Universidad Nacional Autonoma de México

Facultad de Ciencias

NOTAS PARA UN CURSO DE ALGEBRA LINEAL

TESIS

que para obtener el titulo de

MATEMATICO

PRESENTA

ELSA CAMPUZANO VAZQUEZ

México, D.F. 1983

INDICE

CAPITULO II : DEPENDENCIA LINEAL EN IR^m

CAPITULO III : TEORIA DE ECUACIONES

CAPITULO IV : ESPACIOS VECTORIALES

CAPITULO V : FUNCIONES CUADRATICAS Y FORMAS BILINEALES

CAPITULO VI : ESPACIOS CON PRODUCTO INTERIOR

INTRODUCCION

Al mismo tiempo de cumplir con el requisito que señala la Universidad Nacional Autónoma de México para poder aspirar al grado de licenciatura en Matemáticas, concebí la idea junto con mi maestro el Dr. Pablo Barrera Sánchez, de elaborar mi tesis con las características tales que pudiera servir de texto para los estudiantes en los cursos de Algebra Lineal que se imparten en la Facultad de Ciencias.

El texto es autocontenido y cualquier estudiante lo puede leer y estudiar fácilmente.

El curso de Algebra Lineal que el dr. Barrera imparte ha ido evolucionando, y estas notas son el resultado de este desarrollo. Se da un curso de Algebra Lineal basado en el algebra matricial, tratando de dar un enfoque dinámico a las matrices -- una matriz opera sobre otra y la transforma según nosotros queramos. El Algebra Lineal no solamente es una bella materia sino también una materia útil.

Las notas que sirvieron para redactar este texto fueron tomadas en los cursos de Algebra Lineal del maestro Barrera. No hay libros que contengan el material que expongo en mi tesis, cuidadosamente revisado por él.

Deseamos cooperar con ella para resolver el problema que resulta de que la mayoría de los libros de Matemáticas en la actualidad vengan del extranjero y por tanto resulten demasiado caros, además de que se encuentran redactados en Inglés, razones por las cuales no son accesibles a todos los estudiantes.

Elsa Campozano

México D.F. agosto 1983

CAPITULO I

VECTORES
Y
MATRICES

VECTORES

DEFINICION

$$\mathbb{R}^n = \left\{ \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} : \xi_1, \xi_2, \ldots, \xi_n \in \mathbb{R} \right\}$$

A los elementos de $\mathbb{R}^n$ los llamamos VECTORES COLUMNA DE ORDEN ó DIMENSION n. Los números $\xi_1, \xi_2, \ldots, \xi_n$ son los ELEMENTOS ó COMPONENTES del vector.

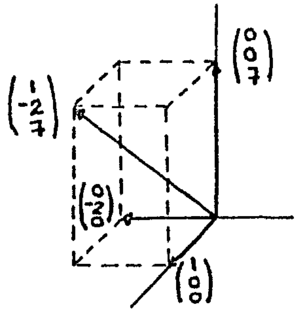
Siempre que nos refiramos a un vector de orden n estaremos pensando en un vector columna de orden n, y cuando n sea fijo, ó se pueda determinar del contexto, emplearemos únicamente la palabra 'vector'.

El uso del término 'dimensión' es por analogía con la geometría - un vector de orden n se puede ver como un punto en el espacio n-dimensional.

Por ejemplo, el vector

$$\begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$$

quedaría representado así:



Escribiremos $[\xi_i]$ para denotar al vector cuya i-ésima componente es ξ_i.

Las operaciones básicas son la suma de dos vectores y la multiplicación de un vector por un escalar:

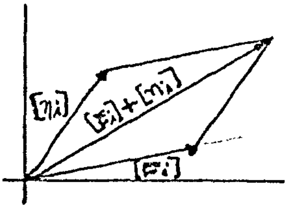
DEFINICION

Sean $[\xi_i], [\eta_i] \in \mathbb{R}^n$. La SUMA de $[\xi_i]$ y $[\eta_i]$, denotada por $[\xi_i] + [\eta_i]$, es el vector $[\gamma_i]$ dado por

$$[\gamma_i] = [\xi_i + \eta_i]$$

La suma de dos vectores tiene la siguiente interpretación geométrica:

Si formamos el paralelogramo cuyos lados son $[\xi_i]$ y $[\eta_i]$, entonces la suma de $[\xi_i]$ y $[\eta_i]$ es la diagonal del paralelogramo que parte del origen.

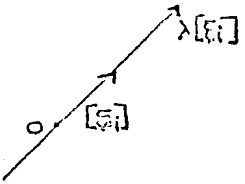


DEFINICION

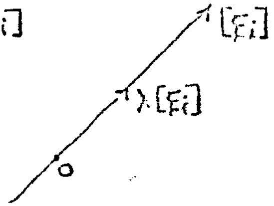
Sea $\lambda \in \mathbb{R}$ y $[\xi_i] \in \mathbb{R}^n$. El PRODUCTO de λ y $[\xi_i]$, escrito $\lambda[\xi_i]$, es el vector $[\beta_i]$ dado por:

$$[\beta_i] = [\lambda \xi_i]$$

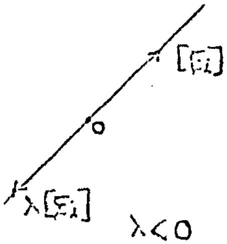
Geométricamente, la operación de multiplicar el vector $[\xi_i]$ por el escalar λ cambia la longitud de $[\xi_i]$. por un factor de $|\lambda|$. si $\lambda<0$, $\lambda[\xi_i]$ tiene dirección opuesta a la de $[\xi_i]$



$\lambda > 1$



$0 < \lambda < 1$



$\lambda < 0$

DEFINICION

El VECTOR CERO de $\mathbb{R}^n$, denotado por 0, es el vector cuyas n componentes valen cero.

Según el contexto en que se use el símbolo "0" sabremos si se trata del escalar cero ó del vector cero.

PROPIEDADES

Sean $[\xi_i], [\eta_i], [\zeta_i] \in \mathbb{R}^n$, $\lambda, \mu \in \mathbb{R}$.

1. $[\xi_i] + [\eta_i] = [\eta_i] + [\xi_i]$

2. $([\xi_i] + [\eta_i]) + [\zeta_i] = [\xi_i] + ([\eta_i] + [\zeta_i])$

3. $[\xi_i] + 0 = [\xi_i]$

4. Para todo $[\xi_i] \in \mathbb{R}^n$ existe un vector, denotado por $-[\xi_i]$, tal que

$$[\xi_i] + (-[\xi_i]) = 0$$

5. $1 \cdot [\xi_i] = [\xi_i]$

6. $\lambda(\mu[\xi_i]) = (\lambda\mu)[\xi_i]$

7. $(\lambda + \mu)[\xi_i] = \lambda[\xi_i] + \mu[\xi_i]$

8. $\lambda([\xi_i]+[\eta_i]) = \lambda[\xi_i] + \lambda[\eta_i]$

Dos vectores de orden n son iguales si y solo si sus componentes correspondientes son iguales.

DEFINICION

$$\mathbb{R}^{n*} = \{[\xi_1 \ \xi_2 \ \dots \ \xi_n] : \xi_1, \xi_2, \dots, \xi_n \in \mathbb{R}\}$$

A los elementos de $\mathbb{R}^{n*}$ los llamamos VECTORES RENGLON DE ORDEN n. $\xi_1, \xi_2, \dots, \xi_n$ son los ELEMENTOS ó COMPONENTES del vector renglón.

Nos referiremos a un vector renglón de orden n como un "renglón" ó un "renglón de orden n", cuando sea necesario especificar su dimensión.

DEFINICION

Sean $[\xi_1 \ \xi_2 \ \dots \ \xi_n]$ y $[\eta_1 \ \eta_2 \ \dots \ \eta_n] \in \mathbb{R}^{n*}$ y $\lambda \in \mathbb{R}$. Entonces

La SUMA de $[\xi_1 \ \xi_2 \ \dots \ \xi_n]$ y $[\eta_1 \ \eta_2 \dots \eta_n]$ es el vector renglón $[\gamma_1 \ \gamma_2 \dots \gamma_n]$, denotado por $[\xi_1 \ \xi_2 \dots \xi_n] + [\eta_1 \ \eta_2 \dots \eta_n]$, dado por :

$$[\gamma_1\ \gamma_2 \ldots \gamma_n] = [\xi_1+\eta_1\ \ \xi_2+\eta_2\ \ldots\ \xi_n+\eta_n]$$

El PRODUCTO de λ y $[\xi_1\ \xi_2 \ldots \xi_n]$ es el vector renglón $[\beta_1\ \beta_2 \ldots \beta_n]$, denotado por $\lambda[\xi_1\ \xi_2 \ldots \xi_n]$, dado por

$$[\beta_1\ \beta_2 \ldots \beta_n] = [\lambda\xi_1\ \ \lambda\xi_2\ \ldots \lambda\xi_n]$$

PROPIEDADES

Sean $[\xi_1\ \xi_2 \ldots \xi_n], [\eta_1\ \eta_2 \ldots \eta_n], [\zeta_1\ \zeta_2 \ldots \zeta_n] \in \mathbb{R}^{n*}$ y $\lambda, \mu \in \mathbb{R}$. Entonces

1. $[\xi_1\ \xi_2 \ldots \xi_n] + [\eta_1\ \eta_2 \ldots \eta_n] = [\eta_1 \ldots \eta_n] + [\xi_1 \ldots \xi_n]$

2. $([\xi_1 \ldots \xi_n] + [\eta_1 \ldots \eta_n]) + [\zeta_1 \ldots \zeta_n] = [\xi_1 \ldots \xi_n] + ([\eta_1 \ldots \eta_n] + [\zeta_1 \ldots \zeta_n])$

3. El renglón $[0\ 0 \ldots 0]$ tiene la propiedad de que

$$[0\ 0 \ldots 0] + [\xi_1\ \xi_2 \ldots \xi_n] = [\xi_1\ \xi_2 \ldots \xi_n]$$

4. Para todo $[\xi_1\ \xi_2 \ldots \xi_n] \in \mathbb{R}^{n*}$ existe un renglón, denotado por $-[\xi_1\ \xi_2 \ldots \xi_n]$, tal que

$$[\xi_1\ \xi_2 \ldots \xi_n] + (-[\xi_1\ \xi_2 \ldots \xi_n]) = [0\ 0 \ldots 0]$$

5. $1\cdot[\xi_1\ \xi_2 \ldots \xi_n] = [\xi_1\ \xi_2 \ldots \xi_n]$

6. $\lambda(\mu[\xi_1\ \xi_2 \ldots \xi_n] = (\lambda\mu)[\xi_1\ \xi_2 \ldots \xi_n]$

7. $(\lambda+\mu)[\xi_1\ \xi_2 \ldots \xi_n] = \lambda[\xi_1\ \xi_2 \ldots \xi_n] + \mu[\xi_1\ \xi_2 \cdots \xi_n]$

8. $\lambda([\xi_1\ \xi_2 \ldots \xi_n] + [\eta_1\ \eta_2 \ldots \eta_n]) = \lambda[\xi_1\ \xi_2 \cdots \xi_n] + \lambda[\eta_1\ \eta_2 \cdots \eta_n]$

Al igual que para los vectores columna, dos vectores renglón son iguales si y solo si son iguales componente a componente.

TRANSPOSICION DE VECTORES

Sea $\varphi : \mathbb{R}^n \to \mathbb{R}^{n*}$ dada por

$$\varphi\begin{pmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{pmatrix} = [\xi_1\ \xi_2 \ldots \xi_n]$$

A φ la podemos ver como la función que transforma columnas en renglones. A esta operación la llamaremos TRANSPOSICION

Si denotamos a los vectores columna

por

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$ y a $\varphi(x)$ por $\phi(x) = x^t$,

(x^t se lee "x transpuesto")

entonces x^t es un vector renglón.

Como se puede ver, φ es una función uno-uno y sobre, ya que si

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}, \quad y = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix} \in \mathbb{R}^n$$, entonces

1) $\varphi(x) = \varphi(y) \Rightarrow [\xi_1 \; \xi_2 \; \dots \xi_n] = [\eta_1 \; \eta_2 \; \dots \; \eta_n]$

$$\Rightarrow \begin{array}{l} \xi_1 = \eta_1 \\ \xi_2 = \eta_2 \\ \vdots \\ \xi_n = \eta_n \end{array}$$

$$\Rightarrow \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{pmatrix}$$

$$\Rightarrow x = y$$

$\therefore$ φ inyectiva

2) Sea $[\xi_1 \ \xi_2 \ \ldots \ \xi_n] \in \mathbb{R}^{n*}$, entonces

$$\varphi \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = [\xi_1 \ \xi_2 \ \ldots \ \xi_n]$$

$\therefore$ φ sobre.

Por lo tanto φ es una correspondencia biunivoca entre $\mathbb{R}^n$ y $\mathbb{R}^{n*}$; ésto nos sugiere introducir la siguiente notación:

A los vectores columna los denotaremos por x:

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} \in \mathbb{R}^n$$

A los vectores renglón los denotaremos por x^t:

$$x^t = [\xi_1 \ \xi_2 \ \ldots \ \xi_n] \in \mathbb{R}^{n*}$$

De acuerdo con esto, al renglón $[0 \ 0 \ \ldots \ 0]$ lo llamaremos 0^t.

Sea $\varphi^{-1} : \mathbb{R}^{n*} \longrightarrow \mathbb{R}^n$ la función inversa de φ :

$$\varphi^{-1}([\xi_1\ \xi_2 \cdots \xi_n]) = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

Es decir, $\varphi^{-1}(x^t) = x \quad \forall\ x^t \in \mathbb{R}^{n*}$.

Podemos ver a φ^{-1} como la función que transforma renglones en columnas.

Si nos permitimos ampliar el significado del término transponer - entendiendo por transposición TANTO la operación que transforma columnas en renglones (φ) COMO la que transforma renglones en columnas (φ^{-1}), entonces podemos pensar en la transposición como la función que transforma columnas en renglones Y renglones en columnas.

Con éste sentido del término transponer, es lícito usar la siguiente notación:

Si $\quad x \in \mathbb{R}^n, \quad \varphi(x) = x^t$
Si $\quad x^t \in \mathbb{R}^{n*}, \quad \varphi^{-1}(x^t) = (x^t)^t$

Entonces tendremos que

$$\varphi^{-1}(x^t) = \varphi^{-1}(\varphi x) = x$$

Pero $\varphi^{-1}(x^t) = (x^t)^t \quad \therefore \quad x = (x^t)^t$

PROPIEDADES DE LA TRANSPOSICION :

1) $(x^t)^t = x \quad \forall\, x \in \mathbb{R}^n$ (Ver la hoja anterior)

2) $(x+y)^t = x^t + y^t \quad \forall\, x, y \in \mathbb{R}^n$

Demostración:

$$x + y = \begin{pmatrix} \xi_1 + \eta_1 \\ \xi_2 + \eta_2 \\ \vdots \\ \xi_n + \eta_n \end{pmatrix}$$

$$\therefore\ (x+y)^t = \left[\xi_1 + \eta_1 \quad \xi_2 + \eta_2 \quad \ldots \quad \xi_n + \eta_n\right]$$

$$= \left[\xi_1 \ \xi_2 \cdots \xi_n\right] + \left[\eta_1 \ \eta_2 \ \ldots \ \eta_n\right]$$

$$= x^t + y^t \ \blacksquare$$

3) $(\alpha x)^t = \alpha x^t \quad \forall\, x \in \mathbb{R}^n,\ \alpha \in \mathbb{R}$

Demostración:

$$\alpha x = \begin{pmatrix} \alpha \xi_1 \\ \alpha \xi_2 \\ \vdots \\ \alpha \xi_n \end{pmatrix}$$

$$\therefore (\alpha x)^t = [\alpha \xi_1 \quad \alpha \xi_2 \quad \dots \quad \alpha \xi_n]$$
$$= \alpha [\xi_1 \, \xi_2 \, \cdots \, \xi_n]$$
$$= \alpha \, x^t$$

■

PRODUCTO PUNTO

DEFINICION

Sean $x, y \in \mathbb{R}^n$. El PRODUCTO PUNTO ó PRODUCTO ESCALAR de x y y es el escalar denotado por $x \cdot y$ dado por:

$$x \cdot y = \xi_1 \eta_1 + \xi_2 \eta_2 + \cdots + \xi_n \eta_n = \sum_{i=1}^{n} \xi_i \eta_i$$

NOTA:

En lo sucesivo siempre que nos refiramos al vector x estaremos pensando en el vector $x = [\xi_i]$, e igualmente para el vector $y = [\eta_i]$.

OBSERVAR que el producto punto de dos vectores es un número real

PROPIEDADES DEL PRODUCTO PUNTO

1. $x \cdot y = y \cdot x \quad \forall\, x, y \in \mathbb{R}^n$

2. $(\lambda x) \cdot y = \lambda (x \cdot y) = x \cdot (\lambda y) \quad \forall\, x, y \in \mathbb{R}^n,\ \lambda \in \mathbb{R}$

3. $(x_1 + x_2) \cdot y = x_1 y + x_2 y \quad \forall\, x_1, x_2, y \in \mathbb{R}^n$

4. $x \cdot (y_1 + y_2) = x \cdot y_1 + x \cdot y_2 \quad \forall\, x, y_1, y_2 \in \mathbb{R}^n$

5. $x \cdot x \geq 0 \quad \forall\, x \in \mathbb{R}^n$
$x \cdot x = 0 \iff x = 0$

Demostración :

1. $$x \cdot y = \sum_{i=1}^{n} \xi_i \eta_i = \sum_{i=1}^{n} \eta_i \xi_i = y \cdot x \;\blacksquare$$

2. $$(\lambda x) \cdot y = \sum_{i=1}^{n} (\lambda \xi_i) \eta_i = \sum_{i=1}^{n} \lambda (\xi_i \eta_i) = \lambda \sum_{i=1}^{n} \xi_i \eta_i$$
$$= \sum_{i=1}^{n} \xi_i (\lambda \eta_i)$$
$$\therefore\ (\lambda x) \cdot y = \lambda (x \cdot y) = x \cdot (\lambda y) \;\blacksquare$$

3. Si $x_1 = [\xi_{i1}]$ $x_2 = [\xi_{i2}]$ entonces
$$(x_1 + x_2) \cdot y = \sum_{i=1}^{n} (\xi_{i1} + \xi_{i2}) \eta_i = \sum_{i=1}^{n} (\xi_{i1} \eta_i + \xi_{i2} \eta_i)$$
$$= \sum_{i=1}^{n} \xi_{i1} \eta_i + \sum_{i=1}^{n} \xi_{i2} \eta_i$$
$$= x_1 \cdot y + x_2 \cdot y \;\blacksquare$$

4. $$\begin{aligned} x\cdot(y_1+y_2) &= (y_1+y_2)\cdot x && \text{(por 1)} \\ &= (y_1\cdot x) + (y_2\cdot x) && \text{(por 3)} \\ &= x\cdot y_1 + x\cdot y_2 && \text{(por 1)} \end{aligned}$$

$\blacksquare$

5. $$x\cdot x = \xi_1^2+\xi_2^2+\cdots+\xi_n^2$$

Como $\xi_i^2 \geq 0 \qquad i=1,2,\ldots,n$

$$\Rightarrow \quad \xi_1^2+\xi_2^2+\cdots+\xi_n^2 \geq 0$$

$$x=0 \quad\Rightarrow\quad x\cdot x = \sum_{i=1}^{n}(0)^2 = 0$$

$$x\cdot x = 0 \quad\Rightarrow\quad \sum_{i=1}^{n}\xi_i^2 = 0$$

$$\Rightarrow \quad \xi_i = 0 \qquad i=1,2,\ldots,n$$

$$\Rightarrow \quad x=0$$

$\square$

EJEMPLOS

1. SI $x=\begin{pmatrix}1\\0\end{pmatrix}$, $y=\begin{pmatrix}1\\1\end{pmatrix}$ $\Rightarrow$ $x\cdot y = 1(1)+(0)(1)=1$

2. SI $x=\begin{pmatrix}1\\0\end{pmatrix}$, $y=\begin{pmatrix}0\\1\end{pmatrix}$ $\Rightarrow$ $x\cdot y = 1(0)+(0)(1)=0$

Observar que, como se puede ver en el ejemplo 2

$$x \cdot y = 0 \not\Rightarrow x = 0 \text{ ó } y = 0$$

DEFINICION

Sea $x \in \mathbb{R}^n$. Definimos la NORMA EUCLIDEANA ó LONGITUD de x, denotada por $\| x \|$, como

$$\| x \| = \sqrt{x \cdot x}$$

PROPIEDADES

1) $\| \alpha x \| = |\alpha| \cdot \| x \|$

Demostración:

$$\| \alpha x \|^2 = \sum_{i=1}^{n} (\alpha \xi_i)^2 = \alpha^2 \cdot \sum_{i=1}^{n} \xi_i^2$$

$$\therefore \; \| \alpha x \| = |\alpha| \sqrt{\sum_{i=1}^{n} \xi_i^2} = |\alpha| \cdot \| x \| \quad \square$$

2) $\|x-y\|^2 = \|x\|^2 + \|y\|^2 - 2x\cdot y$

Demostración :

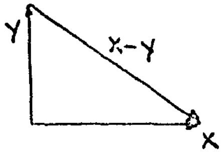
$$\|x-y\|^2 = \sum_{i=1}^{n}(\xi_i-\eta_i)^2 = \sum_{i=1}^{n}\xi_i^2 - \sum_{i=1}^{n}2\xi_i\eta_i + \sum_{i=1}^{n}\eta_i^2$$

$$= \|x\|^2 + \|y\|^2 - 2\sum_{i=1}^{n}\xi_i\eta_i$$

$$= \|x\|^2 + \|y\|^2 - 2x\cdot y \quad \blacksquare$$

3) TEOREMA DE PITAGORAS

$$x\cdot y = 0 \iff \|x-y\|^2 = \|x+y\|^2 = \|x\|^2 + \|y\|^2$$



Demostración :

Por el resultado (2)

$$x\cdot y = 0 \iff \|x-y\|^2 = \|x\|^2 + \|y\|^2$$

De manera análoga se puede ver que

$$x \cdot y = 0 \iff \|x+y\|^2 = \|x\|^2 + \|y\|^2$$

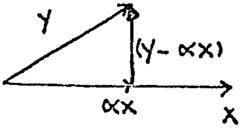
∎

DEFINICION

Sean $x, y \in \mathbb{R}^n$. Decimos que x es ORTOGONAL a y si $x \cdot y = 0$.

Si x es ortogonal a y, escribimos $x \perp y$.

Supongamos ahora que tenemos dos vectores no nulos $x, y \in \mathbb{R}^n$ tales que x no es un múltiplo de y, y y no es un múltiplo de x.



¿ Podemos encontrar siempre un vector que sea ortogonal a x?

Veremos que podemos encontrar $\alpha \in \mathbb{R}$ tal que $(y - \alpha x) \perp x$:

RESULTADO

Sean $x, y \in \mathbb{R}^n$, $x, y \neq 0$ no colineales. Entonces:

$y - \alpha x$ es ortogonal a x $\Leftrightarrow$ $\alpha = \dfrac{x \cdot y}{\|x\|^2}$.

Demostración:

$$
\begin{aligned}
(y - \alpha x) \perp x &\Leftrightarrow x \cdot (y - \alpha x) = 0 \\
&\Leftrightarrow x \cdot y + x \cdot (-\alpha x) = 0 \\
&\Leftrightarrow x \cdot y - \alpha (x \cdot x) = 0 \\
&\Leftrightarrow \alpha = \frac{x \cdot y}{x \cdot x} \quad \text{(esto es posible ya que } x \neq 0\text{)} \\
&\Leftrightarrow \alpha = \frac{x \cdot y}{\|x\|^2}
\end{aligned}
$$

$\blacksquare$

RESULTADO : DESIGUALDAD DE SCHWARTZ

Sean $x, y \in \mathbb{R}^n$. Entonces

$$|x \cdot y| \leq \|x\| \, \|y\|.$$

Demostración :

Si $x = 0$ entonces $|x \cdot y| = \|x\| \, \|y\| = 0$.

Supongamos $x \neq 0$ y sea $\alpha = \dfrac{x \cdot y}{x \cdot x}$

$$\therefore \quad 0 \leq \left\| y - \left(\frac{x\cdot y}{x\cdot x}\right) x \right\|^2 = \left(y - \left(\frac{x\cdot y}{x\cdot x}\right) x \right) \cdot \left(y - \left(\frac{x\cdot y}{x\cdot x}\right) x \right)$$

$$= y\cdot y - 2\,\frac{(x\cdot y)^2}{x\cdot x} + \left(\frac{x\cdot y}{x\cdot x}\right)^2 x\cdot x$$

$$= \frac{(x\cdot x)(y\cdot y) - (x\cdot y)^2}{x\cdot x}$$

$$\Rightarrow \quad 0 \leq (x\cdot x)(y\cdot y) - (x\cdot y)^2$$

$$\Rightarrow \quad (x\cdot y)^2 \leq (x\cdot x)(y\cdot y) = \|x\|^2 \, \|y\|^2$$

$$\Rightarrow \quad |x\cdot y| \leq \|x\| \, \|y\| \qquad \square$$

Como consecuencia de la desigualdad de Schwartz podemos introducir el siguiente concepto

DEFINICION

Sean $x, y \in \mathbb{R}^n$. Definimos

$$\cos\theta = \frac{x\cdot y}{\|x\| \, \|y\|}$$

= coseno del ángulo formado por x y y.

MATRICES

DEFINICION

Tenemos tres formas diferentes y equivalentes de definir una matriz.

1. Convencionalmente definimos una MATRIZ m x n SOBRE IR como un arreglo rectangular de mn números : $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1n}, \alpha_{21}, \alpha_{22}, \ldots, \alpha_{2n}, \ldots, \alpha_{m1}, \alpha_{m2}, \ldots, \alpha_{mn}$, que consta de m renglones y n columnas. Los números que forman el arreglo se llaman ELEMENTOS de la matriz.

Denotaremos a las matrices con letras mayúsculas.

$\therefore$ Si A es una matriz m x n , A es de la forma

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & & \alpha_{mn} \end{bmatrix}$$

donde $\alpha_{ij} \in \mathbb{R}$ $1 \leq i \leq m$, $1 \leq j \leq n$ y

α_{ij} es el elemento de A que está en el i-ésimo renglón y la j-ésima columna de A.

• MATRICES

DEFINICION

Tenemos tres formas diferentes y equivalentes de definir una matriz.

1. Convencionalmente definimos una MATRIZ m x n SOBRE IR como un arreglo rectangular de elementos de IR que consta de m renglones y n columnas. Los números que forman el arreglo se llaman los ELEMENTOS de la matriz.

Denotaremos a las matrices con letras mayúsculas.

∴ si A es una matriz m x n, entonces A es de la forma:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & & \alpha_{mn} \end{bmatrix}$$

donde $\alpha_{ij} \in \mathbb{R}$ $1 \leq i \leq m$, $1 \leq j \leq m$ y

α_{ij} es el elemento de A que está en el i-ésimo renglón y la j-ésima columna.

Escribimos $A = [\alpha_{ij}]$.

2. Podemos definir a la matriz A como un conjunto ordenado de n vectores columna de orden m:

$$A = [a_1 | a_2 | \cdots | a_n]$$

donde

$$a_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix}, \quad a_2 = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix}, \quad \ldots, \quad a_n = \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}$$

Generalmente utilizaremos el índice j para referirnos a las columnas de A.

3. También podemos definir a A como una colección ordenada de m vectores renglón de orden n:

$$A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$$

donde

$$\begin{aligned} r_1^t &= [\alpha_{11} \ \alpha_{12} \ \cdots \ \alpha_{1n}] \\ r_2^t &= [\alpha_{21} \ \alpha_{22} \ \cdots \ \alpha_{2n}] \\ &\vdots \qquad\qquad\qquad \vdots \\ r_m^t &= [\alpha_{m1} \ \alpha_{m2} \ \cdots \ \alpha_{mn}] \end{aligned}$$

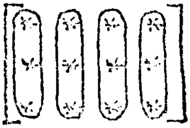
Generalmente utilizaremos el índice i para referirnos a los renglones de A.

Al conjunto de todas las matrices $m \times n$ lo denotaremos por $M(m,n)$

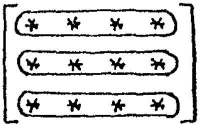
De acuerdo con la definición de matriz que dimos, podemos ver a una matriz como una colección ordenada de :

1. Elementos

2. Columnas



3. Renglones



Es decir, tenemos tres formas distintas de trabajar con una matriz -- escogeremos la representación que más nos convenga según la operación u operaciones a efectuar sobre la matriz

NOTAR que dos matrices m x n son iguales si y solo si sus elementos correspondientes son iguales.

DEFINICION

Una matriz A m x n es CUADRADA si m=n. En este caso decimos que A es de ORDEN n.

Denotaremos al conjunto de todas las matrices cuadradas de orden n por Mn .

DEFINICION

Cualquier matriz cuyos elementos sean todos igual a cero se llamará matriz CERO, y se denotará por el símbolo 0 .

Las operaciones básicas son la suma de dos matrices y la multiplicación de una matriz por un escalar:

DEFINICION

Sean $A, B \in M(m,n)$. La SUMA de A y B es la matriz $C = [\gamma_{ij}]$, denotada por $A+B$ y dada por

$$[\gamma_{ij}] = [\alpha_{ij} + \beta_{ij}]$$

OBSERVACIONES

1. Podríamos haber definido A+B por columnas ó por renglones.

1.a Es decir, si $A = [a_1|a_2|\cdots|a_n]$, $B=[b_1|\cdots b_n]$

$$A+B = [a_1+b_1|a_2+b_2|\cdots|a_n+b_n]$$

Ya que si $C=[c_1|c_2|\cdots|c_n]$ entonces

$$c_j = \begin{pmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{mj} \end{pmatrix} = \begin{pmatrix} \alpha_{1j}+\beta_{1j} \\ \alpha_{2j}+\beta_{2j} \\ \vdots \\ \alpha_{mj}+\beta_{mj} \end{pmatrix} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} + \begin{pmatrix} \beta_{1j} \\ \beta_{2j} \\ \vdots \\ \beta_{mj} \end{pmatrix}$$

$$= a_j + b_j \qquad 1 \leq j \leq n$$

1.b. Análogamente, si

$$A = \begin{bmatrix} r_1^* \\ r_2^* \\ \vdots \\ r_m^* \end{bmatrix} \qquad B = \begin{bmatrix} s_1^* \\ s_2^* \\ \vdots \\ s_m^* \end{bmatrix}$$

Podríamos haber definido

$$A+B = \begin{bmatrix} r_1^* + s_1^* \\ r_2^* + s_2^* \\ \vdots \\ r_m^* + s_m^* \end{bmatrix}$$

2. La suma de dos matrices solamente está definida cuando las dimensiones de ambas son iguales - es decir, cuando tienen el mismo número de renglones y columnas.

DEFINICION

Sea $A \in M(m,n)$ y $\lambda \in \mathbb{R}$. El producto de λ y A, denotado por $\lambda \cdot A$ ó λA, es la matriz $m \times n$ dada por

$$\lambda A = [\lambda \alpha_{ij}]$$

Notar que si $\lambda \neq 0$, entonces A y λA tienen los mismos elementos no nulos.

OBSERVACIONES

Podríamos haber definido λA por columnas ó por renglones :

$$\lambda A = [\lambda a_1 | \lambda a_2 | \cdots | \lambda a_n]$$

$$\lambda A = \begin{bmatrix} \lambda r_1{}^t \\ \lambda r_2{}^t \\ \vdots \\ \lambda r_m{}^t \end{bmatrix}$$

PROPIEDADES

Sean $A, B, C \in M(m,n)$ y $\lambda, \mu \in \mathbb{R}$. Entonces

1. $A+B = B+A$

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

3. $A+0 = A$

4. Para toda $A \in M(m,n)$ existe una matriz $m \times n$, denotada por $-A$ tal que

$$A + (-A) = 0$$

5. $1 \cdot A = A$

6. $\lambda(\mu A) = (\lambda\mu)A$

7. $(\lambda+\mu)A = \lambda A + \mu A$

8. $\lambda(A+B) = \lambda A + \lambda B$

CASOS ESPECIALES DE MATRICES

1. VECTORES COLUMNA

Si A es una matriz $n\times 1$, entonces A es de la forma

$$A = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{n1} \end{bmatrix}$$

que es exactamente la estructura de los miembros de $\mathbb{R}^n$.

No distinguiremos entre una matriz $A \in M(n,1)$ y un vector $a \in \mathbb{R}^n$.

2. VECTORES RENGLON

si A es una matriz $1 \times n$, entonces A es de la forma

$$A = [\alpha_{11} \ \alpha_{12} \cdots \alpha_{1n}]$$

y podemos ver a A como un miembro de $\mathbb{R}^{n*}$.

No distinguiremos entre una matriz $A \in M(1,n)$ y un renglón $a^t \in \mathbb{R}^{n*}$.

3. ESCALARES

si A es una matriz 1×1 entonces

$$A = [\alpha_{11}] \qquad \alpha_{11} \in \mathbb{R}$$

Existe una correspondencia biunívoca entre $\mathbb{R}$ y las matrices 1×1. Abusando de ésto, cuando sea conveniente manejaremos a los escalares como matrices 1×1.

Notación:

Si $A \in M(m,n)$ escribiremos $A \ m \times n$.

PRODUCTOS

PRODUCTO DE UN RENGON POR UN VECTOR

Supongamos que tenemos la siguiente ecuación lineal:

$$\alpha_1 \xi_1 + \alpha_2 \xi_2 + \cdots + \alpha_n \xi_n = \beta \quad \cdots \quad (*)$$

En esta ecuación podemos distinguir entre

a) Los coeficientes de las incógnitas: $\alpha_1, \alpha_2, \ldots, \alpha_n$

b) Las incógnitas: $\xi_1, \xi_2, \ldots, \xi_n$

c) El valor β

Podemos formar, de acuerdo con lo mencionado arriba:

a) La matriz $1 \times n$ de coeficientes

$$r^t = \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_n \end{bmatrix} \quad , \text{ con } \quad r = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

b) El vector cuyas componentes son las incógnitas

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

c) El vector 1×1 cuya componente es β

$$b = (\beta)$$

Observamos que la ecuación (*) es otra forma de escribir

$$r \cdot x = b$$

donde $r \cdot x$ es el producto punto de r y x.

Esto nos sugiere definir la multiplicación de un renglón por un vector como sigue

DEFINICION

Sea $r^t \in \mathbb{R}^{n*}$ y $x \in \mathbb{R}^n$. Entonces el producto de r^t y x, escrito $r^t x$, es el vector $b \in \mathbb{R}^1$ dado por

$$\underset{1\times n}{r^t}\,\underset{n\times 1}{x} = \underset{1\times 1}{b} = r \cdot x$$

EJEMPLO

si $r^t = [1 \ 1 \ 1]$ $x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix}$ entonces

$r^t x = r \cdot x = \xi_1 + \xi_2 + \xi_3$

PROPIEDADES

De las propiedades del producto punto se sigue que:

1. $r^t x = x^t r \quad \forall\, r, x \in \mathbb{R}^n$
2. $(\lambda x^t) y = (\lambda x)^t y = \lambda (x^t y) = x^t (\lambda y) \quad \forall\, \lambda \in \mathbb{R},\ x, y \in \mathbb{R}^n$
3. $(x_1 + x_2)^t y = x_1^t y + x_2^t y \quad \forall\, x_1, x_2, y \in \mathbb{R}^n$
4. $x^t (y_1 + y_2) = x^t y_1 + x^t y_2 \quad \forall\, x, y_1, y_2 \in \mathbb{R}^n$
5. $x^t x \geq 0 \quad \forall\, x \in \mathbb{R}^n$,
   $x^t x = 0 \iff x = 0$

PRODUCTO DE UNA MATRIZ POR UN VECTOR

Ahora, supongamos que tenemos un sistema de m ecuaciones lineales con n incógnitas :

$$\begin{array}{l} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \cdots + \alpha_{1n}\xi_n = \beta_1 \\ \alpha_{21}\xi_1 + \alpha_{22}\xi_2 + \cdots + \alpha_{2n}\xi_n = \beta_2 \\ \quad\vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \cdots + \alpha_{mn}\xi_n = \beta_m \end{array} \qquad \cdots (**)$$

Al igual que en el caso anterior, podemos distinguir entre

a) Los coeficientes α_{ij} $\quad 1 \leq i \leq m \quad 1 \leq j \leq n$

b) Las incógnitas ξ_j $\quad 1 \leq j \leq n$

c) Los valores β_i $\quad 1 \leq i \leq m$

y los podemos agrupar en 3 estructuras :

a) La matriz $A_{m \times n}$ de coeficientes

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}$$

b) El vector $X_{n \times 1}$ de incógnitas :

$$x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$$

c) El vector $b_{m \times 1}$ de valores

$$b = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$$

Queremos expresar el sistema de ecuaciones (**) como una multiplicación $Ax=b$, para lo cual hemos de definir el producto de una matriz por un vector de una manera que respete el producto de un renglón por un vector.

OBSERVAR que en todo sistema de ecuaciones el número de columnas de la matriz de coeficientes es igual al número de incógnitas.

Habíamos visto anteriormente que una matriz puede tener las representaciones

$$A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix} \qquad \text{donde} \quad r_i^t = [\alpha_{i1} \ \alpha_{i2} \ \cdots \ \alpha_{in}]$$
$$1 \le i \le m$$

$$A = [a_1 | a_2 | \cdots | a_n] \qquad \text{donde} \qquad a_j = \begin{bmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{bmatrix} \qquad 1 \leq j \leq n$$

De acuerdo con ésto, podemos reescribir de dos formas al sistema (**)

1)

$$\begin{aligned} r_1^t x &= \beta_1 \\ r_2^t x &= \beta_2 \\ &\vdots \\ r_m^t x &= \beta_m \end{aligned} \qquad \text{ó} \qquad \text{sea}$$

$$\begin{pmatrix} r_1^t x \\ r_2^t x \\ \vdots \\ r_m^t x \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$$

Por ejemplo, el sistema

$$\begin{aligned} \xi_1 + \xi_2 + \xi_3 &= 1 \\ 2\xi_1 - \xi_2 - \xi_3 &= 0 \end{aligned}$$

se puede escribir como

$$\begin{bmatrix} [1 \quad 1 \quad 1]\, x \\ [2 \quad -1 \quad -1]\, x \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

2) $$\begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix} \xi_1 + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix} \xi_2 + \cdots + \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{pmatrix} \xi_n = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_m \end{pmatrix}$$

es decir

$$a_1 \xi_1 + a_2 \xi_2 + \cdots + a_n \xi_n = b$$

Por ejemplo, si tenemos

$$\begin{aligned} \xi_1 + \xi_2 + \xi_3 &= 1 \\ 2\xi_1 - \xi_2 - \xi_3 &= 0 \end{aligned}$$

$$\Rightarrow \begin{pmatrix} \xi_1 + \xi_2 + \xi_3 \\ 2\xi_1 - \xi_2 - \xi_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} 1 \\ 2 \end{pmatrix} \xi_1 + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xi_2 + \begin{pmatrix} 1 \\ -1 \end{pmatrix} \xi_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Esto nos sugiere definir la multiplicación de una matriz por un vector :

DEFINICION 1.

sea $A \in M(m,n)$ y $x \in \mathbb{R}^n$. Entonces la multiplicación de A y x, denotada Ax, es el vector $b \in \mathbb{R}^m$ dado por

$$Ax = \begin{pmatrix} r_1^t\ x \\ r_2^t\ x \\ \vdots \\ r_m^t\ x \end{pmatrix}$$

donde $r_i^t = [\alpha_{i1}\ \alpha_{i2}\ \cdots\ \alpha_{in}]$ $1 \leq i \leq m$ es el i-ésimo renglón de A.

DEFINICION 2.

Sea $A \in M(m,n)$, $x \in \mathbb{R}^n$, entonces Ax es el vector $m \times 1$ dado por

$$Ax = a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n$$

donde $a_j = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$, $1 \leq j \leq n$ es la j-ésima columna de A.

OBSERVACION

Estas dos definiciones son equivalentes, ya que :

$$b_1 = \begin{pmatrix} r_1^t x \\ r_2^t x \\ \vdots \\ r_m^t x \end{pmatrix} \quad y \quad b_2 = a_1\bar{e}_1 + a_2\bar{e}_2 + \cdots + a_n\bar{e}_n$$

entonces

$$b_1 = \begin{pmatrix} \alpha_{11}\bar{e}_1 + \alpha_{12}\bar{e}_2 + \cdots + \alpha_{1n}\bar{e}_n \\ \alpha_{21}\bar{e}_1 + \alpha_{22}\bar{e}_2 + \cdots + \alpha_{2n}\bar{e}_n \\ \vdots \\ \alpha_{m1}\bar{e}_1 + \alpha_{m2}\bar{e}_2 + \cdots + \alpha_{mn}\bar{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{11}\bar{e}_1 \\ \alpha_{21}\bar{e}_1 \\ \vdots \\ \alpha_{m1}\bar{e}_1 \end{pmatrix} + \begin{pmatrix} \alpha_{12}\bar{e}_2 \\ \alpha_{22}\bar{e}_2 \\ \vdots \\ \alpha_{m2}\bar{e}_2 \end{pmatrix} + \cdots + \begin{pmatrix} \alpha_{1n}\bar{e}_n \\ \alpha_{2n}\bar{e}_n \\ \vdots \\ \alpha_{mn}\bar{e}_n \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \vdots \\ \alpha_{m1} \end{pmatrix}\bar{e}_1 + \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix}\bar{e}_2 + \cdots + \begin{pmatrix} \alpha_{1n} \\ \alpha_{2n} \\ \vdots \\ \alpha_{mn} \end{pmatrix}\bar{e}_n$$

$$= a_1\bar{e}_1 + a_2\bar{e}_2 + \cdots + a_n\bar{e}_n$$

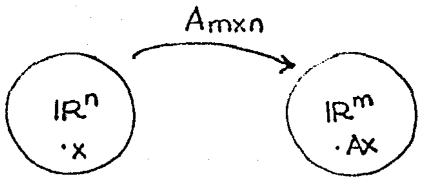
$$= b_2$$

OBSERVACION

No hay ninguna restricción sobre m (= número de renglones de A), pero es necesario que la dimensión del vector x y el número de columnas de A (n) sean iguales para que los productos $r_i^t x$ estén bien definidos.

Ahora, si nos fijamos en que al multiplicar una matriz $m x n$ por un vector $n x 1$ obtenemos un vector $m x 1$, esto nos permite ver a la multiplicación por A_{mxn} como una función que a cada $x \in \mathbb{R}^n$ le asocia una $b \in \mathbb{R}^m$:

$$A_{mxn} : \mathbb{R}^n \to \mathbb{R}^m$$



Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \quad , \quad y \quad x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad \text{entonces}$$

$$Ax = (1)\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + (0)\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\therefore \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xmapsto{A} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

PROPIEDADES

1. $A(x_1 + x_2) = Ax_1 + Ax_2 \qquad \forall\, A_{m \times n}\,,\ x_1, x_2 \in \mathbb{R}^n$

2. $A(\lambda x) = \lambda\,(Ax) \qquad \forall\, A_{m \times n}\,,\ x \in \mathbb{R}^n,\ \lambda \in \mathbb{R}$

3. $(A+B)\,x = Ax + Bx \qquad \forall\, A_{m \times n}, B_{m \times n},\ x \in \mathbb{R}^n$

Demostración :

1.
$$A(x_1+x_2) = \begin{pmatrix} r_1^t\,(x_1+x_2) \\ r_2^t\,(x_1+x_2) \\ \vdots \\ r_m^t\,(x_1+x_2) \end{pmatrix} = \begin{pmatrix} r_1^t x_1 + r_1^t x_2 \\ r_2^t x_1 + r_2^t x_2 \\ \vdots \\ r_m^t x_1 + r_m^t x_2 \end{pmatrix}$$

$$= \begin{pmatrix} r_1^t x_1 \\ r_2^t x_1 \\ \vdots \\ r_m^t x_1 \end{pmatrix} + \begin{pmatrix} r_1^t x_2 \\ r_2^t x_2 \\ \vdots \\ r_m^t x_2 \end{pmatrix} = Ax_1 + Ax_2 \quad \blacksquare$$

2.
$$A(\lambda x) = \begin{pmatrix} r_1^t(\lambda x) \\ r_2^t(\lambda x) \\ \vdots \\ r_m^t(\lambda x) \end{pmatrix} = \begin{pmatrix} \lambda(r_1^t x) \\ \lambda(r_2^t x) \\ \vdots \\ \lambda(r_m^t x) \end{pmatrix} = \lambda \begin{pmatrix} r_1^t x \\ r_2^t x \\ \vdots \\ r_m^t x \end{pmatrix} = \lambda\,(Ax)$$

si $B = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_m^t \end{bmatrix}$ entonces

$$(A+B)X = \begin{pmatrix} (r_1^t + s_1^t)X \\ (r_2^t + s_2^t)X \\ \vdots \\ (r_m^t + s_m^t)X \end{pmatrix} = \begin{pmatrix} r_1^t X + s_1^t X \\ r_2^t X + s_2^t X \\ \vdots \\ r_m^t X + s_m^t X \end{pmatrix}$$

$$= \begin{pmatrix} r_1^t X \\ r_2^t X \\ \vdots \\ r_m^t X \end{pmatrix} + \begin{pmatrix} s_1^t X \\ s_2^t X \\ \vdots \\ s_m^t X \end{pmatrix} = AX + BX \quad \blacksquare$$

OBSERVACION IMPORTANTE

¿ qué más podemos decir acerca de $b = AX$?

si nos fijamos en que

$$b = a_1 \bar{\varepsilon}_1 + a_2 \bar{\varepsilon}_2 + \cdots + a_n \bar{\varepsilon}_n$$

veremos que b es una "combinación"⁽¹⁾ de las columnas de A

$\underset{m\times 1}{b} = \underset{m\times n}{A}\underset{n\times 1}{X}$ SI Y SOLO SI b es una COMBINACION DE LAS COLUMNAS DE A .

(1) Este tipo de "combinaciones" se llaman combinaciones lineales. En el sig. capítulo introduciremos formalmente el concepto, y analizaremos sus propiedades.

Uno de los problemas más importantes del algebra lineal es poder contestar la siguiente pregunta

¿ cuándo es $b_{m\times 1}$ una combinación de $a_1, a_2, \cdots, a_n \in \mathbb{R}^m$?

Esta pregunta la podemos plantear así:

¿ Dado $b \in \mathbb{R}^m$, cuándo existe $x \in \mathbb{R}^n$ tal que $\underset{m\times n}{A}x = b$?

De aquí la importancia de entender bien la interpretación sobre la multiplicación de una matriz por un vector.

MULTIPLICACION DE UN RENGLON POR UNA MATRIZ

Así como definimos el producto $A_{m\times n} X_{n\times 1} = b_{m\times 1}$ ahora queremos definir el producto de un renglón por una matriz

$$Y^t_{1\times m} \, A_{m\times n} = ?$$

DEFINICION 1.

