \in \mathbb{C}$ verifica que $\det(A - \lambda I_n) = 0$ y $x \in \mathbb{C}^n$ un vector propio asociado, entonces por ser x vector propio, se verifica que

$$(Ax^t|x) = (\lambda x|x) = \lambda\|x\|_2^2 \ ;$$

además sabemos que

$$(Ax^t|x) = (x|\overline{A}^t x^t) = (x|Ax^t) = \overline{(Ax^t|x)}$$

por tanto $(Ax^t|x) = \lambda\|x\|_2^2$ es real y como x es no nulo, entonces $\|x\|_2^2 > 0$, luego λ es también real.

Tipos de matrices y valores propios

- Dos valores propios distintos de A tienen asociados vectores propios ortogonales.

  Sean s, t valores propios distintos y x, y vectores propios asociados a s y t, respectivamente. Se verifica entonces que

  $$(Ax^t|y) = (sx|y) = s(x|y) \; .$$

  Usando que A es simétrica obtenemos esta otra igualdad

  $$(Ax^t|y) = (x|Ay^t) = (x|ty) =$$

  $$\bar{t}(x|y) = t(x|y) \; .$$

  Hemos obtenido que $s(x|y) = t(x|y)$ y como $s \neq t$, entonces $(x|y)$ ha de ser cero, luego x e y son ortogonales.

- Para cada autovalor de A, la dimensión del subespacio propio asociado coincide con la multiplicidad de ese valor propio.
  Por ser A simétrica, la ecuación característica tiene todas sus soluciones reales. Sean éstas $s_1, \cdots, s_r$ con multiplicidades $m_1, \cdots, m_r$. Sabemos que $\sum_{i=1}^{r} m_i = n$, ya que el polinomio característico tiene grado n y se puede escribir como $\prod_{i=1}^{r}(s_i - \lambda)^{m_i}$. Consideramos el subespacio

  $$M = V(s_1) \oplus V(s_2) \oplus \cdots \oplus V(s_r),$$

  donde $V(s_i)$ es el subespacio propio asociado a s_i para cada i. Sabemos que los subespacios $V(s_i)$ son ortogonales dos a dos y que $\dim V(s_i) \leq m_i$ para cada i. Basta comprobar que M coincide con $\mathbb{R}^n$ ya que entonces se tendrá que $n = \sum_{i=1}^{n} m_i \geq \sum_{i=1}^{n} \dim V(s_i) = n$, luego $\dim V(s_i) = m_i$ para cada i.

Propiedades de los vectores propios

Supongamos que M no es $\mathbb{R}^n$, entonces el subespacio $M^\perp := \{x \in \mathbb{R}^n : (x|m) = 0\}$ no es nulo. Veamos que el operador asociado a A conserva el subespacio $M^\perp$. En efecto, sea $z \in M^\perp$, entonces, para $m \in M$, tenemos que

$$(Az^t|m) = (z|\overline{A}^t m^t) = (z|Am^t) = 0,$$

ya que $Am^t \in M$, luego el operador asociado a A conserva el subespacio $M^\perp$. Como la restricción de T a $M^\perp$ viene dado por una matriz simétrica, ha de tener algún valor propio y algún vector propio v asociado en $M^\perp$, pero entonces $v \in M$, ya que M contiene a todos los vectores propios de A. Ahora bien $M \cap M^\perp = \{0\}$, luego $v = 0$, lo cual es una contradicción.
Por tanto, $M = \mathbb{R}^n$ y $m_i = \dim V(s_i)$ para cada i.

Teorema de Perron-Frobenius

Sean A y B dos matrices de orden n no nulas y supongamos que $a_{ij} > 0, b_{ij} \geq 0$ para todo i, j. Entonces

- A tiene un valor propio real y positivo mayor que el módulo de todos los demás y tal que tiene un vector propio asociado con todas sus componentes positivas.
- B tiene un valor propio real no negativo mayor que el módulo de todos los demás y tal que tiene un vector propio asociado con todas sus componentes no negativas.
- Si alguna potencia de B es una matriz estrictamente positiva, entonces B tiene un autovalor positivo y un vector propio asociado con todas sus componentes positivas.

Si $x \in \mathbb{R}^n$, escribiremos $x \geq 0$ cuando $x_i \geq 0$ para todo $1 \leq i \leq n$. $x > 0$ significa que $x \geq 0$ y $x \neq 0$.

Demostración.
Probaremos únicamente la primera afirmación. Para ello consideramos el conjunto

$$C := \{t \in \mathbb{R} : \text{existe } x > 0, Ax^t \geq tx\} .$$

Probaremos que C contiene algún positivo. Sea $x \in \mathbb{R}^n, x_i > 0\ \forall i$, entonces $y = Ax^t$ verifica que $y_i > 0$ para todo i. Entonces $y_i = \varepsilon_i x_i$ para conveniente $\varepsilon_i > 0$; basta tomar $\varepsilon = \min_i\{\varepsilon_i\}$ y entonces $Ax^t = y \geq \varepsilon x$, luego $\varepsilon \in C$.
Como las aplicaciones $x \mapsto Ax, x \mapsto tx$ son lineales en $\mathbb{R}^n$, en la definición de C podemos imponer que $\sum_i x_i = 1$.
Probaremos que C es cerrado. Si $t_m \in C$ para cada m y $\{t_m\} \to t$, para cada m existe $x_m > 0$ tal que $\sum_i x_m(i) = 1$ y $Ax_m^t \geq t_m x_m$ para cada m. Pasando a una subsucesión podemos suponer que $\{x_m\} \to x$. Tomando límite ($m \to \infty$) en la expresión $\sum_{i=1}^n x_m(i) = 1$, obtenemos que $\sum_{i=1}^n x(i) = 1$ y es claro que $x(i) \geq 0$ para cada i. Por ser $Ax_m^t \geq t_m x_m$ obtenemos que $Ax^t \geq tx$, luego $t \in C$ y hemos probado que C es cerrado.

Tipos de matrices y valores propios

Además C está mayorado. Sea $t \in C$, x el vector asociado con $\sum_{i=1}^{n} x(i) = 1$. Sumando las coordenadas tenemos que

$$t = \sum_i (tx)(i) \leq \sum_i (Ax^t)(i) =$$

$$\sum_{i=1}^{n} \Big(\sum_{j=1}^{n} (a_{ij} x(j) \Big) = \sum_{j=1}^{n} x(j) \sum_{i=1}^{n} a_{ij} =$$

$$\sum_j \sum_{i=1}^{n} a_{ij} := M \; .$$

Luego C está mayorado (por M).
Como C es cerrado y mayorado, tiene máximo. Sea λ_0 el máximo de C.

Nótese que por ser $a_{ij} > 0 \forall i, j$, entonces $Ax^t(i) > 0$ para cada i si $x \geq 0, x \neq 0$.

Probaremos que λ_0 es el autovalor que buscamos. Por ser $\lambda_0 \in C$, existe $x > 0$ tal que $Ax^t \geq \lambda_0 x$. Veremos que $Ax^t = \lambda_0 x$. Si $Ax^t \neq \lambda_0 x$, entonces $y = Ax^t > 0$, por tanto

$$0 < (A(Ax^t - \lambda_0 x)^t)(j) \forall j \ ,$$

esto es

$$(Ay^t)(j) > \lambda_0 y(j) \forall j \ .$$

Como en cada desigualdad se da la desigualdad estricta, existe $\varepsilon > 0$ tal que

$$Ay^t > (\lambda_0 + \varepsilon) y \ ,$$

luego $\lambda_0 + \varepsilon \in C$, pero esto es imposible, ya que $\lambda_0 = \max C$.

Falta comprobar que λ_0 es mayor que el módulo del resto de los autovalores.

Sea λ otro autovalor de A (real o complejo) e y un vector propio asociado a y, denotamos por $|y| := (|y_1|, \cdots, |y_n|)$; se verifica que

$$A(|y|)^t \geq |Ay^t| = |\lambda y| = |\lambda||y| \ ,$$

luego $|\lambda| \in C$, por tanto $|\lambda| \leq \lambda_0$. De hecho, probaremos que si $\lambda \neq \lambda_0$, se verifica que $|\lambda| < \lambda_0$.
Por ser A estrictamente positiva, para $\delta > 0$ suficientemente pequeño, la matriz $A_\delta = A - \delta I_n$ también es estrictamente positiva. Sabemos que $\lambda - \delta$ y $\lambda_0 - \delta$ son autovalores de A_δ para los que se verifica, en virtud de la primera parte de la demostración que

$$|\lambda - \delta| \leq \lambda_0 - \delta \ . \tag{1}$$

Comprobaremos que la desigualdad anterior no se verifica si $|\lambda| = |\lambda_0|$ y $\lambda \neq \lambda_0$.
En tal caso, usando las propiedades del módulo obtenemos que

$$|\lambda - \delta| \geq ||\lambda| - \delta| = ||\lambda_0| - \delta| = \lambda_0 - \delta \ .$$

Por tanto, tendríamos que $|\lambda - \delta| = ||\lambda| - \delta| = \lambda_0 - \delta =$.

Si $\lambda \in \mathbb{R}$, como $\lambda_0 > 0$, habría de ser $\lambda < 0$ y en este caso

$$|\lambda - \delta| = -\lambda + \delta = \lambda_0 + \delta > \lambda_0 \ ,$$

lo que contradice la desigualdad (1).

Si $\lambda \in \mathbb{C}$, escribimos $\lambda = a + ib$, para $a, b \in \mathbb{R}$, entonces

$$|\lambda - \delta|^2 = (a - \delta)^2 + b^2 = a^2 + b^2 + \delta^2 - 2a\delta$$

mientras que

$$||\lambda| - \delta|^2 = a^2 + b^2 + \delta^2 - 2\sqrt{a^2 + b^2}\delta \ .$$

Como ambos módulos coinciden, entonces ha de ser

$$a = \sqrt{a^2 + b^2} \quad \Rightarrow \quad b = 0 \ ,$$

luego $\lambda \in \mathbb{R}$, y si λ es real, sabemos ya que es imposible.

Proposición

Sea $A \in \mathcal{M}_n$ una matriz positiva.

- A es estocástica y, y sólo si, 1 es un autovalor de A con vector propio asociado $v = (1, 1, \cdots, 1)$.
- A es doblemente estocástica si, y sólo si, 1 es un autovalor de A con vector propio asociado $v = (1, 1, \cdots, 1)$ y lo mismo le ocurre a A^t.

Demostración.
La segunda afirmación es consecuencia inmediata de la primera. Si A es estocástica, como estamos suponiendo que es positiva, esto significa que

$$\sum_{j=1}^{n} a_{ij} = 1, \quad \forall 1 \leq i \leq n \,.$$

La condición anterior equivale a que $Av^t = v$

Ejemplo

Consideramos la matriz A dada por

$$A := \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/4 & 1/4 & 1/2 \end{pmatrix} .$$

La ecuación característica es

$$\det(A - \lambda I_3) = \det \begin{pmatrix} \frac{1}{2} - \lambda & \frac{1}{2} & 0 \\ \frac{1}{3} & \frac{1}{3} - \lambda & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} - \lambda \end{pmatrix} =$$

$$(\frac{1}{2} - \lambda)\Big((\frac{1}{3} - \lambda)(\frac{1}{2} - \lambda) - \frac{1}{12}\Big) - \frac{1}{2}\Big(\frac{1}{3}(\frac{1}{2} - \lambda) - \frac{1}{12}\Big) =$$

$$-\frac{1}{3}\lambda\Big(1 - 4\lambda + 3\lambda^2\Big) = -\lambda(\lambda - 1)\Big(\lambda - \frac{1}{3}\Big) .$$

Tipos de matrices y autovalores

Proposición

Sea $A \in \mathcal{M}_n$ una matriz idempotente. Entonces

- Cada uno de sus autovalores es 0 ó 1.
- Si m_0 es la multiplicidad algebraica de 0 y m_1 la de 1, entonces se verifica que $m_0 = \dim V(0), m_1 = \dim V(1)$.
- $\mathrm{tr}\,(A) = \mathrm{rg}\,(A)$

Demostración.

- Cada uno de sus autovalores (reales ó complejos) es 0 ó 1.
  Si λ es un valor propio de A y x un vector propio asociado a λ y T el operador asociado a A, como $T^2 = T$, obtenemos que

$$\lambda x = T(x) = T^2(x) =$$

$$T(Tx) = T(\lambda x) = \lambda^2 x \ .$$

  Como $x \neq 0$ y $(\lambda - \lambda^2)x = 0$ ha de ser $\lambda = 0$ ó $\lambda = 1$.

Tipos de matrices y autovalores

- Si m_0 es la multiplicidad algebraica de 0 y m_1 la de 1, entonces se verifica que $m_0 = \dim V(0), m_1 = \dim V(1)$.
  Si 0 tiene multiplicidad algebraica m_0, entonces 1 tiene multiplicidad algebraica $n - m_0$, ya que el polinomio característico tiene grado n y en este caso se puede escribir como

  $$P(\lambda) = (-1)^n \lambda^{m_0} (\lambda - 1)^{m_1} \ .$$

  Sea Q el operador asociado a la matriz A en términos de la base canónica. Comprobaremos que

  $$\mathbb{K}^n = \mathrm{Ker}\ Q \oplus \mathrm{Im}\ Q = V(0) \oplus V(1) \ .$$

  Para ello, basta probar que $\mathbb{K}^n \subset \mathrm{Ker}\ Q + \mathrm{Im}\ Q$ y $\mathrm{Ker}\ Q = V(0), \mathrm{Im}\ Q = V(1)$, ya que sabemos que dos vectores propios asociados a valores propios distintos son independientes.
  Sea $x \in \mathbb{K}^n$, escribimos $x = (x - Q(x)) + Q(x)$ y se verifica que

  $$Q(x - Q(x)) = Q(x) - Q^2(x) = 0 \ ,$$

  por tanto, $x - Q(x) \in \mathrm{Ker}\ Q = V(0)$, $Q(x) \in \mathrm{Im}\ Q$. Es claro que $V(0) = \mathrm{Ker}\ Q$ y comprobaremos que $\mathrm{Im}\ Q = V(1)$.

En efecto, si $y \in \mathrm{Im}\, Q$, entonces existe $x \in \mathbb{K}^n$ con $Qx = y$, por tanto, por ser $Q^2 = Q$, tenemos

$$Q(y) = Q^2(x) = Q(x) = y \ ,$$

luego $y \in V(1)$. Recíprocamente, si $y \in V(1)$, tenemos $Q(y) = y$, luego $y \in \mathrm{Im}\, Q$.

Hemos probado que $\mathbb{K}^n = V(0) \oplus V(1)$ y sabemos además que $\dim V(0) \leq m_0, \dim V(1) \leq m_1 = n - m_0$. La igualdad anterior nos asegura que de hecho $\dim V(0) = m_0, \dim V(1) = m_1$ por ser

$$n = \dim \mathbb{K}^n = \dim V(0) + \dim V(1) \leq m_0 + m_1 = n \ .$$

- $\mathrm{tr}\,(A) = \mathrm{rg}\,(A)$
  Por el Teorema de la dimensión sabemos que

$$m_0 + m_1 = n = \dim \mathrm{Ker}\ Q + \dim \mathrm{Im}\ Q = \dim V(0) + \mathrm{rg}\ A =$$

$$m_0 + \mathrm{rg}\ A\ ,$$

por tanto $\mathrm{rg}\ (A) = m_1$.

Por otra parte, sabemos que cualquier matriz verifica que su traza coincide con la suma de sus valores propios, luego $\mathrm{tr}\ (A) = m_1 = \mathrm{rg}\ (A)$.

Definición

Dos matrices cuadradas A y B de orden n son semejantes si existe una matriz inversible P tal que $A = PBP^{-1}$.
A es diagonalizable si A es semejante a una matriz diagonal. Una matriz D de orden n es diagonal si $d_{ij} = 0, \forall 1 \leq i, j \leq n, i \neq j$.

Ejemplo

La matriz A dada por

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

es diagonalizable.

