Álgebra Lineal.

♣

Grado de Economía.

Alfredo Bautista Santa-Cruz

Versión 27-5-2016

Alfredo Bautista Santa-Cruz
Dpto. Análisis Económico: Economía Cuantitativa - Universidad Autónoma de Madrid

Índice general

Capítulo 1

Espacios vectoriales

1.1. Introducción

El conjunto $\mathbb{R}$ no es suficiente para describir o representar ciertas magnitudes:

- Fuerza del viento
- Velocidad de una corriente de agua
- En general, cualquier magnitud que necesite indicar una dirección

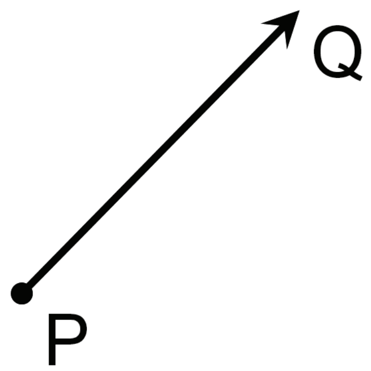
Necesitaremos dotar a $\mathbb{R}^n$ de una estructura que permita el manejo de tales direcciones.

1.1.1. Concepto de vector como segmento orientado

Dados dos puntos $P, Q \in \mathbb{R}^n$ definimos el vector

$$\overline{PQ}$$

como el segmento <u>orientado</u> que une P con Q.



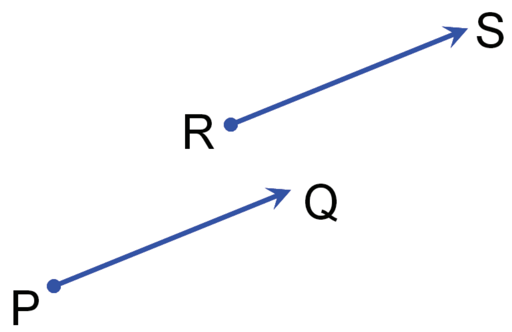
Como $\overline{PQ}$ está orientado entonces

$$\overline{PQ} \neq \overline{QP}$$

Además consideraremos que

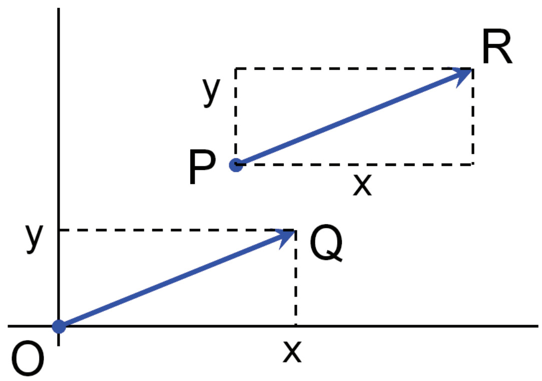
$$\overline{PQ} = \overline{RS}$$

siempre que los segmentos tengan la misma longitud y estén orientados en la misma dirección y sentido.



De esta manera, dado un vector $\overline{PR}$ existirá un punto Q tal que el vector $\overline{OQ}$ verifica que $\boxed{\overline{PR} = \overline{OQ}}$, donde O es el origen de coordenadas.

Ahora podemos definir las **componentes** de un vector de $\mathbb{R}^n$ como los valores de las proyecciones del punto Q sobre los ejes.



Los valores de x e y son las componentes del vector $\overline{OQ}$ (y por tanto, también de $\overline{PR}$).

El módulo de un vector representa la longitud del segmento y se define de la siguiente manera:

Si $(x_1, \ldots, x_n)$ son las componentes de un vector $\overline{OQ}$ entonces utilizando el Teorema de Pitágoras definimos el **módulo** de $\overline{OQ}$ como

$$|\overline{OQ}| = \sqrt{x_1^2 + \ldots + x_n^2}$$

1.1.2. Concepto de vector como n-upla a partir de sus componentes

Como todos los segmentos orientados tienen, como vectores, las mismas componentes, podemos definir los vectores como elementos de $\mathbb{R}^n$.

Así $(x_1, \ldots, x_n) \in \mathbb{R}^n$ puede verse como un segmento orientado además de como un punto.

Hay que tener en cuenta que la orientación implica que

$$(a, b) \neq (b, a) \neq (-b, a) \neq \cdots$$

1.1.3. Operaciones en un espacio vectorial

Dados dos vectores $\overline{OQ_1}$ y $\overline{OQ_2}$ de $\mathbb{R}^n$ con componentes $(x_1, \ldots, x_n)$ y $(y_1, \ldots, y_n)$ respectivamente, se define:

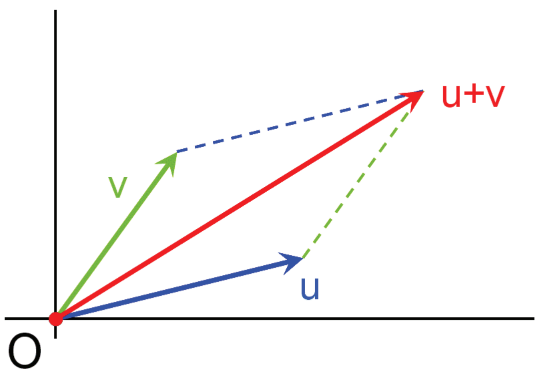
(+) El vector suma de $\overline{OQ_1}$ y $\overline{OQ_2}$ como

$$\overline{OQ_1} + \overline{OQ_2} = (x_1 + y_1, \ldots, x_n + y_n)$$

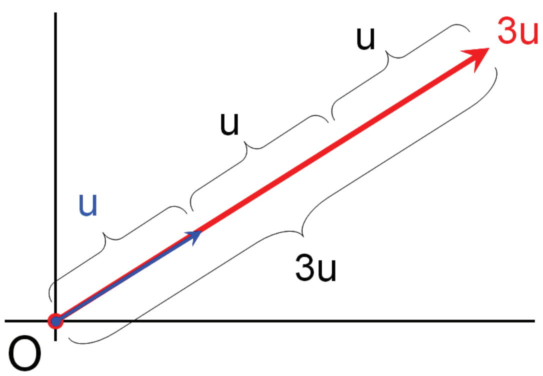
($\cdot$) El vector producto de un escalar $\lambda \in \mathbb{R}$ por $\overline{OQ_1}$ como

$$\lambda \cdot \overline{OQ_1} = (\lambda x_1, \ldots, \lambda x_n)$$

Interpretación geométrica de la suma de vectores



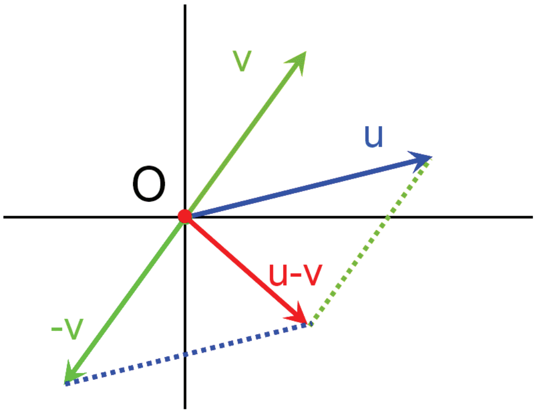
Interpretación geométrica del producto de un vector por un escalar



A partir de estas operaciones podemos definir la diferencia de vectores $\vec{u}$ y $\vec{v}$ como

$$\vec{u} - \vec{v} = \vec{u} + (-1) \cdot \vec{v}$$

cuya interpretación es la siguiente



Proposición 1 *(Propiedades)*

1. *Sea* $\vec{u} \in \mathbb{R}^n$ *un vector cualquiera, entonces se verifica para todo* $\lambda \in \mathbb{R}$ *que*

$$|\lambda \vec{u}| = |\lambda|\,|\vec{u}|$$

2. *Para todo vector* $\vec{u} \in \mathbb{R}^n$ *se tiene que*

$$|\vec{u}| \geq 0$$

   siendo $|\vec{u}| = 0$ *únicamente cuando* $\vec{u} = (0, \ldots, 0)$.

3. *Para todos los vectores* $\vec{u}, \vec{v} \in \mathbb{R}^n$ *se tiene*

$$|\vec{u} + \vec{v}| \leq |\vec{u}| + |\vec{v}|$$

   Esta es la llamada ***Desigualdad triangular***.

Ejercicio 2 *Dados los vectores de* $\mathbb{R}^2$

$$\vec{u} = (1, 2) \qquad y \qquad \vec{v} = (-2, 3)$$

calcular

$$\vec{u} + \vec{v}$$
$$\vec{u} - \vec{v}$$
$$3\vec{u}$$
$$\frac{1}{2}\vec{u} - \vec{v}$$

y representarlos gráficamente.

Vector opuesto

Dado $\vec{u} \in \mathbb{R}^n$ un vector cualquiera, podemos definir su **vector opuesto** $-\vec{u}$ como

$$-\vec{u} = -1 \cdot \vec{u}$$

Vector unitario

Decimos que $\vec{u} \in \mathbb{R}^n$ es un vector **unitario** si se verifica que

$$|\vec{u}| = 1$$

Si un vector $\vec{v} \in \mathbb{R}^n$ no es unitario y verifica que

$$|\vec{v}| = a \neq 0$$

entonces el vector $\frac{1}{a}\vec{v}$ es unitario ya que

$$\left|\frac{1}{a}\vec{v}\right| = \left|\frac{1}{a}\right| |\vec{v}| = \frac{1}{a}a = 1$$

Producto escalar

El producto escalar de dos vectores $\vec{u} = (x_1, \ldots, x_n)$ y $\vec{v} = (y_1, \ldots, y_n)$, que denotaremos como

$$\vec{u} \cdot \vec{v}$$

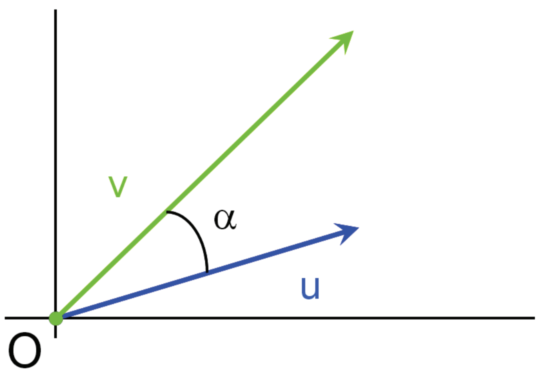
es un *número* que se puede definir como:

1. En función de las componentes:

$$\vec{u} \cdot \vec{v} = x_1 y_1 + \ldots + x_n y_n$$

2. En función del ángulo α que forman los vectores:

$$\vec{u} \cdot \vec{v} = |\vec{u}|\,|\vec{v}| \cos \alpha$$



1.2. Espacios vectoriales

Definición 3 *Un conjunto V no vacío dotado de dos operaciones:*

+) *SUMA de dos elementos de V (operación interna).*

·) *PRODUCTO POR UN ESCALAR de un elemento de V (operación externa).*

Se dice que es un ***espacio vectorial*** *si para todo $u, v, w \in V$ y todo $a, b \in \mathbb{R}$ se verifican las siguientes 8 propiedades:*

Para la SUMA:

(+) **1.** *Asociativa:*

$$(u+v)+w = u+(v+w)$$

(+) **2.** *Conmutativa:*

$$u+v = v+u$$

(+) **3.** *Existe elemento neutro $0_V \in V$:*

$$u+0_V = u$$

(+) **4.** *Existe elemento $-u \in V$ opuesto a u:*

$$u+(-u) = 0_V$$

Para el PRODUCTO POR ESCALARES:

$(\cdot)$ **1.** *Distributiva respecto de la suma de vectores:*

$$a \cdot (u+v) = a \cdot u + a \cdot v$$

$(\cdot)$ **2.** *Distributiva respecto de la suma de escalares:*

$$(a+b) \cdot u = a \cdot u + b \cdot u$$

$(\cdot)$ **3.** *Pseudoasociativa:*

$$(ab) \cdot u = a \cdot (b \cdot u)$$

$(\cdot)$ **4.** *Existe elemento unidad $1 \in \mathbb{R}$:*

$$1 \cdot u = u$$

Ejemplo 4 $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, $\ldots$, $\mathbb{R}^n$ *son espacios vectoriales con las operaciones habituales, es decir, si $\bar{x} = (x_1, \ldots, x_n), \bar{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ y $a \in \mathbb{R}$ entonces las operaciones son:*

1. *SUMA: $\bar{x} + \bar{y} = (x_1 + y_1, \ldots, x_n + y_n)$*

2. *PRODUCTO POR ESCALARES: $a \cdot \bar{x} = (ax_1, \ldots, ax_n)$*

Ejemplo 5 $\mathrm{P}(x)$ *el conjunto de los polinomios de una variable x con las siguientes operaciones. Si $p(x) = \alpha_0 + \alpha_1 x + \cdots + \alpha_n x^n$ y $q(x) = \beta_0 + \beta_1 x + \cdots + \beta_k x^k$ son polinomios y $a \in \mathbb{R}$ entonces las operaciones son:*

1. *SUMA: $(p+q)(x) = p(x) + q(x)$*

2. *PRODUCTO POR ESCALARES: $(a \cdot p)(x) = ap(x) = a\alpha_0 + a\alpha_1 x + \cdots + a\alpha_n x^n$*

Ejemplo 6 $\mathrm{P}_k(x)$ *el conjunto de los polinomios de una variable x de grado menor o igual que k con las mismas operaciones que el espacio vectorial $\mathrm{P}_k(x)$ es un espacio vectorial.*

Ejemplo 7 $\overline{\mathrm{P}_k}(x)$ *el conjunto de los polinomios de una variable x de grado exactamente igual que k con las mismas operaciones que el espacio vectorial $\mathrm{P}_k(x)$* ***no*** *es un espacio vectorial, ya que si $p(x) = \alpha x^k$ y $q(x) = -\alpha x^k$ son dos polinomios de grado exactamente k pero su suma*

$$(p+q)(x) = p(x) + q(x) = \alpha x^k + \left(-\alpha x^k\right) = 0$$

no es un polinomio de grado exactamente igual a k.

Proposición 8 *(Propiedades) Sea V un espacio vectorial*

1. *Si $u, v, w \in V$ tales que $u + v = u + w$ entonces $v = w$.*

2. *Si 0_V es el elemento neutro de V y $a \in \mathbb{R}$ entonces $a \cdot 0_V = 0_V$.*

3. *Si $u \in V$ y $0 \in \mathbb{R}$ entonces $0 \cdot u = 0_V$.*

4. *Si $u \in V$ entonces $(-1) \cdot u = -u$ el elemento opuesto de u.*

5. *Si $u \in V$ y $a \in \mathbb{R}$ tales que $a \cdot u = 0_V$ entonces $u = 0_V$ o $a = 0$.*

Proposición 9 *Sean V_1 y V_2 dos espacios vectoriales y sea*

$$V = V_1 \times V_2 = \{(v_1, v_2) : v_1 \in V_1, v_2 \in V_2\}$$

su producto cartesiano, entonces se verifica que V es un espacio vectorial con las siguientes operaciones:

$$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2)$$

$$a \cdot (u_1, u_2) = (a \cdot u_1, a \cdot u_2)$$

donde $u_1, v_1 \in V_1$ y $u_2, v_2 \in V_2$ y además $a \in \mathbb{R}$.

*Al espacio vectorial V se le denomina **espacio vectorial producto**.*

Ejemplo 10 *Así*

$$\mathbb{R} \times \mathbb{R}^3$$

$$\mathbb{R}^2 \times \mathrm{P}(x)$$

$$\mathbb{R}^7 \times \mathrm{P}_4(x) \times \mathbb{R} \times \mathrm{P}(x)$$

son espacios vectoriales con las operaciones correspondientes.

1.3. Subespacios vectoriales

Dado un espacio vectorial V ¿cualquier subconjunto suyo W es también espacio vectorial con las mismas operaciones?

A veces sí

$$P_k(x) \subset P(x)$$

y a veces no

$$\overline{P_k}(x) \subset P(x)$$

Definición 11 *Sea V un espacio vectorial y W un subconjunto no vacío de V. Se dice que W es un **subespacio vectorial** de V si W dotado de las mismas operaciones definidas en V es, asu vez, un espacio vectorial. Por lo tanto, para que se verifique que W es subespacio vectorial de V tiene que verificarse que*

1. *$u + v \in W$ para todo $u, v \in W$,*

2. *$a \cdot u \in W$ para todo $u \in W$ y todo $a \in \mathbb{R}$,*

que equivalentemente se pueden resumir en que

$$a \cdot u + b \cdot v \in W$$

para todo $u, v \in W$ y todo $a, b \in \mathbb{R}$.

Observación 12 *Dado un espacio vectorial V siempre existen dos subespacios vectoriales suyos:*

1. *El propio V (como subconjunto de sí mismo).*
2. *$W = \{0_V\}$.*

Ejemplo 13 *Estudiar si los siguientes subconjuntos de $\mathbb{R}^3$ son subespacios vectoriales de $\mathbb{R}^3$:*

$$W_1 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\big\}$$

$$W_2 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 - x_2 = 0\big\}$$

$$W_3 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 \leq 2\big\}$$

$\boxed{W_1 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\big\}}$

Si $u, v \in W_1$ entonces son de la forma

$$u = (x_1, x_2, 0) \qquad\qquad v = (y_1, y_2, 0)$$

y dados $a, b \in \mathbb{R}$ entonces

$$a \cdot u + b \cdot v = a \cdot (x_1, x_2, 0) + b \cdot (y_1, y_2, 0) =$$

$$= (ax_1, ax_2, 0) + (by_1, by_2, 0) =$$

$$= (ax_1 + by_1, ax_2 + by_2, 0) \in W_1$$

ya que su última componente es igual a 0.

$\boxed{W_2 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 - x_2 = 0\big\}}$

Si $u, v \in W_1$ entonces son de la forma

$$u = (x_1, x_1, x_3) \qquad\qquad v = (y_1, y_1, y_3)$$

ya que $x_1 - x_2 = 0$ y por tanto $x_1 = x_2$.

Dados $a, b \in \mathbb{R}$ entonces

$$a \cdot u + b \cdot v = a \cdot (x_1, x_1, x_3) + b \cdot (y_1, y_1, y_3) =$$

$$= (ax_1, ax_1, ax_3) + (by_1, by_1, by_3) =$$

$$= (ax_1 + by_1, ax_1 + by_1, ax_3 + by_3) \in W_2$$

ya que sus dos primeras componentes coinciden.

$\boxed{W_3 = \big\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 \leq 2\big\}}$

Tomemos por ejemplo $u = (-3, x_2, x_3)$ que pertenece a W_3 ya que

$$x_1 = -3 \leq 2$$

pero este elemento de W_3 no tiene su opuesto en W_3 ya que

$$-u = (3, -x_2, -x_3)$$

no pertenece a W_3 ya que su primera componente es estrictamente mayor que 2.

Ejercicio 14 *Sea W un subespacio vectorial de $\mathbb{R}^3$, entonces definimos el conjunto*

$$W^{\perp} = \big\{u \in \mathbb{R}^3 : u \cdot v = 0 \text{ para todo } v \in W\big\}$$

Demostrar que este conjunto es un subespacio vectorial de $\mathbb{R}^3$. A este subespacio le llamaremos ***subespacio ortogonal*** *de W.*

1.3.1. Operaciones entre subespacios vectoriales

Podemos observar mediante los ejemplos anteriores si la **intersección** y la **unión** de subespacios vectoriales es, a su vez, otro subespacio vectorial.

La intersección de los subespacios vectoriales

$$W_1 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\right\}$$

$$W_2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 - x_2 = 0\right\}$$

es la siguiente

$$W_1 \cap W_2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0, x_1 - x_2 = 0\right\}$$

cuyos elementos son de la forma

$$u = (x_1, x_2, x_3) = (x_1, x_1, 0)$$

Así si tenemos $u = (x_1, x_1, 0)$ y $v = (y_1, y_1, 0)$ y además $a, b \in \mathbb{R}$ entonces

$$au + bv = a\,(x_1, x_1, 0) + b\,(y_1, y_1, 0) =$$

$$= (ax_1, ax_1, 0) + (by_1, by_1, 0) = (ax_1 + by_1, ax + by_1, 0)$$

que pertenece a $W_1 \cap W_2$ ya que tiene las dos primeras componentes iguales y la tercera es 0. Por lo tanto $W_1 \cap W_2$ es un subespacio vectorial de $\mathbb{R}^3$.

Proposición 15 *Sea V un espacio vectorial. Si $\{W_i : i \in I\}$ es una familia de subespacios vectoriales de V, entonces se verifica que la intersección*

$$W = \bigcap_{i \in I} W_i$$

también es un subespacio vectorial.

Observación 16 *El índice $i \in I$ es arbitrario, puede ser finito o infinito.*

Volvamos a las operaciones entre conjuntos, en cuanto a la unión

$$W_1 \cup W_2 = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0 \text{ o bien } x_1 - x_2 = 0\right\}$$

tenemos que

$$(2, 3, 0) \in W_1 \cup W_2$$

ya que pertenece a W_1 y

$$(1, 1, 1) \in W_1 \cup W_2$$

ya que pertenece a W_2. Así estudiamos su suma

$$(2, 3, 0) + (1, 1, 1) = (3, 4, 1) \notin W_1 \cup W_2$$

ya que no pertenece a ni a W_1 ni a W_2.

La unión de subespacios vectoriales, en general, NO es subespacio vectorial.

Vamos a definir otra operación entre subespacios: la SUMA de subespacios vectoriales.

Definición 17 *Sea X un conjunto dotado de una operación suma y sean X_i con $i = 1, \ldots, n$ una familia de subconjuntos de X. Se define el* ***conjunto suma*** *como:*

$$\sum_{i=1}^{n} X_i = X_1 + \cdots + X_n = \left\{ u \in X : u = \sum_{i=1}^{n} u_i \quad \textit{con } u_i \in X_i \quad i = 1, \ldots, n \right\}$$

Observación 18 *El índice $i = 1, \dots, n$ tiene que ser finito.*

Con los subespacios del ejemplo anterior

$$W_1 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = 0\}$$
$$W_2 = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 - x_2 = 0\}$$

se tiene que su suma es un espacio vectorial ya que

$$W_1 + W_2 = \mathbb{R}^3$$

porque todo vector $(x_1, x_2, x_3) \in \mathbb{R}^3$ puede escribirse como

$$(x_1, x_2, x_3) = (x_1, x_2, 0) + (0, 0, x_3) \in W_1 + W_2$$

que es suma de un vector de W_1 y otro de W_2.

Siguiendo con este ejemplo vemos que no hay unicidad en la forma de escribir un vector de $W_1 + W_2$ como suma de uno de W_1 y otro de W_2, por ejemplo:

$$(1, 1, 1) \in \mathbb{R}^3 = W_1 + W_2$$

puede escribirse como

$$(1, 1, 1) = (1, 1, 0) + (0, 0, 1) \in W_1 + W_2 = \mathbb{R}^3$$
$$(1, 1, 1) = (2, 2, 0) + (-1, -1, 1) \in W_1 + W_2 = \mathbb{R}^3$$

En general

$$(x_1, x_2, x_3) \in \mathbb{R}^3 = W_1 + W_2$$
$$(x_1, x_2, x_3) = (x_1 - a, x_2 - a, 0) + (a, a, x_3) \in W_1 + W_2 = \mathbb{R}^3$$

para cualquier $a \in \mathbb{R}$.

Para evitar esta falta de unicidad en la expresión hacemos la siguiente

Definición 19 *Sea V un espacio vectorial y sean W_1 y W_2 subespacios vectoriales de V. Se dice que W es* ***suma directa*** *de W_1 y W_2 y se denotará como*

$$W = W_1 \oplus W_2$$

si y sólo si se verifica que

1. $W = W_1 + W_2$
2. $W_1 \cap W_2 = \{0_V\}$

Proposición 20 *Sean W_1 y W_2 subespacios vectoriales de V. El conjunto*

$$W = W_1 + W_2$$

es suma directa de los subespacios W_1 y W_2 si y sólo si para todo $u \in W$ existe un único $v \in W_1$ y un único $w \in W_2$ tale que

$$u = v + w$$

Observación 21 *Si la suma directa contiene a todo el espacio vectorial, es decir,*

$$V = W_1 \oplus W_2$$

entonces diremos que W_1 y W_2 son ***subespacios suplementarios****.*

Así

$$\mathbb{R}^n = W \oplus W^{\perp}$$

para cualquier subespacio vectorial W. Diremos que $W^{\perp}$ es el ***suplemento ortogonal*** *de W.*

1.4. Sistemas de generadores

Al igual que un pintor necesita únicamente tres colores para conseguir cualquier otro color, vamos a buscar un conjunto de finito de vectores con el que podamos construir cualquier otro vector de un espacio vectorial.

Definición 22 *Sea V un espacio vectorial. Se dice que $v \in V$ es* ***combinación lineal*** *de $u_1, \ldots, u_m \in V$ si existen escalares $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$ tales que*

$$v = \alpha_1 u_1 + \ldots + \alpha_m u_m$$

Ejemplo 23 *Sean $u_1 = (1,1,1)$, $u_2 = (1,1,0)$ y $u_3 = (0,0,1)$. El vector*

$$w = (-2,-2,1)$$

se puede ecribir como

$$w = u_1 - 3u_2 + u_3$$

es decir, w es combinación lineal de u_1, u_2 y u_3. Pero el vector

$$v = (0,1,0)$$

no puede escribirse como combinación lineal ya que

$$\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3 = v$$

implica que

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} \alpha_1 + \alpha_2 \\ \alpha_1 + \alpha_2 \\ \alpha_1 + \alpha_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

que no tiene solución.

¿Qué vectores se pueden escribir como combinación lineal de u_1, u_2 y u_3? Todos los vectores (x_1, x_2, x_3) que cumplan que $x_1 = x_2$.

Proposición 24 *Sea V un espacio vectorial y sean $u_1, \ldots, u_m \in V$. El conjunto de vectores de V que pueden ponerse como combinación lineal de $u_1, \ldots, u_m$ es un subespacio vectorial de V. Este subespacio lo denotaremos como*

$$\mathcal{L}\{u_1, \ldots, u_m\}$$

Definición 25 *Sea W un subespacio vectorial de V. Se dice que los vectores $u_1, \ldots, u_m \in W$ son un* ***sistema de generadores*** *de W si y sólo si todos los elementos de W se pueden expresar como combinación lineal de $u_1, \ldots, u_m$.*

Ejemplo 26 *Comprobemos si*

$$G_1 = \{u_1 = (1,0,-1), u_2 = (1,-1,0)\}$$

$$G_2 = \{u_1, u_2, u_3 = (0,1,-1)\}$$

son generadores de

$$W = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\right\}$$

o de

$$V = \mathbb{R}^3$$

Los vectores generados por G_1 son de la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_2 \\ -\alpha_1 \end{pmatrix}$$

es decir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_2 \\ -\alpha_1 \end{pmatrix}$$

por lo que tienen que verificar que

$$x_1 + x_2 + x_3 = 0$$

Esto implica que G_1 genera W, pero no genera V.

Los vectores generados por G_2 son de la forma

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_2 + \alpha_3 \\ -\alpha_1 - \alpha_3 \end{pmatrix}$$

es decir

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_2 \\ -\alpha_2 + \alpha_3 \\ -\alpha_1 - \alpha_3 \end{pmatrix}$$

por lo que tienen que verificar la misma ecuación

$$x_1 + x_2 + x_3 = 0$$

y así G_2 genera W, pero no genera V.

Observación 27 *Si G es un sistema de generadores de W y G' es tal que*

$$G \subset G'$$

entonces G' también es sistema de generadores de W.

1.5. Dependencia e independencia lineal

Sea W un subespacio vectorial, ¿existe algún subconjunto de un sistema de generadores de W que sea sistema de generadores de W?

Definición 28 *Sea V un espacio vectorial y sean $u_1, \ldots, u_m \in V$ con $m > 1$. Se dice que:*

1. *Estos vectores son* ***linealmente dependientes*** *si al menos uno de ellos puede escribirse como combinación lineal de los restantes.*

2. *Estos vectores son* ***linealmente independientes*** *cuando ninguno de ellos puede escribirse como combinación lineal de los restantes.*

Un conjunto de vectores se dice linealmente dependiente (resp. independiente) si los vectores que lo integran son linealmente dependientes (resp. independientes).

Proposición 29 *(Propiedades)*

1. *Si* $m = 1$ *entonces, por convenio, si* $u \in V$ *es tal que* $u \neq 0_V$ *entonces es linealmente independiente.*
2. *No pueden verificarse a la vez 1 y 2 de la definición anterior.*
3. *El vector* 0_V *hace que cualquier conjunto de vectores que lo contenga sea linealmente dependiente.*
4. *Si* M *es linealmente dependiente y* $M \subset M'$ *entonces* M' *es linealmente dependiente.*
5. *Si* M *es linealmente independiente y* $M' \subset M$ *entonces* M' *es linealmente independiente.*
6. *Si* $u \in V$ *es linealmente dependiente (resp. independiente) de un conjunto* $S = \{u_1, \ldots, u_m\} \subset V$ *si* u *puede expresarse (resp. no puede expresarse) como combinación lineal de elementos de* S*.*

Ejemplo 30 *En* $\mathrm{P}_3(x)$ *tenemos los conjuntos*

$$L_1 = \left\{p_1(x) = 1 + x, \quad p_2(x) = x^3, \quad p_3(x) = x^3 - 2x - 2\right\}$$

$$L_2 = \left\{q_1(x) = 1, \quad q_2(x) = 1 + x, \quad q_3(x) = 1 + x^2\right\}$$

Como se tiene que

$$p_3(x) = p_2(x) - 2p_1(x)$$

entonces p_3 *es combinación lineal de* p_1 *y* p_2 *y por tanto* L_1 *es linealmente dependiente.*

Por otro lado, como no existen α *y* β *tales que se verifique que*

$$q_1(x) = \alpha q_2(x) + \beta q_3(x)$$

es decir, que

$$1 = \alpha(1 + x) + \beta\left(1 + x^2\right)$$

entonces q_1 *no es combinación lineal de* q_2 *y* q_3*.*

De igual modo podemos ver que q_2 *no es combinación lineal de* q_1 *y* q_3*, ni* q_3 *es combinación lineal de* q_1 *y* q_2*. Por lo tanto* L_2 *es linealmente independiente.*

Si se incrementa el número de vectores, la definición anterior de dependencia lineal deja de ser operativa. Para decidir si un conjunto es linealmente dependiente o independiente utilizaremos la siguiente proposición.

Proposición 31 *Sea* V *un espacio vectorial y sean* $u_1, \ldots, u_m \in V$*. El conjunto de vectores* $\{u_1, \ldots, u_m\}$ *es*

1. *linealmente dependiente si y sólo si existen* $\alpha_1, \ldots, \alpha_m \in \mathbb{R}$*, no todos nulos, tales que*

$$\alpha_1 u_1 + \ldots + \alpha_m u_m = 0_V$$

2. *linealmente independiente si y sólo para que se verifique*

$$\alpha_1 u_1 + \ldots + \alpha_m u_m = 0_V$$

se tiene que cumplir que $\alpha_i = 0$ *para todo* $i = 1, \ldots, m$*.*

Definición 32 *Dado un conjunto de vectores* $S = \{u_1, \ldots, u_m\}$ *de un espacio vectorial* V*, se denomina* ***rango*** *de* S *al número máximo de vectores de* S *linealmente independientes. Lo denotaremos por* $rg(S)$*.*

Ejemplo 33 *El conjunto de vectores*

$$G = \{(1,0,0), (0,1,0), (1,1,0)\}$$

¿qué subespacio genera?

Los vectores generados por G vendrán dados por

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \alpha_3 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha_1 + \alpha_3 \\ \alpha_2 + \alpha_3 \\ 0 \end{pmatrix}$$

es decir, deben verificar la ecuación

$$\boxed{z = 0}$$

No podemos eliminar los parámetros α_1 y α_2.

Si consideramos el conjunto

$$G' = \{(1,0,0), (0,1,0)\}$$

podemos ver que genera el mismo subespacio vectorial. En efecto,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ 0 \end{pmatrix}$$

es decir, deben verificar la misma ecuación

$$\boxed{z = 0}$$

Proposición 34 *Sea V un espacio vectorial y sea $G = \{u_1, \ldots, u_m\} \subset V$ un sistema de generadores de V. Se verifica que*

1. *Si G es un conjunto de vectores linealmente dependiente, entonces existe al menos un subconjunto $G' \subset G$ tal que $G' \neq G$ que también genera V.*

2. *Si G es un conjunto de vectores linealmente independiente, entonces no existe ningún subconjunto $G' \subset G$ tal que $G' \neq G$ que sea además sistema de generadores de V.*

Ejemplo 35 *En el ejemplo anterior hemos visto conjuntos que verifican esta proposición.*

Proposición 36 *Sea V un espacio vectorial y sean $G = \{u_1, \ldots, u_m\}$ y $G' = \{u'_1, \ldots, u'_{m'}\}$ sistemas de generadores de V. Si los vectores de G' son linealmente independentes, entonces $m' \leq m$.*

1.6. Base y dimensión de un espacio vectorial

¿Cuándo tenemos el mínimo número de vectores que generan un espacio vectorial?
¿Cuándo tenemos el máximo número de vectores linealmente independientes?
La última proposición que hemos visto implica que

n^o de vectores de V linealmente independientes	$\leq$	n^o de vectores en un sistema de generadores de V

Cuando en un sistema de generadores hay dependencia lineal, la representación no es única. En efecto, si $G = \{u_1, u_2, u_3, u_4\}$ es un sistema de generadores de V tal que $u_4 = \alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3$. Si w es un vector tal que

$$w = \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \beta_4 u_4$$

entonces también

$$\begin{aligned} w &= \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \beta_4 u_4 = \\ &= \beta_1 u_1 + \beta_2 u_2 + \beta_3 u_3 + \beta_4 (\alpha_1 u_1 + \alpha_2 u_2 + \alpha_3 u_3) = \\ &= (\beta_1 + \beta_4 \alpha_1) u_1 + (\beta_2 + \beta_4 \alpha_2) u_2 + (\beta_3 + \beta_4 \alpha_3) u_3 \end{aligned}$$

es otra representación.

Cuando en un sistema de generadores no hay esta dependencia lineal, la representación es única.