Sea $A \in M(m,n)$ y $y^t \in \mathbb{R}^{m*}$. Entonces el producto de y^t por A, denotado por y^tA, es el renglón $1 \times n$ dado por

$$y^t A = [y^t a_1 \quad y^t a_2 \quad \ldots \quad y^t a_n]$$

donde $a_j = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix}$ es la j-ésima columna de A. $1 \leq j \leq n$

DEFINICION 2.

Sea $A \in M(m,n)$, $y^t \in \mathbb{R}^{m*}$. Entonces y^tA es el renglón de $\mathbb{R}^{n*}$ dado por

$$y^t A = \eta_1 r_1^t + \eta_2 r_2^t + \cdots + \eta_m r_m^t$$

donde $r_i^t = [\alpha_{i1} \ \alpha_{i2} \ \cdots \ \alpha_{in}]$ $\qquad 1 \leq i \leq m$

es el i-ésimo renglón de A.

OBSERVACION

Estas dos definiciones son equivalentes, ya que si

$$d_1{}^t = \eta_1 r_1{}^t + \eta_2 r_2{}^t + \cdots + \eta_m r_m{}^t$$ y

$$d_2{}^t = [y^t a_1 \quad y^t a_2 \quad \cdots \quad y^t a_n]$$ entonces

$$d_1{}^t = \eta_1 [\alpha_{11}\ \alpha_{12} \cdots \alpha_{1n}] + \eta_2 [\alpha_{21}\ \alpha_{22} \cdots \alpha_{2n}] + \cdots + \eta_m [\alpha_{m1}\ \alpha_{m2} \cdots \alpha_{mn}]$$

$$= \left[\sum_{i=1}^{m} \eta_i \alpha_{i1} \quad \sum_{i=1}^{m} \eta_i \alpha_{i2} \quad \cdots \quad \sum_{i=1}^{m} \eta_i \alpha_{in} \right]$$

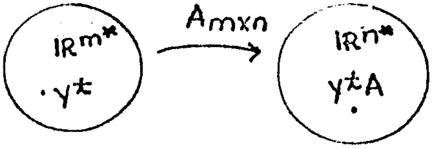
$$= [y^t a_1 \quad y^t a_2 \quad \cdots \quad y^t a_n] = d_2^t$$

OBSERVACION

No hay ninguna restricción sobre n (= número de columnas de A), pero es necesario que la dimensión del renglón y^t y m (= número de renglones de A) sean iguales, para que los productos $y^t a_j$ estén definidos.

Podemos ver la multiplicación de $y^t{}_{1\times m}$ por $A_{m\times n}$ como una función que actua sobre y^t, y que le asocia un $d^t \in \mathbb{R}^{n*}$

$$A_{m\times n} : \mathbb{R}^{m*} \longrightarrow \mathbb{R}^{n*}$$



Por ejemplo, si

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad y^t = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \quad \text{entonces}$$

$$\begin{aligned} y^tA &= (1)\begin{bmatrix} 1 & 2 \end{bmatrix} + (0)\begin{bmatrix} 0 & 1 \end{bmatrix} + (-1)\begin{bmatrix} 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 \end{bmatrix} + \begin{bmatrix} -1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 2 \end{bmatrix} \end{aligned}$$

$$\therefore \begin{bmatrix} 1 & 0 & -1 \end{bmatrix} \longmapsto \begin{bmatrix} 0 & 2 \end{bmatrix}$$

PROPIEDADES

1. $(y_1+y_2)^t A = y_1^tA + y_2^t A \qquad \forall\, y_1, y_2 \in \mathbb{R}^m,\ A_{m\times n}$
2. $(\lambda y)^t A = \lambda\,(y^t A) \qquad \forall\, y \in \mathbb{R}^m,\ A_{m\times n},\ \lambda \in \mathbb{R}$
3. $y^t(A+B) = y^tA + y^tB \qquad \forall\, y \in \mathbb{R}^m,\ A, B \in M(m,n)$

Demostración:

1. $$(y_1+y_2)^t A = [(y_1+y_2)^t a_1 \quad (y_1+y_2)^t a_2 \quad \cdots \quad (y_1+y_2)^t a_n]$$
$$= [y_1^t a_1 + y_2^t a_1, \quad y_1^t a_2 + y_2^t a_2 \quad \cdots y_1^t a_n + y_2^t a_n]$$
$$= [y_1^t a_1 \quad y_1^t a_2 \cdots y_1^t a_n] + [y_2^t a_1 \quad y_2^t a_2 \cdots y_2^t a_n]$$
$$= y_1^t A + y_2^t A \ \blacksquare$$

2. $$(\lambda y)^t A = (\lambda y^t) A$$
$$= [(\lambda y^t) a_1 \quad (\lambda y^t) a_2 \quad \cdots \quad (\lambda y^t) a_n]$$
$$= [\lambda (y^t a_1) \quad \lambda (y^t a_2) \quad \cdots \quad \lambda (y^t a_n)]$$
$$= \lambda [y^t a_1 \quad y^t a_2 \quad \cdots \quad y^t a_n]$$
$$= \lambda (y^t A) \ \blacksquare$$

3. $$y^t (A+B) = [y^t (a_1+b_1) \quad y^t (a_2+b_2) \quad \cdots \quad y^t (a_n+b_n)]$$
$$= [y^t a_1 + y^t b_1 \quad y^t a_2 + y^t b_2 \quad \cdots \quad y^t a_n + y^t b_n]$$
$$= [y^t a_1 \quad y^t a_2 \cdots y^t a_n] + [y^t b_1 \quad y^t b_2 \cdots y^t b_n]$$
$$= y^t A + y^t B \ \blacksquare$$

OBSERVACION IMPORTANTE

Por definición

$$d^t = \eta_1 r_1^t + \eta_2 r_2^t + \cdots + \eta_m r_m^t$$

es decir :

$d^t = y^t A$ SI Y SOLO SI d^t ES UNA COMBINACION[(1)] DE LOS RENGLONES DE A.

Si nos planteamos el problema :

¿ cuándo es $d^t_{1\times n}$ una combinación de $r_1^t, r_2^t, \ldots, r_m^t \in \mathbb{R}^{n*}$?

Entonces una forma equivalente de hacernos esa pregunta sería:

¿ Dado $d^t \in \mathbb{R}^{n*}$, cuándo existe $y^t \in \mathbb{R}^{m*}$ tal que $y^t A = d^t$

donde $A = \begin{bmatrix} r_1^t \\ \vdots \\ r_m^t \end{bmatrix}$?

Este problema es análogo al problema que nos planteamos sobre las columnas de A.

---

(1) Ver nota aclaratoria pag. I.44

MULTIPLICACION DE MATRICES

Supongamos que tenemos dos matrices $A_{m \times n}$, $B_{n \times k}$.

¿Cómo definimos el producto AB de modo que respete los productos que hemos definido anteriormente?

$$\text{si} \quad A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix} = [a_1 | a_2 | \cdots a_n]$$

$$B = [b_1 | b_2 | \cdots | b_k]$$

$$\therefore \quad AB = A[b_1 | b_2 | \cdots b_k]$$

La multiplicación de una matriz por un vector nos llevaría a definir

$$AB = [Ab_1 | Ab_2 | \cdots | Ab_k]$$

Por otro lado

$$AB = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix} B$$

La multiplicación de un renglón por una matriz nos conduce a definir

$$AB = \begin{bmatrix} r_1^t B \\ r_2^t B \\ \vdots \\ r_m^t B \end{bmatrix}$$

O también podríamos definir AB elemento a elemento :

$$AB = \begin{bmatrix} r_1^t b_1 & r_1^t b_2 & \cdots & r_1^t b_K \\ r_2^t b_1 & r_2^t b_2 & \cdots & r_2^t b_K \\ \vdots & \vdots & & \vdots \\ r_m^t b_1 & r_m^t b_2 & \cdots & r_m^t b_K \end{bmatrix}$$

Veremos que podemos definir AB por cualquiera de estas 3 formas ; ya que

$$[Ab_1 | Ab_2 | \cdots | Ab_K] = \begin{bmatrix} r_1^t b_1 & r_1^t b_2 & \cdots & r_1^t b_K \\ r_2^t b_1 & r_2^t b_2 & \cdots & r_2^t b_K \\ \vdots & \vdots & & \vdots \\ r_m^t b_1 & r_m^t b_2 \cdots & & r_m^t b_K \end{bmatrix}$$

$$= \begin{bmatrix} r_1^t B \\ r_2^t B \\ \vdots \\ r_m^t B \end{bmatrix}$$

Tenemos las tres definiciones del producto de dos matrices

DEFINICION

Sea $A \in M(m,n)$, $B \in M(n,K)$. Entonces el producto de A y B, denotado por AB, es la matriz $C = [\gamma_{ij}] \in M(m,K)$ dada por

1. $AB = [Ab_1 \mid Ab_2 \mid \cdots \mid Ab_K]$

2. $$AB = \begin{bmatrix} r_1^t B \\ r_2^t B \\ \vdots \\ r_m^t B \end{bmatrix}$$

3. $$AB = \begin{bmatrix} r_1^t b_1 & r_1^t b_2 & \cdots & r_1^t b_K \\ r_2^t b_1 & r_2^t b_2 & \cdots & r_2^t b_K \\ \vdots & & & \vdots \\ r_m^t b_1 & r_m^t b_2 & \cdots & r_m^t b_K \end{bmatrix}$$

es decir $\gamma_{ij} = r_i^t b_j \quad 1 \le i \le m, \; 1 \le j \le K$

Como acabamos de ver 1., 2. y 3. definen a la misma matriz:

$$A_{m\times n}\, B_{n\times K} = C_{m\times K}$$

OBSERVACIONES

1. Cada columna de AB es el producto de una matriz y una columna:

columna j-ésima de AB = A · (j-ésima columna de B)

2. Cada renglón de AB es el producto de un renglón y una matriz:

i-ésimo renglón de AB = (i-ésimo renglón de A) · B

3. Cada entrada de AB es el producto de un renglón y una columna:

γ_{ij} = (i-ésimo renglón de A) · (j-ésima col. de B).

Gráficamente:

1. $A_{2\times 3}\begin{bmatrix} * & \boxed{\begin{matrix} * \\ * \\ * \end{matrix}} & * \\ * & & * \\ * & & * \end{bmatrix} = \begin{bmatrix} * & \boxed{\begin{matrix} * \\ * \end{matrix}} & * \\ * & & * \end{bmatrix}$

2. $\begin{bmatrix} * & * & * \\ \multicolumn{3}{c}{\boxed{* \quad * \quad *}} \end{bmatrix} B_{3\times 3} = \begin{bmatrix} * & * & * \\ \multicolumn{3}{c}{\boxed{* \quad * \quad *}} \end{bmatrix}$

3. $\begin{bmatrix} * & * & * \\ \multicolumn{3}{c}{\boxed{* \quad * \quad *}} \end{bmatrix} \begin{bmatrix} * & \boxed{\begin{matrix} * \\ * \\ * \end{matrix}} & * \\ * & & * \\ * & & * \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & \boxed{*} & * \end{bmatrix}$

OBSERVACION

Para que AB esté definido necesitamos que el número de columnas de A sea igual al número de renglones de B.

PROPIEDADES

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

2. $A(B+C) = AB+AC$
   $(B+C)D = BD+CD$

3. $AB \neq BA$ en general:

La conmutatividad puede fallar por varias razones

Si A es $m \times n$ y B es $n \times K$, con $K \neq m$, entonces AB está definido pero BA no está definido $\therefore$ $AB \neq BA$

Si A es $m \times n$ y B es $n \times m$, con $m \neq n$, entonces tanto AB como BA están definidos, pero no pueden ser iguales ya que AB es $m \times m$ y BA es $n \times n$, $n \neq m$

Si tanto A como B son $n \times n$, entonces AB y BA son de orden n, pero aún en este caso la conmutatividad puede fallar, como en el sig. ejemplo:

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix};$$

Nos podríamos preguntar para qué nos sirve haber definido el producto de dos matrices. Una forma fructífera de considerar la multiplicación de matrices es viendo a una de las matrices como una función que opera sobre la otra matriz:

1. Tenemos $AB = [Ab_1 | Ab_2 | \cdots | Ab_K]$

Ya vimos que Ab_j – la j-ésima columna de AB – es una combinación de las columnas de A

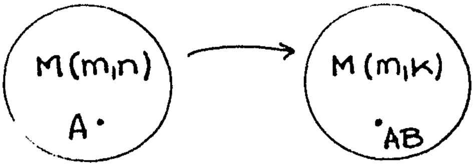
$$Ab_j = a_1\beta_{1j} + a_2\beta_{2j} + \cdots + a_n\beta_{nj} \qquad 1 \leq j \leq K$$

Es decir, LAS COLUMNAS DE AB SON COMBINACIONES DE LAS COLUMNAS DE A.

Podemos pensar en que B está operando sobre A y transformando sus columnas:

$$B_{n\times K} : M(m,n) \longrightarrow M(m,K)$$

$$A_{m\times n} \xmapsto{B_{n\times K}} (AB)_{m\times K}$$



A veces necesitaremos realizar algunas operaciones sobre las columnas de A, ó desearemos transformar a A en otra matriz con la cual sea más fácil trabajar; eso lo podremos lograr escogiendo la matriz $B_{n\times n}$ adecuada y multiplicando (posmultiplicando) por B.

EJEMPLOS . Sea $\underset{m\times n}{A} = [a_1|\cdots|a_n]$

1a. Queremos $AB = [\lambda_1 a_1|\lambda_2 a_2|\cdots|\lambda_n a_n]$

$$B = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Ya que

$$\begin{aligned} Ab_1 &= \lambda_1 a_1 + (0)(a_2) + \cdots + 0\cdot a_n \\ Ab_2 &= 0\cdot a_1 + \lambda_2\cdot a_2 + \cdots + 0\cdot a_n \\ &\vdots \\ Ab_n &= 0\cdot a_1 + 0\cdot a_2 + \cdots + \lambda_n a_n \end{aligned}$$

1b. Queremos $AB = [a_2|a_1|a_3|\cdots|a_n]$, es decir, si queremos cambiar las columnas uno y dos de A entonces

$$B = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

ya que
$$\begin{aligned} Ab_1 &= 0\cdot a_1 + 1\cdot a_2 + 0\cdot a_3 + \cdots + 0\cdot a_n \\ Ab_2 &= 1\cdot a_1 + 0\cdot a_2 + \cdots + 0\cdot a_n \\ Ab_j &= 0\cdot a_1 + \cdots + 1\cdot a_j + \cdots + 0\cdot a_n \\ & \qquad\qquad 3 \leq j \leq n \end{aligned}$$

2. Tenemos

$$AB = \begin{bmatrix} r_1^t B \\ r_2^t B \\ \vdots \\ r_m^t B \end{bmatrix}$$

Habíamos visto que $r_i^t B$ es una combinación de los renglones de B, $1 \leq i \leq m$:

$$r_i^t B = \alpha_{i1} s_1^t + \alpha_{i2} s_2^t + \cdots + \alpha_{in} s_n^t$$

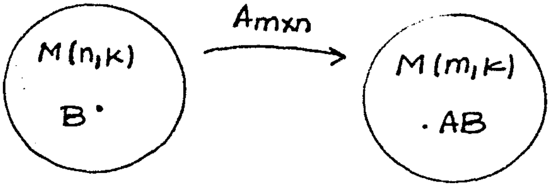
donde $B = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_n^t \end{bmatrix}$

Es decir, LOS RENGLONES DE AB SON COMBINACIONES DE LOS RENGLONES DE B

Podemos pensar que A está operando sobre B:

$$\underset{m\times n}{A}: M(n,K) \longrightarrow M(m,K)$$

$$B_{n\times K} \xmapsto{A_{m\times n}} (AB)_{m\times K}$$



A veces necesitaremos realizar operaciones sobre los renglones de una matriz, por ejemplo de la matriz B, ó transformar una matriz en otra matriz que tenga propiedades deseables, y eso lo podremos lograr multiplicando (premultiplicando) por la matriz, digamos $A_{n\times n}$, adecuada.

EJEMPLOS

sea $B_{n \times K} = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_n^t \end{bmatrix}$

2a. Dar A tal que

$$AB = \begin{bmatrix} s_1^t \\ \lambda_1 s_1^t + \lambda_2 s_2^t \\ s_3^t \\ \vdots \\ s_n^t \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \lambda_1 & \lambda_2 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

2b. Dar A tal que

$$AB = \begin{bmatrix} s_2^t \\ s_1^t \\ s_3^t \\ \vdots \\ s_n^t \end{bmatrix} ; \quad A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

2c Dar A tal que :
i-ésimo renglón de AB = i-ésimo renglón de B, $i \neq j$
y j-ésimo renglón de AB =
λ(i-ésimo de B) + (j-ésimo renglón de B)

$$A = \begin{array}{c|c} \begin{array}{ccccc} & i & & j & \end{array} & \\ \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ \cdots & 1 & \cdots & 0 & \\ & \vdots & & & \\ \cdots & \lambda & \cdots & 1 & \\ & & & & \ddots \quad 1 \end{bmatrix} & \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \end{array} \end{array}$$

Por ejemplo

$$\text{Sea} \quad B = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & 2 \end{bmatrix}.$$

Entonces, si queremos sumar un múltiplo de [1 2 1] al renglón [2 0 1], premultiplicamos por la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

En particular, podemos escoger λ de manera que $\lambda[1\ 2\ 1] + [2\ 0\ 1] = [0\ *\ *]$

Si escogemos $\lambda = -2$ entonces

$$\lambda [1 \quad 2 \quad 1] + [2 \quad 0 \quad 1] = [-2 \quad -4 \quad -2] + [2 \quad 0 \quad 1]$$

$$= [0 \quad -4 \quad -1]$$

$$\therefore AB = \begin{bmatrix} 1 & 2 & 1 \\ 0 & -4 & -1 \\ 1 & 0 & 2 \end{bmatrix}$$

Más adelante veremos que necesitaremos efectuar este tipo de operación, la cual es muy importante.

CASOS ESPECIALES DEL PRODUCTO DE MATRICES

1. Sea $x^t \in M(1,n)$, $y \in M(n,1)$. Entonces

$x^t y \in M(1,1) \sim \mathbb{R}$

$x^t y = x \cdot y$

O sea, que el producto escalar es un caso particular del producto de matrices.

2. Sean $a \in IR^m$, $b \in IR^n$. Entonces el producto ab^t está definido y ab^t es una matriz $m \times n$. Este tipo de producto es muy importante.

OBSERVAR que

$$ab^t = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} b^t = \begin{bmatrix} \alpha_1 b^t \\ \alpha_2 b^t \\ \vdots \\ \alpha_n b^t \end{bmatrix}$$

Es decir, CADA RENGLON DE ab^t ES UN MULTIPLO DE b^t.

Por otra parte,

$$ab^t = a \cdot [\beta_1 \ \beta_2 \ \ldots \ \beta_n] = [\beta_1 a \ \ \beta_2 a \ \cdots \ \beta_n a]$$

Es decir, CADA COLUMNA DE ab^t ES UN MULTIPLO DEL VECTOR a.

Ejemplo

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

Por medio de productos de este tipo vamos a expresar cada matriz como una suma de matrices.

Denotaremos por e_i al vector cuyas componentes valen todas cero, excepto la i-ésima componente, cuyo valor es 1:

$$e_i = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{lugar } i$$

Sean $e_i \in \mathbb{R}^m$ y $e_j \in \mathbb{R}^n$. ¿Qué forma tiene la matriz $e_i e_j^t$?

$$e_i e_j^t = \begin{array}{cc} & \begin{bmatrix} 0 \\ \vdots \\ \end{bmatrix} \\ i & \begin{bmatrix} 1 \end{bmatrix} \\ & \begin{bmatrix} \vdots \\ 0 \end{bmatrix} \end{array} \begin{bmatrix} 0 \cdots \overset{j}{1} \cdots 0 \end{bmatrix} = \begin{array}{c} \\ \\ \\ \end{array} \begin{bmatrix} 0 & 0 & \cdots & \overset{j}{0} & \cdots & 0 \\ \vdots & \vdots & & & & \\ 0 & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & & & \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix} \leftarrow i$$

$e_i e_j^t$ es la matriz $m \times n$ cuyos elementos todos valen cero, excepto la componente que está en el renglón i-ésimo y la columna j-ésima, cuyo valor es 1.

Denotaremos a $e_i e_j^t$ por

$$E_{ij} = e_i e_j^t$$

Ahora expresaremos a la matriz A como suma de matrices de 3 formas distintas:

1. $\underset{m\times n}{A} = [a_1|a_2|\cdots|a_n]$

$$A = [a_1|0|\cdots|0] + [0|a_2|0|\cdots 0] + \cdots + [0|\cdots|0|a_n]$$

Ahora

$$\begin{aligned} [a_1|0|\cdots|0] &= a_1 e_1^t \\ [0|a_2|\cdots|0] &= a_2 e_2^t \\ &\vdots \\ [0|\cdots|0|a_n] &= a_n e_n^t \end{aligned}$$

$$\therefore \quad A = a_1 e_1^t + a_2 e_2^t + \cdots + a_n e_n^t$$

$$\therefore \quad \boxed{A = \sum_{j=1}^{n} a_j e_j^t}$$

Nota: $a_j \in \mathbb{R}^m$, $e_j^t \in \mathbb{R}^{n*}$ $\therefore$ $\underset{m\times 1}{a_j}\,\underset{1\times n}{e_j^t} \in \underset{m\times n}{M(m,n)}$

2. $$A_{m\times n} = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$$

$$A = \begin{bmatrix} r_1^t \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ r_2^t \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r_m^t \end{bmatrix}$$

Pero

$$\begin{bmatrix} r_1^t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_1 r_1^t \; ; \quad \begin{bmatrix} 0 \\ r_2 \\ \vdots \\ 0 \end{bmatrix} = e_2 r_2^t \; ; \; \cdots \; ; \quad \begin{bmatrix} 0 \\ \vdots \\ 0 \\ r_m^t \end{bmatrix} = e_m r_m^t$$

donde $e_i \in \mathbb{R}^m$, $r_i^t \in \mathbb{R}^{n*}$ y

$$\therefore \underset{m\times 1}{e_i} \underset{1\times n}{r_i^t} \in M(m,n)$$

$$\therefore A = e_1 r_1^t + e_2 r_2^t + \cdots + e_m r_m^t$$

$$\boxed{A = \sum_{i=1}^{m} e_i r_i^t}$$

3.

$$A = \begin{bmatrix} \alpha_{11} & 0 \cdots & 0 \\ 0 & 0 \cdots & 0 \\ & \vdots & \\ 0 & 0 \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \alpha_{12} \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & 0 & \alpha_{1n} \\ 0 & \cdots & 0 & 0 \\ & \vdots & & \\ 0 & \cdots & & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} 0 & \cdots & 0 \\ \alpha_{21} & 0 \cdots & 0 \\ & \vdots & \\ 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \alpha_{22} \cdots & 0 \\ & \vdots & \\ 0 & \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & \alpha_{2n} \\ & \vdots & \\ 0 & \cdots & 0 \end{bmatrix}$$

$$+ \cdots + \begin{bmatrix} 0 & \cdots & 0 \\ & \vdots & \\ 0 & \cdots & 0 \\ \alpha_{m1} & \cdots & 0 \end{bmatrix} + \begin{bmatrix} 0 & \vdots & 0 \\ 0 & \cdots & 0 \\ 0 & \alpha_{m2} \cdots & 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 & \vdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots 0 & \alpha_{mn} \end{bmatrix}$$

$$\begin{aligned} = {} & \alpha_{11} E_{11} + \alpha_{12} E_{12} + \cdots + \alpha_{1n} E_{1n} \\ & + \alpha_{21} E_{21} + \alpha_{22} E_{22} + \cdots + \alpha_{2n} E_{2n} \\ & + \cdots + \alpha_{m1} E_{m1} + \alpha_{m2} E_{m2} + \cdots + \alpha_{mn} E_{mn} \end{aligned}$$

$$A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} E_{ij}$$

$$\boxed{A = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} e_i e_j^t}$$

donde $e_i \in \mathbb{R}^m$, $e_j \in \mathbb{R}^n$

OBSERVACION

$$e_i^t e_j = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

Expresaremos el producto de dos matrices como una suma de matrices:

Sea $A_{m \times n} = [a_1 | a_2 | \cdots | a_n]$

$$B_{n \times k} = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_n^t \end{bmatrix}$$

$$\therefore \quad A = \sum_{j=1}^{n} a_j e_j^t \qquad B = \sum_{i=1}^{n} e_i s_i^t$$

$$\therefore \quad AB = \left(\sum_{j=1}^{n} a_j e_j^t \right) \left(\sum_{i=1}^{n} e_i s_i^t \right)$$

$$= \sum_{j=1}^{n} a_j \left(\sum_{i=1}^{n} e_j^t e_i s_i^t \right) = \sum_{j=1}^{n} a_j \left(\sum_{i=1}^{n} (e_j^t e_i) s_i^t \right)$$

$$= \sum_{j=1}^{n} a_j s_j^t$$

$$\therefore \quad \boxed{AB = \sum_{j=1}^{n} a_j s_j^t}$$

TRANSPOSICION DE MATRICES

Recordemos que si $x \in \mathbb{R}^n$, entonces al TRANSPONER x obtenemos el renglón $x^t \in \mathbb{R}^{n*}$

$$x_{n\times 1} \longmapsto x^t_{1\times n}$$

y que si $x^t \in \mathbb{R}^{n*}$, entonces su transpuesto es $x \in \mathbb{R}^n$:

$$x^t_{1\times n} \longmapsto x_{n\times 1}$$

Nos proponemos extender el concepto de transposición de vectores, al de transposición de matrices.

Lo más natural es pensar en la transposición como la función que transforma los renglones de A en columnas, y las columnas de A en renglones.

DEFINICION

Sea $A \in M(m,n)$. Entonces, la MATRIZ TRANSPUESTA de A, denotada por A^t, es la matriz $n \times m$ que se obtiene de A intercambiando sus renglones y columnas:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix} \Longrightarrow A^t = \begin{bmatrix} \alpha_{11} & \alpha_{21} & \cdots & \alpha_{m1} \\ \alpha_{12} & \alpha_{22} & & \alpha_{m2} \\ \vdots & \vdots & & \vdots \\ \alpha_{1n} & \alpha_{2n} & & \alpha_{mn} \end{bmatrix}$$

O sea que si $\underset{m\times n}{A} = \underset{\substack{1\leq i\leq m \\ 1\leq j\leq n}}{[\alpha_{ij}]} = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix} = [a_1|a_2|\cdots|a_n]$

entonces

1) $A^t = [r_1|r_2|\cdots|r_m]$

2) $A^t = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{bmatrix}$

3) $A^t = [\alpha_{ji}] \qquad 1\leq j\leq n, \ 1\leq i\leq m$

EJEMPLOS

1.

$$X = \begin{bmatrix} E_1 \\ E_2 \\ \vdots \\ E_n \end{bmatrix} \Rightarrow X^t = [E_1 \ \ E_2 \ \cdots \ E_n]$$

2. $y^t = [\eta_1 \ \eta_2 \cdots \eta_n] \Rightarrow Y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 3 \end{bmatrix} \Rightarrow A^t = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 3 \end{bmatrix}$

4. $A = \begin{bmatrix} 1 & 2 & 1 \\ -1 & 1 & 3 \end{bmatrix} \Rightarrow A^t = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ 1 & 3 \end{bmatrix}$

5. $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & -3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \Rightarrow A^t = A$

6. $A = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 2 \\ 1 & -2 & 0 \end{bmatrix} \Rightarrow A^t = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -2 \\ -1 & 2 & 0 \end{bmatrix}$

es decir $A^t = -A$.

Las matrices tales que $A^t = A$ ó $A^t = -A$ merecen nombres especiales :

DEFINICION

Sea $A_{n\times n}$. Decimos que A es SIMETRICA si $A^t = A$; decimos que A es ANTI-SIMETRICA si $A^t = -A$.

La matriz del ejemplo 5 es simétrica, y la matriz del ejemplo 6. es anti-simétrica.

Más tarde veremos que trabajar con matrices simétricas es provechoso; por ejemplo, basta con tener datos sobre la mitad de los elementos de la matriz para conocer toda la matriz.

PROPIEDADES DE LA TRANSPOSICION

1. $(A^t)^t = A \qquad \forall A_{m\times n}$
2. $(A+B)^t = A^t + B^t \qquad \forall A, B \in M(m,n)$
3. $(Ax)^t = x^t A^t \qquad \forall A_{m\times n}, \; x \in \mathbb{R}^n$
4. $(y^t A)^t = A^t y \qquad \forall A_{m\times n}, \; y \in \mathbb{R}^m$
5. $(AB)^t = B^t A^t \qquad \forall A_{m\times n}, \; B_{n\times k}$

Demostración :

1. Sea $A = [a_1 | a_2 | \cdots | a_n]$

$$\Rightarrow \quad A^t = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{bmatrix}$$

$$\Rightarrow \quad (A^t)^t = [a_1 | a_2 | \cdots | a_n] = A$$

$$\therefore \quad (A^t)^t = A \quad \blacksquare$$

2. $$(A+B)^t = [a_1+b_1 | a_2+b_2 | \cdots | a_n+b_n]^t$$

$$= \begin{bmatrix} (a_1+b_1)^t \\ (a_2+b_2)^t \\ \vdots \\ (a_n+b_n)^t \end{bmatrix} = \begin{bmatrix} a_1^t + b_1^t \\ a_2^t + b_2^t \\ \vdots \\ a_n^t + b_n^t \end{bmatrix}$$

$$= \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{bmatrix} + \begin{bmatrix} b_1^t \\ b_2^t \\ \vdots \\ b_n^t \end{bmatrix} = A^t + B^t$$

$$\therefore \quad (A+B)^t = A^t + B^t \quad \square$$

3.
$$Ax = \begin{bmatrix} r_1^t x \\ r_2^t x \\ \vdots \\ r_m^t x \end{bmatrix}$$

$$\therefore\ (Ax)^t = \begin{bmatrix} r_1^t x & r_2^t x & \cdots & r_m^t x \end{bmatrix}$$

$$= \begin{bmatrix} x^t r_1 & x^t r_2 & \cdots & x^t r_m \end{bmatrix} = x^t A^t$$

$$\therefore\ (Ax)^t = x^t A^t \ \blacksquare$$

4. $$y^t A = \begin{bmatrix} y^t a_1 & y^t a_2 & \cdots & y^t a_n \end{bmatrix}$$

$$\therefore (y^t A)^t = \begin{bmatrix} y^t a_1 \\ y^t a_2 \\ \vdots \\ y^t a_n \end{bmatrix} = \begin{bmatrix} a_1^t y \\ a_2^t y \\ \vdots \\ a_n^t y \end{bmatrix} = A^t y \qquad \therefore (y^t A) = A^t y \ \blacksquare$$

5 $$AB = \begin{bmatrix} Ab_1 | Ab_2 | \cdots | Ab_k \end{bmatrix}$$

$$\therefore (AB)^t = \begin{bmatrix} (Ab_1)^t \\ (Ab_2)^t \\ \vdots \\ (Ab_k)^t \end{bmatrix} = \begin{bmatrix} b_1^t A^t \\ b_2^t A^t \\ \vdots \\ b_k^t A^t \end{bmatrix} = B^t A^t$$

$$\therefore\ (AB)^t = B^t A^t \ \blacksquare$$

Ejemplo

Sea $A_{m \times n}$. Entonces A^tA es simétrica, ya que

$$(A^tA)^t = A^t(A^t)^t = A^tA \quad \blacksquare$$

PROPIEDAD

$$A = A^t \iff A = [a_1|a_2|\cdots|a_n] = \begin{bmatrix} a_1{}^t \\ a_2{}^t \\ \vdots \\ a_n{}^t \end{bmatrix} .$$

MATRICES ESPECIALES

La presencia de elementos $\alpha_{ij} = 0$ en una matriz generalmente simplifica los cálculos que involucran a dicha matriz.

Algunas matrices con determinada distribución de ceros son tan importantes que merecen nombres especiales.

MATRICES DIAGONALES

DEFINICIÓN

Sea $D = [\delta_{ij}]$ una matriz cuadrada de orden n. D es DIAGONAL si $i \neq j \Rightarrow \delta_{ij} = 0$.

Es decir, D es diagonal cuando sus únicos elementos diferentes de cero están en la diagonal.

A veces denotaremos a una matriz diagonal por

$D = \text{diag}(\delta_i)$, donde δ_i es el i-ésimo elemento diagonal , ó por

$$D = \text{diag}(\delta_1, \delta_2, \cdots, \delta_n)$$

$$D = \begin{bmatrix} \delta_1 & & & \\ & \delta_2 & & \\ & & \ddots & \\ & & & \delta_n \end{bmatrix}$$

¿ Cuál es el efecto de multiplicar una matriz $A \in M(m,n)$ por una matriz diagonal ?

al pre-multiplicar A por una matriz diagonal de orden n, el i-ésimo renglón de A se multiplica por δ_i :

$$DA = \begin{bmatrix} \delta_1 & & & \\ & \delta_2 & & \\ & & \ddots & \\ & & & \delta_m \end{bmatrix} \begin{bmatrix} r_1{}^t \\ r_2{}^t \\ \vdots \\ r_m{}^t \end{bmatrix} = \begin{bmatrix} \delta_1 r_1{}^t \\ \delta_2 r_2{}^t \\ \vdots \\ \delta_m r_m{}^t \end{bmatrix}$$

al pos-multiplicar A por una matriz diagonal de orden n, la i-ésima columna de A se multiplica por δ_i :

$$AD = [a_1 | a_2 | \cdots | a_n] \begin{bmatrix} \delta_1 & & & \\ & \delta_2 & & \\ & & \ddots & \\ & & & \delta_n \end{bmatrix} = [\delta_1 a_1 | \delta_2 a_2 | \cdots | \delta_n a_n]$$

Ejemplos

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 6 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 4 & 8 & 2 \\ 0 & 18 & 15 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 6 & 5 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 0 \\ 2 & 8 & 3 \\ 0 & 12 & 15 \end{bmatrix}$$

En general vamos a tener que $DA \neq AD$. Sin embargo, si D_1 y D_2 son diagonales entonces $D_1 D_2 = D_2 D_1$.

Si $D = \operatorname{diag}(\delta_i)$, con $\delta_i = \delta \quad \forall i = 1, \ldots, n$ entonces $DA = AD \quad \forall A_{n \times n}$.

Ejemplos

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix} \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 6 & 5 \end{bmatrix} = \begin{bmatrix} \delta & 2\delta & 0 \\ 2\delta & 4\delta & \delta \\ 0 & 6\delta & 5\delta \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 6 & 5 \end{bmatrix} \begin{bmatrix} \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}$$

MATRIZ IDENTIDAD

DEFINICION

La matriz de orden n

$$I_n = \operatorname{diag}(1, 1, \ldots, 1)$$

se llama MATRIZ IDENTIDAD DE ORDEN n.

Generalmente el orden de I_n estará determinado por el contexto. En tales casos escribiremos $I_n = I$.

OBSERVAR que

$$I_n = [e_1 | e_2 | \ldots | e_n] = \begin{bmatrix} e_1^t \\ e_2^t \\ \vdots \\ e_n^t \end{bmatrix}$$

donde $e_i \in \mathbb{R}^n$ $\quad 1 \leq i \leq n$.

PROPIEDAD

Sea A $m \times n$. Entonces

$$I_m A = A I_n = A$$

Demostración

$$I_m A = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix} = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$$

$$A I_n = [a_1 | a_2 | \cdots | a_n] \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = [a_1 | a_2 | \cdots | a_n] \blacksquare$$

PERMUTACIONES

DEFINICION 1

Una matriz de orden n se dice que es una PERMUTACION ELEMENTAL si es de la forma

$$P_{ij} = [e_1 | e_2 | \dots | e_{i-1} | \underset{\substack{\uparrow \\ \text{lugar } i}}{e_j} | e_{i+1} | \cdots | e_{j-1} | \underset{\substack{\uparrow \\ \text{lugar } j}}{e_i} | e_{j+1} | \cdots | e_n]$$

donde $i<j$.

Es decir, P_{ij} es la matriz que se obtiene a partir de la identidad intercambiando la columna i-ésima con la j-ésima.

Es decir

$$P_{ij} = \left[\begin{array}{ccccccccccccc|c}
 & & & & i & & & & j & & & & & \\
1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \\
0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & i \\
0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 & \\
0 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & j \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 & \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 & \\
0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 &
\end{array}\right]$$

OBSERVACION

P_{ij} es también la matriz que se obtiene a partir de la matriz identidad intercambiando el renglón i-ésimo con el j-ésimo.

Podríamos haber definido una matriz de permutación en función de sus renglones.

DEFINICION 2

Una matriz de orden n se dice que es una PERMUTACION ELEMENTAL si es de la forma

$$P_{ij} = \begin{array}{c} \left[\begin{array}{c} e_1^t \\ e_2^t \\ \vdots \\ e_{i-1}^t \\ e_j^t \\ e_{i+1}^t \\ \vdots \\ e_{j-1}^t \\ e_i^t \\ e_{j+1}^t \\ \vdots \\ e_{n-1}^t \\ e_n^t \end{array}\right] \end{array} \begin{array}{l} \\ \\ \\ \\ \leftarrow i \\ \\ \\ \\ \leftarrow j \\ \\ \\ \\ \\ \end{array}$$

donde $i < j$.

Por ejemplo

$$P_{25} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

PROPIEDADES DE LAS PERMUTACIONES ELEMENTALES

1. $P_{ij}{}^{t} = P_{ij}$

2. $(P_{ij})^2 = I$

Demostración

1. $$P_{ij}{}^{t} = \left[e_1 | e_2 | \cdots | e_j | \cdots | e_i | \cdots | e_n\right]^{t} = \begin{bmatrix} e_1{}^{t} \\ \vdots \\ e_j{}^{t} \\ \vdots \\ e_i{}^{t} \\ \vdots \\ e_n{}^{t} \end{bmatrix} = P_{ij}$$ ■

2. $$\begin{bmatrix} e_1{}^{t} \\ \vdots \\ e_j{}^{t} \\ \vdots \\ e_i{}^{t} \\ \vdots \\ e_n{}^{t} \end{bmatrix} \left[e_1 | \cdots | e_j | \cdots | e_i | \cdots | e_n\right] =$$

$$= \begin{bmatrix} e_1{}^{t}e_1 & e_1{}^{t}e_2 \cdots & e_1{}^{t}e_j \cdots & e_1{}^{t}e_i \cdots & e_1{}^{t}e_n \\ e_2{}^{t}e_1 & e_2{}^{t}e_2 \cdots & e_2{}^{t}e_j \cdots & e_2{}^{t}e_i \cdots & e_2{}^{t}e_n \\ \vdots & \vdots & & & \vdots \\ e_j{}^{t}e_1 & e_j{}^{t}e_2 \cdots & e_j{}^{t}e_j \cdots & e_j{}^{t}e_i \cdots & e_j{}^{t}e_n \\ e_i{}^{t}e_1 & e_i{}^{t}e_2 \cdots & e_i{}^{t}e_j \cdots & e_i{}^{t}e_i \cdots & e_i{}^{t}e_n \\ & & & & \\ e_n{}^{t}e_1 & e_n{}^{t}e_2 \cdots & e_n{}^{t}e_j \cdots & e_n{}^{t}e_i \cdots & e_n{}^{t}e_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} = I$$

El punto no. 2 de las propiedades nos está diciendo que si intercambiamos dos renglones de la matriz identidad y después los volvemos a intercambiar, entonces al final tendremos a la matriz original - la identidad. Esto también se puede pensar por columnas: Si intercambiamos dos columnas de la identidad y después las volvemos a intercambiar, el efecto sobre la identidad es nulo

Por ejemplo

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

¿ Qué pasa si aplicamos P_{ij} a $x \in \mathbb{R}^n$?

$$P_{ij}\,x = \begin{bmatrix} e_1^t \\ \vdots \\ e_j^t \\ \vdots \\ e_i^t \\ \vdots \\ e_n^t \end{bmatrix} x = \begin{bmatrix} e_1^t x \\ \vdots \\ e_j^t x \\ \vdots \\ e_i^t x \\ \vdots \\ e_n^t x \end{bmatrix} = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_j \\ \vdots \\ \xi_i \\ \vdots \\ \xi_n \end{bmatrix}$$

Es decir, si P_{ij} es la permutación en la que se han intercambiado el renglón i-ésimo con el j-ésimo, entonces al aplicarla al vector x, produce el efecto de intercambiar la i-ésima componente de x con la j-ésima.

Ejemplo

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

¿ Qué pasa al aplicar P_{ij} a $y^t \in \mathbb{R}^n$?

$$y^t P_{ij} = y^t [e_1 | \cdots | e_j | \cdots | e_i | \cdots e_n]$$

$$= [y^t e_1 | \cdots | y^t e_j | \cdots | y^t e_i | \cdots | y^t e_n]$$

$$= [\eta_1 \cdots \eta_j \cdots \eta_i \cdots \eta_n]$$

O sea que al aplicarla a y^t, P_{ij} intercambia la i-ésima y la j-ésima componentes de y^t.

Ejemplo.

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

Ahora, sea A una matriz $n \times k$. ¿Qué efecto produce sobre A la premultiplicación por P_{ij}?

$$P_{ij} A = \begin{bmatrix} e_1^t \\ \vdots \\ e_j^t \\ \vdots \\ e_i^t \\ \vdots \\ e_n^t \end{bmatrix} \cdot A = \begin{bmatrix} e_1^t A \\ \vdots \\ e_j^t A \\ \vdots \\ e_i^t A \\ \vdots \\ e_n^t A \end{bmatrix} = \begin{bmatrix} r_1^t \\ \vdots \\ r_j^t \\ \vdots \\ r_i^t \\ \vdots \\ r_n^t \end{bmatrix}$$

O sea, que si P_{ij} es la permutación en la cual están intercambiados el i-ésimo y el j-ésimo renglón de la identidad, entonces al premultiplicar A por P_{ij} se intercambian el i-ésimo renglón de A

con el j-ésimo.

Ejemplo

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

Análogamente, si pos-multiplicamos P_{ij} por A, entonces P_{ij} actúa sobre A intercambiando la i-ésima columna de A con la j-ésima columna de A:

$$A P_{ij} = [a_1 | a_2 | \cdots | a_n] [e_1 | \cdots | e_j | \cdots | e_i | \cdots | e_n]$$

$$= [a_1 | \cdots | a_j | \cdots | a_i | \cdots | a_n]$$

Ejemplo

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

DEFINICION

La matriz P es una PERMUTACION si es

cuadrada y es tal que cada columna y cada renglón tienen un único elemento distinto de cero, cuyo valor es 1.

Es decir, una matriz de permutación tiene una de las siguientes representaciones

1. $P = [e_{i_1} | e_{i_2} | \dots | e_{i_n}]$ donde $i_1, i_2, \dots, i_n \in \{1, \dots, n\}$ y $i_k \neq i_l$ si $k \neq l$

2. $P = \begin{bmatrix} e_{j_1}^t \\ e_{j_2}^t \\ \vdots \\ e_{j_n}^t \end{bmatrix}$ donde $j_1, j_2, \dots, j_n \in \{1, \dots, n\}$ y $j_k \neq j_l$ si $k \neq l$

Una permutación se obtiene a partir de la matriz identidad intercambiando varios renglones (ó varias columnas)

Por ejemplo, si primero intercambiamos la segunda y la tercera columna, y después intercambiamos la primera columna con la segunda, obtenemos la permutación.

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

y podemos expresar a P como

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Si primero intercambiamos el segundo renglón con el tercer renglón y después intercambiamos el primer renglón con el segundo renglón entonces obtenemos la permutación

$$P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

OBSERVACIONES

Para matrices de permutación vamos a tener que en general $P^t \neq P$ y $P^2 \neq I$

Por ejemplo

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} e_3^t \\ e_1^t \\ e_2^t \end{bmatrix} = [e_2 | e_3 | e_1]$$

PROPIEDADES

1. Si $P = \begin{bmatrix} e_{i_1} \\ e_{i_2} \\ \vdots \\ e_{i_n} \end{bmatrix}$ entonces $PA = \begin{bmatrix} r_{i_1}^t \\ r_{i_2}^t \\ \vdots \\ r_{i_n}^t \end{bmatrix}$

2. Si $P = [e_{j_1} | e_{j_2} | \cdots | e_{j_n}]$ entonces

$$AP = [a_{j_1} | a_{j_2} | \cdots | a_{j_n}]$$

MATRICES TRAPEZOIDALES

DEFINICION

Sea A una matriz mxn. Decimos que A es TRAPEZOIDAL SUPERIOR si $i > j \rightarrow \alpha_{ij} = 0$.

Una matriz trapezoidal superior que además es cuadrada se dice que es TRIANGULAR SUPERIOR

Las matrices trapezoidales superiores tienen una de las siguientes formas

$$\begin{bmatrix} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & * & * \\ 0 & 0 & \cdots & 0 & * \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

(m>n)

$$\begin{bmatrix} * & * & \cdots & * & * & \cdots & * \\ 0 & * & \cdots & * & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & * & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & \cdots & * \end{bmatrix}$$

(n>m)

$$\begin{bmatrix} * & * & \cdots & * \\ 0 & * & & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{bmatrix}$$

(m=n)

Siempre que utilizemos la letra U nos estaremos refiriendo a una matriz triangular superior.

DEFINICION

Sea A una matriz $m \times n$. Decimos que A es TRAPEZOIDAL INFERIOR si $i < j \Rightarrow \alpha_{ij} = 0$.

Una matriz trapezoidal inferior que además es cuadrada se dice que es TRIANGULAR INFERIOR.

Las matrices trapezoidales inferiores tienen una de las siguientes formas:

$$\begin{bmatrix} \ast & 0 & 0 & \cdots & 0 & 0 \\ \ast & \ast & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \ast & \ast & \ast & \cdots & \ast & 0 \\ \ast & \ast & \ast & \cdots & \ast & \ast \\ \ast & \ast & \ast & \cdots & \ast & \ast \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \ast & \ast & \ast & \cdots & \ast & \ast \end{bmatrix}$$

$(m>n)$

$$\begin{bmatrix} \ast & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \ast & \ast & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \ast & \ast & \cdots & \ast & 0 & 0 & \cdots & 0 \\ \ast & \ast & \cdots & \ast & \ast & 0 & \cdots & 0 \end{bmatrix}$$

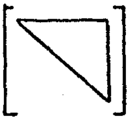
$(m<n)$

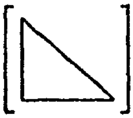
$$\begin{bmatrix} \ast & 0 & 0 & \cdots & 0 \\ \ast & \ast & 0 & \cdots & 0 \\ \ast & \ast & \ast & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \ast & \ast & \ast & \cdots & \ast \end{bmatrix}$$

$(m=n)$

Siempre que denotemos a una matriz con la letra L se sobre entenderá que L es triangular inferior.