El polinomio característico de A viene dado por

$$\det(A-\lambda I_3) = \det\begin{pmatrix} 6-\lambda & -6 & 2 \\ -1 & -1-\lambda & 1 \\ 7 & 3 & 1-\lambda \end{pmatrix} =$$

$$-64+24t+6t^2-t^3 = -(t+4)(t-2)(t-8)$$

Como los tres valores propios son reales y distintos, cada subespacio propio asociado a cada uno de ellos tiene dimensión 1, luego A es diagonalizable. Para encontrar la matriz P que aparece en la definición hemos de encontrar un vector propio asociado a cada valor propio. En este caso tenemos que

$$V(-4) = \mathrm{L}\,\{(-1,-1,2)\}, \quad V(2) = \mathrm{L}\,\{(0,1,3)\}, \quad V(8) = \mathrm{L}\,\{(1,0,1)\}\,.$$

Por tanto, si A es la matriz asociada a T en términos de la base canónica de $\mathbb{R}^3$, en términos de la base $B' := \{(-1,-1,2),(0,1,3),(1,0,1)\}$, la matriz asociada a T es

$$D = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

y se tiene que $P^{-1}AP = D$ siendo

$$P = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & 3 & 1 \end{pmatrix} \quad \text{y} \quad P^{-1} = \frac{1}{6}\begin{pmatrix} -1 & -3 & 1 \\ -1 & 3 & 1 \\ 5 & -3 & 1 \end{pmatrix}$$

Diagonalización de matrices

Observación

- Sean B_1 y B_2 dos bases de $\mathbb{K}^n$ y $T \in L(\mathbb{K}^n)$. Si A es la matriz asociada a T para la base B_1 y C es la matriz asociada a T para B_2, entonces A y C son semejantes. En efecto, si P es la matriz que da el cambio de coordenadas, esto es, si $X_1 = P(X_2)^t$, donde X_1 y X_2 representan las coordenadas del mismo vector en términos de B_1 y B_2, respectivamente, entonces sabemos que P es inversible y se tiene que

  $$P^{-1}AP = C \ ,$$

  luego A y C son semejantes.

- Dos matrices semejantes tienen el mismo polinomio característico y los mismos valores propios.

Proposición

Sea A una matriz cuadrada de orden n y supongamos que A tiene n valores propios distintos $\lambda_1, \cdots, \lambda_n$. Entonces A es diagonalizable. Además, existe una matriz P cuyas columnas son vectores propios $v_1, \cdots, v_n$ asociados a los valores propios $\lambda_1, \cdots, \lambda_n$ tal que

$$D = P^{-1}AP\,,$$

donde

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

La matriz D es única, salvo reordenación de los valores propios $\lambda_1, \cdots, \lambda_n$.

Proposición

Sea A una matriz cuadrada de orden n. Entonces A es diagonalizable si, y sólo si, existe una base de vectores propios del espacio en que está definida la aplicación lineal T que tiene por matriz asociada A respecto de las bases canónicas.

Como consecuencia, A es diagonalizable si, y sólo si, para cada valor propio λ_i de A, se verifica que la dimensión del subespacio propio asociado a λ_i coincide con la multiplicidad de λ_i.

Corolario

- Toda matriz simétrica con coeficientes reales es diagonalizable y sus autovalores son reales. De hecho, si $A \in \mathcal{M}_n$ es simétrica, y sus coef. son reales, existen una matriz ortogonal P y una diagonal D tales que $A = P^{-1}DP$.
- Toda matriz idempotente es diagonalizable.

Ejemplo

La matriz $A = \begin{pmatrix} 4 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 4 \end{pmatrix}$ tiene coef. reales y es simétrica, luego es diagonalizable.

Para diagonalizarla, calculamos los valores propios.

$$\det(A - \lambda I_3) = \det \begin{pmatrix} 4-\lambda & 0 & -1 \\ 0 & 3-\lambda & 0 \\ -1 & 0 & 4-\lambda \end{pmatrix} =$$

$$(4-\lambda)^2(3-\lambda) - (3-\lambda) = (3-\lambda)\Big((4-\lambda)^2 - 1\Big) =$$

$$= -(\lambda - 3)^2(\lambda - 5)\ .$$

Por tanto, los valores propios son $\lambda_1 = 3$, $m_1 = 2$, $\lambda_2 = 5, m_2 = 1$.

Para dar la matriz del cambio de base, describimos los subespacios propios:

$$V(3) = \{X \in \mathbb{R}^3 : AX^t = 3X^t\} =$$

$$\{(x, y, x) : x, y \in \mathbb{R}\}$$

Por tanto, $\{(0, 1, 0), (1, 0, 1)\}$ es una base de $V(3)$ ($\dim V(3) = m_1 = 2$).
Obtenemos también que

$$V(5) = \{X \in \mathbb{R}^3 : AX^t = 5X^t\} =$$

$$\{(x, 0, -x) : x \in \mathbb{R}\} ,$$

luego $(1, 0, -1)$ es una base de $V(5)$.

Diagonalización de matrices

La matriz M que da el cambio de base (de la formada por los vectores propios a la base canónica) es la matriz M cuyas columnas son los tres vectores propios que generan $V(3)$ y $V(5)$. La matriz diagonal D semejante a A tiene en la diagonal principal los valores propios de A, con cada valor propio repetido tantas veces como indique su multiplicidad. Nótese, además, que si primero colocamos en M los vectores que generan $V(3)$, en D primero colocamos la submatriz $3I_2$. En resumen, en este caso, tenemos

$$M = \left(\begin{array}{cc|c} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{array}\right), \quad D = \left(\begin{array}{cc|c} 3 & 0 & 0 \\ 0 & 3 & 0 \\ \hline 0 & 0 & 5 \end{array}\right)$$

Por último, se puede comprobar que $A = MDM^{-1}$. En este caso,
$M^{-1} = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & -1 \end{pmatrix}$

Observación

Si D es una matriz diagonal de la forma

$$D = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix},$$

es inmediato comprobar que para cada natural k se verifica que

$$D^k = \begin{pmatrix} a_1^k & 0 & \cdots & 0 \\ 0 & a_2^k & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n^k \end{pmatrix}.$$

Diagonalización de matrices

Observación

Si A es una matriz diagonalizable, existe una matriz inversible M y una matriz diagonal D tal que $A = MDM^{-1}$, por tanto

$$A^2 = (MDM^{-1})(MDM^{-1}) = MD^2M^{-1} .$$

Usando la expresión anterior de A^2, obtenemos que

$$A^3 = A^2A = (MD^2M^{-1})(MDM^{-1}) = MD^3M^{-1} .$$

Una sencilla inducción permite comprobar que

$$A^k = MD^kM^{-1}, \quad \forall k \in \mathbb{N}.$$

Ejemplo

Consideramos la matriz $A = \begin{pmatrix} 2 & 1 \\ -2 & 0 \end{pmatrix}$ cuyo polinomio característico es

$$\det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ -2 & -\lambda \end{pmatrix} = -\lambda(2-\lambda) + 2 = \lambda^2 - 2\lambda + 2$$

cuyos ceros son $1+i$ y $1-i$. Por tanto, A es diagonalizable, pero la matriz diagonal y los vectores propios asociados a los valores propios no están en $\mathbb{R}^2$, sino en $\mathbb{C}^2$. Los vectores $(1+i, -2)$ y $(1-i, -2)$ son vectores propios asociados a $1+i$ y $1-i$, respectivamente. En $\mathbb{C}^2$ la matriz P es, en este caso,

$$P = \begin{pmatrix} 1+i & 1-i \\ -2 & -2 \end{pmatrix}$$

La matriz $B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ es también semejante a A y se verifica que $MBM^{-1} = A$, siendo

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

que se obtiene al tomar parte real e imaginaria de las componentes del vector propio $(1+i, -2)$. Este procedimiento se puede usar en matrices de orden mayor, cada vez que haya un valor propio que sea un complejo no real. La matriz B se puede escribir también en términos del valor propio $\lambda = 1 + i$ como sigue

$$B = B = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = B = \begin{pmatrix} \mathrm{Re}\,\lambda & \mathrm{Im}\,\lambda \\ -\mathrm{Im}\,\lambda & \mathrm{Re}\,\lambda \end{pmatrix}$$

Proposición

Sea A una matriz cuadrada de orden 2 con coeficientes reales y supongamos que $\lambda = a + ib$ es un autovalor de A con a y b reales y $b \neq 0$. Entonces si $w = x + iy$, para $x, y \in \mathbb{R}^2$ es un vector propio asociado a λ, se verifica que la matriz A es semejante a B, siendo

$$B = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} .$$

De hecho $B = M^{-1}AM$ para la matriz

$$M = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = \begin{pmatrix} \mathrm{Re}\; w_1 & \mathrm{Im}\; w_1 \\ \mathrm{Re}\; w_2 & \mathrm{Im}\; w_2 \end{pmatrix} .$$

Demostración.

Comprobaremos en primer lugar que $M = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix}$ es una matriz inversible. Si no lo fuese, los vectores (x_1, x_2) e (y_1, y_2) son linealmente dependientes. Supongamos, por ejemplo, que $(x_1, x_2) = t(y_1, y_2)$ para algún real t. Entonces tenemos que

$$w = x + iy = (ty_1 + iy_1, ty_2 + iy_2) = (t + i)(y_1, y_2),$$

mientras que

$$\overline{w} = x - iy = (ty_1 - iy_1, ty_2 - iy_2) = (t - i)(y_1, y_2),$$

por tanto, w y $\overline{w}$ son linealmente dependientes.
Pero esto es imposible, ya que w es el vector propio de A asociado a λ, esto es, $Aw^t = \lambda w$.
Como A tiene coeficientes reales, sabemos que

$$A\overline{w^t} = \overline{Aw^t} = \overline{\lambda w} = \overline{\lambda}\overline{w}$$

luego $\overline{\lambda}$ es un valor propio de A y $\overline{w}$ un vector propio asociado.
Por hipótesis $\lambda \notin \mathbb{R}$, luego $\lambda \neq \overline{\lambda}$, por tanto w y $\overline{w}$ son linealmente independientes. Luego M es inversible.

La igualdad $B = M^{-1}AM$ equivale a $MB = AM$.
Sabemos que $Aw^t = \lambda w$. Escribiendo w en términos de x e y, obtenemos

$$Aw^t = Ax^t + iAy^t,$$

$$\lambda w = (a + ib)(x + iy) = ax - by + i(bx + ay)\ .$$

Igualando partes real e imaginaria de sus componentes deducimos

$$Ax^t = ax^t - by^t, \quad Ay^t = bx^t + ay^t\ .$$

Por tanto,

$$AM = A\begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = (Ax^t, Ay^t) = (ax^t - by^t, bx^t + ay^t)\ .$$

Hemos probado la igualdad

$$AM = (ax^t - by^t, bx^t + ay^t) \ .$$

Por otra parte, tenemos que

$$(ax^t - by^t, bx^t + ay^t) = \begin{pmatrix} ax_1 - by_1 & bx_1 + ay_1 \\ ax_2 - by_2 & bx_2 + ay_2 \end{pmatrix} = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} =$$

$$MB$$

Por tanto, $AM = MB$.

Forma canónica de Jordan

En caso de que las matrices no sean diagonalizables, son semejantes a matrices “casi diagonales”. Así aparecen las llamadas “matrices de Jordan”.

Definición

Una matriz cuadrada B de orden n con coef. en $\mathbb{K}$ es un bloque elemental de Jordan si existe $b \in \mathbb{K}$ tal que

$$B = \begin{pmatrix} b & 1 & 0 & \cdots & 0 & 0 \\ 0 & b & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & b & 1 \\ 0 & 0 & 0 & \cdots & 0 & b \end{pmatrix}.$$

Esto es,

$$b_{ij} = \begin{cases} b & \text{si } \; i = j \\ 1 & \text{si } \; i+1 = j \\ 0 & \text{en otro caso} \end{cases}$$

Forma canónica de Jordan

Una generalización de los bloques elementales de Jordan es el tipo de matrices que definimos ahora.

Definición

Una matriz cuadrada B de orden n con coef. en $\mathbb{K}$ es una matriz de Jordan si es diagonal por bloques, y cada bloque es un bloque elemental de Jordan, esto es, si existen bloques de Jordan $J_1, J_2, \cdots J_k$, tales que

$$B = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}.$$

Cuando no se puede diagonalizar un operador, sí es cierto que la matriz asociada es semejante a una matriz de Jordan. Por supuesto, para conocer la matriz semejante y la del cambio de base, es necesario averiguar:

- El número de bloques de la matriz de Jordan
- El orden de cada uno de los bloques.
- Base respecto de la cual la nueva matriz asociada al operador es una matriz de Jordan.

Ejemplo

El ejemplo más sencillo de matriz de Jordan es un bloque de Jordan en el que la diagonal principal sea cero, como

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

Identificando A con la aplicación lineal T asociada en $\mathbb{R}^3$ tenemos que

$$T(e_1) = 0, \quad T(e_2) = e_1, \quad T(e_3) = e_2 .$$

Por tanto

$$T^2(e_1) = 0, \quad T^2(e_2) = T(e_1) = 0, \quad T^2(e_3) = T(e_2) = e_1 ,$$

luego

$$T^3(e_1) = 0, \quad T^3(e_2) = 0, \quad T^3(e_3) = T(e_1) = 0 ,$$

esto es, $T^3 = 0$, y, por supuesto, $A^3 = 0$ y 3 es el primer natural que verifica esa condición. En este caso, diremos que T (ó A) es nilpotente de orden 3.

Ejemplo

La matriz

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

es nilpotente de orden 2, esto es, $A \neq 0, A^2 = 0$.

Definición

Sea X un espacio vectorial sobre K de dimensión finita y $T \in L(X)$; diremos que T es nilpotente de orden p si $T^p = 0$ y p es el primer natural que verifica esa condición.

Ejemplo

La matriz de orden n

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

es nilpotente de orden n.

Forma canónica de Jordan

El argumento para comprobar la afirmación anterior es, esencialmente, el usado antes para $n = 3$. Si T es el operador asociado a A, sabemos que se verifica

$$T(e_1) = 0, \quad T(e_2) = e_1, \quad T(e_3) = e_2, \cdots, T(e_n) = e_{n-1} \ .$$

Aplicando de nuevo T tenemos

$$T^2(e_1) = 0, \quad T^2(e_2) = 0, \quad T^2(e_3) = e_1, \cdots, T^2(e_n) = e_{n-2} \ .$$

Repitiendo el procedimiento k veces ($k \leq n-1$), se tiene que

$$T^k(e_1) = 0, \quad T^k(e_2) = 0, \cdots, T^k(e_k) = 0, \quad T^k(e_{k+1}) = e_1, \cdots, T^k(e_n) = e_{n-k}$$

Por tanto,

$$T^{n-1}(e_1) = 0, \quad T^{n-1}(e_2) = 0, \quad T^{n-1}(e_{n-1}) = 0, \quad T^{n-1}(e_n) = e_1 \ .$$

Así, $T^n = 0$ y n es el primer natural que verifica esa condición.

Observación

Si $T \in L(\mathbb{K}^n)$ es nilpotente de orden p y $v \in \mathbb{K}^n$ verifica que $T^{p-1}(v) \neq 0$, entonces los vectores

$$\{v, T(v), T^2(v), \cdots, T^{p-1}(v)\}$$

son linealmente independientes (luego $p \leq n$).

Demostración.
Lo probaremos para $p = 3$. El mismo argumento funciona en general.

Supongamos que T es nilpotente de orden 3 y que $T^2(v) \neq 0$. Si a, b, c son escalares tales que

$$0 = av + bT(v) + cT^2(v) \ . \tag{2}$$

Aplicando T obtenemos

$$0 = aT(v) + bT^2(v) + cT^3(v) = aT(v) + bT^2(v) \ . \tag{3}$$

Y aplicando de nuevo T queda

$$0 = aT^2(v) + bT^3(v) = aT^2(v) \ .$$

Como $T^2(v) \neq 0$, ha de ser $a = 0$. De la igualdad (3) deducimos entonces $b = 0$ y de (2) $c = 0$, luego $\{v, T(v), T^2(v)\}$ son linealmente independientes.

Teorema

Sea $A \in \mathcal{M}_n$ una matriz nilpotente de orden p, entonces existen bloques de Jordan $J_1, J_2, \cdots, J_k$ tal que A es semejante a la matriz de Jordan

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix}.$$

Además cada matriz J_i es de la forma

$$J_i = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$$

Ejemplo

La matriz A dada por

$$A = \begin{pmatrix} -46 & -3 & 26 & -48 \\ 70 & 4 & -39 & 76 \\ -40 & -3 & 23 & -41 \\ 18 & 1 & -10 & 19 \end{pmatrix}.$$

es nilpotente de orden $p = 4$, ya que $A^4 = 0$ y $A^3 \neq 0$.
De hecho

$$A^2 = \begin{pmatrix} 2 & 0 & -1 & 2 \\ -12 & -1 & 7 & -13 \\ -28 & -2 & 16 & -30 \\ -16 & -1 & 9 & -17 \end{pmatrix} \quad A^3 = \begin{pmatrix} -16 & -1 & 9 & -17 \\ -32 & -2 & 18 & -34 \\ -32 & -2 & 18 & -34 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

$A^4 = 0$

En este caso concreto, basta elegir un elemento v de $\mathbb{R}^4$ que no pertenezca al núcleo de T^3, donde T es el operador asociado a A para que el conjunto

$$\{T^3(v), T^2(v), T(v), v\}$$

sea una base de $\mathbb{R}^4$ y es claro que para esta base la matriz asociada a T es

$$J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

que es un bloque de Jordan.