Definición 37 *Sea $B = \{u_1, \ldots, u_n\}$ un subconjunto de un espacio vectorial V. Se dice que B es una* ***base*** *de V si y sólo si*

1. *B es un sistema de generadores de V.*
2. *B es linealmente independiente.*

Ejemplo 38 *Veamos que $B_1 = \{(1,0), (0,1)\}$ y $B_2 = \{(1,1), (3,1)\}$ son dos bases de $\mathbb{R}^2$.*

Es claro que B_1 genera $\mathbb{R}^2$ ya que

$$v = \begin{pmatrix} x \\ y \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$$

y por tanto tiene solución para cualquier (x, y):

$$\alpha_1 = x \qquad y \qquad \alpha_2 = y$$

Por otro lado, como los únicos valores de α_1 y α_2 para los que se verifica

$$\alpha_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

son

$$\alpha_1 = 0 \qquad y \qquad \alpha_2 = 0$$

entonces B_1 también es linealmente independiente y así B_1 es base.

En el caso de B_2, veamos que es un sistema de generadores de $\mathbb{R}^2$:

$$v = \begin{pmatrix} x \\ y \end{pmatrix} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \alpha_1 + 3\alpha_2 \\ \alpha_1 + \alpha_2 \end{pmatrix}$$

y por tanto tiene solución para cualquier (x, y):

$$\alpha_1 = \frac{1}{2}(-x + 3y) \qquad y \qquad \alpha_2 = \frac{1}{2}(x - y)$$

Por otro lado, como los únicos valores de α_1 y α_2 para los que se verifica

$$\alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

son

$$\alpha_1 = 0 \qquad y \qquad \alpha_2 = 0$$

entonces B_2 también es linealmente independiente y así B_2 es base.

Definición 39 *Sea V un espacio vectorial y $B = \{u_1, \ldots, u_n\}$ una base de V. Se dice que los números $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ son las* ***coordenadas de un vector*** *$v \in V$* ***respecto de la base*** *B y se denotan como $(\alpha_1, \ldots, \alpha_n)_B$ si y sólo si*

$$v = \alpha_1 u_1 + \ldots + \alpha_n u_n$$

Ejemplo 40 *Vamos a calcular las coordenadas del vector $v = (4, 2)$ respecto de las bases*

$$B_1 = \{e_1 = (1, 0), e_2 = (0, 1)\}$$

$$B_2 = \{u_1 = (1, 1), u_2 = (3, 1)\}$$

de $\mathbb{R}^2$ del ejemplo anterior.

Respecto de B_1, como

$$v = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 4 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

entonces v tiene como coordenadas $(4, 2)_{B_1}$ respecto de B_1, es decir

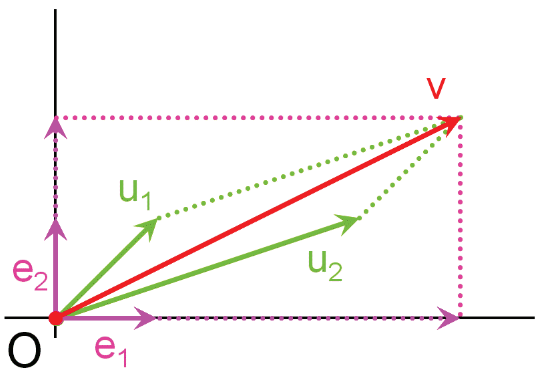
$$v = (4, 2) = (4, 2)_{B_1}$$

Respecto de B_2, como

$$v = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + 1 \begin{pmatrix} 3 \\ 1 \end{pmatrix}$$

entonces v tiene como coordenadas $(1, 1)_{B_2}$ respecto de B_2, es decir

$$v = (4, 2) = (1, 1)_{B_2}$$



Observación 41 *La base B_1 del ejemplo anterior se denomina* ***base canónica*** *de $\mathbb{R}^2$. En general, la base canónica de $\mathbb{R}^n$ es la base*

$$B_C = \{e_1, \ldots, e_n\}$$

con los vectores $e_i = (e_{i1}, \ldots, e_{in})$ dados por

$$e_{ij} = \begin{cases} 1 & \text{si } i = j \\ 0 & \text{si } i \neq j \end{cases}$$

En esta base B_C coinciden las componentes y las coordenadas y los vectores de B_C tienen módulo 1 y son ortogonales dos a dos.

Teorema 42 *Todas las bases de un espacio vectorial engendrado por un número finito de vectores, tienen el mismo número de elementos.*

Este teorema implica que

$$\boxed{\begin{array}{c}\text{Máximo n}^o\text{ de vectores} \\ \text{de } V \text{ linealmente} \\ \text{independientes}\end{array}} = \boxed{\begin{array}{c}\text{Mínimo n}^o\text{ de vectores} \\ \text{de } V \text{ necesarios} \\ \text{para generarlo}\end{array}}$$

Definición 43 *Llamaremos* ***dimensión*** *de un espacio vectorial V (generado por un conjunto finito de vectores) al número de vectores que forman parte de cualquier base de V. Lo denotaremos como*

$$\dim(V)$$

Observación 44

1. $\dim(\mathbb{R}^n) = n$ *ya que la base canónica tiene n elementos.*
2. *Sean V un espacio vectorial y $W \subset V$ un subespacio vectorial de V tal que $\dim(V) = n$ y $\dim(W) = m$ entonces se tiene que*

   $$m = \dim(W) \leq \dim(V) = n$$

   siendo $W = V$ si y sólo si $m = n$.
3. *Si $B_1 = \{u_1, \ldots, u_n\}$ es base de V_1 y $B_2 = \{v_1, \ldots, v_m\}$ es base de V_2 entonces*

   $$B = \{(u_1, 0_{V_2}), \ldots, (u_n, 0_{V_2}), (0_{V_1}, v_1), \ldots, (0_{V_1}, v_m)\}$$

   es base de $V_1 \times V_2$ y además

   $$\dim(V_1 \times V_2) = \dim(V_1) + \dim(V_2) = n + m$$
4. *Si $\dim(V) = n$ entonces todas las bases tienen n elementos y ese es el número máximo de vectores linealmente independientes y el número mínimo de vectores necesarios para generar V.*

Ejemplo 45 *Sea*

$$W = \{(a, a+b, b, 2a-b) : a, b \in \mathbb{R}\} \subset \mathbb{R}^4$$

vamos a encontrar un sistema de generadores de W.

Un vector $(x_1, x_2, x_3, x_4) \in W$ debe cumplir que

$$(x_1, x_2, x_3, x_4) = (a, a+b, b, 2a-b)$$

para algún $a, b \in \mathbb{R}$, entonces

$$(x_1, x_2, x_3, x_4) = (a, a+b, b, 2a-b) =$$

$$= (a, a, 0, 2a) + (0, b, b, -b) =$$
$$= a\,(1, 1, 0, 2) + b\,(0, 1, 1, -1)$$

es decir, (x_1, x_2, x_3, x_4) *es combinación lineal de los vectores* $(1, 1, 0, 2)$ *y* $(0, 1, 1, -1)$*, por lo tanto*

$$G = \{(1, 1, 0, 2), (0, 1, 1, -1)\}$$

es un sistema de generadores de W*.*

Así, podemos expresar que

$$W = \mathcal{L}\{(1, 1, 0, 2), (0, 1, 1, -1)\}$$

Como G *es linealmente independiente, entonces* G *es una base de* W*.*

A continuación vamos a encontrar la expresión analítica de W*, es decir, las ecuaciones que caracterizan a* W*. Como*

$$\left.\begin{array}{l} x_1 = a \\ x_2 = a + b \\ x_3 = b \\ x_4 = 2a - b \end{array}\right\} \Longrightarrow \left.\begin{array}{l} x_2 = x_1 + x_3 \\ x_4 = 2x_1 - x_3 \end{array}\right\}$$

por lo tanto

$$W = \left\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_2 = x_1 + x_3, \quad x_4 = 2x_1 - x_3\right\}$$

¿Podemos formar una base B *de* $\mathbb{R}^4$ *añadiendo vectores a la base* G *de* W*?*

Sí, basta con añadir vectores que no estén en W *y que vayan siendo linealmente independientes con los que vamos teniendo. Por ejemplo, añadiendo* $(1, 0, 0, 0)$ *y* $(0, 1, 0, 0)$ *obtenemos una base de* $\mathbb{R}^4$*, es decir*

$$B_1 = \{(1, 1, 0, 2), (0, 1, 1, -1), (1, 0, 0, 0), (0, 1, 0, 0)\}$$

es base de $\mathbb{R}^4$*, aunque también lo es*

$$B_2 = \{(1, 1, 0, 2), (0, 1, 1, -1), (1, 0, 1, 0), (0, 1, 0, 1)\}$$

Ejercicio 46 *Calcular las coordenadas del vector de* W

$$v = (2, 1, -1, 5)$$

respecto de la base G *como vector de* W *y respecto de las base* B_1 *y* B_2 *como vector de* $\mathbb{R}^4$*.*

Observación 47 *Si* W *es un subespacio vectorial de* V *y* B_W *es una base de* W*, entonces se puede prolongar* B_W *a una base* B *de* V *de manera que* $B_W \subset B$*.*

Llamaremos a B *PROLONGACIÓN de* B_W*. Como hemos visto en el ejemplo anterior, eta prolongación no es única.*

Proposición 48 *Si* W_1 *y* W_2 *son subespacios vectoriales de* V *de dimensión finita, entonces*

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

Si $V = W_1 \oplus W_2$ *entonces como* $W_1 \cap W_2 = \{0_V\}$ *se tiene que*

$$\dim(V) = \dim(W_1) + \dim(W_2)$$

Ejercicio 49 *Dados los subespacios vectoriales de* $\mathbb{R}^3$

$$\begin{array}{l} W_1 = \left\{(x, y, z) \in \mathbb{R}^3 : x = 0\right\} \\ W_2 = \left\{(x, y, z) \in \mathbb{R}^3 : x + y = 0\right\} \\ W_3 = \left\{(x, y, z) \in \mathbb{R}^3 : x - z = 0, \quad y + z = 0\right\} \end{array}$$

estudiar si los subespacios $W_1 + W_2$*,* $W_2 + W_3$ *y* $W_1 + W_3$ *son sumas directas y verificar las ecuaciones de las dimensiones de la proposición anterior.*

Capítulo 2

Matrices y aplicaciones lineales

2.1. Matrices

El sistema

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

tiene 3 tipos de elementos:

- x_j es la variable j-ésima.
- a_{ij} es el coeficiente que multiplica a la variable j-ésima en la i-ésima ecuación.
- b_i es el término independiente de la ecuación i-ésima.

¿Cómo se pueden agrupar estas operaciones para escribirlas de forma más compacta?

Para cada $i = 1, \ldots, m$ podemos denotar $a_{i\bullet} = (a_{i1}, a_{i2}, \ldots, a_{in})$ y $\bar{x} = (x_1, x_2, \ldots, x_n)$. Si utilizamos el producto escalar con estos vectores entonces

$$a_{i\bullet} \cdot \bar{x} = a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i$$

así pues

$$a_{i\bullet} \cdot \bar{x} = b_i$$

De esta manera podemos expresar el sistema como

$$\left\{\begin{array}{c} a_{1\bullet} \cdot \bar{x} = b_1 \\ a_{2\bullet} \cdot \bar{x} = b_2 \\ \vdots \\ a_{m\bullet} \cdot \bar{x} = b_m \end{array}\right.$$

Si llamamos $\bar{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$ entonces

$$\begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{m\bullet} \end{pmatrix} \cdot \bar{x} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = \bar{b}$$

Podemos escribir el sistema anterior como

$$\begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{m\bullet} \end{pmatrix} \cdot \bar{x} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \cdot \bar{x} = \bar{b}$$

por lo que si llamamos

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

entonces el sistema anterior se puede escribir como

$$A\bar{x} = \bar{b}$$

Definición 50 *Dados los números $a_{ij} \in \mathbb{R}$ con $i = 1, \ldots, m$ y $j = 1, \ldots, n$. Al rectángulo de $m \times n$ números ordenados en una tabla con m filas y n columnas de la forma*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

*se le denomina **matriz** de orden $m \times n$.*

Habitualmente denotaremos con letras mayúsculas a las matrices y con letras minúsculas a los elementos que la componen, así podemos escribir

$$A = (a_{ij}) \qquad \text{con } i = 1, \ldots, m \text{ y } j = 1, \ldots, n$$

A veces es conveniente trabajar con filas o con columnas, así escribiremos la i-ésima fila de A como

$$a_{i\bullet} = (a_{i1}, \ldots, a_{in})$$

y la j-ésima columna como

$$a_{\bullet j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

por lo que podemos escribir

$$A = (a_{\bullet 1}, \ldots, a_{\bullet n}) = \begin{pmatrix} a_{1\bullet} \\ \vdots \\ a_{m\bullet} \end{pmatrix}$$

La asignación de los números i y j es fija y, por tanto, también podemos definir el concepto de matriz de la siguiente manera:

Definición 51 *Sea S un conjunto cualquiera y sean $I = \{i : i = 1, \ldots, m\}$ y $J = \{j : j = 1, \ldots, n\}$. Sea*

$$\begin{array}{rccl} f: & I \times J & \to & S \\ & (i,j) & \mapsto & f(i,j) = a_{ij} \end{array}$$

*una aplicación, entonces $A = \{f(i,j) : i \in I, \quad j \in J\} = (a_{ij})$ con $i = 1, \ldots, m$ y $j = 1, \ldots, n$ es una **matriz** de orden $m \times n$, es decir*

$$A = \begin{pmatrix} f(1,1) & \cdots & f(1,n) \\ \vdots & \ddots & \vdots \\ f(m,1) & \cdots & f(m,n) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

*Cuando A tiene m filas y n columnas se dice que tiene **dimensión** u **orden** $m \times n$.*

Al conjunto de matrices de dimensión $m \times n$ lo denotaremos como $\mathcal{M}_{m\times n}$. En particular, si $m = n$ entonces se dice que $\mathcal{M}_{n\times n}$ o abreviando $\mathcal{M}_n$ es el conjunto de matrices cuadradas de orden n.

Observación 52 *Podemos considerar a los vectores como matrices, así el vector fila $u = (u_1, \ldots, u_n)$ es una matriz de orden $1 \times n$ y de igual manera el vector columna $v = \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}$ es una matriz de orden $m \times 1$.*

2.2. Aplicaciones lineales

Entre dos espacios vectoriales V y V' se pueden establecer multitud de correspondencias y aplicaciones. De todas estas correspondencias, vamos a estudiar las que conservan la estructura de espacio vectorial.

Definición 53 *Sean V y V' dos espacios vectoriales. Una aplicación*

$$f : V \to V'$$

*se dice que es una **aplicación lineal** si para cualesquier $u, w \in V$ y $\alpha, \beta \in \mathbb{R}$ se verifica que*

1. $f(u + w) = f(u) + f(w)$
2. $f(\alpha u) = \alpha f(u)$

o de manera equivalente

$$f(\alpha u + \beta w) = \alpha f(u) + \beta f(w)$$

Ejemplo 54 *La aplicación*

$$\begin{array}{rccc} f: & \mathbb{R}^2 & \to & \mathbb{R}^2 \\ & (x_1, x_2) & \mapsto & (e^{x_1}, e^{x_2}) \end{array}$$

no es lineal ya que si

$$u = (u_1, u_2) \qquad\qquad w = (w_1, w_2)$$

entonces

$$f(u + w) = \left(e^{u_1+w_1}, e^{u_2+w_2}\right) \neq \left(e^{u_1} + e^{w_1}, e^{u_2} + e^{w_2}\right) = f(u) + f(w)$$

Ejemplo 55 *La aplicación*

$$\begin{array}{rccc} f: & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \mapsto & (2x_2, x_1 + x_3) \end{array}$$

es lineal ya que si

$$u = (u_1, u_2, u_3) \qquad\qquad w = (w_1, w_2, w_3) \qquad\qquad \alpha, \beta \in \mathbb{R}$$

entonces

$$\begin{aligned} f(\alpha u + \beta w) &= f(\alpha u_1 + \beta w_1, \alpha u_2 + \beta w_2, \alpha u_3 + \beta w_3) = \\ &= (2(\alpha u_2 + \beta w_2), (\alpha u_1 + \beta w_1) + (\alpha u_3 + \beta w_3)) = \\ &= (2\alpha u_2 + 2\beta w_2, \alpha(u_1 + u_3) + \beta(w_1 + w_3)) = \\ &= \alpha(2u_2, u_1 + u_3) + \beta(2w_2, w_1 + w_3) = \\ &= \alpha f(u) + \beta f(w) \end{aligned}$$

Ejemplo 56 *La aplicación* $f(x_1,x_2,x_3,x_4) = (1, x_1+x_2, x_3+x_4)$ *no es lineal ya que si*

$$u = (u_1,u_2,u_3,u_4) \qquad w = (w_1,w_2,w_3,w_4) \qquad \alpha,\beta \in \mathbb{R}$$

entonces

$$\begin{aligned} f(\alpha u + \beta w) &= (1, \alpha u_1 + \beta w_1 + \alpha u_2 + \beta w_2, \alpha u_3 + \beta w_3 + \alpha u_4 + \beta w_4) = \\ &= (1, \alpha(u_1+u_2) + \beta(w_1+w_2), \alpha(u_3+u_4) + \beta(w_3+w_4)) \end{aligned}$$

y como

$$\begin{aligned} \alpha f(u) + \beta f(w) &= \alpha(1, u_1+u_2, u_3+u_4) + \beta(1, w_1+w_2, w_3+w_4) = \\ &= (\alpha+\beta, \alpha(u_1+u_2) + \beta(w_1+w_2), \alpha(u_3+u_4) + \beta(w_3+w_4)) \end{aligned}$$

e igualando las primeras componentes de estos vectores tenemos que

$$\alpha + \beta = 1$$

que no es cierto en general ya que α *y* β *pueden ser cualesquier números.*

Ejemplo 57 *Sea* $f : \mathbb{R}^2 \to \mathbb{R}^3$ *una aplicación lineal tal que*

$$f(-1,2) = (1,3,-2) \qquad y \qquad f(1,4) = (1,-2,5)$$

entonces se puede hallar

$$\begin{gathered} f(0,6) \\ f(2,-4) \\ f(2,2) \end{gathered}$$

y en general se puede calcular

$$f(x_1,x_2)$$

para cualquier (x_1,x_2).

En efecto, como $\{(-1,2),(1,4)\}$ *es una base de* $\mathbb{R}^2$ *entonces, por ser* f *una aplicación lineal, sólo hace falta calcular las coordenadas de los vectores en esta base.*

- *Como*
$$(0,6) = 1 \cdot (-1,2) + 1 \cdot (1,4)$$
entonces
$$f(0,6) = 1 \cdot f(-1,2) + 1 \cdot f(1,4) = (2,1,3)$$
- *Como*
$$(2,-4) = -2 \cdot (-1,2) + 0 \cdot (1,4)$$
entonces
$$f(2,-4) = -2 \cdot f(-1,2) + 0 \cdot f(1,4) = (-2,-6,-4)$$
- *Como*
$$(2,2) = -1 \cdot (-1,2) + 1 \cdot (1,4)$$
entonces
$$f(2,2) = -1 \cdot f(-1,2) + 1 \cdot f(1,4) = (0,-5,7)$$
- *Se deja como ejercicio calcular la expresión de* $f(x_1,x_2)$.

2.2.1. Propiedades de las aplicaciones lineales

Proposición 58 *Sean f y g aplicaciones lineales, V, V', V'' espacios vectoriales y $\lambda \in \mathbb{R}$.*

1. *Si f y g son aplicaciones lineales de V en V', entonces $f+g$ y λf también son aplicaciones lineales de V en V'.*

2. *Si f y g son aplicaciones lineales de V en V', entonces fg en general no es lineal.*

3. *Si f es una aplicación lineal de V en V' y g una aplicación lineal de V' en V'', entonces $g \circ f$ es también una aplicación lineal de V en V''.*

Observación 59 *La aplicación $-f$ definida como $(-f)(u) = -f(u)$ es la* ***aplicación opuesta*** *de f.*

Observación 60 *La* ***aplicación nula*** *se obtiene de la suma de una aplicación y su opuesta.*

Ejemplo 61 *Si $f : \mathbb{R} \to \mathbb{R}$ es tal que $f(x) = x$ y $g : \mathbb{R} \to \mathbb{R}$ es tal que $g(x) = 2x$ entonces la aplicación $fg : \mathbb{R} \to \mathbb{R}$ que se define como $fg(x) = 2x^2$ no es lineal.*

Definición 62 *Dada una aplicación lineal $f : V \to V$, si existe una aplicación $g : V \to V$ tal que para todo $u \in V$ se tiene que*

$$(f \circ g)(u) = (g \circ f)(u) = i(u)$$

donde $i(u) = u$ para todo $u \in V$ es la llamada ***aplicación identidad*** *y g es la* ***aplicación inversa*** *de f y se denota como $g = f^{-1}$.*

Proposición 63 *Sea f una aplicación lineal de V en V, si existe f^{-1} entonces f^{-1} es única y lineal.*

Proposición 64 *Sea f una aplicación lineal de V en V'. Entonces se verifica que:*

1. *$f(0_V) = 0_{V'}$ si 0_V y $0_{V'}$ son los elementos neutros de V y V' respectivamente.*

2. *Si $u_1, \dots, u_p \in V$ son linealmente dependientes entonces $f(u_1), \dots, f(u_p) \in V'$ tambien son linealmente dependientes.*

3. *Si $W \subset V$ es un subespacio vectorial entonces*

$$f(W) = \left\{ u' \in V' : u' = f(u) \text{ para algún } u \in W \right\}$$

es subespacio vectorial de V'.

4. *Si $W' \subset V'$ es un subespacio vectorial entonces*

$$f^{-1}(W') = \left\{ u \in V : f(u) \in W' \right\}$$

es subespacio vectorial de V.

Observación 65 *La dependencia lineal se conserva mediante una aplicación lineal, pero no así la independencia lineal. Por ejemplo,*

$$\begin{array}{cccc} f: & \mathbb{R}^3 & \to & \mathbb{R}^2 \\ & (x_1, x_2, x_3) & \mapsto & (x_1, x_2) \end{array}$$

si tomamos $u_1 = (1,2,1)$ y $u_2 = (2,4,0)$ son linealmente independientes, pero $f(u_1) = (1,2)$ y $f(u_2) = (2,4)$ son linealmente dependientes.

Definición 66 *Dada $f : V \to V'$ una aplicación lineal, entonces se denomina:*

1. ***núcleo*** *de la aplicación lineal f al subespacio de V, que se denota como*

$$\ker(f) = \{u \in V : f(u) = 0_{V'}\}$$

2. ***imagen*** *de la aplicación lineal f al subespacio de V'*

$$\mathrm{Im}(f) = \left\{u' \in V' : f(u) = u' \text{ para algún } u \in V\right\}$$

Es decir, el nucleo de f son todos los vectores de V que mediante f van a parar al $0_{V'}$, y la imagen de f es el conjunto $f(V)$.

Observación 67 *1. f transforma vectores linealmente independientes en vectores linealmente independientes si y sólo si $\ker(f) = \{0_V\}$ (f es inyectiva).*

2. f transforma un sistema de generadores de V en un sistema de generadores de V' si y sólo si $\mathrm{Im}(f) = V'$ (f es suprayectiva).

Ejemplo 68 *Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $f(x_1, x_2, x_3) = (x_1 + x_2, 0)$ y sean*

$$S_1 = \{(1,0,1), (2,0,2)\}$$

$$S_2 = \{(1,1,0), (3,0,0)\}$$

Es claro que S_1 es linealmente dependiente y así

$$f(S_1) = \{(1,0), (2,0)\}$$

es también linealmente dependiente.

Por otro lado S_2 es linealmente independiente pero

$$f(S_2) = \{(2,0), (3,0)\}$$

es linealmente dependiente.

Vamos a calcular $\ker\ f$ *e* $\mathrm{Im}\ f$.

$$\ker\ f = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : f(x_1, x_2, x_3) = (0,0)\right\}$$

entonces si

$$u = (x_1, x_2, x_3) \in \ker\ f$$

se tiene que

$$f(x_1, x_2, x_3) = (x_1 + x_2, 0) = (0,0)$$

es decir que

$$x_2 = -x_1$$

y así

$$\begin{aligned} u = (x_1, -x_1, x_3) &= (x_1, -x_1, 0) + (0,0,x_3) = \\ &= x_1(1,-1,0) + x_3(0,0,1) \end{aligned}$$

por lo que cualquier vector de $\ker\ f$ *es combinación lineal de estos dos vectores y por tanto podemos escribir que*

$$\ker\ f = \left\{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1 + x_2 = 0\right\}$$

o bien

$$\ker\ f = \mathcal{L}\{(1,-1,0), (0,0,1)\}$$

La imagen de f se define como

$$\text{Im } f = \left\{(y_1, y_2) \in \mathbb{R}^2 : f(x_1, x_2, x_3) = (y_1, y_2) \text{ para algún } (x_1, x_2, x_3)\right\}$$

por lo que

$$x_1 + x_2 = y_1$$

$$0 = y_2$$

para algún (x_1, x_2, x_3), por lo que

$$\text{Im } f = \left\{(y_1, y_2) \in \mathbb{R}^2 : y_2 = 0\right\}$$

o bien

$$\text{Im } f = \mathcal{L}\{(1, 0)\}$$

Teorema 69 *Sea $f : V \to V'$ una aplicación lineal entonces se verifica que*

$$\dim(V) = \dim(\ker\ f) + \dim(\text{Im} f)$$

Ejemplo 70 *Del ejemplo anterior con $f(x_1, x_2, x_3) = (x_1 + x_2, 0)$ tenemos que*

$$\ker\ f = \mathcal{L}\{(1, -1, 0), (0, 0, 1)\}$$

$$\text{Im } f = \mathcal{L}\{(1, 0)\}$$

que son tales que

$$\dim(\ker\ f) = 2$$

$$\dim(\text{Im } f) = 1$$

y como

$$\dim\left(\mathbb{R}^3\right) = 3$$

entonces se verifica la ecuación del teorema anterior

$$3 = \dim\left(\mathbb{R}^3\right) = \dim(\ker\ f) + \dim(\text{Im } f) = 2 + 1$$

2.3. Relación entre matriz y aplicación lineal

Si $\mathcal{L}(V, V')$ es el conjunto de aplicaciones lineales entre los espacios vectoriales V y V' tales que $\dim(V) = n$ y $\dim(V') = m$, entonces hay una correspondencia uno a uno entre $\mathcal{L}(V, V')$ y $\mathcal{M}_{m \times n}$. Podemos representar una aplicación lineal mediante una matriz.

2.3.1. Matriz asociada a una aplicación lineal

Vamos a encontrar una representación matricial de una aplicación lineal entre V y V' de dimensión finita.

Sean

$$B_V = \{u_1, \ldots, u_n\}$$

$$B_{V'} = \{v_1, \ldots, v_m\}$$

bases de V y V' respectivamente y sea

$$f : V \to V'$$

una aplicación lineal.

Dado $u \in V$ denotamos como

$$x = (x_1, \ldots x_n)$$

a las coordenadas de u respecto de B_V y como

$$y = (y_1, \dots y_m)$$

a las coordenadas de $f(u)$ respecto de $B_{V'}$.

Como

- $u = x_1 u_1 + \ldots + x_n u_n$
- $y = (y_1, \dots y_m)$ son las coordenadas de $f(u)$ respecto de $B_{V'}$
- f es lineal

entonces se tiene que

$$f(u) = x_1 f(u_1) + \ldots + x_n f(u_n) = y_1 v_1 + \ldots y_m v_m$$

Como ahora

$$f(u_j) = a_{1j} v_1 + \ldots + a_{mj} v_m$$

para cada $j = 1, \ldots, n$ entonces

$$a_{\bullet j} = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

son las coordenadas de $f(u_j)$ respecto de $B_{V'}$, por lo tanto

$$f(u) = x_1 (a_{11} v_1 + \ldots + a_{m1} v_m) + \ldots + x_n (a_{1n} v_1 + \ldots + a_{mn} v_m) =$$

$$= (a_{11} x_1 + \ldots + a_{1n} x_n) v_1 + \ldots + (a_{m1} x_1 + \ldots + a_{mn} x_n) v_m =$$

$$= y_1 v_1 + \ldots y_m v_m$$

entonces

$$\left\{ \begin{array}{c} a_{11} x_1 + \ldots + a_{1n} x_n = y_1 \\ \vdots \\ a_{m1} x_1 + \ldots + a_{mn} x_n = y_m \end{array} \right.$$

y así

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

Por tanto, dada una aplicación lineal $f : V \to V'$ y bases B_V de V y $B_{V'}$ de V', la matriz asociada a f respecto de las bases B_V y $B_{V'}$ es la matriz

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

es decir, la que sus columnas son las coordenadas respecto de la base $B_{V'}$ de los vectores $f(u_i) \in V'$ con $i = 1, \ldots, n$, que son los transformados mediante f de los vectores de la base B_V del espacio vectorial de partida.

Observación 71 *1. Así pues, la matriz asociada a una aplicación lineal no sólo depende de la propia aplicación f sino también de la bases B_V y $B_{V'}$ en las que se expresa. Luego para cualquier $u \in V$ cuyas coordenadas respecto de la base B son $x = (x_1, \ldots, x_n)_B$ y el vector $v \in V'$ cuyas coordenadas respecto de la base B' son $y = (y_1, \ldots, y_m)_{B'}$ tal que $f(u) = v$ se tiene que*

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$$

es decir

$$\boxed{Ax = y}$$

2. Es por ello importante tener presente la diferencia entre componentes y coordenadas.

Ejemplo 72 *Mediante el siguiente ejemplo vamos a ver como calcular la matriz asociada a una aplicación lineal dependiendo de sus bases.*

Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que

$$f(u_1, u_2, u_3) = (u_1 - u_2 + 2u_3, u_1 + 3u_2)$$

y consideremos las bases

$$B_1 = \{(1,0,-1), (2,1,0), (0,1,1)\} \text{ base de } \mathbb{R}^3$$

$$B_c(\mathbb{R}^3) \text{ la base canónica de } \mathbb{R}^3$$

$$B_2 = \{(-2,1), (1,-1)\} \text{ base de } \mathbb{R}^2$$

$$B_c(\mathbb{R}^2) \text{ la base canónica de } \mathbb{R}^2$$

<u>Caso 1</u> Matriz de f respecto de las bases canónicas $B_c(\mathbb{R}^3)$ y $B_c(\mathbb{R}^2)$.

Como en las bases canónicas las coordenadas y las componentes coinciden entonces

$$\begin{aligned} f(1,0,0) &= (1,1) = (1,1)_{B_c(\mathbb{R}^2)} \\ f(0,1,0) &= (-1,3) = (-1,3)_{B_c(\mathbb{R}^2)} \\ f(0,0,1) &= (2,0) = (2,0)_{B_c(\mathbb{R}^2)} \end{aligned}$$

así pues, situamos estas coordenadas como columnas de la matriz A y por tanto la matriz de f respecto de las bases canónicas $B_c(\mathbb{R}^3)$ y $B_c(\mathbb{R}^2)$ es

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

<u>Caso 2</u> Matriz de f respecto de las bases B_1 y $B_c(\mathbb{R}^2)$.

Transformamos la base B_1 y como de nuevo en la base $B_c(\mathbb{R}^2)$ las coordenadas y las componentes coinciden entonces

$$\begin{aligned} f(1,0,-1) &= (-1,1) = (-1,1)_{B_c(\mathbb{R}^2)} \\ f(2,1,0) &= (1,5) = (1,5)_{B_c(\mathbb{R}^2)} \\ f(0,1,1) &= (1,3) = (1,3)_{B_c(\mathbb{R}^2)} \end{aligned}$$

y otra vez colocando estos vectores de coordenadas como columnas de la matriz A se tiene que la matriz de f respecto de las bases canónicas B_1 y $B_c(\mathbb{R}^2)$ es

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

Caso 3 *Matriz de f respecto de las bases $B_c\left(\mathbb{R}^3\right)$ y B_2.*

Transformamos la base canónica de $\mathbb{R}^3$

$$\begin{aligned} f(1,0,0) &= (1,1) \\ f(0,1,0) &= (-1,3) \\ f(0,0,1) &= (2,0) \end{aligned}$$

y ahora buscamos las coordenadas de los transformados respecto de la base B_2.

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

Por otro lado

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

y además

$$(2,0) = \alpha(-2,1) + \beta(1,-1) \Rightarrow \left\{ \begin{array}{c} -2\alpha + \beta = 2 \\ \alpha - \beta = 0 \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \alpha = -2 \\ \beta = -2 \end{array} \right.$$

por tanto tenemos que

$$\begin{gathered} (1,1) = (-2,-3)_{B_2} \\ (-1,3) = (-2,-5)_{B_2} \\ (2,0) = (-2,-2)_{B_2} \end{gathered}$$

por lo que la matriz de f respecto de las bases $B_c\left(\mathbb{R}^3\right)$ y B_2 es

$$A = \left(\begin{array}{ccc} -2 & -2 & -2 \\ -3 & -5 & -2 \end{array} \right)$$

Caso 4 *Matriz de f respecto de las bases B_1 y B_2.*

Transformamos la base B_1 de $\mathbb{R}^3$

$$\begin{aligned} f(1,0,-1) &= (-1,1) \\ f(2,1,0) &= (1,5) \\ f(0,1,1) &= (1,3) \end{aligned}$$

y ahora buscamos las coordenadas de los transformados respecto de la base B_2.

$$(-1,1) = \alpha(-2,1) + \beta(1,-1) \Rightarrow \left\{ \begin{array}{c} -2\alpha + \beta = -1 \\ \alpha - \beta = 1 \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \alpha = 0 \\ \beta = -1 \end{array} \right.$$

Por otro lado

$$(1,5) = \alpha(-2,1) + \beta(1,-1) \Rightarrow \left\{ \begin{array}{c} -2\alpha + \beta = 1 \\ \alpha - \beta = 5 \end{array} \right. \Rightarrow \left\{ \begin{array}{c} \alpha = -6 \\ \beta = -11 \end{array} \right.$$

y además

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

por tanto tenemos que

$$\begin{gathered} (-1,1) = (0,-1)_{B_2} \\ (1,5) = (-6,-11)_{B_2} \\ (1,3) = (-4,-7)_{B_2} \end{gathered}$$

por lo que la matriz de f respecto de las bases $B_c\left(\mathbb{R}^3\right)$ y B_2 es

$$A = \left(\begin{array}{ccc} 0 & -6 & -4 \\ -1 & -11 & -7 \end{array} \right)$$

Observación 73 *Por lo observado en este ejemplo el procedimiento es el siguiente:*
Sea $f : V \to V'$ con base B de V y B' de V' respectivamente.