A veces dibujaremos la forma de una matriz para referirnos a ella. Por ejemplo

U =


L =


PROPIEDAD

Si L es triangular inferior, entonces

$$L = \begin{bmatrix} \lambda_{11} e_1^t \\ \lambda_{21} e_1^t + \lambda_{22} e_2^t \\ \lambda_{31} e_1^t + \lambda_{32} e_2^t + \lambda_{33} e_3^t \\ \vdots \\ \lambda_{n1} e_1^t + \lambda_{n2} e_2^t + \lambda_{n3} e_3^t + \cdots + \lambda_{nn} e_n^t \end{bmatrix} :$$

Demostración

$$L = \begin{bmatrix} \lambda_{11} & & & \\ \lambda_{21} & \lambda_{22} & & \\ \vdots & \vdots & & \\ \lambda_{n1} & \lambda_{n2} & \cdots & \lambda_{nn} \end{bmatrix} \begin{bmatrix} e_1^t \\ e_2^t \\ \vdots \\ e_n^t \end{bmatrix} = \begin{bmatrix} \lambda_{11} e_1^t \\ \lambda_{21} e_1^t + \lambda_{22} e_2^t \\ \vdots \\ \lambda_{n1} e_1^t + \lambda_{n2} e_2^t + \cdots + \lambda_{nn} e_n^t \end{bmatrix}$$

PROPIEDAD

Si U es triangular superior entonces

$$U = \left[\mu_{11} e_1 \mid \mu_{12} e_1 + \mu_{22} e_2 \mid \mu_{13} e_1 + \mu_{23} e_2 + \mu_{33} e_3 \mid \cdots \mid \mu_{1n} e_1 + \mu_{2n} e_2 + \cdots + \mu_{nn} e_n\right]$$

Demostración

$$U = \left[e_1 | e_2 | \cdots e_n\right] \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \cdots & \mu_{1n} \\ & \mu_{22} & \mu_{23} & \cdots & \mu_{2n} \\ & & \mu_{33} & \cdots & \mu_{3n} \\ & & & & \vdots \\ & & & & \mu_{nn} \end{bmatrix}$$

$$= \left[\mu_{11} e_1 \mid \mu_{12} e_1 + \mu_{22} e_2 \mid \cdots \mid \mu_{1n} e_1 + \mu_{2n} e_2 + \cdots + \mu_{nn} e_n\right] \blacksquare$$

PROPIEDAD

El producto de dos matrices triangulares inferiores es triangular inferior.

Demostración :

Supongamos que

$$L = \begin{bmatrix} \lambda_{11} & & & \\ \lambda_{21} & \lambda_{22} & & \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & \\ \vdots & & & \\ \lambda_{n1} & \lambda_{n2} & \lambda_{n3} & \cdots \ \lambda_{nn} \end{bmatrix} \qquad A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_n^t \end{bmatrix}$$

Entonces

$$LA = \begin{bmatrix} \lambda_{11} r_1^t \\ \lambda_{21} r_1^t + \lambda_{22} r_2^t \\ \lambda_{31} r_1^t + \lambda_{32} r_2^t + \lambda_{33} r_3^t \\ \vdots \\ \lambda_{n1} r_1^t + \lambda_{n2} r_2^t + \cdots + \lambda_{nn} r_n^t \end{bmatrix} = \begin{bmatrix} z_1^t \\ z_2^t \\ z_3^t \\ \vdots \\ z_n^t \end{bmatrix}$$

Si A es triangular inferior entonces

$$r_1^t = \alpha_{11} e_1^t$$

$$r_2^t = \alpha_{21} e_1^t + \alpha_{22} e_2^t$$

$$r_3^t = \alpha_{31} e_1^t + \alpha_{32} e_2^t + \alpha_{33} e_3^t$$

$$\vdots$$

$$r_4^t = \alpha_{n1} e_1^t + \alpha_{n2} e_2^t + \alpha_{n3} e_3^t + \cdots + \alpha_{nn} e_n^t$$

por lo tanto

$$z_1^t = \lambda_{11}(\alpha_{11} e_1^t) = (\lambda_{11}\alpha_{11}) e_1^t$$

$$\begin{aligned} z_2^t &= \lambda_{21}(\alpha_{11} e_1^t) + \lambda_{22}(\alpha_{21} e_1^t + \alpha_{22} e_2^t) \\ &= (\lambda_{21}\alpha_{11} + \lambda_{22}\alpha_{21}) e_1^t + (\lambda_{22}\alpha_{22}) e_2^t \end{aligned}$$

$$\begin{aligned} z_3^t &= \lambda_{31}(\alpha_{11} e_1^t) + \lambda_{32}(\alpha_{21} e_1^t + \alpha_{22} e_2^t) + \\ &\quad \lambda_{33}(\alpha_{31} e_1^t + \alpha_{32} e_2^t + \alpha_{33} e_3^t) \\ &= (\lambda_{31}\alpha_{11} + \lambda_{32}\alpha_{21} + \lambda_{33}\alpha_{31}) e_1^t + \\ &\quad (\lambda_{32}\alpha_{22} + \lambda_{33}\alpha_{32}) e_2^t + \lambda_{33}\alpha_{33} e_3^t \end{aligned}$$

$$\vdots$$

$$\begin{aligned} z_n^t &= \lambda_{n1}(\alpha_{11} e_1^t) + \lambda_{n2}(\alpha_{21} e_1^t + \alpha_{22} e_2^t) + \cdots \\ &\quad + \lambda_{nn}(\alpha_{n1} e_1^t + \alpha_{n2} e_2^t + \alpha_{n3} e_3^t + \cdots + \alpha_{nn} e_n^t) \\ &= (\lambda_{n1}\alpha_{11} + \lambda_{n2}\alpha_{21} + \cdots + \lambda_{nn}\alpha_{n1}) e_1^t + \\ &\quad (\lambda_{n2}\alpha_{22} + \cdots + \lambda_{nn}\alpha_{n2}) e_2^t + \cdots + (\lambda_{nn}\alpha_{nn}) e_n^t \end{aligned}$$

$$\begin{aligned} \therefore \quad z_1^t &= \zeta_{11} e_1^t \\ z_2^t &= \zeta_{21} e_1^t + \zeta_{22} e_2^t \\ &\vdots \\ z_n^t &= \zeta_{n1} e_1^t + \zeta_{n2} e_2^t + \cdots + \zeta_{nn} e_n^t \end{aligned}$$

$\therefore$ LA es triangular inferior ■

PROPIEDAD

El producto de dos matrices triangulares superiores es triangular superior.

Demostración

Supongamos

$$U = \begin{bmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \cdots & \mu_{1n} \\ & \mu_{22} & \mu_{23} & \cdots & \mu_{2n} \\ & & \mu_{33} & \cdots & \mu_{3n} \\ & & & & \\ & & & & \mu_{nn} \end{bmatrix} \quad ; \quad A = [a_1 | a_2 | \cdots | a_n]$$

Entonces

$$AU = [\mu_{11} a_1 | \mu_{12} a_1 + \mu_{22} a_2 | \mu_{13} a_1 + \mu_{23} a_2 + \mu_{33} a_3 | \cdots$$

$$| \mu_{1n} a_1 + \mu_{2n} a_2 + \cdots + \mu_{nn} a_n]$$

$$= [b_1 | b_2 | \cdots | b_n]$$

Si A es triangular superior entonces

$a_1 = \alpha_{11} e_1$

$a_2 = \alpha_{12} e_1 + \alpha_{22} e_2$

$$a_3 = \alpha_{13} e_1 + \alpha_{23} e_2 + \alpha_{33} e_3$$

$$\vdots$$

$$a_n = \alpha_{1n} e_1 + \alpha_{2n} e_2 + \alpha_{3n} e_3 + \cdots + \alpha_{nn} e_n$$

Por lo tanto

$$\begin{aligned} b_1 &= \mu_{11} (\alpha_{11} e_1) = (\mu_{11} \alpha_{11}) e_1 \\ &= \beta_{11} e_1 \end{aligned}$$

$$\begin{aligned} b_2 &= \mu_{12} (\alpha_{11} e_1) + \mu_{22} (\alpha_{12} e_1 + \alpha_{22} e_2) \\ &= (\mu_{12} \alpha_{11} + \mu_{22} \alpha_{12}) e_1 + (\mu_{22} \alpha_{22}) e_2 \\ &= \beta_{12} e_1 + \beta_{22} e_2 \end{aligned}$$

$$\begin{aligned} b_3 &= \mu_{13} (\alpha_{11} e_1) + \mu_{23} (\alpha_{12} e_1 + \alpha_{22} e_2) \\ &\qquad + \mu_{33} (\alpha_{13} e_1 + \alpha_{23} e_2 + \alpha_{33} e_3) \\ &= (\mu_{13} \alpha_{11} + \mu_{23} \alpha_{12} + \mu_{33} \alpha_{13}) e_1 + (\mu_{23} \alpha_{22} + \mu_{33} \alpha_{23}) e_2 \\ &\qquad + (\mu_{33} \alpha_{33}) e_3 \\ &= \beta_{13} e_1 + \beta_{23} e_2 + \beta_{33} e_3 \end{aligned}$$

$$\vdots$$

$$\begin{aligned} b_n &= \mu_{1n} (\alpha_{11} e_1) + \mu_{2n} (\alpha_{12} e_1 + \alpha_{22} e_2) \\ &\qquad + \mu_{3n} (\alpha_{13} e_1 + \alpha_{23} e_2 + \alpha_{33} e_3) + \cdots \\ &\qquad + \mu_{nn} (\alpha_{1n} e_1 + \alpha_{2n} e_2 + \alpha_{3n} e_3 + \cdots + \alpha_{nn} e_n) \\ &= (\mu_{1n} \alpha_{11} + \mu_{2n} \alpha_{12} + \mu_{3n} \alpha_{13} + \cdots + \mu_{nn} \alpha_{1n}) e_1 \\ &\qquad + (\mu_{2n} \alpha_{22} + \mu_{3n} \alpha_{23} + \cdots + \mu_{nn} \alpha_{2n}) e_2 \\ &\qquad + (\mu_{3n} \alpha_{33} + \cdots + \mu_{nn} \alpha_{3n}) e_3 + \cdots \\ &\qquad + (\mu_{nn} \alpha_{nn}) e_n \\ &= \beta_{1n} e_1 + \beta_{2n} e_2 + \cdots + \beta_{nn} e_n \end{aligned}$$

Por lo tanto AU es triangular superior.

MATRICES POR BLOQUES

Hemos visto que una matriz puede considerarse como un arreglo de escalares, vectores columna ó vectores renglón. Es natural que ahora nos preguntemos si también es factible ver a una matriz como un arreglo de matrices. Veremos que ésto no solamente es posible sino que además nos será de utilidad para construir algoritmos prácticos para manipular matrices. Para introducir la idea de partir una matriz, primero consideraremos esta operación aplicada a vectores.

DEFINICION

Sea $x = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix}$ un vector de orden n. Sean

$$x_1 = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_r \end{pmatrix} \quad \text{y} \quad x_2 = \begin{pmatrix} \xi_{r+1} \\ \xi_{r+2} \\ \vdots \\ \xi_n \end{pmatrix} \quad \text{donde } 1 \leq r < n.$$

Si escribimos $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$, entonces decimos que los vectores x_1 y x_2 son VECTORES ó PARTICIONES de x, y el acto de dividir a un vector en esta forma se llama PARTIR POR BLOQUES.

Ejemplo

$$x = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 3 \end{bmatrix} \qquad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad x_2 = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$ es una partición de x.

$$x = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \end{bmatrix}$$

Ahora, sea $y \in \mathbb{R}^n$ y sea $y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ una partición de y_1 tal que y_1 y x_1 son ambos de orden r. Entonces si $u = x + y$ y u está partido en forma similar a x y y

se puede ver que

$$u_1 = x_1 + y_1$$
$$u_2 = x_2 + y_2$$

ya que

$$\begin{bmatrix} \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_r \end{bmatrix} \\ \begin{bmatrix} \xi_{r+1} \\ \vdots \\ \xi_n \end{bmatrix} \end{bmatrix} + \begin{bmatrix} \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_r \end{bmatrix} \\ \begin{bmatrix} \eta_{r+1} \\ \vdots \\ \eta_n \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} \xi_1+\eta_1 \\ \vdots \\ \xi_n+\eta_n \end{bmatrix} \\ \begin{bmatrix} \xi_{r+1}+\eta_{r+1} \\ \vdots \\ \xi_n+\eta_n \end{bmatrix} \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \end{bmatrix}$$

Además

$$x^t y = \begin{bmatrix} x_1^t & x_2^t \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1^t y_1 + x_2^t y_2$$

Por ejemplo :

$$\begin{bmatrix} \begin{bmatrix} 1 & 2 \end{bmatrix} & \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \end{bmatrix} \begin{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Esto nos hace ver que si x y y están divididos por bloques, entonces al sumar x con y sus particiones se comportan como si fueran elementos. Las particiones de x y y exhiben este mismo comportamiento cuando formamos el producto punto de x y y.

También se puede dividir a un vector en más de dos subvectores, y en algunas aplicaciones más sofisticadas se utiliza. Sin embargo, el objeto de la partición por bloques es la simplificación, mientras dividamos a un vector en menos componentes, tanto mejor.

Consideremos ahora particiones de matrices.

sea $A = [\alpha_{ij}] \in M(m,n)$ y sean

$$A_1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} \cdots & \alpha_{2n} \\ \vdots & & \\ \alpha_{s1} & \alpha_{s2} \cdots & \alpha_{sn} \end{bmatrix}$$

la matriz de los primeros s renglones de A.

$$A_2 = \begin{bmatrix} \alpha_{s+1,1} & \alpha_{s+1,2} & \cdots \alpha_{s+1,n} \\ \alpha_{s+2,1} & \alpha_{s+2,2} & \cdots \alpha_{s+2,n} \\ & & \\ \alpha_{m,1} & \alpha_{m,2} & \alpha_{mn} \end{bmatrix}, \quad 1 \leq s < m.$$

los últimos m-s renglones de A.

Entonces podemos escribir

$$A = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]$$ y decimos que A_1 y A_2

son SUBMATRICES ó BLOQUES de A.

Indicaremos las dimensiones de los bloques de A de la siguiente manera:

$$A_{m \times n} = \begin{array}{c} \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \begin{array}{l} s \\ m-s \end{array} \\ n \end{array}$$

Es decir, A está dividida en dos bloques:

A_1 $s \times n$

y A_2 $(m-s) \times n$

Si $$A = \begin{array}{c} \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \begin{array}{l} s \\ m-s \end{array} \\ n \end{array} \qquad B = \begin{array}{c} \left[\begin{array}{c} B_1 \\ \hline B_2 \end{array}\right] \begin{array}{l} s \\ m-s \end{array} \\ n \end{array}$$

entonces $$A+B = \left[\begin{array}{c} A_1+B_1 \\ \hline A_2+B_2 \end{array}\right]$$

Aquí vemos otra vez que los bloques se comportan como si fueran elementos

Ahora, sea $x \in \mathbb{R}^n$ y A dividida por bloques tal que

$$A = \begin{array}{c|l} \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\ n & \end{array}$$

Entonces

$$Ax = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] x = \left[\begin{array}{c} A_1 x \\ \hline A_2 x \end{array}\right]$$

Si $y^t \in \mathbb{R}^{m*}$ y lo dividimos de la siguiente manera

$$y^t = \underset{}{\left[\underset{s}{y_1^t} \,\middle|\, \underset{m-s}{y_2^t}\right]}$$

entonces

$$y^t A = y^t \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] =$$

$$\left[\underset{s}{y_1^t} \,\middle|\, \underset{m-s}{y_2^t}\right] \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]\begin{array}{l} s \\ m-s \end{array} = \underset{1\times n}{y_1^t A_1} + \underset{1\times n}{y_2^t A_2}$$

Es decir, al multiplicar un renglón por una matriz por bloques, entonces los subrenglones y las submatrices se comportan como elementos escalares.

Un ejemplo de matriz por bloques sería

$$\left[\begin{array}{c}\begin{bmatrix} * & & & & \\ * & * & & & \\ * & * & * & & \\ \vdots & & & \ddots & \\ * & * & * & \cdots & * \end{bmatrix} \\ \begin{bmatrix} * & * & * & \cdots & * \\ \vdots & & & & \vdots \\ * & * & * & \cdots & * \end{bmatrix}\end{array}\right]$$

También podríamos dividir a A por medio de las submatrices

$$A_1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2r} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mr} \end{bmatrix}$$

– matriz con las primeras r columnas de A

$$A_2 = \begin{bmatrix} \alpha_{1,r+1} & \alpha_{1,r+2} & \cdots & \alpha_{1n} \\ \alpha_{2,r+1} & \alpha_{2,r+2} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \\ \alpha_{m,r+1} & \alpha_{m,r+2} & \cdots & \alpha_{mn} \end{bmatrix}$$

– últimas $n-r$ columnas de A

donde $1 \leq r < n$.

En este caso escribimos

$$A = \begin{array}{c} \left[A_1 \mid A_2 \right] m \\ \quad r \quad n-r \end{array}$$

Para multiplicar A por $x \in \mathbb{R}^n$ tenemos que partir al vector x en

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{array}{c} r \\ n-r \end{array}$$

Entonces

$$Ax = {}^{m}\underset{\quad r \qquad n-r}{\left[A_1 \mid A_2 \right]} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{array}{c} r \\ n-r \end{array} = A_1 x_1 + A_2 x_2$$

Si $y^t \in \mathbb{R}^{m*}$, entonces la multiplicación por y^t está dada por

$$y^t \left[A_1 \mid A_2 \right] = \left[y^t A_1 \mid y^t A_2 \right]$$

El caso más general que nos interesa es cuando A está dividida en cuatro bloques:

$$A = \left[\begin{array}{cccc|cccc}
\alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} & \alpha_{1,r+1} & \alpha_{1,r+2} & \cdots & \alpha_{1n} \\
\alpha_{21} & \alpha_{22} & \cdots & \alpha_{2r} & \alpha_{2,r+1} & \alpha_{2,r+2} & \cdots & \alpha_{2n} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\alpha_{s1} & \alpha_{s2} & \cdots & \alpha_{sr} & \alpha_{s,r+1} & \alpha_{s,r+2} & \cdots & \alpha_{sn} \\
\hline
\alpha_{s+1,1} & \alpha_{s+1,2} & \cdots & \alpha_{s+1,r} & \alpha_{s+1,r+1} & \alpha_{s+1,r+2} & \cdots & \alpha_{s+1,n} \\
\alpha_{s+2,1} & \alpha_{s+2,2} & \cdots & \alpha_{s+2,r} & \alpha_{s+2,r+1} & \alpha_{s+2,r+2} & \cdots & \alpha_{s+2,n} \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\alpha_{m1} & \alpha_{m2} & & \alpha_{m,r} & \alpha_{m,r+1} & \alpha_{m,r+2} & \cdots & \alpha_{m,n}
\end{array}\right]$$

$$= \begin{array}{cc}
\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\
\begin{array}{cc} r & n-r \end{array} &
\end{array}$$

Entonces

$$\text{Si} \qquad X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \begin{array}{l} r \\ n-r \end{array} \qquad Y^t = \begin{array}{cc} s & m-s \\ \multicolumn{2}{c}{\begin{bmatrix} Y_1^t & Y_2^t \end{bmatrix}} \end{array}$$

$$AX = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \left[\begin{array}{c} A_{11}X_1 + A_{12}X_2 \\ \hline A_{21}X_1 + A_{22}X_2 \end{array}\right]$$

$$Y^t A = \begin{bmatrix} Y_1^t & Y_2^t \end{bmatrix} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = \begin{bmatrix} Y_1^t A_{11} + Y_2^t A_{21} & Y_1^t A_{12} + Y_2^t A_{22} \end{bmatrix}$$

Finalmente, si A y B son dos matrices divididas por bloques adecuadamente, entonces podemos operar con sus bloques como si fueran escalares:

$$\text{si} \qquad \underset{m\times n}{A} = \begin{array}{cc} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\ \begin{array}{cc} r & \quad n-r \end{array} & \end{array} \qquad \underset{n\times k}{B} = \begin{array}{cc} \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} t & \quad k-t \end{array} & \end{array}$$

entonces

$$AB = \begin{array}{cc} \left[\begin{array}{c|c} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ \hline A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\ \begin{array}{cc} t & \qquad\qquad k-t \end{array} & \end{array}$$

Decimos que un conjunto de matrices está dividido CONFORMEMENTE si las operaciones a realizar entre los bloques correspondientes están definidas.

$$\left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \hline * & * & * & * \\ * & * & * & * \end{array}\right] \qquad \left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \hline * & * & * & * \\ * & * & * & * \end{array}\right]$$

Estas matrices están divididas conformemente con respecto a la suma (mas no c.r.a. multiplicación)

Las sig. matrices están divididas conformemente con respecto a la multiplicación (mas no con respecto a la suma).

$$\left[\begin{array}{ccc|c} * & * & * & * \\ * & * & * & * \\ \hline * & * & * & * \\ * & * & * & * \end{array}\right] \left[\begin{array}{c|ccc} * & * & * & * \\ * & * & * & * \\ * & * & * & * \\ \hline * & * & * & * \end{array}\right] = \left[\begin{array}{c|ccc} * & * & * & * \\ * & * & * & * \\ \hline * & * & * & * \\ * & * & * & * \end{array}\right]$$

PROPIEDAD

$$\text{Si} \quad A = \begin{array}{cc} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array} \quad \text{entonces} \quad A^t = \begin{array}{cc} \left[\begin{array}{c|c} A_{11}^t & A_{21}^t \\ \hline A_{12}^t & A_{22}^t \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} s & m-s \end{array} & \end{array}$$

Demostración

$$A = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & 0 \end{array}\right] \end{array} \begin{array}{c} s \\ + \\ m-s \end{array} \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & A_{12} \\ \hline 0 & 0 \end{array}\right] \end{array} + \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & 0 \\ \hline A_{21} & 0 \end{array}\right] \end{array} \begin{array}{c} s \\ + \\ m-s \end{array} \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A_{22} \end{array}\right] \end{array}$$

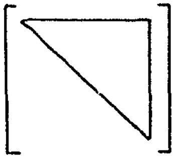
$$\therefore \quad A^t = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} A_{11}^t & 0 \\ \hline 0 & 0 \end{array}\right] \end{array} \begin{array}{c} r \\ + \\ n-r \end{array} \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & 0 \\ \hline A_{12}^t & 0 \end{array}\right] \end{array} \begin{array}{c} r \\ + \\ n-r \end{array} \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & A_{21}^t \\ \hline 0 & 0 \end{array}\right] \end{array} \begin{array}{c} r \\ + \\ n-r \end{array} \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A_{22}^t \end{array}\right] \end{array}$$

$$= \begin{array}{cc} \left[\begin{array}{c|c} A_{11}^t & A_{21}^t \\ \hline A_{12}^t & A_{22}^t \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} s & m-s \end{array} & \end{array}$$

Por medio del siguiente problema solo queremos hacer ver la utilidad de las matrices como herramienta para resolver ciertos problemas. En los capítulos que siguen presentamos la teoría que nos permitirá establecer las condiciones bajo las cuales se pueden aplicar los procedimientos que aquí se presentan.

PROBLEMA

Dada una matriz $A_{n \times n}$ encontrar una matriz $B_{n \times n}$ tal que

$BA =$ 

donde B no debe ser trivial, es decir, no debe tener renglones con puros ceros.

Podríamos resolver el problema de llevar a una matriz a la forma triangular por pasos:

En el primer paso nuestro objetivo sería encontrar una matriz B_1 tal que

$$B_1 A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix} = A_1$$

En el segundo paso buscaríamos una matriz B_2 tal que

$$B_2 A_1 = \begin{bmatrix} * & * & * & \cdots & * & * \\ 0 & * & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ 0 & 0 & * & \cdots & * & * \\ \vdots & & & & & \\ 0 & 0 & * & \cdots & * & * \end{bmatrix} = A_2$$

Y en general, en el k-ésimo paso buscaríamos una matriz B_k tal que

$$B_k A_{k-1} = \begin{array}{r c} & \begin{array}{ccccccc} & & & \overset{k}{\downarrow} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ k \rightarrow \\ \\ \\ \\ \end{array} & \left[\begin{array}{ccccccc} * & * & \cdots & * & * & \cdots & * \\ 0 & * & \cdots & * & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \\ 0 & 0 & \cdots & * & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & * & \cdots & * \end{array}\right] \end{array} = A_k$$

al terminar el paso $n-1$, el proceso acaba, con A_{n-1} la matriz triangular superior.

Si nos fijamos en las matrices $A_1, A_2, \cdots, A_{n-1}$, vemos que nos conviene dividirlas por bloques.

Después del primer paso tenemos

$$A_1 = \begin{array}{r|c|ccc|} & 1 & & n-1 & \\ \hline 1 & * & * & \cdots & * \\ \hline & 0 & * & \cdots & * \\ & \vdots & \vdots & & \vdots \\ n-1 & 0 & * & \cdots & * \end{array} = \begin{array}{|c|c|l} 1 & n-1 & \\ \hline * & * \cdots * & 1 \\ \hline 0 & \bar{A}_1 & n-1 \end{array} = B_1 A$$

En el segundo paso, para multiplicar B_2 por A_1 tenemos que dividir a B_2 por bloques

$$B_2 = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right]$$

Por tanto, tenemos

$$B_2 A_1 = \begin{array}{r|c|c|} & 1 & n-1 \\ \hline 1 & B_{11} & B_{12} \\ \hline n-1 & B_{21} & B_{22} \end{array} \begin{array}{|c|c|l} 1 & n-1 & \\ \hline * & * \cdots * & 1 \\ \hline 0 & \bar{A}_1 & n-1 \end{array} = \begin{array}{|c|cccc|l} 1 & & n-1 & & & \\ \hline * & * & * & \cdots & * & 1 \\ \hline 0 & * & * & \cdots & * & \\ 0 & 0 & * & \cdots & * & \\ \vdots & \vdots & & & \vdots & \\ 0 & 0 & * & \cdots & * & n-1 \end{array} = A_2$$

Como queremos que el primer renglón de A_1 no se altere, nos conviene poner $B_{11} = I_1$ y $B_{12} = 0_{1\times(n-1)}$. B_{21} queda forzado a ser $B_{21} = 0_{(n-1)\times 1}$ si llamamos a B_{22}, $B_{22} = \bar{B}_2$ entonces

$$B_2 = \begin{array}{cc} \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & \bar{B}_2 \end{array}\right] & \begin{array}{l} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & \quad n-1 \end{array} & \end{array}$$

$$\therefore \quad B_2 A_1 = \begin{array}{cc} \begin{array}{r} 1 \\ n-1 \end{array} & \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & \bar{B}_2 \end{array}\right] \\ & \begin{array}{cc} 1 & \quad n-1 \end{array} \end{array} \begin{array}{cc} \left[\begin{array}{c|c} * & * \cdots * \\ \hline 0 & \bar{A}_1 \end{array}\right] & \begin{array}{l} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & \quad n-1 \end{array} & \end{array} = \begin{array}{cc} \left[\begin{array}{c|c} * & * \cdots * \\ \hline 0 & \bar{B}_2\bar{A}_1 \end{array}\right] & \begin{array}{l} 1 \\ n-1 \end{array} \\ \begin{array}{cc} 1 & \quad n-1 \end{array} & \end{array}$$

De aquí vemos que $\bar{B}_2$ tiene ser tal que

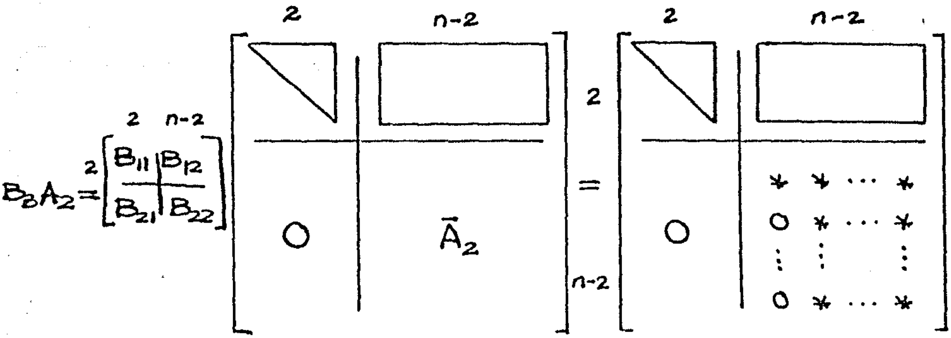
$$\bar{B}_2 \bar{A}_1 = \begin{array}{cc} \left[\begin{array}{ccccc} * & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ 0 & * & \cdots & * & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & * & \cdots & * & * \end{array}\right] & n-1 \\ n-1 & \end{array} = \begin{array}{cc} \left[\begin{array}{c|c} * & * * \cdots * \\ \hline 0 & \bar{A}_2 \end{array}\right] & \begin{array}{l} 1 \\ n-2 \end{array} \\ \begin{array}{cc} 1 & \quad n-2 \end{array} & \end{array}$$

Es decir, al escoger $\bar{B}_2$ estamos resolviendo otra vez el problema de hacer ceros abajo de un elemento, pero en este caso las matrices con las que estamos trabajando son de una dimensión menor.

Veremos ahora qué forma tiene la matriz B_3.

$$\text{Sea} \quad B_3 = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right]$$

Por lo tanto



Para que los dos primeros renglones de A_2 no se modifiquen necesitamos que $B_{11} = I_2$ y que $B_{12} = O_{2\times(n-2)}$

Además, B_{21} está forzado a ser $O_{(n-2)\times 2}$

$$\therefore \quad B_3 = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & \bar{B}_3 \end{array}\right] \qquad \text{donde}$$

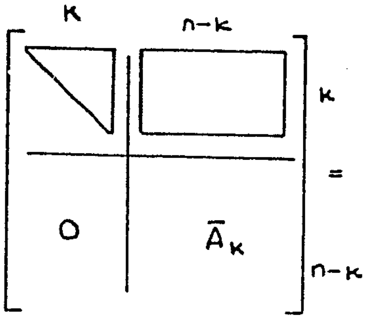
$$\bar{B}_3 \bar{A}_2 = \begin{array}{c} \left[\begin{array}{cccc} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{array}\right] n-2 \\ n-2 \end{array}$$

En general, en el k-ésimo paso B_k es de la forma

$$B_K = \begin{array}{cc} \left[\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & \bar{B}_k \end{array}\right] & \begin{array}{l} k-1 \\ n-(k-1) \end{array} \\ \begin{array}{cc} k-1 & \quad n-(k-1) \end{array} & \end{array}$$

y tal que

$$B_k A_{k-1} =$$



$$= A_K$$

$$\bar{B}_k \bar{A}_{k-1} = \left[\begin{array}{c|ccc} * & * & \cdots & * \\ \hline 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{array}\right] = \begin{array}{cc} \left[\begin{array}{c|c} * & * \cdots * \\ \hline 0 & \bar{A}_k \end{array}\right] & \begin{array}{l} 1 \\ n-k \end{array} \\ \begin{array}{cc} 1 & \quad n-k \end{array} & \end{array}$$

Hemos visto que el problema de encontrar una matriz B tal que $BA = [\triangledown]$ se reduce a encontrar matrices $\bar{B}_k$ tal que

$$\bar{B}_k \bar{A}_{k-1} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

donde $\bar{A}_{k-1}$ es una matriz de orden $n-(k-1)$, y la $\bar{B}_k$ que buscamos también lo es.

La matriz B que buscamos será $B = B_{n-1}\, B_{n-2} \cdots B_2\, B_1$ y tendremos

$$BA = A_{n-1}.$$

Nos queda ahora el siguiente

PROBLEMA

Dada una matriz $A_{m\times n}$ encontrar una matriz $B_{m\times m}$ tal que

$$BA = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \cdots & * \end{bmatrix}$$

Supongamos $A = [a_1 | a_2 | \cdots | a_n]$ con $a_1 \neq 0$, ya que si $a_1 = 0$ entonces no tenemos nada por hacer.

Sea $|\alpha_{k1}| = \max\{|\alpha_{11}|, |\alpha_{21}|, \ldots, |\alpha_{m1}|\}$

Como estamos suponiendo $a_1 \neq 0$, entonces $\alpha_{k1} \neq 0$.

Vamos a intercambiar el primer renglón con el k-ésimo multiplicando a A por la matriz de permutación P_{1k}:

$$P_{1k} A = \begin{bmatrix} r_k^t \\ r_2^t \\ \vdots \\ r_{k-1}^t \\ r_1^t \\ r_{k+1}^t \\ \vdots \\ r_{m-1}^t \\ r_m^t \end{bmatrix} = \bar{A}$$

$$\bar{A} = [\bar{\alpha}_{ij}] = [\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n] = \begin{bmatrix} \bar{r}_1^t \\ \bar{r}_2^t \\ \vdots \\ \bar{r}_m^t \end{bmatrix}$$

$\therefore \ \bar{a}_1 \neq 0$ y $\bar{\alpha}_{11} \neq 0$.

Sea $\bar{B} = \begin{bmatrix} \bar{s}_1^t \\ \bar{s}_2^t \\ \vdots \\ \bar{s}_m^t \end{bmatrix}$

Encontraremos $\bar{s}_1^t, \bar{s}_2^t, \ldots, \bar{s}_m^t$ tal que

$$\bar{B}\bar{A} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

Tenemos que

$$\bar{B}\bar{A} = \begin{bmatrix} \bar{s}_1^t \\ \bar{s}_2^t \\ \vdots \\ \bar{s}_m^t \end{bmatrix} \bar{A} = \begin{bmatrix} \bar{s}_1^t \bar{A} \\ \bar{s}_2^t \bar{A} \\ \vdots \\ \bar{s}_m^t \bar{A} \end{bmatrix} = \bar{B} \begin{bmatrix} \bar{r}_1^t \\ \bar{r}_2^t \\ \vdots \\ \bar{r}_m^t \end{bmatrix}$$

Ahora

$$\bar{s}_1^t A = \bar{\beta}_{11} \bar{r}_1^t + \bar{\beta}_{12} \bar{r}_2^t + \cdots + \bar{\beta}_{1m} \bar{r}_m^t$$

Como en el primer renglón no nos interesa hacer ceros, no tiene ningún caso cambiarlo.

$\therefore$ Nos conviene escoger

$$\bar{\beta}_{11} = 1$$

$$\bar{\beta}_{12} = \bar{\beta}_{13} = \cdots \bar{\beta}_{1m} = 0$$

$$\therefore \quad \bar{\bar{s}}_1^t = [1 \ 0 \cdots 0]$$

Para el segundo renglón tenemos

$$\bar{\bar{s}}_2^t \bar{A} = \bar{\beta}_{21} \bar{r}_1^t + \bar{\beta}_{22} \bar{r}_2^t + \cdots + \bar{\beta}_{2m} \bar{r}_m^t$$

ahora, $\bar{r}_1^t = [\bar{\alpha}_{11} \ \bar{\alpha}_{12} \cdots \bar{\alpha}_{1n}]$

$\bar{r}_2^t = [\bar{\alpha}_{21} \ \bar{\alpha}_{22} \cdots \bar{\alpha}_{2n}]$

$$\therefore \quad \bar{\beta}_{21} \bar{r}_1^t = [\bar{\beta}_{21} \bar{\alpha}_{11} \quad \bar{\beta}_{21} \bar{\alpha}_{12} \quad \cdots \quad \bar{\beta}_{21} \alpha_{1n}]$$

$$\bar{\beta}_{22} \bar{r}_2^t = [\bar{\beta}_{22} \bar{\alpha}_{21} \quad \bar{\beta}_{22} \bar{\alpha}_{22} \cdots \bar{\beta}_{22} \bar{\alpha}_{2n}]$$

$$\therefore \bar{\beta}_{21} \bar{r}_1^t + \bar{\beta}_{22} \bar{r}_2^t = [\bar{\beta}_{21} \bar{\alpha}_{11} + \bar{\beta}_{22} \bar{\alpha}_{21} \quad * \ * \cdots *]$$

De aquí vemos que si podemos hacer $\bar{\beta}_{21} \bar{\alpha}_{11} + \bar{\beta}_{22} \bar{\alpha}_{21} = 0$, entonces podemos hacer

$\bar{\bar{s}}_2^t \bar{A} = \bar{\beta}_{21} \bar{r}_1^t + \bar{\beta}_{22} \bar{r}_2^t$, tomando

$$\bar{\beta}_{23} = \bar{\beta}_{24} = \cdots = \bar{\beta}_{2m} = 0$$

En efecto, podemos escoger $\bar{\beta}_{21}$ y $\bar{\beta}_{22}$ tales que $\bar{\beta}_{21}\bar{\alpha}_{11} + \bar{\beta}_{22}\bar{\alpha}_{21} = 0$

ya que $\bar{\beta}_{21}\bar{\alpha}_{11} = -\bar{\beta}_{22}\bar{\alpha}_{21}$ y como $\bar{\alpha}_{11} \neq 0$ entonces

$$\bar{\beta}_{21} = -\bar{\beta}_{22}\frac{\bar{\alpha}_{21}}{\bar{\alpha}_{11}}$$

Escojamos $\bar{\beta}_{22} = 1$

$$\therefore \quad \bar{\beta}_{21} = -\frac{\bar{\alpha}_{21}}{\bar{\alpha}_{11}}$$

$$\therefore \quad \bar{s}_2^t = \left[\bar{\beta}_{21} \;\; \bar{\beta}_{22} \;\; \cdots \;\; \dot{\beta}_{2m}\right]$$

$$\therefore \quad \bar{s}_2^t = \left[-\frac{\bar{\alpha}_{21}}{\bar{\alpha}_{11}} \;\; 1 \;\; 0 \;\; \cdots \;\; 0\right]$$

Similarmente se puede ver que

$$\bar{s}_3^t = \left[-\frac{\bar{\alpha}_{31}}{\bar{\alpha}_{11}} \;\; 0 \;\; 1 \;\; 0 \;\; \cdots \;\; 0\right]$$

$$\vdots$$

$$\bar{s}_m^t = \left[-\frac{\bar{\alpha}_m}{\bar{\alpha}_{11}} \;\; 0 \;\; 0 \cdots 0 \cdots 1\right] \cdots$$

Por lo tanto la matriz $\bar{B}$ es de la forma

$$\bar{B} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\bar{\alpha}_{21}/\bar{\alpha}_{11} & 1 & 0 & \cdots & 0 \\ -\bar{\alpha}_{31}/\bar{\alpha}_{11} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & & \\ -\bar{\alpha}_{m1}/\bar{\alpha}_{11} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

OBSERVAR que $\bar{B}$ es una matriz triangular inferior.

Como se puede ver, el objeto de permutar el primer renglón con el k-ésimo fue asegurarnos de que el elemento $\bar{\alpha}_{11}$ fuera distinto de cero (para poder dividir entre $\bar{\alpha}_{11}$). Esta operación era necesaria únicamente en el caso en que $\alpha_{11} = 0$. Sin embargo, en la práctica lo más eficiente es tener en el lugar de α_{11} al elemento de a_1 cuyo valor absoluto es máximo.

La matriz B que buscamos es el producto de $\bar{B}$ y P_{1k}, ya que

$$\bar{B}\bar{A} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

pero como $\bar{A} = P_{IK}A$ tenemos que

$$(\bar{B}P_{IK})\,A = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

EJEMPLO

Llevar a la forma triangular a la matriz

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

$$P_{13}A = \begin{bmatrix} 4 & 5 & -4 \\ 3 & 0 & -3 \\ 0 & 2 & -1 \end{bmatrix} \quad ; \quad \bar{B}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -3/4 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\bar{B}_1 P_{13} A = \begin{bmatrix} 4. & 5 & -4 \\ 0 & -15/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

CASO ESPECIAL.

Supongamos que al llevar a A a la forma triangular superior no es necesario intercambiar renglones.

Veremos qué forma tienen las matrices que transforman a la matriz A en $A_{n-1} =$ ◹.

Por lo que acabamos de ver, B_1 es de la forma

$$B_1 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \lambda_{21} & 1 & 0 & \cdots & 0 \\ \lambda_{31} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

donde $\lambda_{i1} = -\dfrac{\alpha_{i1}}{\alpha_{11}} \qquad i = 2, \ldots, n$.

$$B_1 A = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \cdots & * \end{bmatrix} = A_1 = \begin{bmatrix} \alpha'_{11} & \alpha'_{12} & \cdots & \alpha'_{1n} \\ 0 & \alpha'_{22} & \cdots & \alpha'_{2n} \\ \vdots & & & \\ 0 & \alpha'_{n2} & \cdots & \alpha'_{nn} \end{bmatrix}$$

En el segundo paso tendremos que

$$B_2 = \begin{bmatrix} I_1 & 0 \\ 0 & \bar{B}_2 \end{bmatrix}$$ donde

$$\bar{B}_2 \cdot \begin{bmatrix} \alpha'_{22} & \alpha'_{23} & \cdots & \alpha'_{2n} \\ \alpha'_{32} & \alpha'_{33} & \cdots & \alpha'_{3n} \\ \vdots & & & \\ \alpha'_{n2} & \alpha'_{n3} & \cdots & \alpha'_{nn} \end{bmatrix} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ & & & \\ 0 & * & \cdots & * \end{bmatrix}$$

Por lo tanto $\bar{B}_2$ es de la forma

$$\bar{B}_2 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \lambda_{32} & 1 & 0 & \cdots & 0 \\ \lambda_{i2} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n2} & 0 & 0 & & 1 \end{bmatrix}, \quad \lambda_{i2} = -\frac{\alpha'_{i2}}{\alpha'_{22}}$$ para

$i = 3, 4, \ldots, n$

Por lo tanto tenemos

$$B_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_{32} & 1 & 0 & \cdots & 0 \\ 0 & \lambda_{42} & 0 & 1 & \cdots & 0 \\ \vdots & & & & & \\ 0 & \lambda_{n2} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$B_2 A_1 = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix} = A_2 = \begin{bmatrix} \alpha^2_{11} & \alpha^2_{12} & \alpha^2_{13} & \cdots & \alpha^2_{1n} \\ 0 & \alpha^2_{22} & \alpha^2_{23} & \cdots & \alpha^2_{2n} \\ 0 & 0 & \alpha^2_{33} & \cdots & \alpha^2_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \alpha^2_{n3} & \cdots & \alpha^2_{nn} \end{bmatrix}$$

Para el tercer paso

$$B_3 = \begin{bmatrix} I_2 & 0 \\ 0 & \overline{B}_3 \end{bmatrix} \qquad \text{con}$$

$$\overline{B}_3 \cdot \begin{bmatrix} \alpha^2_{33} & \alpha^2_{34} \cdots & \alpha^2_{3n} \\ \alpha^2_{43} & \alpha^2_{44} \cdots & \alpha^2_{4n} \\ \vdots & \vdots & \vdots \\ \alpha^2_{n3} & \alpha^2_{n4} \cdots & \alpha^2_{nn} \end{bmatrix} = \begin{bmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix}$$

Por lo tanto $\bar{B}_3$ es de la forma

$$\bar{B}_3 = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \lambda_{43} & 1 & 0 & \cdots & 0 \\ \lambda_{53} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n3} & 0 & 0 & & 1 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_{43} & 1 & 0 & \cdots & 0 \\ 0 & 0 & \lambda_{53} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \lambda_{n3} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

En general, en el k-ésimo paso vamos a tener

$$B_K = \left[\begin{array}{c|ccccc} I_{k-1} & & & 0 & & \\ \hline & 1 & 0 & 0 & \cdots & 0 \\ & \lambda_{k+1,k} & 1 & 0 & \cdots & 0 \\ 0 & \lambda_{k+2,k} & 0 & 1 & \cdots & 0 \\ & \vdots & \vdots & \vdots & & \vdots \\ & \lambda_{nk} & 0 & 0 & & 1 \end{array}\right] n-(k-1)$$

Por lo tanto tenemos

$$B_{n-1} B_{n-2} \cdots B_2 B_1 A = \begin{bmatrix} \diagdown \end{bmatrix} = U$$

donde cada una de las $B's$ es una matriz triangular inferior.

Si llamamos $L_i = B_i$ para $i = 1, \ldots, n-1$ entonces tenemos que :

$$L_{n-1} L_{n-2} \cdots L_2 L_1 A = U$$

Nos interesa interpretar esta igualdad, y escribirla de manera diferente. Para ello necesitamos observar algunas propiedades de las matrices L_i.

Notemos que cada una de las operaciones efectuadas con cada una de las matrices L_i se puede invertir, es decir, podemos encontrar una L_i^{-1} tal que el efecto combinado $L_i^{-1} L_i$ sea el de dejar la matriz sobre la que opera en forma inalterada.

Empezemos por observar L_1 :

$$A_1 = L_1 A = \begin{bmatrix} 1 & & & & \\ \lambda_{21} & 1 & & & \\ \lambda_{31} & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \\ \lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix} A = \begin{bmatrix} r_1^t \\ \lambda_{21} r_1^t + r_2^t \\ \lambda_{31} r_1^t + r_3^t \\ \\ \lambda_{n1} r_1^t + r_n^t \end{bmatrix}$$

Por lo tanto, si quisieramos recuperar a A a partir de A_1, lo único que tenemos que hacer es:

1) restar al 2° renglón de A_1 λ_{21} veces el primer renglón de A_1

2) restar al 3^er renglón de A_1 λ_{31} veces el primer renglón de A_1

⋮

n-1) restar al n-ésimo renglón de A_1 λ_{n1} veces el primer renglón de A_1

pero ésto es precisamente lo que hace la matriz L_1, solo que en vez de restar suma.

Podemos concluir que L_1^{-1} debe de obtenerse a partir de L_1, cambiando los signos de $\lambda_{21}, \lambda_{31}, \ldots, \lambda_{n1}$, es decir:

$$L_1^{-1} = \begin{bmatrix} 1 & & & & \\ -\lambda_{21} & 1 & & & \\ -\lambda_{31} & 0 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ -\lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

Para confirmar lo anterior observemos que :

$$L_1^{-1} L_1 = \begin{bmatrix} 1 & & & & \\ -\lambda_{21} & 1 & & & \\ -\lambda_{31} & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \\ -\lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ \lambda_{21} & 1 & & & \\ \lambda_{31} & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \\ \lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ -\lambda_{21}+\lambda_{21} & 1 & 0 & \cdots & 0 \\ -\lambda_{31}+\lambda_{31} & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\lambda_{n1}+\lambda_{n1} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 \end{bmatrix} = I_n$$

y por lo tanto

$$L_1^{-1} A_1 = L_1^{-1} (L_1 A) = (L_1^{-1} L_1) A = I_n A = A$$

que es lo que queríamos.