Forma canónica de Jordan

Calculamos $K_3 = \mathrm{Ker}\ T^3$

$\mathrm{Ker}\ T^3 := \{(x, y, u, v) \in \mathbb{R}^4 : -16x - y + 9u - 17v = 0, -32x - 2y + 18u - 34v = 0\}$

Por ejemplo, $e_2 \notin \mathrm{Ker}\ T^3$, por tanto, aplicando T tenemos

$T(e_2) = (-3, 4, -3, 1), \quad T^2(e_1) = (0, -1, -2, -1), \quad T^3(e_2) = (-1, -2, -2, 0)$

Así, la matriz que da el cambio de base es

$$M = \begin{pmatrix} T^3(e_2) & T^2(e_2) & T(e_2) & e_2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & -3 & 0 \\ -2 & -1 & 4 & 1 \\ -2 & -2 & -3 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix}.$$

Esto es, $A = MJM^{-1}$.

Forma canónica de Jordan

En este caso, se verifica que

$$\mathrm{rg}\ A = 3, \quad \mathrm{rg}\ A^2 = 2 \quad \mathrm{rg}\ A^3 = 1 \ .$$

Por el Teorema de la dimensión, sabemos que, si llamamos $K_i = \mathrm{Ker}\ T^i$, $i = 0, 1, 2, 3, 4$, donde tomamos $T^0 = I$, luego $K_0 = \{0\}$.

$$\dim K_0 = 0, \quad \dim K_1 = 1, \quad \dim K_2 = 2, \quad \dim K_3 = 3, \quad \dim K_4 = 4, \quad .$$

Notamos $n_i = \dim K_i$ y $d_i = n_i - n_{i-1}$, $i = 1, 2, 3, 4$. Por tanto, en este caso, obtenemos

$$d_1 = 1 - 0 = 1, \quad d_2 = 2 - 1 = 1, \quad d_3 = 3 - 2 = 1, \quad d_4 = 4 - 3 = 1 \ .$$

d_1 indica el número de bloques de Jordan que aparece en la matriz de Jordan semejante a A. En este caso, sólo hay un bloque. En general, para saber el número de bloques de cada dimensión que aparecen, se tiene en cuenta las diferencias $d_i - d_{i+1}$, que son, en este caso,

$$d_1 - d_2 = 0, \quad d_2 - d_3 = 0, \quad d_3 - d_4 = 0, \quad d_4 = 1 \ .$$

Como $d_1 - d_2 = 0$, no hay bloques de orden 1. Por el mismo argumento, tampoco hay bloques de órdenes 2 ($d_2 - d_3 = 0$) ni 3 ($d_3 - d_4 = 0$). Como $d_4 = 1$, hay un bloque de orden 4.

Ejemplo

La matriz A dada por

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

es nilpotente de orden $p = 3$, ya que $A^3 = 0$ y $A^2 \neq 0$. De hecho,

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

Además se verifica que $\mathrm{rg}\, A = 3, \mathrm{rg}\, A^2 = 1$.

Forma canónica de Jordan

Como

$$\text{rg } A = 3, \quad \text{rg } A^2 = 1 \quad \text{rg } A^3 = 0 \ .$$

Por el Teorema de la dimensión, sabemos que si T es el operador asociado a A, y llamamos $K_i = \text{Ker } T^i$, $i = 0, 1, 2, 3$, donde tomamos $T^0 = I$, luego $K_0 = \{0\}$.

$$\dim K_0 = 0, \quad \dim K_1 = 2, \quad \dim K_2 = 4, \quad \dim K_3 = 5 \ .$$

Notamos $n_i = \dim K_i$ y $d_i = n_i - n_{i-1}$, $i = 1, 2, 3$. Por tanto, en este caso, obtenemos

$$d_1 = 2 - 0 = 2, \quad d_2 = 4 - 2 = 2, \quad d_3 = 5 - 4 = 1 \ .$$

d_1 indica el número de bloques de Jordan que aparece en la matriz de Jordan semejante a A. En este caso, hay dos bloques. En general, para saber el número de bloques de cada dimensión que aparecen, se tiene en cuenta las diferencias $d_i - d_{i+1}$, que son, en este caso,

$$d_1 - d_2 = 0, \quad d_2 - d_3 = 1, \quad d_3 = 1 \ .$$

Como $d_1 - d_2 = 0$, no hay bloques de orden 1. Por el mismo argumento, hay un bloque de orden dos ($d_2 - d_3 = 1$) y otro de orden 3 ($d_3 = 1$).

Forma canónica de Jordan

En este caso, la matriz A es semejante a la siguiente matriz de Jordan

$$J = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right),$$

Forma canónica de Jordan

Ahora daremos la matriz del cambio de base. Para ello, recordamos la cadena de inclusiones de los núcleos de T^i, que son, en este caso

$$\{0\} = K_0 \subset K_1 \subset K_2 \subset K_3 = \mathbb{R}^5 \ ,$$

con dimensiones

$$0 < 2 < 4 < 5 \ .$$

Elegimos un vector $v_1 \notin K_2$ y otro v_2 tal que $v_2 \in K_2 \backslash K_1$, tal que $T(v_2) \in K_2 \backslash K_1$. J será entonces la matriz asociada a T respecto de la base formada por los vectores

$$\{T^2(v_1), T(v_1), v_1, T(v_2), v_2\} \ .$$

En este caso, basta tomar $v_1 = e_2$, ya que $T(e_2) = e_1$, $T^2(e_2) = (0, 1, 1, 1, 0)$. El siguiente vector v_2 es tal que $v_2, T(v_2) \in K_2 \backslash K_1$. En este caso, los subespacios K_1 y K_2 están dados por

$$K_1 = \{(0, y, z, u, 0) \in \mathbb{R}^5 : y - 3z + 2u = 0\},$$

$$K_2 = \{(x, y, z, u, v) \in \mathbb{R}^5 : y - 3z + 2u = 0\}$$

El vector $v_2 = (1, 3, 1, 0, 1)$ verifica que $v_2 \in K_2 \backslash K_1$ y $T(v_2) = (0, 1, 3, 4, 0)$.

Forma canónica de Jordan

Podemos usar la siguiente base de $\mathbb{R}^5$

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

La matriz M que da el cambio de base es

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 3 \\ 1 & 0 & 0 & 3 & 1 \\ 1 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} .$$

Se verifica entonces que $A = MJM^{-1}$.

Forma canónica de Jordan

Nótese que el único valor propio de una matriz nilpotente es 0. Para matrices que tienen un único valor propio, existe un resultado similar al que enunciamos antes para nilpotentes.

Proposición

Sea $A \in \mathcal{M}_n$ tal que λ es un autovalor de A de multiplicidad n. Entonces existe una matriz de Jordan J semejante a A dada por

$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_k \end{pmatrix},$$

donde cada submatriz J_i es un bloque elemental de Jordan de la forma

$$J_i = \begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}$$

El orden de las submatrices J_i es decreciente y éste puede tomar los valores $p, p-1, \cdots, 1$, siendo p el menor natural tal que $(A - \lambda I_n)^p = 0$. Además el número de bloques (k) es $n - \mathrm{rg}\ (A - \lambda I_n)$ y existen r_m bloques de orden $m \in \{1, 2, \cdots, p\}$ siendo

$$r_m = \mathrm{rg}\ (A - \lambda I_n)^{m-1} + \mathrm{rg}\ (A - \lambda I_n)^{m+1} - 2\mathrm{rg}\ (A - \lambda I_n)^m .$$

Idea de la demostración:
Se prueba que $A - \lambda I_n$ es nilpotente de orden p, luego se puede aplicar el resultado válido para idempotentes con lo cual, existe una base de $\mathbb{K}^n$ para la que la aplicación lineal asociada a $A - \lambda I_n$ viene dada por una matriz de Jordan con ceros en la diagonal principal.
Sumando λI, se obtiene que el operador asociado a A viene dado por J en términos de esa misma base.

Ejemplo

La matriz A dada por

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

tiene por polinomio característico

$$P(\lambda) = \lambda^4 - 12t^3 + 54t^2 - 108t + 81 = (t-3)^4$$

por tanto, el único valor propio es 3 (con multiplicidad 4). En este caso llamamos $B = A - 3I_4$, esto es,

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

Averiguamos el mínimo p tal que $B^p = 0$. En este caso tenemos

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

y $B^3 = 0$, luego $p = 3$ y los bloques de Jordan que aparecen pueden ser de orden $1, 2$ ó 3. En este caso, $\mathrm{rg}\,(B) = 2$, luego aparecen $4 - 2 = 2$ bloques de Jordan en la diagonalización. Veamos cuántos aparecen de cada orden. Sabemos que $\mathrm{rg}\,(B^2) = 1, \mathrm{rg}\, B^3 = 0$. Si r_m es el número de bloques elementales de Jordan de orden m, sabemos que

$$r_1 = 4 + 1 - 2.2 = 1, r_2 = 2 + 0 - 2.1 = 0, r_3 = 1 + 0 - 2.0 = 1 \ ,$$

por tanto, hay un bloque de orden 3 y otro de orden 1. Así, B es semejante a la matriz de Jordan J dada por

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

Para dar la matriz del cambio de base, necesitamos en este caso, un vector v_1 que no esté en el núcleo de T^2 y otro vector v_2 de $\mathrm{Ker}\ T$ que no pertenezca al subespacio vectorial generado por v_1, $T(v_1)$, $T^2(v_1)$, donde T es el operador asociado a B. Describimos entonces el núcleo de T y de T^2.

$$K_1 = \mathrm{Ker}\ T = \{(x, y, z, t) \in \mathbb{R}^4 : x + z = 0, 2y + t = 0\},$$

$$K_2 = \mathrm{Ker}\ T^2 = \{(x, y, z, t) \in \mathbb{R}^4 : x - 2y + z - t = 0\}\ .$$

Por ejemplo, podemos tomar $v_1 = e_1$, entonces

$$T^2(e_1) = (1, 0, -1, 0), \quad T(e_1) = (1, 1, 0, -1)\ .$$

Podemos tomar $v_2 = (0, 1, 0, -2)$ y la matriz del cambio de base es

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

Se verifica que $MJM^{-1} = B$, luego, si llamamos $C := J + 3I_4$, esto es,

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

se verifica que C es semejante a A y además $MCM^{-1} = A$, ya que

$$MCM^{-1} = M(J + 3I_4)M^{-1} = B + 3I_4 = A\,.$$

Forma canónica de Jordan

Ejemplo

Consideramos la siguiente matriz

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

cuyo polinomio característico es

$$P(\lambda) = -\lambda^5 + 5i\lambda^4 + 10\lambda^3 - 10i\lambda^2 - 5\lambda + i = -(\lambda - i)^5 .$$

Se verifica que $\mathrm{rg}\,(A - iI_5) = 2$, por tanto $\dim \mathrm{Ker}\,(T - iI) = 3$, donde T es la aplicación lineal de $\mathbb{C}^5$ en $\mathbb{C}^5$ asociada a A, luego A no es diagonalizable, pues $3 < 5$. Como i es el único valor propio de A, entonces $B = A - iI_5$ es nilpotente.

De hecho

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

y $B^2 = 0$, luego B es semejante a una matriz de Jordan, donde los bloques elementales tienen orden 1 ó 2. Como $\mathrm{rg}\ B = 2$, entonces aparecerán $5 - 2 = 3$ bloques elementales de Jordan. Ahora miramos los índices que nos dicen cuántos bloques aparecen de cada orden. En este caso, tenemos que

$$r_1 = \mathrm{rg}\ B^0 + \mathrm{rg}\ B^2 - 2\mathrm{rg}\ B = 5+0-2.2 = 1, \quad r_2 = \mathrm{rg}\ B + \mathrm{rg}\ B^3 - 2\mathrm{rg}\ B^2 = 2+0-2.0$$

Forma canónica de Jordan

Por tanto, B es semejante a la matriz de Jordan J dada por

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

Para dar la matriz del cambio de base, necesitamos en este caso, dos vectores $v_1, v_2 \notin \mathrm{Ker}\ T$ linealmente independientes y otro vector $v_3 \in \mathrm{Ker}\ T$, que no esté generado por

$$\{T(v_1), v_1, T(v_2), v_2\}$$

Calculamos el núcleo del operador T, dado por

$$K_1 = \mathrm{Ker}\ T = \{(0, y, z, z, v) : y, z, v \in \mathbb{C}\}$$

Podemos tomar $v_1 = e_1, v_2 = e_3$, con lo cual

$$T(v_1) = (0, 1, 1, 1, 0), \quad T(v_2) = (0, -2, -1, -1, 0)$$

El vector $v_3 = e_5 \in \operatorname{Ker} T$ y no está en el subespacio generado por $\{e_1, e_3, T(e_1), T(e_3)\}$, con lo que la matriz del cambio de base viene dada por

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

Se verifica que $MJM^{-1} = B$. Si llamamos $C := J + iI_5$, esto es,

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

se verifica que C es semejante a A. Además M da el cambio de base, ya que

$$MCM^{-1} = M(J + iI_4)M^{-1} = B + iI_4 = A \, .$$

Forma canónica de Jordan

Teorema

Sea $T : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ una aplicación lineal con matriz asociada A. Supongamos que los autovalores de A son $\lambda_1, \cdots, \lambda_k \in \mathbb{K}$ y se verifica que $\sum_{i=1}^{k} m_i = n$, donde m_i es la multiplicidad de λ_i. Entonces se verifica que

- A es semejante a una matriz triangular superior que tiene en la diagonal principal los autovalores de A.
- La matriz A es semejante a una matriz diagonal por bloques de la forma

$$B = \begin{pmatrix} B_1 & 0 & \cdots & 0 \\ 0 & B_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k \end{pmatrix},$$

  donde, para cada i, B_i es una matriz triangular superior de orden m_i cuyos elementos de la diagonal principal son λ_i.

Idea de la demostración

El apartado primero se prueba por inducción sobre la dimensión del espacio. Para $n = 1$ es trivial, y para pasar de n a $n + 1$, se separa un valor propio y se elige una base con un vector propio asociado a ese valor propio. Basta aplicar la hipótesis de inducción.

En virtud de la primera parte, puede suponerse que A es semejante a una matriz triangular superior que tiene k bloques en la diagonal principal que son matrices triangulares superiores de órdenes $m_1, m_2, \cdots, m_k$ tales que la matriz i-ésima tiene a λ_i en la diagonal principal. Ahora hay que hacer cero el resto de los coeficientes de la matriz que están fuera de esos bloques. Para eso se usa un procedimiento similar al usado para calcular la inversa de una matriz usando el método de Gauss-Jordan. En este caso, se hace un cambio de base que sólo cambia uno de los vectores de la base de partida (v_i) por otro del tipo $v_i + tv_j$, de forma que al aplicar este cambio de base, se consigue anular alguno de los coeficientes de la matriz anterior.

De esta forma se consiguen anular coeficientes, de abajo a arriba y de izquierda a derecha.

Un ejemplo de cómo cambian las matrices al aplicar ese tipo de cambios de base.
Consideramos la matriz A dada por

$$A = \begin{pmatrix} a & b & c & d & e \\ 0 & f & g & h & i \\ 0 & 0 & j & k & l \\ 0 & 0 & 0 & m & n \\ 0 & 0 & 0 & 0 & o \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -t \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

Forma canónica de Jordan

Se tiene que

$$M^{-1}AM = \begin{pmatrix} a & b & c & d & e+dt \\ 0 & f & g & h & i+ht \\ 0 & 0 & j & k & l+kt \\ 0 & 0 & 0 & m & n+(m-o)t \\ 0 & 0 & 0 & 0 & o \end{pmatrix}, \quad A = \begin{pmatrix} a & b & c & d & e \\ 0 & f & g & h & i \\ 0 & 0 & j & k & l \\ 0 & 0 & 0 & m & n \\ 0 & 0 & 0 & 0 & o \end{pmatrix},$$

Otro ejemplo

Consideramos la matriz A dada por

$$A = \begin{pmatrix} a & b & c & d & e \\ 0 & f & g & h & i \\ 0 & 0 & j & k & l \\ 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & 0 & o \end{pmatrix},$$

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & t & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad M^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -t & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

Se tiene que

$$M^{-1}AM = \begin{pmatrix} a & b & c & d+ct & e \\ 0 & f & g & h+gt & i \\ 0 & 0 & j & k+(j-m)t & l \\ 0 & 0 & 0 & m & 0 \\ 0 & 0 & 0 & 0 & o \end{pmatrix},$$

mientras que

$$A = \begin{pmatrix} a & b & c & d & e \\ 0 & f & g & h & i \\ 0 & 0 & j & k & l \\ 0 & 0 & 0 & m & n \\ 0 & 0 & 0 & 0 & o \end{pmatrix},$$

Forma canónica de Jordan

Nótese que una vez que se obtiene una matriz diagonal por bloques como en el enunciado del Teorema, cada uno de los bloques B_i es una matriz con un único autovalor, luego se puede aplicar el método descrito en uno de los enunciados a cada uno de los bloques de la diagonal.