1. *Transformamos mediante f los vectores de la base B de V.*

2. *Calculamos las coordenadas de estos transformados respecto de la base B' de V'.*

3. *Estas coordenadas son las columnas de la matriz A asociada a f respecto de las bases indicadas.*

Proposición 74 *Sean V y V' dos espacios vectoriales de dimensión n y m respectivamente. Dada una aplicación lineal $f : V \to V'$ con matriz asociada A respecto de las bases $B_1 = \{u_1, \ldots, u_n\}$ de V y la canónica de V', entonces se verifica que:*

1. *Si $\{e_1, \ldots, e_n\}$ es la base canónica de V entonces el conjunto de vectores $\{f(e_1), \ldots, f(e_n)\}$ es un sistema de generadores de* $\operatorname{Im} f$.

2. *El conjunto de vectores $\{a_{\bullet 1}, \ldots, a_{\bullet n}\}$ donde $a_{\bullet j}$ con $j = 1, \ldots, n$ son las columnas de la matriz A, es también un sistema de generadores de* $\operatorname{Im} f$.

2.3.2. Aplicación lineal asociada a una matriz (sólo respecto de las bases canónicas)

Sea $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ una matriz y consideremos V y V' dos espacios vectoriales tales que $\dim V = n$ y $\dim V' = m$dotados de sus bases canónicas. Si $(x_1, \ldots, x_n)$ son las coordenads de $u \in V$ entonces

$$Au = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} =$$

$$= \left(a_{1\bullet} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ldots, a_{m\bullet} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right)$$

Así pues, si $f : V \to V'$ es la aplicación lineal que tiene a A como matriz asociada respecto de las bases canónicas de V y V' entonces

$$f(u) = Au$$

y por tanto

$$f(u) = (a_{1\bullet} u, \ldots, a_{m\bullet} u)$$

Ejemplo 75 *Sea $A = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 1 & 0 \end{pmatrix}$ la matriz de la aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^2$ que tiene a A como matriz asociada respecto de las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente. Entonces si*

$$u = (x_1, x_2, x_3) \in \mathbb{R}^3$$

tenemos que

$$f(x_1, x_2, x_3) = \begin{pmatrix} 3 & 1 & 2 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} =$$

$$= (3x_1 + x_2 + 2x_3, -x_1 + x_2)$$

2.4. Rango de una matriz

Por la proposición anterior tenemos que

$$\begin{aligned}\dim(\mathrm{Im}\ f) &= \text{máximo número de vectores columna de } A\\ &\quad\ \text{linealmente independientes}\end{aligned}$$

es decir, coincide con el rango de dicho conjunto de vectores.

Definición 76 *Dada una matriz A cualquiera de orden $m \times n$, se denomina **rango de la matriz** A al máximo número de vectores columna de A linealmente independientes.*

Se denota como $rg(A)$.

Observación 77 *Todas las matrices asociadas a una misma aplicación lineal f tienen el mismo rango.*

Observación 78 *Sea f una aplicación lineal y sea A una matriz asociada a f, entonces*

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

2.5. Tipos de matrices según su rango

Definición 79 *Dada una matriz $A \in \mathcal{M}_{m\times n}$ se tiene que:*

1. *Si $m \neq n$ y $rg(A) = \min\{m, n\}$ se dice que A es de **rango completo**.*
2. *Si $m = n$ y $rg(A) = n$ se dice que A es **no singular** o **regular**.*
3. *Si $m = n$ y $rg(A) < n$ se dice que A es **singular**.*

Ejemplo 80 *Sea $f : \mathbb{R}^3 \to \mathbb{R}^2$ tal que $f(x, y, z) = (2x, x + z)$ que tiene como matriz respecto de las bases canónicas de $\mathbb{R}^3$ y $\mathbb{R}^2$ respectivamente*

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

Así la

$$\mathrm{Im}\ f = \mathcal{L}\left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\} = \mathcal{L}\left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}.$$

Es claro que $\begin{pmatrix}2\\1\end{pmatrix}$ y $\begin{pmatrix}0\\1\end{pmatrix}$ son linealmente independientes ya que

$$\alpha\begin{pmatrix}2\\1\end{pmatrix} + \beta\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

$$\left.\begin{array}{c} 2\alpha = 0\\ \alpha + \beta = 0\end{array}\right\} \Rightarrow \alpha = \beta = 0$$

Por tanto

$$rg(A) = 2$$

Ejemplo 81 *Sea $f : \mathbb{R}^2 \to \mathbb{R}^4$ una aplicación lineal tal que*

$$\begin{aligned} f(3, -5) &= (1, 1, 1, 1)\\ f(-1, 2) &= (2, 1, 0, -2)\end{aligned}$$

se pide hallar:

1. La expresión analítica de f.

2. La matriz asociada a f respecto de las bases canónicas.

3. Determinar $\operatorname{Im}\ f$, $\ker\ f$ y sus dimensiones $\dim(\operatorname{Im} f)$ y $\dim(\ker\ f)$.

La expresión analítica de f.

Observar que $B = \left\{ \begin{pmatrix} 3 \\ -5 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\}$ *es una base de* $\mathbb{R}^2$.

Un vector $v = (v_1, v_2) \in \mathbb{R}^2$ *tendrá coordenadas* $(\alpha, \beta)_B$ *respecto de la base* B. *Es decir*

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \alpha \begin{pmatrix} 3 \\ -5 \end{pmatrix} + \beta \begin{pmatrix} -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3\alpha - \beta \\ -5\alpha + 2\beta \end{pmatrix}$$

$$\left. \begin{array}{c} v_1 = 3\alpha - \beta \\ v_2 = -5\alpha + 2\beta \end{array} \right\} \Rightarrow \textit{Resolviendo} \quad \left. \begin{array}{c} \alpha = 2v_1 + v_2 \\ \beta = 5v_1 + 3v_2 \end{array} \right\}$$

entonces tenemos que

$$\begin{aligned} f(v) = f(v_1, v_2) = f(\alpha(3, -5) + \beta(-1, 2)) = \\ = \alpha \cdot f(3, -5) + \beta \cdot f(-1, 2) = \\ = (2v_1 + v_2) \cdot (1, 1, 1, 1) + (5v_1 + 3v_2) \cdot (2, 1, 0, -2) = \\ = (12v_1 + 7v_2, 7v_1 + 4v_2, 2v_1 + v_2, -8v_1 - 5v_2) \end{aligned}$$

y por lo tanto

$$\boxed{f(v_1, v_2) = (12v_1 + 7v_2, 7v_1 + 4v_2, 2v_1 + v_2, -8v_1 - 5v_2)}$$

La matriz asociada a f respecto de las bases canónicas.

Por el apartado anterior tenemos que la matriz A asociada a f respecto de las bases canónicas de $\mathbb{R}^2$ y $\mathbb{R}^4$ es

$$A = \begin{pmatrix} 12 & 7 \\ 7 & 4 \\ 2 & 1 \\ -8 & -5 \end{pmatrix}$$

Determinar $\operatorname{Im}\ f$, $\ker\ f$ y sus dimensiones $\dim(\operatorname{Im}\ f)$ y $\dim(\ker\ f)$.

Tenemos que

$$\operatorname{Im}\ f = \mathcal{L} \left\{ \begin{pmatrix} 12 \\ 7 \\ 2 \\ -8 \end{pmatrix} \begin{pmatrix} 7 \\ 4 \\ 1 \\ -5 \end{pmatrix} \right\}$$

y como estos vectores generadores son linealmente independientes entonces

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

y como

$$\dim\left(\mathbb{R}^2\right) = \dim(\operatorname{Im}\ f) + \dim(\ker\ f)$$

se tiene que

$$\dim(\ker\ f) = 0$$

por lo que

$$\ker\ f = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

Ejercicio 82 *Demostrar que la expresión analítica de $\operatorname{Im}\ f$ del ejemplo anterior es*

$$\operatorname{Im}\ f = \left\{ (x, y, z, t) \in \mathbb{R}^4 : x - 2y + z = 0, \quad x - 2z + t = 0 \right\}$$

2.6. Operaciones con matrices

Ya hemos visto como operar el producto de una matriz $A \in \mathcal{M}_{m\times n}$ por un vector columna $v \in \mathbb{R}^n$,

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + \ldots + a_{1n}v_n \\ \vdots \\ a_{m1}v_1 + \ldots + a_{mn}v_n \end{pmatrix}$$

El resultado es otro vector columna de $\mathbb{R}^m$, que podemos ver como una matriz de orden $m \times 1$.

Veamos a continuación la representación de operaciones entre aplicaciones lineales como operaciones entre matrices.

2.6.1. Suma de matrices (suma de aplicaciones lineales)

Sean $f : \mathbb{R}^n \to \mathbb{R}^m$ y $g : \mathbb{R}^n \to \mathbb{R}^m$ dos aplicaciones lineales con matrices asociadas $A = (a_{ij})$ y $B = (b_{ij})$ respectivamente de orden $m \times n$.

Sea $\bar{x} \in \mathbb{R}^n$, entonces

$$f(\bar{x}) = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = A\bar{x}$$

$$g(\bar{x}) = \begin{pmatrix} b_{11}x_1 + \ldots + b_{1n}x_n \\ \vdots \\ b_{m1}x_1 + \ldots + b_{mn}x_n \end{pmatrix} = B\bar{x}$$

así tenemos que

$$(f+g)(\bar{x}) = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} + \begin{pmatrix} b_{11}x_1 + \ldots + b_{1n}x_n \\ \vdots \\ b_{m1}x_1 + \ldots + b_{mn}x_n \end{pmatrix} =$$

$$= \begin{pmatrix} (a_{11}+b_{11})x_1 + \ldots + (a_{1n}+b_{1n})x_n \\ \vdots \\ (a_{m1}+b_{m1})x_1 + \ldots + (a_{mn}+b_{mn})x_n \end{pmatrix} =$$

$$= \begin{pmatrix} a_{11}+b_{11} & \cdots & a_{1n}+b_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & \cdots & a_{mn}+b_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (A+B)\bar{x}$$

2.6.2. Producto de una matriz por un escalar (producto de una aplicación lineal por un escalar)

Sea $f : \mathbb{R}^n \to \mathbb{R}^m$ una aplicación lineal con matriz asociada $A = (a_{ij})$ de orden $m \times n$ y sea $\lambda \in \mathbb{R}$.
Sea $\bar{x} \in \mathbb{R}^n$, entonces

$$f(\bar{x}) = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = A\bar{x}$$

así tenemos que

$$(\lambda f)(\bar{x}) = \lambda f(\bar{x}) = \lambda \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda a_{11}x_1 + \ldots + \lambda a_{1n}x_n \\ \vdots \\ \lambda a_{m1}x_1 + \ldots + \lambda a_{mn}x_n \end{pmatrix} =$$

$$= \begin{pmatrix} \lambda a_{11} & \cdots & \lambda a_{1n} \\ \vdots & \ddots & \vdots \\ \lambda a_{m1} & \cdots & \lambda a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (\lambda A)\, \bar{x}$$

Proposición 83 *(Propiedades) Sean $A, B, C \in \mathcal{M}_{m\times n}$ y $\alpha, \beta \in \mathbb{R}$:*

1. *La suma de matrices tiene las propiedades:*

   a) *Conmutativa: $A + B = B + A$.*

   b) *Asociativa: $(A + B) + C = A + (B + C)$.*

   c) *Existe elemento neutro $O_m \in \mathcal{M}_{m\times n}$ tal que $A + O_m = A$.*

   d) *Existe elemento opuesto $-A$ de $A \in \mathcal{M}_{m\times n}$ tal que $A + (-A) = O_m$.*

2. *El producto de una matriz por un escalar tiene las propiedades:*

   a) *Distributiva de la suma de matrices: $\alpha (A + B) = \alpha A + \alpha B$.*

   b) *Distributiva de la suma de escalares: $(\alpha + \beta) A = \alpha A + \beta A$.*

   c) *Pseudoasociativa: $(\alpha\beta) A = \alpha (\beta A)$.*

   d) *Existe elemento unidad: $1 \in \mathbb{R}$ es tal que $1A = A$.*

Observación 84 *Estas 8 propiedades hacen que $\mathcal{M}_{m\times n}$ sea un espacio vectorial. La base canónica de $\mathcal{M}_{m\times n}$ será*

$$\left\{ E_{i_0 j_0} = (e_{ij}) : i_0 = 1, \ldots, m; \quad j_0 = 1, \ldots, m; \quad e_{ij} = \left\{ \begin{array}{ll} 0 & si\ i \neq i_0\ o\ j \neq j_0 \\ 1 & si\ i = i_0\ y\ j = j_0 \end{array} \right. \right\}$$

que tiene mn elementos. Por lo tanto $\dim (\mathcal{M}_{m\times n}) = mn$.

De la misma manera y por la correspondencia entre matrices y aplicaciones lineales se tiene que $\mathcal{L}(V, V')$ es un espacio vectorial de dimesión mn.

Ejemplo 85 *Sean $A = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 3 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -1 & 3 \\ 2 & 1 & 5 \end{pmatrix}$ y $C = \begin{pmatrix} 1 & 1 \\ -5 & 0 \\ 1 & -7 \end{pmatrix}$ matrices y sea $\alpha = 5$, entonces*

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

y además

$$\alpha C = 5 \begin{pmatrix} 1 & 1 \\ -5 & 0 \\ 1 & -7 \end{pmatrix} = \begin{pmatrix} 5 & 5 \\ -25 & 0 \\ 5 & -35 \end{pmatrix}$$

2.6.3. Producto de matrices (composición de aplicaciones lineales)

Sean $f : \mathbb{R}^n \to \mathbb{R}^m$ y $g : \mathbb{R}^m \to \mathbb{R}^p$ dos aplicaciones lineales con matrices asociadas $A = (a_{ij}) \in \mathcal{M}_{m\times n}$ y $B = (b_{ij}) \in \mathcal{M}_{p\times m}$.

Sea $\bar{x} \in \mathbb{R}^n$ e $\bar{y} \in \mathbb{R}^m$, entonces como

$$f(\bar{x}) = \begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix} = A\bar{x}$$

$$g(\bar{y}) = \begin{pmatrix} b_{11}y_1 + \ldots + b_{1m}y_m \\ \vdots \\ b_{p1}y_1 + \ldots + b_{pm}y_m \end{pmatrix} = B\bar{y}$$

tenemos que

$$g \circ f(\bar{x}) = g(f(\bar{x})) = g\left(\begin{pmatrix} a_{11}x_1 + \ldots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n \end{pmatrix}\right) =$$

$$= \begin{pmatrix} b_{11}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + b_{1m}(a_{m1}x_1 + \ldots + a_{mn}x_n) \\ \vdots \\ b_{p1}(a_{11}x_1 + \ldots + a_{1n}x_n) + \ldots + b_{pm}(a_{m1}x_1 + \ldots + a_{mn}x_n) \end{pmatrix} =$$

$$= \begin{pmatrix} (b_{11}a_{11} + \ldots + b_{1m}a_{m1})x_1 + \ldots + (b_{11}a_{1n} + \ldots + b_{1m}a_{mn})x_n \\ \vdots \\ (b_{p1}a_{11} + \ldots + b_{pm}a_{m1})x_1 + \ldots + (b_{p1}a_{1n} + \ldots + b_{pm}a_{mn})x_n \end{pmatrix} =$$

$$= \begin{pmatrix} b_{11}a_{11} + \ldots + b_{1m}a_{m1} & \cdots & b_{11}a_{1n} + \ldots + b_{1m}a_{mn} \\ \vdots & \ddots & \vdots \\ b_{p1}a_{11} + \ldots + b_{pm}a_{m1} & \cdots & b_{p1}a_{1n} + \ldots + b_{pm}a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C\bar{x}$$

es decir si $C = (c_{ij}) \in \mathcal{M}_{p\times n}$ se tiene que el elemento c_{ij} es

$$c_{ij} = b_{i\bullet}a_{\bullet j}$$

y matricialmente se expresa como

$$C = BA = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{p1} & \cdots & b_{pm} \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

Proposición 86 *(Propiedades) Sean A, B, C, D matrices tales que $A \in \mathcal{M}_{m\times p}$, $B \in \mathcal{M}_{p\times q}$ y $C, D \in \mathcal{M}_{q\times n}$, entonces*

1. *En general el producto no es conmutativo, es decir, en general*

$$AB \neq BA$$

2. *El producto es asociativo*

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

3. *Existen matrices $I_m \in \mathcal{M}_{m\times m}$ e $I_p \in \mathcal{M}_{p\times p}$ tales que para toda $A \in \mathcal{M}_{m\times p}$ se tiene que*

$$I_mA = A \qquad AI_p = A$$

4. *El producto de A por cualquier matriz nula de dimensión adecuada es una matriz nula*

$$AO_{p\times s} = O_{m\times s} \qquad\qquad O_{s\times m}A = O_{s\times p}$$

5. *El producto es distributivo (con las dimensiones adecuadas)*

$$B\left(C + D\right) = BC + BD$$

6. *Se verifica que*

$$rg\left(AB\right) \leq \text{mín}\left\{rg\left(A\right), rg\left(B\right)\right\}$$

7. *Si $rg\left(A\right) = r \leq \text{mín}\left\{m, p\right\}$ entonces existen C_1 y C_2 de rango r y órdenes $m \times r$ y $r \times p$ respectivamente tales que*

$$A = C_1C_2$$

Observación 87 *En general **no se verifica** algo tan natural en $\mathbb{R}$ como*

1. *Si $ab = 0$ entonces $a = 0$ o $b = 0$.*

2. *Si $ab = ac$ y $a \neq 0$ entonces $b = c$.*

Como contraejemplos tenemos

$$A = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \text{ y } B = \begin{pmatrix} -2 & 4 \\ 1 & -2 \end{pmatrix}$$

que son tales que

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

y además

$$A = \begin{pmatrix} 2 & 3 \\ 6 & 9 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \text{ y } C = \begin{pmatrix} -2 & 1 \\ 3 & 2 \end{pmatrix}$$

que verifican

$$AB = AC = \begin{pmatrix} 5 & 8 \\ 15 & 24 \end{pmatrix}$$

pero $B \neq C$.

Ejercicio 88 *Sean $A = \begin{pmatrix} 1 & -4 & 2 \\ -1 & 4 & -2 \end{pmatrix}$ y $B = \begin{pmatrix} 1 & 2 \\ -1 & 3 \\ 5 & -2 \end{pmatrix}$, calcule AB y BA.*

Ejercicio 89 *Sean*

$$f\left(x_1, x_2, x_3\right) = \left(x_1, x_1 + x_2, x_2 + x_3\right)$$
$$g\left(x_1, x_2, x_3\right) = \left(x_2, 2x_3 - x_1, x_1 - x_2\right)$$

1. *Halle las matrices asociadas respecto de las bases canónicas.*

2. *Determine, utilizando las matrices del apartado anterior, las expresiones de las aplicaciones $f + g$, $3g$, $g \circ f$ y $f \circ g$.*

2.7. Matriz inversa (aplicación lineal inversa)

Sea $f : \mathbb{R}^n \to \mathbb{R}^n$ una aplicación lineal con matiz asociada $A \in \mathcal{M}_{n\times n}$. Supongamos que existe f^{-1} con matriz $B \in \mathcal{M}_{n\times n}$, entonces como

$$f \circ f^{-1}(\bar{x}) = id(\bar{x}) \qquad \text{y} \qquad f^{-1} \circ f(\bar{x}) = id(\bar{x})$$

se tiene que

$$AB\bar{x} = I_n\bar{x} \qquad \text{y} \qquad BA\bar{x} = I_n\bar{x}$$

así pues cuando existe una matriz $B \in \mathcal{M}_{n\times n}$ que cumple que

$$AB = BA = I_n$$

entonces B es la matriz inversa de A y se denota como A^{-1}.

Como f^{-1}, si existe, es única, entonces A^{-1} también es única.

Observación 90 *Las matrices no cuadradas carecen de inversa y no todas las matrices cuadradas tienen inversa.*

Proposición 91 *Dada $A \in \mathcal{M}_{n\times n}$ se verifica que existe A^{-1} si y sólo si la matriz A es de rango completo.*

Ejemplo 92 *La matriz $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}^{-1}$ es de rango completo y $A^{-1} = \begin{pmatrix} -\frac{2}{5} & -\frac{1}{5} & \frac{3}{5} \\ 1 & 1 & -1 \\ \frac{1}{5} & -\frac{2}{5} & \frac{1}{5} \end{pmatrix}$ es su matriz inversa.*

Proposición 93 *(Propiedades de la inversa) Sean $A, B \in \mathcal{M}_{n\times n}$, tales que $rg(A) = rg(B) = n$. Se verifica que*

1. *Existe A^{-1} matriz inversa de A y es única.*
2. *La matriz A^{-1} es invertible y $\left(A^{-1}\right)^{-1} = A$.*
3. *La matriz AB tiene inversa y además $(AB)^{-1} = B^{-1}A^{-1}$.*
4. *Si $C \in \mathcal{M}_{n\times n}$ tal que $rg(C) = n$ y $AC = I_n$ entonces se tiene que $CA = I_n$ y $C = A^{-1}$. Recíprocamente si $CA = I_n$ entonces $AC = I_n$ y $C = A^{-1}$.*

2.7.1. Transformaciones elementales

Llamaremos transformaciones elementales entre filas de una matriz A no singular a las siguientes operaciones:

- Intercambiar filas.
- Multiplicar una fila por un escalar.
- Sumar a una fila una combinación lineal de las restantes filas.

Estas transformaciones se pueden representar mediante matrices no singulares $P_1, \ldots, P_r$ del mismo orden que A, es decir $P_1, \ldots, P_r \in \mathcal{M}_{n\times n}$.

Definición 94 *Dos matrices $A, B \in \mathcal{M}_{n\times n}$ son **semejantes** o **equivalentes** si existe una transformación elemental $P \in \mathcal{M}_{n\times n}$ tal que*

$$A = PB$$

2.7.2. Cálculo de la inversa mediante el método de Gauss-Jordan

Para encontrar la matriz inversa de $A \in \mathcal{M}_{n\times n}$, por la propiedad 4 anterior, buscamos una matriz C formada por transformaciones elementales tal que

$$C = P_r P_{r-1} \cdots P_2 P_1$$

y además

$$P_r P_{r-1} \cdots P_2 P_1 A = I_n$$

por lo tanto

$$A^{-1} = P_r P_{r-1} \cdots P_2 P_1$$

El mismo cálculo puede hacerse con transformaciones elementales de columnas pero en este curso lo haremos siempre de filas.

Ejemplo 95 *Sea* $A = \begin{pmatrix} 1 & 1 & 2 \\ 2 & 1 & -1 \\ 3 & 1 & 1 \end{pmatrix}$*, entonces construimos la matriz*

$$B = (A\,|I_3\,) = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 2 & 1 & -1 & 0 & 1 & 0 \\ 3 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

y trataremos de buscar una matriz C tal que $C = \left(I_3\,\middle|A^{-1}\right)$, es decir, transformaremos las columnas de B para que sean las columnas de la matriz identidad I_3.

Así

$$\left\{\begin{array}{c} 1^a \\ -2\cdot 1^a + 2^a \\ -3\cdot 1^a + 3^a \end{array}\right. \qquad P_1 = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -3 & 0 & 1 \end{pmatrix} \qquad B_1 = \left(\begin{array}{ccc|ccc} 1 & 1 & 2 & 1 & 0 & 0 \\ 0 & -1 & -5 & -2 & 1 & 0 \\ 0 & -2 & -5 & -3 & 0 & 1 \end{array}\right)$$

$$\left\{\begin{array}{c} 1^a + 2^a \\ -1\cdot 2^a \\ -2\cdot 2^a + 3^a \end{array}\right. \qquad P_2 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -2 & 1 \end{pmatrix} \qquad B_2 = \left(\begin{array}{ccc|ccc} 1 & 0 & -3 & -1 & 1 & 0 \\ 0 & 1 & 5 & 2 & -1 & 0 \\ 0 & 0 & 5 & 1 & -2 & 1 \end{array}\right)$$

$$\left\{\begin{array}{c} 1^a + \frac{3}{5}\cdot 3^a \\ 2^a - 3^a \\ \left(\frac{1}{5}\right)\cdot 3^a \end{array}\right. \qquad P_3 = \begin{pmatrix} 1 & 0 & \frac{3}{5} \\ 0 & 1 & -1 \\ 0 & 0 & \frac{1}{5} \end{pmatrix} \qquad B_3 = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{-2}{5} & \frac{-1}{5} & \frac{3}{5} \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & \frac{1}{5} & \frac{-2}{5} & \frac{1}{5} \end{array}\right)$$

Por lo tanto la matriz inversa de A es

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

que resulta que es precisamente

$$A^{-1} = P_3 P_2 P_1$$

aunque no hace falta calcular las matrices de las transformaciones elementales ya que A^{-1} aparece a la derecha de la matriz B_3 que tiene en su parte izquierda la matriz identidad.

Ejemplo 96 *Vamos a calcular la inversa de* $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$. *Así tenemos que*

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{array}\right) \sim_{P_1} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & 1 & 0 \\ 0 & -1 & -1 & -1 & 0 & 1 \end{array}\right) \sim_{P_2}$$

$$\sim_{P_2} \left(\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \sim_{P_3} \left(\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right) \sim_{P_4}$$

$$\sim_{P_4} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right)$$

donde las transformaciones elementales son

$$P_1 \to \begin{cases} 1^a \\ -1^a + 2^a \\ -1^a + 3^a \end{cases} \quad P_2 \to \begin{cases} 1^a \\ -3^a \\ -2^a \end{cases} \quad P_3 \to \begin{cases} 1^a - 3^a \\ 2^a - 3^a \\ 3^a \end{cases} \quad P_4 \to \begin{cases} 1^a - 2^a \\ 2^a \\ 3^a \end{cases}$$

Observación 97 *Si la matriz A no es invertible, entonces no es posible obtener la matriz identidad en la parte izquierda de la matriz $B = (A\,|\,I\,)$.*

Ejercicio 98 *Calcular las matrices inversas de*

$$A_1 = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 2 & 4 \\ 0 & 1 & 1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ 1 & 2 & 1 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 2 & 2 & 1 \\ 1 & 1 & 3 \\ 1 & -1 & 0 \end{pmatrix}$$

2.8. Tipos de matrices

Podemos clasificar las matrices cuadradas según:

1. La disposición de sus elementos
   - *a*) Matrices triangulares.
   - *b*) Matrices diagonales.
   - *c*) Matrices simétricas.
   - *d*) Matrices antisimétricas.
2. Su comportamiento ante algunas operaciones
   - *a*) Matrices ortogonales.
   - *b*) Matrices idempotentes.
   - *c*) Matrices nilpotentes.
   - *d*) Matrices unipotentes.

2.8.1. Matrices triangulares superiores

La matriz $A = (a_{ij})$ es **triangular superior** si todos los elementos de A situados por debajo de la diagonal principal son nulos, es decir

$$a_{ij} = 0 \text{ para todo } i > j \text{ con } i, j = 1, \ldots, n$$

Como ejemplo tenemos $A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

2.8.2. Matrices triangulares inferiores

La matriz $A = (a_{ij})$ es **triangular inferior** si todos los elementos de A situados por encima de la diagonal principal son nulos, es decir

$$a_{ij} = 0 \text{ para todo } i < j \text{ con } i, j = 1, \ldots, n$$

Como ejemplo tenemos $A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \\ 3 & 6 & 8 & 0 \\ 4 & 7 & 9 & 1 \end{pmatrix}$

2.8.3. Matrices diagonales

La matriz $A = (a_{ij})$ es **diagonal** si es triangular superior e inferior, es decir

$$a_{ij} = 0 \text{ para todo } i \neq j \text{ con } i, j = 1, \ldots, n$$

Como ejemplos tenemos $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

2.8.4. Matrices simétricas

La matriz $A = (a_{ij})$ es **simétrica** si para todo $i = 1, \ldots, n$ la fila i-ésima y la columna i-ésima son iguales, es decir

$$a_{ij} = a_{ji} \text{ para todo } i, j = 1, \ldots, n$$

Como ejemplo tenemos $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 0 & 6 \\ 3 & 6 & 8 \end{pmatrix}$ que son simétricas respecto de la diagonal principal.

2.8.5. Matrices antisimétricas

La matriz $A = (a_{ij})$ es **antisimétrica** si

$$a_{ij} = -a_{ji} \text{ para todo } i, j = 1, \ldots, n$$

Así tenemos que los elementos de la diagonal a_{ii} son todos nulos ya que $a_{ii} = -a_{ii}$ y entonces $2a_{ii} = 0$ implica que $a_{ii} = 0$.

Como ejemplo tenemos $A_1 = \begin{pmatrix} 0 & -2 & 1 \\ 2 & 0 & -3 \\ -1 & 3 & 0 \end{pmatrix}$ y $A_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

2.8.6. Matrices ortogonales

La matriz A es **ortogonal** si la matriz A^t transpuesta de A (que también se denota como A' y es la que resulta de cambiar filas por columnas) es la inversa de A, es decir, que

$$A^t = A^{-1}$$

En este caso los vectores columna de A son ortogonales dos a dos y de módulo igual a 1.

Ejemplos son $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} \frac{1}{2} & \frac{-\sqrt{3}}{2} & 0 \\ \frac{\sqrt{3}}{2} & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

2.8.7. Matrices idempotentes

La matriz A es **idempotente** si se verifica que

$$A^2 = A$$

Ejemplos de este tipo son $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & \frac{1}{3} \end{pmatrix}$.

A veces, también se dice que una matriz A es idempotente si existe $k \in \mathbb{N}$ tal que $k > 1$ que verifica que $A^k = A$.

2.8.8. Matrices nilpotentes

La matriz A es **nilpotente** si se verifica que

$$A^2 = 0_n \text{ la matriz nula}$$

Ejemplos de este tipo son $A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ y $A_2 = \begin{pmatrix} 2 & -4 \\ 1 & -2 \end{pmatrix}$.

A veces, también se dice que una matriz A es nilpotente si existe $k \in \mathbb{N}$ tal que $k > 1$ que verifica que $A^k = 0_n$.

2.8.9. Matrices unipotentes

La matriz A es **unipotente** si se verifica que

$$A^2 = I_n \text{ la matriz identidad}$$

Ejemplos de este tipo son $A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ y $A_2 = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{pmatrix}$.

A veces, también se dice que una matriz A es nilpotente si existe $k \in \mathbb{N}$ tal que $k > 1$ que verifica que $A^k = I_n$.

Ejercicio 99 *Clasificar las siguientes matrices según las categorías anteriores.*

$$A_1 = \begin{pmatrix} 2 & 3 & 4 \\ 1 & 0 & 0 \\ -1 & 0 & -1 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Capítulo 3

Traza y determinante de una matriz

3.1. Introducción

Hemos visto que $rg\,(A)$ es un número asociado a la matriz A. Este número nos da información sobre las columnas (o filas) linealmente independientes. Nos da cuenta de la información valiosa y de la cantidad de la información redundante.

Existen otras funciones asociadas a las matrices que nos dan información sobre su invertibilidad, etc.

3.2. Traza de una matriz

Definición 100 *Dada una matriz cuadrada A de orden $n \times n$, se define la* ***traza*** *de A, como el valor de la suma de todos los elementos de la diagonal principal de A. Se denota como $tr\,(A)$. Así pues*

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

Ejemplo 101 *Sean* $A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 3 & 0 \\ 1 & 1 & -2 \end{pmatrix}$ *y* $B = \begin{pmatrix} 2 & 4 \\ 1 & 0 \end{pmatrix}$ *entonces se tiene que*

$$tr\,(A) = 1 + 3 - 2 = 2 \qquad y \qquad tr\,(B) = 2 + 0$$

Proposición 102 *1. Sean A, B matrices cuadradas de orden $n \times n$, y sea $\alpha \in \mathbb{R}$, entonces*

$$tr\,(A + B) = tr\,(A) + tr\,(B)$$

$$tr\,(\alpha A) = \alpha \cdot tr\,(A)$$

2. Sean A, B matrices de órdenes $m \times n$ y $n \times m$ respectivamente, entonces

$$tr\,(AB) = tr\,(BA)$$

Observación 103 *1. La aplicación entre espacios vectoriales*

$$tr : \mathcal{M}_{n\times n} \to \mathbb{R}$$

virtud de la propiedad 1 anterior.

2. En general, ***no es cierto*** *que*

$$tr\,(AB) = tr\,(A) \cdot tr\,(B)$$

$$tr\,\left(A^{-1}\right) = \frac{1}{tr\,(A)}$$

3.3. Determinante de una matriz

Definición 104 *Dada una matriz A de orden $n \times n$, el* ***determinante*** *de A es la suma de los $n!$ productos con signo de n factores que se obtienen considerando los elementos de A de forma que cada producto tenga un elemento y sólo uno de cada fila y cada columna de A. Lo denotaremos como $|A|$ o $\det(A)$.*

Analíticamente

$$|A| = \sum_{k=1}^{n!} (-1)^{s_k} a_{1j_1} a_{2j_2} \dots a_{nj_n}$$

donde $\{j_1, \dots, j_n\}$ es una de las $n!$ permutaciones de los elementos del conjunto $\{1, 2, \dots, n\}$ y s_k es el número de transposiciones o "cambios"necesarios para reordenar la permutación $\{j_1, \dots, j_n\}$ en el orden de $\{1, 2, \dots, n\}$.

Observación 105 *1. A partir de la definición tenemos inmediatamente que*

$$|0_n| = 0$$

y además

$$|I_n| = 1$$

2. Esta definición no es la más operativa, pero encontraremos fórmulas y propiedades para calcular los determinantes de las matrices sin necesidad de construir cada vez la fórmula adecuada.

3.3.1. Cálculo de determinantes

Matriz de orden 1 Si $A = (a_{11}) \in \mathcal{M}_{1\times 1}$ entonces el cálculo del determinante es trivial ya que

$$|A| = a_{11}$$

Matriz de orden 2 Dada $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathcal{M}_{2\times 2}$ y aplicando la definición anterior tenemos que

$$|A| = (-1)^{s_1} a_{1j_1} a_{2j_2} + (-1)^{s_2} a_{1i_1} a_{2i_2}$$

Para encontrar la fórmula debemos identificar las permutaciones $\{j_1, j_2\}$ y $\{i_1, i_2\}$ así como los valores de s_1 y s_2.