Similarmente obtenemos que si

$$L_2 = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & \lambda_{32} & 1 & & \\ \vdots & \vdots & \vdots & & \\ 0 & \lambda_{n2} & 0 & \cdots & 1 \end{bmatrix}, \quad L_2^{-1} = \begin{bmatrix} 1 & & & & \\ 0 & 1 & & & \\ 0 & -\lambda_{32} & 1 & & \\ \vdots & \vdots & \vdots & & \\ 0 & -\lambda_{n2} & 0 & \cdots & 1 \end{bmatrix}$$

$$L_3 = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & 0 & 1 & & & \\ 0 & 0 & \lambda_{43} & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & \lambda_{n3} & 0 & \cdots & 1 \end{bmatrix}, \quad L_3^{-1} = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & 0 & 1 & & & \\ 0 & 0 & -\lambda_{43} & 1 & & \\ \vdots & \vdots & & \vdots & \ddots & \\ 0 & 0 & -\lambda_{n3} & 0 & \cdots & 1 \end{bmatrix},$$

etc.

$$L_{n-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & & \\ \vdots & \vdots & & \\ 0 & 0 & 1 & \\ 0 & 0 & \lambda_{n,n-1} & 1 \end{bmatrix}, \quad L_{n-1}^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & & \\ \vdots & \vdots & & \\ 0 & 0 & 1 & \\ 0 & 0 & -\lambda_{n,n-1} & 1 \end{bmatrix}$$

entonces

$$L_2^{-1}L_2 = I_n \quad , L_3^{-1}L_3 = I_n, \dots, \quad L_{n-1}^{-1}L_{n-1} = I_n \quad .$$

Utilizaremos lo anterior para reescribir la relación

$$L_{n-1}L_{n-2}\cdots L_2L_1A = U$$

de otra manera muy sugestiva.

Como acabamos de ver, para cada L_i existe una matriz L_i^{-1} tal que

$$L_i^{-1}L_i = I_n \qquad i = 1, 2, \dots, n-1$$

Por lo tanto

$$(L_{n-1}^{-1}L_{n-1})L_{n-2}\cdots L_2L_1A = I_nL_{n-2}\cdots L_2L_1A = L_{n-1}^{-1}U$$
$$(L_{n-2}^{-1}L_{n-2})L_{n-3}\dots L_2L_1A = I_nL_{n-3}\cdots L_2L_1A = L_{n-2}^{-1}L_{n-1}^{-1}U$$
$$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$$
$$(L_1^{-1}L_1)A = I_nA = L_1^{-1}L_2^{-1}\cdots L_{n-2}^{-1}L_{n-1}^{-1}U$$

$$\therefore \quad A = L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1}U$$

Como cada una de las matrices L_i es triangular inferior, vamos a tener que

$$L_1^{-1}L_2^{-1}\cdots L_{n-1}^{-1} \quad \text{es triangular inferior}$$

Por lo tanto a la matriz A la podemos factorizar de la siguiente manera:

$$A = LU$$

NOTA :

Hasta ahora, nuestro principal interés ha sido ilustrar el poder de las operaciones matriciales como herramienta.

Por ejemplo, supongamos que tenemos el sistema

$$Ax = b$$

y que podemos factorizar

$$A = LU$$

entonces el sistema $Ax = b$ es equivalente a

$$LUx = b \qquad (*)$$

Si hacemos $y = Ux$, entonces obtenemos

$$Ly = b .$$

al llegar a (*) prácticamente ya tenemos resuelto el sistema $Ax=b$, ya que los sistemas de ecuaciones cuyas matrices asociadas son triangulares son especialmente fáciles de manejar.

Para el caso en que sea necesario intercambiar varios renglones de A, entonces los podemos intercambiar desde un principio por medio de una matriz de permutación P, y obtener la factorización

$$PA = \tilde{L}\tilde{U}.$$

CAPITULO II

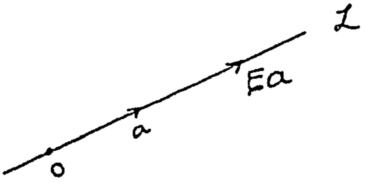
DEPENDENCIA LINEAL EN $\mathbb{R}^m$

DEPENDENCIA LINEAL EN $\mathbb{R}^m$

MOTIVACION

Supongamos que estamos trabajando con vectores en $\mathbb{R}^3$.

Si $\mathcal{L}$ es una recta que pasa por el origen y $a \in \mathbb{R}^3$ está sobre la recta, entonces todos los puntos sobre $\mathcal{L}$ son de la forma $y = \xi a$, para alguna $\xi \in \mathbb{R}$.



Es decir, $\mathcal{L} = \{ y \in \mathbb{R}^3 : y = \xi a, \xi \in \mathbb{R}, a \neq 0 \}$

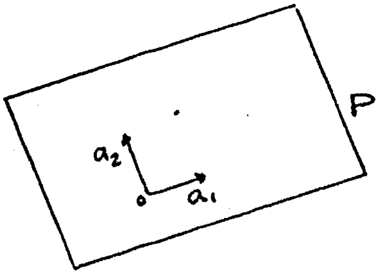
Observar que el vector a no está unívocamente determinado, ya que si $y \in \mathcal{L} \Rightarrow$

$y = \xi a$ para alguna $\xi \in \mathbb{R} \Rightarrow$

$y = (\frac{\xi}{\lambda})(\lambda a) = \xi' a' \qquad (\lambda \neq 0)$

Por lo tanto vemos que si $\lambda \neq 0$, podemos sustituir a a por (λa) y generar la misma recta $\mathcal{L}$ $\therefore$ el vector a no está unívocamente determinado por la recta $\mathcal{L}$.

Ahora, supongamos que tenemos un plano P y dos vectores no colineales a_1 y a_2 sobre P. Entonces cualquier vector del plano se puede obtener a partir de a_1 y a_2, si exigimos que P pase por el orígen.



Es decir

$$P = \left\{ y \in \mathbb{R}^3 : y = \xi_1 a_1 + \xi_2 a_2 \,,\ a_1, a_2 \text{ no colineales} \right\}$$

Observemos también que en vez de a_1 y a_2 podríamos haber usado cualesquiera dos vectores no colineales del plano - por ejemplo en vez de a_1 podríamos tomar $\lambda_1 a_1$ para cualquier $\lambda_1 \neq 0$, y en vez de a_2 podríamos tomar $\lambda_2 a_2$ para $\lambda_2 \neq 0$.

De aquí vemos que los vectores que determinan un plano no están unívocamente determinados.

Ahora, supongamos que estamos trabajando con vectores en $\mathbb{R}^m$. Entonces, al igual que para $\mathbb{R}^3$ el conjunto formado por los multiplos de un vector no nulo es una recta, y el conjunto de vectores de la forma $\xi_1 a_1 + \xi_2 a_2$ es un plano si a_1 y a_2 no son colineales.

¿Qué pasa si en lugar de tomar uno ó dos vectores tomamos n vectores de $\mathbb{R}^m$?

En algebra Lineal nos interesa estudiar a los subconjuntos $\mathcal{S} \subseteq \mathbb{R}^m$ de la forma

$$\mathcal{S} = \{ y \in \mathbb{R}^m : \ y = \xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n \ ; \ \xi_1, \xi_2, \dots, \xi_n \in \mathbb{R} \ ; \ a_1, a_2, \dots, a_n \in \mathbb{R}^m \}$$

$$= \{ y \in \mathbb{R}^m : \ y = \sum_{j=1}^{n} \xi_j a_j \ ; \ \xi_j \in \mathbb{R}, \ a_j \in \mathbb{R}^m \}$$

DEFINICION

Sean $a_1, a_2, \dots, a_n \in \mathbb{R}^n$. Decimos que $y \in \mathbb{R}^m$ es una COMBINACION LINEAL de $a_1, a_2, \dots, a_n$ si existen $\xi_1, \xi_2, \dots, \xi_n \in \mathbb{R}$ tales que y se puede expresar de la forma

$$y = \xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n$$

Es decir,

y es una combinación lineal de $a_1, a_2, \ldots, a_n$. $a_{\bar{n}}$ si y solo si existe una $x \in \mathbb{R}^n$ tal que

$$y = Ax$$

donde $A = [a_1 | a_2 | \cdots | a_n]$ es la matriz cuyas columnas son los vectores $a_1, a_2, \ldots, a_n$ y

$$x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

La generalización de los conceptos de línea, plano, etc., que pasan por el origen está dada por

DEFINICION

Sea $\mathcal{S} \subseteq \mathbb{R}^m$. Decimos que $\mathcal{S}$ es un SUBESPACIO LINEAL de $\mathbb{R}^m$ si

1) $\mathcal{S} \neq \emptyset$
2) $s_1, s_2 \in \mathcal{S} \Rightarrow s_1 + s_2 \in \mathcal{S}$
3) $\lambda \in \mathbb{R}, \; s \in \mathcal{S} \Rightarrow \lambda s \in \mathcal{S}$

EJEMPLOS

1) $\mathbb{R}^m$ es un subespacio lineal de $\mathbb{R}^m$.

2) El conjunto $\{0\}$, cuyo único elemento es el vector cero es el subespacio TRIVIAL de $\mathbb{R}^m$.

RESULTADO

Sean $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$. Entonces el conjunto de todas las combinaciones lineales de $a_1, a_2, \ldots, a_n$, al cual denotaremos por $[a_1, a_2, \ldots, a_n]$ es un subespacio lineal de $\mathbb{R}^m$

Demostración:

Antes de proseguir, hay que observar que $[a_1|a_2|\cdots|a_n]$ es la matriz cuyas columnas son $a_1, a_2, \ldots, a_n$, y que por lo tanto no hay que confundir $[a_1, a_2, \ldots, a_n]$ con $[a_1|a_2|\cdots|a_n]$.

Supongamos $y_1, y_2 \in [a_1, a_2, \ldots, a_n]$

$\therefore \; y_1 = Ax_1, \; y_2 = Ax_2$ donde $A = [a_1|a_2|\cdots|a_n]$

$\therefore \; y_1 + y_2 = Ax_1 + Ax_2 = A(x_1 + x_2) \in [a_1, a_2, \ldots, a_n]$.

Si $y \in [a_1, a_2, \ldots, a_n]$ entonces $y = Ax$

$\therefore \; \lambda y = \lambda Ax = A\lambda x \in [a_1, a_2, \ldots, a_n]$.

$\therefore$ $[a_1, a_2, \ldots, a_n]$ es un subespacio lineal de $\mathbb{R}^m$. ▪

DEFINICION

Si $\mathcal{S}$ es un subespacio lineal de $\mathbb{R}^m$ y $\mathcal{S} = [a_1, a_2, \ldots, a_n]$ entonces diremos que $[a_1, a_2, \ldots, a_n]$ es <u>UN</u> CONJUNTO DE GENERADORES DE $\mathcal{S}$.

EJEMPLO

Hasta ahora hemos visto ejemplos en los cuales los vectores que generan un subespacio lineal están dados explícitamente. Un ejemplo en el cual ésto no sucede es el siguiente:

$$\mathcal{S} = [s \in \mathbb{R}^n : r^t s = 0 \quad \text{para alguna} \quad r \in \mathbb{R}^n]$$

es un subespacio lineal de $\mathbb{R}^n$, ya que

si $s_1, s_2 \in \mathcal{S}$ entonces

$$r^t(s_1 + s_2) = r^t s_1 + r^t s_2 = 0 + 0 = 0$$

$\therefore$ $s_1 + s_2 \in \mathcal{S}$

Si $s \in \mathcal{S}$ entonces $r^t(\lambda s) = \lambda(r^t s) = \lambda \cdot 0 = 0$

$\therefore$ $\lambda s \in \mathcal{S}$ para $\lambda \in \mathbb{R}$.

$\therefore$ $\mathcal{S}$ es un subespacio lineal.

OBSERVAR que $\mathcal{S}$ es el conjunto de vectores ortogonales a r.

Es decir, dado $r \in \mathbb{R}^n$, entonces el conjunto de vectores ortogonales a r es un subespacio lineal de $\mathbb{R}^n$.

PROPIEDAD

Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ dos subespacios lineales de $\mathbb{R}^n$. Entonces $\mathcal{S}_1 \cap \mathcal{S}_2$ es un subespacio lineal de $\mathbb{R}^n$.

Demostración:

Supongamos $s_1, s_2 \in \mathcal{S}_1 \cap \mathcal{S}_2 \Rightarrow$
$s_1, s_2 \in \mathcal{S}_1$ y $s_1, s_2 \in \mathcal{S}_2$

$\therefore$ $s_1 + s_2 \in \mathcal{S}_1$ y $s_1 + s_2 \in \mathcal{S}_2$
$\therefore$ $s_1 + s_2 \in \mathcal{S}_1 \cap \mathcal{S}_2$

Y si $s \in \mathcal{S}_1 \cap \mathcal{S}_2$ entonces $s \in \mathcal{S}_1$ y $s \in \mathcal{S}_2$
$\therefore$ $\lambda s \in \mathcal{S}_1$ y $\lambda s \in \mathcal{S}_2$
$\therefore$ $\lambda s \in \mathcal{S}_1 \cap \mathcal{S}_2$

$\therefore$ $\mathcal{S}_1 \cap \mathcal{S}_2$ es un subespacio lineal de $\mathbb{R}^n$ ▫

EJEMPLO

Si $\mathcal{S}_1 = \{ x \in \mathbb{R}^n : r_1^t x = 0 \quad ; \ r_1 \in \mathbb{R}^n \}$

$\mathcal{S}_2 = \{ x \in \mathbb{R}^n : r_2^t x = 0 \ ; \ r_2 \in \mathbb{R}^n \}$

entonces $\mathcal{S}_1 \cap \mathcal{S}_2 = \{ x \in \mathbb{R}^n : r_1^t x = 0 \ \text{ y } \ r_2^t x = 0 \}$ es un subespacio lineal de $\mathbb{R}^m$.

Como consecuencia de la propiedad también tenemos que si $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_m$ son subespacios lineales de $\mathbb{R}^n$ entonces

$\mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_m$ es un subespacio lineal

de $\mathbb{R}^n$.

EJEMPLO

Sean

$$\begin{aligned}
\mathcal{S}_1 &= \{ x \in \mathbb{R}^n : r_1^t x = 0 \ ; \ r_1 \in \mathbb{R}^n \} \\
\mathcal{S}_2 &= \{ x \in \mathbb{R}^n : r_2^t x = 0 \ ; \ r_2 \in \mathbb{R}^n \} \\
&\ \vdots \qquad\qquad\qquad\qquad\qquad\quad \vdots \\
\mathcal{S}_m &= \{ x \in \mathbb{R}^n : r_m^t x = 0 \ ; \ r_m \in \mathbb{R}^n \}
\end{aligned}$$

Entonces $\mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_m$ es un subespacio lineal de $\mathbb{R}^n$.

Denotemos a $\mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_m$ por $\mathcal{S}$. Entonces $\mathcal{S}$ está dado por el sistema de ecuaciones

$$\begin{gathered} r_1^t x = 0 \\ r_2^t x = 0 \\ \vdots \\ r_m^t x = 0 \end{gathered}$$

Si A es la matriz $m \times n$ dada por

$$A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$$

entonces $x \in \mathcal{S} \Leftrightarrow Ax = 0$. Es decir

$\mathcal{S} = \{x \in \mathbb{R}^n : Ax = 0\}$ es un subespacio lineal de $\mathbb{R}^n$:

El conjunto de soluciones del sistema $Ax = 0$ es un subespacio lineal de $\mathbb{R}^n$.

Resumiendo, dada una matriz A $m\times n$, entonces tenemos los siguientes subespacios asociados a ella:

1) El subespacio de $\mathbb{R}^m$ generado por las columnas de A.

2) El subespacio de $\mathbb{R}^n$ dado por el conjunto de soluciones de $AX=0$.

Estos dos subespacios son tan importantes que merecen nombres especiales:

DEFINICION

Sea A una matriz $m\times n$. Entonces

1) La IMAGEN ó RANGO de A, denotado por Im (A) ó R(A), es el subespacio de $\mathbb{R}^m$ generado por las columnas de A

$$Im(A)=[a_1,a_2,\dots,a_n]$$

2) El SUBESPACIO NULO, NUCLEO ó KERNEL de A, denotado por N(A) ó Ker (A), es el conjunto de $x\in\mathbb{R}^n$ tal que $Ax=0$

$$N(A) = \{ X \in \mathbb{R}^n : \; AX = 0 \}$$

Hemos visto que dada A $m \times n$, entonces el rango y el núcleo de A son subespacios lineales. Es natural que nos planteemos:

¿ Todo subespacio lineal de $\mathbb{R}^m$ se puede ver como el núcleo ó rango de una matriz ? Es decir, ¿ son todos los subespacios de $\mathbb{R}^m$ finitamente generados ?

Para poder hacer ver que los subespacios de $\mathbb{R}^m$ sí son finitamente generados necesitamos introducir algunos conceptos ; estos conceptos los motivaremos en base a las siguientes propiedades de los subespacios finitamente generados.

PROPIEDADES

1) $[0] = \{0\}$

A partir del vector cero no podemos generar ningún vector diferente del vector cero.

2) $[\lambda a] = [a] \qquad \forall\, a \in \mathbb{R}^m \;, \; \lambda \neq 0$

La recta generada por el vector a y la recta generada por cualquier múltiplo no nulo de a son la misma.

3) $a \in [a] \quad \forall a \in \mathbb{R}^m$

El vector a siempre está en la recta generada por el mismo.

4) $[a,0] = [0,a] = [a]$

El "plano" generado por el vector a y el vector 0 no es en realidad un plano - sino que, se reduce a la recta generada por el vector a (¡y a un punto si $a=0$). O sea que no generamos más vectores si consideramos las combinaciones lineales de a y 0 que si únicamente tomaramos combinaciones lineales de a.

5) $[a,\lambda a] = [\lambda a, a] = [a]$

El "plano" generado por a y cualquier múltiplo de a es una recta (la recta determinada por a). O sea que el conjunto de combinaciones de a no crece si consideramos combinaciones lineales de a y un múltiplo de a.

Por lo tanto para estudiar a $[a,\lambda a]$ basta estudiar al conjunto $[a]$, lo cual es más simple ya que únicamente trabajamos con un vector, y no dos.

6) $[a,b] = [b,a] = [a] \Leftrightarrow b \in [a]$

i) Si b está en la recta generada por a, entonces el subespacio generado por a y b no es un plano, sino que se reduce a la recta generada por a.

ii) Si el subespacio lineal generado por los vectores a y b es la recta generada por a, entonces b es un múltiplo de a; es decir, b está en la recta generada por a.

7) $[a_1, a_2, \ldots, a_n] = [a] \Leftrightarrow a_j \in [a] \quad j = 1, 2, \ldots, n$

Es decir, si el subespacio determinado por n vectores se reduce a una recta, entonces los n vectores necesariamente están en la recta.

E inversamente, si tenemos n vectores en una recta, entonces el subespacio lineal generado por los n vectores es la recta. A partir de vectores en un subespacio lineal no podemos generar vectores fuera del subespacio.

Para estudiar a $[a_1, a_2, \ldots, a_n]$ nos conviene fijarnos en $[a]$.

8) $a \in [a_1, a_2, \ldots, a_n] \Rightarrow [a] \subseteq [a_1, a_2, \ldots, a_n]$

O sea que si un vector a está en el subespacio determinado por $a_1, a_2, \ldots, a_n$ entonces cualquier múltiplo de a también está en el hiperplano; es decir, la recta determinada por a está contenida en el hiperplano.

Demostración:

1) $[0] = \{0\}$

$$[0] = \{y \in \mathbb{R}^m : y = \xi \cdot 0 ,\ \xi \in \mathbb{R}\}$$
$$= \{y \in \mathbb{R} : y = 0\} = \{0\} \ \blacksquare$$

2) $[\lambda a] = [a] \qquad \lambda \neq 0$

$$\begin{aligned} y \in [\lambda a],\ \lambda \neq 0 \quad &\Longleftrightarrow \quad y = \xi(\lambda a) \quad , \quad \xi \in \mathbb{R} \\ &\Longleftrightarrow \quad y = (\xi\lambda)a \\ &\Longleftrightarrow \quad y = \xi' a \qquad \xi' \in \mathbb{R} \\ &\Longleftrightarrow \quad y \in [a] \ \blacksquare \end{aligned}$$

3) $a \in [a]$

$a = 1 \cdot a \qquad \therefore \quad a \in [a]$ ■

4) $[a,0] = [0,a] = [a]$

$$\begin{aligned} y \in [a,0] &\iff y = \xi_1 a + \xi_2 \cdot 0 \iff y = \xi_1 a \\ &\iff y = \xi_2 \cdot 0 + \xi_1 \cdot a \iff y \in [0,a] \\ &\iff y \in [a] \end{aligned}$$ ■

5) $[a,\lambda a] = [\lambda a, a] = [a]$

$$\begin{aligned} y \in [a,\lambda a] &\iff y = \xi_1 a + \xi_2 (\lambda a) \\ &\iff y = \xi_2 (\lambda a) + \xi_1 a \\ &\iff y \in [\lambda a, a] \\ &\iff y = (\xi_1 + \xi_2 \lambda) a \\ &\iff y \in [a] \end{aligned}$$ ■

6) $[a,b] = [b,a] = [a] \iff b \in [a]$

$$\begin{aligned} y \in [a,b] &\iff y = \xi_1 a + \xi_2 b \\ &\iff y = \xi_2 b + \xi_1 a \\ &\iff y \in [b,a] \end{aligned}$$

$\therefore \quad [a,b] = [b,a]$ siempre.

$[\Rightarrow$ Supongamos $[a,b] = [a]$.

$b = 0 \cdot a + 1 \cdot b \qquad \therefore \quad b \in [a,b]$

$\therefore \quad b \in [a]$.

$\Leftarrow]$ supongamos $b \in [a]$ $\therefore$ $b = \lambda a$

ahora, $y \in [a, b] \Leftrightarrow y = \xi_1 a + \xi_2 b$

$$\Leftrightarrow y = \xi_1 a + \xi_2 (\lambda a)$$
$$\Leftrightarrow y = (\xi_1 + \xi_2 \lambda) a$$
$$\Leftrightarrow y \in [a] \,.$$

7) $[a_1, a_2, \ldots, a_n] = [a] \Leftrightarrow a_j \in [a]$ $j = 1, 2, \ldots, n$

$[\Rightarrow$ Si $[a_1, a_2, \ldots, a_n] = [a]$ entonces

$$a_j = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + 1 \cdot a_j + \cdots + 0 \cdot a_n$$

$\therefore$ $a_j \in [a_1, a_2, \cdots, a_n]$ $j = 1, 2, \cdots, n$

$\Leftarrow]$ Si $a_j \in [a]$ $j = 1, 2, \ldots, n$ entonces
$a_j = \lambda_j a$, $\lambda_j \in \mathbb{R}$ $j = 1, 2, \cdots, n$ y:

$y \in [a_1, a_2, \cdots, a_n] \Leftrightarrow y = \xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n$

$$\Leftrightarrow y = \xi_1 (\lambda_1 a) + \xi_2 (\lambda_2 a) + \cdots + \xi_n (\lambda_n a)$$
$$\Leftrightarrow y = (\xi_1 \lambda_1) a + (\xi_2 \lambda_2) a + \cdots + (\xi_n \lambda_n) a$$
$$\Leftrightarrow y = (\xi_1 \lambda_1 + \xi_2 \lambda_2 + \cdots + \xi_n \lambda_n) a$$
$$\Leftrightarrow y \in [a]$$

□

8) $a \in [a_1, a_2, \ldots, a_n] \Rightarrow [a] \subseteq [a_1, a_2, \ldots, a_n]$

supongamos $a \in [a_1, a_2, \ldots a_n]$

$\therefore \quad a = \xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n \qquad \xi_j \in \mathbb{R}$

Sea $\quad y \in [a]$

$\therefore \quad y = \lambda a \qquad \lambda \in \mathbb{R}$

$\therefore \quad y = \lambda(\xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n)$

$\therefore \quad y = (\lambda\xi_1)a_1 + (\lambda\xi_2)a_2 + \cdots + (\lambda\xi_n)a_n$

$\therefore \quad y \in [a_1, a_2, \ldots, a_n]$ ∎

PROPIEDADES (PLANOS)

1) $[a_1, a_2] = [a_2, a_1]$

Es decir, el plano generado por a_1 y a_2 es el mismo que el generado por a_2 y a_1

2) $[a_1] \subseteq [a_1, a_2] \quad ; \quad [a_2] \subseteq [a_1, a_2]$

O sea que si tenemos el plano generado por a_1 y a_2, entonces tanto la recta generada por a_1 como la generada por a_2 están contenidas en el plano.

3) $[a_1, a_2] \neq [a_1] \iff a_2 \notin [a_1]$

$[a_1, a_2] \neq [a_2] \iff a_1 \notin [a_2]$

Por el inciso anterior, podríamos escribir la propiedad (3) como

$$[a_1,a_2] \not\subseteq [a_1] \Leftrightarrow a_2 \notin [a_1]$$
$$[a_1,a_2] \not\subseteq [a_2] \Leftrightarrow a_1 \notin [a_2]$$

Es decir, si el plano generado por los vectores a_1 y a_2 contiene vectores que no están en la recta generada por a_1, entonces a_2 no está en la recta generada por a_1.

Inversamente, si a_2 no está en la recta generada por a_1, entonces el conjunto de combinaciones lineales de a_1 y a_2 contiene vectores que no están en la recta generada por a_1.

La afirmación para $[a_2]$ es análoga.

4) $[a_1,a_2,a] = [a_1,a_2] \Leftrightarrow a \in [a_1,a_2]$

Si un vector a está en el plano determinado por dos vectores a_1 y a_2, entonces el subespacio generado por a_1 y a_2 no crece si consideramos el subespacio generado por $\{a_1,a_2,a\}$.

Inversamente, si el subespacio generado por los vectores a_1, a_2 y a

no se "achica" al reducir los generadores a dos vectores (a_1 y a_2) entonces el vector que sobra (a) necesariamente está en el plano determinado por los otros dos.

5) $a_1, a_2 \in [b_1, b_2, \ldots, b_n] \Rightarrow [a_1, a_2] \subseteq [b_1, b_2, \ldots, b_n]$.

Si dos vectores están en un subespacio, entonces el plano determinado por ellos está contenido en el subespacio.

6) $[b_1, b_2, \ldots, b_n] \subseteq [a_1, a_2] \Leftrightarrow b_j \in [a_1, a_2]$

Es decir, si el subespacio generado por n vectores está contenido en un plano (el plano determinado por a_1 y a_2) entonces los n vectores están en ese plano.

E Inversamente, si tenemos n vectores en un plano, entonces el subespacio generado por ellos está totalmente contenido en el plano .- a partir de vectores en un plano no podemos construir vectores fuera del plano.

7) $y \in [a_1, a_2] \iff y = Ax$ donde A es la matriz $m \times 2$ cuyas columnas son a_1 y a_2 y $x \in \mathbb{R}^2$

Vemos que es equivalente estudiar los productos Ax y los conjuntos $[a_1, a_2]$.

Demostración :

1) $[a_1, a_2] = [a_2, a_1]$

$$
\begin{aligned}
y \in [a_1, a_2] &\iff y = \bar{\xi}_1 a_1 + \bar{\xi}_2 a_2 \\
&\iff y = \bar{\xi}_2 a_2 + \bar{\xi}_1 a_1 \\
&\iff y \in [a_2, a_1] \quad \square
\end{aligned}
$$

2) $[a_1] \subseteq [a_1, a_2]$

$[a_2] \subseteq [a_1, a_2]$

$$y \in [a_1] \Rightarrow y = \xi_1 a_1 \Rightarrow y = \xi_1 a_1 + 0 \cdot a_2$$
$$\Rightarrow y \in [a_1, a_2]$$

análogamente $\quad y \in [a_2] \Rightarrow y \in [a_1, a_2]$ $\blacksquare$

3) $[a_1,a_2] \neq [a_1] \Leftrightarrow a_2 \notin [a_1]$.
$[a_1,a_2] \neq [a_2] \Leftrightarrow a_1 \notin [a_2]$

Observar que como consecuencia de la propiedad (2)

$$[a_1,a_2] \neq [a_1] \Leftrightarrow [a_1,a_2] \not\subseteq [a_1]$$

$$[a_1,a_2] \neq [a_1] \Leftrightarrow a_2 \notin [a_1] \ :$$

$[\Rightarrow$ supongamos $a_2 \in [a_1]$ y sea $y \in [a_1,a_2]$.

$$\Rightarrow y = \xi_1 a_1 + \xi_2 a_2$$

Pero $a_2 \in [a_1] \therefore a_2 = \xi a_1$

$$\therefore y = \xi_1 a_1 + \xi_2 (\xi a_1) = (\xi_1 + \xi_2 \xi) a_1$$

$$\therefore y \in [a_1] \quad \therefore [a_1,a_2] \subseteq [a_1]$$

$$\therefore [a_1] \subsetneq [a_1,a_2] \Rightarrow a_2 \notin [a_1]$$

⇐] Supongamos $[a_1, a_2] \subseteq [a_1]$

$\Rightarrow [a_2] \subseteq [a_1, a_2] \subseteq [a_1]$
$\Rightarrow a_2 \in [a_1]$

$\therefore\ a_2 \notin [a_1] \Rightarrow [a_1, a_2] \not\subseteq [a_1]$.

La demostración de la segunda afirmación es análoga. ▪

4) $[a_1, a_2, a] = [a_1, a_2] \Leftrightarrow a \in [a_1, a_2]$

[⇒ Supongamos $[a_1, a_2, a] = [a_1, a_2]$

$\Rightarrow [a] \subseteq [a_1, a_2, a] \subseteq [a_1, a_2]$
$\Rightarrow a \in [a_1, a_2]$.

⇐] Supongamos $a \in [a_1, a_2]$

$\therefore\ a = \lambda_1 a_1 + \lambda_2 a_2$

$y \in [a_1, a_2, a] \Rightarrow y = \xi_1 a_1 + \xi_2 a_2 + \xi a$

$$\begin{aligned} \therefore\ y &= \xi_1 a_1 + \xi_2 a_2 + \xi(\lambda_1 a_1 + \lambda_2 a_2) \\ &= (\xi_1 + \xi\lambda_1) a_1 + (\xi_2 + \xi\lambda_2) a_2 \end{aligned}$$

$\therefore\ [a_1, a_2, a] \subseteq [a_1, a_2]$

Por otro lado siempre se tiene que

$[a_1, a_2] \subseteq [a_1, a_2, a]$ ya que

$$\begin{aligned} y \in [a_1, a_2] &\Rightarrow y = \xi_1 a_1 + \xi_2 a_2 \\ &\Rightarrow y = \xi_1 a_1 + \xi_2 a_2 + 0 \cdot a \\ &\Rightarrow y \in [a_1, a_2, a]. \end{aligned}$$

$\therefore \quad a \in [a_1, a_2] \Rightarrow [a_1, a_2, a] = [a_1, a_2]$ ∎

5) $a_1, a_2 \in [b_1, \ldots, b_n] \Rightarrow [a_1, a_2] \subseteq [b_1, \ldots, b_n]$

suponqamos $a_1, a_2 \in [b_1, \ldots, b_n]$

$$\therefore \quad \begin{aligned} a_1 &= \lambda_{11} b_1 + \lambda_{21} b_2 + \cdots + \lambda_{n1} b_n \\ a_2 &= \lambda_{12} b_1 + \lambda_{22} b_2 + \cdots + \lambda_{n2} b_n \end{aligned}$$

Sea $y \in [a_1, a_2]$

$$\Rightarrow y = \xi_1 a_1 + \xi_2 a_2$$

$$\Rightarrow y = \xi_1 (\lambda_{11} b_1 + \cdots + \lambda_{n1} b_n) + \xi_2 (\lambda_{12} b_1 + \cdots + \lambda_{n2} b_n)$$

$$\Rightarrow y = (\xi_1 \lambda_{11} + \xi_2 \lambda_{12}) b_1 + (\xi_1 \lambda_{21} + \xi_2 \lambda_{22}) b_2 + \cdots + (\xi_1 \lambda_{n1} + \xi_2 \lambda_{n2}) b_n$$

$$\Rightarrow y \in [b_1, b_2, \ldots, b_n].$$

b) $[a_1, a_2, \ldots, a_n] \subseteq [b_1, b_2] \iff a_j \in [b_1, b_2]$
$j = 1, \ldots, n.$

$[\Rightarrow$ Supongamos $[a_1, a_2, \ldots, a_n] \subseteq [b_1, b_2]$

$\therefore\ [a_j] \subseteq [a_1, a_2, \ldots, a_n] \subseteq [b_1, b_2]\,,\ j = 1, \ldots, n$

$\therefore\ a_j \in [b_1, b_2].$

$\Leftarrow]$ Supongamos $a_j \in [b_1, b_2]\,,\ j = 1, 2, \ldots, n.$

$\therefore\ a_j = \lambda_{1j} b_1 + \lambda_{2j} b_2 \qquad j = 1, \ldots, n$

Sea $y \in [a_1, a_2, \ldots, a_n]$

$$\begin{aligned}
\Rightarrow\ & y = \xi_1 a_1 + \xi_2 a_2 + \cdots + \xi_n a_n \\
\Rightarrow\ & y = \xi_1(\lambda_{11} b_1 + \lambda_{21} b_2) + \cdots + \xi_n(\lambda_{1n} b_1 \\
& \qquad + \lambda_{2n} b_2) \\
\Rightarrow\ & y = (\xi_1 \lambda_{11} + \xi_2 \lambda_{12} + \cdots + \xi_n \lambda_{1n}) b_1 \\
& \qquad + (\xi_1 \lambda_{21} + \xi_2 \lambda_{22} + \cdots + \xi_n \lambda_{2n}) b_2 \\
\Rightarrow\ & y \in [b_1, b_2] \,\square
\end{aligned}$$

7) $y \in [a_1, a_2] \iff y = Ax$, donde $A = [a_1 | a_2]$
y $x \in \mathbb{R}^2$.

$y \in [a_1, a_2] \iff y = \xi_1 a_1 + \xi_2 a_2$

$$\Longleftrightarrow \quad y = [a_1 | a_2] \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = Ax$$

donde $A = [a_1 | a_2]$, $x = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$.

EJEMPLOS (para $\mathbb{R}^2$)

1) $\mathbb{R}^2 = [e_1, e_2]$

ya que $y \in \mathbb{R}^2 \Longleftrightarrow y = \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}$

$$\Longleftrightarrow y = \begin{bmatrix} \eta_1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_2 \end{bmatrix}$$

$$\Longleftrightarrow y = \eta_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \eta_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\Longleftrightarrow y \in [e_1, e_2]$$

2) $0 = 0 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\therefore$ 0 puede ser generado por $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

sin embargo $\begin{bmatrix} 1 \\ 0 \end{bmatrix} \neq \lambda \cdot 0 \quad \forall \lambda \quad \therefore \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ no puede ser generado a partir del 0.

3) Sean $a_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $a_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

entonces $[a_1, a_2] = [e_1, e_2] = \mathbb{R}^2$. Ya que

$$y \in [a_1, a_2] \iff y = \xi_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \xi_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\iff y = \xi_1 \left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right) + \xi_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\iff y = \xi_1 (e_1 + e_2) + \xi_2 e_2$$

$$\iff y = \xi_1 e_1 + (\xi_1 + \xi_2) e_2$$

$$\iff y \in [e_1, e_2]$$

O sea que para generar $\mathbb{R}^2$ no hay una unica pareja de vectores generadores.

4) $[e_1, e_2, a] = [e_1, e_2]$ para cualquier a

ya que si $a = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ entonces

$a = \alpha_1 e_1 + \alpha_2 e_2$,

$$y \in [e_1, e_2, a] \iff y = \xi_1 e_1 + \xi_2 e_2 + \xi a$$

$$\iff y = \xi_1 e_1 + \xi_2 e_2 + \xi (\alpha_1 e_1 + \alpha_2 e_2)$$

$$\Leftrightarrow \; y = (\xi_1 + \xi\lambda_1)\,e_1 + (\xi_2 + \xi\lambda_2)\,e_2$$
$$\Leftrightarrow \; y \in [e_1, e_2]$$

PROPIEDADES

a menos que se indique lo contrario, todos los vectores que se consideren serán elementos de $\mathbb{R}^m$.

1) $y \in [a_1, a_2, \dots, a_n] \Leftrightarrow y = Ax$ donde A es la matriz $m \times n$ cuyas columnas son los vectores $a_1, a_2, \dots, a_n$ y $x \in \mathbb{R}^n$

2) $[a_1, a_2, \dots, a_n] = [a_{j_1}, a_{j_2}, \dots, a_{j_n}]$

donde $j_1, j_2, \dots, j_n \in \{1, 2, \dots, n\}$ y $j_l \neq j_m,\ l \neq m$

Es decir, el conjunto de combinaciones lineales de n vectores es independiente del orden en que tomemos los vectores.

3) $a_j \in [a_1, a_2, \dots, a_n] \qquad j = 1, 2, \dots, n$

4) $y_1, y_2, \dots, y_k \in [a_1, a_2, \dots, a_n] \Rightarrow$

$$[y_1, y_2, \dots, y_k] \subseteq [a_1, a_2, \dots, a_n]$$

5) $[a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_\ell] = [a_1, a_2, \dots, a_n] \iff$

$$b_j \in [a_1, a_2, \dots, a_n] \qquad j = 1, \dots, \ell$$

6) $[a_1, a_2, \dots, a_{j-1}, a_{j+1}, \dots, a_n] \neq [a_1, a_2, \dots, a_{j-1}, a_j, a_{j+1}, \dots, a_n]$

$$\iff a_j \notin [a_1, a_2, \dots, a_{i-1}, a_{i+1}, \dots, a_n]$$

7) $\mathbb{R}^m = [e_1, e_2, \dots, e_m]$

Demostración

1) $y \in [a_1, a_2, \dots, a_n] \iff y = \xi_1 a_1 + \xi_2 a_2 + \dots + \xi_n a_n$

$$\iff y = [a_1 | a_2 | \cdots | a_n] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

$$\iff y = Ax$$

donde $A = [a_1 | a_2 | \cdots | a_n]$ y $x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} \in \mathbb{R}^n$

2) $[a_1, a_2, \ldots, a_n] = [a_{j_1}, a_{j_2}, \ldots, a_{j_n}]$, donde $j_1, j_2, \ldots, j_n \in \{1, 2, \ldots, n\}$ y $j_\ell \neq j_k$ $\ell \neq k$.

Esto es consecuencia de la conmutatividad de la suma:

Si $y = [a_1 | a_2 | \cdots | a_n] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$

$\Rightarrow$ $y = [a_{j_1} | a_{j_2} | \cdots | a_{j_n}] \begin{bmatrix} \xi_{j_1} \\ \xi_{j_2} \\ \vdots \\ \xi_{j_n} \end{bmatrix}$ ▪

3) $a_j \in [a_1, a_2, \ldots, a_n]$ $j = 1, 2, \ldots, n$.

$a_j = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_{j-1} + 1 \cdot a_j + 0 \cdot a_{j+1} + \cdots + 0 \cdot a_n$ ▪

4) $y_1, y_2, \ldots, y_k \in [a_1, a_2, \ldots, a_n] \Rightarrow [y_1, y_2, \ldots, y_k] \subseteq [a_1, a_2, \ldots, a_n]$

supongamos $y_j = Ax_j$; $x_j \in \mathbb{R}^n$ $j = 1, \ldots, k$

sea $z \in [y_1, y_2, \ldots, y_k]$ $\therefore$ $z = \underline{Y} w$
donde $\underline{Y} = [y_1 | y_2 | \cdots | y_k]$ y $w \in \mathbb{R}^k$

$\therefore$ $z = [Ax_1 | Ax_2 | \cdots | Ax_k] w$

$$\begin{aligned} &= A\,[x_1|x_2|\cdots|x_k]\,w \\ &= AX\,w \\ &= A\,(Xw) = A\tilde{w} \end{aligned}$$

donde $\tilde{w} = \underset{n\times k}{X}\;\underset{k\times 1}{w} \in \mathbb{R}^n$

$\therefore\quad [y_1, y_2, \dots, y_k] \subseteq [a_1, a_2, \dots, a_n]$ ∎

5). $[a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_\ell] = [a_1, a_2, \dots, a_n]$

$\Leftrightarrow\quad b_j \in [a_1, a_2, \dots, a_n]$

$\Rightarrow]\quad b_j \in [a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_\ell] \subseteq [a_1, a_2, \dots, a_n]$

$\Leftarrow]$ si $b_j \in [a_1, a_2, \dots, a_n] \Rightarrow b_j = Ax_j$ para alguna $x_j \in \mathbb{R}^n$

sea $y \in [a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_\ell]$

$\therefore\quad y = [a_1|a_2|\cdots|a_n|b_1|b_2|\cdots|b_\ell]\,w,$

para alguna $w \in \mathbb{R}^{n+\ell}$

$$\begin{aligned} \therefore\quad y &= [Ae_1|Ae_2|\cdots|Ae_n|\,Ax_1|\,Ax_2|\cdots|Ax_\ell]\,w \\ &= A\,[e_1|e_2|\cdots|e_n|x_1|x_2|\cdots|x_\ell]\,w \end{aligned}$$

donde $e_i \in \mathbb{R}^n$ y $w \in \mathbb{R}^{n+\ell}$

$\therefore \quad y = A\tilde{\omega}$, donde

$\tilde{\omega} = [e_1|e_2|\cdots|e_n|x_1|x_2|\cdots|x_\ell] \, w \quad \in \mathbb{R}^n$ ∎

b) $[a_1,a_2,\dots,a_{i-1},a_{i+1},\dots,a_n] \neq [a_1,a_2,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n]$

$\Leftrightarrow \quad a_i \notin [a_1,a_2,\dots,a_{i-1},a_{i+1},\dots,a_n]$

[⇒ Demostración por contradicción:

supongamos $a_i \in [a_1,a_2,\dots,a_{i-1},a_{i+1},\dots,a_n]$

$$\therefore \quad a_i = \sum_{j\neq i} \lambda_j a_j$$

Sea $y \in [a_1,a_2,\dots,a_{i-1},a_i,a_{i+1},\dots,a_n]$

$$\therefore \quad y = \sum_{j=1}^{n} \xi_j a_j$$

$$= \sum_{j\neq i} \xi_j a_j + \xi_i \sum_{j\neq i} \lambda_j a_j$$

$$= \sum_{j\neq i} (\xi_j + \xi_i \lambda_j) a_j$$

$$\therefore \quad y \in [a_1,a_2,\dots,a_{i-1},a_{i+1},\dots,a_n]$$

⇐] $a_i \in [a_1,a_2,\dots,a_n] \subseteq [a_1,\dots,a_{i-1},a_{i+1},\dots,a_n]$, de esta relación y por contradicción se obtiene la segunda parte. ∎

7) $\mathbb{R}^m = [e_1, e_2, \dots, e_m]$

$$y \in \mathbb{R}^m \Leftrightarrow y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix}$$

$$\Leftrightarrow y = \begin{bmatrix} \eta_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \eta_2 \\ \vdots \\ 0 \end{bmatrix} + \cdots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ \eta_m \end{bmatrix}$$

$$\Leftrightarrow y = \eta_1 e_1 + \eta_2 e_2 + \cdots + \eta_m e_m$$

$$\Leftrightarrow y \in [e_1, e_2, \dots, e_m] \qquad \blacksquare$$

OBSERVACIONES

Como consecuencia de la propiedad (5) tenemos que

1) si $\mathcal{S}$ es un subespacio lineal de $\mathbb{R}^m$ y $\{a_1, a_2, \dots, a_n\}$ es UN conjunto de generadores de $\mathcal{S}$ entonces si $a \in \mathcal{S}$, $\{a_1, a_2, \dots, a_n, a\}$ es también UN conjunto de generadores.

2) si $\{a_1, a_2, \dots, a_n\}$ es UN conjunto de generadores de $\mathcal{S}$ y si

$$a_j \in [a_1, a_2, \dots, a_{j-1}, a_{j+1}, \dots, a_n]$$

para alguna $j \in \{1,2,\ldots,n\}$, entonces $\{a_1, a_2, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n\}$ es también UN conjunto de generadores de $\mathcal{S}$.

De aquí podemos ver que un subespacio lineal puede tener una infinidad de conjuntos generadores.

Por ejemplo, si $\mathcal{S} = [a]$ entonces $\mathcal{S} = [\lambda_i a]$ para cualquier $\lambda_i \neq 0$, $i = 1,2,\ldots$

3) Sabiendo que $\mathbb{R}^m = [e_1, e_2, \ldots, e_m]$, es natural que nos cuestionemos:

Dado $\mathcal{S} \subseteq \mathbb{R}^m$. ¿Existe un conjunto finito de vectores que generan a $\mathcal{S}$?

Es decir:
Dado $\mathcal{S} \subseteq \mathbb{R}^m$ ¿Existe una matriz $A_{m \times n}$ tal que $\mathcal{S} = Im(A)$?

En el caso en que $\mathcal{S}$ esté generado por un número finito de vectores, ¿Existe un conjunto mínimo de vectores que generan a $\mathcal{S}$?

Para contestar estas preguntas necesitamos una herramienta adecuada, y ésta es el concepto de dependencia lineal que desarrollaremos a continuación. Estas preguntas están ligadas al siguiente

PROBLEMA

Dado $y \in \mathbb{R}^m$ y $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$ ¿cuándo $y \in [a_1, a_2, \ldots, a_n]$?

Como vimos anteriormente

$$y \in [a_1, a_2, \ldots, a_n] \Leftrightarrow y = AX$$

donde $A = [a_1 | a_2 | \cdots | a_n]$ y $X \in \mathbb{R}^n$

Recordemos que $AX = y$ representa un sistema de m ecuaciones con n incógnitas, en el cual conocemos la matriz A de los coeficientes, el vector y, y tenemos que determinar el vector X.

El sistema $AX=y$ no siempre tiene solución. $AX=y$ tiene solución únicamente cuando y es una combinación lineal de las columnas de A.

Si $n<m$, entonces el sistema generalmente no tiene solución.

Por ejemplo si $m=3$ y $n=2$ entonces $\mathbb{R}^m = \mathbb{R}^3$ y $\mathbb{R}^n = \mathbb{R}^2$.

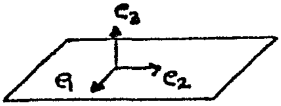
Sean $a_1 = e_1$, $a_2 = e_2$, $y = e_3$

entonces el sistema

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

no tiene solución, ya que el vector $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

no está en el plano generado por e_1 y e_2:



DEFINICION

Sea $y \in \mathbb{R}^m$ y $a_1, a_2, \dots, a_n \in \mathbb{R}^m$. Entonces

Si $y \in [a_1, a_2, \dots, a_n]$ diremos que y es LINEALMENTE DEPENDIENTE DE $\{a_1, a_2, \dots, a_n\}$

Si $y \notin [a_1, a_2, \dots, a_n]$ diremos que y es LINEALMENTE INDEPENDIENTE DE $\{a_1, a_2, \dots, a_n\}$

Por ejemplo, e_3 es linealmente independiente de $\{e_1, e_2\}$

Antes de analizar las propiedades de la dependencia e independencia lineal para el caso general, veremos qué pasa cuando $n=1$ (rectas) y $n=2$ (planos).

PROPIEDADES (RECTAS)

1) El vector cero es linealmente dependiente de $\{a\}$ para todo $a \in \mathbb{R}^m$.

2) λa es linealmente dependiente de $\{a\}$ para todo $a \in \mathbb{R}^m$ y $\lambda \in \mathbb{R}$.

3) b es linealmente dependiente de $\{a\} \leftrightarrow [b] \subseteq [a]$

4) b es linealmente dependiente de $\{a\} \Leftrightarrow$ $[a,b] = [a]$

5) a es linealmente independiente del vector cero $\forall a \neq 0$

6) b es linealmente independiente de a $\Leftrightarrow$ $[a,b] \neq [a]$.