Ejemplo

Consideramos la matriz A dada por

$$A = \begin{pmatrix} 1 & 2 & -7 & 5 & 1 \\ 2 & 2 & -8 & 7 & 0 \\ 2 & 1 & -9 & 9 & 2 \\ 2 & 1 & -12 & 12 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix} .$$

tiene por polinomio característico

$$P(\lambda) = -\lambda^5 + 9\lambda^4 - 30t^3 + 46t^2 - 33 + 9 = -(t-3)^2(t-1)^3 .$$

Luego los valores propios y sus multiplicidades son

$$\lambda_1 = 1, \;\; m_1 = 3, \quad \lambda_2 = 3, \;\; m_2 = 2 .$$

Se verifica que $\dim V(1) = 1, \dim V(3) = 1$, luego A no es diagonalizable.

En este caso tenemos

$$V(1) = \mathrm{L}\,\{(0,1,1,1,0)\}, \quad V(3) = \mathrm{L}\,\{(3,9,11,13,0)\}\,.$$

Intentamos construir una matriz triangular superior semejante a A. Para ello, empezamos por construir una base de $\mathbb{R}^5$ que contenga un vector propio asociado a uno de los dos valores propios. Usamos, para empezar, la base

$$\{(0,1,1,1,0), e_1, e_3, e_4, e_5\}$$

y la matriz P_0 del cambio de base es de la aplicación lineal asociada a T en términos de esta nueva base es

$$P_0 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}\,.$$

Así que la matriz de la aplicación lineal asociada a T en términos de esta nueva base es

$$T_1 := P_0^{-1}AP_0 = \left(\begin{array}{cc|ccc} 1 & 2 & -8 & 7 & 0 \\ 0 & 1 & -7 & 5 & 1 \\ \hline 0 & 0 & -1 & 2 & 2 \\ 0 & 0 & -4 & 5 & 3 \\ 0 & 0 & 0 & 0 & 3 \end{array}\right) = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

Nótese que la submatriz A_{11} es triangular superior y 1 es el único valor propio. Como T_1 tiene los mismos valores propios que A, esto significa que 1 y 3 pueden ser los únicos valores propios de A_{22} (con multiplicidades 1 y 2, respectivamente). Ahora se repite el procedimiento para esta última submatriz (A_{22}).

Forma canónica de Jordan

En efecto, el polinomio característico de A_{22} es

$$-t^3 + 7t^2 - 15t + 9 = -(t-1)(t-3)^2$$

y los subespacios propios en este espacio tres-dimensional son

$$V(1) = \mathrm{L}\,\{(1,1,0)\}, \quad V(3) = \mathrm{L}\,\{(1,2,0)\}$$

Completamos a una base de $\mathbb{R}^3$ que incluya al menos un vector propio, por ejemplo:

$$\{(1,1,0),(1,2,0),(0,0,1)\}$$

con lo que ahora podemos considerar la nueva base en $\mathbb{R}^5$ siguiente:

$$\{e_1, e_2, (0,0,1,1,0),(0,0,1,2,0),(0,0,0,0,1)\}.$$

Por tanto, la matriz P_1 del cambio de base y la matriz $T_2 := P_1^{-1} T_1 P_1$ son

$$P_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad T_2 = P_1^{-1} T_1 P_1 = \begin{pmatrix} 1 & 2 & -1 & 6 & 0 \\ 0 & 1 & -2 & 3 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Nótese que la última matriz es triangular superior y en la diagonal principal aparecen dos bloques, de órdenes 3 y 2, tales que cada uno tiene en la diagonal principal un único elemento.
Ahora hacemos ceros los elementos que están fuera de los dos bloques que aparecen en la diagonal principal. Empezamos por el coeficiente que aparece en el lugar $3-5$. Para ello usamos una matriz del tipo

$$P_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

para conveniente parámetro t, entonces

$$T_3 = P_2^{-1} T_2 P_2 = \begin{pmatrix} 1 & 2 & -1 & 6 & -\frac{1}{2} \\ 0 & 1 & -2 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Ahora repetimos el procedimiento para anular el elemento que ocupa el lugar $2-4$ de la matriz T_3. Para ello tomamos usamos una matriz del tipo

$$P_3 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

para conveniente parámetro t, entonces

$$T_4 = P_3^{-1} T_3 P_3 = \begin{pmatrix} 1 & 2 & -1 & 9 & \frac{-1}{2} \\ 0 & 1 & -2 & 0 & \frac{-3}{2} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

Tomando ahora

$$P_4 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{-3}{4} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad P_5 = \begin{pmatrix} 1 & 0 & 0 & \frac{9}{2} & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

y

$$P_6 = \begin{pmatrix} 1 & 0 & 0 & 0 & \frac{-13}{4} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

obtenemos

$$T_5 = P_4^{-1} T_4 P_4 = \begin{pmatrix} 1 & 2 & -1 & 9 & -2 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix},$$

$$T_6 = P_5^{-1} T_5 P_5 = \begin{pmatrix} 1 & 2 & -1 & 0 & \frac{-13}{2} \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix},$$

$$T_7 = P_6^{-1} T_6 P_6 = \begin{pmatrix} 1 & 2 & -1 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 3 \end{pmatrix}.$$

Por último, ahora cada uno de los bloques que aparece en la diagonal principal son matrices nilpotentes, por tanto, a cada una de ellas se puede aplicar el método que existe para convertirlas en una matriz de Jordan.

Potencias de matrices

Proposición

Sea $A \in \mathcal{M}_n(\mathbb{K})$ y supongamos que sus autovalores son $\lambda_1, \cdots, \lambda_n \in \mathbb{K}$ (no necesariamente distintos).

- Si $A = PDP^{-1}$ para matrices D diagonal y P inversible, entonces
$$A^k = PD^k P^{-1} \quad \forall k \in \mathbb{N} \,.$$
- Si A no es diagonalizable, J es su forma canónica de Jordan y $A = PJP^{-1}$, siendo P una matriz inversible y
$$J = \begin{pmatrix} J_1 & 0 & \cdots & 0 \\ 0 & J_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_r \end{pmatrix} ,$$
donde, para cada i, J_i es un bloque elemental de Jordan.

Entonces se verifica que

$$A^k = PJ^kP^{-1} = P\begin{pmatrix} J_1^k & 0 & \cdots & 0 \\ 0 & J_2^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & J_r^k \end{pmatrix}P^{-1} ,$$

donde, para cada $i \in \{1, 2, \cdots, r\}$,

$$J_i^k = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}^k = \sum_{i=0}^{k} \binom{k}{i} \lambda_i^{k-i} B^i ,$$

siendo $B = J_i - \lambda_i I_s$, para $s \leq m_i$ el orden del bloque J_i y además se verifica que $B_i^s = 0$.

Proposición

Supongamos que $A \in \mathcal{M}_m$ tal que la serie $\sum_n A^n$ converge, entonces $I_m - A$ es inversible y se verifica que

$$(I_m - A)^{-1} = \sum_{k=0}^{\infty} A^k .$$

Demostración.
Se verifica que

$$(I_m - A)\Big(\sum_{k=0}^{n} A^k\Big) = \sum_{k=0}^{n} A^k - \sum_{k=1}^{n} A^{n+1} =$$

$$= I_m - A^{n+1} .$$

Tomando límite en la igualdad $(I_m - A)(\sum_{k=0}^{n} A^k) = I_m - A^{n+1}$ y usando que el producto de matrices es continuo se tiene que

$$(I_m - A)\Big(\sum_{k=0}^{\infty} A^k\Big) = I_m ,$$

Aplicaciones

La hipótesis de la proposición anterior (convergencia de la serie) se verifica, por ejemplo, si A es diagonalizable y todos los valores propios de A tienen módulo menor que uno.

Ejemplo

Sea $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, donde a y b son números reales con valor absoluto menor que uno, entonces la matriz $I_2 - A$ es inversible y se tiene que

$$A^k = \begin{pmatrix} a^k & 0 \\ 0 & b^k \end{pmatrix}, \forall k \in \mathbb{N} .$$

Por tanto,

$$\sum_{k=0}^{n} A^k = \begin{pmatrix} \sum_{k=0}^{n} a^k & 0 \\ 0 & \sum_{k=0}^{n} b^k \end{pmatrix} =$$

$$\begin{pmatrix} \frac{1-a^{n+1}}{1-a} & 0 \\ 0 & \frac{1-b^{n+1}}{1-b} \end{pmatrix} .$$

Por tanto,

$$\sum_{k=0}^{n} A^k = \begin{pmatrix} \frac{1-a^{n+1}}{1-a} & 0 \\ 0 & \frac{1-b^{n+1}}{1-b} \end{pmatrix}.$$

Tomando límite $n \to \infty$ obtenemos que

$$\begin{pmatrix} 1-a & 0 \\ 0 & 1-b \end{pmatrix}^{-1} = (I_2 - A)^{-1} = \begin{pmatrix} \frac{1}{1-a} & 0 \\ 0 & \frac{1}{1-b} \end{pmatrix}.$$

Por supuesto, si la matriz D es diagonal (no importa el orden) y todos los valores propios son menores que uno, se puede aplicar el mismo método.

Aplicaciones

En tal caso, si

$$D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

se tiene que

$$(I_n - D)^{-1} = \begin{pmatrix} \frac{1}{1-\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{1-\lambda_n} \end{pmatrix}.$$

De todos modos el resultado de la aplicación anterior es trivial.

Aplicaciones

Si $A \in \mathcal{M}_n$ es diagonalizable, y se verifica que $A = MDM^{-1}$, para D una matriz diagonal cuyos valores propios tienen módulo menor que uno, entonces sabemos que

$$A^k = MD^kM^{-1}, \quad \forall k \in \mathbb{N}\,.$$

Por tanto

$$\sum_{k=0}^{n} A^k = \sum_{k=0}^{n} MD^kM^{-1} =$$

$$M\left(\sum_{k=0}^{n} D^k\right)M^{-1}\,.$$

Nótese que la matriz $\sum_{k=0}^{n} D^k$ tiene una forma sencilla.
Algo similar se puede hacer si A no es diagonalizable, pero conocemos una matriz de Jordan semejante a A.

Observación

- Si x es un número real (o complejo) se verifica que

$$e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n, \qquad \operatorname{sen}\ x = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} x^{2n+1},$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} x^{2n} .$$

- El espacio vectorial $L(\mathbb{K}^n)$ (luego $M_n(\mathbb{K})$) se puede dotar de una norma que verifica

$$\|S \circ T\| \leq \|S\| \, \|T\| \, , \forall S, T \in L(\mathbb{K}^n)$$

luego

$$\|T^k\| \leq \|T\|^k, \quad \forall k \in \mathbb{N}.$$

- $L(\mathbb{K}^n)$ es completo con esa norma, luego toda serie $\sum_m T_m$ en $L(X)$ tal que $\sum_m \|T_m\|$ converja es, de hecho, convergente en $L(\mathbb{K}^n)$.

Aplicaciones

Definición

Supongamos que $A \in \mathcal{M}_m$, entonces se define la exponencial de A, el seno de A y el coseno de A de la siguiente forma:

$$e^A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n, \qquad \operatorname{sen} A = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} A^{2n+1},$$

$$\cos A = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n)!} A^{2n}.$$

Las series anteriores son convergentes, ya que en norma están dominadas por una serie del tipo

$$\sum_n \frac{1}{n!} r^n$$

para algún real positivo r y esta última serie converge siempre.

Ejemplo

Sea $A = MJM^{-1}$ donde $M \in \mathcal{M}_3$ es una matriz inversible y

$$J = \begin{pmatrix} \frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{4} \end{pmatrix} .$$

Pretendemos calcular e^A (en función de M, claro).
Para ello calculamos las potencias de J, para ello usamos las submatrices

$$J_1 = \begin{pmatrix} \frac{1}{2} & 1 \\ 0 & \frac{1}{2} \end{pmatrix} , \quad J_2 = \left(\frac{1}{4}\right) .$$

Se tiene que

$$A^k = MJ^k M^{-1} = M \begin{pmatrix} J_1^k & 0 \\ 0 & J_2^k \end{pmatrix} M^{-1}$$

Aplicaciones

Ahora bien

$$J_1 = \frac{1}{2} I_2 + B \ ,$$

donde $B := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, luego $B^2 = 0$. Por tanto, si $k \geq 2$, tenemos que

$$J_1^k = \left(\frac{1}{2} I_2 + B\right)^k = \sum_{i=0}^{k} \binom{k}{i} \frac{1}{2^{k-i}} B^i \quad (\text{como } B^p = 0 \text{ para } p \geq 2)$$

$$\binom{k}{0} \frac{1}{2^{k-0}} B^0 + \binom{k}{1} \frac{1}{2^{k-1}} B^1 =$$

$$\frac{1}{2^k} I_2 + k \frac{1}{2^{k-1}} B \ .$$

Aplicaciones

Por tanto,

$$e^A = \sum_{k=0}^{\infty} \frac{1}{k!} A^k =$$

$$\sum_{k=0}^{\infty} \frac{1}{k!} M \begin{pmatrix} J_1^k & 0 \\ 0 & J_2^k \end{pmatrix} M^{-1} =$$

$$M \left(\sum_{k=0}^{\infty} \frac{1}{k!} \begin{pmatrix} J_1^k & 0 \\ 0 & J_2^k \end{pmatrix} \right) M^{-1}$$

Ahora calculamos de forma separada la suma de los bloques que aparecen en la diagonal principal de la matriz anterior. Empezamos por el segundo

$$\sum_{k=0}^{\infty} \frac{1}{k!} J_2^k =$$

$$\sum_{k=0}^{\infty} \frac{1}{k!} \frac{1}{4^k} = e^{\frac{1}{4}}$$

Aplicaciones

Ahora nos ocupamos de la submatriz de orden 2, esto es,

$$\sum_{k=0}^{\infty} \frac{1}{k!} J_1^k =$$

$$I_2 + J_1 + \sum_{k=2}^{\infty} \frac{1}{k!} \left(\frac{1}{2^k} I_2 + k \frac{1}{2^{k-1}} B \right) =$$

$$I_2 + \left(\frac{1}{2} I_2 + B \right) + \sum_{k=2}^{\infty} \frac{1}{k!} \left(\frac{1}{2^k} I_2 + k \frac{1}{2^{k-1}} B \right) =$$

$$\sum_{k=0}^{\infty} \frac{1}{k!} \frac{1}{2^k} I_2 + \sum_{k=1}^{\infty} \frac{1}{k!} k \frac{1}{2^{k-1}} B =$$

$$e^{\frac{1}{2}} I_2 + \sum_{k=0}^{\infty} \frac{1}{k!} \frac{1}{2^k} B = e^{\frac{1}{2}} \left(I_2 + B \right) .$$

Aplicaciones

Sólo falta sustituir en la expresión de e^A, y se obtiene que

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

Tema 4. Formas cuadráticas.

Análisis Matemático I
1^o Licenciatura de Estadística

Forma cuadráticas

Definición

Una aplicación $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ es una forma cuadrática si es un polinomio homogéneo de grado dos en n variables, esto es, si existen números reales $\{a_{ij} : 1 \leq i, j \leq n\}$ tales que

$$P(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j, \quad \forall x \in \mathbb{R}^n .$$

Una forma cuadrática Q en $\mathbb{R}^n$ siempre procede de una forma bilineal $B : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ de manera que

$$P(x) = B(x, x), \forall x \in \mathbb{R}^n .$$

Formas cuadrátricas

Ejemplos

- En $\mathbb{R}$ una forma cuadrática tiene la forma

$$Q(x) = ax^2 \quad (x \in \mathbb{R})$$

  para conveniente número real a.

- La aplicación P dada por

$$P(x,y,z) = x^2 + 2xy + 6y^2 + 4yz + z^2 \quad ((x,y,z) \in \mathbb{R}^3)$$

  es cuadrática y la forma bilineal asociada B viene dada por

$$B((x,y,z),(u,v,w)) = xu+2xv+6yv+4yw+zw \quad \forall (x,y,z),(u,v,w) \in \mathbb{R}^3$$

Observación

Para cada forma cuadrática Q en $\mathbb{R}^n$ existe una matriz A cuadrada de orden n tal que

$$Q(x) = xAx^t, \quad \forall x \in \mathbb{R}^n .$$

Ejemplo

Si P es la forma cuadrática dada por

$$P(x,y,z) = x^2 + 2xy + 6y^2 + 4yz + z^2 \quad ((x,y,z) \in \mathbb{R}^3),$$

y

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 6 & 2 \\ 0 & 2 & 1 \end{pmatrix},$$

es inmediato comprobar que $P(X) = XAX^t$ para cada $X \in \mathbb{R}^3$.

Lema

Dada una forma cuadrática Q en $\mathbb{R}^n$, existe una única matriz simétrica B de orden n tal que

$$Q(x) = xBx^t, \quad \forall x \in \mathbb{R}^n \ .$$

Diremos en tal caso que B es la matriz asociada a la forma cuadrática Q.