Las únicas permutaciones de $\{1, 2\}$ son $\{1, 2\}$ y $\{2, 1\}$, por lo tanto, elegimos

$$\{j_1, j_2\} = \{1, 2\} \qquad \{i_1, i_2\} = \{2, 1\}$$

luego

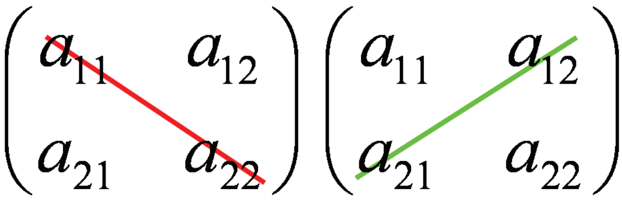
$$s_1 = 0 \qquad s_2 = 1$$

Así pues la fórmula quedará

$$|A| = (-1)^0 a_{11} a_{22} + (-1)^1 a_{12} a_{21}$$

y entonces

$$\boxed{|A| = a_{11} a_{22} - a_{12} a_{21}}$$



$$|A| = \underline{a_{11}a_{22}} - \underline{a_{12}a_{21}}$$

Ejemplo 106 *Sea* $A = \begin{pmatrix} 1 & 2 \\ 4 & -3 \end{pmatrix}$ *entonces* $|A| = 1 \cdot (-3) - 2 \cdot 4 = -11.$

Matriz de orden 3 Calculemos el determinante de $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

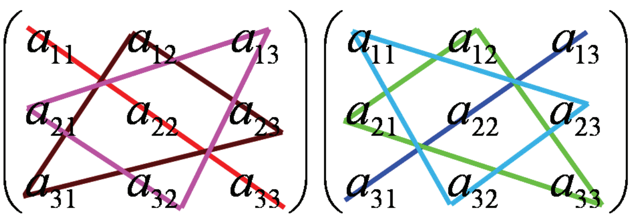
Las permutaciones de $\{1, 2, 3\}$ son las siguientes:

Permutación	Indice	Término
$\{1, 2, 3\}$	$\Rightarrow s_1 = 0$	$\Rightarrow +a_{11}a_{22}a_{33}$
$\{3, 2, 1\}$	$\Rightarrow s_2 = 1$	$\Rightarrow -a_{13}a_{22}a_{31}$
$\{1, 3, 2\}$	$\Rightarrow s_3 = 1$	$\Rightarrow -a_{11}a_{23}a_{32}$
$\{3, 1, 2\}$	$\Rightarrow s_4 = 2$	$\Rightarrow +a_{13}a_{21}a_{32}$
$\{2, 3, 1\}$	$\Rightarrow s_5 = 2$	$\Rightarrow +a_{12}a_{23}a_{31}$
$\{2, 1, 3\}$	$\Rightarrow s_6 = 1$	$\Rightarrow -a_{12}a_{21}a_{33}$

Por lo tanto, la fórmula del determinante de A es

$$\begin{aligned} |A| &= a_{11}a_{22}a_{33} + a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} - \\ & -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \end{aligned}$$

Esta fórmula se puede recordar mediante **la regla de Sarrus**:



$$|A| = \underline{a_{11}a_{22}a_{33}} + \underline{a_{13}a_{21}a_{32}} + \underline{a_{12}a_{23}a_{31}} -$$
$$- \underline{a_{13}a_{22}a_{31}} - \underline{a_{11}a_{23}a_{32}} - \underline{a_{12}a_{21}a_{33}}$$

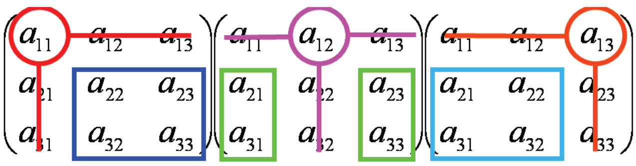
Ejemplo 107 *Sea* $A = \begin{pmatrix} 1 & 3 & 0 \\ 2 & -1 & -2 \\ 4 & 5 & -3 \end{pmatrix}$ *entonces tenemos que*

$$\begin{gathered} |A| = 1 \cdot (-1) \cdot (-3) + 2 \cdot 5 \cdot 0 + 3 \cdot (-2) \cdot 4 - \\ -0 \cdot (-1) \cdot 4 - 1 \cdot (-2) \cdot 5 - (-3) \cdot 2 \cdot 3 = \\ = 3 + 0 - 24 - 0 + 10 + 18 = 7 \end{gathered}$$

Observación 108 *Si A es de orden 3 entonces agrupando algunos términos tenemos que*

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

donde estos determinantes son las partes de A que resultan de eliminar la fila y la columna correspondiente al elemento que los acompaña multiplicándolos.



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

Definición 109 *Dada una matriz $A = (a_{ij}) \in \mathcal{M}_{n\times n}$*

1. *Se denomina* ***menor complementario del elemento*** *a_{ij} de A para cada $i, j = 1, \ldots, n$ al determinante de la matriz M_{ij} de orden $n-1$ que resulta de A al eliminar la fila i-ésima y la columna j-ésima, es decir*

$$|M_{ij}| = \begin{vmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

2. *El* **adjunto** *o* **cofactor del elemento** a_{ij} *para cada* $i, j = 1, \ldots, n$ *que se denota como* A_{ij} *y se define como*

$$A_{ij} = (-1)^{i+j} |M_{ij}|$$

Una manera de recordar los signos de los adjuntos ($(-1)^{i+j}$) es mediante la siguiente matriz de signos

$$\begin{pmatrix} + & - & + & \cdots & & \\ - & + & - & \cdots & & \\ + & - & + & \cdots & & \\ \vdots & \vdots & \vdots & \ddots & & \\ & & & & + & - \\ & & & & - & + \end{pmatrix}$$

Observación 110 *Según las definiciones anteriores, la expresión de* $|A|$ *queda como*

$$|A| = a_{11}A_{11} + a_{12}A_{12} + a_{13}A_{13}$$

aunque operando se puede ver que

$$|A| = a_{12}A_{12} + a_{22}A_{22} + a_{32}A_{32}$$

En general podemos desarrollar el determinante de una matriz mediante la suma de los productos de los elementos de una fila (o columna) por sus respectivos adjuntos. Así si A *es una matriz de orden* n *se tiene que para cualquier* $k = 1, \ldots, n$

$$|A| = \sum_{j=1}^{n} a_{kj}A_{kj} \qquad \textit{(por filas)}$$

o de igual forma

$$|A| = \sum_{i=1}^{n} a_{ik}A_{ik} \qquad \textit{(por columnas)}$$

De esta manera, mediante el desarrollo por adjuntos, el problema de calcular el determinante de una matriz A *de orden* n *se reduce a calcular* n *determinantes de orden* $n-1$.

Incluso se puede reiterar el proceso hasta reducirlo al cálculo de determinantes de orden 2 o 3 de los que ya conocemos su fórmula.

Ejemplo 111 *Sea* $A = \begin{pmatrix} 1 & 2 & 0 & 9 \\ 2 & 3 & 4 & 6 \\ 1 & 6 & 0 & -1 \\ 0 & -5 & 0 & 8 \end{pmatrix}$, *entonces desarrollando por la segunda fila tenemos que*

$$|A| = 2 \cdot (-1) \begin{vmatrix} 2 & 0 & 9 \\ 6 & 0 & -1 \\ -5 & 0 & 8 \end{vmatrix} + 3 \cdot (1) \begin{vmatrix} 1 & 0 & 9 \\ 1 & 0 & -1 \\ 0 & 0 & 8 \end{vmatrix} +$$

$$+4 \cdot (-1) \begin{vmatrix} 1 & 2 & 9 \\ 1 & 6 & -1 \\ 0 & -5 & 8 \end{vmatrix} + 6 \cdot (1) \begin{vmatrix} 1 & 2 & 0 \\ 1 & 6 & 0 \\ 0 & -5 & 0 \end{vmatrix} =$$

$$= -2 \cdot 0 + 3 \cdot 0 - 4 \cdot (-18) + 6 \cdot 0 = 72$$

También podemos desarrollar por la tercera columna que es la que más ceros tiene, así

$$|A| = 0\cdot(1)\begin{vmatrix} 2 & 3 & 6 \\ 1 & 6 & -1 \\ 0 & -5 & 8 \end{vmatrix} + 4\cdot(-1)\begin{vmatrix} 1 & 2 & 9 \\ 1 & 6 & -1 \\ 0 & -5 & 8 \end{vmatrix} +$$

$$+0\cdot(1)\begin{vmatrix} 1 & 2 & 9 \\ 2 & 3 & 6 \\ 0 & -5 & 8 \end{vmatrix} + 0\cdot(-1)\begin{vmatrix} 1 & 2 & 9 \\ 2 & 3 & 6 \\ 1 & 6 & -1 \end{vmatrix} =$$

$$= 0 - 4\cdot(-18) + 0 + 0 = 72$$

3.3.2. Propiedades de los determinantes

1. Sea $A \in \mathcal{M}_{n\times n}$ y A^T su matriz transpuesta, entonces
$$|A| = \left|A^T\right|$$

2. $$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1}+b_{j1} & a_{j2}+b_{j2} & \cdots & a_{jn}+b_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{j1} & b_{j2} & \cdots & b_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

3. Para todo $\lambda \in \mathbb{R}$ se tiene que $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{j1} & \lambda a_{j2} & \cdots & \lambda a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \lambda \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$

4. Sea $A \in \mathcal{M}_{n\times n}$ y $\lambda \in \mathbb{R}$, entonces
$$|\lambda A| = \lambda^n \cdot |A|$$

5. $$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

6. $$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = - \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{in} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix}$$

7. $\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{j1} & a_{j2} & \cdots & a_{jn} \\ \vdots & \vdots & & \vdots \\ \lambda a_{j1} & \lambda a_{j2} & \cdots & \lambda a_{jn} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n,j+1} & \cdots & a_{nn} \end{vmatrix} = 0$

8. Si en una matriz $A \in \mathcal{M}_{n\times n}$, a una de sus filas (respectivamente columnas) se le suma una combinación lineal de las filas (respectivamente columnas) restantes, el determinante resultante también es $|A|$.

9. Sean $A, B \in \mathcal{M}_{n\times n}$ entonces $|AB| = |A|\,|B|$.

10. Sean $A \in \mathcal{M}_{n\times n}$ y $k \in \mathbb{N}$, con $k \neq 0$, entonces $\left|A^k\right| = |A|^k$.

11. Sea $A \in \mathcal{M}_{n\times n}$ no singular, entonces $\left|A^{-1}\right| = \dfrac{1}{|A|}$.

12. La suma de productos de los elementos de una fila (respectivamente columna) de una matriz $A \in \mathcal{M}_{n\times n}$ por los adjuntos de otra fila (respectivamente columna) diferente, es igual a 0, es decir $\sum\limits_{j=1}^{n} a_{ij}A_{kj} = 0$ para todo $i, k = 1, \ldots, n$ con $k \neq i$.

3.3.3. Rango de una matriz

Hemos visto el rango de una matriz como el número de vectores columna (o fila) linealmente independientes.

Vamos a utilizar el determinante para calcular el rango de una matriz.

Definición 112 *Diremos que S es una **submatriz** S de la matriz A si S resulta de eliminar un número de filas y columnas enteras de A.*

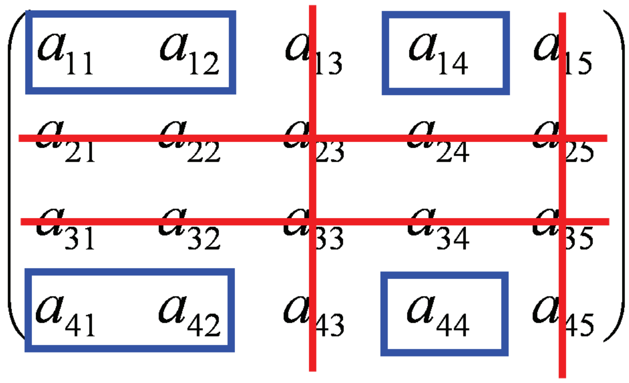
Ejemplo 113 *Si A es la siguiente matriz*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{pmatrix}$$

eliminando, por ejemplo, las columnas tercera y quinta, y las filas segunda y tercera, obtenemos la submatriz

$$S = \begin{pmatrix} a_{11} & a_{12} & a_{14} \\ a_{41} & a_{42} & a_{44} \end{pmatrix}$$

de la matriz A.



Teorema 114 *Dada una matriz $A \in \mathcal{M}_{m\times n}$, se verifica que $rg\,(A) = k$ si y sólo si existe una submatriz $A_k \in \mathcal{M}_{k\times k}$ tal que $|A_k| \neq 0$ y cualquier submatriz S de A de orden mayor que k es tal que $|S| = 0$.*

Corolario 115 *Dada una matriz $A \in \mathcal{M}_{n\times n}$, entonces $rg\,(A) < n$ si y sólo si $|A| = 0$, o equivalentemente, $rg\,(A) = n$ si y sólo si $|A| \neq 0$.*

Corolario 116 *Dada una matriz $A \in \mathcal{M}_{m\times n}$, entonces $rg\,(A) = k$ si y sólo si tiene exactamente k filas linealmente independientes y k columnas linealmente independientes.*

Ejemplo 117 *Si* $A = \begin{pmatrix} 1 & 0 & 1 \\ 3 & 2 & -1 \\ 1 & 1 & 0 \end{pmatrix}$ *entonces* $|A| = 2 \neq 0$ *y por tanto* $rg\,(A) = 3$.

Ejemplo 118 *Si* $A = \begin{pmatrix} 1 & 2 & 0 & 1 & 4 \\ 0 & 1 & -1 & 2 & 5 \\ 4 & 0 & 1 & 0 & 0 \end{pmatrix}$*, como* $rg\,(A) \leq 3$ *por tener 3 filas, si encontramos una submatriz de orden* 3×3 *con determinante distinto de 0 entonces* $rg\,(A) = 3$.

Así $A_3 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \\ 4 & 0 & 0 \end{pmatrix}$ *es tal que* $|A_3| = 4 \cdot \begin{vmatrix} 2 & 1 \\ 1 & 2 \end{vmatrix} = 12 \neq 0$*. Por lo tanto* $rg\,(A) = 3$.

Ejemplo 119 *Si* $A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$*, entonces* $rg\,(A) \leq 4$*. Veamos a ver si se verifica que* $rg\,(A) = 4$*. Para ello calculamos* $|A|$*, así desarrollando por la primera columna*

$$|A| = 1 \cdot \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{vmatrix} = 0$$

por tanto tenemos que A *no tiene rango completo y así* $rg\,(A) < 4$.

Buscamos alguna submatriz de orden 3×3 *(hay 16) tal que su determinante sea distinto de 0. Hay submatrices, como por ejemplo* A_{14} *y* A_{13}*, tales que*

$$A_{14} = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 0 \qquad\qquad A_{13} = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} = 0$$

pero podemos encontrar una submatriz A_{44} *tal que*

$$|A_{44}| = \begin{vmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{vmatrix} = -1 \neq 0$$

por lo tanto $rg\,(A) = 3$.

3.3.4. Cálculo de la inversa de una matriz

Vimos que si $A \in \mathcal{M}_{n\times n}$ tal que $rg\,(A) < n$ entonces no existe A^{-1}, y al contrario, si $rg\,(A) = n$ entonces existe la inversa A^{-1}.

Por lo tanto,

$$\boxed{\text{si } |A| \neq 0 \text{ entonces existe } A^{-1}}$$

$$\boxed{\text{si } |A| = 0 \text{ entonces no existe } A^{-1}}$$

Consideramos la matriz $A = (a_{ij}) \in \mathcal{M}_{n\times n}$ tal que $|A| \neq 0$ y construimos la **matriz adjunta**, que es la formada por los adjuntos de los elementos de A:

$$Adj\,(A) = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

Consideramos entonces el producto

$$A \cdot (Adj\,(A))^T = \begin{pmatrix} \sum\limits_{i=1}^{n} a_{1i}A_{1i} & \cdots & \sum\limits_{i=1}^{n} a_{1i}A_{ni} \\ \vdots & \ddots & \vdots \\ \sum\limits_{i=1}^{n} a_{ni}A_{1n} & \cdots & \sum\limits_{i=1}^{n} a_{ni}A_{ni} \end{pmatrix}$$

y como $\sum\limits_{i=1}^{n} a_{ji}A_{ji} = |A|$ para todo $j = 1, \ldots, n$ y, por la propiedad 12 anterior, $\sum\limits_{i=1}^{n} a_{ji}A_{kn} = 0$ para $k \neq j$ con $j, k = 1, \ldots, n$, entonces

$$A \cdot (Adj\,(A))^T = \begin{pmatrix} |A| & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & |A| \end{pmatrix} = |A| \cdot I_n$$

por lo tanto se tiene que

$$A \cdot \frac{1}{|A|}\,(Adj\,(A))^T = I_n$$

y así

$$\boxed{A^{-1} = \frac{1}{|A|}\,(Adj\,(A))^T}$$

Observación 120 *Tengamos en cuenta las siguientes observaciones:*

1. *Cuando la matriz* A *es de orden grande, esta forma de calcularlo es muy laboriosa.*

2. Se verifica que

$$(Adj\,(A))^T = Adj\,(A^T)$$

Ejemplo 121 *Queremos calcular la matriz inversa de* $A = \begin{pmatrix} 4 & 1 & -1 \\ 0 & 3 & 2 \\ 3 & 0 & 7 \end{pmatrix}$. *Calculamos primero el determinante:* $|A| = 99$.

Ahora calculamos los adjuntos de la matriz A:

$$\begin{array}{lll} A_{11} = \begin{vmatrix} 3 & 2 \\ 0 & 7 \end{vmatrix} = 21 & A_{12} = -\begin{vmatrix} 0 & 2 \\ 3 & 7 \end{vmatrix} = 6 & A_{13} = \begin{vmatrix} 0 & 3 \\ 3 & 0 \end{vmatrix} = -9 \\ A_{21} = -\begin{vmatrix} 1 & -1 \\ 0 & 7 \end{vmatrix} = -7 & A_{22} = \begin{vmatrix} 4 & -1 \\ 3 & 7 \end{vmatrix} = 31 & A_{23} = -\begin{vmatrix} 4 & 1 \\ 3 & 0 \end{vmatrix} = 3 \\ A_{31} = \begin{vmatrix} 1 & -1 \\ 3 & 2 \end{vmatrix} = 5 & A_{32} = -\begin{vmatrix} 4 & -1 \\ 0 & 2 \end{vmatrix} = -8 & A_{33} = \begin{vmatrix} 4 & 1 \\ 0 & 3 \end{vmatrix} = 12 \end{array}$$

Luego,

$$Adj\,(A) = \begin{pmatrix} 21 & 6 & -9 \\ -7 & 31 & 3 \\ 5 & -8 & 12 \end{pmatrix}$$

y por tanto

$$(Adj\,(A))^T = \begin{pmatrix} 21 & -7 & 5 \\ 6 & 31 & -8 \\ -9 & 3 & 12 \end{pmatrix}$$

por lo que se tiene que

$$A^{-1} = \frac{1}{|A|}\,(Adj\,(A))^T = \frac{1}{99}\begin{pmatrix} 21 & -7 & 5 \\ 6 & 31 & -8 \\ -9 & 3 & 12 \end{pmatrix}$$

Ejemplo 122 *Calcular el determinante de la matriz*

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

Vamos a utilizar las propiedades de los determinantes, asi

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

$$= h^2 \begin{vmatrix} 0 & 1 & 1 \\ 0 & h & -h \\ -1 & 0 & h \end{vmatrix} = h^2 \cdot (-1) \begin{vmatrix} 1 & 1 \\ h & -h \end{vmatrix} = -h^2 \cdot (-2h) = 2h^3$$

donde en la primera igualdad hemos desarrollado por la 5ª fila, en la segunda igualdad por la 3ª fila, en la tercera igualdad hemos cambiado la 2ª fila por la combinación lineal 2ª fila - 3ª fila, y en la cuarta desigualdad hemos desarrollado por la 1ª columna.

Ejemplo 123 *Consirerar el conjunto $S = \{(1,1,0), (3,\alpha,2), (0,-1,\alpha)\}$. ¿Para que valor de $\alpha \in \mathbb{R}$ es linealmente independiente S?*

Esto es equivalente a decir preguntarse cuándo se cumple que

$$\begin{vmatrix} 1 & 3 & 0 \\ 1 & \alpha & -1 \\ 0 & 2 & \alpha \end{vmatrix} \neq 0$$

Por lo tanto, como

$$\begin{vmatrix} 1 & 3 & 0 \\ 1 & \alpha & -1 \\ 0 & 2 & \alpha \end{vmatrix} = \alpha^2 - 3\alpha + 2$$

veamos cuando se cumple que

$$\alpha^2 - 3\alpha + 2 = 0$$

por lo que

$$\alpha = 1 \quad y \quad \alpha = 2$$

y así el conjunto S es linealmente independiente para todo $\alpha \in \mathbb{R} \setminus \{1,2\}$.

Ejemplo 124 *Dado el subespacio*

$$W = \left\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 = x_3\right\}$$

estudiar si

$$S_1 = \{(0,0,0,1), (0,1,1,1), (-1,1,0,1)\}$$
$$S_2 = \{(2,1,3,4), (1,0,1,0), (1,1,2,4)\}$$

son sistemas generadores de W.

Como $\dim W = 3$ ya que está descrito por una ecuación, entonces necesitamos al menos 3 vectores para generar W, ya que las bases de W tienen 3 vectores.

Todos los vectores de S_1 y S_2 pertenecen a W ya que verifican la ecuación

$$x_1 + x_2 = x_3$$

(ejercicio: comprobarlo), y por lo tanto sólo necesitamos ver si los vectores de S_1 y los vectores de S_2 son linealmente independientes. Para ello estudiamos el rango de las matrices

$$A_1 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} \qquad y \qquad A_2 = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 0 & 1 \\ 3 & 1 & 2 \\ 4 & 0 & 4 \end{pmatrix}$$

Para A_1 encontramos la submatriz $\begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}$ que verifica que $\begin{vmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{vmatrix} = 1 \neq 0$ y por lo tanto $rg\,(A_3) = 3$ y así deducimos que S_1 es un sistema de generadores.

Por otro lado, para A_2 las cuatro submatrices de orden 3×3

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

tienen determinante igual a 0 (ejercicio: comprobarlo) y como existe una submatriz de orden 2×2, por ejemplo $\begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}$ con determinante no nulo, entonces $rg\,(A_2) = 2$ y por tanto S_2 no es un sistema generador de W ya que uno de sus vectores es combinación lineal de los otros dos.

Ejemplo 125 *¿Para qué valores de $a \in \mathbb{R}$ es invertible la matriz* $A = \begin{pmatrix} a & 1 & 2 \\ 2 & a & 2 \\ 1 & a & 1 \end{pmatrix}$*?*

Tenemos que buscar los valores de $a \in \mathbb{R}$ tales que $rg\,(A) = 3$, es decir, tales que $|A| \neq 0$. Como

$$|A| = a^2 + 4a + 2 - 2a - 2a^2 - 2 = 2a - a^2$$

entonces

$$|A| = 0$$

implica que $\boxed{a = 0}$ o $\boxed{a = 2}$ y en ese caso se tiene que $rg\,(A) < 3$. Por lo tanto, podemos concluir que A es invertible si y sólo si

$$\boxed{a \neq 0} \text{ o } \boxed{a \neq 2}$$

Capítulo 4

Sistemas lineales

4.1. Sistemas lineales

Definición 126 *Un* ***sistema*** *es un conjunto de relaciones entre un conjunto de cantidades desconocidas. A estas cantidades desconocidas las llamaremos* ***incógnitas****. Cuando estas relaciones son lineales entre las incógnitas, diremos que el sistema de ecuaciones es* ***lineal****. La forma general de un sistema lineal será:*

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

donde

a_{ij} *son los* ***coeficientes***
x_j *son las* ***incógnitas***
b_i *son los* ***términos independientes***

Si utilizamos la notación matricial el sistema se puede escribir como

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

o bien

$$A\bar{x} = \bar{b}$$

donde

$A = (a_{ij}) \in \mathcal{M}_{m\times n}$ *es la* ***matriz de coeficientes***
$\bar{x} \in \mathbb{R}^n$ *es el vector de* ***variables*** *o* ***incógnitas***
$\bar{b} \in \mathbb{R}^m$ *es el vector de* ***términos independientes***

A la matriz $\tilde{A} \in \mathcal{M}_{m\times(n+1)}$ formada por A y el vector columna $\bar{b}$ de términos independientes de la forma:

$$\tilde{A} = \left(A|\,\bar{b}\right) = \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array} \right)$$

se le llama ***matriz ampliada*** *del sistema.*

Ejemplo 127 *El sistema* $\left\{ \begin{array}{l} x_1 + x_2 + x_3 = 3 \\ x_1 - 8x_2 + 7x_3 = 5 \end{array} \right.$ *es lineal y su expresión matricial es*

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -8 & 7 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$$

y su matriz ampliada es

$$\tilde{A} = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -8 & 7 & 5 \end{array}\right)$$

Ejemplo 128 *El sistema* $\left\{\begin{array}{l} t^2 x_1 + 2t x_2 = 1 + s \\ e^{ts} x_1 - (s-t)^2 x_2 = 0 \end{array}\right.$ *es lineal en las variables* $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ *y su expresión matricial es*

$$\begin{pmatrix} t^2 & 2t \\ e^{ts} & -(s-t)^2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1+s \\ 0 \end{pmatrix}$$

y su matriz ampliada es

$$\tilde{A} = \left(\begin{array}{cc|c} t^2 & 2t & 1+s \\ e^{ts} & -(s-t)^2 & 0 \end{array}\right)$$

Definición 129 *Sea* $A\bar{x} = \bar{b}$ *un sistema de ecuaciones lineales.*

1. *Se dice que el sistema es* ***homogéneo*** *si* $\bar{b} = \bar{0} \in \mathbb{R}^m$.

2. *Se dice que el sistema es* ***no homogéneo*** *si* $\bar{b} \neq \bar{0} \in \mathbb{R}^m$.

3. *El* ***sistema homogéneo asociado*** *al sistema* $A\bar{x} = \bar{b}$ *es el sistema*

$$A\bar{x} = \bar{0}$$

independientemente del valor de $\bar{b}$.

Definición 130 *Dado el sistema lineal* $A\bar{x} = \bar{b}$, *se dice que un vector* $x^* \in \mathbb{R}^n$ *es* ***solución*** *del mismo si y sólo si se verifican las* m *ecuaciones, es decir*

$$Ax^* = \bar{b}$$

Ejemplo 131 *¿El vector* $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -4 \\ 2 \end{pmatrix}$ *es solución del sistema* $\left\{\begin{array}{l} x_1 + 5x_2 = 6 \\ 3x_1 - 2x_2 = 3 \end{array}\right.$ *?*

Para averiguarlo basta con sustituir el vector en el sistema, así

$$\left\{\begin{array}{ll} (-4) + 5(2) = 6 & \checkmark \\ 3(-4) - 2(2) \neq 3 & \times \end{array}\right.$$

y como no se verifican las 2 ecuaciones entonces $\begin{pmatrix} -4 \\ 2 \end{pmatrix}$ *no es solución del sistema.*

Ejemplo 132 *Averiguar si el vector* $(1,1,1)$ *es una solución del sistema*

$$\left\{\begin{array}{r} x + y + z = 3 \\ 2x - y + 3z = 4 \\ 2y - z = 1 \end{array}\right.$$

Sustituimos en forma matricial, por tanto

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 3 \\ 0 & 2 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} \begin{array}{c} \checkmark \\ \checkmark \\ \checkmark \end{array}$$

y se verifican las 3 ecuaciones, por lo tanto sí es solución del sistema.

4.2. Tipos de sistemas

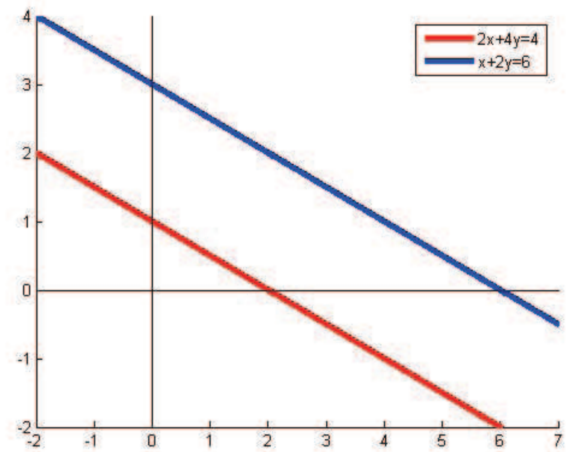
Definición 133 *Sea $A\bar{x} = \bar{b}$ un sistema de ecuaciones lineales tal que $A \in \mathcal{M}_{m\times n}$, $\bar{x} \in \mathbb{R}^n$ y $\bar{b} \in \mathbb{R}^m$. Diremos que el sistema es:*

1. ***Incompatible**: si no tiene solución.*
2. ***Compatible**: si tiene solución.*
   a) ***Compatible determinado**: si la solución es única.*
   b) ***Compatible indeterminado**: si existe más de una solución.*

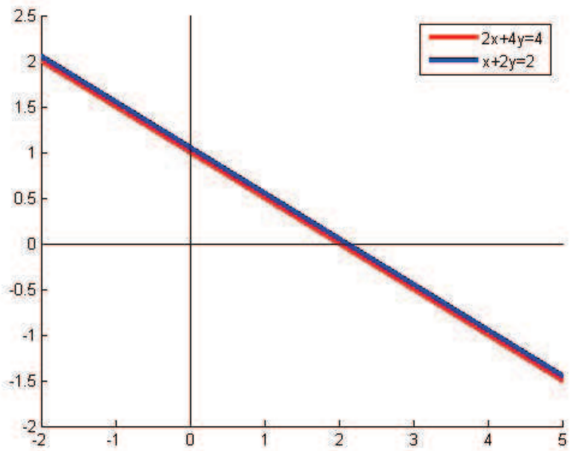
4.2.1. Interpretación geométrica de los sistemas lineales

Si representamos todas las ecuaciones, su intersección será la solución del sistema.
Veámoslo mediante ejemplos en $\mathbb{R}^2$:

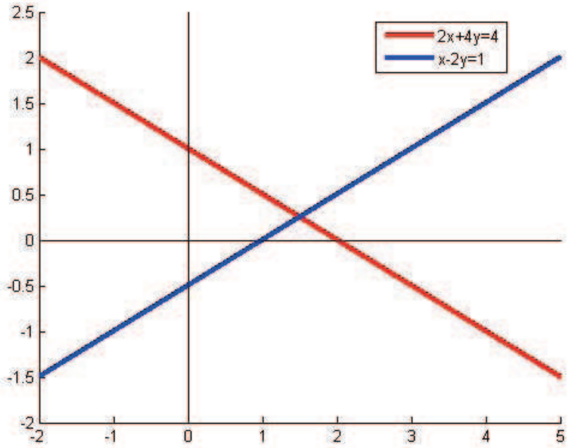
Consideremos el sistema $\left\{ \begin{array}{l} 2x + 4y = 4 \\ x + 2y = 6 \end{array} \right.$. Éste no tiene solución porque lo forman 2 rectas paralelas y no tienen puntos comunes. Luego el sistema es incompatible.



Modifiquemos el sistema como $\left\{ \begin{array}{l} 2x + 4y = 4 \\ x + 2y = 2 \end{array} \right.$. Ahora tiene infinitas soluciones porque las rectas son coincidentes y todos sus puntos son comunes. El sistema es compatible indeterminado.



Si ahora el sistema es $\begin{cases} 2x+4y=4 \\ x-2y=1 \end{cases}$, entonces tiene una única solución ya que las rectas se cortan en un punto. Por lo que el sistema es compatible determinado.



En el caso de sistemas con tres incógnitas, las ecuaciones en $\mathbb{R}^3$ representan planos en el espacio, por lo que la posición relativa entre planos en $\mathbb{R}^3$ contempla más casos distintos que los de las rectas en $\mathbb{R}^2$.

4.3. Planteamiento de sistemas lineales

Dado un problema, será necesario identificar las cantidades desconocidas (variables) y las relaciones entre ellas (ecuaciones) para poder resolverlo.

Ejemplo 134 *Una fábrica de cerámica fabrica floreros y jarrones. Para cada florero y cada jarrón utiliza una cantidad fija de material. Cada artesano invierte 3 minutos en la producción de un florero y 2 minutos en la de un jarrón. El material del florero cuesta 0'25 € y el material del jarrón cuesta 0'20 € ¿Cuántas piezas de cada tipo puede hacer un artesano en una jornada de 8 horas si se gastan 43 € en materiales?*

Las cantidades que queremos calcular son el número de floreros y jarrones que pueden hacerse en las 8 horas de trabajo de cada artesano. Por lo tanto las incógnitas serán:

x_1 *el número de floreros* $\qquad$ x_2 *el número de jarrones*

Los datos que tenemos son:

Se invierten 3 minutos por florero.
Se invierten 2 minutos por jarrón.
Se gastan 0'25 €por florero en material.
Se gastan 0'20 €por jarrón en material.
Cada artesano trabaja 8 horas (= 480 minutos).
Cada artesano utiliza 43 €en material.

Por lo tanto las ecuaciones que se pueden plantear son:

$$\begin{cases} 3x_1 + 2x_2 = 480 & \rightarrow \text{ tiempo (en minutos)} \\ 0'25x_1 + 0'20x_2 = 43 & \rightarrow \text{ dinero del material} \end{cases}$$

Ejercicio 135 *Plantear el siguiente problema:*

Se sabe que a una fiesta van a acudir 60 personas, los organizadores han comprado 100 regalos de manera que los hombres recibirán un regalo y las mujeres 3. ¿Cuántos hombres y mujeres acudirán a la fiesta?

4.4. Sistemas equivalentes

Definición 136 *Dos sistemas de ecuaciones lineales compatibles son* **equivalentes** *si y sólo si tienen el mismo conjunto de soluciones.*

Proposición 137 *Sea el sistema lineal compatible* $A\bar{x} = \bar{b}$ *con* $A \in \mathcal{M}_{m\times n}$, $\bar{x} \in \mathbb{R}^n$ *y* $\bar{b} \in \mathbb{R}^m$, *entonces se verifica que este sistema es equivalente a*

$$BA\bar{x} = B\bar{b}$$

donde B *es cualquier matriz no singular de orden* $m \times n$.