Demostración :

1) $0 = 0 \cdot a \quad \therefore \quad 0 \in [a] \quad \forall a \in \mathbb{R}^m$

2) $(\lambda a) = \lambda \cdot a \quad \therefore \quad \lambda a \in [a]$

3) $\Rightarrow]$ Supongamos b es linealmente dependiente de $\{a\}$.

$\Rightarrow \quad b = \lambda a$ para alguna $\lambda \in \mathbb{R}$

Sea $y \in [b] \Rightarrow y = \xi b$ para $\xi \in \mathbb{R}$

$\therefore \quad y = \xi(\lambda a) = (\xi \lambda) a \quad \in [a]$

$\therefore \quad [b] \subseteq [a]$.

$\Leftarrow]$ supongamos $[b] \subseteq [a]$

$\therefore\ b \in [b] \subseteq [a] \qquad \therefore\ b \in [a]$

4) [⇒ Supongamos b linealmente dependiente de a

$\Rightarrow b = \lambda a$. Sea $y \in [a,b]$

$\therefore\ y = [a|b]x = [a|\lambda a]x$, $x \in \mathbb{R}^2$

$= a[1\ \lambda]x = \lambda' a$

donde $\lambda' = [1\ \lambda]x$

$\therefore\ y \in [a]$

$\therefore\ [a,b] \subseteq [a] \qquad \therefore\ [a,b] = [a]$

⇐] Supongamos $[a,b] \subseteq [a]$

$\therefore\ [b] \subseteq [a,b] \subseteq [a] \qquad \therefore\ b \in [a]$

5) Supongamos a linealmente dependiente de 0 $\Rightarrow a = \lambda \cdot 0 \Rightarrow a = 0$

6) Es consecuencia de (4)

PROPIEDADES (PLANOS)

1) 0 es linealmente dependiente de $\{a_1,a_2\}$

2) a_1 es linealmente dependiente de $\{a_1,a_2\}$
a_2 es linealmente dependiente de $\{a_1,a_2\}$

3) $[a_1,a_2]$ es linealmente dependiente de $\{a_1,a_2\}$

4) a es linealmente dependiente de $\{a_1,a_2\} \Leftrightarrow$
$[a_1,a_2,a] = [a_1,a_2]$

5) a es linealmente independiente de $\{a_1,a_2\}$
$\Leftrightarrow [a_1,a_2,a] \neq [a_1,a_2]$

Demostración

1) $0 = 0\cdot a_1 + 0\cdot a_2 \quad \therefore \quad 0 \in [a_1,a_2]$ ∎

2) $a_1 = 1\cdot a_1 + 0\cdot a_2 \quad \therefore \quad a_1 \in [a_1,a_2]$
$a_2 = 0\cdot a_1 + 1\cdot a_2 \quad \therefore \quad a_2 \in [a_1,a_2]$ ∎

3) $[a_1,a_2] \subseteq [a_1,a_2]$ ∎

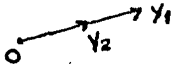
4) $a \in [a_1,a_2] \Leftrightarrow [a_1,a_2,a] = [a_1,a_2]$ ∎

5) $a \notin [a_1,a_2] \Leftrightarrow [a_1,a_2,a] \neq [a_1,a_2]$ ∎

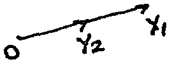
EJEMPLOS

Supongamos que $y_1, y_2 \in \mathbb{R}^m$. Entonces podemos tener los siguientes casos :

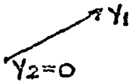
1) y_1 es linealmente dependiente de y_2

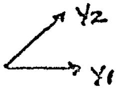


2) y_2 es linealmente dependiente de y_1

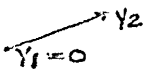


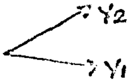
3) y_1 es linealmente independiente de y_2



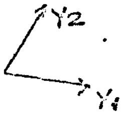


4) y_2 es linealmente independiente de y_1





5) y_1 es linealmente independiente de y_2 y y_2 es linealmente independiente de y_1



OBSERVACIONES

1) y_1 linealmente independiente de $\{y_2\} \Rightarrow$ y_2

linealmente independiente de $\{y_1\}$.

Ya que, por ejemplo, y_1 es linealmente independiente de $\{0\}$, pero 0 es linealmente dependiente de $\{y_1\}$, si $y_1 \neq 0$.

2) y_1 linealmente dependiente de $\{y_2\} \nRightarrow y_2$ linealmente dependiente de $\{y_1\}$.

Si $y_2 \neq 0$, entonces 0 es linealmente dependiente de $\{y_2\}$ y y_2 es linealmente independiente de $\{0\}$.

3) y linealmente dependiente de $\{y_1, y_2\} \nRightarrow y_1$ linealmente dependiente de $\{y, y_2\}$.

0 es linealmente dependiente de $\{e_1, e_2\}$ y sin embargo e_1 es linealmente independiente de $\{0, e_2\}$

4) y linealmente independiente de $\{y_1, y_2\} \nRightarrow$ y_1 linealmente independiente de $\{y, y_2\}$.

e_1 es linealmente independiente de $\{0, e_2\}$ pero 0 es linealmente dependiente de $\{e_1, e_2\}$.

PROPIEDADES

A menos que se indique lo contrario, todos los vectores que se consideren serán elementos de $\mathbb{R}^m$.

1) 0 es linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$

2) a_j es linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$
$j = 1, 2, \ldots, n$

3) $y_1, y_2, \ldots, y_k$ linealmente dependiente de $\{a_1, a_2, \ldots, a_n\} \Rightarrow [y_1, y_2, \ldots, y_k]$ linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$

4) $[a_{j_1}, a_{j_2}, \ldots, a_{j_k}]$ es linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$; $j_1, j_2, \ldots, j_k \in \{1, 2, \ldots, n\}$

Demostración:

1) $0 = 0 \cdot a_1 + 0 \cdot a_2 + \cdots + 0 \cdot a_n \quad \therefore 0 \in [a_1, a_2, \ldots, a_n]$ ■

2) $a_j \in [a_1, a_2, \ldots, a_n] \quad j = 1, 2, \ldots, n$ ■

3) $y_1, y_2, \ldots, y_k \in [a_1, a_2, \ldots, a_n] \Rightarrow$
$[y_1, y_2, \ldots, y_k] \subseteq [a_1, a_2, \ldots, a_n]$ ■

4) Es consecuencia de (3) y (2) ■

RESULTADO

b es linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$ $\Leftrightarrow$ $[a_1, a_2, \ldots, a_n, b] = [a_1, a_2, \ldots, a_n]$

Demostración:

$$b \in [a_1, a_2, \ldots, a_n] \Leftrightarrow [a_1, a_2, \ldots, a_n] = [a_1, a_2, \ldots, a_n, b] \blacksquare$$

RESULTADO

b es linealmente independiente de $\{a_1, a_2, \ldots, a_n\}$ $\Leftrightarrow$ $[a_1, a_2, \ldots, a_n, b] \neq [a_1, a_2, \ldots, a_n]$

Demostración:

$$b \notin [a_1, a_2, \ldots, a_n] \Leftrightarrow [a_1, a_2, \ldots, a_n, b] \neq [a_1, a_2, \ldots, a_n] \blacksquare$$

Ahora estamos en condiciones de poder presentar los conceptos de dependencia e independencia lineal de conjuntos finitos de vectores.

DEFINICION

Sean $a_1, a_2, \ldots, a_n \in \mathbb{R}^m$. Decimos que $\{a_1, a_2, \ldots, a_n\}$ es LINEALMENTE DEPENDIENTE si cuando menos una de las a_j es linealmente dependiente de las restantes. Es decir, si existe $j \in \{1, 2, \ldots, n\}$ tal que

$$a_j \in [a_1, a_2, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n]$$

DEFINICION

Sean $a_1, a_2, \dots, a_n \in \mathbb{R}^m$. Decimos que $\{a_1, a_2, \dots, a_n\}$ es LINEALMENTE INDEPENDIENTE si cada a_j es linealmente independiente de las restantes. Es decir, si

$$a_j \notin [a_1, a_2, \dots, a_{j-1}, a_{j+1}, \dots, a_n] \qquad \forall j \in \{1, 2, \dots, n\}$$

EJEMPLOS

1. $\{e_1, e_2, e_2 + e_1\}$ es linealmente dependiente
2. $\{e_1, e_2\}$ es linealmente independiente.

Los resultados que hemos visto sobre dependencia e independencia lineal los podemos traducir al lenguaje matricial:

RESULTADOS

1) $y \in \mathbb{R}^m$ es linealmente dependiente de $\{a_1, a_2, \dots, a_n\} \subseteq \mathbb{R}^m \iff y \in R(A) \iff y = Ax\,;\ x \in \mathbb{R}^n$

2) $y \in \mathbb{R}^m$ es linealmente independiente de $\{a_1, a_2, \dots, a_n\} \subseteq \mathbb{R}^m \iff y \notin R(A) \iff y \neq Ax \quad \forall x \in \mathbb{R}^n$

Demostración :

1) y linealmente dependiente de $\{a_1, a_2, \ldots, a_n\}$ $\Leftrightarrow$ $y \in [a_1, a_2, \ldots, a_n] \Leftrightarrow y \in R(A) \Leftrightarrow y = AX$ para alguna $X \in \mathbb{R}^n$ $\blacksquare$

2) y linealmente independiente de $\{a_1, a_2, \ldots, a_n\}$ $\Leftrightarrow$ $y \notin [a_1, a_2, \ldots, a_n] \Leftrightarrow y \notin R(A) \Leftrightarrow y \neq AX \ \forall X \in \mathbb{R}^n$ $\blacksquare$

RESULTADOS

1) $\{a_1, a_2, \ldots, a_n\} \subseteq \mathbb{R}^m$ es linealmente dependiente $\Leftrightarrow$ existe $X \neq 0$ tal que $AX = 0 \Leftrightarrow N(A) \neq \{0\}$

2) $\{a_1, a_2, \ldots, a_n\} \subseteq \mathbb{R}^m$ es linealmente independiente $\Leftrightarrow$ $(AX = 0 \Rightarrow X = 0) \Leftrightarrow N(A) = \{0\}$

Demostración :

1) $[\Rightarrow]$ Supongamos $a_j \in [a_1, a_2, \ldots, a_{j-1}, a_{j+1}, \ldots, a_n]$ para alguna $j \in \{1, 2, \ldots, n\}$

$$\therefore \quad a_j = [a_1 | a_2 | \cdots | a_{j-1} | a_{j+1} | \cdots | a_n] \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{j-1} \\ \xi_{j+1} \\ \vdots \\ \xi_n \end{bmatrix}$$

$$\Rightarrow [a_1|a_2|\cdots|a_{j-1}|a_{j+1}|\cdots|a_n]\begin{bmatrix}\xi_1\\ \vdots\\ \xi_{j-1}\\ \xi_{j+1}\\ \vdots\\ \xi_n\end{bmatrix} - a_j = 0$$

$$\Rightarrow [a_1|a_2|\cdots|a_n]\begin{bmatrix}\xi_1\\ \vdots\\ \xi_{j-1}\\ 0\\ \xi_{j+1}\\ \vdots\\ \xi_n\end{bmatrix} + [a_1|a_2|\cdots|a_n]\begin{bmatrix}0\\ \vdots\\ 0\\ -1\\ 0\\ \vdots\\ 0\end{bmatrix}\begin{matrix}\\ \\ \\ \leftarrow j\\ \\ \\ \\ \end{matrix} = 0$$

$$\Rightarrow [a_1|a_2|\cdots|a_n]\begin{bmatrix}\xi_1\\ \vdots\\ \xi_{j-1}\\ -1\\ \xi_{j+1}\\ \vdots\\ 0\end{bmatrix} = 0$$

$\Rightarrow$ existe $x \in \mathbb{R}^n \neq 0$ tal que $Ax = 0$

$\Rightarrow$ $N(A) \neq \{0\}$

$\Leftarrow$] Supongamos $N(A) \neq \{0\}$ $\therefore$ existe $x \in \mathbb{R}^n$, $x \neq 0$ tal que $Ax = 0$

Supongamos $\xi_j \neq 0$

$$Ax = 0 \quad \Rightarrow$$

$$[a_1|a_2|\cdots|a_n] \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_j \\ \vdots \\ \xi_n \end{bmatrix} = 0 \quad \Rightarrow$$

$$[a_1|a_2|\cdots|a_n] \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{j-1} \\ 0 \\ \xi_{j+1} \\ \vdots \\ \xi_n \end{bmatrix} + [a_1|a_2|\cdots a_n] \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \xi_j \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow j \\ \\ \\ \\ \end{matrix} = 0 \quad \Rightarrow$$

$$[a_1|a_2|\cdots|a_{j-1}|a_{j+1}|\cdots|a_n] \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_{j-1} \\ \xi_{j+1} \\ \vdots \\ \xi_n \end{bmatrix} + \xi_j a_j = 0 \quad \Rightarrow$$

$$a_j = [a_1|a_2|\cdots|a_{j-1}|a_{j+1}|\cdots|a_n] \begin{bmatrix} \xi_1/\xi_j \\ \vdots \\ \xi_{j-1}/\xi_j \\ \xi_{j+1}/\xi_j \\ \vdots \\ \xi_n/\xi_j \end{bmatrix}$$

(esto es posible ya que $\xi_j \neq 0$)

$$\Rightarrow a_j \in [a_1|a_2|\cdots|a_{j-1}|a_{j+1}|\cdots|a_n]$$

$$\Rightarrow \{a_1, a_2, \ldots, a_n\} \text{ es linealmente dependiente} \blacksquare$$

2) Es equivalente a (1) ■

Hasta ahora hemos visto los conceptos de combinación lineal, subespacio lineal, etc. para vectores columna. Las definiciones para vectores renglón son análogas.

DEFINICION

Sea $c^t \in \mathbb{R}^{n*}$ y $r_1^t, r_2^t, \ldots, r_m^t \in \mathbb{R}^{n*}$. Decimos que c^t es una combinación lineal de $r_1^t, r_2^t, \ldots, r_m^t$ si existen $\eta_1, \eta_2, \ldots, \eta_m \in \mathbb{R}$ tal que

$$c^t = \eta_1 r_1^t + \eta_2 r_2^t + \cdots + \eta_m r_m^t$$

Es decir c^t es una combinación lineal de $r_1^t, r_2^t, \ldots, r_m^t$ si existe $y \in \mathbb{R}^m$ tal que

$$y^t A = c^t$$

donde A es la matriz cuyos renglones son

$r_1^t, r_2^t, \ldots, r_m^t$.

Y en términos de la matriz A^t tenemos que

$$y^t A = c^t \iff A^t y = c$$

RESULTADOS

1) $c^t \in \mathbb{R}^{n*}$ es linealmente dependiente de $\{r_1^t, r_2^t, \ldots, r_m^t\} \subseteq \mathbb{R}^{n*} \iff y^t A = c^t$ para alguna $y \in \mathbb{R}^m \iff A^t y = c$; $y \in \mathbb{R}^m \iff$ $c \in R(A^t)$

2) $c^t \in \mathbb{R}^{n*}$ es linealmente independiente de $\{r_1^t, r_2^t, \ldots, r_m^t\} \subseteq \mathbb{R}^{n*} \iff y^t A \neq c^t \quad \forall\, y \in \mathbb{R}^m$ $\iff A^t y \neq c \quad \forall\, y \in \mathbb{R}^m \iff c \notin R(A^t)$

3) $\{r_1^t, r_2^t, \ldots, r_m^t\}$ es linealmente dependiente
$\iff$ existe $y \in \mathbb{R}^m$, $y \neq 0$ tal que $y^t A = 0^t$
$\iff$ existe $y \in \mathbb{R}^m$, $y \neq 0$ tal que $A^t y = 0^t$
$\iff N(A^t) \neq \{0\}$

4) $\{r_1^t, r_2^t, \ldots, r_m^t\}$ es linealmente independiente
$\iff y^t A = 0^t \Rightarrow y = 0$
$\iff A^t y = 0 \Rightarrow y = 0$
$\iff N(A^t) = \{0\}$

La demostración es análoga a la demostración para vectores columna.

OBSERVACIONES

al considerar vectores renglón nos hemos topado con otros dos subespacios asociados a la matriz $A_{m\times n}$:

1) $R(A^t) = [r_1, r_2, \ldots, r_m]$; $A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$

2) $N(A^t) = \{ y \in \mathbb{R}^m : A^t y = 0 \} = \{ y \in \mathbb{R}^m : y^t A = 0^t \}$

O sea, que asociados a una matriz tenemos asociados cuatro subespacios lineales:

1) $N(A) \subseteq \mathbb{R}^n$
2) $R(A) \subseteq \mathbb{R}^m$
3) $N(A^t) \subseteq \mathbb{R}^m$
4) $R(A^t) \subseteq \mathbb{R}^n$

Estos subespacios son muy importantes.

OBSERVACION

Si $y^t A = 0^t \Rightarrow y^t \cdot [a_1 | a_2 | \cdots | a_n] = 0^t$

$\Rightarrow [y^t a_1 | y^t a_2 | \cdots | y^t a_n] = 0^t$

$\Rightarrow$ y es ortogonal a las columnas de A.

Nos habíamos planteado la siguiente pregunta:

¿ Dado $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio lineal de $\mathbb{R}^m$, existe una matriz A tal que $\mathcal{S} = R(A)$?

Para $\mathcal{S} = \mathbb{R}^m$ la respuesta es afirmativa, ya que $\mathbb{R}^m = [e_1, e_2, \ldots, e_m] = R(I_m)$.

Para un subespacio lineal $\mathcal{S} \subseteq \mathbb{R}^m$ podríamos seguir el procedimiento

1° sea $a_1 \in \mathcal{S}$, $a_1 \neq 0$

2° si $\mathcal{S} \subseteq [a_1]$ ya acabamos.
si $\mathcal{S} \not\subseteq [a_1]$, sea $a_2 \in \mathcal{S}$, $a_2 \notin [a_1]$

⋮

En el i-ésimo paso

si $\mathcal{S} \subseteq [a_1, a_2, \ldots, a_{i-1}]$ ya acabamos,
si $\mathcal{S} \not\subseteq [a_1, a_2, \ldots, a_{i-1}]$ tomamos $a_i \in \mathcal{S}$ tal que $a_i \notin [a_1, a_2, \ldots, a_{i-1}]$.

Este proceso podría continuar indefinidamente, pero el hecho de que $\mathcal{S}$ sea subespacio de $\mathbb{R}^m$ nos hace pensar que después de un número finito de pasos $\mathcal{S} = [a_1, a_2, \ldots, a_k]$. Veremos que ésto siempre pasa.

Observemos que $\{e_1, e_2, \ldots, e_m\}$ es un conjunto máximo linealmente independiente, ya que $\{e_1, e_2, \ldots, e_m, a\}$ es linealmente dependiente para cualquier $a \in \mathbb{R}^m$.

¿Es ésta una propiedad del conjunto $\{e_1, e_2, \ldots, e_m\}$ ó es una propiedad de $\mathbb{R}^m$?

Que fuera una propiedad intrínseca a $\mathbb{R}^m$ significaría que los subconjuntos de $\mathbb{R}^m$ linealmente independientes son finitos y que tal vez no pueden tener más de m vectores. Para ver las posibilidades de que ésto sea cierto necesitamos encontrar ejemplos de conjuntos de vectores linealmente independientes que generen $\mathbb{R}^m$, es decir, encontrar matrices $A_{m\times n}$ tales que

$$\mathbb{R}^m = R(A)$$

Afortunadamente existen unas matrices simples con esta propiedad, a saber las matrices triangulares.

RESULTADO

Sea $L_{m\times m}$ una matriz triangular inferior, y tal que los elementos de la diagonal sean diferentes de cero. Entonces

1) Las columnas de L generan $\mathbb{R}^m$: $R(L) = \mathbb{R}^m$

2) Las columnas de L son linealmente inde-

pendientes.

Demostración:

$$Si \quad L = \begin{bmatrix} \lambda_{11} & & & & \\ \lambda_{21} & \lambda_{22} & & & \\ \lambda_{31} & \lambda_{32} & \lambda_{33} & & \\ \vdots & \vdots & \vdots & & \\ \lambda_{m1} & \lambda_{m2} & \lambda_{m3} & \cdots & \lambda_{mm} \end{bmatrix} = [\ell_1|\ell_2|\cdots|\ell_m]$$

entonces:

1)
$$\begin{aligned} \ell_1 &= \lambda_{11}e_1 + \lambda_{21}e_2 + \lambda_{31}e_3 + \cdots + \lambda_{m1}e_m \\ \ell_2 &= \lambda_{22}e_2 + \lambda_{32}e_3 + \cdots + \lambda_{m2}e_m \\ \ell_3 &= \lambda_{33}e_3 + \cdots + \lambda_{m3}e_m \\ &\vdots \\ \ell_m &= \lambda_{mm}e_m \end{aligned}$$

Como $\lambda_{11} \neq 0$ entonces

$$\frac{\ell_1}{\lambda_{11}} = e_1 + \frac{\lambda_{21}}{\lambda_{11}}e_2 + \frac{\lambda_{31}}{\lambda_{11}}e_3 + \cdots + \frac{\lambda_{m1}}{\lambda_{11}}e_m$$

es decir $e_1 \in [\ell_1, e_2, e_3, \ldots, e_m]$

Podemos hacer lo mismo para los demás vectores e_j y deducir del sistema de ecuaciones que:

$$
\begin{aligned}
e_1 &\in [\ell_1, e_2, e_3, \dots, e_m] \\
e_2 &\in [\ell_2, e_3, e_4, \dots, e_m] \\
&\vdots \\
e_{m-1} &\in [\ell_{m-1}, e_m] \\
e_m &\in [\ell_m]
\end{aligned}
$$

Esto implica que

$$
\begin{aligned}
e_m &\in [\ell_m] \\
e_{m-1} &\in [\ell_{m-1}, e_m] \subseteq [\ell_{m-1}, \ell_m] \\
&\vdots \\
e_2 &\in [\ell_2, e_3, e_4, \dots, e_m] \subseteq [\ell_2, \ell_3, \ell_4, \dots, \ell_m] \\
e_1 &\in [\ell_1, e_2, e_3, \dots, e_m] \subseteq [\ell_1, \ell_2, \ell_3, \dots, \ell_m]
\end{aligned}
$$

o sea que

$$e_j \in [\ell_1, \ell_2, \dots, \ell_m] \qquad 1 \leq j \leq m$$

y por lo tanto

$$[e_1, e_2, \dots, e_m] \subseteq [\ell_1, \ell_2, \dots, \ell_m]$$

$$\therefore\ [e_1, e_2, \dots, e_m] = [\ell_1, \ell_2, \dots, \ell_m]$$

$$\therefore\ \mathbb{R}^m = R(L)$$

2) Ahora veremos que las columnas de L son linealmente independientes:

Supongamos $LX=0$ para alguna $X=[\xi_i]\in\mathbb{R}^m$. Entonces

$$\begin{aligned}
&\lambda_{11}\xi_1 &&=0\\
&\lambda_{21}\xi_1+\lambda_{22}\xi_2 &&=0\\
&\vdots\\
&\lambda_{m1}\xi_1+\lambda_{m2}\xi_2+\cdots+\lambda_{mm}\xi_m &&=0
\end{aligned}$$

$$\lambda_{11}\xi_1=0,\ \lambda_{11}\neq0\Rightarrow\xi_1=0$$
$$0+\lambda_{22}\xi_2=0,\ \lambda_{22}\neq0\Rightarrow\xi_2=0$$
$$\vdots$$
$$0+0+\cdots+\lambda_{mm}\xi_m=0,\ \lambda_{mm}\neq0\Rightarrow\xi_m=0$$

o sea que $X=0$.

$\therefore\ LX=0\Rightarrow X=0$

$\therefore$ Las columnas de L son linealmente independientes ▪

RESULTADO

Sea $U_{m\times m}$ una matriz triangular superior y tal que los elementos de la diagonal sean diferentes de cero. Entonces

1) Las columnas de U generan $\mathbb{R}^m$: $R(U)=\mathbb{R}^m$

2) Los columnas de U son linealmente independientes.

La demostración es análoga a la demostración para matrices triangulares inferiores. ∎

Hemos demostrado que si L es triangular inferior, entonces

$\lambda_{ii} \neq 0 \Rightarrow L$ genera $\mathbb{R}^m$
$\lambda_{ii} \neq 0 \rightarrow$ las columnas de L son linealmente independientes.

Queremos ver que si L es triangular inferior entonces:
las columnas de L generan $\mathbb{R}^m \Leftrightarrow$ las columnas de L son linealmente independientes.

Primero veremos:

RESULTADO

Sea $L_{m \times m}$ una matriz triangular inferior.

Entonces las columnas de L son linealmente independientes $\Leftrightarrow$ los elementos de la diagonal son diferentes de cero.

Demostración:

$\Leftarrow$] Esta implicación ya la demostramos

$\Rightarrow$] Supongamos que algún elemento de la diagonal es igual a cero :

Sea $\lambda_{kk}=0$ donde k es tal que $\lambda_{ii} \neq 0$ para $i>k$; o sea que en el caso de que haya más de un elemento de la diagonal igual a cero, nos fijaremos en el último de tales elementos.

$$L = \begin{bmatrix} \lambda_{11} & & & & & \\ \lambda_{21} & \lambda_{22} & & & & \\ \vdots & & & & & \\ \lambda_{k1} & \lambda_{k2} & \cdots & 0 & & \\ \vdots & & & & & \\ \lambda_{m1} & \lambda_{m2} & \cdots & \lambda_{mk} & \cdots & \lambda_{mm} \end{bmatrix}$$

$$\therefore \quad \begin{aligned} l_m &= \lambda_{mm} e_m \\ l_{m-1} &= \lambda_{m-1,m-1} e_{m-1} + \lambda_{m,m-1} e_m \\ &\vdots \\ l_{k+1} &= \lambda_{k+1,k+1} e_{k+1} + \cdots + \lambda_{m,k+1} e_m \\ l_k &= \lambda_{k+1,k} e_{k+1} + \cdots + \lambda_{m,k} e_m \end{aligned}$$

Como $\lambda_{ii} \neq 0$ para $i>k \Rightarrow$

$$\begin{aligned} & e_m \in [l_m] \\ & e_{m-1} \in [l_{m-1}, e_m] \subseteq [l_{m-1}, l_m] \\ & \vdots \\ & e_{k+1} \in [l_{k+1}, e_{k+2}, \ldots, e_m] \subseteq [l_{k+1}, l_{k+2}, \ldots, l_m] \end{aligned}$$

Por lo tanto tenemos que

$$\ell_K \in [e_{K+1}, e_{K+2}, \ldots, e_m] \subseteq [\ell_{K+1}, \ell_{K+2}, \ldots, \ell_m]$$

$\therefore$ las columnas de L son linealmente dependientes.

$\therefore$ Si L tiene columnas linealmente independientes $\Rightarrow$ todos los elementos de la diagonal de L son diferentes de cero.

RESULTADO

Sea $L_{m\times m}$ una matriz triangular inferior. Entonces las columnas de L generan $\mathbb{R}^m$ $\Leftrightarrow$ los elementos de la diagonal de L son diferentes de cero.

Demostración :

$\Leftarrow$ Ya demostramos esta implicación.

$\Rightarrow$ Supongamos que existe al menos un elemento de la diagonal de L cuyo valor es cero. Sea λ_{kk} el primero de tales elementos:

$$\lambda_{kk} = 0 \quad \text{y} \quad \lambda_{ii} \neq 0 \quad \text{para} \quad i < k.$$

Entonces afirmamos que $e_k \notin R(L)$. Veremos que no existe $x \in \mathbb{R}^m$ tal que $Lx = e_k$:

$$Lx = \begin{bmatrix} \lambda_{11} & & & & \\ \lambda_{21} & \lambda_{22} & & & \\ \vdots & \vdots & & & \\ \lambda_{k1} & \lambda_{k2} & \cdots\ 0 & & \\ \vdots & \vdots & \vdots & & \\ \lambda_{m1} & \lambda_{m2} & \cdots\ \lambda_{mk} & \cdots & \lambda_{mm} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \\ \vdots \\ \xi_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$$

Entonces, para $i < k$ ξ_i tiene que ser igual a cero:

$$0 = \lambda_{11}\xi_1, \quad \lambda_{11} \neq 0 \Rightarrow \xi_1 = 0$$

$$0 = \lambda_{21}\xi_1 + \lambda_{22}\xi_2, \quad \xi_1 = 0, \quad \lambda_{22} \neq 0 \Rightarrow \lambda_{22}\xi_2 = 0$$
$$\Rightarrow \xi_2 = 0$$
$$\vdots$$
$$0 = \lambda_{k-1,1}\xi_1 + \lambda_{k-1,2}\xi_2 + \cdots + \lambda_{k-1,k-1}\xi_{k-1}$$
$$= 0 + 0 + \cdots + \lambda_{k-1,k-1}\xi_{k-1} = \lambda_{k-1,k-1}\xi_{k-1}$$
$$\therefore \quad \xi_{k-1} = 0$$

$\therefore$ Tenemos que

$$\lambda_{k1}\xi_1 + \lambda_{k2}\xi_2 + \cdots + \lambda_{k,k-1}\xi_{k-1} + 0 \cdot \xi_k = 1$$

Es decir $0 + 0 + \cdots + 0 + 0 = 1$ $\underset{0}{\nabla}$

$\therefore$ La suposición de que existe $x \in \mathbb{R}^m$

tal que $LX = e_k$ es absurda

$\therefore \quad LX \neq e_k \quad \forall\, X \in \mathbb{R}^m$

$\therefore \quad e_k \notin R(L)$

$\therefore$ Las columnas de L no generan $\mathbb{R}^m$

$\therefore$ Las columnas de L generan $\mathbb{R}^m \Rightarrow$ los elementos de la diagonal de L son diferentes de cero ■

Hemos demostrado el siguiente

RESULTADO

Sea L $m\times m$ una matriz triangular inferior. Entonces son equivalentes

1) Las columnas de L generan $\mathbb{R}^m$
2) Las columnas de L son linealmente independientes
3) Los elementos de la diagonal de L son diferentes de cero.

Para las matrices triangulares superiores tenemos el resultado, que se demuestra de forma similar :

RESULTADO

Sea $U_{m \times m}$ una matriz triangular superior. Entonces son equivalentes :

1) Las columnas de U generan $\mathbb{R}^m$
2) Las columnas de U son linealmente independientes
3) Los elementos de la diagonal de L son diferentes de cero.

■

Para los renglones de T tenemos el siguiente resultado, donde T denota a una matriz triangular superior ó inferior :

RESULTADO

Sea $T_{m \times m}$ una matriz triangular. Entonces son equivalentes :

1) Los renglones de T generan $\mathbb{R}^{m*}$
2) Los renglones de T son linealmente independientes.
3) Los elementos de la diagonal de T son diferentes de cero.

Este resultado se demuestra fácilmente a partir de la matriz transpuesta T^t.

RESULTADO

Los renglones de T son linealmente independientes si y solo si las columnas de T son linealmente independientes.

La demostración es inmediata y la omitimos.

■

De aquí podemos ver que es muy fácil saber cuando las columnas ó renglones de las matrices triangulares de orden m generan $\mathbb{R}^m$.

Ahora, supongamos que tenemos una matriz triangular por bloques. Es decir, A mxm de la forma

$$A = \begin{array}{cc} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array}$$

Entonces tenemos:

RESULTADO

Sea A mxn una matriz triangular por bloques:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

Entonces, si las columnas de A_{11} y las columnas de A_{22} son linealmente independientes $\Rightarrow$ las columnas de A son linealmente independientes: Es decir:

$$\begin{pmatrix} A_{11}X_1 = 0 \Rightarrow X_1 = 0 \\ A_{22}X_2 = 0 \Rightarrow X_2 = 0 \end{pmatrix} \Rightarrow \left(AX = 0 \Rightarrow X = 0 \right)$$

donde $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathbb{R}^n$

Demostración:

Supongamos $A_{11}X_1 = 0 \Rightarrow X_1 = 0$ y
$A_{22}X_2 = 0 \Rightarrow X_2 = 0$

Veremos $AX = 0 \Rightarrow X = 0$. Sea $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathbb{R}^n$

Supongamos $AX = 0$:

$$Ax = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11}x_1 + A_{12}x_2 \\ A_{22}x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \quad \begin{matrix} A_{22}x_2 = 0 \\ A_{11}x_1 + A_{12}x_2 = 0 \end{matrix} \quad \Rightarrow x_2 = 0$$

$$\Rightarrow \quad A_{11}x_1 = 0 \quad \Rightarrow \quad x_1 = 0$$

$$\therefore \quad x_1 = x_2 = 0 \quad \therefore \quad x = 0_{\blacksquare}$$

RESULTADO

Sea A $m \times n$ triangular por bloques.

$$A = \begin{array}{cc} r & n-r & \\ \left[\begin{array}{c} A_{11} \\ 0 \end{array}\right. & \left.\begin{array}{c} A_{12} \\ A_{22} \end{array}\right] & \begin{array}{l} s \\ m-s \end{array} \end{array}$$

Entonces, si las columnas de A son linealmente independientes -> las columnas de A_{11} son linealmente independientes:

$$(Ax = 0 \Rightarrow x = 0) \Rightarrow (A_{11}x_1 = 0 \Rightarrow x_1 = 0)$$

donde $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^n$.

Demostración :

Supongamos que existe alguna $x_1 \in \mathbb{R}^r$, $x_1 \neq 0$ tal que $A_{11} x_1 = 0$.

sea $x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \in \mathbb{R}^n$. Entonces $x \neq 0$, ya que $x_1 \neq 0$ y además

$$Ax = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} A_{11} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\therefore$ Existe una $x \neq 0$ tal que $Ax = 0$

$\therefore$ $(Ax = 0 \Rightarrow x = 0) \Rightarrow (A_{11} x_1 = 0 \Rightarrow x_1 = 0)$ ∎

Observemos que el que A tenga columnas independientes no implica que A_{22} tenga columnas independientes, ya que, por ejemplo, si

$$A = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

entonces las columnas de A son linealmente independientes, pero las columnas de A_{22} son linealmente dependientes.

Sin embargo, podríamos haber dividido por bloques a A de la siguiente manera:

$$A = \left[\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] = \begin{bmatrix} A'_{11} & A'_{12} \\ 0 & A'_{22} \end{bmatrix}$$

Estos bloques se diferencian de los primeros en que A'_{11} es triangular y en que las columnas de A'_{22} sí son linealmente independientes.

RESULTADO

Sea $A_{m\times n}$ una matriz triangular por bloques, de la siguiente forma:

$$A = \begin{array}{cc} \begin{array}{cc} r & n-r \end{array} & \\ \begin{bmatrix} T & A_{12} \\ 0 & A_{22} \end{bmatrix} & \begin{array}{l} r \\ m-r \end{array} \end{array}$$

donde T es una matriz triangular de orden r, $N(T) = \{0\}$

Entonces, si las columnas de A son linealmente independientes → A_{22} tiene columnas independientes.

Demostración :

Supongamos que existe $x_2 \in \mathbb{R}^{n-r}$, $x_2 \neq 0$ tal que $A_{22} x_2 = 0$

$$\text{Sea} \quad b = -\underset{r\times(n-r)}{A_{12}} \underset{(n-r)\times 1}{x_2} \in \mathbb{R}^r$$

Como T es una matriz triangular de orden r, entonces las columnas de T generan $\mathbb{R}^r$, por el resultado anterior.

Como $b \in \mathbb{R}^r$, $b \in \operatorname{Im}(T)$ y por lo tanto existe un $x_1 \in \mathbb{R}^r$ tal que

$$Tx_1 = b = -A_{12} x_2$$

Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, donde x_1 y x_2 son las descritas arriba

$\therefore$ $x \neq 0$, ya que $x_2 \neq 0$, y además

$$Ax = \begin{bmatrix} T & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} Tx_1 + A_{12}x_2 \\ A_{22}x_2 \end{bmatrix} = \begin{bmatrix} -A_{12}x_2 + A_{12}x_2 \\ A_{22}x_2 \end{bmatrix}$$

$= \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Por tanto existe $x \neq 0$ tal que $Ax = 0$ ■

Por lo tanto concluimos que, para matrices triangulares y triangulares por bloques, necesitamos cuando menos m vectores independientes para generar $\mathbb{R}^m$.

Es decir, parece ser que para que

$$\mathbb{R}^m = R(A)$$

entonces A debe de tener cuando menos m columnas linealmente independientes.

Antes de continuar veremos qué es necesario para que

$$\mathbb{R}^m = [a_1, a_2, \ldots, a_n] = R(A)$$

Observemos que ésto implica

$$[e_1, e_2, \ldots, e_m] = [a_1, a_2, \ldots, a_n] = \mathbb{R}^m$$

sabemos que $[e_1, e_2, \ldots, e_m] = [a_1, a_2, \ldots, a_n]$
$\iff$ $e_i \in [a_1, a_2, \ldots, a_n]$, $1 \leq i \leq m$.

O sea que

$$\mathbb{R}^m = R(A) \iff e_i = A x_i \;, \quad x_i \in \mathbb{R}^n \quad 1 \leq i \leq m.$$

Esto lo podemos escribir en notación matricial :

$$[e_1|e_2|\cdots|e_m] = [Ax_1|Ax_2|\cdots|Ax_m]$$

$$= A[x_1|x_2|\cdots|x_m]$$

$$\therefore \qquad I_m = AX$$

donde X es la matriz $n \times m$ cuyas columnas son $x_1, x_2, \ldots, x_m$.

Es decir $\quad \mathbb{R}^m = [a_1, a_2, \ldots, a_n]$

$\Leftrightarrow \quad I = AX \quad$ para alguna matriz $X_{n\times m}$

Ahora, esto nos da información sobre la matriz X, ya que si tomamos algun vector $y \in \mathbb{R}^m$, $y \neq 0$ entonces:

$$y = Iy = AXy \neq 0$$

$$\Rightarrow A(Xy) \neq 0 \quad \Rightarrow Xy \neq 0$$

O sea que si X es tal que

$\underset{m}{I} = \underset{m\times n}{A}\underset{n\times m}{X}$ entonces las columnas de X

son linealmente independientes.

Podemos ver a las matrices A y X como funciones que operan sobre vectores de $\mathbb{R}^n$ y $\mathbb{R}^m$ respectivamente:

$$\begin{tikzcd}\mathbb{R}^m \arrow[r, bend left, "X"] & \mathbb{R}^n \arrow[r, bend left, "A"] & \mathbb{R}^m\end{tikzcd}$$

O sea que si $R(A) = \mathbb{R}^m \Rightarrow$ existe $X : \mathbb{R}^m \rightarrow \mathbb{R}^n$ con columnas independientes tal que $AX = I_m$.

Esto no nos dice nada acerca de la dependencia ó independencia lineal de las columnas de A. Esta situación es un caso particular del producto de dos matrices:

$$C_{m\times k} = A_{m\times n} B_{n\times k}$$

RESULTADO

Si $C_{m\times k} = A_{m\times n} B_{n\times k}$. Entonces:

Las columnas de C son linealmente independientes $\Rightarrow$ las columnas de B son linealmente independientes.

Demostración:

Supongamos $y \in \mathbb{R}^k$, $y \neq 0$

entonces $Cy \neq 0$ ya que C tiene columnas independientes.

$$\therefore \quad Cy = A(By) \neq 0$$

$\Rightarrow$ $By \neq 0$, ya que si By fuera igual a cero entonces tendríamos que

$$A(By) = A(0) = 0 \;!$$

$\therefore$ $By \neq 0$ $\therefore$ las columnas de B son linealmente independientes.

■

Sin embargo, si reflexionamos un poco más sobre el resultado anterior, veremos que de él podemos extraer una herramienta muy útil para investigar si las columnas de una matriz son linealmente independientes, puesto que el producto AB se puede interpretar como acción de A sobre B ó de B sobre A.

RESULTADO

Supongamos $C = AB$, con $N(B) = \{0\}$, es decir, tal que B tiene columnas independientes. Entonces

$$N(A) = \{0\} \Rightarrow N(C) = \{0\}$$

Demostración :

Supongamos $N(B) = \{0\}$ y $N(C) \neq \{0\}$

$\therefore$ $CX = 0$ para alguna $X \in \mathbb{R}^k$, $X \neq 0$

$\therefore$ $ABX = 0$, $X \neq 0$ $\therefore$ $A(BX) = 0$ $X \neq 0$

Como $N(B) = \{0\}$ y $X \neq 0 \Rightarrow BX \neq 0$

Sea $\bar{X} = BX$ $\therefore$ $A\bar{X} = 0$ con $\bar{X} \neq 0$

$\therefore$ $N(A) \neq \{0\}$

$\therefore$ $N(A) = \{0\} \Rightarrow N(C) = \{0\}$ cuando $N(B) = \{0\}$ ■

RESULTADO

Sea $\underset{m\times k}{C} = \underset{m\times n}{A} \; \underset{n\times k}{B}$. Supongamos $N(B) = \{0\}$.

Entonces $N(C) \neq \{0\} \Rightarrow N(A) \neq \{0\}$.

Es decir, si $C = AB$ y B tiene columnas independientes entonces, en el caso de que C tenga columnas dependientes, la "responsable" de esto es la matriz A.

Demostración :

Es la misma que para el resultado anterior.

RESULTADO

Sea $\underset{m\times k}{C} = \underset{m\times n}{A} \cdot \underset{n\times k}{B}$. Entonces

$N(C) \neq \{0\} \Rightarrow N(A) \neq \{0\}$ ó $N(B) \neq \{0\}$. Es decir,

si C tiene columnas dependientes entonces A tiene columnas dependientes ó B tiene columnas dependientes.

Demostración :

suponqamos $N(C) \neq \{0\}$. por lo tanto existe alquna $X \in \mathbb{R}^K$, $X \neq 0$ tal que

$$CX = 0 \quad \therefore \quad ABX = 0 \quad X \neq 0$$

Si $N(A) = \{0\} \Rightarrow A(BX) = 0 \quad X \neq 0$
$\Rightarrow BX = 0 \quad X \neq 0$
$\Rightarrow N(B) \neq \{0\}$

Si $N(B) = \{0\} \Rightarrow A(BX) = 0 \quad X \neq 0$
$\Rightarrow A\bar{X} = 0$ con $\bar{X} = BX \neq 0$
$\Rightarrow N(A) \neq \{0\}$.

OBSERVACIONES

El resultado que acabamos de demostrar nos permite estudiar la dependencia lineal desde dos puntos de vista.

Al multiplicar a la matriz A por la derecha por la matriz B, entonces estamos operando sobre las columnas de A:

Las columnas de $C = AB$ son combinaciones lineales de las columnas de A.

Si escogemos B de tal manera que sus columnas sean linealmente independientes, entonces C tiene columnas dependientes cuando A tiene columnas dependientes.

Supongamos que queremos investigar la dependencia lineal de las columnas de A. Entonces esta discusión nos sugiere a multiplicar a A por una matriz B con columnas independientes y tal que para $C = AB$ sea fácil contestar si sus columnas son linealmente independientes ó no. Si $N(C) \neq \{0\}$ entonces sabremos que las columnas de A son dependientes.

Ahora, supongamos que nos interesa examinar la dependencia lineal de las columnas de la matriz B.

Si premultiplicamos a B por la matriz A entonces :

Si A tiene columnas independientes y C=AB tiene columnas dependientes entonces necesariamente B tiene columnas dependientes.

Ahora, notemos que al premultiplicar a B por la matriz A, estamos operando sobre los renglones de B .

Es decir :

¡ Para investigar la dependencia lineal de las columnas de B , realizamos operaciones sobre los renglones de B !

Veremos que ésta es una forma muy adecuada de saber si las columnas de B son linealmente independientes ó no.

RESULTADO

Sea $\underset{m\times k}{C} = \underset{m\times n}{A} \cdot \underset{n\times k}{B}$. Supongamos $N(A)=\{0\}$.

Entonces $N(C)=\{0\} \iff N(B)=\{0\}$

Es decir, las columnas de C son independientes $\iff$ las columnas de B son independientes, cuando $N(A)=\{0\}$.

Demostración :

Supongamos $N(A) = \{0\}$. Entonces

$$\begin{aligned} N(C) = \{0\} &\iff Cx \neq 0 \ , \ x \neq 0 \\ &\iff A(Bx) \neq 0 \ , \ x \neq 0 \\ &\iff Bx \neq 0 \ , \ x \neq 0 \\ &\iff N(B) = \{0\} \end{aligned}$$

■

OBSERVACION

Si $C = AB$, entonces, el que B tenga columnas Independientes $\not\Rightarrow$ C tenga columnas independientes.

Ya que si $N(A) \neq \{0\}$ y $y \in \mathbb{R}^k$ es diferente de cero entonces $By \neq 0$, pero puede pasar que $Cy = A(By) = 0$.

Regresemos a nuestro problema original :

Dada A $m \times n$, cuándo $R(A) = \mathbb{R}^m$

Como hemos visto, tenemos razones para sospechar que A debe tener al menos m columnas linealmente independientes, y en tal caso tendríamos que $n \geq m$; pero si $n > m$ entonces las columnas restantes serían dependientes de las m columnas independientes, y entonces A debería tener columnas dependientes.

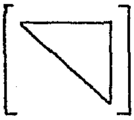
Observemos que si lo anterior es cierto, entonces toda matriz A $m\times n$, con $n > m$ debería tener columnas dependientes, puesto que m vectores independientes parecen ser suficientes para generar $\mathbb{R}^m$. Antes de demostrar este resultado veremos cómo transformar a A en una matriz para la cual es inmediato.

La forma en que abordaremos el problema será la siguiente:

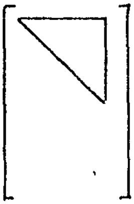
Transformaremos a A en una matriz en la que la dependencia lineal de sus columnas sea fácil de investigar, es decir, transformaremos a A en una matriz triangular.

RESULTADO

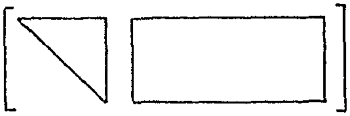
Sea $A_{m\times n}$ una matriz. Entonces existe una matriz $B_{m\times m}$, con $N(B) = \{0\}$, tal que BA tiene una de las siguientes formas:



$m = n$



$m > n$



$m < n$

Demostración:

sea $A = [a_1|a_2|\ldots|a_n]$, $a_j \in \mathbb{R}^m$ $1 \leq j \leq n$.

En el primer paso haremos lo siguiente:

Si $a_1 = 0$ entonces no hacemos nada.
Si $a_1 \neq 0$, sea $|\alpha_{k_1}| = \max\,[|\alpha_{11}|, |\alpha_{21}|, \ldots, |\alpha_{m1}|]$.

$\therefore$ $\alpha_{k_1} \neq 0$.

Intercambiemos el primer renglón de A con el k_1-ésimo :

$$C_1 = P_{1K_1} A = [\gamma_{ij}]$$

P_{1K_1} es la permutación que efectúa el intercambio de renglones.