Formas cuadráticas

Tipos de formas cuadráticas

Sea $Q : \mathbb{R}^n \longrightarrow \mathbb{R}$ una forma cuadrática.

- Q es definida positiva si $Q(x) > 0$ para todo vector x de $\mathbb{R}^n$ no nulo.
- Q es definida negativa si $Q(x) < 0$ para todo vector x de $\mathbb{R}^n$ no nulo.
- Q es semidefinida positiva si $Q(x) \geq 0$ para todo vector x de $\mathbb{R}^n$.
- Q es semidefinida negativa si $Q(x) \leq 0$ para todo vector x de $\mathbb{R}^n$.
- Q es indefinida si existen vectores $x, y \in \mathbb{R}^n$ tales que $Q(x) < 0$ y $Q(y) > 0$.

Formas cuadráticas

Ejemplos

- La forma cuadrática $Q : \mathbb{R}^3 \longrightarrow \mathbb{R}$ dada por $Q(x, y, z) = x^2 + 2y^2 + z^2 + 2xy$ es definida positiva ya que

$$Q(x, y, z) = (x + y)^2 + y^2 + z^2,$$

  luego Q toma valores no negativos y sólo se anula en $(0, 0, 0)$.

- La forma cuadrática $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $Q(x, y) = x^2 - 4xy + 4y^2$ es semidefinida positiva ya que

$$Q(x, y) = (x - 2y)^2$$

  luego Q toma valores no negativos y se anula en vectores no nulos.

- La forma cuadrática $Q : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $Q(x, y) = xy$ es indefinida, ya que

$$Q(1, 1) = 1 > 0, \quad Q(1, -2) = -2 < 0 .$$

Forma canónica

Supongamos que la matriz D asociada a la forma cuadrática Q en $\mathbb{R}^n$ es diagonal.

- Q es definida positiva si $d_{ii} > 0$ para todo i.
- Q es definida negativa si $d_{ii} < 0$ para todo i.
- Q es semidefinida positiva si $d_{ii} \geq 0$ para todo i.
- Q es semidefinida negativa si $d_{ii} \leq 0$ para todo i.
- Q es indefinida si existen i, j tales que $d_{ii}d_{jj} < 0$.

Criterios de clasificación de formas cuadráticas

Autovalores

Supongamos que la matriz A asociada a la forma cuadrática Q en $\mathbb{R}^n$ tiene por autovalores $\lambda_1, \lambda_2, \cdots, \lambda_n$ (no necesariamente distintos).

- Q es definida positiva si $\lambda_i > 0$ para todo $i \in \{1, 2, \cdots, n\}$.
- Q es definida negativa si $\lambda_i < 0$ para todo $1 \leq i \leq n$.
- Q es semidefinida positiva si $\lambda_i \geq 0$ para todo $1 \leq i \leq n$.
- Q es semidefinida negativa si $\lambda_i \leq 0$ para todo $i \in \{1, 2, \cdots, n\}$.
- Q es indefinida si existen i, j tales que $\lambda_i \lambda_j < 0$.

Justificación del criterio de los menores

Supongamos que Q es una forma cuadrática en $\mathbb{R}^2$ con matriz asociada A y que $a_{11} \neq 0$. Sabemos que

$$Q(x,y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2, \quad \forall (x,y) \in \mathbb{R}^2 .$$

Completamos cuadrados en la expresión anterior, con lo que obtenemos

$$Q(x,y) = a_{11}\left(x^2 + 2\frac{a_{12}}{a_{11}}xy + \frac{a_{12}^2}{a_{11}^2}y^2\right) + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)y^2 =$$

$$a_{11}\left(x + \frac{a_{12}}{a_{11}}y\right)^2 + \frac{\det(A)}{a_{11}}y^2 .$$

Criterios de clasificación de formas cuadráticas

Justificación del criterio de los menores

Supongamos que Q es definida positiva, entonces

$$Q(e_1) = a_{11} > 0 \; .$$

Como

$$Q(x,y) = a_{11}\left(x + \frac{a_{12}}{a_{11}}y\right)^2 + \frac{\det(A)}{a_{11}}y^2 \; ,$$

si restringimos Q al subespacio vectorial dado por $\{(x,y) \in \mathbb{R}^2 : x + \frac{a_{12}}{a_{11}}y = 0\}$, usando que Q es definida positiva y que $a_{11} > 0$, obtenemos que $\det A > 0$.

Recíprocamente, usando la expresión anterior de Q, si suponemos que $a_{11} > 0$ y que $\det(A) > 0$, entonces Q es definida positiva.

De manera similar, aunque usando expresiones más complicadas, se obtienen criterios para clasificar las formas cuadráticas de más de dos variables en términos de los menores principales de la matriz asociada.

Formas cuadráticas

Notación

Sea A una matriz cuadrada de orden n. Para $i \in \{1, 2, \cdots, n\}$, notaremos por D_i al determinante de la submatriz de orden i que se obtiene a partir de las primeras i filas de A y de las primeras i columnas de A, esto es,

$$D_i = \det \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1i} \\ a_{11} & a_{12} & \cdots & a_{1i} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ii} \end{pmatrix}$$

Ejemplo

Si A es la matriz dada por

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

entonces $D_1 = 1, \quad D_2 = \det \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} = -6, \quad D_3 = \det A = -74.$

Criterios de clasificación de formas cuadráticas

Teorema

Sea Q una forma cuadrática en $\mathbb{R}^n$ y A la matriz cuadrada de orden n asociada. Entonces se verifica que

- Q es definida positiva si, y sólo si, $D_i > 0, \forall 1 \leq i \leq n$.
- Q es definida negativa si, y sólo si, $(-1)^i D_i > 0, \forall 1 \leq i \leq n$.
- Q es semidefinida positiva si $D_i > 0, \forall 1 \leq i \leq n-1$ y $D_n = 0$.
- Q es semidefinida negativa si $(-1)^i D_i > 0, \forall 1 \leq i \leq n-1$ y $D_n = 0$.

Criterios de clasificación de formas cuadráticas

Ejemplos

- La forma cuadrática en $\mathbb{R}^3$ dada por
$$Q(x,y,z) = (x\ y\ z)\begin{pmatrix} 1 & 1 & 1 \\ 1 & 4 & 0 \\ 1 & 0 & 7 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
es definida positiva, ya que $D_1 = 1 > 0$, $D_2 = \det\begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} = 3 > 0$ y $D_3 = \det A = 17 > 0$.
- En cambio, si
$$Q(x,y,z) = (x\ y\ z)\begin{pmatrix} -4 & 0 & 2 \\ 0 & -5 & 0 \\ 2 & 0 & -2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
entonces $D_1 = -4 < 0$, $D_2 = \det\begin{pmatrix} -4 & 0 \\ 0 & -5 \end{pmatrix} = 20 > 0$ y $D_3 = \det A = -20 < 0$, luego Q es definida negativa.

Criterios de clasificación de formas cuadráticas

Ejemplos

- La forma cuadrática en $\mathbb{R}^3$ dada por
$$Q(x,y,z) = (x\ y\ z)\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
verifica que $D_1 = 1 > 0$, $D_2 = \det\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1 > 0$ y $D_3 = \det A = 0$, luego es semidefinida positiva (y no es definida positiva).
- En el ejemplo siguiente
$$Q(x,y,z) = (x\ y\ z)\begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & -3 \\ 0 & -3 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$
se tiene $D_1 = -1 < 0$, $D_2 = \det\begin{pmatrix} -1 & 1 \\ 1 & -4 \end{pmatrix} = 3 > 0$ y $D_3 = 0$, luego Q es semidefinida negativa y no es definida negativa.

Criterios de clasificación de formas cuadráticas

Nótese que el resultado anterior no da una equivalencia en los últimos casos (en los que se obtiene forma semidefinida). Existe una forma de reconocer cuando una forma cuadrática es semidefinida a partir de la matriz asociada, pero el criterio es algo más complicando que en el caso de formas definidas. Para dar ese criterio usaremos determinantes de otras submatrices.

Definición

Dada una matriz A cuadrada de orden n, se llama menor principal primario de A de orden r, y lo notaremos por H_r al valor del determinante de alguna submatriz de A de orden r que se obtiene a partir de A suprimiendo $n-r$ filas y las $n-r$ columnas que ocupan las mismas posiciones que las filas suprimidas.

Criterios de clasificación de formas cuadráticas

Ejemplos

- Si A es la matriz dada por
$$A = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix},$$
los menores principales primarios de orden 1 son los siguientes
$$H_1^1 = 1, \;\; H_1^2 = 4,$$
y se han obtenido de A suprimiendo la fila 2 y la columna 2 (H_1^1) ó bien suprimiendo la fila 1 y la columna 1 (H_1^2). Sólo hay un menor primario de orden 2, que coincide con $\det A$.

Ejemplos

- Si A es la matriz dada por

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 5 & 4 & 6 \\ 8 & 9 & 0 \end{pmatrix},$$

los menores principales primarios de orden 1 son los siguientes

$$H_1^1 = 1, \quad H_1^2 = 4, \quad H_1^3 = 0\ .$$

Los de orden 2 se obtienen suprimiendo una fila y una columna (3-3, 2-2 ó 1-1); son, por tanto,

$$H_2^{12} = det \begin{pmatrix} 1 & 2 \\ 5 & 4 \end{pmatrix} = -6, \qquad H_2^{13} = det \begin{pmatrix} 1 & 3 \\ 8 & 0 \end{pmatrix} = -24,$$

$$H_2^{23} = det \begin{pmatrix} 4 & 6 \\ 9 & 0 \end{pmatrix} = -36\ .$$

Sólo hay un menor primario de orden 3, que coincide con det A.

Criterios de clasificación de formas cuadráticas

Teorema

Sea Q una forma cuadrática en $\mathbb{R}^n$ y A la matriz cuadrada de orden n asociada. Entonces se verifica que

- Q es semidefinida positiva si, y sólo si, $H_i \geq 0, \forall 1 \leq i \leq n$.
- Q es semidefinida negativa si, y sólo si, $(-1)^i H_i \geq 0, \forall 1 \leq i \leq n$.

Ejemplos

- La forma cuadrática en $\mathbb{R}^3$ dada por

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

verifica que la matriz asociada tiene determinante 0. Por tanto, no puede ser definida. En este caso, estudiamos los menores principales, empezando por los de orden 1, que son,

$$H_1^1 = 1, \quad H_1^2 = 4, \quad H_1^3 = 4 \ .$$

Los de orden 2 son en este caso

$$H_2^{12} = det \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} = 0 \quad H_2^{13} = det \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix} = 4,$$

$$H_2^{23} = det \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix} = 16 \ .$$

Sólo hay un menor primario de orden 3, que coincide con $\det A = 0$. Como todos los menores primarios son no negativos, entonces Q es semidefinida positiva.

Ejemplos

- Consideramos ahora la forma cuadrática en $\mathbb{R}^3$ dada por

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

que verifica que la matriz asociada tiene determinante 0. Por tanto, no puede ser definida. Como tiene menores principales primarios de orden 1 de distinto signo, entonces no puede ser semidefinida, luego es indefinida, ya que

$$H_1^1 = -1,\ \ H_1^2 = -1,\ \ H_1^3 = 2\ .$$

Tema 5. Espacios métricos. Espacios normados.

Análisis Matemático I
1^o Licenciatura de Estadística

Definición

Sea E un conjunto no vacío. Una distancia (o métrica) en E es una función $d : E \times E \longrightarrow \mathbb{R}_0^+$ que verifica:

- $d(x, y) = 0 \Leftrightarrow x = y$ (no degeneración).
- $d(x, y) = d(y, x)$, $\forall x, y \in E$ (propiedad simétrica).
- $d(x, z) \leq d(x, y) + d(y, z)$, $\forall x, y, z \in E$ (desigualdad triangular).

Al par ordenado (E, d) se le denomina espacio métrico.

Ejemplos

- Si $E = \mathbb{R}$, la función

$$d(x, y) = |x - y| \quad (x, y \in \mathbb{R})$$

define una distancia en $\mathbb{R}$.

- Si $1 \leq p < \infty$, definimos

$$d_p(x, y) = \left(\sum_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}} \quad (x, y \in \mathbb{R}^n) \, .$$

La función d_p es una distancia en $\mathbb{R}^n$. La comprobación es inmediata para $p = 1$, mientras que en el resto de los casos, la no degeneración y la simetría son inmediatas, pero la comprobación de la desigualdad triangular requiere más trabajo.
Para $p = 2$, la distancia d_2 se llama distancia euclídea en $\mathbb{R}^n$.

Ejemplos

- La función dada por

$$d_\infty(x, y) = \max\{|x_i - y_i| : 1 \leq i \leq n\} \quad (x, y \in \mathbb{R}^n)$$

  es una distancia en $\mathbb{R}^n$. Para comprobarlo basta usar las propiedades del valor absoluto.
- Si E es un conjunto no vacío, la función

$$d(x, y) = \begin{cases} 0 & \text{si } x = y \\ 1 & \text{si } x \neq y \end{cases} \qquad (x, y \in E)$$

  define una distancia en E (la distancia discreta).

Ejemplos

- Sea $E = C[a, b]$ (el espacio de las funciones reales continuas en el intervalo $[a, b]$ y

$$d_\infty(f, g) = \max\{|f(t) - g(t)| : t \in [a, b]\} \quad (f, g \in E) \, .$$

Es inmediato comprobar que d_∞ es una distancia en E

Observación

Si (E, d) es un espacio métrico, se verifica que

$$|d(x,y) - d(x,z)| \leq d(y,z), \quad \forall x, y, z \in E .$$

Demostración.
Sean x, y, z elementos de E. Basta usar que en vista de la desigualdad triangular de d se verifica que

$$d(x,y) \leq d(x,z) + d(z,y),$$

y usando que d es simétrica, y, por tanto, $d(z,y) = d(y,z)$, obtenemos

$$d(x,y) - d(x,z) \leq d(y,z) .$$

Como la desigualdad anterior se verifica para elementos arbitrarios x, y, z en E, cambiando y por z y usando de nuevo la simetría, tenemos que

$$d(x,z) - d(x,y) \leq d(z,y) = d(y,z) .$$

Como consecuencia de las dos desigualdades anteriores, tenemos que

$$|d(x,y) - d(x,z)| \leq d(y,z), \quad \forall x, y, z \in E .$$

Observaciones

- Si (E, d) es un espacio métrico, cualquier subconjunto $\varnothing \neq A \subset E$ es un espacio métrico con la restricción de la distancia d.
- Si (X, d_1) e (Y, d_2) son dos espacios métricos, entonces la función d definida como sigue es una distancia en $X \times Y$

$$d((x, y), (u, v)) = \max\{d_1(x, u), d_2(y, v)\} \quad (\forall x, u \in X, y, v \in Y)$$

Bolas abiertas y cerradas

Sea (E, d) un espacio métrico y $a \in E$.

- Si $r > 0$, la bola abierta de centro a y radio r es el conjunto dado por
$$B(a, r) := \{x \in E : d(x, a) < r\},$$
- Si $r \geq 0$, bola cerrada de centro a y radio r es el conjunto dado por
$$\overline{B}(a, r) := \{x \in E : d(x, a) \leq r\},$$

Espacios métricos

Conjuntos abiertos y cerrados

Sea (E, d) un espacio métrico y $A \subset E$.

- Diremos que A es abierto si para cada elemento $a \in A$, existe un positivo r tal que $B(a, r) \subset A$. Equivalentemente, un conjunto abierto es una unión de bolas abiertas.
- Se dice que un punto $x \in E$ está en la adherencia de A (o clausura de A) si se verifica la siguiente condición:

  $$\forall r > 0, B(x, r) \cap A \neq \varnothing \ .$$

  Es decir, x es adherente a A si A contiene elementos arbitrariamente próximos al elemento x. Se suele notar por $\overline{A}$ a la adherencia de A. Siempre se verifica que $A \subset \overline{A}$.
- Diremos que A es cerrado si $\overline{A} \subset A$, esto es, si A contiene todos los elementos que están arbitrariamente próximos a elementos de A.

Proposición

Sea (E, d) un espacio métrico y $A \subset E$. Entonces A es abierto si, y sólo si, $E \backslash A$ es cerrado, donde

$$E \backslash A := \{x \in E : x \notin A\} \ .$$

Demostración.
Supongamos que $C = E \backslash A$ es cerrado, esto es $C = \overline{C}$. Probaremos que A es abierto. Sea $a \in A$, luego $a \notin C$, por tanto $a \notin \overline{C}$. Como consecuencia, existe $r > 0$, tal que $B(a, r) \cap C \neq \varnothing$. Por tanto, como obviamente $B(a, r) \subset E$ y $B(a, r) \cap C \neq \varnothing$, entonces $B(a, r) \subset E \backslash C = A$. Hemos probado que A es abierto.
Recíprocamente, supongamos que A es abierto y probaremos que C es cerrado, esto es, que $\overline{C} \subset C$. Sea por tanto $x \in \overline{C}$; si $x \in A$, entonces, por ser A abierto, existe $r > 0$, tal que $B(x, r) \subset A$, por tanto, $B(x, r) \cap C = \varnothing$, y x no sería adherente a C, lo que contradice nuestra hipótesis. Hemos obtenido que $x \notin A$, como obviamente $x \in E$, entonces $x \in E \backslash A = C$. Hemos probado que $\overline{C} \subset C$, luego C es cerrado.