Observación 138 *Esto implica que las operaciones que nos dan lugar a sistemas equivalentes son:*

1. *Permutar 2 ecuaciones.*
2. *Multiplicar una ecuación por un escalar no nulo.*
3. *Sustituir una ecuación cualquiera por la suma de ella misma con una combinación lineal de las restantes.*

Ejemplo 139 *Los sistemas*

$$\left\{\begin{array}{l} 2x+y+z=6 \\ x+y-z=4 \\ 3x+2y=10 \end{array}\right. \qquad y \qquad \left\{\begin{array}{l} 2x+2y-2z=8 \\ 3x+2y=10 \end{array}\right.$$

son equivalentes ya que sus soluciones son:

$$\{(2-2\alpha, 2+3\alpha, \alpha) : \alpha \in \mathbb{R}\}$$

Ejemplo 140 *En el sistema* $A\bar{x} = \bar{b}$ *dado por*

$$\left\{\begin{array}{r} -2y+2z=0 \\ x+3y-2z=2 \\ x+y=2 \end{array}\right.$$

podemos observar que la 2ª ecuación es igual a la 3ª menos la 1ª, es decir

$$1^a \text{ ec.} + 2^a \text{ ec.} - 3^a \text{ ec.} = 0$$

por lo tanto si

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

entonces como B *es no singular entonces* $BA\bar{x} = B\bar{b}$, *esto es*

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

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

por lo que el sistema original es equivalente a

$$\left\{\begin{array}{r} -2y+2z=0 \\ x+y=2 \end{array}\right.$$

4.5. Existencia de soluciones. Teorema de Rouché-Fröbenius

Hemos visto ejemplos de sistemas con solución única, infinitas soluciones y sin solución. ¿Es posible saber cuántas soluciones va a tener un sistema sin necesidad de resolverlo?

Teorema 141 *(**Rouché-Fröbenius** - Versión 1) Dado un sistema lineal con m ecuaciones y n incógnitas*

$$A\bar{x} = \bar{b}$$

con $A \in \mathcal{M}_{m\times n}$, $\bar{x} \in \mathbb{R}^n$ y $\bar{b} \in \mathbb{R}^m$ donde la matriz ampliada es

$$\tilde{A} = \left(A|\,\bar{b}\right) \in \mathcal{M}_{m\times(n+1)}$$

entonces se verifica que:

1. *El sistema es incompatible si y sólo si $rg\,(A) \neq rg\left(\tilde{A}\right)$.*

2. *El sistema es compatible si y sólo si $rg\,(A) = rg\left(\tilde{A}\right)$.*

   a) *Compatible determinado si y sólo si $rg\,(A) = rg\left(\tilde{A}\right) = n$.*

   b) *Compatible indeterminado si y sólo si $rg\,(A) = rg\left(\tilde{A}\right) < n$.*

Ejemplo 142 *En el siguiente sistema, determinar la compatibilidad según los distintos valores de $a \in \mathbb{R}$*

$$\left\{ \begin{aligned} ax - 2y + z &= 1 \\ x + ay + 2z &= a \\ x + z &= 1 \end{aligned} \right.$$

Tendremos que estudiar los $rg\,(A)$ y $rg\left(\tilde{A}\right)$ siendo

$$A = \begin{pmatrix} a & -2 & 1 \\ 1 & a & 2 \\ 1 & 0 & 1 \end{pmatrix} \qquad y \qquad \tilde{A} = \left(\begin{array}{ccc|c} a & -2 & 1 & 1 \\ 1 & a & 2 & a \\ 1 & 0 & 1 & 1 \end{array}\right)$$

Calculamos entonces $|A|$, así

$$|A| = \begin{vmatrix} a & -2 & 1 \\ 1 & a & 2 \\ 1 & 0 & 1 \end{vmatrix} = a^2 - a - 2$$

entonces si $|A| = 0$ se tiene que $rg\,(A) < 3$.

Así $a^2 - a - 2 = 0$ implica que $a = 2$ o $a = -1$, y en este caso como encontramos una submatriz de A de orden 2×2 con determinante no nulo

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

entonces tenemos que $rg\,(A) = 2$.

Para calcular $rg\left(\tilde{A}\right)$ completamos la matriz $\begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$ con la que hemos obtenido que $rg\,(A) = 2$, con una fila y una columna de $\tilde{A}$ para ver si la submatriz 3×3 así resultante tiene determinante no nulo. Por tanto, añadimos los elementos de la primera fila y la cuarta columna, con lo que su determinante queda:

$$\begin{vmatrix} a & -2 & 1 \\ 1 & a & a \\ 1 & 0 & 1 \end{vmatrix} = a^2 - 3a + 2$$

que se anula con $a = 2$ o $a = 1$.

Por lo tanto, resumiendo

$$\begin{array}{rc|l} rg\,(A) = 3 \ si & \boxed{\begin{array}{c} a \neq 2 \\ a \neq -1 \end{array}} & rg\left(\tilde{A}\right) = 3 \ si \ \boxed{a \neq 2} \\ rg\,(A) = 2 \ si & \boxed{\begin{array}{c} a = 2 \\ a = -1 \end{array}} & rg\left(\tilde{A}\right) = 2 \ si \ \boxed{a = 2} \end{array}$$

entonces

$$\begin{array}{rcl} 2 = rg\,(A) \neq rg\left(\tilde{A}\right) = 3 \quad si \ \boxed{a = -1} & \Rightarrow & Incompatible \\ rg\,(A) = rg\left(\tilde{A}\right) = 2 \quad si \ \boxed{a = 2} & \Rightarrow & Comp.\ Indeterminado \\ rg\,(A) = rg\left(\tilde{A}\right) = 3 \quad si \ \boxed{a \neq -1} \ y \ \boxed{a \neq 2} & \Rightarrow & Comp.\ Determinado \end{array}$$

Observación 143 *Cualquier sistema homogéneo $A\bar{x} = \bar{0}$ es siempre compatible ya que*

$$\tilde{A} = (\,A|\,\bar{0})$$

y así

$$rg\,(A) = rg\left(\tilde{A}\right)$$

La solución $\boxed{\bar{x} = \bar{0}}$ es siempre solución del sistema homogéneo, aunque si $rg\,(A) = rg\left(\tilde{A}\right) < n$ entonces existen más soluciones. A la solución $\boxed{\bar{x} = \bar{0}}$ la llamaremos ***solución trivial****.*

Teorema 144 *(**Rouché-Fröbenius** - Versión 2) Dado un sistema lineal con m ecuaciones y n incógnitas*

$$A\bar{x} = \bar{b}$$

supuesto que f es la aplicación lineal

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

que tiene a A como matriz asociada respecto de las bases canónicas, se verifica que:

1. *El sistema es incompatible si y sólo si $\bar{b} \notin \mathrm{Im}\,(f)$.*
2. *El sistema es compatible si y sólo si $\bar{b} \in \mathrm{Im}\,(f)$.*
    - a) *Compatible determinado si y sólo si $\dim\left(\mathrm{Im}\,(f)\right) = n$.*
    - b) *Compatible indeterminado si y sólo si $\dim\left(\mathrm{Im}\,(f)\right) < n$.*

Ejemplo 145 *Consideremos el siguiente sistema*

$$\left\{ \begin{array}{r} x + y + z = 1 \\ x - y + 2z = 2 \\ x + 3y = 5 \end{array} \right.$$

que tiene matrices

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 1 & 3 & 0 \end{pmatrix} \qquad y \qquad \tilde{A} = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 1 & -1 & 2 & 2 \\ 1 & 3 & 0 & 5 \end{array}\right)$$

La aplicación lineal $f : \mathbb{R}^3 \to \mathbb{R}^3$ que tiene a A como matriz asociada respecto de la base canónica es

$$f\,(x, y, z) = (x + y + z, x - y + 2z, x + 3y)$$

y su imagen, por lo tanto es

$$\operatorname{Im}(f) = \mathcal{L}\{(1,1,1), (1,-1,3), (1,2,0)\} = \mathcal{L}\{(1,1,1), (1,2,0)\}$$

y su expresión analítica se calcula como

$$(u_1, u_2, u_3) = \alpha(1,1,1) + \beta(1,2,0) = (\alpha+\beta, \alpha+2\beta, \alpha)$$

así

$$\begin{cases} u_1 = \alpha + \beta \\ u_2 = \alpha + 2\beta \\ u_3 = \alpha \end{cases}$$

entonces podemos eliminar los parámetros α, β y queda

$$\boxed{u_2 = 2u_1 - u_3}$$

por lo que

$$\operatorname{Im}(f) = \left\{(u_1, u_2, u_3) \in \mathbb{R}^3 : u_2 = 2u_1 - u_3\right\}$$

Así, como el vector $\bar{b} = (1,2,5)$ no verifica la ecuación de la imagen, entonces

$$\bar{b} = (1,2,5) \notin \operatorname{Im}(f)$$

por lo que el sistema es incompatible.

Ejemplo 146 *Estudiemos el sistema*

$$\begin{cases} 7x - 3y - 3z = 7 \\ 2x + 4y + z = 0 \\ -2y - z = 2 \end{cases}$$

Sus matrices son

$$A = \begin{pmatrix} 7 & -3 & -3 \\ 2 & 4 & 1 \\ 0 & -2 & -1 \end{pmatrix} \qquad y \qquad \tilde{A} = \left(\begin{array}{ccc|c} 7 & -3 & -3 & 7 \\ 2 & 4 & 1 & 0 \\ 0 & -2 & -1 & 2 \end{array}\right)$$

Si $f : \mathbb{R}^3 \to \mathbb{R}^3$ es la aplicación lineal que tiene a A como matriz asociada respecto de la base canónica y como

$$rg(A) = \dim(\operatorname{Im}(f)) = 3$$

entonces se tiene que

$$\operatorname{Im}(f) = \mathbb{R}^3$$

y entonces como $\bar{b} = (7,0,2) \in \operatorname{Im}(f)$ tenemos que el sistema es compatible determinado.

4.6. Propiedades de las soluciones de un sistema de ecuaciones lineales

Proposición 147 *Dado el sistema de ecuaciones homogéneo*

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0 \end{cases}$$

se verifica que

1. *El conjunto de soluciones S es un subespacio vectorial de $\mathbb{R}^n$.*

2. *Si A es la matriz de coeficientes del sistema y $rg(A) = r$ entonces $\dim(S) = n - r$.*

Observación 148 *Observemos que*

1. *Como S es un subespacio vectorial de $\mathbb{R}^n$, entonces la suma de soluciones y el producto por un escalar de una solución es también una solución.*

2. *La ecuación $\dim(S) = n - r$ no es otra que*

$$\dim(\ker(f)) = \dim(\mathbb{R}^n) - \dim(\mathrm{Im}(f))$$

ya que

$$\begin{aligned} S &= \ker(f) \\ rg(A) &= \dim(\mathrm{Im}(f)) \end{aligned}$$

donde f es la aplicación lineal asociada a la matriz de coeficientes del sistema respecto de las bases canónicas.

Ejemplo 149 *Consideremos el sistema*

$$\left\{ \begin{aligned} x + 3y + 2z = 0 \\ -x + 5y + 3z = 0 \\ -3x + 7y + 4z = 0 \end{aligned} \right.$$

Su matriz de coeficientes es

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

Como $|A| = 0$ y como $\begin{vmatrix} 7 & -3 \\ 2 & 4 \end{vmatrix} = 34 \neq 0$ entonces

$$rg(A) = 2$$

Por lo tanto

$$\dim(S) = 3 - 2 = 1$$

donde S es el espacio de soluciones del sistema. Se deja como ejercicio demostrar que $S = \mathcal{L}\{(-1, -5, 8)\}$.

¿Qué sucede cuando el sistema no es homogéneo?

Proposición 150 *Sea el sistema de ecuaciones lineales*

$$A\bar{x} = \bar{b} \qquad [1]$$

y su sistema homogéneo asociado

$$A\bar{x} = \bar{0} \qquad [2]$$

Si S es el conjunto de soluciones de [2], el sistema [1] es compatible y una de sus soluciones es $x^* \in \mathbb{R}^n$ entonces el conjunto S^* de soluciones del sistema [1] es

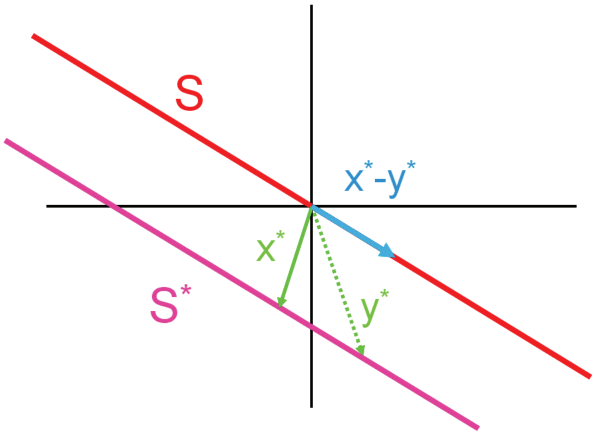
$$S^* = \{\bar{x}^*\} + S = \{\bar{w} \in \mathbb{R}^n : \bar{w} = \bar{x}^* + u \qquad \text{con } u \in S\}$$

Observación 151 *Debemos tener en cueta que*

1. *En este caso S^* no es un subespacio vectorial de $\mathbb{R}^n$ ya que la solución trivial $\bar{x} = (0, \ldots, 0)$ no pertenece a S^*.*

2. *A la solución $\bar{x}^*$ se la llama **solución particular** del sistema.*

4.6.1. Interpretación geométrica



Ejemplo 152 *Estudiemos el sistema*

$$\left\{\begin{array}{r} -2y+2z=0 \\ x+3y-2z=2 \\ x+y=2 \end{array}\right.$$

Sus matrices son

$$A=\left(\begin{array}{ccc} 0 & -2 & 2 \\ 1 & 3 & -2 \\ 1 & 1 & 0 \end{array}\right) \qquad y \qquad \tilde{A}=\left(\begin{array}{ccc|c} 0 & -2 & 2 & 0 \\ 1 & 3 & -2 & 2 \\ 1 & 1 & 0 & 2 \end{array}\right)$$

y son tales que $rg\,(A)=2$ *y* $rg\left(\tilde{A}\right)=2$*, por lo que el sistema es compatible indeterminado.*

Si S *es el conjunto de soluciones del sistema homogéneo asociado*

$$\left\{\begin{array}{r} -2y+2z=0 \\ x+3y-2z=0 \\ x+y=0 \end{array}\right.$$

entonces

$$\dim(S)=3-rg\,(A)=1$$

Como la segunda ecuación se puede obtener de la resta de la tercera menos la primera, entonces el sistema homogéneo es equivalente a

$$\left\{\begin{array}{r} -2y+2z=0 \\ x+y=0 \end{array}\right.$$

y así tenemos que

$$x=-y \qquad\qquad z=y$$

por lo que si $(x,y,z)\in S$ *entonces*

$$(x,y,z)=(-y,y,y)=y\,(-1,1,1)$$

y así tenemos que

$$S=\mathcal{L}\{(-1,1,1)\}=\{(-\alpha,\alpha,\alpha):\alpha\in\mathbb{R}\}$$

y como una solución particular del sistema no homogéneo es

$$\bar{x}^*=(1,1,1)$$

entonces el conjunto de soluciones S^* *es*

$$S^*=(1,1,1)+S=\{(1-\alpha,1+\alpha,1+\alpha):\alpha\in\mathbb{R}\}$$

4.7. Regla de Cramer

Si $A\bar{x} = \bar{b}$ es un sistema compatible determinado con $A \in \mathcal{M}_{n\times n}$, su solución única será

$$\bar{x}^* = A^{-1}\bar{b}$$

Para este tipo de sistemas tenemos un método que ofrece la solución de forma sencilla.

Proposición 153 *Dado el sistema de ecuaciones lineales $A\bar{x} = \bar{b}$ con $A \in \mathcal{M}_{n\times n}$ y $|A| \neq 0$, entonces se verifica que su única solución $\bar{x}^* = (x_1^*, \ldots, x_n^*)$ viene dado por*

$$x_i^* = \frac{|B_i|}{|A|}$$

donde B_i es la matriz de orden $n \times n$ que se obtiene sustituyendo la columna i-ésima de A por el vector $\bar{b}$ de términos independientes.

Ejemplo 154 *El sistema*

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

es compatible determinado ya que $|A| = -3 \neq 0$ y así se tiene que $rg\,(A) = 3 = rg\left(\tilde{A}\right)$. Podemos calcular su solución como

$$\bar{x}^* = A^{-1}\bar{b} =$$

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

Este cálculo necesita bastante trabajo al tener que calcular la inversa.

Por la regla de Cramer tenemos que

$$x_1^* = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & -2 & 1 \end{vmatrix}}{|A|} = -1$$

$$x_2^* = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 1 \\ 1 & 1 & 1 \end{vmatrix}}{|A|} = 0$$

$$x_3^* = \frac{\begin{vmatrix} 1 & 1 & 1 \\ 2 & 0 & 0 \\ 1 & -2 & 1 \end{vmatrix}}{|A|} = 2$$

y así obtenemos también $\bar{x}^* = (-1, 0, 2)$.

Ejemplo 155 *El sistema*

$$\left\{ \begin{aligned} x + z &= 2 \\ 2x + 3y + z &= 3 \\ 3x + 3y + 2z &= 5 \end{aligned} \right.$$

es compatible indeterminado con rango 2 (¡ejercicio!).

¿Podemos utilizar la Regla de Cramer? ¡SI!

Primero eliminamos las ecuaciones que pueden obtenerse mediante combinaciones lineales de otras y pasamos el mismo número de incógnitas al término independiente. En este caso eliminamos la 3ª ecuación y pasamos la incógnita z al termino independiente, así

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

es compatible determinado, aunque su solución (x, y) dependerá del valor de z. Por tanto, por la regla de Cramer tenemos que

$$x^* = \frac{\begin{vmatrix} 2-z & 0 \\ 3-z & 3 \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix}} = 2 - z \qquad\qquad y^* = \frac{\begin{vmatrix} 1 & 2-z \\ 2 & 3-z \end{vmatrix}}{\begin{vmatrix} 1 & 0 \\ 2 & 3 \end{vmatrix}} = \frac{z-1}{3}$$

Por lo que la solución del sistema completo es

$$S = \left\{(x, y, z) \in \mathbb{R}^3 : x = 2 - z, y = \frac{z-1}{3}\right\} = \left\{\left(2 - \alpha, \frac{\alpha - 1}{3}, \alpha\right) : \alpha \in \mathbb{R}\right\}$$

4.8. Método de Gauss-Jordan

Ya vimos como utilizar el método de Gauss-Jordan para calcular la inversa de una matriz.

A continuación vamos a adaptarlo para encontrar la solución del sistema de ecuaciones lineales $A\bar{x} = \bar{b}$.

De nuevo tenemos ciertas operaciones válidas:

1. Intercambiar ecaciones.

2. Multiplicar una ecuación por un escalar.

3. Sumar una ecuación a un múltiplo de otra.

Buscaremos mediante este método unas transformaciones elementales

$$P_1, \ldots, P_r$$

tales que

$$B = P_r \cdot \cdots \cdot P_1 \cdot A$$

es una matriz triangular.

Así, los tenemos la equivalencia entre los sistemas de matrices

$$\left(A |\, \bar{b}\right) \sim \left(B |\, P_r \cdots P_1 \bar{b}\right)$$

Ejemplo 156 *Trataremos de resolver el sistema $A\bar{x} = \bar{b}$ de un ejemplo anterior dado por:*

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

Hemos visto que $rg\,(A) = rg\left(\tilde{A}\right) = 3$, por lo que el sistema es compatible determinado.

Así, buscamos transformaciones elementales para que el sistema sea equivalente a uno triangular

$$\left\{\begin{array}{c} 1^a \\ 2 \cdot 1^a - 2^a \\ 1^a - 3^a \end{array}\right. \qquad\qquad P_1 A = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & 3 & 0 & 0 \end{array}\right)$$

$$\left\{\begin{array}{c} 1^a \\ 2^a \\ 3\cdot 2^a - 2\cdot 3^a \end{array}\right. \qquad P_2P_1A = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 2 \\ 0 & 0 & 3 & 6 \end{array}\right)$$

y esta matriz representa al sistema

$$\left\{\begin{array}{r} x+y+z=1 \\ 2y+z=2 \\ 3z=6 \end{array}\right.$$

por lo que sustituyendo de atrás hacia adelante obtenemos la solución

$$(x,y,z) = (-1,0,2)$$

Ejemplo 157 *De igual modo trataremos de resolver el sistema $A\bar{x}=\bar{b}$ dado por:*

$$\begin{pmatrix} 1 & 0 & 1 \\ 2 & 3 & 1 \\ 3 & 3 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ 5 \end{pmatrix}$$

Ya vimos que $rg\,(A) = rg\left(\tilde{A}\right) = 2$, por lo que el sistema es compatible indeterminado. Así,

$$\left\{\begin{array}{c} 1^a \\ -2\cdot 1^a + 2^a \\ -3\cdot 1^a + 3^a \end{array}\right. \qquad P_1A = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 3 & -1 & -1 \\ 0 & 3 & -1 & -1 \end{array}\right)$$

$$\left\{\begin{array}{c} 1^a \\ 2^a \\ -2^a + 3^a \end{array}\right. \qquad P_2P_1A = \left(\begin{array}{ccc|c} 1 & 0 & 1 & 2 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

y esta matriz representa al sistema

$$\left\{\begin{array}{r} x+z=2 \\ 3y-z=-1 \end{array}\right.$$

por lo que pasando una variable al término independiente

$$\left\{\begin{array}{r} x=2-z \\ 3y=-1+z \end{array}\right.$$

obtenemos la solución

$$(x,y,z) = \left(2-z, \frac{-1+z}{3}, z\right)$$

Ejemplo 158 *Vamos a tratar de resolver el siguiente sistema mediante el método de Gauss-Jordan:*

$$\left\{\begin{array}{r} x+y+z=1 \\ x+2y-z=1 \\ 2x+3y=4 \end{array}\right.$$

Compruebe $rg\,(A) = 2 \neq 3 = rg\left(\tilde{A}\right)$ por lo que el sistema es incompatible. Así,

$$\left\{\begin{array}{c} 1^a \\ -1^a + 2^a \\ -2\cdot 1^a + 3^a \end{array}\right. \qquad P_1A = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & 1 & -2 & 2 \end{array}\right)$$

$$\left\{\begin{array}{c} 1^a \\ 2^a \\ -2^a + 3^a \end{array}\right. \qquad P_2P_1A = \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 2 \end{array}\right)$$

por lo que el sistema es equivalente a

$$\left\{ \begin{array}{r} x+y+z=1 \\ y-2z=0 \\ 0=2 \end{array} \right.$$

que en vista de la última ecuación es evidente que no tiene solución.

Ejemplo 159 *Vamos a estudiar la existencia de solución del sistema*

$$\left\{ \begin{array}{r} x+y+z=3 \\ x-y+z=1 \\ 2x+az=b \end{array} \right.$$

dependiendo de los parámetros a y b. La matriz de coeficientes y la matriz ampliada son:

$$A=\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 2 & 0 & a \end{pmatrix} \qquad \tilde{A}=\left(\begin{array}{ccc|c} 1 & 1 & 1 & 3 \\ 1 & -1 & 1 & 1 \\ 2 & 0 & a & b \end{array}\right)$$

Como $|A|=4-2a$ entonces

$$|A|=0 \Longleftrightarrow a=2$$

por lo que $rg\,(A)=3=rg\left(\tilde{A}\right)$ si $a\neq 2$, y así el sistema será compatible determinado si $a\neq 2$.

Por otro lado, si $a=2$ tenemos que $rg\,(A)=2$ ya que $\begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}=-2\neq 0$. Para saber si, en este caso, el sistema tiene solución tendremos que estudiar el $rg\left(\tilde{A}\right)$. Así, completamos la submatriz $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ con los elementos de la tercera fila y la cuarta columna de $\tilde{A}$, y como

$$\begin{vmatrix} 1 & 1 & 3 \\ 1 & -1 & 1 \\ 2 & 0 & b \end{vmatrix}=8-2b$$

que se anula cuando $b=4$, tenemos que si $b\neq 4$ entonces $rg\left(\tilde{A}\right)=3$, y si $b=4$ entonces $rg\left(\tilde{A}\right)=2$.

Resumiendo:

1. *Si $a\neq 2$ y para todo $b\in\mathbb{R}$ el sistema es compatible determinado ya que $rg\,(A)=rg\left(\tilde{A}\right)=3$.*
2. *Si $a=2$ y $b=4$ el sistema es compatible indeterminado ya que $rg\,(A)=rg\left(\tilde{A}\right)=2$.*
3. *Si $a=2$ y $b\neq 4$ el sistema es incompatible ya que $rg\,(A)=2$ y $rg\left(\tilde{A}\right)=3$.*

Ejercicio 160 *Resolver el sistema anterior para los casos compatibles, es decir para $a\neq 2$ y cualquier $b\in\mathbb{R}$, y para $a=2$ y $b=4$.*

Capítulo 5

Autovalores y autovectores. Diagonalización de matrices

5.1. Introducción

Supongamos que una aplicación lineal f tiene a $A \in \mathcal{M}_{n\times n}$ como matriz asociada respecto de las bases canónicas.

Si A fuese diagonal entonces si $\vec{e}_i$ es el vector i-ésimo de la base canónica entonces

$$f(\vec{e}_i) = \lambda_i \vec{e}_i$$

y esto nos simplificaría el estudio de la aplicación lineal f.

¿Existe una base respecto de la que la matriz de f sea diagonal ?

Ejemplo 161 *Consideremos la aplicación $f(x,y) = (y,4x)$ cuya matriz asociada respecto de las bases canónicas es $A_C = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$.*

Observemos cuáles son los transformados de la base $B = \{(1,2), (1,-2)\}$, en efecto

$$\begin{aligned} f(1,2) &= (2,4) = 2\cdot(1,2) \\ f(1,-2) &= (-2,4) = -2\cdot(1,-2) \end{aligned}$$

vemos que se transforman en un múltiplo suyo.

Entonces la matriz de la aplicación f respecto de la base B será la matriz diagonal

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

También se puede observar que los subespacios

$$\begin{aligned} V_1 &= \mathcal{L}\{(1,2)\} \\ V_2 &= \mathcal{L}\{(1,-2)\} \end{aligned}$$

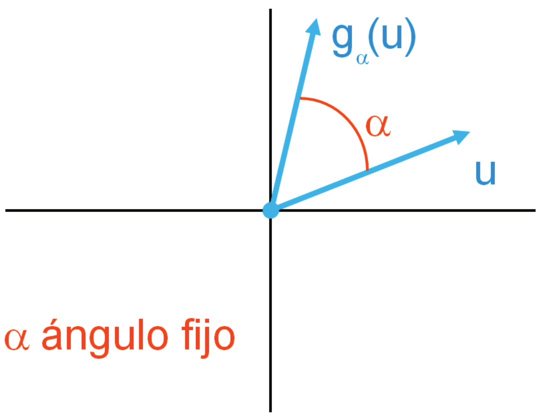
son invariantes mediante f, es decir que

$$\begin{aligned} f(V_1) &= V_1 \\ f(V_2) &= V_2 \end{aligned}$$

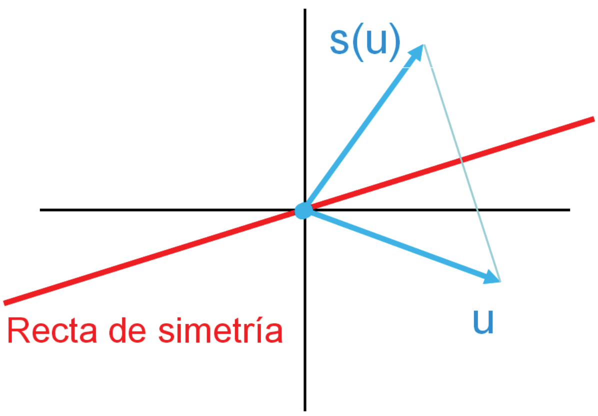
Observación 162 *No siempre existen subespacios (propios) invariantes. Si consideramos las aplicaciones lineales*

$$\begin{aligned} g_\alpha &: \mathbb{R}^2 \to \mathbb{R}^2 \\ s &: \mathbb{R}^2 \to \mathbb{R}^2 \end{aligned}$$

definidas como un giro de ángulo α alrededor del $(0,0)$ y una simetría respecto de una recta que pasa por $(0,0)$ respectivamente, entonces geométricamente se puede observar que g_α no tiene ningún subespacio invariante y s deja invariante la recta de simetría y su perpendicular.



La aplicación g_α no deja ningún subespacio invariante (excepto $\{(0,0)\}$ y $\mathbb{R}^2$).



La recta de simetría y su perpendicular son los únicos subespacios que quedan invariantes mediante la aplicación s (además de $\{(0,0)\}$ y $\mathbb{R}^2$).

Vamos a buscar los subespacios invariantes de las aplicaciones lineales. Para ello tendremos que buscar vectores $v \in V$ tales que $f(v)$ sea proporcional a v, es decir:

Definición 163 *Sea V un espacio vectorial de dimensión finita y $f : V \to V$ una aplicación lineal. Se dice que el escalar $\lambda \in \mathbb{R}$ es un **autovalor** o un **valor propio** de f si existe un vector $\vec{v} \in V$ con $\vec{v} \neq 0_V$ tal que*

$$f(\vec{v}) = \lambda \vec{v}$$

*A tal vector $\vec{v} \in V$ le llamaremos **autovector** o **vector propio** asociado al autovalor λ.*

Observación 164 *Si $\vec{v} \in V$ es un autovector de f asociado al autovalor λ entonces cualquier múltiplo de $\vec{v}$ es también autovector asociado al mismo autovalor. En efecto, si $\alpha \in \mathbb{R}$ y si $\vec{w} = \alpha\vec{v}$ entonces*

$$f(\vec{w}) = f(\alpha\vec{v}) = \alpha f(\vec{v}) = \alpha(\lambda\vec{v}) = \lambda(\alpha\vec{v}) = \lambda\vec{w}$$

Observación 165 *Si $\vec{v} \in V$ es un autovector de f asociado al autovalor $\lambda \in \mathbb{R}$ entonces no existe otro autovalor al que esté asociado, ya que si*

$$f(\vec{v}) = \lambda \vec{v} \qquad y \qquad f(\vec{v}) = \mu \vec{v}$$

entonces

$$0_V = f(\vec{v}) - f(\vec{v}) = \lambda \vec{v} - \mu \vec{v} = (\lambda - \mu)\vec{v}$$

Por lo que $(\lambda - \mu) = 0$ y así tenemos que $\lambda = \mu$, luego los dos autovalores coinciden.

Desde el punto de vista de la matriz de la aplicación lineal tenemos la siguiente:

Proposición 166 *Sea $f : V \to V$ una aplicación lineal con matriz $A \in \mathcal{M}_{n\times n}$ respecto de la base $B = \{\vec{u}_1, \dots, \vec{u}_n\}$ de V. Entonces se verifica que una condición necesaria y suficiente para que $\lambda \in \mathbb{R}$ sea una autovalor de f con autovector asociado $\vec{v} \in V$ es que*

$$A\vec{x} = \lambda \vec{x}$$

donde $\vec{x} \in \mathbb{R}^n$ son las coordenadas de $\vec{v}$ respecto de B.

Ejemplo 167 *Hemos visto que la aplicación lineal $f(x,y) = (y, 4x)$ tiene como matriz asociada respecto de la base canónica a $A = \begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix}$.*

Si consideramos los vectores $\vec{v}_1 = (1,2)$ y $\vec{v}_2 = (1,-2)$, que tiene como coordenadas respecto de la base canónica $B_C\left(\mathbb{R}^3\right)$

$$\vec{v}_1 = (1,2)_{B_C} \qquad y \qquad \vec{v}_2 = (1,-2)_{B_C}$$

y entonces

$$\begin{pmatrix} 0 & 1 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{B_C} = 2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}_{B_C}$$

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

Si consideramos la base $B = \{\vec{v}_1, \vec{v}_2\} = \{(1,2), (1,-2)\}$ entonces la matriz de f respecto a B es

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

y las coordenadas de $\vec{v}_1$ y $\vec{v}_2$ respecto de la base B son

$$\vec{v}_1 = (1,0)_B \qquad y \qquad \vec{v}_2 = (0,1)_B$$

Por tanto,

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

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

Ejemplo 168 *Dada la matriz $A = \begin{pmatrix} 1 & a \\ 2 & b \end{pmatrix}$, vamos a calcular los valores de $a, b \in \mathbb{R}$ para que el vector $\vec{v} = (-2,1)$ sea un autovector asociado al autovalor $\lambda = 5$.*

Así

$$\begin{pmatrix} 1 & a \\ 2 & b \end{pmatrix} \begin{pmatrix} -2 \\ 1 \end{pmatrix} = 5 \begin{pmatrix} -2 \\ 1 \end{pmatrix}$$

esto implica que

$$\left\{ \begin{array}{c} -2 + a = -10 \\ -4 + b = 5 \end{array} \right. \Longrightarrow \boxed{\begin{array}{c} a = -8 \\ b = 9 \end{array}}$$

5.2. Cálculo de autovalores y autovectores de una matriz.

Sea $A \in \mathcal{M}_{n\times n}$ la matriz asociada a una aplicación lineal $f : \mathbb{R}^n \to \mathbb{R}^n$ respecto de la base canónica de $\mathbb{R}^n$, entonces tenemos que

Proposición 169 *Las siguientes condiciones son equivalentes:*

1. *λ es autovalor de A.*
2. *El sistema homogéneo $(A - \lambda I)\vec{x} = \vec{0}$ es compatible indeterminado.*
3. *$|A - \lambda I| = 0$*

Así para calcular los autovalores tenemos la condición 3.

Definición 170 *Dada una matriz $A \in \mathcal{M}_{n\times n}$, se denomina **polinomio característico** de A al polinomio de grado n definido por*

$$P_n(\lambda) = |A - \lambda I|$$

siendo

$$P_n(\lambda) = 0$$

*la **ecuación característica** de la matriz A.*

Observación 171

1. *Es claro que los autovalores son las raices del polinomio característico.*
2. *Si $P_n(\lambda) = (\lambda - \lambda_i)^r Q_{n-r}(\lambda)$ donde $Q_{n-r}(\lambda_i) \neq 0$ entonces se dice que $\lambda = \lambda_i$ es un **autovalor múltiple** con **multiplicidad** r.*
3. *Para el valor λ^*, todos las soluciones del sistema $(A - \lambda^* I)\vec{x} = \vec{0}$ son autovectores asociados al autovalor $\lambda = \lambda^*$.*

Ejemplo 172 *Dada $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, su polinomio característico es*

$$P_2(\lambda) = |A - \lambda I| = \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 3 = (\lambda - 1)(\lambda - 3)$$

por lo que sus autovalores son $\boxed{\lambda_1 = 1}$ y $\boxed{\lambda_2 = 3}$ ya que son las raices de $P_2(\lambda)$.