$\therefore$ $\gamma_{11} \neq 0$, ya que $\gamma_{11} = \alpha_{k_1 1} \neq 0$

La matriz

$$L_1 = \begin{bmatrix} 1 & & & & \\ -\frac{\gamma_{21}}{\gamma_{11}} & 1 & & & \\ -\frac{\gamma_{31}}{\gamma_{11}} & 0 & 1 & & \\ \vdots & & & & \\ -\frac{\gamma_{m1}}{\gamma_{11}} & 0 & 0 & \cdots & 1 \end{bmatrix}$$

es tal que

$$L_1 P_{1k_1} A = \begin{bmatrix} * & * & * & \cdots & * & * & \cdots & * \\ 0 & * & * & \cdots & * & * & \cdots & * \\ 0 & * & * & \cdots & * & * & \cdots & * \\ \vdots & & & & & & & \\ 0 & * & * & \cdots & * & * & \cdots & * \end{bmatrix} = A_1 = \left[\begin{array}{c|c} * & * \; * \; \cdots \; * \; * \; \cdots \; * \\ \hline 0 & \\ 0 & \\ \vdots & \bar{A}_1 \\ 0 & \end{array}\right]$$

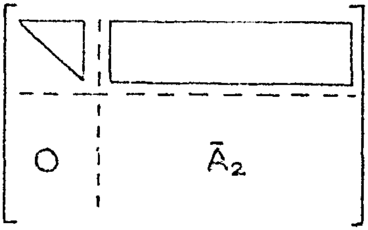
$\bar{A}_1$ es una matriz $(m-1) \times (n-1)$ a la cual le podemos aplicar el mismo proceso que a la matriz A.

Si la primera columna de $\bar{A}_1$ es el vector $0 \in \mathbb{R}^{m-1}$, entonces no hacemos nada.

Si la primera columna de $\bar{A}_1$ es diferente de cero, entonces intercambiamos el segundo renglón de A_1 - que es el primer renglón de $\bar{A}_1$ - con el renglón k_2 de A_1. El renglón k_2 de A_1 es tal que contiene al mayor elemento en valor absoluto de la primera columna de $\bar{A}_1$. Esto lo hacemos multiplicando por la permutación de orden m P_{2k_2}

Por lo tanto existe una matriz triangular inferior L_2 de orden m tal que

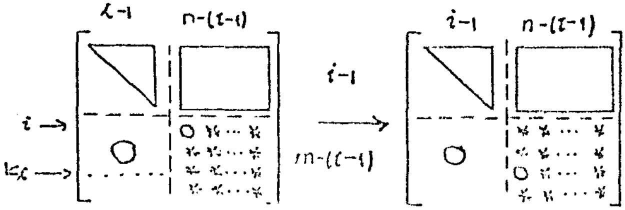
$$L_2 P_{2k_2} A_1 = \begin{bmatrix} * & * & * & \cdots & * & * & \cdots & * \\ 0 & * & * & \cdots & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \cdots & * & * & \cdots & * \end{bmatrix} = A_2 =$$



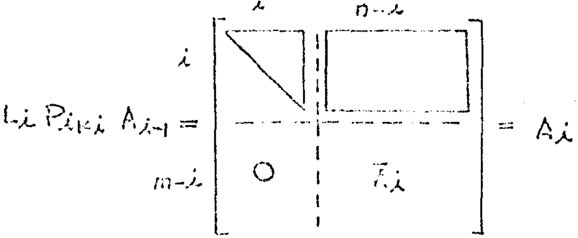
En el i-ésimo paso vamos a aplicarle el proceso a la matriz $\bar{A}_{i-1}$, que es una matriz $(m-(i-1)) \times (n-(i-1))$.

Si la primera columna de $\hat{A}_{i-1}$ es el vector cero de $\mathbb{R}^{m-(i-1)}$ entonces no hacemos nada.

Si la primera columna de $\hat{A}_{i-1}$ es diferente de cero entonces intercambiamos el i-ésimo renglón de A_{i-1}, que es el primer renglón de $\bar{A}_{i-1}$, con el k_i-ésimo renglón de A_{i-1}; k_i es tal que el k_i-ésimo renglón de A_{i-1} contiene al mayor elemento, en valor absoluto, de la primera columna de $\bar{A}_{i-1}$. Esta operación la llevamos a cabo multiplicando por la permutación de orden m P_{ik_i}



Después premultiplicamos por una matriz triangular inferior L_i de orden m, con 1's en la diagonal tal que



al cabo del paso j-ésimo termina este proceso, donde

si $m < n$ entonces $j = m-1$
si $m = n$ entonces $j = m-1$
y si $m > n$ entonces $j = n-1$

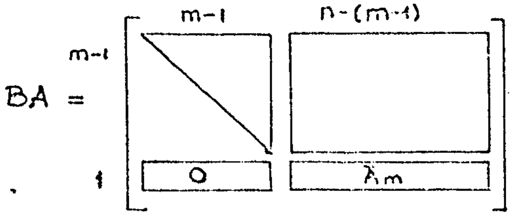
Sea $B = L_j P_{jk_j} \cdots L_2 P_{2k_2} L_1 P_{1k_1}$

El factor $L_i P_{ik_i}$, $1 \le i \le j$, aparece únicamente cuando la primera columna de $\bar{A}_{i-1}$ es diferente de cero. Además $L_i P_{ik_i}$ es el producto de dos matrices cuyas columnas son linealmente independientes, $\therefore$ $N(L_i P_{ik_i}) = \{0\}$ $1 \le i \le j$.

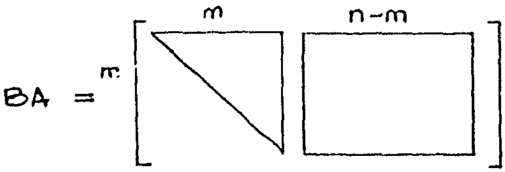
$\therefore$ B tiene columnas independientes: $N(B) = \{0\}$.

Veremos ahora qué forma tiene la matriz BA

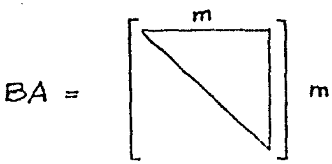
Si $m < n$ entonces el proceso termina en el paso $m-1$, y tenemos



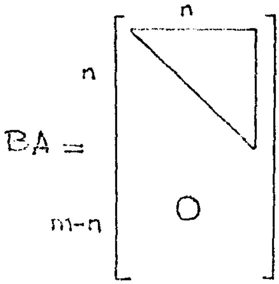
Podemos reagrupar a BA por bloques para obtener



Si $m=n$ entonces también terminamos el proceso el paso $m-1$, pero en este caso BA es triangular superior:



Si $m>n$, entonces concluimos el proceso al cabo de $n-1$ pasos y BA es de la forma



EJEMPLO

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

Primero intercambiamos el primer renglón con el cuarto renglón:

$$P_{14} A = \begin{bmatrix} 2 & 1 & 4 & 0 & 0 & 2 \\ 1 & 0 & 1 & 3 & 2 & 1 \\ 1 & 1 & -1 & 4 & 5 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Sea $$L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1/2 & 1 & 0 & 0 \\ -1/2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\therefore \quad L_1 P_{14} A = \begin{bmatrix} 2 & 1 & 4 & 0 & 1 & 2 \\ 0 & -1/2 & -1 & 3 & 2 & 0 \\ 0 & 1/2 & -3 & 4 & 5 & -1 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} = A_1$$

Ahora intercambiamos el 2° y el 4° renglón:

$$P_{24}\,A_1 = \begin{bmatrix} 2 & 1 & 4 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & -3 & 4 & 5 & -1 \\ 0 & -1/2 & -1 & 3 & 2 & 0 \end{bmatrix}$$

Sea $$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1/2 & 1 & 0 \\ 0 & 1/2 & 0 & 1 \end{bmatrix}$$

$$L_2 P_{24} A_1 = \begin{bmatrix} 2 & 1 & 4 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 4 & 5 & -1 \\ 0 & 0 & -1 & 3 & 2 & 0 \end{bmatrix} = A_2$$

Sea $$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/3 & 1 \end{bmatrix}$$

$$\therefore\quad L_3 A_2 = \begin{bmatrix} 2 & 1 & 4 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 4 & 5 & -1 \\ 0 & 0 & 0 & \frac{5}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

$$\therefore \quad B = L_3 L_2 P_{24} L_1 P_{14}$$

$$\therefore \quad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{3} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 0 & \frac{1}{3} & 0 & 1 \end{bmatrix} P_{24} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} P_{14}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -\frac{1}{5} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 1 & -\frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & -\frac{1}{2} \\ \frac{2}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$

$$\therefore \quad B = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & 1 & -\frac{1}{2} \\ \frac{2}{3} & 1 & -\frac{1}{3} & -\frac{1}{3} \end{bmatrix}$$ es tal que

$$BA = \begin{bmatrix} 2 & 1 & 4 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3 & 4 & 5 & -1 \\ 0 & 0 & 0 & 5/3 & 1/3 & 1/3 \end{bmatrix}$$

Ahora demostraremos que si tenemos n vectores en $\mathbb{R}^m$, con $n > m$, entonces éstos son linealmente dependientes:

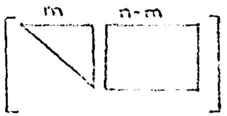
RESULTADO FUNDAMENTAL

Sea $A = [a_1 | a_2 | \cdots | a_n]$ con $a_j \in \mathbb{R}^m$, $1 \leq j \leq n$.
Si $n > m$ entonces $\{a_1, a_2, \ldots, a_n\}$ es linealmente dependiente.

Demostración:

Por el resultado anterior existe una matriz B $m \times m$, con $N(B) = \{0\}$ tal que

$BA =$



$$\overset{m}{=} \left[\begin{array}{c|c} \overset{m}{U} & \overset{n-m}{A_{12}} \end{array}\right] = C$$

Si las columnas de U son linealmente dependientes $\Rightarrow$ las columnas de C también lo son, y por lo tanto los de A también.

Si las columnas de U son linealmente independientes, entonces generan $\mathbb{R}^m$, y las $n-m>0$ restantes columnas de C están en $R(U)=\mathbb{R}^m$. De aquí concluimos que las columnas de C son linealmente dependientes, y por lo tanto que las columnas de A son dependientes ∎

COROLARIO

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$. Entonces existe una matriz A $m \times n$ tal que $\mathcal{S} = R(A)$

Demostración :

1° Sea $a_1 \in \mathcal{S}$, $a_1 \neq 0$

2° Si $\mathcal{S} \subseteq [a_1]$ ya acabamos.

Si $\mathcal{S} \not\subseteq [a_1]$, sea $a_2 \notin [a_1]$

3° Si $\mathcal{S} \subseteq [a_1, a_2]$ ya acabamos.

Si $\mathcal{S} \not\subseteq [a_1, a_2]$ sea $a_3 \notin [a_1, a_2]$

⋮

etc.

Este proceso tiene que acabar después de un número finito de pasos, digamos n, con $n \leq m$, ya que de lo contrario tendríamos más de m vectores de $\mathbb{R}^m$ linealmente independientes.■

RESULTADO

Supongamos $\mathcal{S} = R(A_1) = R(A_2)$, donde A_1 es una matriz $m \times n_1$ y A_2 es $m \times n_2$. Si A_1 y A_2 tienen columnas independientes entonces $n_1 = n_2$.

Demostración:

Supongamos $n_1 < n_2$. Como $\mathcal{S} = R(A_1) = R(A_2) \Rightarrow$ las columnas de A_2 son combinaciones lineales de las columnas de A_1:

$$\underset{m\times n_2}{A_2} = \underset{m\times n_1}{A_1}\ \underset{n_1\times n_2}{X_2} \qquad \text{para } X_2 \ n_1\times n_2$$

Como $N(A_2) = N(A_1) = \{0\} \Rightarrow N(X_2) = \{0\}$!

$\Rightarrow$ las columnas de X_2 son linealmente independientes, lo cual es imposible ya que $n_1 < n_2$.

$\therefore \quad n_2 \leq n_1$

andlogamente, si $n_2 < n_1$ entonces

$$\underset{m\times n_1}{A_1} = \underset{m\times n_2}{A_2}\ \underset{n_2\times n_1}{X_1}\ , \text{ para alguna } X_1\ n_2\times n_1$$

$$N(A_1) = N(A_2) = \{0\} \Rightarrow N(X_1) = \{0\} \ !$$

Esto es imposible ya que $n_2 < n_1$

$$\therefore \quad n_1 \leq n_2$$

$$\therefore \quad n_1 \leq n_2 \quad y \quad n_2 \leq n_1 \quad \therefore \quad n_1 = n_2 \ \blacksquare$$

DEFINICION

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$. Decimos que $\{a_1, a_2, \dots, a_r\}$ es una BASE de $\mathcal{S}$ si

1) $\mathcal{S} = [a_1, a_2, \dots, a_r]$

2) $\{a_1, a_2, \dots, a_r\}$ es linealmente independiente.

RESULTADO

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$. Si $\{a_1, a_2, \dots, a_{n_1}\}$ y $\{b_1, b_2, \dots, b_{n_2}\}$ son dos bases de $\mathcal{S}$, entonces $n_1 = n_2$. Es decir, cualesquiera dos bases de $\mathcal{S} \subseteq \mathbb{R}^m$ tienen el mismo número de elementos.

Demostración:

Si $A_1 = [a_1|a_2|\cdots|a_{n_1}]$ y $A_2 = [b_1|b_2|\cdots|b_{n_2}]$

$\Rightarrow \quad \mathcal{S} = R(A_1) = R(A_2)$

y como las columnas de A_1 y A_2 son linealmente independientes $\rightarrow n_1 = n_2$

■

DEFINICION

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$. El número de elementos de cualquier base de $\mathcal{S}$ se llama DIMENSION de $\mathcal{S}$. A la dimensión de $\mathcal{S}$ la denotaremos por dim $(\mathcal{S})$.

COROLARIO

Si $\{a_1, a_2, \ldots, a_m\}$ son m vectores linealmente independientes de $\mathbb{R}^m$, entonces

$R(A) = [a_1, a_2, \ldots, a_m] = \mathbb{R}^m$, donde

$A = [a_1|a_2|\cdots|a_m]$.

Es decir, si A es de orden m con columnas independientes, entonces éstas forman una base de $\mathbb{R}^m$.

Demostración:

Sea $b \in \mathbb{R}^m$ y sea

$$B = [A \mid b] = [a_1 | a_2 | \cdots | a_m | b]$$

B es $m \times (m+1)$ $\not\Rightarrow$ tiene columnas independientes

$\Rightarrow$ existe $y \in \mathbb{R}^{m+1}$, $y \neq 0$ tal que $By = 0$

Si $y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_{m+1} \end{bmatrix}$ entonces

$$\eta_1 a_1 + \eta_2 a_2 + \cdots + \eta_m a_m + \eta_{m+1} b = 0$$

Si $\eta_{m+1} = 0$ entonces

$$\eta_1 a_1 + \cdots + \eta_m a_m = A \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix} = 0$$

Como A tiene columnas independientes $\Rightarrow$
$\eta_1 = \eta_2 = \cdots = \eta_m = 0$ $\Rightarrow$ $y = 0$!, pero esto no es posible $\therefore$ $\eta_{m+1} \neq 0$.

$$\Rightarrow \quad b = \left(\frac{-\eta_1}{\eta_{m+1}}\right) a_1 + \left(\frac{-\eta_2}{\eta_{m+1}}\right) a_2 + \cdots + \left(-\frac{\eta_m}{\eta_{m+1}}\right) a_m$$

$\therefore \quad b = Ax$, donde $x \in \mathbb{R}^m$

$\therefore \quad b \in \mathbb{R}^m \geqq R(A)$

$\therefore \quad R(A) = \mathbb{R}^m$ ■

COROLARIO

Sea A $m \times n$. Si B es $n \times n$ con columnas independientes (i.e. $N(B) = \{0\}$) entonces

$$R(AB) = R(A)$$

Demostración:

Es claro $R(AB) \subseteq R(A)$.

Si $y \in R(A) \Rightarrow y = Ax$ para alguna $x \in \mathbb{R}^n$.

Por el corolario anterior $\mathbb{R}^n = R(B)$ $\therefore x \in R(B)$

$\therefore \quad x = Bs$ para alguna $s \in \mathbb{R}^n$

$\Rightarrow \quad y = A(Bs) = (AB)s \Rightarrow y \in R(AB)$

$\therefore \quad R(AB) = R(A)$ □

COROLARIO

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$ tal que $\dim(\mathcal{S}) = k < m$. Entonces $(k+1)$ vectores de $\mathcal{S}$ son linealmente dependientes. Es decir :

si $\{s_1, s_2, \ldots, s_{k+1}\} \subseteq \mathcal{S} \Rightarrow \{s_1, s_2, \ldots, s_{k+1}\}$ es linealmente dependiente.

Demostración

Sea $\{\tilde{s}_1, \tilde{s}_2, \ldots, \tilde{s}_k\}$ una base de $\mathcal{S}$, y sean

$$S = [s_1 | s_2 | \cdots | s_{k+1}] \quad , \quad \tilde{S} = [\tilde{s}_1 | \tilde{s}_2 | \cdots | \tilde{s}_k] \;.$$

$\therefore$ $s_i \in \mathcal{S} \Rightarrow s_i = \tilde{S} x_i$ para alguna $x_i \in \mathbb{R}^k$

$$\therefore \quad \underset{m\times(k+1)}{S} = [\tilde{S}x_1 | \tilde{S}x_2 | \cdots | \tilde{S}x_{k+1}] = \underset{m\times k}{\tilde{S}} \; \underset{k\times(k+1)}{X}$$

Como X es $k\times(k+1) \Rightarrow$ tiene columnas dependientes

$\Rightarrow$ $\tilde{S}X$ tiene columnas dependientes

$\therefore$ S tiene columnas dependientes.

□

COROLARIO

Sea $\mathcal{S} \subseteq \mathbb{R}^m$ un subespacio de $\mathbb{R}^m$ de dimensión $k \leq m$.

Si $\{s_1, s_2, \ldots, s_k\} \subseteq \mathcal{S}$ es linealmente independiente, entonces

$$\mathcal{S} = [s_1, s_2, \ldots, s_k]$$

Ya vimos que si A es $m \times n$, con $n > m$, entonces las columnas de A son linealmente dependientes. Por lo tanto, si A $m \times n$ tiene columnas independientes $\Rightarrow$ $n \leq m$.

RESULTADO

Si A $m \times n$ tiene columnas independientes, entonces existe una matriz $B_{m \times m}$, con $N(B) = \{0\}$ tal que

$$BA = \begin{array}{cc} & n \\ \begin{array}{c} n \\ \\ m-n \end{array} & \left[\begin{array}{c} \diagdown \\ \\ 0 \end{array}\right] \end{array} = \left[\begin{array}{c} T \\ \hline 0 \end{array}\right] , \quad \text{con } N(T) = \{0\},$$

Es consecuencia del resultado fundamental.

RESULTADO

Si A es $m \times n$, con $m > n$, entonces los renglones de A son linealmente dependientes.

Demostración:

Los renglones de A son linealmente independientes $\Leftrightarrow$ las columnas de A^t son linealmente independientes.

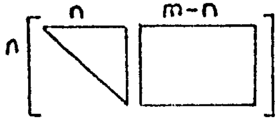
Si A es $m \times n$ entonces A^t es $n \times m$

$\therefore$ Si $m > n \Rightarrow$ las columnas de A^t son dependientes $\Rightarrow$ los renglones de A son dependientes ▫

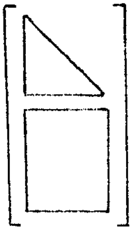
Para saber si los renglones de A son linealmente independientes o no podemos aplicar el procedimiento que hemos estado usando a la matriz A^t

Si A es una matriz $m \times n$ entonces A^t es una matriz $n \times m$.

Por ejemplo, si $m > n$ entonces sabemos que existe una matriz $\bar{B}_{n \times n}$ tal que $N(\bar{B}) = \{0\}$ y

$\bar{B} A^t =$ 

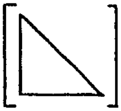
si transponemos, vemos que

$A \bar{B}^t =$  $= AB \quad ; \quad B = \bar{B}^t$

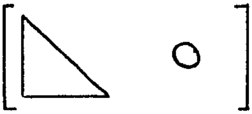
De aquí vemos que para investigar la dependencia lineal de los renglones de A podemos operar sobre las columnas de A; este enfoque nos puede ser de utilidad también.

Desarrollaremos una técnica para

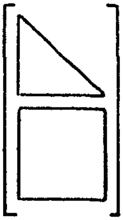
reducir a la matriz A a una de las siguientes formas :



$m=n$



$m<n$



$m>n$

El procedimiento es el mismo que introducimos para operar con renglones, pero aplicado a las columnas. Lo presentamos explícitamente para mayor claridad.

Primero resolveremos el siguiente problema

PROBLEMA

Dada una matriz A $m \times n$ encontrar una matriz B $n \times n$, con $N(B) = \{0\}$, tal que

$$AB = \begin{bmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ * & * & * & \cdots & * \end{bmatrix}$$

Supongamos $A = \begin{bmatrix} r_1{}^t \\ r_2{}^t \\ \vdots \\ r_m{}^t \end{bmatrix}$

Si $r_1^t = 0^t$, entonces ya acabamos.

Si $r_1^t \neq 0^t$, entonces sea

$$|\alpha_{1k}| = \max\{|\alpha_{11}|, |\alpha_{12}|, \ldots, |\alpha_{1n}|\}$$

$\therefore \quad \alpha_{1k} \neq 0$.

Vamos a intercambiar la primera columna de A con la k-ésima, posmultiplicando a A por la permutación P_{1k}:

$$AP_{1k} = [a_k | a_2 | \cdots | a_{k-1} | a_1 | a_{k+1} | \cdots | a_{n-1} | a_n]$$
$$= [\tilde{\alpha}_{ij}] = [\tilde{a}_1 | \tilde{a}_2 | \ldots | \tilde{a}_n] = \begin{bmatrix} \tilde{r}_1^t \\ \tilde{r}_2^t \\ \vdots \\ \tilde{r}_m^t \end{bmatrix}$$

Como $a_k \neq 0 \Rightarrow \tilde{a}_1 \neq 0$ y $\tilde{\alpha}_{11} \neq 0$

sea $\tilde{B} = [\tilde{b}_1 | \tilde{b}_2 | \ldots | \tilde{b}_n]$

Encontraremos $b_1, b_2, \ldots, b_n$ tal que

$$AP_{1k}\tilde{B} = \begin{bmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & * & \cdots & * \\ * & * & * & \cdots & * \\ \vdots & & & & \\ * & * & * & \cdots & * \end{bmatrix}$$

Tenemos :

$$AP_{1K}\tilde{B} = \left[AP_{1K}\tilde{b}_1 \,|\, AP_{1K}\tilde{b}_2 \,|\cdots|\, AP_{1K}\tilde{b}_n \right]$$

$$= \left[\tilde{a}_1 \,|\tilde{a}_2|\cdots|\tilde{a}_n\right]\tilde{B} = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}$$

· En la primera columna no nos interesa hacer ceros, por lo tanto no tiene ningún caso alterarla :

$$AP_{1K}\tilde{b}_1 = \tilde{\beta}_{11}\tilde{a}_1 + \tilde{\beta}_{21}\tilde{a}_2 + \cdots + \tilde{\beta}_{n1}\tilde{a}_n$$

∴ Nos conviene escoger

$$\tilde{\beta}_{11} = 1$$
$$\tilde{\beta}_{21} = \tilde{\beta}_{31} = \cdots = \tilde{\beta}_{n1} = 0$$

$$\therefore \quad \tilde{b}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Para la segunda columna tenemos

$$AP_{1K}\tilde{b}_2 = \tilde{\beta}_{12}\tilde{a}_1 + \tilde{\beta}_{22}a_2 + \cdots + \tilde{\beta}_{n2}\tilde{a}_n$$

ahora , teníamos que

$$\tilde{a}_1 = \begin{bmatrix} \tilde{\alpha}_{11} \\ \tilde{\alpha}_{21} \\ \vdots \\ \tilde{\alpha}_{m1} \end{bmatrix} \qquad \tilde{a}_2 = \begin{bmatrix} \tilde{\alpha}_{12} \\ \tilde{\alpha}_{22} \\ \vdots \\ \tilde{\alpha}_{m2} \end{bmatrix}$$

$$\therefore \; \tilde{\beta}_{12}\tilde{a}_1 + \beta_{22}\tilde{a}_2 = \begin{bmatrix} \tilde{\beta}_{12}\tilde{\alpha}_{11} + \tilde{\beta}_{22}\tilde{\alpha}_{12} \\ * \\ \vdots \\ * \end{bmatrix}$$

$\therefore$ Si podemos lograr que

$$\tilde{\beta}_{12}\tilde{\alpha}_{11} + \tilde{\beta}_{22}\tilde{\alpha}_{12} = 0$$

entonces podemos tomar

$$AP_{1K}\tilde{b}_2 = \tilde{\beta}_{12}\tilde{a}_1 + \tilde{\beta}_{22}\tilde{a}_2 \,,$$

tomando $\tilde{\beta}_{32} = \tilde{\beta}_{42} = \cdots = \tilde{\beta}_{n2} = 0$

Sí podemos tomar $\tilde{\beta}_{12}$ y $\tilde{\beta}_{22}$ tal que $\tilde{\beta}_{12}\tilde{\alpha}_{11} + \tilde{\beta}_{22}\tilde{\alpha}_{12} = 0$, ya que

$\tilde{\beta}_{12}\tilde{\alpha}_{11} + \tilde{\beta}_{22}\tilde{\alpha}_{12} = 0$ y tomando $\tilde{\beta}_{22} = 1$

$$\Rightarrow \tilde{\beta}_{12}\tilde{\alpha}_{11} = -\tilde{\alpha}_{12}$$

$\Rightarrow \tilde{\beta}_{12} = -\dfrac{\tilde{\alpha}_{12}}{\tilde{\alpha}_{11}}$, y esto es posible ya que $\tilde{\alpha}_{11} \neq 0$

$$\therefore \quad \tilde{b}_2 = \begin{bmatrix} -\dfrac{\tilde{\alpha}_{12}}{\tilde{\alpha}_{11}} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

similarmente, se puede ver que

$$\tilde{b}_3 = \begin{bmatrix} -\dfrac{\tilde{\alpha}_{13}}{\tilde{\alpha}_{11}} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \tilde{b}_n = \begin{bmatrix} -\dfrac{\tilde{\alpha}_{1n}}{\tilde{\alpha}_{11}} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Por lo tanto, la matriz $\tilde{B}$ es de la forma

$$\tilde{B} = \begin{bmatrix} 1 & \beta_{12} & \beta_{13} & \cdots & \beta_{1n} \\ & 1 & 0 & & 0 \\ & & 1 & & 0 \\ & & & & \vdots \\ & & & & 1 \end{bmatrix}$$

donde $\beta_{1j} = -\dfrac{\tilde{\alpha}_{1j}}{\tilde{\alpha}_{11}} \qquad 1 \leq j \leq n$

observar que $\tilde{B}$ es una matriz triangular superior, con columnas y renglones linealmente independientes.

$$\text{sea} \quad B = P_{1k}\tilde{B}$$

$$\therefore \quad N(B) = \{0\} \quad \text{y}$$

$$AB = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}$$

Ahora, supongamos que tenemos una matriz A mxn.

Por el resultado que acabamos de probar, existe una matriz B_1 tal que

$$AB_1 = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix} = \begin{array}{c} \\ 1 \\ m-1 \end{array} \!\! \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} * & 0 \\ \hline * & \\ \vdots & \bar{A}_1 \\ * & \end{array}\right] \end{array} = A_1$$

donde B_1 puede ser la matriz I_n, en el caso de que el primer renglon de A sea 0^t, ó puede ser el producto de una permutación y una matriz triangular superior.

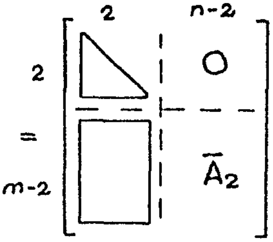
$\bar{A}_1$ es una matriz $(m-1)\times(n-1)$ y le podemos aplicar el mismo proceso que a A:

Sea $\bar{B}_2$ una matriz $(n-1)\times(n-1)$ tal que

$$\bar{A}_1\bar{B}_2 = \begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & & & \\ * & * & \cdots & * \end{bmatrix}$$

Si $B_2 = \begin{bmatrix} I_1 & 0 \\ 0 & \bar{B}_2 \end{bmatrix}$ entonces

$$A_1B_2 = \begin{bmatrix} * & 0 & 0 & \cdots & 0 \\ * & * & 0 & \cdots & 0 \\ * & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \\ * & * & * & \cdots & * \end{bmatrix} =$$



En general, para el i-ésimo paso vamos a tener que

$$B_i = \begin{array}{c|c|l} & & n-(i-1) \\ \hline \\ & I_{i-1} \quad 0 & i-1 \\ & 0 \quad \bar{B}_i & n-(i-1) \end{array}$$

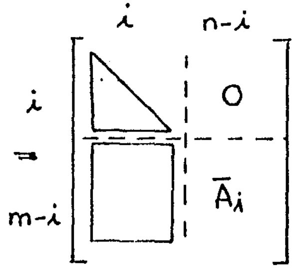
$$B_i = \left[\begin{array}{c|c} I_{i-1} & 0 \\ \hline 0 & \bar{B}_i \end{array}\right]\begin{array}{l} i-1 \\ n-(i-1) \end{array}$$

donde

$$\bar{A}_{i-1}\bar{B}_i = \overset{n-(i-1)}{\begin{bmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{bmatrix}} m-(i-1)$$

Y

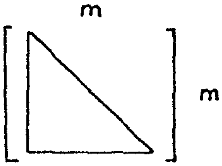
$A_{i+1}B_i =$



$= A_i$

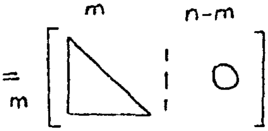
Si $m=n$ entonces después de $m-1$ pasos llegamos a que existe $B_{n\times n}$, $N(B)=\{0\}$:

$AB =$



Si $m<n$, entonces después de $m-1$ pasos tenemos que existe una matriz $B_{n\times n}$ tal que

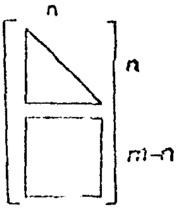
$AB =$



con $N(B)=\{0\}$

Finalmente si $m>n$ entonces para alguna $B_{n\times n}$; $N(B)=\{0\}$

$AB =$



RESULTADO

Si A $m \times n$ tiene renglones independientes $\Rightarrow$ $m \leq n$, y $\therefore$ existe una matriz B $n \times n$ tal que $N(B) = \{0\}$ y AB es de la forma

$$AB = \left[\begin{array}{c|c} \triangle & 0 \end{array}\right]$$

CAPITULO III

TEORIA DE ECUACIONES

TEORIA DE ECUACIONES

Supongamos que tenemos el sistema de ecuaciones

$$Ax=b$$

para cierta matriz A $m \times n$ y un vector $b \in \mathbb{R}^m$.

Sabemos que esta ecuación tiene solución cuando b pertenece a la imagen de A:

$$\text{existe } x \in \mathbb{R}^n \text{ tal que } Ax=b \iff b \in R(A)$$

¿ Cuando tendremos que $Ax=b$ tiene solución para cualquier $b \in \mathbb{R}^m$? Es decir ¿ cuando $R(A)=\mathbb{R}^m$?

Podemos ver a la matriz A como una función

$$A: \mathbb{R}^n \to \mathbb{R}^m$$

si $m>n$ entonces no puede pasar que $\mathbb{R}^m=R(A)$

si $m \leq n$ entonces es posible que $\mathbb{R}^m=R(A)$

También podemos ver a la matriz A como una función

$$A : \mathbb{R}^{m*} \to \mathbb{R}^{n*}$$

y analizar el sistema

$$y^t A = c^t \qquad \text{para} \quad c \in \mathbb{R}^{n*}$$

si $m < n$ entonces no puede pasar que $\mathbb{R}^n = R(A^t)$
pero
si $m \geq n$ entonces es posible que $\mathbb{R}^n = R(A^t)$

Por lo tanto vemos que únicamente para el caso $m = n$ puede darse al mismo tiempo que tengamos

$$R(A) = \mathbb{R}^m$$
$$R(A^t) = \mathbb{R}^n = \mathbb{R}^m$$

Veremos si es que hay alguna liga más fuerte entre los renglones y las columnas de una matriz.

En el capítulo anterior vimos que para matrices triangulares T tenemos :

Las columnas de T son linealmente independientes
$\Leftrightarrow$ Los renglones de T son linealmente independientes
$\Leftrightarrow$ Los elementos de la diagonal de T son diferentes de cero.

De aquí deducimos que si T es una matriz triangular con $N(T)=\{0\}$, entonces el número de renglones y de columnas linealmente independientes de T es el mismo.

Veremos si ésto se cumple para matrices cuadradas en general.

RESULTADO

Sea A una matriz de orden n. Entonces

$$R(A)=\mathbb{R}^n \Leftrightarrow R(A^t)=\mathbb{R}^n$$

Demostración:

$R(A)=\mathbb{R}^n \Rightarrow I_n = AX$ para alguna $X_{n\times n}$, $N(X)=\{0\}$

$\Rightarrow y^t I_n = y^t = (y^t A)X \neq 0^t$ si $y\neq 0$

$\Rightarrow y^t A \neq 0^t$ si $y \neq 0$

$\Rightarrow A^t y \neq 0$ si $y\neq 0$

$\Rightarrow$ las columnas de A^t son independientes, y como son $n \Rightarrow$ generan $\mathbb{R}^n$

$R(A^t)=\mathbb{R}^n \Rightarrow YA = I_n$ para alguna $Y_{n\times n}$, $N(Y)=\{0\}$

$\Rightarrow Y(Ax) = I_n x = x \neq 0$ si $x \neq 0$

$$\Rightarrow \quad Ax \neq 0 \qquad \text{si} \quad x \neq 0$$

$$\Rightarrow \quad R(A) = \mathbb{R}^n$$

■

RESULTADO

Sea A una matriz de orden n. Entonces

$$N(A) = \{0\} \Leftrightarrow N(A^t) = \{0\}.$$

Este resultado es equivalente al anterior.

■

Por lo tanto, si A es una matriz cuadrada, con $N(A) = \{0\}$, entonces A tiene igual número de columnas y renglones linealmente independientes.

Seguiremos analizando algunos casos para ver si ésto se cumple para otros tipos de matrices ó si únicamente vale para las matrices cuadradas cuyo núcleo es cero.

RESULTADO

Sea A una matriz $m \times n$. Si $R(B) = \mathbb{R}^n$ entonces

$$R(A^t) = R((BA)^t) = R(A^t B^t).$$

donde B es $n \times n$

Demostración:

Es claro que $R((BA)^t) \subseteq R(A^t)$.

Si $y \in R(A^t) \Rightarrow y = A^t x'$ para alguna $x' \in \mathbb{R}^n$

Pero $\mathbb{R}^n = R(B) = R(B^t)$ por uno de los resultados que acabamos de probar

$\therefore$ $x' \in R(B^t)$ $\therefore$ $x' = B^t x$ para alguna $x \in \mathbb{R}^n$

$$\therefore \quad y = A^t x' = A^t B^t x = (BA)^t x$$

$$\therefore \quad y \in R((BA)^t) \quad \therefore \quad R(A^t) \subseteq R((BA)^t)$$

$$\therefore \quad R(A^t) = R((BA)^t)$$

■

LEMA 1

Si $A = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[A_1 \mid A_2 \right] \end{array} m$ y $R(A_2) \subseteq R(A_1)$

entonces $R(A) = R(A_1)$.

Demostración:

$R(A_1) \subseteq R(A)$ siempre.

Si $y \in R(A) \Rightarrow y = A_1 x_1 + A_2 x_2$ con $x_1 \in \mathbb{R}^r$, $x_2 \in \mathbb{R}^{n-r}$

como $A_2X_2 \in R(A_2) \subseteq R(A_1) \Rightarrow A_2X_2 = A_1X_1'$, $X_1' \in \mathbb{R}^r$

$$\therefore \; y = A_1X_1 + A_1X_1' = A_1(X_1 + X_1') \Rightarrow y \in R(A_1)$$

$$\therefore \; R(A) \subseteq R(A_1) \qquad \therefore \; R(A) = R(A_1)$$

■

COROLARIO

Si $A = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]$ y $R(A_2^t) \subseteq R(A_1^t)$

entonces $R(A^t) = R(A_1^t)$

■

RESULTADO

Sea $A \in M(m,n)$. Si B es una matriz de orden n tal que $N(B) = \{0\}$ entonces

$$\dim(R(AB)^t) = \dim(R(A^t))$$

Es decir, A y AB tienen el mismo número de renglones independientes cuando B es cuadrada y $N(B) = \{0\}$. Observar que B está operando sobre las columnas, y no los renglones, de la matriz A.

Demostración:

Sean $r_{i_1}^t, r_{i_2}^t, \ldots, r_{i_k}^t$ renglones de A, entonces

$$\{r_{i_1}^t, r_{i_2}^t, \ldots, r_{i_k}^t\} \text{ es linealmente independiente}$$

$$\Leftrightarrow \{r_{i_1}^t B, r_{i_2}^t B, \ldots, r_{i_k}^t B\} \text{ es linealmente independiente.}$$

ya que como B es de orden y $N(B) = \{0\} \Rightarrow N(B^t) = \{0\}$ y por lo tanto

$$\sum_{l=i_1}^{i_k} \alpha_l r_l^t = 0^t \Leftrightarrow \left(\sum_{j=i_1}^{i_k} \alpha_l r_l^t \right) B = 0^t$$

$$\Leftrightarrow \sum_{j=i_1}^{i_k} \alpha_l (r_l^t B) = 0^t$$

Por lo tanto

$$\dim (R(AB)^t) = \dim (R(A^t))$$

$\boxtimes$

RESULTADO

Sea A una matriz $m \times n$ con columnas independientes y tal que $m > n$. Entonces A tiene n renglones independientes.

Demostración :

Hay que observar primero que pedimos $m > n$, ya que si $m < n$ entonces las columnas de A son dependientes. El caso $m = n$ lo excluímos porque ya lo estudiamos.

Sabemos que existe una matriz B $m \times m$, con $N(B) = \{0\}$ tal que

$$BA = \begin{array}{c} n \\ \left[\begin{array}{c} T \\ \hline 0 \end{array}\right]\begin{array}{l} n \\ m-n \end{array} \end{array} \qquad \text{con } N(T) = \{0\}$$

T tiene n renglones linealmente independientes, ya que sus columnas son independientes.

$$\text{Pero} \quad R(A^t) = R((BA)^t) = R([T^t \mid 0]) = R(T^t)$$

$$\therefore \dim(R(A^t)) = \dim(R(T^t)) = n = \dim(R(A))$$

$$\therefore \dim(R(A^t)) = \dim(R(A)) = n$$

■

RESULTADO

Sea A una matriz $m \times n$ con renglones independientes y tal que $m < n$. Entonces A tiene m columnas linealmente independientes.

Demostración :

Existe una matriz B nxn , con $N(B) = \{0\}$ tal que

$$AB = \begin{array}{cc} & \begin{array}{cc} m & n-m \end{array} \\ m & \left[\begin{array}{c|c} T & 0 \end{array}\right] \end{array} \quad , \text{ con } \quad N(T) = \{0\}$$

Como estamos operando sobre las columnas de A entonces $R(A) = R(AB)$ y como $N(T) = \{0\}$ $\Rightarrow$ T tiene m columnas independientes.

como $R(A) = R(T) \Rightarrow$

$$\dim(R(A)) = \dim(R(T)) = m = \dim(R(A^t))$$

$$\therefore \quad \dim(R(A)) = \dim(R(A^t))$$

∎

Después de ver estos resultados, es natural esperar que cualquier matriz tiene el mismo número de columnas y renglones linealmente independientes.

TEOREMA DEL RANGO

Sea A una matriz $m \times n$. Entonces A tiene igual número de columnas y renglones independientes:

$$\dim(R(A)) = \dim(R(A^t)).$$

Demostración:

Supongamos $\dim (R(A)) = r < n \quad , \quad n \leq m.$

Si $\tilde{a}_1, \tilde{a}_2, \ldots, \tilde{a}_r$ son r columnas independientes de A, entonces existe una permutación P $n\times n$ tal que

$$AP = [\tilde{a}_1 | \tilde{a}_2 | \cdots | \tilde{a}_r | \tilde{a}_{r+1} | \cdots | \tilde{a}_n] = [\underset{r}{A_1} | \underset{n-r}{A_2}]$$

donde $R(A_2) \subseteq R(A_1)$ y $N(A_1) = \{0\}$, y por lo tanto existe una matriz B $n\times n$, con $N(B) = \{0\}$ tal que

$$APB = [\underset{r}{A_1} | \underset{n-r}{0}] = A(PB) = AC \quad , \text{ con } N(C) = \{0\}, \text{ donde } C = PB .$$

Por lo tanto $\dim (R (AC)^t) = \dim (R(A^t))$

Calcularemos $\dim (R(AC)^t)$:

$$(AC)^t = \begin{bmatrix} A_1^t \\ \hline 0 \end{bmatrix} \begin{matrix} r \\ n-r \end{matrix}$$

$$\therefore \quad R((AC)^t) = \left\{ y \in \mathbb{R}^n : \quad y = \begin{bmatrix} z \\ 0 \end{bmatrix} \begin{matrix} r \\ n-r \end{matrix} , \text{ donde } z \in R(A_1^t) \right\}$$

claramente $\dim (R (AC)^t) = \dim(R(A_1^t))$, y como demostramos anteriormente $\dim (R(A_1^t)) = \dim (R(A_1))$

$$\therefore \quad \dim (R(A_1)) = \dim (R(A)) = \dim (R(A^t)) \quad \square$$

Análogamente se puede ver que si la matriz A $m \times n$ tiene $r < m$ renglones independientes, entonces A tiene r columnas independientes. Notar que para que sea posible que A tenga r renglones independientes es necesario que $n \geq r$.

DEFINICION

Sea A una matriz $m \times n$. El RANGO de A, denotado por $r(A)$, se define como

$$r(A) = \dim(R(A)) = \dim(R(A^t)).$$

OBSERVACION

$$r(A) \leq \min\{m, n\}$$

Ya que el número de renglones independientes es $\leq m$, y el número de columnas independientes es $\leq n$.

Decimos que A $m \times n$ tiene RANGO MAXIMO cuando $r(A) = \min\{m, n\}$.

Ahora que sabemos que una matriz tiene igual número de columnas y renglones independientes, nos gustaría encontrar, si es posible, un tipo especial de matrices para las cuales fuera inmediato saber el rango, y cuáles son las columnas y renglones independientes.

Antes de seguir adelante veremos para qué nos sirve conocer el rango de una matriz para resolver el problema que nos habiamos planteado.

Recordemos que estamos interesados en contestar la siguiente pregunta:

Dada una matriz A $m \times n$:

¿ cuándo tiene solución el sistema

$$Ax = b$$

para toda $b \in \mathbb{R}^m$?

Sabemos que $Ax = b$ tiene solución para toda $b \in \mathbb{R}^m$ cuando $R(A) = \mathbb{R}^m$

Por lo tanto es necesario que $r(A) = m$

Para matrices triangulares T de orden n sabemos que si algun elemento de la diagonal vale cero, entonces $r(T) < n$. Esto nos induce a pensar que para matrices triangulares los renglones y las columnas linealmente independientes de T son aquellas cuyo elemento, que pertenece a la diagonal de T, es diferente de cero.

El siguiente ejemplo nos muestra que ésto no es válido en general

$$T = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

Es más, de este ejemplo podemos concluir que ni siquiera el rango de T está dado por el número de elementos diagonales diferentes de cero.

Trabajaremos con matrices divididas por bloques para tratar de resolver nuestro problema.

Empezaremos por generalizar un resultado que probamos en el capítulo anterior:

RESULTADO

Sea $$A = \begin{array}{c|cc|} & s & n-s \\ s & A_{11} & A_{12} \\ m-s & 0 & A_{22} \end{array}$$ una matriz $m \times n$

con A_{11} cuadrada.

III.14

$$N(A)=\{0\} \iff N(A_{11})=\{0\} \text{ y } N(A_{22})=\{0\}$$

Demostración:

Basta con demostrar $N(A)=\{0\} \Rightarrow N(A_{22})=\{0\}$

Supongamos $N(A_{11})=\{0\}$, ya que si $N(A_{11})\neq\{0\} \Rightarrow N(A)\neq\{0\}$.

Si $N(A_{22})\neq 0 \Rightarrow$ existe $x_2 \in \mathbb{R}^{n-s}$, $x_2 \neq 0$ tal que $A_{22}x_2 = 0$

$$\text{Sea } b = -\underset{s\times(n-s)}{A_{12}} \cdot \underset{(n-s)\times 1}{x_2} \in \mathbb{R}^s$$

Como A_{11} es de orden s y $N(A_{11})=\{0\}$ $\Rightarrow R(A_{11}) = \mathbb{R}^s$

$$\therefore \quad b \in R(A_{11})$$

$\therefore$ existe una $x_1 \in \mathbb{R}^s$ tal que

$$A_{11}x_1 = b = -A_{12}x_2$$

$$\text{Sea} \quad x = \begin{array}{c} \left[\begin{array}{c} x_1 \\ x_2 \end{array}\right] \end{array}\begin{array}{l} {\scriptstyle s} \\ {\scriptstyle n-s} \end{array} \in \mathbb{R}^n$$

$\therefore$ $x \neq 0$ ya que $x_2 \neq 0$ y además

$$Ax = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} A_{11}x_1 + A_{12}x_2 \\ A_{22}x_2 \end{bmatrix}$$

$$= \begin{bmatrix} -A_{12}x_2 + A_{12}x_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$\therefore \quad N(A) \neq \{0\}$

$\blacksquare$

PROPIEDADES DE $r(A)$ CON RESPECTO A BLOQUES

1) Sean $A_{m \times n}$ y B $m \times m$ tal que $N(B) = \{0\}$.
Entonces $r(A) = r(BA)$:

$$R(A^t) = R((BA)^t) \Rightarrow r(A) = r(BA)$$

$\blacksquare$

2) Sean A $m \times n$ y B $n \times n$ tal que $N(B) = \{0\}$
Entonces $r(A) = r(AB)$:

$$R(A) = R(AB) \Rightarrow r(A) = r(AB) \quad \blacksquare$$

3) Si $A = [A_1 \mid 0] \Rightarrow r(A) = r(A_1)$:

$$R(A) = R(A_1) \Rightarrow r(A) = r(A_1) \blacksquare$$

4) Si $A = [A_1 | A_2]$ y $R(A_2) \subseteq R(A_1) \Rightarrow r(A) = r(A_1)$

Demostración:

$$R(A) \subseteq R(A_1) \subseteq R(A) \Rightarrow R(A) = R(A_1)$$
$$\Rightarrow r(A) = r(A_1) \blacksquare$$

5) Si $A = [A_1 | A_2]$ y $R(A_1) \cap R(A_2) = \{0\} \Rightarrow$
$r(A) = r(A_1) + r(A_2)$.