Espacios métricos

Usando la definición de conjunto abierto, y el hecho de que los conjuntos abiertos son los complementarios de los cerrados, es inmediato probar el siguiente enunciado:

Proposición

Sea (E, d) un espacio métrico.

- El total (E) y el vacío son abiertos y cerrados.
- La familia de los conjuntos abiertos es estable por uniones arbitrarias, esto es, la unión de una familia de conjuntos abiertos es un conjunto abierto.
- La intersección de dos conjuntos abiertos es un abierto.
- La familia de los conjuntos cerrados es estable por intersecciones arbitrarias, esto es, la intersección de una familia de conjuntos cerrados es un conjunto cerrado.
- La unión de dos conjuntos cerrados es un cerrado, por tanto, la familia de los conjuntos cerrados es estable por uniones finitas.

Ejemplos

- Los conjuntos que contienen un único elemento son conjuntos cerrados. Un conjunto finito es siempre cerrado.
- Las bolas abiertas son conjuntos abiertos y las bolas cerradas son conjuntos cerrados.
- Si $E = \mathbb{R}^2$ dotado de la métrica euclídea (o de cualquiera de las métricas d_p (para algún $1 \leq p \leq \infty$), los conjuntos siguientes

  $$A = \{(x, y) \in \mathbb{R}^2 : x > 0\}, \qquad B = \{(x, y) \in \mathbb{R}^2 : x < y\}$$

  y

  $$D = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 < 20\}$$

  son abiertos.

Ejemplos

- Si $E = \mathbb{R}^2$ dotado de la métrica euclídea, los conjuntos siguientes
$$C = \{(x,y) \in \mathbb{R}^2 : x = 0\}, \quad E = \{(x,y) \in \mathbb{R}^2 : x \leq y\}$$
y
$$F = \{(x,y) \in \mathbb{R}^2 : 0 \leq x^2 \leq 20\}$$
son cerrados.
- Si $E = \mathbb{R}^2$ dotado de la métrica euclídea (o de cualquiera de las métricas d_p (para algún $1 \leq p \leq \infty$), el conjunto
$$G = \{(x,y) \in \mathbb{R}^2 : 0 < x^2 \leq 20\}$$
no es abierto ni cerrado.
- En general, conjuntos descritos por un número finito de desigualdades estrictas donde intervienen funciones continuas son abiertos (son intersecciones finitas de abiertos).
- Sin embargo, conjuntos descritos por un número finito de desigualdades no estrictas o igualdades donde intervienen funciones continuas son cerrados.

Sucesión convergente

Sea (E, d) un espacio métrico y $\{x_n\}$ una sucesión en E. Diremos que $\{x_n\}$ es una sucesión convergente si existe un elemento $x \in E$ que verifica

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : n \geq m \Rightarrow d(x_n, x) < \varepsilon,$$

equivalentemente, si $\{d(x_n, x)\} \to 0$. Si se verifica la condición anterior, diremos que $\{x_n\}$ converge a x y en tal caso escribiremos $\{x_n\} \to x$.
Es fácil comprobar (ejercicio) que el elemento x que verifica la condición de convergencia es único y se llama límite de la sucesión $\{x_n\}$, y escribiremos $x = \lim\{x_n\}$.

En espacios métricos, muchos de los conceptos y condiciones (como conjunto abierto, cerrado, función continua, etc.) se pueden expresar en términos de la convergencia de sucesiones, que depende, obviamente de la métrica del espacio. A continuación veremos un ejemplo de estas caracterizaciones.

Caracterización secuencial de la adherencia

Sea (E, d) un espacio métrico, $A \subset E$ y $x \in E$. Entonces $x \in \overline{A}$ si, y sólo si, existe una sucesión $\{a_n\}$ de elementos de A convergente a x.

Demostración.
(Ejercicio).

Definición

Una sucesión $\{x_n\}$ de elementos de un espacio métrico (E, d) es de Cauchy si se verifica

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : p, q \geq m \Rightarrow d(x_p, x_q) \leq \varepsilon,$$

equivalentemente,

$$\forall \varepsilon > 0, \ \exists m \in \mathbb{N} : \ [n \geq m, \ p \in \mathbb{N}] \Rightarrow d(x_{n+p}, x_n) \leq \varepsilon.$$

Espacios métricos

Observación

Es inmediato comprobar que en un espacio métrico toda sucesión convergente es de Cauchy.
$\mathbb{Q}$ es un ejemplo de que el recíproco de esta afirmación no es cierto.

Espacio métrico completo

Un espacio métrico es completo si las sucesiones de Cauchy en ese espacio son convergentes.

Ejemplos

- $\mathbb{R}$ es completo (con la distancia asociada al valor absoluto) y $\mathbb{Q}$ no lo es.
- De hecho, $(\mathbb{R}^n, d_p)$ es completo, para todo $1 \leq p \leq \infty$. Además las sucesiones convergentes en $\mathbb{R}^n$ son aquellas tales que todas sus sucesiones coordenadas convergen en $\mathbb{R}$. Lo mismo ocurre para las sucesiones de Cauchy.
- $(C[a, b], d_\infty)$ es completo.

Continuidad en espacios métricos

Para funciones reales de variable real, una función $f : \mathbb{R} \longrightarrow \mathbb{R}$ es continua en un punto $a \in \mathbb{R}$ si se verifica la siguiente condición:

$$\forall \varepsilon > 0 \exists \delta > 0 : x \in \mathbb{R}, |x - a| < \delta \quad \Rightarrow |f(x) - f(a)| < \varepsilon \ .$$

Escribiendo la expresión anterior en términos de la distancia, obtenemos la siguiente reformulación equivalente:

$$\forall \varepsilon > 0 \exists \delta > 0 : x \in \mathbb{R}, d(x, a) < \delta \quad \Rightarrow \quad d(f(x), f(a)) < \varepsilon \ .$$

Para espacios métricos la definición que daremos es ésta última expresión, la única que tiene sentido si disponemos de una distancia d.

Definición

Sean (E, d), (F, ρ) espacios métricos, $f : E \longrightarrow F$ y $a \in E$. Diremos que f es continua en a si se verifica la siguiente condición:

$$\forall \varepsilon > 0, \ \exists \delta > 0 : \left\{ \begin{array}{c} x \in E \\ d(x, a) < \delta \end{array} \right\} \Rightarrow \ \rho(f(x), f(a)) < \varepsilon \ ,$$

$$\forall \varepsilon > 0, \ \exists \delta > 0 : \ f(B(a, \delta)) \subset B(f(a), \varepsilon).$$

f es continua en E si lo es en cada punto de E.

Continuidad

Ejemplos

- Toda función constante entre dos espacios métricos es continua.
- Si (E, d) es un espacio métrico, la función identidad en E es continua.
- La función valor absoluto es continua en $\mathbb{R}$.
- Toda aplicación lineal de $\mathbb{R}^n$ en $\mathbb{R}^m$ es continua, considerando en $\mathbb{R}^n$ y en $\mathbb{R}^m$ cualquiera de las distancias d_p.

Proposición (caracterización secuencial de la continuidad)

Sean (E, d), (F, ρ) espacios métricos, $f : E \longrightarrow F$ y $a \in E$. Entonces f es continua en a si, y sólo si, se verifica la siguiente condición:

$$\forall \{x_n\} \to a, x_n \in E, \forall n \quad \Rightarrow \quad \{f(x_n)\} \to f(a) \ .$$

Demostración.
Supongamos que f es continua en a. Sea $\{x_n\}$ una sucesión en E convergente hacia a. Fijemos $\varepsilon > 0$ y tomemos $\delta > 0$ tal que

$$[x \in E, \ d(x, a) < \delta] \Rightarrow \rho(f(x), f(a)) < \varepsilon.$$

Como $\{x_n\} \to a$, existe un natural m tal que si $n \geq m$ se verifica que $d(a_n, a) < \delta$, y por tanto

$$\rho(f(a_n), f(a)) < \varepsilon.$$

Hemos probado que $\{f(x_n\} \to f(a)$.

Supongamos que f no es continua en a.
Entonces existe $\varepsilon_0 > 0$ verificando la siguiente condición:

$$\forall \delta > 0, f(B(a, \delta)) \not\subseteq B(f(a), \varepsilon_0) \ .$$

Para cada natural n, aplicamos la condición anterior para $\delta = \frac{1}{n}$, luego existe un elemento $x_n \in B(a, \frac{1}{n})$ tal que $\rho(f(x_n), f(a) \geq \varepsilon_0$.

Por la forma de elegir la sucesión $\{x_n\}$ tenemos que $d(x_n, a) < \frac{1}{n}$ para cada n, luego $\{x_n\} \to a$ y, sin embargo, $\{f(x_n)\}$ no converge a $f(a)$, ya que $\rho(f(x_n), f(a)) \geq \varepsilon_0$ para cada natural n.

Hemos probado que si f no es continua en a, entonces no conserva las sucesiones convergentes hacia a. Como consecuencia, si f conserva las sucesiones convergentes hacia a, entonces f es continua en a.

Teorema (caracterización de la continuidad global)

Sean (E, d), (F, ρ) espacios métricos y $f : E \longrightarrow F$. Equivalen las siguientes afirmaciones:

- f es continua, esto es,

$$\forall a \in E, \ \forall \varepsilon > 0, \ \exists \delta_a > 0 : \left\{ \begin{array}{c} x \in E \\ d(x, a) < \delta_a \end{array} \right\} \Rightarrow \ \rho(f(x), f(a)) < \varepsilon.$$

- La imagen inversa por f de cualquier abierto de F es un abierto de E:
- La imagen inversa por f de cualquier cerrado de F es un cerrado de E:
- La función f aplica valores adherentes de cualquier subconjunto de E en valores adherentes de la imagen de dicho conjunto, esto es:

$$f(\overline{A}) \subset \overline{f(A)}, \ \forall A \subset E.$$

Corolario

Sea (E, d) un espacio métrico, $f : E \longrightarrow \mathbb{R}$ una aplicación continua y $r \in \mathbb{R}$. Entonces

- El conjunto $\{x \in E : f(x) < r\}$ es abierto.
- El conjunto $\{x \in E : f(x) > r\}$ es abierto.
- El conjunto $\{x \in E : f(x) \geq r\}$ es cerrado.
- El conjunto $\{x \in E : f(x) \leq r\}$ es cerrado.
- El conjunto $\{x \in E : f(x) = r\}$ es cerrado.

Ejemplos

- Si $E = \mathbb{R}^2$ dotado de la métrica euclídea (o de cualquiera de las métricas d_p (para algún $1 \leq p \leq \infty$), el conjunto

$$D = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 < 20\}$$

  es abierto, por ser intersección de los dos conjuntos siguientes

$$A = \{(x, y) \in \mathbb{R}^2 : 0 < x^2\}, \quad B = \{(x, y) \in \mathbb{R}^2 : x^2 < 20\}$$

  y cada uno de ellos es abierto, por ser imagen inversa mediante la función $f(x, y) = x^2 \quad ((x, y) \in \mathbb{R}^2)$ (que es continua) de un intervalo abierto.

- Si $E = \mathbb{R}^2$, el conjunto

$$F = \{(x, y) \in \mathbb{R}^2 : 0 \leq x^2 \leq 20\}$$

  es cerrado (intersección de dos cerrados).

Espacios normados

Es bien conocido que la función valor absoluto $|\ |:\mathbb{R}\longrightarrow\mathbb{R}_0^+$ verifica las siguientes condiciones:

- $x\in\mathbb{R},\ |x|=0 \quad\Leftrightarrow\quad x=0.$
- $|xy|=|x||y|, \quad \forall x,y\in\mathbb{R}.$
- $|x+y|\leq|x|+|y|, \quad \forall x,y\in\mathbb{R}.$

Una norma en un espacio vectorial es una función real que verifica las tres condiciones anteriores.

Norma y espacio normado

Sea X un espacio vectorial sobre un cuerpo $\mathbb{K}$ ($\mathbb{R}$ ó $\mathbb{C}$). Una función $\|\ \|:X\longrightarrow\mathbb{R}_0^+$ es una norma en X si verifica

- Si $x\in X, \|x\|=0 \quad\Rightarrow x=0$ (no degeneración).
- $\|tx\|=|t|\ \|x\|, \quad \forall t\in\mathbb{K}, x\in X$ (homogeneidad)
- $\|x+y\|\leq\|x\|+\|y\|, \quad \forall x,y\in X$ (subaditividad o desigualdad triangular)

Un espacio normado es un espacio vectorial real o complejo, dotado de una norma.

Espacios normados

Para un espacio normado X, notaremos por B_X y S_X a la bola cerrada unidad y a la esfera unidad, respectivamente, esto es,

$$B_X := \{x \in X : \|x\| \leq 1\}, \quad S_X := \{x \in X : \|x\| = 1\} .$$

Es sencillo probar el siguiente enunciado:

Proposición

Si $(X, \| \ \|)$ es un espacio normado, la aplicación $d : X \times X \longrightarrow \mathbb{R}_0^+$ dada por

$$d(x, y) = \|x - y\| \qquad (x, y \in X)$$

es una distancia en X.

A partir de ahora consideraremos a todo espacio normado como un espacio métrico dotado de la distancia definida en la proposición anterior.

Espacios normados

Proposición

Sea X un espacio normado, entonces se verifica:

- $\big|\|x\| - \|y\|\big| \leq \|x-y\|, \ \ \forall x, y \in X$.
- La suma en X es continua.
- El producto por escalares es continuo.

Demostración.

- $$\big|\|x\| - \|y\|\big| \leq \|x-y\|, \ \ \forall x, y \in X$$

  Usando la desigualdad triangular de la norma, obtenemos que

  $$\|x\| = \|(x-y)+y\| \leq \|x-y\| + \|y\|,$$

  de donde

  $$\|x\| - \|y\| \leq \|x-y\| \ . \tag{4}$$

  Intercambiando los papeles de x e y, y usando la homogeneidad de la norma, tenemos

  $$\|y\| - \|x\| \leq \|y-x\| = \|x-y\| \ . \tag{5}$$

  De las desigualdades (4) y (5) se deduce la que pretendíamos probar.

- La suma es continua
  Sea $(x_0, y_0) \in X \times X$; comprobaremos que la suma es continua en (x_0, y_0). Usando la definición, hemos de probar que

  $$\forall \varepsilon > 0 : \exists \delta > 0,\ x, y \in X, \|x - x_0\| < \delta, \|y - y_0\| < \delta \ \Rightarrow \|(x+y)-(x_0+y_0)\| < \varepsilon$$

  Dado $\varepsilon > 0$, tomamos $\delta = \frac{\varepsilon}{2}$. Entonces si $x, y \in X$ y $\|x - x_0\| < \delta, \|y - y_0\| < \delta$ tenemos que

  $$\|(x + y) - (x_0 + y_0)\| \leq \|x - x_0\| + \|y - y_0\| < \delta + \delta = \varepsilon \, .$$

  Hemos probado que la suma es continua en un elemento arbitrario de $X \times X$.

Espacios normados

- El producto por escalares es continuo

  Sea $(t_0, x_0) \in \mathbb{K} \times X$; comprobaremos que el producto es continuo en (t_0, x_0); , esto significa que

$$\forall \varepsilon > 0 : \exists \delta > 0,\ t \in \mathbb{K}, x \in X, |t - t_0| < \delta, \|x - x_0\| < \delta \ \Rightarrow \|tx - t_0 x_0\| < \varepsilon \,.$$

Dado $\varepsilon > 0$, tomamos $\delta = \min\{\frac{\varepsilon}{2(1+|t_0|)}, \frac{\varepsilon}{2\|x_0\|+1}, 1\}$. Entonces si $t \in \mathbb{K}$, $x \in X$, $|t - t_0| < \delta$ y $\|x - x_0\| < \delta$, obtenemos que

$$\|tx - t_0 x_0\| \leq \|(tx - tx_0) + (tx_0 - t_0 x_0)\| \leq$$

$$\|tx - tx_0\| + \|tx_0 - t_0 x_0\| \leq$$

$$|t|\ \|x - x_0\| + |t - t_0|\|x_0\| \leq$$

$$(|t_0| + \delta)\delta + \delta\|x_0\| <$$

$$(|t_0| + 1)\delta + \frac{\varepsilon}{2} \leq$$

$$\frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \,.$$

Luego el producto es continuo.