Para calcular sus autovectores plantearemos los sistemas

$$(A - I)\vec{x} = \vec{0} \qquad y \qquad (A - 3I)\vec{x} = \vec{0}$$

Así

$$(A - I)\vec{x} = \begin{pmatrix} 2-\lambda_1 & 1 \\ 1 & 2-\lambda_1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\{\begin{array}{l} x + y = 0 \\ x + y = 0 \end{array}\right. \Longrightarrow \boxed{y = -x}$$

Por lo tanto, los autovectores asociados al autovalor $\lambda_1 = 1$ son de la forma

$$\vec{v} = (x, -x)$$

Luego el subespacio invariante de autovectores asociados a $\lambda_1 = 1$ es

$$V(1) = \mathcal{L}\{(1, -1)\}$$

Por otro lado,

$$(A-3I)\vec{x} = \begin{pmatrix} 2-\lambda_2 & 1 \\ 1 & 2-\lambda_2 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left\{ \begin{array}{c} -x+y=0 \\ x-y=0 \end{array} \right. \Longrightarrow \boxed{y=x}$$

Por lo tanto, los autovectores asociados al autovalor $\lambda_2 = 3$ *son de la forma*

$$\vec{v} = (x, x)$$

Luego el subespacio invariante de autovectores asociados a $\lambda_2 = 3$ *es*

$$V(3) = \mathcal{L}\{(1,1)\}$$

Ejemplo 173 *Consideremos la matriz* $B = \begin{pmatrix} 2 & 0 & 0 \\ -1 & -1 & 1 \\ 3 & 0 & -1 \end{pmatrix}$, *su polinomio característico es*

$$P_3(\lambda) = |B - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 0 \\ -1 & -1-\lambda & 1 \\ 3 & 0 & -1-\lambda \end{vmatrix} = (2-\lambda)(1+\lambda)^2$$

por lo que sus autovalores son $\boxed{\lambda_1 = 2}$ *(simple) y* $\boxed{\lambda_2 = -1}$ *(doble).*

Calcularemos los autovalores de B *resolviendo los sistemas*

$$(B-2I)\vec{x} = \vec{0} \qquad y \qquad (B+I)\vec{x} = \vec{0}$$

Calculamos ahora los autovectores asociados a $\lambda_1 = 2$*:*

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

$$\left\{ \begin{array}{c} -x-3y+z=0 \\ 3x-3z=0 \end{array} \right. \Longrightarrow \begin{array}{c} \boxed{y=0} \\ \boxed{z=x} \end{array}$$

Así, los autovectores asociados al autovalor $\lambda_1 = 2$ *son de la forma*

$$\vec{v} = (x, 0, x)$$

Luego el subespacio invariante de autovectores asociados a $\lambda_1 = 2$ *es*

$$V(2) = \mathcal{L}\{(1,0,1)\}$$

Por otro lado, los asociados a $\lambda_2 = -1$ *se calculan como*

$$(B+I)\vec{x} = \begin{pmatrix} 3 & 0 & 0 \\ -1 & 0 & 1 \\ 3 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\left\{ \begin{array}{c} 3x=0 \\ -x+z=0 \\ 3x=0 \end{array} \right. \Longrightarrow \begin{array}{c} \boxed{z=x=0} \\ \boxed{y \text{ libre}} \end{array}$$

Así, los autovectores asociados al autovalor $\lambda_2 = -1$ *son de la forma*

$$\vec{v} = (0, y, 0)$$

por lo que el subespacio invariante de autovectores asociados a $\lambda_2 = -1$ *es*

$$V(-1) = \mathcal{L}\{(0,1,0)\}$$

Ejemplo 174 *Un giro de ángulo* $\dfrac{\pi}{2}$ *en* $\mathbb{R}^2$ *tiene como matriz*

$$G = \begin{pmatrix} \cos\frac{\pi}{2} & -\operatorname{sen}\frac{\pi}{2} \\ \operatorname{sen}\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

por lo que su polinomio característico es

$$P_2(\lambda) = |G - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

que no tiene raices reales, por lo que la matriz G *no tiene autovalores reales ni, como vimos, hay subespacios invariantes mediante la matriz* G.

Observación 175 *Como se puede ver en los ejemplos anteriores, no siempre vamos a poder encontrar autovalores reales, ni siempre encontraremos tantos autovectores como dimensión tenga el espacio vectorial. Este hecho será fundamental a la hora de diagonalizar una matriz y debemos tenerlo presente.*

Proposición 176 *Si* A *es una matriz de orden* $n \times n$, *se tiene que*

$$|A - \lambda I| = (-\lambda)^n + (-\lambda)^{n-1}\, tr(A) + \cdots + (-\lambda)^{n-k}\, tr_k(A) + \cdots + |A|$$

siendo $tr_k(A)$ *con* $k = 2, \ldots, n-1$ *la suma de todos los menores de orden* k *que contienen en su diagonal principal* k *elementos de la diagonal principal de* A.

Ejemplo 177 *Dada* $A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}$ *tenemos que* $tr(A) = 3$ *y* $|A| = -1$. *Ahora tenemos que* $tr_2(A) = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} + \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$, *por lo que*

$$P_3(\lambda) = -\lambda^3 + 3\lambda^2 - \lambda - 1$$

5.2.1. Propiedades de los autovalores

Proposición 178 *Sea* $A \in \mathcal{M}_{n\times n}$ *una matriz cuyos autovalores son* $\lambda_1, \ldots, \lambda_n$. *estos autovalores pueden ser reales o complejos, iguales o distintos. Se verifica que:*

1. *Los autovalores de* A^t *coinciden con los de* A.
2. *El número de autovalores nulos de una matriz* A *de rango* r *tal que* $r < n$ *es mayor o igual que* $n - r$.
3. $|A| = \prod_{i=1}^{n} \lambda_i$
4. $tr(A) = \sum_{i=1}^{n} \lambda_i$
5. *Si* $\alpha \in \mathbb{R}$ *con* $\alpha \neq 0$ *entonces los autovalores de la matriz* αA *son* $\alpha\lambda_1, \ldots, \alpha\lambda_n$.
6. *Si* A *es no singular entonces* $\dfrac{1}{\lambda_1}, \ldots, \dfrac{1}{\lambda_n}$ *son los autovalores de la matriz inversa* A^{-1}.
7. *Para cualquier número natural* $k \neq 0$ *los autovalores de* A^k *son* $\lambda_1^k, \ldots, \lambda_n^k$.

Ejemplo 179 *Sea* $A = \begin{pmatrix} 1 & 3 & 3 & -3 \\ 0 & 4 & 0 & 0 \\ 2 & -2 & 2 & -2 \\ -1 & 1 & 1 & -1 \end{pmatrix}$*, su polinomio característico es*

$$P_4(\lambda) = \lambda^4 - 6\lambda^3 + 32\lambda \qquad \textit{¡Calcularlo!}$$

y por tanto sus autovalores son

$$\lambda_1 = 4 \text{ (doble)} \qquad \lambda_2 = 0 \qquad \lambda_3 = -2$$

La propiedad 3 anterior implica que

$$|A| = 4 \cdot 4 \cdot 0 \cdot (-2) = 0$$

y así $rg(A) = 3$*. Además por la propiedad 4 tenemos que*

$$tr(A) = 4 + 4 + 0 - 2 = 6$$

y por la propiedad 7 se tiene que A^3 *tiene como autovalores*

$$\mu_1 = 4^3 = 64 \text{ (doble)} \qquad \mu_2 = 0 \qquad \mu_3 = (-2)^3 = -8$$

5.2.2. Propiedades de los autovalores de algunos tipos de matrices

Matrices triangulares

Para una matriz triangular $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{pmatrix}$ se tiene que

$$|A - \lambda I| = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} - \lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} - \lambda \end{vmatrix} = (a_{11} - \lambda) \cdot \ldots \cdot (a_{nn} - \lambda)$$

entonces **los autovalores son los elementos de la diagonal**.

Análogamente se puede razonar para matrices triangulares inferiores.

Matrices simétricas $(A^t = A)$

Si A es una matriz simétrica de orden $n \times n$ entonces los n autovalores de A son números reales.

Matrices ortogonales $(A^{-1} = A^t)$

Si los autovalores de una matriz ortogonal A son números reales entonces son $\lambda = 1$ o $\lambda = -1$. Esto se debe a que si $A^{-1} = A^t$ entonces $\dfrac{1}{\lambda} = \lambda$ y esto implica que $\lambda = \pm 1$.

Matrices idempotentes $(A^2 = A)$

Los autovalores de una matriz idempotente son $\lambda = 0$ o $\lambda = 1$.

Esto es debido a que si $A^2 = A$ entonces $\lambda^2 = \lambda$ y esto implica que $\lambda(\lambda - 1) = 0$, por lo que $\lambda = 0$ o $\lambda = 1$.

5.2.3. Propiedades de los autovectores

Proposición 180 *Sea $A \in \mathcal{M}_{n\times n}$ una matriz cuyos autovalores son $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ con multiplicidades $m_1, \ldots, m_r$ respectivamente, tales que $\sum_{i=1}^{n} m_i = n$. Se verifica que:*

1. *Para cada $i = 1, \ldots, r$, el conjunto*
$$V(\lambda_i) = \{\vec{v} \in \mathbb{R}^n : A\vec{v} = \lambda_i \vec{v}\}$$
es un subespacio vectorial de $\mathbb{R}^n$ tal que $\dim(V(\lambda_i)) \leq m_i$. Además si $m_{i_0} = 1$ entonces $\dim(V(\lambda_{i_0})) = 1$.

2. *Los autovectores correspondientes a autovalores distintos son linealmente independientes.*

3. *Si $r = n$, es decir, si todos los autovectores son distintos, entonces $\mathbb{R}^n = V(\lambda_1) \oplus \ldots \oplus V(\lambda_n)$.*

5.2.4. Propiedades de los autovectores de algunos tipos de matrices.

Matrices simétricas ($A^t = A$)

Si A es una matriz simétrica, entonces para cada autovalor λ_i de A se tiene que $\dim(V(\lambda_i)) = m_i$.

Matrices idempotentes ($A^2 = A$)

Si m_0 es la multiplicidad del autovalor $\lambda = 0$ y m_1 la multiplicidad del autovalor $\lambda = 1$ entonces se tiene que $\dim(V(0)) = m_0$ y $\dim(V(1)) = m_1$.

Ejemplo 181 *Sea la matriz $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$. Su polinomio característico es*

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & -3 & 3-\lambda \end{vmatrix} = (1-\lambda)^3$$

por lo que tiene un único autovalor $\boxed{\lambda = 1}$ con multiplicidad $m = 3$ (autovalor triple).

Sus autovectores los calculamos como:

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

$$\left\{ \begin{array}{r} -x + y = 0 \\ -y + z = 0 \\ x - 3y + 2z = 0 \end{array} \right. \Longrightarrow \boxed{x = y = z}$$

Por lo tanto, los autovectores asociados al autovalor $\lambda = 1$ son de la forma

$$\vec{v} = (x, x, x)$$

Luego el subespacio invariante de autovectores asociados a $\lambda = 1$ es

$$V(1) = \mathcal{L}\{(1,1,1)\}$$

y por lo tanto

$$\dim(V(1)) = 1 < 3$$

5.3. Diagonalización de matrices

Definición 182 *Dadas las matrices $A, B \in \mathcal{M}_{n\times n}$, se dice que son* ***semejantes*** *si y sólo si existe una matriz $P \in \mathcal{M}_{n\times n}$* ***no singular*** *tal que*

$$B = PAP^{-1} \qquad \text{o bien} \qquad A = P^{-1}BP$$

Definición 183 *Dada una matriz $A \in \mathcal{M}_{n\times n}$, se dice que A es* ***diagonalizable*** *si y sólo si existe una matriz $M \in \mathcal{M}_{n\times n}$* ***no singular*** *tal que*

$$D = M^{-1}AM$$

donde $D \in \mathcal{M}_{n\times n}$ es una matriz diagonal. Es decir, A es semejante a una matriz diagonal D.

Ejemplo 184 *La matriz $A = \begin{pmatrix} 6 & -6 & 2 \\ -1 & -1 & 1 \\ 7 & 3 & 1 \end{pmatrix}$ es diagonalizable ya que la matriz M no singular*

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ -2 & 3 & 1 \end{pmatrix}, \text{ con inversa } M^{-1} = \begin{pmatrix} \frac{1}{6} & \frac{1}{2} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{2} & \frac{1}{6} \\ \frac{5}{6} & -\frac{1}{2} & \frac{1}{6} \end{pmatrix}$$

verifica que

$$M^{-1}AM = \begin{pmatrix} -4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{pmatrix} = D \qquad \text{¡Comprobadlo!}$$

La aplicación

$$f(x, y, z) = (6x - 6y + 2z, -x - y + z, 7x + 3y + z)$$

tiene a A como matriz asociada respecto de las bases canónicas. Si consideramos la base de autovectores

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

como

$$\begin{aligned} f(1, 1, -2) &= (-4, -4, 8) = -4 \cdot (1, 1, -2) \\ f(0, 1, 3) &= (0, 2, 6) = 2 \cdot (0, 1, 3) \\ f(1, 0, 1) &= (8, 0, 8) = 8 \cdot (1, 0, 1) \end{aligned}$$

entonces la matriz asociada a f respecto de la base B es la matriz diagonal D, por lo que A y D representan la misma aplicación lineal f y ambas matrices tienen los mismos autovalores

$$\lambda_1 = -4 \qquad \lambda_2 = 2 \qquad \lambda_3 = 8$$

5.3.1. Caracterización de matrices diagonalizables

No todas las matrices cuadradas son diagonalizables, como hemos podido ver en algunos ejemplos anteriores, pero existen casos en los que podemos asegurar la diagonalización.

Proposición 185 *Sea una matriz $A \in \mathcal{M}_{n\times n}$. Se verifica que si A tiene n autovalores $\lambda_1, \dots \lambda_n$ distintos dos a dos, entonces A es diagonalizable. Además existe una matriz $P \in \mathcal{M}_{n\times n}$ cuyas columnas son los autovectores $\vec{v}_1, \dots \vec{v}_n$ asociados a los autovalores $\lambda_1, \dots \lambda_n$ de A para la que se cumple*

$$\boxed{A = PDP^{-1}} \text{ o equivalentemente } \boxed{D = P^{-1}AP}$$

siendo D la matriz diagonal $D = \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$.

También existen matrices diagonalizables que tienen autovalores iguales, ¡pero no todas lo son! El siguiente teorema nos da la condición.

Teorema 186 *Una matriz $A \in \mathcal{M}_{n\times n}$ es diagonalizable si y sólo si existe una base de autovectores del espacio en el que está definida la aplicación lineal f que tiene a A por matriz asociada respecto de las bases canónicas.*

Observación 187 *El teorema anterior se traduce en que si $A \in \mathcal{M}_{n\times n}$ tiene autovalores λ_i con multiplicidades algebraicas m_i para $i = 1, \ldots, p$, siendo*

$$n = \sum_{i=1}^{p} m_i$$

entonces A es diagonalizable si y sólo si

$$\dim \left(V\left(\lambda_i\right)\right) = m_i$$

para todo $i = 1, \ldots, p$. Si esto sucede entonces se tiene que

$$V = V\left(\lambda_1\right) \oplus \cdots \oplus V\left(\lambda_p\right)$$

Ejemplo 188 *(Matriz con autovalor múltiple diagonalizable)*

Consideremos la matriz $A = \begin{pmatrix} -11 & -4 & 8 \\ 0 & 1 & 0 \\ -12 & -4 & 9 \end{pmatrix}$. Vamos a calcular los autovalores:

$$P_3\left(\lambda\right) = \left|A - \lambda I\right| = \begin{vmatrix} -11-\lambda & -4 & 8 \\ 0 & 1-\lambda & 0 \\ -12 & -4 & 9-\lambda \end{vmatrix} =$$

$$= \left(1-\lambda\right)\left[\lambda^2 + 2\lambda - 3\right] = \left(1-\lambda\right)\left(1-\lambda\right)\left(-3-\lambda\right)$$

por lo que los autovalores son $\lambda_1 = 1$ con multiplicidad $m_1 = 2$, y $\lambda_2 = -3$ con multiplicidad $m_2 = 1$.

A continuación calcularemos los autovectores asociados a $\lambda_1 = 1$. Resolvemos el sistema $(A - I)\vec{x} = \vec{0}$, así

$$\left.\begin{array}{c} -12x - 4y + 8z = 0 \\ 0 = 0 \\ -12x - 4y + 8z = 0 \end{array}\right\} \Rightarrow \boxed{y = -3x + 2z}$$

y por lo tanto

$$V\left(1\right) = \left\{(x, y, z) \in \mathbb{R}^3 : y = -3x + 2z\right\} = \mathcal{L}\left\{\left(1, -3, 0\right), \left(0, 2, 1\right)\right\}$$

por lo que

$$\dim\left(V\left(1\right)\right) = 2$$

Por otro lado, para calcular los autovectores asociados a $\lambda_2 = -3$ resolvemos el sistema $(A + 3I)\vec{x} = \vec{0}$, así

$$\left.\begin{array}{c} -8x - 4y + 8z = 0 \\ 4y = 0 \\ -12x - 4y + 12z = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{y = 0} \\ \boxed{z = x} \end{array}$$

y por lo tanto

$$V\left(-3\right) = \left\{(x, y, z) \in \mathbb{R}^3 : y = 0, x - z = 0\right\} = \mathcal{L}\left\{\left(1, 0, 1\right)\right\}$$

por lo que

$$\dim\left(V\left(-3\right)\right) = 1$$

Como $\dim(V(\lambda_1)) = m_1$ *y* $\dim(V(\lambda_2)) = m_2$*, entonces* A *es diagonalizable ya que*

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

es una base de autovectores y por lo tanto

$$A = MDM^{-1}$$

siendo $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$ *y* $M = \begin{pmatrix} 1 & 0 & 1 \\ -3 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$.

Ejemplo 189 *(Matriz con autovalor múltiple no diagonalizable)*

Sea $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$*. Vamos a calcular sus autovalores:*

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & 1 \\ 1 & -3 & 3-\lambda \end{vmatrix} =$$

$$= -\lambda^3 + 3\lambda^2 - 3\lambda + 1 = (1-\lambda)^3$$

por lo que hay un único autovalor $\lambda = 1$ *con multiplicidad* $m = 3$.

A continuación calcularemos los autovectores asociados a $\lambda = 1$*. Resolvemos el sistema* $(A - I)\vec{x} = \vec{0}$*, así*

$$\left.\begin{array}{r} -x+y=0 \\ -y+z=0 \\ x-3y+2z=0 \end{array}\right\} \Rightarrow \begin{array}{l} \boxed{y=x} \\ \boxed{z=y} \end{array}$$

y por lo tanto

$$V(1) = \left\{(x,y,z) \in \mathbb{R}^3 : x-y=0, y-z=0\right\} = \mathcal{L}\{(1,1,1)\}$$

por lo que

$$\dim(V(1)) = 1$$

y como $\dim(V(\lambda)) \neq m$ *entonces* A *es no diagonalizable ya que*

$$B = \{(1,1,1)\}$$

no forma una base.

5.3.2. Diagonalización de algunos tipos de matrices especiales

Por sus características y propiedades, algunos tipos de matrices son siempre diagonalizables.

Matrices simétricas ($A = A^T$)

Proposición 190 *Dada una matriz* $A \in \mathcal{M}_{n\times n}$ ***simétrica****, se verifica que:*

1. *A es diagonalizable.*
2. *Existe al menos una base de autovectores ortonormal y por tanto existe una matriz* $Q \in \mathcal{M}_{n\times n}$ *ortogonal tal que* $A = QDQ^T$.

Ejemplo 191 *Sea* $A = \begin{pmatrix} 2 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & -1 \end{pmatrix}$*. Calculamos el polinomio característico:*

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 & 2 \\ 0 & -1-\lambda & 0 \\ 2 & 0 & -1-\lambda \end{vmatrix} =$$

$$= (-1-\lambda)\left[\lambda^2 - \lambda - 6\right] = (-1-\lambda)(-2-\lambda)(3-\lambda)$$

por lo que los autovalores son $\lambda_1 = -1$*,* $\lambda_2 = -2$ *y* $\lambda_3 = 3$ *todos con multiplicidad* $m = 1$*.*

A continuación calcularemos los autovectores asociados a $\lambda_1 = -1$*. Resolvemos el sistema* $(A + I)\vec{x} = \vec{0}$*, así*

$$\left.\begin{array}{c} 3x + 2z = 0 \\ 0 = 0 \\ 2x = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{y \text{ libre}} \\ \boxed{x = z = 0} \end{array}$$

y por lo tanto

$$V(-1) = \left\{(x, y, z) \in \mathbb{R}^3 : x = 0, z = 0\right\} = \mathcal{L}\{(0, 1, 0)\}$$

por lo que

$$\dim(V(1)) = 1 = m$$

Por otro lado, para calcular los autovectores asociados a $\lambda_2 = -2$ *resolvemos el sistema* $(A + 2I)\vec{x} = \vec{0}$*, así*

$$\left.\begin{array}{c} 4x + 2z = 0 \\ y = 0 \\ 2x + z = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{y = 0} \\ \boxed{z = -2x} \end{array}$$

y por lo tanto

$$V(-2) = \left\{(x, y, z) \in \mathbb{R}^3 : y = 0, 2x + z = 0\right\} = \mathcal{L}\{(1, 0, -2)\}$$

por lo que

$$\dim(V(-2)) = 1 = m$$

Por último, para los autovectores asociados a $\lambda_3 = 3$ *resolvemos* $(A - 3I)\vec{x} = \vec{0}$*, así*

$$\left.\begin{array}{c} -x + 2z = 0 \\ -4y = 0 \\ 2x - 4z = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{y = 0} \\ \boxed{x = 2z} \end{array}$$

y por lo tanto

$$V(3) = \left\{(x, y, z) \in \mathbb{R}^3 : y = 0, x - 2z = 0\right\} = \mathcal{L}\{(2, 0, 1)\}$$

por lo que

$$\dim(V(3)) = 1 = m$$

Como $\dim(V(\lambda_i)) = m$ *con* $i = 1, 2, 3$*, entonces* A *es diagonalizable ya que*

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

es una base de autovectores y por lo tanto

$$A = MDM^{-1}$$

siendo $D = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ *y* $M = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & -2 & 1 \end{pmatrix}$.

Observemos que los autovectores de B son ortogonales dos a dos calculando los productos escalares:

$$(0,1,0)\cdot(1,0,-2)=0$$
$$(0,1,0)\cdot(2,0,1)=0$$
$$(1,0,-2)\cdot(2,0,1)=0$$

pero no son ortonormales ya que no son todos unitarios:

$$\|(0,1,0)\|=1$$
$$\|(1,0,-2)\|=\sqrt{5}$$
$$\|(2,0,1)\|=\sqrt{5}$$

Pero si los normalizamos, entonces sí formarán una base ortonormal de autovectores:

$$B_N=\left\{(0,1,0),\left(\frac{1}{\sqrt{5}},0,\frac{-2}{\sqrt{5}}\right),\left(\frac{2}{\sqrt{5}},0,\frac{1}{\sqrt{5}}\right)\right\}$$

y ahora la matriz de autovectores $Q=\begin{pmatrix}0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}}\\ 1 & 0 & 0\\ 0 & \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}}\end{pmatrix}$ *es una matriz ortogonal y por lo tanto* $A=QDQ^T$.

Ejercicio 192 *Estudiar si la matriz* $A=\begin{pmatrix}4 & 0 & -1\\ 0 & 3 & 0\\ -1 & 0 & 4\end{pmatrix}$ *es diagonalizable. En caso afirmativo, encontrar la matriz D diagonal y M no singular tal que $A=MDM^{-1}$. Si es posible, encontrar una matriz Q ortogonal tal que $A=QDQ^T$.*

Matrices idempotentes ($A^2=A$)

Proposición 193 *Dada una matriz $A\in\mathcal{M}_{n\times n}$ tal que $A\neq 0$, entonces si A es* ***idempotente*** *se tiene que entonces A es diagonalizable.*

Ejemplo 194 *Sea* $A=\begin{pmatrix}-1 & 1 & 2\\ 2 & 0 & -2\\ -2 & 1 & 3\end{pmatrix}$ *que es idempotente. Vamos a calcular los autovalores:*

$$P_3(\lambda)=|A-\lambda I|=\begin{vmatrix}-1-\lambda & 1 & 2\\ 2 & -\lambda & -2\\ -2 & 1 & 3-\lambda\end{vmatrix}=$$

$$=-\lambda^3+2\lambda^2-\lambda=-\lambda(1-\lambda)^2$$

por lo que los autovalores son $\lambda_1=0$ con multiplicidad $m_1=1$, y $\lambda_2=1$ con multiplicidad $m_2=2$.

Por la proposición anterior podremos encontrar una base de autovectores. Para calcular los autovectores asociados a $\lambda_1=0$ resolvemos el sistema $A\vec{x}=\vec{0}$, así

$$\left.\begin{array}{c}-x+y+2z=0\\ 2x-2z=0\\ -2x+y+3z=0\end{array}\right\}\Rightarrow\begin{array}{c}\boxed{z=x}\\ \boxed{y=-x}\end{array}$$

y por lo tanto

$$V(0)=\left\{(x,y,z)\in\mathbb{R}^3 : x+y=0, x-z=0\right\}=\mathcal{L}\left\{(1,-1,1)\right\}$$

por lo que

$$\dim(V(0))=1=m_1$$

Calculamos ahora los autovectores asociados a $\lambda_2 = -3$ *resolviendo* $(A - I)\,\vec{x} = \vec{0}$*, así*

$$\left.\begin{array}{c} -2x + y + 2z = 0 \\ 2x - y - 2z = 0 \\ -2x + y + 2z = 0 \end{array}\right\} \Rightarrow \boxed{y = 2x - 2z}$$

y por lo tanto

$$V\,(1) = \left\{(x, y, z) \in \mathbb{R}^3 : y = 2x - 2z\right\} = \mathcal{L}\left\{(1, 2, 0)\,, (0, -2, 1)\right\}$$

por lo que

$$\dim\,(V\,(1)) = 2 = m_2$$

Como $\dim\,(V\,(\lambda_1)) = m_1$ *y* $\dim\,(V\,(\lambda_2)) = m_2$*, entonces* A *es diagonalizable como ya esperábamos y*

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

es una base de autovectores y por lo tanto

$$A = MDM^{-1}$$

siendo $D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ *y* $M = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 2 & -2 \\ 1 & 0 & 1 \end{pmatrix}$.

5.4. Aplicaciones

5.4.1. Cálculo de la k-ésima potencia de una matriz

Proposición 195 *Sea* $A \in \mathcal{M}_{n\times n}$ *una matriz con autovalores* $\lambda_1, \ldots, \lambda_n$ *iguales o distintos. Si* A *es diagonalizable con* $A = MDM^{-1}$ *donde* M *es la matriz de autovectores y* D *la matriz diagonal de autovalores, entonces para todo* $k \in \mathbb{N}$ *se tiene que*

$$A^k = MD^kM^{-1} = M \begin{pmatrix} \lambda_1^k & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^k \end{pmatrix} M^{-1}$$

Ejemplo 196 *Vamos a calcular la potencia* k *de la matriz* $A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ *que tiene como autovalores* $\lambda_1 = 0$*,* $\lambda_2 = 1$ *y* $\lambda_3 = 3$ *y los subespacios de autovalores asociados son*

$$V\,(0) = \mathcal{L}\left\{\left(\tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}, \tfrac{1}{\sqrt{3}}\right)\right\}$$

$$V\,(1) = \mathcal{L}\left\{\left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{-1}{\sqrt{2}}\right)\right\}$$

$$V\,(3) = \mathcal{L}\left\{\left(\tfrac{1}{\sqrt{6}}, \tfrac{-2}{\sqrt{6}}, \tfrac{1}{\sqrt{6}}\right)\right\}$$

tales que

$$Q = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \qquad y \qquad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

son las matrices Q *ortogonal y* D *diagonal tales que* $A = QDQ^T$.

Así se tiene que

$$A^k = QD^kQ^T = Q \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^k \end{pmatrix} Q^T =$$

$$= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^k \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{pmatrix}^{-1} =$$

$$= \begin{pmatrix} \frac{1}{2} + \frac{1}{6}3^k & -\frac{1}{3}3^k & -\frac{1}{2} + \frac{1}{6}3^k \\ -\frac{1}{3}3^k & \frac{2}{3}3^k & -\frac{1}{3}3^k \\ -\frac{1}{2} + \frac{1}{6}3^k & -\frac{1}{3}3^k & \frac{1}{2} + \frac{1}{6}3^k \end{pmatrix}$$

5.4.2. Planteamiento de un problema económico

El siguiente problema aparece en .Algebra lineal y teoría de matrices”de R. Barbolla y P. Sanz (Ed. Prentice Hall):

La población activa de un país se clasifica en tres categorías profesionales: técnicos superiores (x), obreros espacializados (y), y obreros no especializados (z). Así en cada generación k, la fuerza de trabajo del país está caracterizada por el número de personas incluidas en las tres categorías, es decir $(x(k), y(k), z(k))$.

Supóngase que:

- Cada trabajador activo sólo tiene un hijo.
- El 50 % de los hijos de los técnicos superiores lo son también, el 25 % pasa a ser obrero especializado y el 25 % restante es obrero no especializado.
- Los hijos de los obreros espacializados se reparten en las mismas tres categorias según los porcentajes 30 %, 40 % y 30 %.
- Para los hijos de los obreros no especializados las proporciones de reparto entre las categorías son 50 %, 25 % y 25 %.

Se pide:

1. Plantear en forma matricial un modelo que represente la distribución de la fuerza de trabajo del país de generación en generación.
2. ¿Cuál será la distribución de los trabajadores a largo plazo, independientemente de la distribución inicial?

Para responder al primer apartado, basta con plantear un sistema que nos diga como cambian de categoría los trabajadores de la generación k a la generación $k+1$.

Así, este sistema viene dado por:

$$\begin{cases} x(k+1) = \frac{1}{2}x(k) + \frac{3}{10}y(k) + \frac{1}{2}z(k) \\ y(k+1) = \frac{1}{4}x(k) + \frac{2}{5}y(k) + \frac{1}{4}z(k) \\ x(k+1) = \frac{1}{4}x(k) + \frac{3}{10}y(k) + \frac{1}{4}z(k) \end{cases}$$

por lo que la matriz A del sistema será

$$A = \begin{pmatrix} \frac{1}{2} & \frac{3}{10} & \frac{1}{2} \\ \frac{1}{4} & \frac{2}{5} & \frac{1}{4} \\ \frac{1}{4} & \frac{3}{10} & \frac{1}{4} \end{pmatrix}$$

Si tenemos un dato inicial $\begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix}$ entonces el sistema se puede escribir para la generación k-ésima como:

$$\begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} = A \begin{pmatrix} x(k-1) \\ y(k-1) \\ z(k-1) \end{pmatrix} = A^2 \begin{pmatrix} x(k-2) \\ y(k-2) \\ z(k-2) \end{pmatrix} = \cdots = A^k \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix}$$

por lo tanto, como la matriz A se puede escribir como $A = MDM^{-1}$ con

$$M = \begin{pmatrix} -1 & 15 & -4 \\ 0 & 10 & 3 \\ 1 & 9 & 1 \end{pmatrix} \qquad \text{y} \qquad D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{3}{20} \end{pmatrix} \qquad \text{¡Comprobarlo!}$$

entonces

$$A^k = MD^kM^{-1}$$

por lo que

$$\begin{pmatrix} x(k) \\ y(k) \\ z(k) \end{pmatrix} = MD^kM^{-1} \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix} = M \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \left(\frac{3}{20}\right)^k \end{pmatrix} M^{-1} \begin{pmatrix} x(0) \\ y(0) \\ z(0) \end{pmatrix}$$

Se dejan *como ejercicio* los cálculos de este problema.

Ejemplo 197 *(Diagonalización con parámetros)*

Determinar si la matriz $A = \begin{pmatrix} a & b & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \end{pmatrix}$ *es diagonalizable en función de los parámetros* $a, b \in \mathbb{R}$.

Como la matriz A es triangular, los autovalores serán los elementos de la diagonal $\lambda_1 = a$, $\lambda_2 = 1$ y $\lambda_3 = 2$. De esta manera si el parámetro $a \neq 1$ y $a \neq 2$, como los tres autovalores serían distintos, la matriz A sería diagonalizable.

Si $a = 1$ entonces $\lambda = 1$ será un autovalor doble y tendremos que estudiar la dimensión del subespacio $V(1)$. Por tanto, resolvemos el sistema $(A - I)\vec{x} = \vec{0}$ para encontrar los autovectores de $V(1)$:

$$\left.\begin{array}{r} by = 0 \\ 2z = 0 \\ z = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{z = 0} \\ \boxed{by = 0} \end{array}$$

por lo que si $b = 0$ entonces la solución sería $\boxed{z = 0}$ y por tanto $V(1) = \mathcal{L}\{(1,0,0),(0,1,0)\}$ y así $\dim(V(1)) = 2$, por lo que la matriz A sería diagonalizable.

Por otro lado si, $b \neq 0$, entonces la solución será $\boxed{z = 0}$ y $\boxed{y = 0}$ y así tendríamos que $V(1) = \mathcal{L}\{(1,0,0)\}$ y $\dim(V(1)) = 1$, por lo que la matriz A no sería diagonalizable.

En el caso en que $a = 2$ entonces $\lambda = 2$ sería un autovalor doble y tendríamos que estudiar la dimensión del subespacio $V(2)$. Así, resolvemos el sistema $(A - 2I)\vec{x} = \vec{0}$ para encontrar los autovectores de $V(2)$:

$$\left.\begin{array}{r} by = 0 \\ -y + 2z = 0 \\ 0 = 0 \end{array}\right\} \Rightarrow \begin{array}{c} \boxed{by = 0} \\ \boxed{y = 2z} \end{array}$$

por lo que si $b = 0$ entonces la solución será $\boxed{y = 2z}$ y por lo tanto se tiene que $V(2) = \mathcal{L}\{(1,0,0),(0,2,1)\}$ por lo que $\dim(V(2)) = 2$ y la matriz A sería diagonalizable.

Si por el contrario, $b \neq 0$, entonces la solución sería $\boxed{y = 0}$ y $\boxed{z = 0}$, y tendríamos que $V(2) = \mathcal{L}\{(1,0,0)\}$ por lo que $\dim(V(2)) = 1$ y la matriz A no sería diagonalizable.