Demostración:

Sea $y \in R(A)$ entonces tenemos que

$$y \in R(A) \Leftrightarrow y = Ax$$
$$\Leftrightarrow y = A_1x_1 + A_2x_2$$
$$\Leftrightarrow y = y_1 + y_2, \text{ con } y_1 \in R(A_1) \text{ y } y_2 \in R(A_2).$$

Es decir:

$$R(A) = \{y \in \mathbb{R}^m : y = y_1 + y_2, \; y_1 \in R(A_1), y_2 \in R(A_2)\}$$

Ahora, supongamos que $\dim(R(A_1)) = r_1$
$\dim(R(A_2)) = r_2$

Podemos permutar las columnas de A_1 de modo que

$$A_1 P_1 = [\overset{r_1}{A_1^{(1)}} \mid A_1^{(2)}] \qquad \text{con} \quad R(A_1^{(2)}) \subseteq R(A_1^{(1)})$$

y también podemos permutar las columnas de A_2 para obtener

$$A_2 P_2 = [\overset{r_2}{A_2^{(1)}} \mid A_2^{(2)}] \qquad \text{con} \quad R(A_2^{(2)}) \subseteq R(A_2^{(1)}).$$

$$\therefore \quad \begin{aligned} R(A_1 P_1) &= R(A_1^{(1)}) = R(A_1) \\ R(A_2 P_2) &= R(A_2^{(1)}) = R(A_2) \end{aligned}$$

Nota: $N(A_1^{(1)}) = \{0\}$ y $N(A_2^{(1)}) = \{0\}$.

$\therefore$ veremos que

$$N([A_1^{(1)} \mid A_2^{(1)}]) = \{0\} \quad :$$

Supongamos

$$[A_1^{(1)} \mid A_2^{(1)}]\, X = 0$$

$$\Rightarrow \quad A_1^{(1)} X_1 + A_2^{(1)} X_2 = 0$$

$$\Rightarrow \quad A_1^{(1)} X_1 = - A_2^{(1)} X_2$$

$$\Rightarrow \quad y_1 = y_2, \qquad \text{con} \quad \begin{aligned} & y_1 \in R(A_1^{(1)}) = R(A_1) \\ \text{y} \quad & y_2 \in R(A_2^{(1)}) = R(A_2) \end{aligned}$$

como $R(A_1) \cap R(A_2) = \{0\} \Rightarrow$ el único vector que puede estar en ambos subespacios es el vector cero

$$\therefore \quad y_1 = y_2 = 0$$

$$\Rightarrow \quad \begin{array}{l} A_1^{(1)} x_1 = 0 \\ A_2^{(1)} x_2 = 0 \end{array}$$

pero como $N(A_1^{(1)}) = \{0\}$ y $N(A_2^{(1)}) = \{0\}$

$\Rightarrow x_1 = 0, \quad x_2 = 0 \quad \Rightarrow x = 0$

$$\Rightarrow \quad N([A_1^{(1)} \mid A_2^{(1)}]) = \{0\}$$

$\therefore$ A tiene $r_1 + r_2$ columnas independientes

$$\therefore \quad r(A) = r(A_1) + r(A_2) \blacksquare$$

6) Si $A = \left[\begin{array}{c} A_1 \\ \hline 0 \end{array}\right] \Rightarrow r(A) = r(A_1)$

7) Si $A = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]$ y $R(A_2^*) \subseteq R(A_1^*)$

$$\Rightarrow \quad r(A) = r(A_1)$$

8) Si $A = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]$ y $R(A_1^*) \cap R(A_2^*) = \{0^*\}$

entonces $r(A) = r(A_1) + r(A_2)$

9) Si $A = \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] \Rightarrow r(A) = r(A_{11}) + r(A_{22})$

Demostración:

Sean $A_1 = \left[\begin{array}{c} A_{11} \\ \hline 0 \end{array}\right]$, $A_2 = \left[\begin{array}{c} 0 \\ \hline A_{22} \end{array}\right]$ $\therefore A = [A_1 | A_2]$

y $R(A_1) \cap R(A_2) = \{0\}$.

$\therefore r(A) = r(A_1) + r(A_2) = r(A_{11}) + r(A_{22})$ ■

10) Si $A = \begin{array}{c} \\ s \\ m-s \end{array}\!\!\begin{array}{c} \begin{array}{cc} s & n-s \end{array} \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \end{array}$ con A_{11} de orden s y tal que $N(A_{11}) = \{0\}$

entonces $r(A) = r(A_{11}) + r(A_{22})$

Demostración:

Si A_{11} es de orden s con columnas independientes $\Rightarrow R(A_{11}) = \mathbb{R}^s \Rightarrow R(A_{12}) \subseteq R(A_{11})$

$\therefore$ existe una $\bar{B}_{(n-s)\times(n-s)}$ tal que
$A_{12} = A_{11}\bar{B}$

Sea $B = \left[\begin{array}{c|c} I_s & -\bar{B} \\ \hline 0 & I_{n-s} \end{array}\right]$ $\therefore N(B) = \{0\}$

y además

$$AB = \begin{array}{c} \\ s \\ m-s \end{array}\!\!\begin{array}{cc} s & n-s \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \end{array} \begin{array}{c} n-s \\ \left[\begin{array}{c|c} I_s & -\bar{B} \\ \hline 0 & I_{n-s} \end{array}\right] \end{array} = \left[\begin{array}{c|c} A_{11} & -A_{11}\bar{B} + A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

$$= \left[\begin{array}{c|c} A_{11} & -A_{12} + A_{12} \\ \hline 0 & A_{22} \end{array}\right] = \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right].$$

$$\therefore \quad r(A) = r(A_{11}) + r(A_{22}) \blacksquare$$

Observar que la condición importante para deducir este resultado es que $R(A_{12}) \subseteq R(A_{11})$, y no que A sea cuadrada.

RESULTADO

Si $\quad A = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \quad$ y $\quad R(A_{12}) \subseteq R(A_{11})$

entonces $\quad r(A) = r(A_{11}) + r(A_{22})$

Demostración :

Si $R(A_{12}) \subseteq R(A_{11}) \Rightarrow$ existe una matriz $B_{n \times n}$, con $N(B) = \{0\}$ tal que

$$AB = \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right]$$

■

Este resultado es el que nos da la clave para transformar a la matriz A en una matriz para la cual sea fácil "leer" cuales son los renglones y las columnas linealmente independientes.

Supongamos que tenemos un método para transformar a una matriz en

$$\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

tal que para A_{11} ya tenemos resuelto el problema de saber fácilmente cuáles son sus columnas y renglones independientes, y tal que $R(A_{12}) \subseteq R(A_{11})$.

Esto nos permite olvidarnos de los bloques A_{11} y A_{12}, y dedicarnos a investigar a la matriz A_{22}, solo que con la ventaja de que A_{22} tiene dimensión menor que la matriz original, es decir, tenemos el mismo problema a resolver, pero en una dimensión menor, cuando menos.

Nosotros conocemos una técnica para reducir una matriz a la forma

$$\begin{bmatrix} U & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

donde U es triangular superior.

Si $N(U) = \{0\}$ entonces podemos asegurar que $R(A_{12}) \subseteq R(U)$ y que por lo tanto $r(A) = r(U) + r(A_{22})$.

Qué podemos decir en el caso en que

$$A = \left[\begin{array}{ccccc|c} \circledast & * & \cdots & * & * & \\ & \circledast & \cdots & * & * & \\ & & \ddots & \vdots & \vdots & A_{12} \\ & & & \circledast & * & \\ & & & & 0 & \\ \hline & & 0 & & & A_{22} \end{array}\right], \quad \text{donde } \circledast \neq 0$$

En este caso basta con que el último renglón de A_{12} tenga un elemento diferente de cero para que $R(A_{12}) \not\subseteq R(U)$.

si exigimos que A_{11} sea una matriz cuadrada, entonces ya no podemos asegurar que $r(A) = r(A_{11}) + r(A_{22})$, y para investigar cuál es $r(A)$ tendremos que considerar también a A_{12}.

Intentaremos solucionar el problema para un caso particular:

Sea $$A = \begin{bmatrix} 1 & -1 & 2 & 3 & 5 & 2 \\ 3 & -3 & 0 & 9 & 5 & 0 \\ 4 & -4 & 5 & 12 & 15 & 0 \\ 5 & -5 & 7 & 15 & 20 & 10 \\ 2 & -2 & 1 & 6 & 10 & 2 \end{bmatrix}$$

En un primer paso

$$A \rightarrow A_1 = \left[\begin{array}{c:ccccc} 1 & -1 & 2 & 3 & 5 & 2 \\ \hdashline 0 & 0 & -6 & 0 & -10 & -6 \\ 0 & 0 & -3 & 0 & -5 & -8 \\ 0 & 0 & -3 & 0 & -5 & 0 \\ 0 & 0 & -3 & 0 & 0 & -2 \end{array}\right]$$

En el segundo paso tendríamos

$$A_1 \rightarrow A_2 = \left[\begin{array}{cc:cccc} 1 & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & -6 & 0 & -10 & -6 \\ \hdashline 0 & 0 & -3 & 0 & -5 & -8 \\ 0 & 0 & -3 & 0 & -5 & 0 \\ 0 & 0 & -3 & 0 & 0 & -2 \end{array}\right]$$

Esta es precisamente la situación que nos interesa evitar, por lo cual nos conviene modificar la partición de A_2:

$$A_2 = \left[\begin{array}{cc|cccc} 1 & -1 & 2 & 3 & 5 & 2 \\ \hline 0 & 0 & -6 & 0 & -10 & -6 \\ 0 & 0 & -3 & 0 & -5 & -8 \\ 0 & 0 & -3 & 0 & -5 & 0 \\ 0 & 0 & -3 & 0 & 0 & -2 \end{array}\right] = \left[\begin{array}{c|c} A_{11}^{(2)} & A_{12}^{(2)} \\ \hline 0 & A_{22}^{(2)} \end{array}\right]$$

ya que entonces $R(A_{12}^{(2)}) \subseteq R(A_{11}^{(2)})$.

En el siguiente paso nos tocaría aplicar el procedimiento a $A_{22}^{(2)}$ y aumentar a $A_{11}^{(2)}$ un renglón y una columna para obtener:

$$A_3 = \left[\begin{array}{ccc|ccc} 1 & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & -6 & 0 & -10 & -6 \\ \hline 0 & 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 5 & 1 \end{array}\right] = \left[\begin{array}{c|c} A_{11}^{(3)} & A_{12}^{(3)} \\ \hline 0 & A_{22}^{(3)} \end{array}\right]$$

donde $R(A_{12}^{(3)}) \subseteq R(A_{11}^{(3)})$

Ahora, al siguiente paso tendríamos

$$A_4 = \left[\begin{array}{cccc|cc} 1 & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & -6 & 0 & -10 & -6 \\ 0 & 0 & 0 & 0 & 0 & -5 \\ \hline 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 5 & 1 \end{array}\right] = \left[\begin{array}{c|c} A_{11}^{(4)} & A_{12}^{(4)} \\ \hline 0 & A_{22}^{(4)} \end{array}\right]$$

Estamos otra vez en la situación de que $R(A_{12}^{(4)}) \nsubseteq R(A_{11}^{(4)})$, circunstancia que está determinada por el hecho de que la primera columna de $A_{22}^{(3)}$ sea el vector cero. Por lo tanto, lo más indicado es que si la primera columna de $A_{22}^{(j)}$ vale cero, entonces nuestro procedimiento aplicado a A_j debe consistir en aumentar una columna a $A_{11}^{(j)}$, y el bloque resultante será el primer bloque de A_{j+1}. En este caso tendríamos:

$$A_3 \to A_4 = \left[\begin{array}{cccc|cc} 1 & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & -6 & 0 & -10 & -6 \\ \hline 0 & 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 5 & 1 \end{array}\right] = \left[\begin{array}{c|c} A_{11}^{(4)} & A_{12}^{(4)} \\ \hline 0 & A_{22}^{(4)} \end{array}\right]$$

Con esta partición recuperamos la propiedad de que $R(A_{12}^{(4)}) \subseteq R(A_{11}^{(4)})$.

Al aplicar el procedimiento a $A_{22}^{(4)}$ obtenemos

$$A_4 \to A_5 = \left[\begin{array}{ccccc|c} 1 & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & -6 & 0 & -10 & -6 \\ 0 & 0 & 0 & 0 & 5 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & -5 \end{array}\right]$$

Y finalmente:

$$A_5 \rightarrow A_6 = \begin{bmatrix} \textcircled{1} & -1 & 2 & 3 & 5 & 2 \\ 0 & 0 & \textcircled{-6} & 0 & -10 & -6 \\ 0 & 0 & 0 & 0 & \textcircled{5} & 1 \\ 0 & 0 & 0 & 0 & 0 & \textcircled{3} \\ \hdashline 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Notar que en este proceso siempre avanzamos por columnas.

Hemos llevado a A a una forma para la cual es muy fácil determinar las columnas y los renglones linealmente independientes:

Los cuatro primeros renglones son independientes; los renglones de A que eran linealmente dependientes fueron transformados en el vector renglón 0^t - en este caso la matriz original únicamente tenía un renglón linealmente dependiente de los restantes.

Cada renglón independiente (no nulo) de A_6 contiene un "pivote"; las columnas linealmente independientes de A_6 son aquellas que contienen un pivote.

Al transformar a A en A_6 nos cuidamos en cada paso de no perder la propiedad $R(A_{12}) \subseteq R(A_{11})$ donde

$$A_k = \begin{array}{cc} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \end{array}$$

. Veremos porqué este procedimiento funciona, .

permitiendo que A_{11} no sea cuadrada.

Supongamos que A es de la forma

$$A = \begin{array}{cc} \begin{array}{cc} \quad s \quad & \quad n-s \quad \end{array} & \\ \left[\begin{array}{c|c} A_{11} & \begin{matrix} * & * & \cdots & * \\ * & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ * & * & \cdots & * \end{matrix} \\ \hline \begin{matrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{matrix} & \begin{matrix} \circledast & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{matrix} \end{array}\right] & \begin{array}{c} r \\ \\ m-r \end{array} \end{array} = \begin{array}{c} \begin{array}{ccc} \quad s \quad & 1 & n-(s+1) \end{array} \\ \left[\begin{array}{c|c|c} A_{11} & b & \begin{matrix} * & \cdots & * \\ * & \cdots & * \\ \vdots & \bar{A}_{12} & \vdots \\ * & \cdots & * \end{matrix} \\ \hline \begin{matrix} 0 & \cdots & 0 \end{matrix} & \sigma & c^t \\ \hdashline 0 & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} & \begin{matrix} * & \cdots & * \\ \vdots & & \vdots \\ * & \cdots & * \end{matrix} \end{array}\right] \end{array} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

con $R(A_{12}) \subseteq R(A_{11})$ y donde la primera columna de A_{12} es $b \in R(A_{11})$ y el primer renglón de A_{22} es $[\sigma \mid c^t]$, con $\sigma \in \mathbb{R}$, $\sigma \neq 0$ y $c \in \mathbb{R}^{n-(s+1)}$.

Veremos si podemos aumentar el bloque A_{11} un renglón y una columna para reducir nuestro trabajo a un bloque más pequeño.

Es decir, queremos saber si la siguiente partición de A:

$$A = \begin{array}{cc} \begin{array}{cc} s+1 & n-(s+1) \end{array} & \\ \left[\begin{array}{c|c} \tilde{A}_{11} & \tilde{A}_{12} \\ \hline 0 & \tilde{A}_{22} \end{array}\right] & \begin{array}{c} r+1 \\ m-(r+1) \end{array} \end{array}$$

donde

$$\tilde{A}_{11} = \begin{array}{c c c} s & 1 & \\ \left[\begin{array}{c|c} A_{11} & b \\ \hline 0 & \sigma \end{array}\right. & \left.\right] & \begin{array}{c} r \\ 1 \end{array} \end{array}$$

$$\tilde{A}_{12} = \begin{array}{c c} n-(s+1) & \\ \left[\begin{array}{c} \bar{A}_{12} \\ \hline c^t \end{array}\right] & \begin{array}{c} r \\ 1 \end{array} \end{array}$$

tiene la propiedad de que $R(\tilde{A}_{12}) \subseteq R(\tilde{A}_{11})$.

Para demostrar que esta propiedad se cumple tenemos que ver que para toda $y \in R(\tilde{A}_{12})$ existe una $x \in \mathbb{R}^{s+1}$ tal que $\tilde{A}_{11} x = y$:

Supongamos que $y \in R(\tilde{A}_{12})$, por lo tanto queremos encontrar una $x \in \mathbb{R}^{s+1}$ que satisfaga el sistema

$$\begin{array}{c c} & \begin{array}{cc} s & 1 \end{array} \\ \begin{array}{c} r \\ 1 \end{array} & \left[\begin{array}{c|c} A_{11} & b \\ \hline 0 & \sigma \end{array}\right] \end{array} x = y$$

Como $y \in R(\tilde{A}_{12}) \Rightarrow$ $y = \left[\begin{array}{c} \bar{A}_{12} \\ \hline c^t \end{array}\right] z = \left[\begin{array}{c} \bar{A}_{12} z \\ \hline c^t z \end{array}\right]$.

para alguna $z \in \mathbb{R}^{n-s}$

Como $R(\bar{A}_{12}) \subseteq R(A_{11})$ tenemos que $\underset{r\times(n-s)}{\bar{A}_{12}} = \underset{r\times s}{A_{11}} \underset{s\times(n-s)}{B}$ para alguna matriz B $s\times(n-s)$

$\therefore$ $\bar{A}_{12} z = A_{11} B z = A_{11} z_1$, donde $z_1 = Bz \in \mathbb{R}^s$.

$$\therefore \quad y = \begin{bmatrix} A_{11} z_1 \\ c^t z \end{bmatrix}$$

$$\therefore \quad y \in R(\tilde{A}_{11}) \iff \begin{array}{cc} \begin{matrix} s & \quad 1 \end{matrix} & \\ \left[\begin{array}{c|c} A_{11} & b \\ \hline 0 & \sigma \end{array}\right] & \begin{matrix} r \\ 1 \end{matrix} \end{array} \cdot x = \left[\begin{array}{c} A_{11} z_1 \\ \hline c^t z \end{array}\right]$$

para alguna $x \in \mathbb{R}^{s+1}$

$$\iff \begin{array}{cc} & \begin{matrix} s & \quad 1 \end{matrix} \\ \begin{matrix} r \\ 1 \end{matrix} & \left[\begin{array}{c|c} A_{11} & b \\ \hline 0 & \sigma \end{array}\right] \end{array} \begin{array}{cc} \left[\begin{array}{c} x_1 \\ \hline \lambda \end{array}\right] & \begin{matrix} s \\ 1 \end{matrix} \end{array} = \begin{array}{cc} \left[\begin{array}{c} A_{11} z_1 \\ \hline c^t z \end{array}\right] & \begin{matrix} r \\ 1 \end{matrix} \end{array}$$

$$\iff \begin{aligned} A_{11} x_1 + \lambda b &= A_{11} z_1 \\ \sigma \lambda &= c^t z \end{aligned}$$

Como $\sigma \neq 0$ entonces podemos despejar λ: $\lambda = \frac{1}{\sigma} \cdot c^t z$.

Además, si $\lambda = \frac{1}{\sigma} \cdot c^t z$ entonces

$A_{11} z_1 - \lambda b \in \mathbb{R}^r$, ya que $b \in \mathbb{R}^r$ y $R(A_{11}) \subseteq \mathbb{R}^r$

$\therefore$ existe una $x_1^* \in \mathbb{R}^s$ tal que $A_{11} x_1^* = A_{11} z_1 - \lambda b$

$$\therefore \quad \tilde{A}_{11} \begin{array}{c} \left[\begin{array}{c} x_1^* \\ \hdashline \frac{c^t z}{\sigma} \end{array}\right] \end{array}\begin{array}{l} s \\ 1 \end{array} = \begin{array}{c c} & \begin{array}{cc} s & 1 \end{array} \\ \begin{array}{c} r \\ 1 \end{array} & \left[\begin{array}{c:c} A_{11} & b \\ \hdashline 0 & \sigma \end{array}\right] \end{array} \left[\begin{array}{c} x_1^* \\ \hdashline \frac{c^t z}{\sigma} \end{array}\right]\begin{array}{l} s \\ 1 \end{array} = \left[\begin{array}{c} A_{11} x_1^* + b \frac{c^t z}{\sigma} \\ c^t z \end{array}\right]\begin{array}{l} s \\ 1 \end{array}$$

$$\left[\begin{array}{c} A_{11} z_1 - \left(\frac{c^t z}{\sigma}\right) b + b \frac{c^t z}{\sigma} \\ c^t z \end{array}\right] = \left[\begin{array}{c} A_{11} z_1 \\ c^t z \end{array}\right] = y$$

$\therefore$ $y \in R(\tilde{A}_{11})$, y además $r(\tilde{A}_{11}) = r(A_{11}) + 1$ = número de renglones de $\tilde{A}_{11}$ ∎

Con esta discusión podemos ahora describir el procedimiento completamente.

Dada A $m \times n$ encontrar B $m \times m$, con $N(B) = \{0\}$ tal que

$$BA = \begin{array}{c c} \begin{array}{cc} \quad k \quad & \quad n-k \quad \end{array} & \\ \left[\begin{array}{cccc|cccc} 0 & 0 & \cdots & \circledast & * & * & \cdots & * \\ \hline & & & & * & * & \cdots & * \\ & & & & \vdots & \vdots & & \vdots \\ & & & & * & * & \cdots & * \end{array}\right] & \begin{array}{l} 1 \\ \\ m-1 \\ \end{array} \end{array} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

donde $R(A_{12}) \subseteq R(A_{11})$.

Sea $A = [a_1 | a_2 | \cdots | a_n]$, y sea a_k la primera columna de A diferente de cero :

$$A = \begin{bmatrix} 0 & 0 & \cdots & 0 & \overset{\overset{k}{\downarrow}}{*} & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & * & \cdots & * \\ 0 & 0 & \cdots & 0 & \circledast & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & * & * & \cdots & * \end{bmatrix}$$

Entonces podemos permutar el primer renglón de A con otro renglón para obtener

$$PA = \begin{bmatrix} 0 & 0 & \cdots & 0 & \circledast & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & * & \cdots & * \\ 0 & 0 & \cdots & 0 & * & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & * & * & \cdots & * \end{bmatrix} , \quad \text{con } \circledast \neq 0$$

Usando al elemento $\circledast$ como pivote, podemos multiplicar por una matriz triangular inferior L, con $N(L) = \{0\}$ tal que

$$A_K = LPA = \begin{array}{cc} \left[\begin{array}{ccccc|ccc} 0 & 0 & \cdots & 0 & \circledast & * & \cdots & * \\ \hline 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & * & \cdots & * \end{array}\right] & \begin{array}{l} 1 \\ \\ m-1 \\ \\ \\ \end{array} \\ \begin{array}{cc} \qquad k \qquad & \qquad n-k \end{array} & \end{array} = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right]$$

Es importante observar que utilizamos el subíndice K para indicar que el primer bloque de A_K tiene K columnas.

Como $\circledast \neq 0$ entonces $R(A_{12}) \subseteq R(A_{11})$.

Esto reduce nuestro problema a investigar la dependencia lineal para los renglones y columnas de la matriz A_{22}, que es de dimensión menor que A:

supongamos que hemos transformado a la matriz A en

$$A_k = \begin{array}{cc} \begin{array}{cc} \quad k & \qquad n-k \end{array} & \\ \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] & \begin{array}{l} \ell \\ m-\ell \end{array} \end{array}$$

donde $R(A_{12}) \subseteq R(A_{11})$ y $r(A_{11}) = \ell$

Entonces supongamos $A_{22} = [\tilde{a}_1 | \tilde{a}_2 | \cdots | \tilde{a}_{n-k}]$:

1) si $\tilde{a}_1 = 0$ entonces modificamos la partición de A_k:

$$A_{k+1} = A_k = \begin{array}{cc} \begin{array}{cc} k+1 & n-(k+1) \end{array} & \\ \left[\begin{array}{c|c} \bar{A}_{11} & \bar{A}_{12} \\ \hline 0 & \bar{A}_{22} \end{array}\right] & \begin{array}{l} \ell \\ m-\ell \end{array} \end{array}$$

y notamos que se sigue cumpliendo

$r(\bar{A}_{11}) = \ell$ y $R(\bar{A}_{12}) \subseteq R(\bar{A}_{11})$.

2) si $\tilde{a}_1 \neq 0$ entonces podemos obtener una

matriz $\bar{B}$ de orden $(m-\ell)$, con $N(B)=\{0\}$ tal que

$$\underset{(m-\ell)\times(n-k)}{\bar{B}A_{22}} = \begin{array}{c} \begin{array}{cc} 1 & \quad n-(k+1) \end{array} \\ \left[\begin{array}{c|cccc} \circledast & * & * & \cdots & * \\ \hline 0 & * & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \cdots & * \end{array}\right] \end{array} = \begin{array}{c} 1 \\ m-(\ell+1) \end{array} \left[\begin{array}{c|c} \circledast & * \; * \cdots * \\ \hline 0 & \bar{A}_{22} \end{array}\right]$$

sea $B = \left[\begin{array}{c|c} I_\ell & 0 \\ \hline 0 & \bar{B} \end{array}\right]$ y sea

$$A_{K+1} = B\,A_K = \begin{array}{c|c|c|c} & k & n-k & \\ \ell & A_{11} & A_{12} & \ell \\ \hline m-\ell & 0 & \left[\begin{array}{c|c} \circledast & * \cdots * \\ \hline 0 & \bar{A}_{22} \end{array}\right] & \begin{array}{c} 1 \\ m-(\ell+1) \end{array} \\ & k & 1 \quad n-(k+1) & \end{array}$$

Modificamos la partición de A_{K+1}, aumentando un renglón y una columna a A_{11}:

si $\bar{A}_{11} = \begin{array}{c} \\ \ell \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} k & \quad 1 \end{array} \\ \left[\begin{array}{c|c} A_{11} & \begin{array}{c} * \\ \vdots \\ * \end{array} \\ \hline 0\;0 \cdots 0 & \circledast \end{array}\right] \end{array}$, entonces

$\bar{A}_{22}$ es la submatriz determinada arriba, y tenemos que

$$A_{k+1} = \begin{array}{cc} \begin{array}{cc} k+1 & n-(k+1) \end{array} & \\ \left[\begin{array}{c|c} \bar{A}_{11} & \bar{A}_{12} \\ \hline 0 & \bar{A}_{22} \end{array}\right] & \begin{array}{c} \ell+1 \\ m-(\ell+1) \end{array} \end{array}$$

con $r(\bar{A}_{11}) = \ell+1$; $R(\bar{A}_{12}) \subseteq R(\bar{A}_{11})$

Con este procedimiento siempre avanzamos por columnas, y terminamos cuando obtenemos A_n.

Veremos qué forma puede tener la matriz A_n.

Si $r(A_n) = m = n$ entonces

$$BA = \begin{bmatrix} \circledast & * & \cdots & * \\ & \circledast & \cdots & * \\ & & \ddots & \vdots \\ & & & \circledast \end{bmatrix}$$

Si $r(A_n) = n \leq m$ entonces

$$BA = \left[\begin{array}{cccc} \circledast & * & \cdots & * \\ & \circledast & & * \\ & & \ddots & \vdots \\ & & & \circledast \\ \hline & & 0 & \end{array}\right] \begin{array}{c} n \\ \\ m-n \end{array}$$

Si $r(A_n) = m \leq n$ entonces

$$BA = \left[\begin{array}{ccccccccccccccccc}
 & & & k_1 & & & k_2 & & & k_3 & & & & k_m & & & \\
0 & \cdots & 0 & \circledast & * & \cdots & * & * & \cdots & * & * & \cdots & \cdots & * & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & \circledast & * & \cdots & * & * & \cdots & \cdots & * & * & \cdots & * \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & \circledast & * & \cdots & \cdots & * & * & \cdots & * \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & & \vdots & \vdots & & \vdots \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & \circledast & * & \cdots & *
\end{array}\right]$$

y finalmente, si $r(A_n) = r < \min(m,n)$ entonces

$$BA = \left[\begin{array}{ccccccccccccccccc|c}
 & & & k_1 & & & k_2 & & & k_3 & & & & k_r & & & & \\
0 & \cdots & 0 & \circledast & * & \cdots & * & * & \cdots & * & * & \cdots & \cdots & * & * & \cdots & * & \\
0 & \cdots & 0 & 0 & 0 & \cdots & \circledast & * & \cdots & * & * & \cdots & \cdots & * & * & \cdots & * & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & \circledast & * & \cdots & \cdots & * & * & \cdots & * & r \\
\vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots & & & \vdots & \vdots & & \vdots & \\
0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & \cdots & \circledast & * & \cdots & * & \\
\hline
\multicolumn{17}{c|}{0} & m-r
\end{array}\right]$$

Veamos qué características tiene la matriz A_n para el caso más general:

A_n tiene r pivotes, y cada pivote está a la derecha del pivote del renglón superior; una columna puede tener a lo más un pivote.

A_n tiene r renglones no nulos, que están antes que los renglones de ceros.

Cada renglón no nulo contiene un pivote, y abajo de cada pivote hay una columna de ceros obtenida por eliminación.

Este tipo de matrices nos serán muy útiles, y les daremos un nombre:

DEFINICION

Sea $G = [\gamma_{ij}]$ una matriz $m \times n$. Decimos que G es una matriz ESCALON (ó ESCALONADA) reducida por filas si existen $r > 0$, con $r \leq \min\{m, n\}$, y r enteros positivos $k_1, k_2, \ldots, k_r$, $1 \leq k_i \leq n$, tal que

1) $k_1 < k_2 < \cdots < k_r$.

Cada columna puede tener a lo más un pivote.

2) $\gamma_{ij} = 0$ si $i > r$

Los últimos m-r renglones son cero

3) $\gamma_{ij} = 0$ si $j < k_i$

En cada renglón no nulo el pivote está precedido de ceros.

En el caso más general vamos a tener que una matriz escalonada es de la forma

$$\begin{array}{c|l} \left[\begin{array}{c} G \\ \hline O \end{array}\right] & \begin{array}{l} r \\ m-r \end{array} \end{array}$$

donde G es una matriz escalonada de rango máximo, es decir $r(G) = r$.

PROPIEDADES DE LAS MATRICES ESCALONADAS

1) Los renglones no nulos de una matriz escalonada son independientes.

Demostración :

Sea $G = [\gamma_{ij}]$ una matriz escalonada $m \times n$.

Sean $z_1^t, z_2^t, \ldots, z_r^t$ los renglones no nulos de G.

Veremos que $\{z_1^t, z_2^t, \ldots, z_1^t\}$ es linealmente independiente :

Tenemos que :

$$\begin{aligned}
z_1^t &= \begin{bmatrix} 0 & \cdots & 0 & \gamma_{1k_1} & * & \cdots & * & * & \cdots & * & * & \cdots & * \end{bmatrix}\\
z_2^t &= \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & \gamma_{2k_2} & * & \cdots & * & * & \cdots & * \end{bmatrix}\\
&\ \vdots \\
z_r^t &= \begin{bmatrix} 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \gamma_{rk_r} & \cdots & * \end{bmatrix}
\end{aligned}$$

donde $\gamma_{ik_i} \neq 0 \quad 1 \leq i \leq r$ y $k_1 < k_2 < \cdots < k_r$

Es fácil ver que ninguno de estos renglones se puede obtener como combinación lineal de los restantes ∎

2) El rango de una matriz escalonada es igual al número de renglones no nulos de la misma.

3) Por medio de una permutación P de orden n podemos reducir a G a la forma

$$GP = \begin{array}{c|c} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \left[\begin{array}{c|c} U & A_{12} \\ \hline 0 & 0 \end{array}\right] \end{array} \qquad \text{con} \quad r(U) = r$$

Demostración:

Las r columnas con pivote son de la forma ;

$$g_{k_1} = \begin{bmatrix} \circledast \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad g_{k_2} = \begin{bmatrix} * \\ \circledast \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad g_{k_3} = \begin{bmatrix} * \\ * \\ \circledast \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad g_{k_r} = \begin{bmatrix} * \\ * \\ * \\ \vdots \\ \circledast \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \\ \leftarrow r \\ \\ \\ \\ \end{array}$$

Llevaremos a G a la forma deseada por pasos :

En el primer paso intercambiamos la primera columna de $G = G_0$ con la k_1-ésima :

$$GP_1 = \begin{array}{r} \\ \\ \\ \\ \\ \\ m-1 \\ \\ \\ \end{array} \left[\begin{array}{c|cccccccccccc} & & & k_1 & & k_2 & & k_3 & & k_r & & & \\ \circledast & 0 & \cdots & 0 & \cdots & * & \cdots & * & \cdots & * & * & \cdots & * \\ \hline 0 & 0 & \cdots & 0 & \cdots & \circledast & \cdots & * & \cdots & * & * & \cdots & * \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & \circledast & \cdots & * & * & \cdots & * \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & \circledast & * & \cdots & * \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & 0 & & 0 \end{array}\right] \begin{array}{l} \\ \\ \\ \\ \\ \leftarrow r \\ \\ \\ \\ \end{array} = G_1$$

$$n-1$$

$$= \left[\begin{array}{c|c} \overset{1}{*} & \overset{n-1}{* \cdots *} \\ \hline 0 & \boxed{} \end{array}\right] = \begin{array}{r} 1 \\ m-1 \end{array} \left[\begin{array}{c|c} \overset{1}{U_1} & \overset{n-1}{\boxed{}} \\ \hline 0 & \boxed{} \end{array}\right] ,$$

donde $r(U_1) = 1$

En el segundo paso intercambiamos la 2ª columna de G_1 con la k_2-ésima

$$P_2G_1 = \begin{array}{c|cc|cccccccccccc|c}
 & & & & & k_1 & & k_2 & & k_3 & & k_r & & & & \\
\hline
2 & \circledast & * & 0 & \cdots & 0 & \cdots & * & \cdots & * & \cdots\ \cdots & * & * & \cdots & * & \\
 & & \circledast & 0 & \cdots & 0 & \cdots & 0 & \cdots & * & \cdots\ \cdots & * & * & \cdots & * & \\
\hline
 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & \circledast & \cdots\ \cdots & * & * & \cdots & * & \\
 & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \\
 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots\ \cdots & \circledast & * & \cdots & * & \leftarrow r \\
 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots\ \cdots & 0 & 0 & \cdots & 0 & \\
m-2 & \vdots & \vdots & \vdots & & \vdots & & \vdots & & \vdots & & \vdots & \vdots & & \vdots & \\
 & 0 & 0 & 0 & \cdots & 0 & \cdots & 0 & \cdots & 0 & \cdots\ \cdots & 0 & 0 & \cdots & 0 & \\
\hline
 & \multicolumn{2}{c|}{2} & \multicolumn{12}{c|}{n-2} & \\
\end{array}$$

$$= \begin{array}{cc|c|c}
\circledast & * & \boxed{} & 2 \\
 & \circledast & & \\
\hline
\multicolumn{2}{c|}{0} & \boxed{} & m-2 \\
\multicolumn{2}{c|}{2} & n-2 &
\end{array}
= \begin{array}{c|c}
U_2 & \boxed{} \\
\hline
0 & \boxed{} \\
2 & n-2
\end{array}
= G_2 \quad ; \quad r(U_2) = 2$$

En el i-ésimo paso, si P_i es la permutación de orden n que intercambia la i-ésima y la k_i-ésima columnas de una matriz $m \times n$ entonces

$$G_{i-1}P_i = \begin{array}{c|cccc|c|}
 & & i & & & n-i \\
\hline
 & \circledast & * & \cdots & * & \\
i & & \circledast & \cdots & * & \boxed{} \\
 & & & \ddots & \vdots & \\
 & & & & \circledast & \\
\hline
m-i & & & 0 & & \boxed{} \\
\end{array} =$$

$$= \begin{array}{cc} & \begin{array}{cc} \quad i \quad & \quad n-i \quad \end{array} \\ \begin{array}{c} i \\ m-i \end{array} & \left[\begin{array}{c|c} U_i & \boxed{} \\ \hline 0 & \boxed{} \end{array}\right] \end{array} = G_i \; ; \; r(U_i)=i$$

En el r-ésimo paso obtenemos

$$GP = \begin{array}{cc} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \left[\begin{array}{c|c} U & A_{12} \\ \hline 0 & 0 \end{array}\right] \end{array} \qquad \text{donde} \quad r(U)=r$$

y donde $P = P_1 P_2 \cdots P_r$ y $U = U_r$.

Observar que de paso hemos demostrado que si $g_{k_1}, g_{k_2}, \ldots, g_{k_\ell}$ son las primeras ℓ columnas de G que tienen pivote entonces, si G_ℓ es la matriz

$$G_\ell = [g_{k_1} | g_{k_2} | \cdots | g_{k_\ell}]$$

G_ℓ es de la forma

$$G_\ell = \begin{array}{cc} & \ell \\ \begin{array}{c} \ell \\ m-\ell \end{array} & \left[\begin{array}{c} U_\ell \\ \hline 0 \end{array}\right] \end{array} \qquad \text{donde} \quad r(U_\ell) = \ell$$

4) Las columnas de G que tienen pivote son linealmente independientes.

Demostración:

Sea $G_r = [g_{k_1} | g_{k_2} | \cdots | g_{k_r}]$ la matriz cuyas columnas son las r columnas con pivote de G.

$$\therefore \quad G_r = \begin{array}{c} r \\ \left[\begin{array}{c} U_r \\ \hline 0 \end{array}\right] \end{array}\begin{array}{l} r \\ m-r \end{array} \qquad \text{donde} \qquad r(U_r) = r$$

$\therefore \; r(G_r) = r(U_r)$

$\therefore$ las columnas de G_r son independientes ■

5) Sean $g_{k_1}, g_{k_2}, \ldots, g_{k_r}$ las columnas de G que tienen pivote.
Entonces
$\{g_{k_1}, g_{k_2}, \ldots, g_{k_r}\}$ es una base de $R(G)$.

Demostración:

Como $r(G) = r$ y $\{g_{k_1}, g_{k_2}, \ldots, g_{k_r}\}$ es linealmente independiente

$\Rightarrow \{g_{k_1}, g_{k_2}, \ldots, g_{k_r}\}$ es una base de $R(G)$ ■

6) Sean $g_{k_1}, g_{k_2}, \dots, g_{k_\ell}$ las primeras ℓ columnas con pivote de G, y sea

$$G_\ell = [g_{k_1} | g_{k_2} | \cdots | g_{k_\ell}].$$ Entonces

$$R(G_\ell) = \{ y \in \mathbb{R}^m : \eta_{\ell+1} = \eta_{\ell+2} = \cdots = \eta_m = 0 \}$$

Demostración :

Sea $C = \{ y \in \mathbb{R}^m : \eta_{\ell+1} = \eta_{\ell+2} = \cdots = \eta_m = 0 \}$

Sea $y \in R(G_\ell) \Rightarrow y = G_\ell x$ para alguna $x \in \mathbb{R}^\ell$

Pero G_ℓ es de la forma

$$G_\ell = \begin{bmatrix} U_\ell \\ \hline 0 \end{bmatrix} \begin{matrix} \ell \\ m-\ell \end{matrix}, \quad \text{con } r(U_\ell) = \ell$$

$$\therefore \quad y = \begin{bmatrix} U_\ell \\ \hline 0 \end{bmatrix} x = \begin{bmatrix} U_\ell x \\ \hline 0 \end{bmatrix} \begin{matrix} \ell \\ m-\ell \end{matrix} \Rightarrow y \in C$$

$$\therefore \quad R(G_\ell) \subseteq C$$

sea $y \in C \Rightarrow y = \begin{bmatrix} y_1 \\ \hline 0 \end{bmatrix} \begin{matrix} \ell \\ m-\ell \end{matrix}, \quad y_1 \in \mathbb{R}^\ell$

Como $r(U_\ell) = \ell \Rightarrow R(U_\ell) = \mathbb{R}^\ell \quad \therefore$ existe una $x \in \mathbb{R}^\ell$ tal que $U_\ell x = y_1$

$$\therefore \quad G_\ell x = \begin{bmatrix} U_\ell \\ 0 \end{bmatrix} x = \begin{bmatrix} U_\ell x \\ 0 \end{bmatrix} = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} = y$$

$$\therefore \quad y \in R(G_\ell) \qquad \therefore \quad C \subseteq R(G_\ell)$$

$$\therefore \qquad C \subseteq R(G_\ell) \subseteq C$$

$$\therefore \quad R(G_\ell) = C \; \blacksquare$$

7) $$R(G) = \{ y \in \mathbb{R}^m : \eta_{r+1} = \eta_{r+2} = \cdots = \eta_m = 0 \}$$

$$= \left\{ y \in \mathbb{R}^m : \quad y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix} \quad ; \; y_1 \in \mathbb{R}^r \right\}$$

8) Las columnas que no tienen pivote son dependientes de las anteriores con pivote.

Demostración :

Sea q una columna sin pivote de G. Si q está entre las columnas q_{k_ℓ} y $q_{k_{\ell+1}}$ (que tienen pivote) entonces

$$g = \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_\ell \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \text{y por el resultado (6)}$$

$$g \in [g_{k1}, g_{k2}, \ldots, g_{k\ell}] .$$

OBSERVACIONES

1) Sean A una matriz $m \times n$, y B una matriz de orden m, con $N(B) = \{0\}$ tal que $BA = G$ es una matriz escalonada.

Sean $g_{k1}, g_{k2}, \ldots, g_{kr}$ las columnas con pivote de G. Entonces

$$\{a_{k1}, a_{k2}, \ldots, a_{kr}\} \text{ es una base de } R(A).$$

Demostración:

Sean $A_r = [a_{k1} | a_{k2} | \cdots | a_{kr}]$

$G_r = [g_{k1} | g_{k2} | \cdots | g_{kr}]$

Sabemos que $r(A) = r(G) = r$

$\therefore \quad r(A_r) \leq r$.

Además $BA = [Ba_1|Ba_2|\cdots|Ba_n] = [g_1|g_2|\cdots|g_n]$

$\therefore$ $BA_r = [Ba_{k_1}|Ba_{k_2}|\cdots|Ba_{k_r}] = [g_{k_1}|g_{k_2}|\cdots|g_{k_r}] = G_r$

$\therefore$ $BA_r = G_r$, y como $N(B) = \{0\}$ y $N(G_r) = \{0\} \to N(A_r) = \{0\}$

$\therefore$ Las columnas de A_r son independientes
$\therefore$ $r(A_r) = r = r(A)$.

$\therefore$ Las columnas de A_r forman una base de $R(A)$ ■

2) Sean A, B y G como en (1). Entonces una base de $R(A^t)$ es

$$\{z_1^t, z_2^t, \dots, z_r^t\}$$

donde $z_1^t, z_2^t, \dots, z_r^t$ son los r renglones no nulos de G.

Demostración:

$$R(A^t) = R((BA)^t) = R(G^t)$$ ■

3) Contamos con un método muy poderoso para

(1) Calcular $r(A)$
(2) Dar una base de $R(A)$
(3) Dar una base de $R(A^t)$

Trataremos de sacar el mayor provecho de esta técnica que hemos desarrollado.

Nos habíamos planteado la pregunta

¿ Cuándo $Ax=b$ tiene solución $\forall\ b \in \mathbb{R}^m$?

$$
\begin{aligned}
Ax=b \text{ tiene solución } \forall\ b \in \mathbb{R}^m \\
&\iff R(A) = \mathbb{R}^m \\
&\iff r(A) = m \leq \min\{m,n\} \\
&\Rightarrow m \leq n .
\end{aligned}
$$

Es decir, para que $Ax=b$ tenga solución para cualquier $b \in \mathbb{R}^m$, no puede haber más de n ecuaciones, y cuando menos m de ellas deben ser independientes.

También nos habíamos planteado:

Dada $b \in \mathbb{R}^m$ ¿ cuándo tiene solución el sistema $Ax = b$?

Sabemos que $Ax=b$ tiene solución $\iff b \in R(A)$

En la práctica este problema se resuelve transformando a A y a b en $\breve{A}$ y $\tilde{b}$ de tal forma que

$$b \in R(A) \iff \tilde{b} \in R(\breve{A})$$

y tal que $\tilde{b} \in R(\tilde{A})$ se pueda contestar fácilmente.

Veremos que las soluciones de un sistema de ecuaciones se conservan si transformamos a A y b en BA y Bb respectivamente, donde B es cuadrada y $N(B) = \{0\}$

RESULTADO

Sean A $m \times n$, $b \in \mathbb{R}^m$ y $B_{m \times m}$ tal que $N(B) = \{0\}$ Entonces

$$Ax_0 = b \iff BAx_0 = Bb$$

Es decir, x_0 es solución del sistema $Ax = b$ $\iff$ x_0 es solución del sistema $\tilde{A}x = \tilde{b}$, donde $\tilde{A} = BA$ y $\tilde{b} = Bb$.

Demostración :

$\Rightarrow]$ $Ax_0 = b \Rightarrow (BA)x_0 = Bb$

$\Leftarrow]$ Supongamos $(BA)x_0 = Bb$

$$\Rightarrow B(Ax_0 - b) = 0$$

$$\Rightarrow Ax_0 - b = 0 \qquad (N(B) = \{0\})$$

$$\Rightarrow Ax_0 = b$$

RESULTADO

Sean $A_{m\times n}$, $b \in \mathbb{R}^m$ y $B_{m\times m}$, con $N(B) = \{0\}$. Entonces

$$b \in R(A) \iff \tilde{b} \in R(\tilde{A})$$

donde $\tilde{A} = BA$ y $\tilde{b} = Bb$.

Demostración:

Por el resultado anterior $Ax = b \iff \tilde{A}x = \tilde{b}$ ∎

Estos dos resultados nos dan la pauta a seguir, y es natural que nos preguntemos si la matriz B que transforma a A en una matriz escalonada nos sirve.

Como habíamos visto anteriormente, si G es una matriz escalonada de rango r entonces

$$R(G) = \left\{ y \in \mathbb{R}^m : \; y = \begin{bmatrix} y_1 \\ 0 \end{bmatrix}\begin{matrix} r \\ m-r \end{matrix} \right\}$$

Es decir:

si $\tilde{b} = \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \end{bmatrix}$ entonces $\tilde{b} \in R(G) \iff b_2 = 0$

Por lo tanto para las matrices escalonadas sí es muy fácil decidir cuándo un vector está en el subespacio generado por las columnas de G.

En la práctica es usual trabajar con la matriz

$$[A \mid b]$$

que se conoce como la MATRIZ AUMENTADA del sistema $Ax=b$, haciendo uso del siguiente resultado :

RESULTADO

El sistema $Ax=b$ tiene solución $\Leftrightarrow r([A|b])=r(A)$.

Demostración :

$$Ax=b \text{ tiene solución } \Leftrightarrow b \in R(A)$$
$$\Leftrightarrow r([A|b]) = r(A) \;\blacksquare$$

Ya sabemos que el conjunto de soluciones de $Ax=0$ es un subespacio $(N(A))$.
¿ Qué forma tendrá el conjunto de soluciones del sistema $Ax=b$?