Ejemplos

- $(\mathbb{R}, |\ |)$ es un espacio normado.
- En el espacio $\mathbb{R}^N$, es inmediato comprobar que las funciones siguientes son normas:

$$\|x\|_1 := \sum_{i=1}^{N} |x(i)|, \quad \|x\|_\infty := \max\{|x(i)| : 1 \leq i \leq N\} \quad (x \in \mathbb{R}^N)$$

- En el espacio $\mathbb{R}^N$, definimos para $1 \leq p < \infty$ la siguiente generalización de $\|\ \|_1$ (y de la norma euclídea):

$$\|x\|_p := \left(\sum_{i=1}^{N} |x(i)|^p\right)^{\frac{1}{p}} \qquad (x \in \mathbb{R}^N).$$

  La función anterior es una norma y al espacio $\mathbb{R}^N$ dotado de la norma anterior se le suele notar por ℓ_p^N. Es inmediato comprobar la no degeneración y la homogeneidad de la aplicación anterior. Resulta más complicado comprobar la desigualdad triangular (salvo en los casos $p = 1, 2$).

Ejemplos

- La versión infinito-dimensional de los espacios del apartado anterior son los espacios que se suelen notar por ℓ_p. Si $1 \leq p < \infty$, se considera el conjunto

$$\ell_p := \left\{ x : \mathbb{N} \longrightarrow \mathbb{R} : \sum_n |x(n)|^p \text{ converge} \right\}.$$

Usando la desigualdad triangular de la norma del espacio ℓ_p^N, es fácil comprobar que el conjunto anterior es un espacio vectorial y que la función

$$\|x\|_p := \left(\sum_{n=1}^{\infty} |x(n)|^p \right)^{\frac{1}{p}} \qquad (x \in \ell_p)$$

es una norma.

Espacios normados

Ejemplos

- El espacio ℓ_∞ (real) es el conjunto de las sucesiones de reales acotadas con la norma dada por

$$\|x\|_\infty = \sup_n |x(n)| \qquad (x \in \ell_\infty) .$$

  c_0 es el subespacio de ℓ_∞ de las sucesiones convergentes a cero, con la norma que hereda de ℓ_∞.

- Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida y $1 \leq p < \infty$, el espacio $\mathcal{L}_p(\mu)$ viene dado por

$$\mathcal{L}_p(\mu) := \{f : \Omega \longrightarrow \mathbb{R} : f \text{ es medible y } \int_\Omega |f|^p < \infty\}$$

  que es un espacio vectorial. Definimos la función dada por

$$\|f\|_p := \left(\int_\Omega |f|^p \, d\mu\right)^{\frac{1}{p}} \quad (f \in \mathcal{L}_p(\mu)) .$$

Ejemplos

- La función anterior es homogénea (inmediato) y verifica la desigualdad triangular, pero puede valer cero en funciones no nulas. De hecho

$$N(\mu) := \{f \in \mathcal{L}_p(\mu) : \|f\|_p = 0\} = \{f : \Omega \longrightarrow \mathbb{R} : f = 0 \text{ c.p.d.}\}$$

En este caso, se define el espacio vectorial $L_p(\mu) = \mathcal{L}_p(\mu)/N(\mu)$. Los elementos son clases de equivalencia de la forma $f + N(\mu)$ y se verifica que $f + N(\mu) = g + N(\mu)$ cuando $f - g \in N(\mu)$. En $L_p(\mu)$ se define la suma y el producto por reales de manera natural (usando la suma y el producto usual sobre representantes de cada clase). La función $\| \ \|$ induce una norma en el cociente (que notaremos igual) definiendo

$$\|f + N(\mu)\|_p := \left(\int_\Omega |f|^p \, d\mu\right)^{\frac{1}{p}} \quad (f \in \mathcal{L}_p(\mu)) .$$

Se verifica que $(L_p(\mu), \| \ \|_p)$ es un espacio normado.

Ejemplos

- Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida, una función $f : \Omega \longrightarrow \mathbb{R}$ es esencialmente acotada si es medible y existe $M > 0$ tal que

  $$\mu(\{t \in \Omega : |f(t)| > M\}) = 0 \ .$$

  En tal caso diremos que M es una cota esencial de f

  Se define el espacio $\mathcal{L}_\infty(\mu)$ como el espacio de las funciones reales en Ω esencialmente acotadas y la función dada por

  $$\|f\|_\infty := \inf\{M \geq 0 : M \text{ es cota esencial de } f\} \qquad (f \in \mathcal{L}_\infty(\mu)) \ .$$

  Es fácil probar que el ínfimo anterior es, de hecho, un mínimo y que el espacio $\mathcal{L}_\infty(\mu)$ es un espacio vectorial.
  La función anterior es homogénea y verifica la desigualdad triangular, y

  $$\{f \in \mathcal{L}_\infty(\mu) : \|f\|_\infty = 0\} = N(\mu) \ .$$

  Pasando al espacio cociente $L_\infty(\mu) := \mathcal{L}_\infty(\mu)/N(\mu)$, se obtiene un espacio normado.

Ejemplos

- Si $K \subset \mathbb{R}^n$ es un compacto (cerrado y acotado), entonces definimos en el espacio $\mathcal{C}(K)$ de las funciones reales definidas en K que son continuas, la función dada por

  $$\|f\| = \max\{|f(t)| : t \in K\} \qquad (f \in \mathcal{C}(K)) \ ,$$

  que es una norma en $\mathcal{C}(K)$.

Espacios normados

Proposición

Si $1 \leq p \leq q < \infty$, se verifica que

$$\|x\|_\infty \leq \|x\|_q \leq \|x\|_p \leq \|x\|_1 \leq N\|x\|_\infty \qquad \forall x \in \mathbb{R}^N .$$

Equivalentemente,

$$B_\infty \subset NB_1, \quad B_1 \subset B_p \subset B_q \subset B_\infty,$$

donde B_p es la bola unidad cerrada de centro 0 y radio 1 para la norma $\| \ \|_p$ en $\mathbb{R}^N$ para $1 \leq p \leq \infty$.

Demostración.
Sea $x \in \mathbb{R}^N$, entonces

$$\|x\|_\infty = \max\{|x(i)| : 1 \leq i \leq n\} \leq$$

$$\left(\sum_{i=1}^{n} |x(i)|^q\right)^{\frac{1}{q}} = \|x\|_q .$$

Espacios normados

Probamos ahora la desigualdad entre las normas $\| \ \|_p$ y $\| \ \|_q$.
Para ello, usando la homogeneidad de la norma, probamos primero que $B_p \subset B_q$ para $p \leq q$. Si $x \in B_p$, entonces $\sum_{i=1}^N |x(i)|^p \leq 1$, luego $|x(i)| \leq 1$ para todo i, de donde $|x(i)|^q \leq |x(i)|^p$ y obtenemos que

$$\sum_{i=1}^N |x(i)|^q \leq \sum_{i=1}^N |x(i)|^p \leq 1.$$

En consecuencia, $\|x\|_q \leq 1$ y $x \in B_q$. Ahora comprobaremos que $\|x\|_q \leq \|x\|_p$ si $x \in \mathbb{R}^N$, desigualdad que es clara si $x = 0$. Si $x \neq 0$, el elemento $y = \dfrac{x}{\|x\|_p}$ verifica que $\|y\|_p = \dfrac{\|x\|_p}{\|x\|_p} = 1$, luego $y \in B_p \subset B_q$, esto es,

$$\frac{\|x\|_q}{\|x\|_p} = \|y\|_q \leq 1,$$

equivalentemente,

$$\|x\|_q \leq \|x\|_p$$

como queríamos probar.

La desigualdad $\|x\|_p \leq \|x\|_1$ para $1 \leq p < \infty$ es un caso particular de la que acabamos de probar.

Por último, si $x \in \mathbb{R}^N$, es claro que

$$\|x\|_1 = \sum_{k=1}^{N} |x(i)| \leq$$

$$\sum_{k=1}^{N} \max\{|x(i)| : 1 \leq i \leq N\} \leq$$

$$N\|x\|_\infty \, .$$

Las inclusiones entre las bolas unidad para las distintas normas se deducen de manera inmediata de las desigualdades entre las distintas normas en $\mathbb{R}^N$ que acabamos de probar.

Merece la pena recalcar que si X es un espacio vectorial con dos normas $\|\ \|$ y $|||\ |||$ tales que se verifica

$$\|x\| \leq |||\, x\, |||\,, \qquad \forall x \in X\ , \tag{6}$$

entonces se tiene la siguiente inclusión para las bolas unidad correspondientes

$$B_{|||\ |||} \subset B_{\|\ \|}\ ,$$

siendo esta última condición equivalente a (6).

Algo un poco más general (de comprobación inmediata): si $M \geq 0$, se obtiene que

$$\|x\| \leq M\, |||\, x\, |||\,, \qquad \forall x \in X \quad \Leftrightarrow \quad B_{|||\ |||} \subset M B_{\|\ \|}\ .$$

Ahora relacionaremos los distintos espacios de sucesiones ℓ_p.

Proposición

Si $1 \leq p \leq q \leq \infty$, entonces $\ell_p \subset \ell_q$ y además se verifica que

$$\|x\|_q \leq \|x\|_p, \quad \forall x \in \ell_p \ .$$

Demostración.

Supongamos primero que $q = \infty$ y que $p < q$. Si $x \in \ell_p$, entonces la serie $\sum_n |x(n)|^p$ es convergente, luego la sucesión $\{x(n)\} \to 0$, por tanto está acotada, luego $x \in \ell_\infty$. Además, es claro que

$$\|x\|_\infty = \sup\{|x(n)| : n \in \mathbb{N}\} \leq$$

$$\left(\sum_{n=1}^{\infty} |x(n)|^p\right)^{\frac{1}{p}} = \|x\|_p \ .$$

Espacios normados

Supongamos ahora que $1 \leq p \leq q < \infty$. Si $x \in \ell_p$, entonces $\sum_n |x(n)|^p$ es convergente, luego existe un natural N verificando que

$$n \geq N \;\Rightarrow\; |x(n)| < 1 \; .$$

Por ser $p \leq q$, se obtiene que $n \geq N$, entonces $|x(n)|^q \leq |x(n)|^p$, luego, aplicando el criterio de comparación para series, por ser $\sum_n |x(n)|^p$ convergente, también lo es la serie $\sum_n |x(n)|^q$ es convergente, luego $x \in \ell_q$.
Ahora comprobamos la desigualdad

$$\|x\|_q \leq \|x\|_p, \forall x \in \ell_p \; . \tag{7}$$

Por la homogeneidad de las normas, basta comprobar la desigualdad (7) en la esfera unidad para alguna de las dos normas.
En efecto, sea $x \in B_{\ell_p}$, entonces $\sum_{n=1}^{\infty} |x(n)|^p \leq 1$, luego $|x(n)| \leq 1$ para cada n, por tanto, al ser $p \leq q$, tenemos

$$|x(n)|^q \leq |x(n)|^p, \forall n \;\Rightarrow\; \sum_{n=1}^{\infty} |x(n)|^q \leq \sum_{n=1}^{\infty} |x(n)|^p \leq 1 \; .$$

Hemos probado que $\|x\|_q \leq 1$, que es (7) en este caso.
La homogeneidad de las normas permite concluir la demostración.

Para espacios de medida finita, la situación es bien distinta a la de los espacios de sucesiones, como muestra el siguiente enunciado:

Proposición

Sea $(\Omega, \mathcal{A}, \mu)$ un espacio de medida y supongamos que μ es finita. Si $1 \leq p \leq q \leq \infty$, entonces $\mathcal{L}_q(\mu) \subset \mathcal{L}_p(\mu)$ y además si $q < \infty$ se verifica que

$$\|f\|_p \leq \Big(\mu(\Omega) + 1\Big)^{\frac{1}{p}} \|f\|_q, \quad \forall f \in \mathcal{L}_q(\mu)$$

y

$$\|f\|_p \leq \Big(\mu(\Omega)\Big)^{\frac{1}{p}} \|f\|_\infty, \quad \forall f \in \mathcal{L}_\infty(\mu)$$

Demostración.
Supongamos primero que $q = \infty$ y que $p \neq \infty$. Si $f \in \mathcal{L}_\infty(\mu)$, entonces tenemos que

$$\int_\Omega |f|^p \, d\mu \leq \int_\Omega \|f\|_\infty^p \, d\mu \leq \|f\|_\infty^p \mu(\Omega) < \infty \ ,$$

por tanto, $f \in \mathcal{L}_p(\mu)$ y además $\|f\|_p \leq \|f\|_\infty \Big(\mu(\Omega)\Big)^{\frac{1}{p}}$.

Supongamos ahora que $1 \leq p \leq q < \infty$. Si $f \in \mathcal{L}_q(\mu)$ y $\|f\|_q \leq 1$, consideramos los conjuntos

$$P := \{t \in \Omega : |f(t)| \leq 1\}, \quad G := \{t \in \Omega : |f(t)| > 1\},$$

que son medibles, disjuntos y verifican $P \cup N = \Omega$.
Obviamente tenemos que

$$\int_\Omega |f|^p \, d\mu = \int_P |f|^p \, d\mu + \int_G |f|^p \, d\mu \leq$$

$$\mu(P) + \int_G |f|^q \, d\mu \leq$$

$$\mu(\Omega) + \Big(\|f\|_q\|\Big)^q \leq \mu(\Omega) + 1$$

Como consecuencia, en vista de la homogeneidad de la norma, se obtiene que $\mathcal{L}_q(\mu) \subset \mathcal{L}_p(\mu)$ y además

$$\|f\|_p \leq \Big(\mu(\Omega) + 1\Big)^{\frac{1}{p}} \|f\|_q, \qquad \forall f \in \mathcal{L}_q(\mu) \; .$$

Espacios normados

Definición

Un espacio de Banach es un espacio normado completo, esto es, la distancia asociada a la norma es completa.

Prácticamente todos los ejemplos de normados que hemos visto son espacios de Banach. He ahí una lista de algunos:

Ejemplos

- $(\mathbb{R}, |\ |)$ es un espacio de Banach.
- Si $1 \leq p \leq \infty$, $(\mathbb{R}^N, \|\ \|_p)$ es un espacio de Banach.
- Si $1 \leq p \leq \infty$, $(\ell_p, \|\ \|_p)$ es un espacio de Banach.
- $(c_0, \|\ \|_\infty)$ es un espacio de Banach.
- Si $(\Omega, \mathcal{A}, \mu)$ es un espacio de medida y $1 \leq p \leq \infty$, $(L_p(\mu), \|\ \|_p)$ es un espacio de Banach.
- Si $K \subset \mathbb{R}^n$ es un compacto (cerrado y acotado), $(\mathcal{C}(K), \|\ \|_\infty)$ es un espacio de Banach.

Espacios normados

Ejemplos

- $(c_{00}, \| \ \|_\infty)$ no es un espacio de Banach, donde

$$c_{00} := \{x : \mathbb{N} \longrightarrow \mathbb{R} : \exists N \in \mathbb{N}, x(k) = 0 \forall k \geq N\} \ .$$

Espacios normados

Daremos una caracterización de la complitud en espacios normados que a veces resulta útil y que enunciamos a continuación.

Definición

Si X es un espacio normado, una serie $\sum x_n$ es X es absolutamente convergente si la serie $\sum \|x_n\|$ es convergente.

Proposición

Un espacio normado X es completo si, y sólo si, toda serie absolutamente convergente en X es convergente.

Idea de la demostración:

1. Para probar que una sucesión de Cauchy es convergente, basta que una subsucesión suya lo sea. Si $\{x_n\}$ es una sucesión en X, una subsucesión de $\{x_n\}$ es una sucesión de la forma $\{x_{\sigma(n)}\}$, donde σ es una aplicación de $\mathbb{N}$ en $\mathbb{N}$ estrictamente creciente.
2. Toda sucesión $\{a_n\}$ en un espacio vectorial puede verse como la sucesión de sumas parciales de una serie. Basta tomar $x_1 := a_1, \ x_{n+1} := a_{n+1} - a_n$ para $n \in \mathbb{N}$, entonces $\sum_{k=1}^{n} x_i = a_k$ para cada natural k.

Espacios normados

Para aplicaciones lineales entre espacios normados, la continuidad se puede reformular de varias formas sencillas. Para dar esta caracterización, recordamos el siguiente concepto:

Definición

Si X es un espacio normado, un subconjunto $A \subset X$ es acotado si está contenido en una bola, equivalentemente si existe $M \geq 0$ tal que

$$\|a\| \leq M, \qquad \forall a \in A\,.$$

Ejemplos

Si $(X, \|\ \|)$ es un espacio normado, B_X es un conjunto acotado, mientras que si $x_0 \in X, x_0 \neq 0$, el conjunto

$$R := \{tx_0 : t \in \mathbb{R}\}$$

no lo es.