Resumiendo:

- A *es diagonalizable si y sólo si* $\begin{cases} a \neq 1 \text{ y } a \neq 2 \text{ para cualquier } b \in \mathbb{R} \\ a = 1 \text{ y } b = 0 \\ a = 2 \text{ y } b = 0 \end{cases}$.

- A *no es diagonalizable si y sólo si* $\begin{cases} a = 1 \text{ y } b \neq 0 \\ a = 2 \text{ y } b \neq 0 \end{cases}$.

Ejercicio 198 *Estudiar si la matriz* $A = \begin{pmatrix} a & 1 & 0 \\ 0 & 3 & 0 \\ b & 0 & 1 \end{pmatrix}$ *es diagonalizable en función de los parámetros* $a, b \in \mathbb{R}$.

Capítulo 6

Formas cuadráticas

6.1. Formas cuadráticas. Tipos.

Las formas cuadráticas se emplean en campos tan diversos como el Algebra, el Análisis, la Estadística, la Economía,... y en esta última se utilizan en Econometría, Optimización, Teoría de Consumidores, ...

Definición 199 *Se dice que un polinomio p en las variables $x_1, \ldots, x_n$ es un* ***polinomio cuadrático****, si cada uno de sus términos tiene grado dos, es decir,*

$$p(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

siendo $a_{ij} \in \mathbb{R}$, $i, j = 1, \ldots, n$ los coeficientes y x_i, $i = 1, \ldots, n$ las variables con valores en $\mathbb{R}$.

Por lo tanto, puede verse que dado un vector $\vec{x} = (x_1, \ldots, x_n)$, un polinomio cuadrático se puede interpretar como una aplicación (que no es lineal) que al vector $\vec{x} \in \mathbb{R}^n$ le asocia el número real $p(x_1, \ldots, x_n) = p(\vec{x})$.

Definición 200 *Se denomina* ***forma cuadrática*** *q a toda aplicación de $\mathbb{R}^n$ en $\mathbb{R}$ que a cada vector $\vec{x} \in \mathbb{R}^n$ le hace corresponder el valor numérico dado por un polinomio cuadrático.*

Definición 201 *Cuando una forma cuadrática q, está expresada en los siguientes términos*

$$q(y_1, \ldots, y_n) = \sum_{i=1}^{n} d_i y_i^2$$

con $d_i \in \mathbb{R}$, $i = 1, \ldots, n$ se dice que está en ***forma canónica****.*

Ejemplo 202 *El polinomio*

$$p_1(x_1, x_2, x_3) = 2x_1^2 + 3x_1 x_2 + 5x_3^2 + x_1$$

no es un polinomio cuadrático.

Ejemplo 203 *El polinomio*

$$p_2(x_1, x_2, x_3) = x_1^2 + 2x_1 x_2 + 6x_2^2 + 4x_2 x_3 + x_3^2$$

es un polinomio cuadrático y la forma cuadrática p_2 puede expresarse en forma canónica mediante

$$p_2(y_1, y_2, y_3) = y_1^2 + 5y_2^2 + \frac{1}{5} y_3^2$$

siendo $\left\{ \begin{array}{l} y_1 = x_1 + x_2 \\ y_2 = x_2 + \frac{2}{5} x_3 \\ y_3 = x_3 \end{array} \right.$.

6.2. Matriz asociada a una forma cuadrática

Es posible utilizar la notación matricial para describir una forma cuadrática. Así

$$q(x_1, \ldots, x_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

se puede escribir como

$$q(\vec{x}) = \vec{x}^T A \vec{x}$$

donde $\vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ y por tanto $\vec{x}^T = (x_1, \ldots, x_n)$, y además $A = (a_{ij})$ con $i, j = 1, \ldots, n$.

Por lo tanto

$$q(\vec{x}) = q(x_1, \ldots, x_n) = (x_1, \ldots, x_n) \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Ejemplo 204 *La forma cuadrática $p_2(x_1, x_2, x_3) = x_1^2 + 2x_1x_2 + 6x_2^2 + 4x_2x_3 + x_3^2$ puede escribirse matricialmente como:*

$$p_2(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 2 & 0 \\ 0 & 6 & 4 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

De igual forma, si agrupamos los términos como

$$p_2(x_1, x_2, x_3) = x_1^2 + x_1x_2 + x_2x_1 + 6x_2^2 + 3x_2x_3 + x_3x_2 + x_3^2$$

entonces en notación matricial se puede escribir como

$$p_2(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 1 & 0 \\ 1 & 6 & 3 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Por tanto la notación matricial de una forma cuadrática NO siempre es única.

Proposición 205 *Dada una forma cuadrática $q : \mathbb{R}^n \to \mathbb{R}$, existe una UNICA matriz simétrica $Q \in \mathcal{M}_{n\times n}$ tal que*

$$q(\vec{x}) = \vec{x}^T Q \vec{x}$$

Definición 206 *Dada una forma cuadrática $q : \mathbb{R}^n \to \mathbb{R}$, la única matriz simétrica $Q \in \mathcal{M}_{n\times n}$ para la que para todo $\vec{x} \in \mathbb{R}^n$ se verifica que*

$$q(\vec{x}) = \vec{x}^T Q \vec{x}$$

es a la que llamaremos ***matriz asociada*** *a la forma cuadrática q, denominándose* ***expresión matricial*** *de q, precisamente a la dada a partir de la matriz Q.*

Ejemplo 207 *La expresión matricial de la forma cuadrática*

$$q_1(x_1, x_2, x_3, x_4) = 2x_1^2 + 3x_1x_4 - 5x_2x_3 + 7x_3^2 - 5x_4^2 + x_2x_4$$

es

$$q_1(x_1,x_2,x_3,x_4) = (x_1,x_2,x_3,x_4)\begin{pmatrix} 2 & 0 & 0 & \frac{3}{2} \\ 0 & 0 & -\frac{5}{2} & \frac{1}{2} \\ 0 & -\frac{5}{2} & 7 & 0 \\ \frac{3}{2} & \frac{1}{2} & 0 & -5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Por otro lado, si q_2 es una forma cuadrática con matriz asociada Q tal que

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

entonces

$$q_2(x_1,x_2,x_3) = x_1^2 + 2x_1x_2 + 4x_1x_3 - x_2^2 + 4x_3^2$$

6.3. Tipos de formas cuadráticas

Definición 208 *Dada una forma cuadrática $q : \mathbb{R}^n \to \mathbb{R}$, se dice que:*

1. *q es* ***definida positiva*** *(dp) si y sólo si para todo $\vec{x} \in \mathbb{R}^n$ con $\vec{x} \neq \vec{0}$ se verifica que $q(\vec{x}) > 0$.*
2. *q es* ***definida negativa*** *(dn) si y sólo si para todo $\vec{x} \in \mathbb{R}^n$ con $\vec{x} \neq \vec{0}$ se verifica que $q(\vec{x}) < 0$.*
3. *q es* ***semidefinida positiva*** *(sdp) si y sólo si para todo $\vec{x} \in \mathbb{R}^n$ se verifica que $q(\vec{x}) \geq 0$ y además existe con $\vec{x}_0 \in \mathbb{R}^n$ con $\vec{x}_0 \neq \vec{0}$ tal que $q(\vec{x}_0) = 0$.*
4. *q es* ***semidefinida negativa*** *(sdn) si y sólo si para todo $\vec{x} \in \mathbb{R}^n$ se verifica que $q(\vec{x}) \leq 0$ y además existe con $\vec{x}_0 \in \mathbb{R}^n$ con $\vec{x}_0 \neq \vec{0}$ tal que $q(\vec{x}_0) = 0$.*
5. *q es* ***indefinida*** *si y sólo si existen $\vec{x}_1, \vec{x}_2 \in \mathbb{R}^n$ tales que $q(\vec{x}_1) > 0$ y $q(\vec{x}_2) < 0$.*

Observación 209 *Esta clasificación es exhaustiva y excluyente. Es decir cada forma cuadrática debe pertenecer a uno de estos tipos y sólo a uno.*

Ejemplo 210 *Las formas cuadráticas*

$$\begin{gathered} q_1(x_1,x_2,x_3) = x_1^2 + 2x_1x_2 + 2x_2^2 + 3x_3^2 \\ q_2(x_1,x_2,x_3) = -4x_1^2 - 5x_2^2 - 2x_3^2 + 4x_1x_3 \\ q_3(x_1,x_2) = x_1^2 - 4x_1x_2 + 4x_2^2 \\ q_4(x_1,x_2,x_3,x_4) = -4x_1^2 + 4x_1x_2 - x_2^2 - 9x_3^2 + 6x_3x_4 - x_4^2 \\ q_5(x_1,x_2) = x_1x_2 \end{gathered}$$

se pueden escribir como

$$\begin{gathered} q_1(x_1,x_2,x_3) = (x_1+x_2)^2 + x_2^2 + 3x_3^2 \qquad \boxed{dp} \\ q_2(x_1,x_2,x_3) = -\left[(2x_1-x_3)^2 + 5x_2^2 + x_3^2\right] \qquad \boxed{dn} \\ q_3(x_1,x_2) = (x_1-2x_2)^2 \qquad \boxed{sdp} \\ q_4(x_1,x_2,x_3,x_4) = -\left[(2x_1-x_2)^2 + (3x_3-x_4)^2\right] \qquad \boxed{sdn} \\ q_5(x_1,x_2) = x_1x_2 \qquad \boxed{Indefinida} \end{gathered}$$

Observación 211 *Si tenemos la expresión canónica de la forma cuadrática, la clasificación es inmediata:*

$$q(y_1,\ldots,y_n) = \sum_{i=1}^{n} d_i y_i^2$$

entonces

1. *q es definida positiva $\iff d_i > 0$ para todo $i = 1, \ldots, n$.*

2. *q es definida negativa $\iff d_i < 0$ para todo $i = 1, \ldots, n$.*

3. *q es semidefinida positiva $\iff d_i \geq 0$ para todo $i = 1, \ldots, n$ y existe i_0 tal que $d_{i_0} = 0$.*

4. *q es semidefinida negativa $\iff d_i \leq 0$ para todo $i = 1, \ldots, n$ y existe i_0 tal que $d_{i_0} = 0$.*

5. *q es indefinida $\iff$ existen i_0 e i_1 tales que $d_{i_0} > 0$ y $d_{i_1} < 0$.*

Observación 212 *Sean $q(\vec{x}) = \vec{x}^T A \vec{x}$ y $p(\vec{x}) = \vec{x}^T B \vec{x}$ dos formas cuadráticas en las variables $\vec{x} = (x_1, \ldots, x_n)$, entonces se tiene que:*

1. *Para cualquier $\alpha > 0$, las formas cuadráticas $q(\vec{x})$ y $\alpha q(\vec{x}) = \vec{x}^T (\alpha A) \vec{x}$ tienen el mismo carácter.*

2. *Si q es definida positiva, semidefinida positiva o indefinida, entonces $-q$ es definida negativa, semidefinida negativa o indefinida respectivamente.*

3. *Si p y q son dos formas cuadráticas definidas positivas o definidas negativas, entonces $p + q$ tiene el mismo carácter que p y q.*

4. *Si q es definida positiva (negativa) y p es semidefinida positiva (negativa) entonces $p + q$ es definida positiva (negativa).*

Ejercicio 213 *Investigar el carácter de la forma $p + q$ sabiendo que p y q son ambas semidefinidas positivas o semidefinidas negativas.*

6.4. Criterios de clasificación de formas cuadráticas

Para los siguientes criterios sólo vamos a considerar que la matriz asociada a la forma cuadrática es una matriz **simétrica**.

6.4.1. Criterio de los menores principales

En el ejemplo anterior, hemos transformado las formas cuadráticas completando cuadrados. Veámoslo en el siguiente ejemplo. Sea la forma cuadrática q tal que $a_{11} \neq 0$ entonces

$$
\begin{aligned}
q(x_1, x_2) &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = \\
&= a_{11}\left[x_1^2 + \frac{2a_{12}}{a_{11}}x_1x_2 + \frac{a_{12}^2}{a_{11}^2}x_2^2\right] + \left(a_{22} - \frac{a_{12}^2}{a_{11}}\right)x_2^2 = \\
&= a_{11}\left(x_1 + \frac{a_{12}}{a_{11}}x_2\right)^2 + \frac{|A|}{a_{11}}x_2^2
\end{aligned}
$$

llamando $D_1 = a_{11}$ y $D_2 = |A|$ entonces si

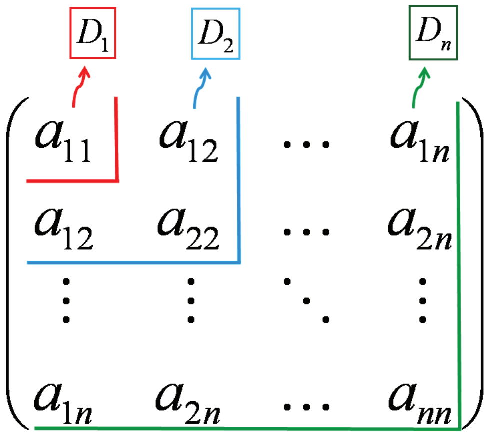
$$a_{11} > 0 \text{ y } \frac{|A|}{a_{11}} > 0 \text{ entonces } |A| > 0 \text{ y además } q \text{ es dp.}$$

y si

$$a_{11} < 0 \text{ y } \frac{|A|}{a_{11}} < 0 \text{ entonces } |A| > 0 \text{ y además } q \text{ es dn.}$$

Este hecho motivará el siguiente teorema.

Llamemos $D_1, D_2, \ldots, D_n$ a los **menores principales** de la matriz asociada A, esto es



Teorema 214 ***(Criterio de menores principales)*** *Sea $q(\vec{x}) = \vec{x}^T A \vec{x}$ una forma cuadrática en las variables $\vec{x} = (x_1, \ldots, x_n)$. Entonces se verifica que:*

1. *q es definida positiva $\Longleftrightarrow$ $D_i > 0$ para todo $i = 1, \ldots, n$.*
2. *q es definida negativa $\Longleftrightarrow$ $(-1)^i D_i > 0$ para todo $i = 1, \ldots, n$.*
3. *Si $D_i > 0$ para todo $i = 1, \ldots, n-1$ y $D_n = |A| = 0$ entonces q es semidefinida positiva.*
4. *Si $(-1)^i D_i > 0$ para todo $i = 1, \ldots, n-1$ y $D_n = |A| = 0$ entonces q es semidefinida negativa.*

Observación 215 ***(Importante)*** *Si $|A| \neq 0$ entonces la forma cuadrática q sólo puede ser definida positiva, definida negativa o indefinida. Si además no cumple los puntos 1 y 2 del teorema anterior, entonces la forma cuadrática q es indefinida.*

Ejemplo 216 *Vamos a clasificar mediante este criterio de menores principales las siguientes formas cuadráticas:*

1. $$q_1(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$$

   Como tenemos que

   $$D_1 = 5 > 0$$
   $$D_2 = \begin{vmatrix} 5 & 4 \\ 4 & 5 \end{vmatrix} = 9 > 0$$
   $$D_3 = |A| = 10 > 0$$

   entonces por el criterio de menores principales se tiene que q_1 es definida positiva.

2. $$q_2(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} -4 & 0 & 2 \\ 0 & -5 & 0 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$$

   Como tenemos que

   $$D_1 = -4 < 0$$
   $$D_2 = \begin{vmatrix} -4 & 0 \\ 0 & -5 \end{vmatrix} = 3 > 0$$
   $$D_3 = |A| = -20 < 0$$

   entonces por el criterio de menores principales se tiene que q_2 es definida negativa.

3. $q_3(x_1,x_2,x_3) = (x_1,x_2,x_3)\begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$

Como tenemos que

$$D_1 = 1 > 0$$
$$D_2 = \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} = 1 > 0$$
$$D_3 = |A| = 0$$

entonces por el criterio de menores principales se tiene que q_3 es semidefinida positiva.

4. $q_4(x_1,x_2,x_3) = (x_1,x_2,x_3)\begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & -3 \\ 0 & -3 & -3 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$

Como tenemos que

$$D_1 = -1 < 0$$
$$D_2 = \begin{vmatrix} -1 & 1 \\ 1 & -4 \end{vmatrix} = 3 > 0$$
$$D_3 = |A| = 0$$

entonces por el criterio de menores principales se tiene que q_4 es semidefinida negativa.

Ejemplo 217 *A veces podremos clasificar las formas cuadráticas indefinidas utilizando el criterio de menores principales y la observación importante anterior:*

$$q_5(x_1,x_2,x_3) = (x_1,x_2,x_3)\begin{pmatrix} -1 & -1 & 2 \\ -1 & 3 & 0 \\ 2 & 0 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$$

y se tiene que

$$D_1 = -1 < 0$$
$$D_2 = \begin{vmatrix} -1 & -1 \\ -1 & 3 \end{vmatrix} = -2 < 0$$
$$D_3 = |A| = -20 < 0$$

por lo que q_5 no es ni definida positiva ni negativa, ya que la condición del criterio de menores principales para formas cuadráticas definidas es necesaria y suficiente.

Por otro lado, como $D_3 = |A| \neq 0$ entonces tampoco será semidefinida (ni negativa ni positiva) por lo tanto la forma cuadrática q_5 será indefinida.

De hecho $q_5(1,0,0) = -1 < 0$ y $q_5(0,1,0) = 2 > 0$.

Observación 218 *El criterio de menores principales no siempre es capaz de clasificar la forma cuadrática. Por lo que necesitaremos un criterio más fuerte para poder clasificar todas las formas cuadráticas.*

Ejemplo 219 *Tratemos de clasificar*

$$q_6(x_1,x_2,x_3) = (x_1,x_2,x_3)\begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$$

Los menores principales son

$$D_1 = 2 > 0$$
$$D_2 = \begin{vmatrix} 2 & -2 \\ -2 & 2 \end{vmatrix} = 0$$
$$D_3 = |A| = 0$$

entonces como $D_3 = |A| = 0$, la forma cuadrática q_6 sólo puede ser semidefinida o indefinida, pero no podemos concluir ya que las condiciones del criterio de menores principales para las formas cuadráticas semidefinidas no son necesarias y suficientes sino únicamente suficientes. Es decir, para que una forma cuadrática sea semidefinida no es necesario verificar tales condiciones. Por lo tanto no la podemos clasificar mediante este criterio.

6.4.2. Criterio de los autovalores

Teorema 220 ***(Criterio de autovalores)*** *Sea $q(\vec{x}) = \vec{x}^T A\vec{x}$ una forma cuadrática con matriz asociada A, donde $\lambda_1, \ldots, \lambda_n$ son sus autovalores. Entonces se verifica que:*

1. *q es $\boxed{dp} \Longleftrightarrow \lambda_i > 0$ para todo $i = 1, \ldots, n$.*
2. *q es $\boxed{dn} \Longleftrightarrow \lambda_i < 0$ para todo $i = 1, \ldots, n$.*
3. *q es $\boxed{sdp} \Longleftrightarrow \lambda_i \geq 0$ para todo $i = 1, \ldots, n$ y existe i_0 tal que $\lambda_{i_0} = 0$.*
4. *q es $\boxed{sdn} \Longleftrightarrow \lambda_i \leq 0$ para todo $i = 1, \ldots, n$ y existe i_0 tal que $\lambda_{i_0} = 0$.*
5. *q es $\boxed{indefinida} \Longleftrightarrow$ existen i_0 e i_1 tales que $\lambda_{i_0} < 0$ y $\lambda_{i_1} > 0$.*

Observación 221 *A diferencia del teorema de clasificación por menores principales, este nuevo criterio de autovalores nos da condiciones necesarias y suficientes para cada categoría, pudiendo así clasificar todas las formas cuadráticas. Esto no evita que podamos seguir utilizando el criterio de menores principales como primera opción, ya que es más sencillo y cómodo de utilizar. Cuando el criterio de menores principales no nos proporcione información, será necesario utilizar el de autovalores. Su dificultad vendrá dada por la dificultad de calcular los autovalores de la matriz asociada a la forma cuadrática.*

Ejemplo 222 *Vamos a clasificar mediante este criterio de autovalores las siguientes formas cuadráticas:*

1. $q_1(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A\vec{x}$

   Así
   $$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 5-\lambda & 4 & 2 \\ 4 & 5-\lambda & 2 \\ 2 & 2 & 2-\lambda \end{vmatrix} =$$
   $$= -\lambda^3 + 12\lambda^2 - 21\lambda + 10 = (1-\lambda)(1-\lambda)(10-\lambda)$$

   por lo que
   $$\lambda_1 = 1 > 0 \qquad \lambda_2 = 1 > 0 \qquad \lambda_3 = 10 > 0$$

   y entonces por el criterio de autovalores se tiene que q_1 es definida positiva.

2. $q_2(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} -4 & 0 & 2 \\ 0 & -5 & 0 \\ 2 & 0 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A\vec{x}$

   Así
   $$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} -4-\lambda & 0 & 2 \\ 0 & -5-\lambda & 0 \\ 2 & 0 & -2-\lambda \end{vmatrix} =$$
   $$= -\lambda^3 - 11\lambda^2 - 34\lambda - 20 = (5+\lambda)\left[-\lambda^2 - 6\lambda - 4\right]$$

por lo que

$$\lambda_1 = -5 < 0 \qquad \lambda_2 = -3 - \sqrt{5} < 0 \qquad \lambda_3 = -3 + \sqrt{5} < 0$$

y entonces por el criterio de autovalores se tiene que q_2 *es definida negativa.*

3. $q_3(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} 1 & 1 & 2 \\ 1 & 2 & 3 \\ 2 & 3 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$

Así

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & 2 \\ 1 & 2-\lambda & 3 \\ 2 & 3 & 5-\lambda \end{vmatrix} =$$
$$= -\lambda^3 + 8\lambda^2 - 3\lambda = -\lambda\left[\lambda^2 - 8\lambda + 3\right]$$

por lo que

$$\lambda_1 = 0 \qquad \lambda_2 = 4 + \sqrt{13} > 0 \qquad \lambda_3 = 4 - \sqrt{13} > 0$$

y entonces obtenemos que q_3 *es semidefinida positiva.*

4. $q_4(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & -3 \\ 0 & -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$

Así

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} -1-\lambda & 1 & 0 \\ 1 & -4-\lambda & -3 \\ 0 & -3 & -3-\lambda \end{vmatrix} =$$
$$= -\lambda^3 - 8\lambda^2 - 9\lambda = -\lambda\left[\lambda^2 + 8\lambda + 9\right]$$

por lo que

$$\lambda_1 = 0 \qquad \lambda_2 = -4 + \sqrt{7} < 0 \qquad \lambda_3 = -4 - \sqrt{7} < 0$$

y entonces tenemos que q_4 *es semidefinida negativa.*

5. $q_5(x_1, x_2, x_3) = (x_1, x_2, x_3) \begin{pmatrix} -2 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A \vec{x}$

Así

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} -2-\lambda & 0 & 1 \\ 0 & -\lambda & 0 \\ 1 & 0 & 3-\lambda \end{vmatrix} =$$
$$= -\lambda^3 + \lambda^2 + 7\lambda = -\lambda\left[\lambda^2 - \lambda - 7\right]$$

por lo que

$$\lambda_1 = 0 \qquad \lambda_2 = \frac{1 + \sqrt{29}}{2} > 0 \qquad \lambda_3 = \frac{1 - \sqrt{29}}{2} < 0$$

y entonces por el criterio de autovalores se tiene que q_5 *es indefinida.*

6. *Con el criterio de menores principales no hemos podido clasificar*

$$q_6(x_1,x_2,x_3) = (x_1,x_2,x_3)\begin{pmatrix} 2 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \vec{x}^T A\vec{x}$$

Ahora mediante el criterio de autovalores tenemos que

$$P_3(\lambda) = |A-\lambda I| = \begin{vmatrix} 2-\lambda & -2 & 0 \\ -2 & 2-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} =$$

$$= -\lambda^3 + 5\lambda^2 - 4\lambda = -\lambda\left[\lambda^2 - 5\lambda + 4\right]$$

por lo que

$$\lambda_1 = 0 \qquad \lambda_2 = 1 > 0 \qquad \lambda_3 = 4 > 0$$

y entonces podemos concluir que q_6 es semidefinida positiva.

Observación 223 *Se pueden combinar los dos criterios para clasificar formas cuadráticas de 2 y 3 variables, teniendo en cuenta las propiedades de los autovalores en relación con la <u>traza</u> y el <u>determinante</u>. Veámoslo en los siguientes casos.*

Consideremos la forma cuadrática $q(\vec{x}) = \vec{x}^T A\vec{x}$:

1. *Caso* $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$

   a) *Si $|A| = ac - b^2 = \lambda_1\lambda_2 < 0$ entonces q es $\boxed{\textit{indefinida}}$.*

   b) *Si $|A| = 0$ y* $\begin{cases} tr(A) = a+c = \lambda_1+\lambda_2 > 0 \implies q \text{ es } \boxed{sdp} \\ tr(A) = a+c = \lambda_1+\lambda_2 < 0 \implies q \text{ es } \boxed{sdn} \end{cases}$

   c) *Si $|A| > 0$ y* $\begin{cases} tr(A) > 0 \implies q \text{ es } \boxed{dp} \\ tr(A) < 0 \implies q \text{ es } \boxed{dn} \end{cases}$ *ya que $|A| = \lambda_1\lambda_2$ y $tr(A) = \lambda_1+\lambda_2$ siendo λ_1, λ_2 los autovalores de A.*

2. *Caso* $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$

   a) *Si $|A| > 0$ entonces q no es ni $\boxed{dn}$, ni $\boxed{\textit{semidefinida}}$. Por lo tanto sólo puede ser $\boxed{\textit{indefinida}}$ o $\boxed{dp}$, siendo esto último si $a_{11} > 0$ y $D_2 > 0$.*

   b) *Si $|A| = 0$ entonces q no es ni $\boxed{dp}$, ni $\boxed{dn}$. Por lo tanto debe ser $\boxed{\textit{indefinida}}$ o $\boxed{\textit{semidefinida}}$, ya que A tiene al menos un autovalor nulo. Este caso es más sencillo clasificarlo mediante el criterio de autovalores.*

   c) *Si $|A| < 0$ entonces q no es ni $\boxed{dp}$, ni $\boxed{\textit{semidefinida}}$, pudiendo ser sólo $\boxed{\textit{indefinida}}$ o $\boxed{dn}$. Será esto último si $a_{11} < 0$ y $D_2 > 0$.*

   d) *En el caso de $A \in \mathcal{M}_{3\times 3}$, la traza no proporciona información relevante.*

Ejemplo 224 *Clasificar la forma cuadrática q en función del parámetro $a \in \mathbb{R}$:*

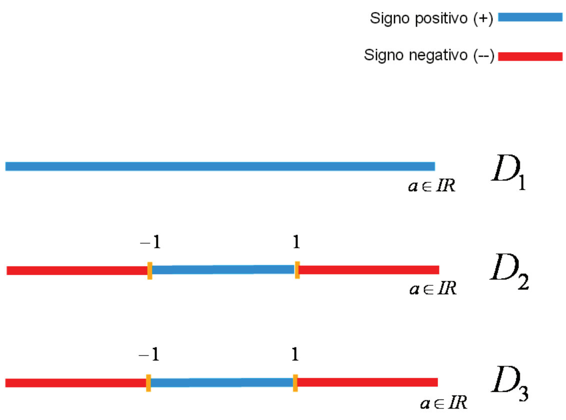
$$q(x,y,z) = (x,y,z)\begin{pmatrix} 1 & -2a & 0 \\ -2a & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Utilizaremos el criterio de menores principales, y si encontramos algún caso en el que no podamos decidir, utilizaremos el criterio de autovalores.

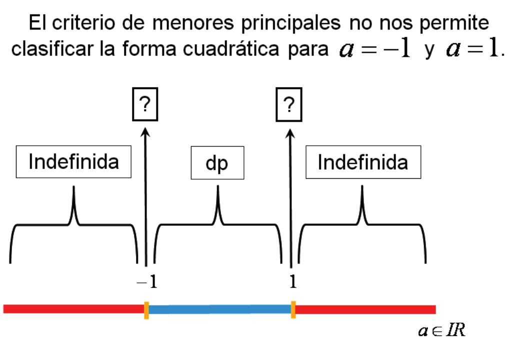
Por tanto sus menores principales toman los valores:

$$D_1 = 1 \qquad D_2 = 4 - 4a^2 \qquad D_3 = 12 - 12a^2$$

Para cada menor principal podemos representar el parámetro $a \in \mathbb{R}$ en una recta indicando su signo, de la siguiente manera



Luego, consideramos los signos de los menores principales en vertical, superponiendo todos los puntos en los que estos se anulen:



Como para $a = -1$ y $a = 1$ tenemos que $D_1 > 0$, $D_2 = D_3 = 0$ entonces el criterio de menores principlaes no nos proporciona información.

Para estos 2 casos particulares tendremos que acudir al criterio de autovalores.

Así, si $a = -1$ entonces

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 & 0 \\ 2 & 4-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{vmatrix} =$$

$$= (3-\lambda)\left[(1-\lambda)(4-\lambda) - 4\right] = \lambda(3-\lambda)(5-\lambda)$$

por lo que los autovalores son

$$\lambda_1 = 0$$
$$\lambda_2 = 3 > 0$$
$$\lambda_3 = 5 > 0$$

por lo que q es semidefinida positiva para $a = -1$.

Por otro lado, si $a = 1$ *entonces*

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & -2 & 0 \\ -2 & 4-\lambda & 0 \\ 0 & 0 & 3-\lambda \end{vmatrix} =$$

$$= (3-\lambda)\left[(1-\lambda)(4-\lambda) - 4\right] = \lambda(3-\lambda)(5-\lambda)$$

por lo que los autovalores son los mismos que anteriormente

$$\begin{aligned} \lambda_1 &= 0 \\ \lambda_2 &= 3 > 0 \\ \lambda_3 &= 5 > 0 \end{aligned}$$

por lo que q también es semidefinida positiva para $a = 1$.

Entonces se puede concluir que

$$\boxed{\begin{array}{rcl} a \in (-\infty, -1) & \Longrightarrow & q \text{ es indefinida} \\ a = -1 & \Longrightarrow & q \text{ es semidefinida positiva (sdp)} \\ a \in (-1, 1) & \Longrightarrow & q \text{ es definida positiva (dp)} \\ a = 1 & \Longrightarrow & q \text{ es semidefinida positiva (sdp)} \\ a \in (1, \infty) & \Longrightarrow & q \text{ es indefinida} \end{array}}$$

Ejemplo 225 *Clasificar la forma cuadrática* $q(x, y, z) = ax^2 + 2xy + 2xz + 2yz + ay^2 + az^2$ *en función del parámetro* $a \in \mathbb{R}$.

La expresión matricial de q es:

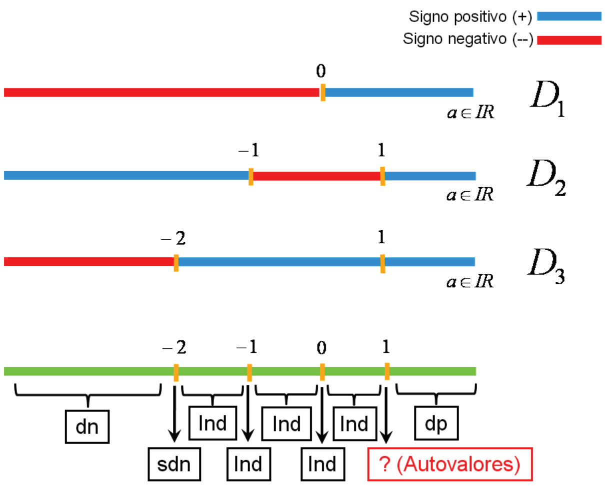
$$q(x, y, z) = (x, y, z) \begin{pmatrix} a & 1 & 1 \\ 1 & a & 1 \\ 1 & 1 & a \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

De igual forma utilizaremos el criterio de menores principales, y si encontramos algún caso en el que éste no nos proporcione información, emplearemos el criterio de autovalores.

Sus menores principales son:

$$D_1 = a \qquad D_2 = a^2 - 1 \qquad D_2 = (a-1)^2(a+2)$$

La representación de los signos de los menores principales es



Como el único valor de $a \in \mathbb{R}$ para el que no se puede clasificar la forma cuadrática q mediante el criterio de menores principales es $a = 1$, ya que $D_1 > 0$, $D_2 = D_3 = 0$, entonces utilizaremos el criterio de autovalores para este caso particular.

Así, si $a = 1$ entonces

$$P_3(\lambda) = |A - \lambda I| = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} = \lambda^2 (3 - \lambda)$$

y como los autovalores son

$$\lambda_1 = 0$$
$$\lambda_2 = 0$$
$$\lambda_3 = 3 > 0$$

tenemos que q es semidefinida positiva para $a = 1$.

En resumen, tenemos que

$a \in (-\infty, -2)$	$\Longrightarrow$	*q es definida negativa (dn)*
$a = -2$	$\Longrightarrow$	*q es semidefinida negativa (sdn)*
$a \in (-2, 1)$	$\Longrightarrow$	*q es indefinida*
$a = 1$	$\Longrightarrow$	*q es semidefinida positiva (sdp)*
$a \in (1, \infty)$	$\Longrightarrow$	*q es definida positiva (dp)*

Ejemplo 226 *Clasificar la forma cuadrática $q(x, y) = bx^2 + 2axy + 2y^2$ en función de los parámetros $a, b \in \mathbb{R}$.*

La expresión matricial de q es:

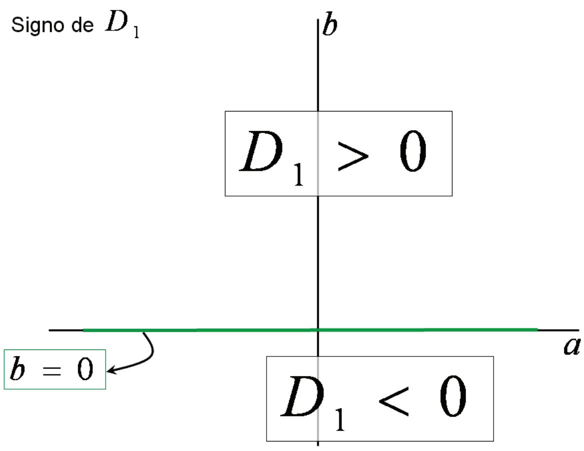
$$q(x, y) = (x, y) \begin{pmatrix} b & a \\ a & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Sus menores principales son:

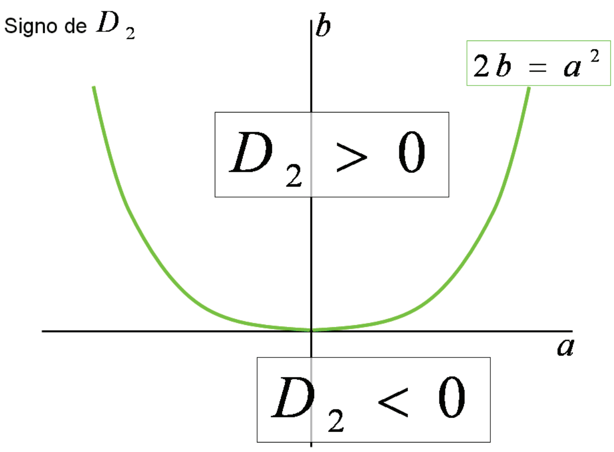
$$D_1 = b \qquad\qquad D_2 = 2b - a^2$$

Como ahora tenemos 2 parámetros, los podemos representar en 2 ejes. Tendremos que representar en este plano los puntos que anulan los menores principales, es decir, las curvas $b = 0$ y $2b - a^2 = 0$ en el plano ab.