Proposición

Sean X e Y espacios normados y $T : X \longrightarrow Y$ una aplicación lineal. Equivalen las siguientes condiciones:

i) Existe una constante $M > 0$ tal que $\|T(x)\| \leq M\|x\|$, $\forall x \in X$.

ii) T es lipschitziana, es decir, existe una constante $C > 0$ tal que $\|T(x) - T(y)\| \leq C\|x - y\|$ para todo $x, y \in X$.

iii) T es continua.

iv) T es continua en 0.

v) Si A es un subconjunto acotado de X, entonces $T(A)$ es un subconjunto acotado de Y.

vi) T está acotada en B_X, esto es, $T(B_X)$ es acotado.

Demostración.

Por ser T lineal, es claro que i) $\Rightarrow$ ii). Claramente toda aplicación lipschitziana es continua, luego continua en cero.

iii) $\Rightarrow$ iv)

Sea $A \subset X$ un conjunto acotado e $\{y_n\}$ una sucesión en $T(A)$, entonces si para cada natural n, x_n es un elemento de A que verifica $T(x_n) = y_n$, obtenemos que para cada sucesión de reales $\{a_n\}$ convergente a cero, entonces $\{a_n x_n\}$ converge a cero en X. Como T es continua en cero, entonces $\{a_n y_n\} = \{T(a_n x_n)\}$ también converge a cero. Como la conclusión anterior es válida para cualquier sucesión de reales $\{a_n\}$ convergente a cero, entonces $\{y_n\}$ es acotada. Luego $T(A)$ es acotado.

Es claro que v) es un caso particular de iv).

v) $\Rightarrow$ i) Sea $M > 0$ tal que $\|T(x)\| \leq M$, para cada $x \in B_X$. Entonces, si $z \in X$ es un vector no nulo, tendremos que $\frac{z}{\|z\|} \in B_X$, luego $\left\|T\left(\frac{z}{\|z\|}\right)\right\| \leq M$, esto es, $\|T(z)\| \leq M\|z\|$, para cada $z \in X$.

Espacios normados

Definición

Si X e Y son dos espacios normados, notamos por $L(X, Y)$ al espacio de los operadores lineales y continuos de X en Y, al que dotaremos de la norma dada por

$$\|T\| := \sup_{x \in B_X} \|Tx\| \qquad (T \in L(X, Y)) .$$

Cuando $X = Y$, escribiremos simplemente $L(X)$ en lugar de $L(X, X)$.

En el caso particular de que Y sea el cuerpo base, entonces notaremos por X^* al espacio $L(X, \mathbb{K})$, el dual topológico de X.

Usando la homogeneidad de la norma, es inmediato comprobar que
$\|Tx\| \leq \|T\|\|x\|$, para cada $T \in L(X, Y)$ y cada $x \in X$.
De hecho, $\|T\|$ es el ínfimo de todos los reales no negativos K que verifican
$\|Tx\| \leq K\|x\|$, para todo $x \in X$.

Corolario

Sean X, Y, Z espacios normados. Si $T \in L(X, Y), S \in L(Y, Z)$, entonces $S \circ T \in L(X, Z)$ y $\|S \circ T\| \leq \|S\|\|T\|$.

Demostración.
Obviamente la composición de aplicaciones lineales y continuas es lineal y continua y tenemos que

$$\|(S \circ T)(x)\| \leq \|S\|\|Tx\| \leq$$

$$\|S\|\|T\|x\|,$$

para cada $x \in X$, luego $\|S \circ T\| \leq \|S\|\|T\|$.

Ejemplo

- Consideramos el operador $T : \ell_\infty \longrightarrow \ell_\infty$ definido por

$$T(x)(1) = 0, \quad T(x)(n+1) = x(n), \quad \forall n \in \mathbb{N} \qquad (x \in \ell_\infty)\ .$$

  Es inmediato que $Tx \in \ell_\infty$ para cada $x \in \ell_\infty$, T es lineal y $\|Tx\|_\infty = \|x\|_\infty$, para cada $x \in \ell_\infty$, luego T es continuo y $\|T\| = 1$. T es el operador de desplazamiento hacia la derecha en ℓ_∞.

- Si $y \in \ell_1$, definimos $f : c_0 \longrightarrow \mathbb{R}$ por

$$f(x) := \sum_{n=1}^{\infty} y(n)x(n) \qquad (x \in c_0)\ .$$

  f es un funcional lineal y continuo en c_0 y se verifica que

$$\|f\| = \|y\|_1\ .$$

Definición

Sea X un espacio vectorial. Un subconjunto $E \subset X$ es equilibrado si verifica

$$x \in E, t \in \mathbb{K}, |t| \leq 1 \;\Rightarrow\; tx \in E\ .$$

Por ejemplo, una bola centrada en cero es un conjunto equilibrado, mientras que el intervalo $[0, 1] \subset \mathbb{R}$ no es equilibrado. Para funcionales lineales, a las condiciones que caracterizan la continuidad de aplicaciones lineales podemos añadir esta otra:

Proposición

Sea X un espacio normado y $f : X \to \mathbb{K}$ lineal. Entonces f es continuo si, y sólo si, $\mathrm{Ker}\, f$ es cerrado.

Demostración.
Obviamente, si f es continuo, entonces $\operatorname{Ker} f = f^{-1}(\{0\})$ es cerrado.

Supongamos que $\operatorname{Ker} f$ es cerrado. Para probar la continuidad de f, basta comprobar que f está acotado en la bola unidad.
Como B_X es equilibrado y f lineal, entonces $f(B_X)$ es equilibrado. Si f no estuviese acotado en la bola unidad, por ser $f(B_X)$ equilibrado, tendríamos $f(B_X) = \mathbb{K}$. Por tanto, la imagen de cualquier abierto no vacío por f es el cuerpo base. Por tanto, $\operatorname{Ker} f$ es denso en X, esto es, $\overline{\operatorname{Ker} f} = X$. Por ser $\operatorname{Ker} f$ cerrado (hipótesis), tenemos $X = \operatorname{Ker} f$, luego f es nulo, pero entonces es trivialmente continuo, luego acotado en la bola unidad.
Hemos probado que si $\operatorname{Ker} f$ es cerrado, f está acotado en la bola unidad, luego es continuo.

Ejemplos

- En el espacio $\mathcal{C}[a,b]$, definimos el funcional φ dado por

$$\varphi(f) = \int_a^b f(t)\, dt \qquad (f \in \mathcal{C}[a,b]) \ ,$$

  que es obviamente lineal y verifica

$$|\varphi(f)| \leq (b-a)\|f\|_\infty, \qquad \forall f \in \mathcal{C}[a,b],$$

  luego φ es continuo y es fácil comprobar que $\|\varphi\| = b - a$.

- En el subespacio c_{00} de las sucesiones casinulas de c_0, el funcional lineal dado por

$$F(x) = \sum_{i=1}^{\infty} x(i) \qquad (x \in c_{00}),$$

  es discontinuo, ya que si llamamos $x_n := \sum_{i=1}^{n} e_i \in B_{c_{00}}$, para cada natural n, entonces $F(x_n) = n$, luego F no está acotado en la bola unidad.

Ejemplos

- Para cada sucesión $a \in \ell_\infty$, y cada sucesión $x \in \ell_1$, la serie $\sum a(n)x(n)$ es convergente y el funcional dado por

$$F(x) = \sum_{n=1}^{\infty} a(n)x(n) \qquad (x \in \ell_1),$$

es lineal y continuo, ya que

$$|F(x)| = \Big|\sum_{n=1}^{\infty} a(n)x(n)\Big| \leq$$

$$\sum_{n=1}^{\infty} |a(n)x(n)| \leq \|a\|_\infty \sum_{n=1}^{\infty} |x(n)| =$$

$$\|a\|_\infty \|x\|_1 \ ,$$

para cada $x \in \ell_1$, luego F es continuo.

Ejemplos

- No es difícil describir las isometrías que permiten conseguir las identificaciones siguientes:

$$c_0^* \equiv \ell_1, \;\; \ell_1^* \equiv \ell_\infty, \;\; \ell_p^* \equiv \ell_q \text{ si } \frac{1}{p} + \frac{1}{q} = 1 \;\; (1 < p < \infty) \,.$$

El hecho de que c_0 sea isométrico a ℓ_1 significa que existe una aplicación lineal $T : \ell_1 \longrightarrow c_0^*$ tal que T es biyectiva y además verifica

$$\|Tx\| = \|x\|_1, \qquad \forall x \in \ell_1 \,.$$

En este caso concreto, se puede usar la aplicación T definida por

$$T(x)(y) = \sum_{n=1}^{\infty} x(n)y(n) \quad (x \in \ell_1, y \in c_0) \,.$$

Espacios normados

Probaremos al menos que en dimensión finita todas las normas son equivalentes.

Definición

Sea X un espacio vectorial y $\|\ \|, ⦀\ ⦀$ dos normas en X. Diremos que $\|\ \|$ y $⦀\ ⦀$ son equivalentes si existen constantes $m, M > 0$ verificando

$$m\|x\| \leq ⦀ x ⦀ \leq M\|x\|, \forall x \in X .$$

Ejemplos

- En $\mathbb{R}^N$ las normas $\|\ \|_1$ y $\|\ \|_\infty$ son equivalentes (de hecho todas las normas $\|\ \|_p$ son equivalentes para $1 \leq p \leq \infty$), ya que sabemos que

$$\|x\|_\infty \leq \|x\|_1 \leq N\|x\|_\infty, \qquad \forall x \in \mathbb{R}^N .$$

Espacios normados

Ejemplos

- En $C[0,1]$, la norma usual y la norma $\| \ \|_1$ dada por

$$\|f\|_1 := \int_0^1 |f(t)|\, dt \qquad (f \in C[0,1])$$

no son equivalentes. Es claro que

$$\|f\|_1 := \int_0^1 |f(t)|\, dt \le \|f\|_\infty, \qquad \forall f \in C[0,1]\ .$$

Sin embargo, no es cierta una desigualdad del tipo

$$\|f\|_\infty \le M\|f\|_1, \qquad \forall f \in C[0,1]\ .$$

para ningún positivo M.

Ejemplos

- Supongamos que

$$\|f\|_\infty \leq M\|f\|_1, \qquad \forall f \in C[0,1]\,. \tag{8}$$

para algún positivo M. Para cada natural n, la función dada por

$$f_n(t) = t^n \qquad (t \in [0,1])$$

es continua, verifica que $\|f_n\|_\infty = 1$ y, sin embargo,

$$\|f_n\|_1 = \int_0^1 t^n\,dt = \frac{t^{n+1}}{n+1}\Big]_0^1 = \frac{1}{n+1}\,.$$

Si la desigualdad anterior entre las normas fuese cierta, tendríamos

$$1 \leq M\frac{1}{n+1}, \quad \forall n \in \mathbb{N}\,,$$

y esto es imposible.

Espacios normados

Proposición

Sea X un espacio vectorial y $\| \ \|, ||| \ |||$ dos normas en X. Entonces existe una constante $M > 0$ tal que

$$||| x ||| \leq M \|x\|, \qquad \forall x \in X$$

si, y sólo si, todo conjunto abierto en $(X, ||| \ |||)$ es abierto en $(X, \| \ \|)$.

Demostración.

Supongamos que existe una constante $M > 0$ que verifica

$$||| x ||| \leq M \|x\|, \qquad \forall x \in X$$

Como el operador identidad $I : (X, \| \ \|) \longrightarrow (X, ||| \ |||)$ es lineal, la desigualdad anterior equivale a la continuidad de I. Por la caracterización (general) de la continuidad, I es continua si, y solo si, la imagen inversa por I de un abierto en $(X, ||| \ |||)$ es abierto en $(X, \| \ \|)$, esto es, si todo abierto de $(X, ||| \ |||)$ es abierto en $(X, \| \ \|)$.

Corolario

En un espacio normado X, dos normas son equivalentes si, y sólo si, los conjuntos abiertos de X coinciden para ambas normas.

Teorema (Hausdorff)

En un espacio normado finito-dimensional dos normas cualesquiera son equivalentes.

Demostración.
Sea X un espacio vectorial de dimensión N. Fijamos una base $\{e_1, \ldots, e_n\}$ de X y definimos

$$\left\| \sum_{i=1}^{N} x(i) e_i \right\|_1 := \sum_{i=1}^{N} |x(i)|$$

para cada elemento $x \in \mathbb{K}^N$. Es inmediato que $\| \ \|_1$ es una norma en X. Sea $\| \ \|$ otra norma en X. Probaremos que ésta es equivalente a $\| \ \|_1$.

Espacios normados

Primero, la desigualdad triangular permite comprobar que $\| \ \|$ es continua en $(X, \| \ \|_1)$, ya que para cualesquiera $x, y \in X$ tenemos

$$\|x - y\| = \left\| \sum_{i=1}^{N} x(i)e_i - \sum_{i=1}^{N} y(i)e_i \right\| \leq$$

$$\sum_{i=1}^{N} |x(i) - y(i)| \|e_i\|_1 \leq$$

$$\max\{\|e_i\|_1 : 1 \leq i \leq N\} \|x - y\|_1 \ .$$

Como la esfera unidad para $\| \ \|_1$ es un compacto de $(X, \| \ \|_1)$, entonces la función $\| \ \|$ alcanza su mínimo y su máximo en dicha esfera. Si llamamos $m := \min\{\|x\| : x \in X, \|x\|_1 = 1\}$, y $M := \max\{\|x\| : x \in X, \|x\|_1 = 1\}$, entonces se tiene obviamente que

$$m \leq \|x\| \leq M, \quad \forall x \in X, \|x\|_1 = 1.$$

Por la homogeneidad de la norma, las desigualdades anteriores implican que

$$m\|x\|_1 \leq \|x\| \leq M\|x\|_1, \quad \forall x \in X \ .$$

Como es claro que m, M son reales positivos, hemos probado que $\| \ \|$ y $\| \ \|_1$ son normas equivalentes.

Corolario

Sean X e Y espacios normados, tal que X es finito-dimensional. Si T es una aplicación lineal de X en Y, entonces T es continua.

Demostración.
La aplicación $|\!|\!|\ |\!|\!|$ dada por

$$|\!|\!| x |\!|\!| := \|x\| + \|Tx\| \qquad (x \in X),$$

es una norma en X, luego equivalente a $\|\ \|$, por tanto T es continuo.

Corolario

Dos espacios normados finito-dimensionales de la misma dimensión son isomorfos, esto es, existe una aplicación entre ambos que es lineal, biyectiva, continua y con inversa continua.

Demostración.
Si X e Y son espacios normados de dimensión N, entonces existe una aplicación lineal T biyectiva entre ellos. Por el corolario anterior, T y T^{-1} son continuas, luego X e Y son isomorfos.

Espacios normados

Es fácil probar el siguiente resultado.

Proposición

Sea (E, d) un espacio métrico completo y $A \subset E$. Entonces A es cerrado si, y sólo si, A es completo.

Corolario

Todos los espacios normados de dimensión finita son completos, luego todo subespacio vectorial finito-dimensional de un normado es cerrado.

Demostración.

Si X es un espacio normado de dimensión N, entonces existe una aplicación lineal y biyectiva $T : X \longrightarrow \mathbb{R}^N$. Por el corolario anterior, T y T^{-1} son continuos, por tanto, existen constantes $m, M > 0$ tales que

$$m\|Tx\|_2 \leq \|x\| \leq M\|Tx\|_2, \quad \forall x \in X .$$

Como $(\mathbb{R}^N, \|\ \|)$ es completo, de la desigualdad anterior, se deduce que X es completo, ya que $\{x_n\} \to x$ si, y sólo si, $\{Tx_n\} \to Tx$ y $\{x_n\}$ es de Cauchy en X si, y sólo si, $\{Tx_n\}$ es de Cauchy en $\mathbb{R}^N$.

Espacios normados

Aplicando el mismo argumento usado en la demostración anterior se obtiene lo siguiente:

Proposición

Sea X un espacio vectorial y $\| \ \|$, $\|| \ \||$ dos normas en X tales que existe una constante $M > 0$ que verifica

$$\|x\| \leq M \||\, x \,\|| \,, \qquad \forall x \in X \,.$$

Sea $\{x_n\}$ una sucesión en X. Se verifica que

1. Si $\{x_n\}$ converge en $(X, \|| \ \||)$, entonces es también convergente en $(X, \| \ \|)$
2. Si $\{x_n\}$ es de Cauchy en $(X, \|| \ \||)$, entonces es de Cauchy en $(X, \| \ \|)$

Demostración.

Comprobamos la relación entre las sucesiones convergentes. Si $\{x_n\}$ converge a x en $(X, \|| \ \||)$, entonces, $\{\||\, x_n - x \,\||\} \to 0$, y de la desigualdad

$$0 \leq \|x_n - x\| \leq M \||\, x_n - x \,\|| \,, \qquad \forall n \in \mathbb{N} \,,$$

obtenemos que $\{\|x_n - x\|\} \to 0$, luego $\{x_n\}$ converge en $(X, \| \ \|)$.