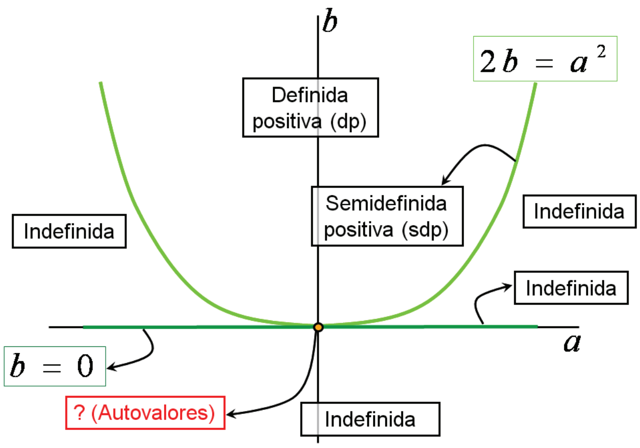
Para el primer menor principal D_1 tenemos:



Para el segundo menor principal D_2 se tiene:



Por lo tanto, superponiendo las dos gráficas y aplicando el criterio de menores principales obtenemos que:



Los únicos valores para los que no podemos clasificar la forma cuadrática mediante el criterio de menores principales son $a = 0$ y $b = 0$. Para estos valores aplicaremos el criterio de autovalores.

Así, si $a = 0$ y $b = 0$ entonces

$$P_2(\lambda) = |A - \lambda I| = \begin{vmatrix} -\lambda & 0 \\ 0 & 2-\lambda \end{vmatrix} = -\lambda(2-\lambda)$$

y como se tiene que los autovalores son

$$\lambda_1 = 0$$
$$\lambda_2 = 2 > 0$$

entonces se tiene que q es semidefinida positiva para $a = 0$ y $b = 0$.

En resumen, la clasificación queda como

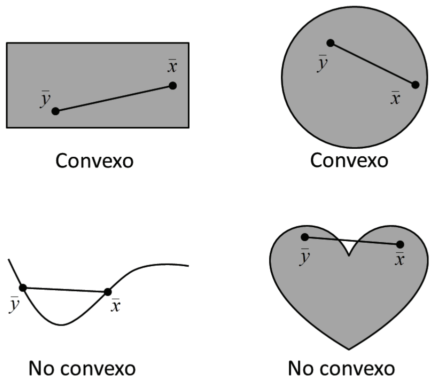
$b > 0$ *y* $2b > a^2$	$\Longrightarrow$	q *es definida positiva (dp)*
$2b = a^2$	$\Longrightarrow$	q *es semidefinida positiva (sdp)*
Resto de valores $a, b \in \mathbb{R}$	$\Longrightarrow$	q *es indefinida*

Capítulo 7

Convexidad de conjuntos y funciones

7.1. Conjuntos convexos

Un conjunto S es **convexo** si para cada par de puntos suyos, el segmento que los une está totalmente contenido en S.



Definición 227 *Un subconjunto $S \subset \mathbb{R}^n$ es **convexo** si para todo par de puntos $\bar{x}, \bar{y} \in S$ y todo $\lambda \in [0,1]$ se verifica que*

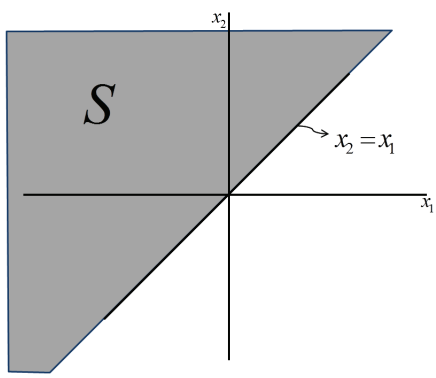
$$\bar{z} = \lambda \bar{x} + (1 - \lambda)\, \bar{y} \in S$$

*es decir, si llamamos **segmento de extremos** $\bar{x}$ e $\bar{y}$ al conjunto*

$$[\bar{x}, \bar{y}] = \{\bar{z} \in \mathbb{R}^n : \bar{z} = \lambda \bar{x} + (1 - \lambda)\, \bar{y}, \lambda \in [0,1]\}$$

entonces decimos que S es convexo si para todo $\bar{x}, \bar{y} \in S$ se verifica que $[\bar{x}, \bar{y}] \subset S$.

Ejemplo 228 *¿Es el conjunto $S = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq x_1\right\}$ convexo?*



Sean $\left\{\begin{array}{l}\bar{x}=(x_1,x_2)\\ \bar{y}=(y_1,y_2)\end{array}\right.$ *tales que* $\bar{x},\bar{y}\in S$, *probaremos que para todo* $\lambda\in[0,1]$ *se tiene que el punto* $\bar{z}=\lambda\bar{x}+(1-\lambda)\,\bar{y}$ *pertenece a* S.

Así,

$$\bar{z}=(z_1,z_2)=\lambda\,(x_1,x_2)+(1-\lambda)\,(y_1,y_2)=(\lambda x_1+(1-\lambda)\,y_1,\lambda x_2+(1-\lambda)\,y_2)$$

por lo que $\left\{\begin{array}{l}z_1=\lambda x_1+(1-\lambda)\,y_1\\ z_2=\lambda x_2+(1-\lambda)\,y_2\end{array}\right.$. *Como* $\bar{x},\bar{y}\in S$ *entonces verifican que*

$$\left\{\begin{array}{l}x_2\geq x_1\\ y_2\geq y_1\end{array}\right.$$

y por lo tanto, como $\lambda\geq 0$ *y* $1-\lambda\geq 0$ *entonces*

$$\left\{\begin{array}{l}\lambda x_2\geq\lambda x_1\\ (1-\lambda)\,y_2\geq(1-\lambda)\,y_1\end{array}\right.$$

Sumando tenemos que

$$\lambda x_2+(1-\lambda)\,y_2\geq\lambda x_1+(1-\lambda)\,y_1$$

es decir,

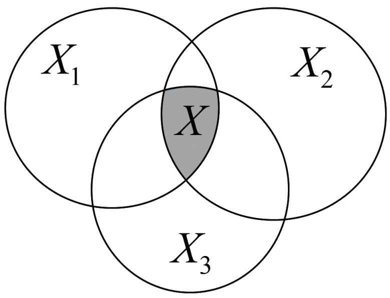
$$z_2\geq z_1$$

luego $\bar{z}\in S$ *y esto demusetra que* S *es convexo.*

7.1.1. Propiedades de los conjuntos convexos

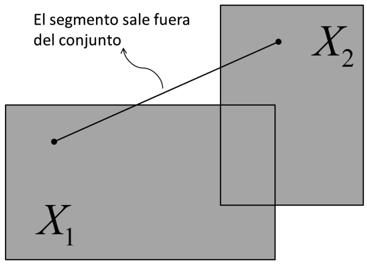
Intersección

Sea $\{X_i : i\in I\}$ una familia de subconjuntos convexos de $\mathbb{R}^n$, entonces $X=\bigcap_{i\in I}X_i$ es convexo.



Unión

La unión de dos conjuntos convexos, en general, no es un conjunto convexo.



Suma de dos conjuntos convexos

Sea $\{X_i : i \in I\}$ una familia de subconjuntos convexos de $\mathbb{R}^n$, entonces si I es un conjunto finito ($I = \{1, \ldots, m\}$), el conjunto suma dado por

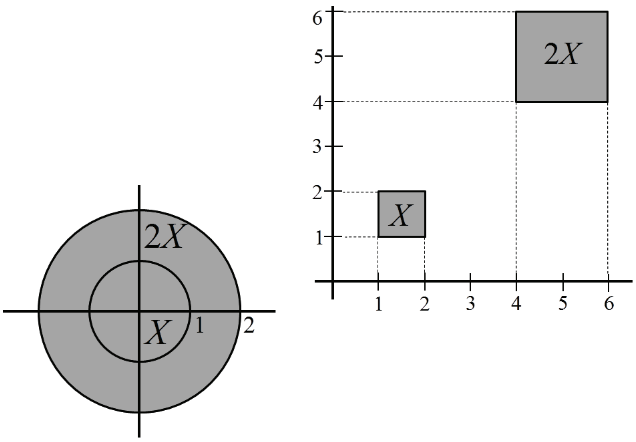
$$X = \sum_{i=1}^{m} X_i = \left\{ \bar{x} \in \mathbb{R}^n : \bar{x} = \sum_{i=1}^{m} \bar{x}_i \ \text{ con } \ \bar{x}_i \in X_i \ , \ i = 1, \ldots, m \right\}$$

es convexo.

Producto por un escalar

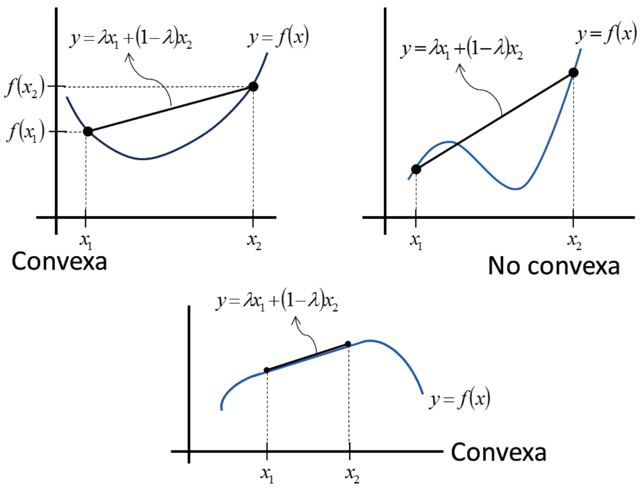
Si $X \subset \mathbb{R}^n$ es un conjunto convexo y $\alpha \in \mathbb{R}$, entonces el conjunto dado por

$$\alpha X = \{\bar{y} \in \mathbb{R}^n : \bar{y} = \alpha \bar{x} \ \text{ con } \ \bar{x} \in X\}$$



7.2. Funciones cóncavas y convexas

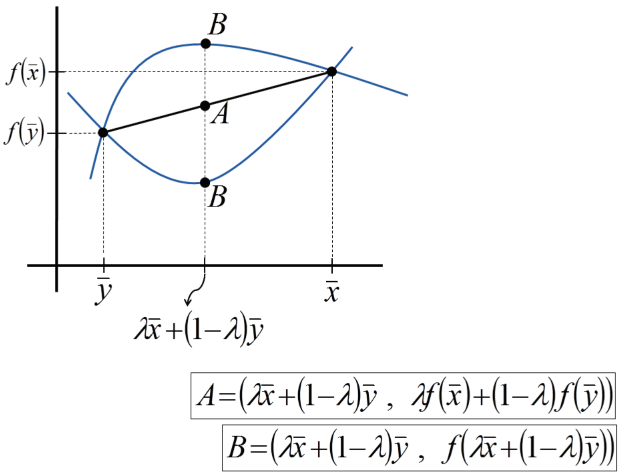
Una **función es convexa** si el segmento que une dos puntos cualesquiera de su gráfica no está por debajo de dicha gráfica. Del mismo modo podemos decir que una **función es cóncava** si el segmento que une dos puntos cualesquiera de su gráfica no está por encima de su gráfica.



Definición 229 *Sea M un subconjuto convexo y no vacío de $\mathbb{R}^n$, y sea $f : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función, entonces se tiene que*

1. *La **función** f **es convexa** en M si y sólo si para cualquier $\bar{x}, \bar{y} \in M$ y para todo $\lambda \in [0,1]$ se verifica que*
$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$
2. *La **función** f **es cóncava** en M si y sólo si para cualquier $\bar{x}, \bar{y} \in M$ y para todo $\lambda \in [0,1]$ se verifica que*
$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) \geq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$
3. *La **función** f **es estrictamente convexa** en M si y sólo si para cualquier $\bar{x}, \bar{y} \in M$ con $\bar{x} \neq \bar{y}$ y para todo $\lambda \in [0,1]$ se verifica que*
$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) < \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$
4. *La **función** f **es estrictamente cóncava** en M si y sólo si para cualquier $\bar{x}, \bar{y} \in M$ con $\bar{x} \neq \bar{y}$ y para todo $\lambda \in [0,1]$ se verifica que*
$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) > \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

7.2.1. Interpretación gráfica de la definición



Si el punto B no está por encima de A entonces

$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) \leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

y en caso de que el punto B no esté por debajo de A se verifica que

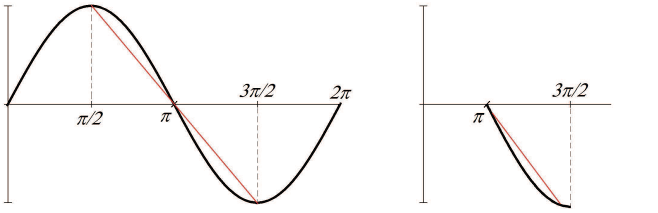
$$f(\lambda \bar{x} + (1-\lambda)\bar{y}) \geq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

Observación 230 *En la definición anterior no se pide la continuidad o derivabilidad de la función f.*

Observación 231 *Sólo tiene sentido hablar de concavidad o convexidad de funciones **sobre conjuntos convexos**.*

Observación 232 *El conjunto sobre el que se define la función es muy importante e influye en el carácter de la función.*

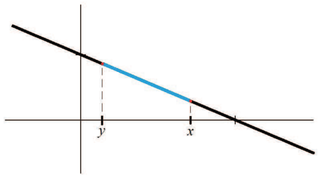
Ejemplo 233 *Sea* $f(x) = \operatorname{sen} x$. *Si consideramos* f *definida en distintos conjuntos* $M_1 = [0, 2\pi]$ *y* $M_2 = \left[\pi, \frac{3\pi}{2}\right]$ *se puede observar que* f *no es ni cóncava ni convexa en* M_1 *pero es estrictamente convexa en* M_2.



Ni cóncava ni convexa *Estrictamente convexa*

Ejemplo 234 *Estudiemos gráficamente la convexidad o concavidad de diversas funciones:*

1. $f(x) = ax + b$ *en* $\mathbb{R}$ *con* $a, b \in \mathbb{R}$.



Cóncava y convexa

Se observa que es cóncava y convexa aunque no estrictamente. La demostración analítica es sencilla, veamos que

$$\begin{aligned} f(\lambda\bar{x} + (1-\lambda)\bar{y}) &\geq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y}) \\ f(\lambda\bar{x} + (1-\lambda)\bar{y}) &\leq \lambda f(\bar{x}) + (1-\lambda) f(\bar{y}) \end{aligned}$$

por lo que

$$f(\lambda\bar{x} + (1-\lambda)\bar{y}) = \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

En efecto, como

$$f(\lambda\bar{x} + (1-\lambda)\bar{y}) = a(\lambda\bar{x} + (1-\lambda)\bar{y}) + b =$$

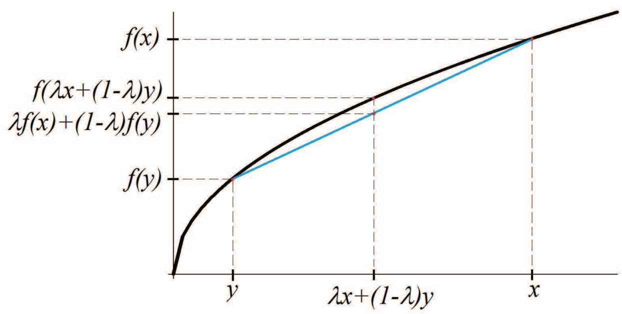
$$= a\lambda\bar{x} + a(1-\lambda)\bar{y} + b = \lambda a\bar{x} + (1-\lambda)a\bar{y} + \lambda b + (1-\lambda)b =$$

$$= \lambda a\bar{x} + \lambda b + (1-\lambda)a\bar{y} + (1-\lambda)b = \lambda(a\bar{x} + b) + (1-\lambda)(a\bar{y} + b) =$$

$$= \lambda f(\bar{x}) + (1-\lambda) f(\bar{y})$$

luego f *es cóncava y convexa.*

2. $f(x) = \sqrt{x}$ *en* $\mathbb{R}$.



Estrictamente cóncava

Se observa que es estrictamente cóncava. Para demostrarlo analíticamente hay que verificar que

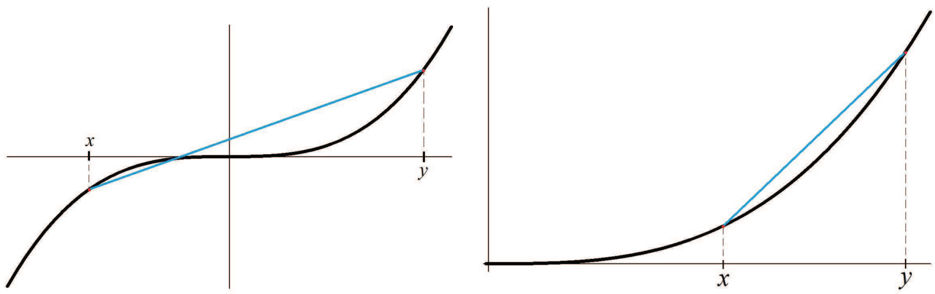
$$f\left(\lambda\bar{x} + (1-\lambda)\,\bar{y}\right) > \lambda f\left(\bar{x}\right) + (1-\lambda)\, f\left(\bar{y}\right)$$

es decir

$$\sqrt{\lambda\bar{x} + (1-\lambda)\,\bar{y}} > \lambda\sqrt{\bar{x}} + (1-\lambda)\sqrt{\bar{y}}$$

que se deja como ejercicio.

3. $f\left(x\right) = x^3$ *en* $\mathbb{R}$ *y en* $\mathbb{R}_+$.



Ni cóncava ni convexa — Estrictamente cóncava

Se observa que f no es ni cóncava ni convexa en $\mathbb{R}$, pero es estrictamente convexa en $\mathbb{R}_+$. Para demostrar esto, es necesario verificar que

$$f\left(\lambda\bar{x} + (1-\lambda)\,\bar{y}\right) > \lambda f\left(\bar{x}\right) + (1-\lambda)\, f\left(\bar{y}\right)$$

es decir

$$\left(\lambda\bar{x} + (1-\lambda)\,\bar{y}\right)^3 < \lambda\bar{x}^3 + (1-\lambda)\,\bar{y}^3$$

Se deja como ejercicio.

7.2.2. Propiedades de las funciones cóncavas y convexas

Proposición 235 *Sea M un subconjunto convexo de $\mathbb{R}^n$ y consideremos $f : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}$. Se verifica que*

1. *Si f es convexa en M, entonces los conjuntos $\Lambda_\alpha = \{\bar{x} \in M : f\left(\bar{x}\right) \leq \alpha\}$ son convexos para todo $\alpha \in \mathbb{R}$.*

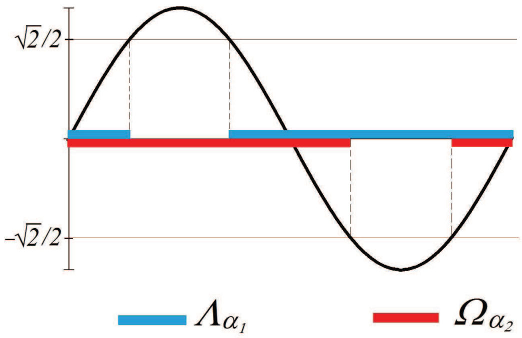
2. *Si f es cóncava en M, entonces los conjuntos $\Omega_\alpha = \{\bar{x} \in M : f\left(\bar{x}\right) \geq \alpha\}$ son convexos para todo $\alpha \in \mathbb{R}$.*

Observación 236 *La proposición anterior se puede utilizar para demostrar que una función no es cóncava o convexa como ilustramos en los siguientes ejemplos.*

Ejemplo 237 *Consideremos los siguientes ejemplos:*

1. *Sea $f\left(x\right) = \operatorname{sen} x$ con $x \in [0, 2\pi]$.*

   Si encontramos α_1 tal que Λ_{α_1} no es convexo, entonces la función f no es convexa. De igual forma, si encontramos α_1 tal que Ω_{α_2} no es convexo, entonces la función f no es cóncava.



En efecto, si $\alpha_1 = \frac{\sqrt{2}}{2}$ *y* $\alpha_2 = \frac{-\sqrt{2}}{2}$ *entonces*

$$\Lambda_{\alpha_1} = \left\{ x \in [0, 2\pi] : \operatorname{sen} x \leq \frac{\sqrt{2}}{2} \right\} = \left[0, \frac{\pi}{4}\right] \cup \left[\frac{3\pi}{4}, 2\pi\right]$$

$$\Omega_{\alpha_2} = \left\{ x \in [0, 2\pi] : \operatorname{sen} x \geq \frac{-\sqrt{2}}{2} \right\} = \left[0, \frac{5\pi}{4}\right] \cup \left[\frac{7\pi}{4}, 2\pi\right]$$

que no son convexos, luego f *no es cóncava ni convexa en* $[0, 2\pi]$.

2. *Sea* $f(x) = \operatorname{sen} x$ *con valores* $x \in \left[\frac{\pi}{2}, \pi\right]$.

   Aplicando las propiedades anteriores no es posible determinar si f *es cóncava o convexa en* $\left[\frac{\pi}{2}, \pi\right]$ *ya que para cualquier* $\alpha \in \mathbb{R}$ *se tiene que tanto* Λ_α *como* Ω_α *son convexos.*

3. *Sea* $f(x) = e^x$ *con* $x \in \mathbb{R}$.

   De nuevo, aplicando las propiedades no es posible determinar si f *es cóncava o convexa en* $\mathbb{R}$ *ya que para cualquier* $\alpha \in \mathbb{R}$ *se tiene que tanto* Λ_α *como* Ω_α *son convexos.*

4. *Sea* $f(x) = \frac{1}{x}$ *con* $x > 0$.

   Como $\Lambda_\alpha = \left\{x \in \mathbb{R} : \frac{1}{x} \leq \alpha\right\} = \left[\frac{1}{\alpha}, \infty\right)$ *y como* $\Omega_\alpha = \left\{x \in \mathbb{R} : \frac{1}{x} \geq \alpha\right\} = \left(0, \frac{1}{\alpha}\right]$ *que son conjuntos convexos, entonces las propiedades anteriores no nos permiten determinar si* f *es convexa o cóncava.*

Proposición 238 *Sea* M *un subconjunto convexo de* $\mathbb{R}^n$ *y* $f : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ *una función. Se verifica que:*

1. *Si* f *es convexa en* M *entonces la función* $-f$ *es cóncava en* M.

2. *Si* f *es estrictamente convexa en* M *entonces la función* $-f$ *es estrictamente cóncava en* M.

3. *Sea* f *es convexa en* M *y* $\alpha \in \mathbb{R}$. *Si* $\alpha \geq 0$ *entonces la función* αf *es convexa en* M *y si* $\alpha \leq 0$ *entonces la función* αf *es cóncava en* M.

4. *Si* $\{f_i : i = 1, \ldots, m\}$ *es una familia de funciones convexas en* M *entonces la función* f *definida como* $f = \sum_{i=1}^{m} \alpha_i f_i$ *con* $\alpha_i \geq 0$ *para* $i = 1, \ldots, m$ *es una función convexa en* M.

Observación 239 *La suma de funciones convexa es convexa, pero esto, en general no es cierto con el producto de funciones. Por ejemplo las funciones* $f(x) = x$ *y* $g(x) = x^2$ *son convexas, pero su producto* $fg(x) = x^3$ *no lo es.*

7.3. Funciones cóncavas y convexas diferenciables

Proposición 240 *Sea M un subconjunto abierto, no vacío y convexo de $\mathbb{R}^n$, y $f : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función diferenciable. Se verifica que:*

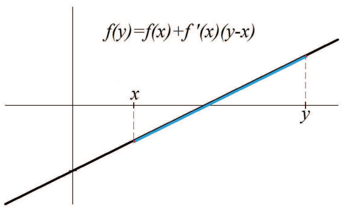
1. *La función f es convexa en M si y sólo si para cualesquier puntos $\bar{x}, \bar{y} \in M$ se tiene que*
$$f(\bar{y}) \geq f(\bar{x}) + \nabla f(\bar{x})(\bar{y} - \bar{x})$$
o bien
$$[\nabla f(\bar{y}) - \nabla f(\bar{x})](\bar{y} - \bar{x}) \geq 0$$
donde ∇f es el gradiente de f.

2. *La función f es estrictamente convexa en M si y sólo si para cualesquier puntos $\bar{x}, \bar{y} \in M$ con $\bar{x} \neq \bar{y}$ se tiene que*
$$f(\bar{y}) > f(\bar{x}) + \nabla f(\bar{x})(\bar{y} - \bar{x})$$
o bien
$$[\nabla f(\bar{y}) - \nabla f(\bar{x})](\bar{y} - \bar{x}) > 0$$
donde ∇f es el gradiente de f.

3. *Los resultados para f cóncava se obtienen invirtiendo el símbolo de las respectivas desigualdades.*

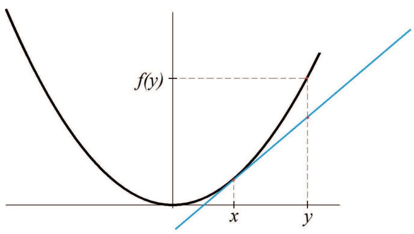
Observación 241 *La proposición anterior nos caracteriza las funciones convexas y cóncavas dependiendo de si la gráfica de f está por encima o por debajo de cualquier recta tangente a ésta.*

Ejemplo 242 *La función $f(x) = ax + b$ en $\mathbb{R}$ es cóncava y convexa.*



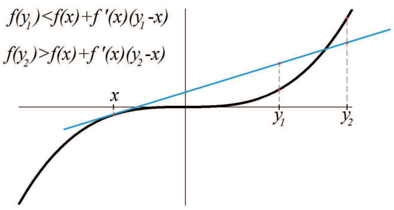
Cóncava y convexa

Ejemplo 243 *La función $f(x) = x^2$ en $\mathbb{R}$ es convexa.*



Estrictamente convexa

Ejemplo 244 *La función $f(x) = x^3$ en $\mathbb{R}$ no es ni cóncava ni convexa.*



Ni cóncava ni convexa

Proposición 245 *Sea M un subconjunto abierto, no vacío y convexo de $\mathbb{R}^n$, y $f : M \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ una función diferenciable con derivadas parciales continuas hasta el orden 2. Se verifica que:*

1. *La función f es convexa en M si y sólo si para todo $\bar{x} \in M$ se verifica que $\bar{y}^T Hf(\bar{x})\,\bar{y} \geq 0$ para cualquier $\bar{y} \in \mathbb{R}^n$, es decir, para todo $\bar{x} \in M$ la forma cuadrática con matriz asociada $Hf(\bar{x})$ es semidefinida positiva o definida positiva.*
2. *Si para todo $\bar{x} \in M$ se verifica que $\bar{y}^T Hf(\bar{x})\,\bar{y} > 0$ para cualquier $\bar{y} \in \mathbb{R}^n$, es decir, para todo $\bar{x} \in M$ la forma cuadrática con matriz asociada $Hf(\bar{x})$ es definida positiva, entonces la función f es estrictamente convexa en M.*
3. *Los resultados para funciones cóncavas se obtienen considerando las formas cuadráticas definidas y semidefinidas negativas.*

Ejemplo 246 *Sea $f(x,y) = \log(xy)$ en $\mathbb{R}^2_{++}$.*

Como la matriz hessiana de f es $Hf(x,y) = \begin{pmatrix} \frac{-1}{x^2} & 0 \\ 0 & \frac{-1}{y^2} \end{pmatrix}$ entonces es definida negativa en $\mathbb{R}^2_{++}$, luego f es estrictamente cóncava.

Ejemplo 247 *Sea $f(x,y) = ax^2 + by^2$ en $\mathbb{R}^2$ con $a, b \in \mathbb{R}$.*

La matriz hessiana de f es $Hf(x,y) = \begin{pmatrix} 2a & 0 \\ 0 & 2b \end{pmatrix}$ entonces se tiene que

si $a > 0$ y $b > 0$	$\Longrightarrow$	Hf es D.P.	$\Longrightarrow$	f es estrictamente convexa
si $\begin{cases} a = 0 \text{ y } b > 0 \\ a > 0 \text{ y } b = 0 \end{cases}$	$\Longrightarrow$	Hf es S.D.P.	$\Longrightarrow$	f es convexa
si $a < 0$ y $b < 0$	$\Longrightarrow$	Hf es D.N.	$\Longrightarrow$	f es estrictamente cóncava
si $\begin{cases} a = 0 \text{ y } b < 0 \\ a < 0 \text{ y } b = 0 \end{cases}$	$\Longrightarrow$	Hf es S.D.N.	$\Longrightarrow$	f es cóncava
si $a = 0$ y $b = 0$	$\Longrightarrow$	$f \equiv 0$	$\Longrightarrow$	f es cóncava y convexa
si $\begin{cases} a > 0 \text{ y } b < 0 \\ a < 0 \text{ y } b > 0 \end{cases}$	$\Longrightarrow$	Hf es indefinida	$\Longrightarrow$	f no es cóncava ni convexa

Ejemplo 248 *Sea $f(x,y,z) = e^{xy+z}$ en $\mathbb{R}^3$.*

La matriz hessiana de f es $Hf(x,y,z) = e^{xy+z}\begin{pmatrix} y^2 & 1+xy & y \\ 1+xy & x^2 & x \\ y & x & 1 \end{pmatrix}$ entonces, clasificando por menores principales se tiene que

$$\begin{array}{ll} D_1 = y^2 \geq 0 & \\ D_2 = -(1+2xy)\,e^{2(xy+z)} & \text{que puede tomar cualquier signo.} \\ D_3 = -1 < 0 & \end{array}$$

luego Hf es indefinida en $\mathbb{R}^3$. Por tanto, f no es ni cóncava ni convexa.

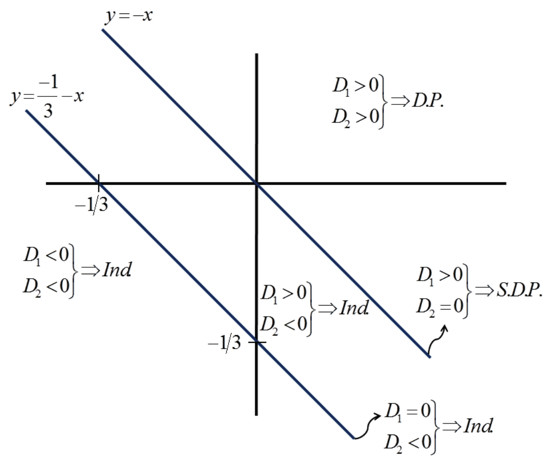
Ejemplo 249 *Dada $f(x,y) = (2+x)^2 + (y+x)^3$ determinar los subconjuntos de $\mathbb{R}^2$ en los que f es convexa o cóncava.*

El hessiano de f es $Hf(x,y) = \begin{pmatrix} 2+6(x+y) & 6(x+y) \\ 6(x+y) & 6(x+y) \end{pmatrix}$. Si aplicamos el criterio de menores principales para clasificar la forma cuadrática asociada obtenemos que

$$D_1 = 2 + 6(x+y) \qquad D_2 = 12(x+y)$$

por lo que

$$\begin{array}{lcl} D_1 \geq 0 & \Longrightarrow & y \geq \frac{-1}{3} - x \\ D_2 \geq 0 & \Longrightarrow & y \geq -x \end{array}$$



Entonces

$$\begin{array}{lll} En\ \left\{(x,y)\in\mathbb{R}^2 : y>-x\right\} & \Longrightarrow & f \text{ es estrictamente convexa} \\ En\ \left\{(x,y)\in\mathbb{R}^2 : y\geq -x\right\} & \Longrightarrow & f \text{ es convexa} \\ En\ \left\{(x,y)\in\mathbb{R}^2 : y<-x\right\} & \Longrightarrow & f \text{ no es ni cóncava ni convexa} \end{array}$$

Ejemplo 250 *Comprobar si el conjunto* $M=\left\{(x,y)\in\mathbb{R}^2 : 3x^2+2xy+2y^2\leq 30\right\}$ *es convexo.*

Si consideramos $f(x,y)=3x^2+2xy+2y^2$ *entonces* $M=\Lambda_{30}=\left\{(x,y)\in\mathbb{R}^2 : f(x,y)\leq 30\right\}$, *por lo que nos basta demostrar que* f *es convexa para concluir que* Λ_{30} *(y por consiguiente* M*) es convexo.*

El hessiano de f *es* $Hf(x,y)=\begin{pmatrix} 6 & 2 \\ 2 & 4 \end{pmatrix}$ *y como sus menores principales son* $D_1=6>0$ *y* $D_2=20>0$ *entonces* Hf *es definida positivapor lo que* f *es convexa. Luego* M *es convexo.*

Ejemplo 251 *Comprobar si el conjunto* $M=\left\{(x,y)\in\mathbb{R}^2 : e^{x+y}\leq 10\right\}$ *es convexo.*

Consideremos $f(x,y)=e^{x+y}$ *entonces* $M=\Lambda_{10}=\left\{(x,y)\in\mathbb{R}^2 : f(x,y)\leq 10\right\}$. *Veamos si* f *es convexa, en cuyo caso concluiremos que* Λ_{30} *(y por consiguiente* M*) es convexo.*

El hessiano de f *es* $Hf(x,y)=\begin{pmatrix} e^{x+y} & e^{x+y} \\ e^{x+y} & e^{x+y} \end{pmatrix}$ *y como sus menores principales son* $D_1=e^{x+y}>0$ *y* $D_2=0$ *entonces* Hf *es semidefinida positiva, luego* f *es convexa. Por consiguiente* M *es convexo.*