Analicemos algunos ejemplos antes de continuar:

Sea $n=2$:

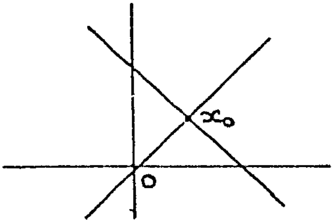
(1) $$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & 0 \\ 1 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

$$\therefore \quad \begin{array}{r} \xi_1 - \xi_2 = 0 \\ 2\,\xi_2 = 2 \end{array} \quad \Rightarrow \quad \begin{array}{l} \xi_2 = 1 \\ \xi_1 = 1 \end{array}$$

$\therefore \quad x_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ es solución del sistema.

Geométricamente $r_1^t x = 0$ y $r_2^t x = 2$ son dos rectas no-paralelas que se intersectan en el punto x_0.



El conjunto de soluciones del sistema es $\{x_0\}$.

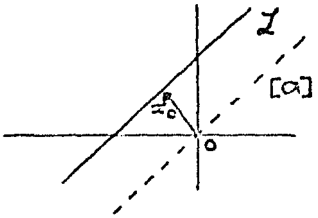
(2) $$\begin{bmatrix} 1 & -1 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} r_1^t x \\ r_2^t x \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 & -1 \\ -3 & 3 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$\therefore$ El conjunto de soluciones es la recta dada por

$$\xi_1 - \xi_2 = -1$$

Geométricamente, al resolver el sistema estamos calculando la intersección de las rectas $r_1^t x = -1$ y $r_2^t x = 3$, solo que en este caso las dos rectas coinciden.



Se puede ver que si x y x_0 son dos puntos en el conjunto de soluciones, entonces

$$x - x_0 = \lambda a \quad , \text{ donde } \quad a = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Es decir, el conjunto de soluciones del sistema es

$$\mathcal{L} = \left\{ x, x_0 \in \mathbb{R}^2 : \; x - x_0 = \lambda a \; ; \; a = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Si fijamos un punto sobre $\mathcal{L}$, por ejemplo el punto x_0, entonces el conjunto de soluciones toma la forma

$$\mathcal{L} = \left\{ x \in \mathbb{R}^2 : \ x = x_0 + \lambda a \ ; \ a = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}$$

Esto nos permite ver al conjunto de soluciones como una recta paralela a $[a]$.

Veremos que para $n = 3$ podemos tratar al conjunto de soluciones de forma análoga.

$$(1) \qquad \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 1 & 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \rightarrow$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix}$$

$$\therefore \quad \xi_3 = 0 = \xi_2 \quad ; \quad \xi_1 = 1$$

Geométricamente, estamos calculando la intersección de 3 planos, ninguno de los cuales es paralelo a los otros dos. Los planos se intersectan en el punto

$$x_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Por lo tanto, el conjunto de soluciones del sistema es $\{x_0\}$.

(2)
$$\begin{bmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 1 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & -1 & 1 & -1 \end{bmatrix} \rightarrow$$

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

El conjunto solución está dado por la intersección de los planos

$$\begin{aligned} 2\xi_1 + \xi_2 + \xi_3 &= 1 \\ \xi_2 - \xi_3 &= 1 \end{aligned}$$

Si tenemos dos puntos en el espacio de soluciones, x y x_0, entonces es fácil ver que

$$\mathcal{I} = \{x, x_0 \in \mathbb{R}^3 : x - x_0 = \lambda a \; ; \; a^t = [-1 \;\; 1 \;\; 1]\}$$

nos da el conjunto de soluciones.

Si fijamos un punto sobre $\mathcal{I}$, digamos x_0, entonces

$$\mathcal{I} = \{x \in \mathbb{R}^3 : x = x_0 + \lambda a \; ; \; a^t = [-1 \;\; 1 \;\; 1]\}.$$

(3) $$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 12 \\ -3 & -6 & -9 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{bmatrix} = \begin{bmatrix} -5 \\ -20 \\ 15 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 & 3 & -5 \\ 4 & 8 & 12 & -20 \\ -3 & -6 & -9 & 15 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 2 & 3 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

La solución del sistema es el plano

$$\xi_1 + 2\xi_2 + 3\xi_3 = -5$$

y corresponde a la intersección de 3 planos - solo que sucede que los 3 planos son en realidad uno solo.

Si x y x_0 son cualesquiera dos puntos sobre el plano, entonces

$$x - x_0 = \lambda_1 a_1 + \lambda_2 a_2$$

donde $a_1^t = [-2 \quad 1 \quad 0]$ y $a_2^t = [-3 \quad 0 \quad 1]$.

Es decir, el conjunto de soluciones está dado por

$$\mathcal{L} = \{x, x_0 \in \mathbb{R}^3 : \begin{array}{l} x - x_0 = \lambda_1 a_1 + \lambda_2 a_2 ; \\ a_1^t = [-2 \quad 1 \quad 0] \\ a_2^t = [-3 \quad 0 \quad 1] \end{array} \}$$

Si fijamos un punto sobre el plano, por ejemplo el punto x_0, entonces

$$\mathcal{L} = \{x \in \mathbb{R}^3 : \begin{array}{l} x = x_0 + \lambda_1 a_1 + \lambda_2 a_2 ; \; x_0 \in \mathcal{L} ; \\ a_1^t = [-2 \quad 1 \quad 0] , \quad a_2^t = [-3 \quad 1 \quad 0] \end{array} \}.$$

Y podemos pensar en $\mathcal{L}$ como un plano paralelo al subespacio determinado por a_1 y a_2, ó como un subespacio trasladado.

Podemos ver que los conjuntos solución son en general puntos, rectas, planos, etc., pero que no tienen que pasar por el origen, y se puede pensar en ellos como conjuntos paralelos a subespacios de $\mathbb{R}^n$ ó como subespacios trasladados. A estos conjuntos les pondremos un nombre, ya que se pueden manejar casi tan fácilmente como los subespacios.

DEFINICION

Un subconjunto $\mathcal{I} \subseteq \mathbb{R}^m$ es una VARIEDAD LINEAL si existen $a_1, a_2, \ldots, a_k \in \mathbb{R}^m$ tales que para todo $x, x_0 \in \mathcal{I}$

$$x - x_0 \in [a_1, a_2, \ldots, a_k]$$

Es decir, si $x - x_0 = \sum_{i=1}^{k} \lambda_i a_i$

Observemos que si $\mathcal{I} \subseteq \mathbb{R}^m$ es una variedad lineal, entonces el subespacio $[a_1, a_2, \ldots, a_k]$ queda determinado unívocamente por $\mathcal{I}$.

Si llamamos $\mathcal{J} = [a_1, a_2, \ldots, a_k]$ entonces

$x - x_0 \in \mathcal{J}$ para cualesquiera $x, x_0 \in \mathcal{I}$

Es decir :

$\mathcal{I}$ es una variedad lineal $\Leftrightarrow$ $\mathcal{I}$ determina unívocamente un subespacio $\mathcal{J} \subseteq \mathbb{R}^m$ tal que si $x, x_0 \in \mathcal{I}$ entonces $x - x_0 \in \mathcal{J}$

Ahora, si fijamos un punto cualquiera sobre $\mathcal{L}$, digamos x_0, entonces para todo $x \in \mathcal{L}$ tenemos que

$$x - x_0 = s \qquad \text{para alguna} \quad s \in \mathcal{S}$$

es decir, x es de la forma

$$x = x_0 + s \qquad \text{para alguna} \quad s \in \mathcal{S}$$

Por lo tanto tenemos que:

$\mathcal{L}$ es una variedad lineal $\Leftrightarrow$ existe un subespacio $\mathcal{S} \subseteq \mathbb{R}^m$ tal que

$$\mathcal{L} = \{ x \in \mathbb{R}^m : x = x_0 + s \, ; \; x_0 \in \mathcal{L} \; , \; s \in \mathcal{S} \}$$

A $\mathcal{L}$ la denotaremos por

$$\mathcal{L} = x_0 + \mathcal{S} \, .$$

Notemos que una variedad lineal puede ser un subespacio también.- no estamos excluyendo el caso de que la variedad pase por el origen. Esto nos permite hablar de la dimensión de una variedad lineal.

DEFINICION

Sea $\mathcal{L} = x_0 + \mathcal{S}$ una variedad lineal. Entonces la DIMENSION de $\mathcal{L}$ es:

$$\dim(\mathcal{L}) = \dim(\mathcal{S}) \quad .$$

RESULTADO

Sean A $m \times n$ y $b \in \mathbb{R}^m$. Si $Ax = b$ tiene solución entonces el conjunto de soluciones es

$$\mathcal{L}(b) = x_0 + N(A) \quad , \quad \text{donde} \quad Ax_0 = b \; .$$

Demostración:

Supongamos que x_0 es solución de $Ax = b$. Entonces:

$$\begin{aligned} Ax = b &\iff Ax - Ax_0 = 0 \\ &\iff A(x - x_0) = 0 \\ &\iff x - x_0 \in N(A) \\ &\iff x \in x_0 + N(A) \end{aligned}$$

∎

Hemos demostrado que el conjunto de soluciones del sistema $Ax = b$ es una variedad lineal.

RESULTADO

El sistema $Ax=b$ tiene solución única $\Leftrightarrow$ $N(A)=\{0\}$.

Demostración :

Es consecuencia inmediata del resultado anterior

Ahora, supongamos que ya sabemos que el sistema $Ax=b$ tiene solución. El siguiente paso es calcular el conjunto de soluciones

$$\mathcal{I}(b)=\{x\in\mathbb{R}^n:\ Ax=b\}$$

Como consecuencia del resultado, este problema se reduce a calcular un punto sobre $\mathcal{I}$ y encontrar el subespacio asociado a $\mathcal{I}$.

Es decir, tenemos que encontrar un Y_0 tal que $AY_0=b$, y encontrar una base de $N(A)$.

Primero veremos cómo encontrar una base de $N(A)$.

Observemos que si $N(A)\neq\{0\}$, entonces al menos una columna de A es dependiente de las demás.

si $\{a_{k_1}, a_{k_2}, \ldots, a_{k_r}\}$ es una base de $R(A)$

y a_ℓ es una columna de A que no está en la base, entonces $a_\ell \in [a_{k_1}, a_{k_2}, \ldots, a_{k_r}]$ y

$$a_\ell = \sum_{\lambda=1}^{r} \lambda_{k_i} a_{k_i}$$

entonces el vector x_ℓ dado por

$$x_\ell = [\xi_i^{(\ell)}] ; \quad \xi_\ell^{(\ell)} = \begin{cases} 0 & i \notin \{\ell, k_1, k_2, \ldots, k_r\} \\ 1 & \lambda = \ell \\ -\lambda_i & i \in \{k_1, k_2, \ldots, k_r\} \end{cases}$$

es tal que $Ax_\ell = 0$. Es decir $x_\ell \in N(A)$.

Por ejemplo, si

$$A \doteq [a_1 | a_2 | a_3 | a_4 | a_5] \quad \text{donde}$$

$$\begin{aligned} a_2 &= 2a_1 - a_3 + a_5 \\ a_4 &= a_1 + a_3 - a_5 \end{aligned} \qquad \text{entonces}$$

$$x_2 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{y} \quad x_5 = \begin{bmatrix} -1 \\ 0 \\ -1 \\ 1 \\ 1 \end{bmatrix} \quad \text{están en } N(A).$$

Observemos que si a_{ℓ_1} y a_{ℓ_2}, $\ell_1 \neq \ell_2$, son dos columnas de A que no están en la base de $R(A)$, entonces

$$\{x_{\ell_1}, x_{\ell_2}\}$$

es independiente.

Ya que si $x_{\ell_1} = [\xi_i^{(\ell_1)}]$ y $x_{\ell_2} = [\xi_i^{(\ell_2)}]$ entonces

$$\xi_{\ell_1}^{(\ell_1)} = 1 \qquad \xi_{\ell_1}^{(\ell_2)} = 0$$

$$\xi_{\ell_2}^{(\ell_1)} = 0 \qquad \xi_{\ell_2}^{(\ell_2)} = 1$$

Es decir x_{ℓ_1} y x_{ℓ_2} son de la forma:

$$x_{\ell_1} = \left[\begin{array}{c} * \\ * \\ \vdots \\ 1 \\ * \\ \vdots \\ 0 \\ \vdots \\ * \end{array}\right] \begin{array}{c} \\ \\ \\ \leftarrow \ell_1 \rightarrow \\ ; \quad x_{\ell_2} = \\ \\ \leftarrow \ell_2 \rightarrow \\ \\ \\ \end{array} \left[\begin{array}{c} * \\ * \\ \vdots \\ 0 \\ * \\ \vdots \\ 1 \\ \\ * \end{array}\right]$$

Por lo tanto, vemos que no hay forma de obtener x_{ℓ_1} a partir de x_{ℓ_2}, y recíprocamente, x_{ℓ_2} no se puede obtener como combinación lineal de x_{ℓ_1}.

Esto nos indica que el camino a seguir para calcular $N(A)$ es el siguiente:

Encontrar primero las columnas de A que sean una base de $R(A)$, y después expresar cada una de las columnas restantes en términos de la base.

Veremos que este procedimiento nos proporciona una base de $N(A)$.

Para que este procedimiento sea práctico, es necesario que el proceso de expresar las columnas en términos de la base sea sencillo, pues expresar una columna en términos de la base equivale a resolver un sistema de ecuaciones.

Veremos para qué matrices es fácil calcular $N(A)$.

Primero hay que observar que si $r(A)=n$, donde A es $m \times n$, entonces $N(A)=\{0\}$, y no tenemos que hacer nada para calcular $N(A)$. Y si $m<n$ entonces, como $r(A) \leq \min\{m,n\}$, vamos a tener que $r(A)<n$ y por lo tanto $N(A) \neq \{0\}$.

RESULTADO

Sean A $m \times n$ y B $n \times n$ tal que $N(B) = \{0\}$. Entonces

$$N(A) = N(BA)$$

Demostración

Por un resultado anterior tenemos que

$$Ax = b \iff (BA)x = (Bb) \qquad \text{si} \quad N(B) = \{0\}.$$

En particular para $b = 0$

$$Ax = 0 \iff (BA)x = 0$$

$$\therefore \quad x \in N(A) \iff x \in N(BA) \qquad \blacksquare$$

Por lo tanto, nos podemos facilitar la tarea de expresar las columnas dependientes de A en función de la base de $R(A)$ encontrando una matriz BA para la cual sea más sencillo resolver el mismo problema.

RESULTADO

Sea $A = \begin{array}{cc} n & \\ \begin{bmatrix} A_1 \\ \\ A_2 \end{bmatrix} & \begin{array}{l} r \\ \\ m-r \end{array} \end{array}$. Entonces $N(A) = N(A_1) \cap N(A_2)$.

Demostración :

$$x \in N(A) \iff Ax = 0$$

$$\iff \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] x = 0$$

$$\iff \begin{array}{c} A_1 x \\ A_2 x \end{array} = 0$$

$$\iff A_1 x = 0 \quad y \quad A_2 x = 0$$

$$\iff x \in N(A_1) \quad y \quad x \in N(A_2)$$

$$\iff x \in N(A_1) \cap N(A_2) \quad \blacksquare$$

RESULTADO

$$\text{Si} \quad A = \begin{array}{c} n \\ \left[\begin{array}{c} A_1 \\ \hline 0 \end{array}\right] \end{array} \begin{array}{l} r \\ m-r \end{array} \quad \text{entonces} \quad N(A) = N(A_1) .$$

Estos resultados nos serán de utilidad, puesto que intuímos que las matrices para las cuales es más fácil calcular $N(A)$ son las matrices escalonadas.

RESULTADO

Si $A = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ [A_1 \mid A_2]_m \end{array}$ entonces

$$N(A) = \left\{ x = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}\begin{matrix} r \\ n-r \end{matrix} \in \mathbb{R}^n : A_1\tilde{x} + A_2\tilde{y} = 0 \right\}$$

Demostración:

$$Ax = 0 \iff [A_1 \mid A_2]\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} = 0$$

$$\iff A_1\tilde{x} + A_2\tilde{y} = 0 \quad \blacksquare$$

RESULTADO

Si $A = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ [A_1 \mid A_2]_m \end{array}$ con $N(A_1) = \{0\}$ y $R(A_2) \subseteq R(A_1)$ entonces

$$\dim(N(A)) = n - r = n - r(A).$$

Demostración:

$$N(A) = \left\{ x = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}\begin{matrix} r \\ n-r \end{matrix} \in \mathbb{R}^n : A_1\tilde{x} + A_2\tilde{y} = 0 \right\}$$

Sean $x = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}$ $x' = \begin{bmatrix} \tilde{x}' \\ \tilde{y}' \end{bmatrix}$ dos vectores en $N(A)$

Entonces afirmamos que

$\{x, x'\}$ es linealmente independiente $\Leftrightarrow$
$\{\tilde{y}, \tilde{y}'\}$ es linealmente independiente :

$\Rightarrow$ Supongamos $\{x, x'\}$ es linealmente independiente.

si $\alpha\tilde{y} + \beta\tilde{y}' = 0$

$\Rightarrow$ $-A_2\ (\alpha\tilde{y} + \beta\tilde{y}') = 0$

$\Rightarrow$ $\alpha\ (-A_2\tilde{y}) + \beta(-A_2\tilde{y}') = 0$

$\Rightarrow$ $\alpha\ (A_1\tilde{x}) + \beta\ (A_1\tilde{x}') = 0$

$\Rightarrow$ $A_1(\alpha\tilde{x} + \beta\tilde{x}') = 0$

$\Rightarrow$ $\alpha\tilde{x} + \beta\tilde{x}' = 0$ $\quad (N(A_1) = \{0\})$

$$\Rightarrow \begin{cases} \alpha\tilde{x} + \beta\tilde{x}' = 0 \\ \alpha\tilde{y} + \beta\tilde{y}' = 0 \end{cases}$$

$\Rightarrow$ $\alpha x + \beta x' = 0$ $\quad (\{x, x'\}$ linealmente ind. $)$

$\Rightarrow$ $\alpha = \beta = 0$ $\quad \therefore$ $\{\tilde{y}, \tilde{y}'\}$ Independiente .

$\Leftarrow$ Supongamos $\{\tilde{y}, \tilde{y}'\}$ es linealmente independiente.

Si $\alpha x + \beta x' = 0$

$$\Rightarrow \alpha \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} + \beta \begin{bmatrix} \tilde{x}' \\ \tilde{y}' \end{bmatrix} = 0$$

$$\Rightarrow \alpha \tilde{y} + \beta \tilde{y}' = 0$$

$$\Rightarrow \alpha = \beta = 0 \quad \therefore \{x, x'\}$$ es linealmente independiente.

Además, como $R(A_2) \subseteq R(A_1)$, dada una $\tilde{y} \in IR^{n-r}$ siempre existe una $\tilde{x} \in IR^r$ tal que el vector

$$x = \begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix} \in N(A)$$

Es decir, el sistema

$$A_1 \tilde{x} = - A_2 \tilde{y}$$

siempre tiene solución.

Hemos demostrado que cada vector $x \in N(A)$ está determinado unívocamente por la correspondiente $\tilde{y} \in IR^{n-r}$, y por lo tanto tenemos tantos vectores independientes en $N(A)$ como vectores independientes hay en IR^{n-r} :

$$\dim(N(A)) = n-r \quad \square$$

RESULTADO

Sean $A_1 \in M(m,n_1)$ y $A_2 \in M(m,n_2)$ tales que $N(A_1) = \{0\}$ y $N(A_2) = \{0\}$. Sea $A = [A_1|A_2]$, entonces

$$\dim (N(A)) = \dim (R(A_1) \cap R(A_2))$$

Demostración:

Notar que $N(A)$ es un subespacio de $\mathbb{R}^{n_1+n_2}$ mientras que $R(A_1) \cap R(A_2)$ es un subespacio de $\mathbb{R}^m$.- demostraremos que estos dos subespacios tienen la misma dimensión.

Tenemos que para cualquier $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \in \mathbb{R}^n$ $(n = n_1+n_2)$:

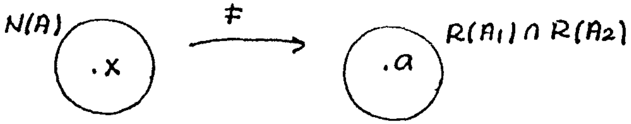
$$X \in N(A) \Leftrightarrow [A_1|A_2] \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = 0$$

$$\Leftrightarrow A_1X_1 + A_2X_2 = 0$$

$$\Leftrightarrow A_1X_1 = A_2(-X_2)$$

$$\Leftrightarrow a = A_1X_1 = A_2(-X_2) \in R(A_1) \cap R(A_2)$$

Cada $X \in N(A)$ determina una $a \in R(A_1) \cap R(A_2)$



$$x \longmapsto F(x) = a$$

Esta correspondencia $F : N(A) \longrightarrow R(A_1) \cap R(A_2)$ es biunívoca, ya que obviamente es sobre y si

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \in N(A) \, , \quad x \neq y$$

entonces $F(x) \neq F(y)$, ya que

$$x \neq y \iff x_1 \neq y_1 \text{ ó } x_2 \neq y_2$$

$$x_1 \neq y_1 \Rightarrow A_1 x_1 \neq A_1 y_1 \qquad (N(A_1) = \{0\})$$
$$\Rightarrow F(x) \neq F(y)$$

$$x_2 \neq y_2 \Rightarrow A_2 x_2 \neq A_2 y_2 \qquad (N(A_2) = \{0\})$$
$$\Rightarrow F(x) \neq F(y)$$

Además, $\{x, y\} \subseteq N(A)$ independiente $\iff$ $\{F(x), F(y)\} \subseteq R(A_1) \cap R(A_2)$ independiente,

ya que

$$\alpha x + \beta y = 0 \quad \Leftrightarrow \quad \alpha \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \beta \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = 0$$

$$\Leftrightarrow \quad \begin{aligned} \alpha x_1 + \beta y_1 &= 0 \\ \alpha x_2 + \beta y_2 &= 0 \end{aligned}$$

$$\Leftrightarrow \quad \begin{aligned} A_1 (\alpha x_1 + \beta y_1) &= 0 \qquad (N(A_1) = \{0\}) \\ A_2 (\alpha x_2 + \beta y_2) &= 0 \qquad (N(A_2) = \{0\}) \end{aligned}$$

$$\Leftrightarrow \quad \begin{aligned} \alpha (A_1 x_1) + \beta (A_1 y_1) &= 0 \\ \alpha (A_2 x_2) + \beta (A_2 y_2) &= 0 \end{aligned}$$

$$\Leftrightarrow \quad \begin{aligned} \alpha (A_1 x_1) + \beta (A_1 y_1) &= 0 \\ \alpha (A_2 (-x_2)) + \beta (A_2 (-y_2)) &= 0 \end{aligned}$$

$$\Leftrightarrow \quad \alpha F(x) + \beta F(y) = 0$$

Por lo tanto

$$\dim (N(A)) = \dim (R(A_1) \cap R(A_2))$$

$\square$

Intuimos que para cualquier matriz $A_{m\times n}$ también debemos de tener

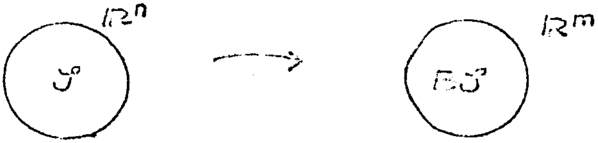
$$\dim\left(N(A)\right) = n-r \qquad \text{donde} \quad r=r(A).$$

Para poder demostrar ésto vamos a tener que probar unos resultados, ya que este problema está ligado al problema de saber cuándo

$$\dim(\mathcal{S}) = \dim(B\mathcal{S})$$

donde $\mathcal{S} \subseteq \mathbb{R}^n$ y B es una matriz $m \times n$.

Hay que observar que la matriz B está jugando el papel de una función que opera sobre $\mathcal{S}$



RESULTADO

Sea $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$. Entonces $B\mathcal{S} \subseteq \mathbb{R}^m$ es un subespacio de $\mathbb{R}^m$ y

$$\dim(\mathcal{S}) \geq \dim(B\mathcal{S}) \qquad , \text{donde } B \text{ es}$$

Demostración :

Sea $\{s_1, s_2, \ldots, s_k\}$ una base de $\mathcal{S}$, $s_i \in \mathbb{R}^n$

y sean $B = [b_1 | b_2 | \cdots | b_n]$, con $b_i \in \mathbb{R}^m$ y

$$S = [s_1 | s_2 | \cdots | s_k]; \quad r(S) = k.$$

Entonces $B\mathcal{S}$ es un subespacio de $\mathbb{R}^m$:

$$B\mathcal{S} = R(BS)$$

Es decir $B\mathcal{S}$ es el subespacio generado por las columnas de BS

Como BS tiene k columnas $\Rightarrow$ $r(BS) \leq k$.

Es decir: $\dim(B\mathcal{S}) \leq k = r(S)$

$\therefore \quad \dim(B\mathcal{S}) \leq \dim(\mathcal{S})$ ▫.

Esta es únicamente otra manera de decir que no podemos generar vectores independientes a partir de vectores dependientes, ó que no podemos aumentar el número de vectores independientes de un subconjunto de $\mathbb{R}^n$ mediante transformaciones lineales.

¿ Cuándo tenemos que $\dim(\mathcal{S}) = \dim(B\mathcal{S})$?

RESULTADO

Sea $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$ y B una matriz $m \times n$. Entonces

$$N(B) = \{0\} \Rightarrow \dim(\mathcal{S}) = \dim(B\mathcal{S}).$$

Demostración:

Si $\mathcal{S} = R(S)$ donde $S_{n \times k}$ es de rango k entonces

$$B\mathcal{S} = R(BS) \quad , \quad (BS \text{ es una matriz } m \times k)$$

como $N(B) = \{0\}$ y $N(S) = \{0\} \Rightarrow N(BS) = \{0\}$

$\therefore \quad r(BS) = k = r(S)$

$$\therefore \quad \dim(B\mathcal{S}) = \dim(\mathcal{S}) \quad \square$$

Observar que $N(B) = \{0\}$ es una condición suficiente mas no necesaria, ya que si, por ejemplo,

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad S = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \qquad \mathcal{S} = R(S) \quad \text{entonces}$$

$$N(B) \neq \{0\} \;, \quad \dim(\mathcal{S}) = 1 \quad \text{pero} \quad BS = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\therefore \quad \dim(B\mathcal{S}) = 1 = \dim(\mathcal{S})$$

RESULTADO

Sean A una matriz $m \times n$ y B una matriz de orden n con $N(B) = \{0\}$. Entonces

$$\dim(N(A)) = \dim(N(AB)).$$

Demostración:

Sean $\mathcal{S}_1 = N(A) = \{x \in \mathbb{R}^n : Ax = 0\}$

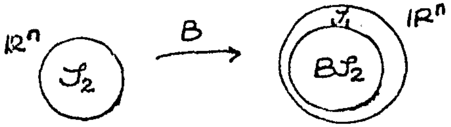
$\mathcal{S}_2 = N(AB) = \{y \in \mathbb{R}^n : (AB)y = 0\}$

Podemos ver a B como una función $B: \mathbb{R}^n \to \mathbb{R}^n$

Veremos que $B(\mathcal{S}_2) = \mathcal{S}_1$

Sea $y \in \mathcal{S}_2 \Rightarrow (AB)y = 0$
$\Rightarrow A(By) = 0$
$\Rightarrow By \in \mathcal{S}_1$

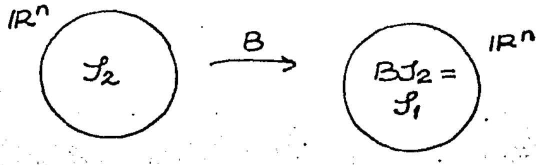
$\therefore \quad B\mathcal{S}_2 \subseteq \mathcal{S}_1$



Ahora, como B es de orden n y $N(B) = \{0\}$ $\Rightarrow R(B) = \mathbb{R}^n$. Por lo tanto dada $x \in \mathcal{S}_1 \subseteq \mathbb{R}^n$ existe $y \in \mathbb{R}^n$ tal que $By = x$, además $y \in \mathcal{S}_2$ ya que

$$ABy = Ax = 0$$

$$\therefore \quad \mathcal{J}_1 \subseteq B\mathcal{J}_2 \qquad \therefore \quad B\mathcal{J}_2 = \mathcal{J}_1$$



Por el resultado anterior tenemos que

$$\dim(\mathcal{J}_2) = \dim(B\mathcal{J}_2) = \dim(\mathcal{J}_1)$$

$$\therefore \quad \dim(N(A)) = \dim(N(AB)) \quad \square$$

Observar que $N(A) \neq N(AB)$ en general, ya que si, por ejemplo

$$A = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad y \quad P = P_{23} P_{24} P_{45}$$

entonces $$AP = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad y$$

$N(A)$ está generado por $x = \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ y

$N(AP)$ por $x = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$

Ahora ya contamos con la herramienta para demostrar el siguiente

RESULTADO

Sea A una matriz $m \times n$. Entonces

$$\dim(N(A)) + \dim(R(A)) = n.$$

Demostración.

Suponqamos $r(A) = r$. Por lo tanto existe una permutación P $n \times n$ tal que

$$AP = \begin{array}{c} \begin{array}{cc} r & n-r \end{array} \\ \left[A_{11} \mid A_{12} \right] \end{array} , \quad \text{con} \quad R(A_{12}) \subseteq R(A_{11})$$

y $r(A_{11}) = r$. (En A_{11} están unas columnas de A que forman una base de $R(A)$).

$\therefore \quad \dim(N(A)) = \dim(N(AP)) = n - r$ ▫

Sea $A = \begin{array}{cc} r & n-r \\ \left[A_{11} \right. & \left. A_{12} \right] \end{array} m$, donde $R(A_{12}) \subseteq R(A_{11})$.

Entonces, para encontrar una base de $N(A)$ tenemos que resolver los sistemas

$$\begin{aligned} A_{11}\tilde{x} &= -A_{12}e_1 \\ A_{11}\tilde{x} &= -A_{12}e_2 \\ &\vdots \\ A_{11}\tilde{x} &= -A_{12}e_{n-r} \end{aligned}$$

donde $e_i \in \mathbb{R}^{n-r}$, $1 \leq \ell \leq n-r$

Es decir, tenemos que resolver los sistemas

$$A_{11}\tilde{x} = -\tilde{a}_i \qquad 1 \leq \ell \leq n-r$$

donde $A_{12} = [\tilde{a}_1 \mid \tilde{a}_2 \mid \cdots \mid \tilde{a}_{n-r}]$.

Si $\tilde{x}_i$ es la solución del sistema

$$A_{11}\tilde{x} = -a_i \qquad 1 \leq i \leq n-r$$

entonces una base de $N(A)$ estará dada por $\{x_1, x_2, \dots, x_{n-r}\}$ donde

$$x_i = \begin{array}{c} \begin{bmatrix} \tilde{x}_i \\ \\ e_i \end{bmatrix} \end{array} \begin{array}{l} r \\ \\ n-r \end{array} \qquad 1 \leq i \leq n-r.$$

Veamos ahora cómo encontrar una base de $N(G)$ para matrices escalonadas.

Sea $\underset{r\times n}{G} = [g_1 | g_2 | \cdots | g_n]$ una matriz escalonada de orden r.

Sean $g_{k_1}, g_{k_2}, \ldots, g_{k_r}$ las columnas con pivote.

Sea g_j la j-ésima columna de G, y supongamos que g_j no tiene pivote, entonces

$$g_j \in [g_{k_1}, g_{k_2}, \ldots, g_{k_\ell}] \subseteq [g_{k_1}, g_{k_2}, \ldots, g_{k_r}]$$

donde $g_{k_1}, g_{k_2}, \ldots, g_{k_\ell}$ son las columnas con pivote que están antes que g_j.

$$\therefore\ g_j = \lambda_{k_1}^{(j)} g_{k_1} + \lambda_{k_2}^{(j)} g_{k_2} + \cdots + \lambda_{k_\ell}^{(j)} g_{k_\ell} + 0 \cdot g_{k_{\ell+1}} + \cdots + 0 \cdot g_{k_r}$$

$$\therefore\ g_j = G_r \tilde{x}_j$$

donde $G_r = [g_{k_1} | g_{k_2} | \cdots | g_{k_r}]$ y

$$\tilde{x}_j = \begin{bmatrix} \lambda_{k_1}^{(j)} \\ \lambda_{k_2}^{(j)} \\ \vdots \\ \lambda_{k_\ell}^{(j)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Notemos que la matriz G_r es triangular y que por lo tanto el sistema

$$G_r \tilde{X} = g_j$$

es muy fácil de resolver

sea $x_j = \left[E_i^{(j)} \right] \in \mathbb{R}^n$ tal que

$$E_i^{(j)} = \begin{cases} 0 & \text{si} \quad i \notin \{ j, k_1, k_2, \ldots, k_l \} \\ 1 & \text{si} \quad i = j \\ -\lambda_{k_i}^{(j)} & \text{si} \quad i \in \{ k_1, k_2, \ldots, k_l \} \end{cases}$$

Entonces $x_j \in N(G)$ y los $(n-r)$ vectores obtenidos de este modo forman una base de $N(G)$:

si $x_j = \left[E_i^{(j)} \right] \in N(G)$ entonces

$$E_j^{(j)} = 1 \quad \text{y} \quad E_i^{(j)} = 0 \quad \text{para} \quad i > j \; .$$

De aquí se puede ver que X_j es independiente de los otros vectores obtenidos de este modo.

Ejemplo

Si $$G = \begin{bmatrix} \textcircled{1} & 1 & 1 & 0 & 1 \\ 0 & \textcircled{1} & 0 & 1 & 1 \\ 0 & 0 & 0 & \textcircled{1} & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = [g_1 | g_2 | g_3 | g_4 | g_5]$$

g_3 es combinación lineal de g_1 y g_2

$$g_3 = 1 \cdot g_1 + 0 \cdot g_2 + 0 \cdot g_4$$

$$\therefore \quad \tilde{x}_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \xi_1^{(3)} \\ \xi_2^{(3)} \\ \xi_4^{(3)} \end{bmatrix} \therefore \quad x_3 = \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

g_5 es combinación lineal de g_1, g_2 y g_4

$$g_5 = 1 \cdot g_1 + 0 \cdot g_2 + 1 \cdot g_4$$

$$\therefore \quad \tilde{x}_5 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \xi_1^{(5)} \\ \xi_2^{(5)} \\ \xi_4^{(5)} \end{bmatrix} \quad \therefore \quad x_5 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -1 \\ 1 \end{bmatrix}$$

$\therefore \{x_3, x_5\}$ es una base de $N(G)$.

Regresemos ahora a nuestro problema original:

Dada $b \in \mathbb{R}^m$ y A $m \times n$ encontrar

$$\mathcal{L}(b) = \{ x \in \mathbb{R}^n : Ax = b \} = x_0 + N(A).$$

Para calcular $N(A)$ hacemos lo siguiente:

Transformamos a A en $G = BA$, donde G es escalonada y B es de orden m con $N(B) = \{0\}$.

Como vimos anteriormente, $N(A) = N(G)$. Por lo tanto, ya sabemos calcular $N(A)$.

Para tener a $L(b)$ únicamente nos falta encontrar un punto $x_0 \in \mathcal{L}(b)$:

Si $\tilde{b} = Bb$ entonces sabemos que

$$Ax = b \Leftrightarrow Gx = \tilde{b}$$

Por lo tanto para encontrar un $x_0 \in \mathcal{L}(b)$ basta con expresar a $\tilde{b}$ como combinación lineal de las columnas de G que tienen pivote, ya que éstas forman una base de $R(G)$.

Es decir, si $x_0 = [\xi_i^{(0)}]$, entonces tomamos

$\xi_j^{(0)} = 0$ si j es un índice no asociado a una columna con pivote :

Más explícitamente

Si $q_{k_1}, q_{k_2}, \ldots, q_{k_r}$ son las columnas con pivote y

$$G_r = [q_{k_1} | q_{k_2} | \cdots | q_{k_r}]$$

entonces sea $\tilde{x}$ tal que

$$G_r \tilde{x} = \tilde{b}$$

Este es un sistema triangular, que es fácil de resolver

Si $\tilde{x} = \begin{bmatrix} \xi_{k_1} \\ \xi_{k_2} \\ \vdots \\ \xi_{k_r} \end{bmatrix}$ entonces

sea $x_0 \in \mathbb{R}^n$ dada por

$$x_0 = [\xi_i^{(0)}] \quad \text{donde} \quad \xi_i^{(0)} = \begin{cases} \xi_{k_i} & i \in \{k_1, k_2, \ldots, k_r\} \\ 0 & i \notin \{k_1, k_2, \ldots, k_r\} \end{cases}$$

entonces $Ax_0 = b$.

Ejemplo

Resolver el sistema $Ax=b$ donde

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

$$\left[\begin{array}{ccccc|c}1&2&-1&4&0&3\\0&1&0&2&0&1\\1&3&-1&6&0&4\\1&1&-1&2&1&3\end{array}\right]=[A\,|\,b]$$

$$\rightarrow\left[\begin{array}{c:cccc|c}1&2&-1&4&0&3\\ \hdashline 0&1&0&2&0&1\\0&1&0&2&0&0\\0&1&0&2&-1&0\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccc:cc|c}1&2&-1&4&0&3\\0&1&0&2&0&1\\ \hdashline 0&0&0&0&0&0\\0&0&0&0&1&1\end{array}\right]$$

$$\rightarrow\left[\begin{array}{ccccc|c}\textcircled{1}&2&-1&4&0&3\\0&\textcircled{1}&0&2&0&1\\0&0&0&0&\textcircled{1}&1\\ \hdashline 0&0&0&0&0&0\end{array}\right]=[E\,|\,\tilde{b}]$$

El sistema sí tiene solución ya que $\tilde{b} \in R(G)$

Primero encontraremos una base de $N(A) = N(G)$:

La 3^a columna es una combinación lineal de la 1^a y la 2^a columnas:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} E_1^{(3)} \\ E_2^{(3)} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \Rightarrow \begin{array}{l} E_2^{(3)} = 0 \\ E_1^{(3)} = -1 \end{array}$$

$$\therefore \quad \tilde{x}_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix}$$

$$\therefore \quad x_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

La 4^a columna está dada por

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} E_1^{(4)} \\ E_2^{(4)} \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} \Rightarrow \begin{array}{l} E_5^{(4)} = 0 \\ E_2^{(4)} = 2 \end{array}$$

$$\therefore \quad \tilde{x}_4 = \begin{bmatrix} 0 \\ 2 \\ 0 \end{bmatrix} \qquad \therefore \quad x_4 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$\therefore$ Una base de $N(A)$ es $\{x_2, x_4\}$

Ahora encontraremos un punto en $\mathcal{I}(b)$:

Hay que resolver el sistema

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1^{(0)} \\ \xi_2^{(0)} \\ \xi_5^{(0)} \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

De donde se ve que $\xi_5 = \xi_2 = \xi_1 = 1$

$$\therefore \quad \tilde{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\therefore \quad x_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{I}(b)$$

$$\therefore \quad \mathcal{I}(b) = x_0 + [x_2, x_4]$$

$$\therefore \quad \tilde{x}_4 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \qquad \therefore \quad x_4 = \begin{bmatrix} 0 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$\therefore$ Una base de $N(A)$ es $\{x_2, x_4\}$

Ahora encontraremos un punto en $\mathcal{I}(b)$:

Hay que resolver el sistema

$$\begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1^{(0)} \\ \xi_2^{(0)} \\ \xi_5^{(0)} \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}$$

Claramente $\xi_5^{(0)} = \xi_2^{(0)} = \xi_1^{(0)} = 1$

$$\therefore \quad \tilde{x}_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\therefore \quad x_0 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \in \mathcal{I}(b)$$

$$\therefore \quad \mathcal{I}(b) = x_0 + [x_2, x_4]$$

FACTORIZACION $PA = LU$

Supongamos que tenemos una matriz $A_{m\times m}$

Entonces, como vimos en el primer capítulo podemos llevar a A a la forma triangular superior, y si en el proceso no es necesario permutar los renglones de A entonces tenemos que

$$A = LU$$

donde L es triangular inferior y U es triangular superior, y $N(L) = \{0\}$.

Veremos que en el caso más general existen una permutación P, una matriz triangular inferior L y una matriz triangular superior U tal que

$$PA = LU \qquad N(L) = \{0\}$$

Para llevar a A a la forma triangular lo hicimos por pasos

$$\begin{aligned} A = A_0 &\to A_1 = L_1 P_1 A_0 \\ A_1 &\to A_2 = L_2 P_2 A_1 \\ A_2 &\to A_3 = L_3 P_3 A_2 \\ &\vdots \qquad\quad \vdots \\ A_{n-2} &\to A_{n-1} = L_{n-1} P_{n-1} A_{n-2} \end{aligned}$$

Para obtener $BA = LU$ donde

$$B = L_{m-1}P_{m-1}L_{m-2}P_{m-2}\cdots L_3P_3L_2P_2L_1P_1$$
$$U = A_{m-1}$$

y donde

P_i es una permutación elemental de la forma

$$P_i = \begin{array}{cccc} & i-1 & & \\ i-1 & \left[\begin{array}{c} I_{i-1} \\ \hline 0 \end{array}\right. & \left.\begin{array}{c} 0 \\ \hline p^{(i)} \end{array}\right] & \\ & & & m-(i-1) \\ & & m-(i-1) & \end{array}$$

con $P^{(i)}$ permutación elemental

y L_i es una matriz triangular inferior de la forma

$$L_i = \begin{array}{cccc} & i-1 & 1 & m-i \\ i-1 & I_{i-1} & \begin{array}{c} 0 \\ \vdots \\ 0 \end{array} & 0 \\ 1 & 0 \cdots 0 & 1 & 0 \; 0 \; 0 \cdots 0 \\ m-i & 0 & \begin{array}{c} * \\ * \\ * \\ \vdots \\ * \end{array} & \begin{array}{ccccc} 1 & & & & \\ 0 & 1 & & & \\ 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & & \\ 0 & 0 & 0 & \cdots & 1 \end{array} \end{array} = \begin{array}{cccc} & i & & \\ i & \left[\begin{array}{c} I_i \\ \hline M^{(i)} \end{array}\right. & \left.\begin{array}{c} 0 \\ \hline I_{m-i} \end{array}\right] & m-i \\ & & m-i & \end{array}$$

Veremos que es posible intercambiar los renglones que sean necesarios ANTES de realizar las operaciones para triangularizar a A ; es decir, queremos ver que existe una permutación P tal que

$$PA = LU \qquad N(L) \neq \{0\}$$

Lo que nosotros tenemos es una matriz B, con $N(B) = \{0\}$ tal que $BA = U$.

Procederemos a descomponer a B en el producto de dos matrices

$$B = \tilde{L} P$$

donde P es una matriz de permutación y $\tilde{L}$ triangular inferior, y entonces veremos que existe una L triangular inferior tal que $L\tilde{L} = I_n$, y por lo tanto

$$L\tilde{L} PA = LBA = LU$$

$$\therefore \quad PA = LU$$

Observemos también que es natural esperar que la permutación P que necesitamos sea la matriz

$$P = P_{n-1} P_{n-2} \cdots P_2 P_1$$

Trataremos de encontrar $\tilde{L}$ tal que

$\tilde{L}P = B$, es decir, tal que

$$\tilde{L}P = L_{m-1}P_{m-1}\,L_{m-2}P_{m-2}\cdots L_3P_3\,L_2P_2\,L_1P_1$$

Observemos que las matrices L_i son de la forma

$$(*)\qquad L_i = \begin{array}{cc} & \begin{array}{cc} i & \quad \end{array} \\ \begin{array}{c} i \\ \\ \end{array} & \left[\begin{array}{c|c} L^{(i)} & 0 \\ \hline M^{(i)} & I_{m-i} \end{array}\right] m-i \\ & \begin{array}{cc} \quad & m-i \end{array} \end{array}$$

donde $L^{(i)}$ es triangular inferior y $N(L^{(i)}) = \{0\}$.

tomando $L^{(i)} = I_i$.

En la discusión que sigue, cada vez que nos refiramos a una matriz L_i ($\tilde{L}_i, \tilde{L}_i$) estará implícito que L_i ($\tilde{L}_i, \tilde{L}_i$) es de la forma arriba indicada.

Esto nos será útil ya que resolveremos el problema por pasos, operando en cada paso sobre una matriz L_i y obteniendo una matriz en la cual hemos aumentado en una columna y un renglón la dimensión de la primera submatriz . En el último paso la matriz resultante es $\tilde{L}$

Observemos que si L_i y $\tilde{L}_i$ son dos matrices de la forma (*) entonces el producto $L_i\tilde{L}_i$ también lo es, ya que

$$L_i\tilde{L}_i = \begin{array}{c} \\ i \\ \\ \end{array}\left[\begin{array}{c|c} \overset{i}{L^{(i)}} & \overset{m-i}{0} \\ \hline M^{(i)} & I_{m-i} \end{array}\right] \begin{array}{c} i \\ \\ \end{array}\left[\begin{array}{c|c} \overset{i}{\tilde{L}^{(i)}} & \overset{m-i}{0} \\ \hline \tilde{M}^{(i)} & I_{m-i} \end{array}\right]_{m-i} = \begin{array}{c} i \\ \\ \end{array}\left[\begin{array}{c|c} \overset{i}{L^{(i)}\tilde{L}^{(i)}} & \overset{m-i}{0} \\ \hline M^{(i)}\tilde{L}^{(i)}+\tilde{M}^{(i)} & I_{m-i} \end{array}\right]$$

$$= \begin{array}{c} i \\ \\ \end{array}\left[\begin{array}{c|c} \overset{i}{\bar{L}^{(i)}} & 0 \\ \hline \bar{M}^{(i)} & \underset{m-i}{I_{m-i}} \end{array}\right]_{m-i} = \bar{L}_i \quad , N(\bar{L}^{(i)}) = \{0\}$$

Esto es consecuencia de que el producto de matrices triangulares inferiores es triangular inferior y $\therefore$ $\bar{L}^{(i)} = L^{(i)}\tilde{L}^{(i)}$ es triangular inferior.

Recordemos que si P es una permutación elemental entonces $P^2 = I$.

Veamos qué pasa si hacemos la siguiente operación:

$$\begin{aligned} B &= L_{m-1}P_{m-1}L_{m-2}P_{m-2}\cdots L_3P_3L_2P_2L_1P_1 \\ &= L_{m-1}P_{m-1}L_{m-2}P_{m-2}\cdots L_3P_3L_2P_2L_1(P_2P_2)P_1 \\ &= L_{m-1}P_{m-1}L_{m-2}P_{m-2}\cdots L_3P_3L_2(P_2L_1P_2)P_2P_1 \end{aligned}$$

Y si $P_2L_1P_2 = \tilde{L}_2$ entonces tendríamos que

$$B = L_{m-1} P_{m-1} L_{m-2} P_{m-2} \cdots L_3 P_3 L_2 \tilde{L}_2 \quad P_2 P_1$$
$$= L_{m-1} P_{m-1} \; L_{m-2} \; P_{m-2} \cdots L_3 P_3 \bar{L}_2 \quad P_2 P_1 ,$$ donde
$$\bar{L}_2 = L_2 \tilde{L}_2$$

En el siguiente paso

$$B = L_{m-1} P_{m-1} L_{m-2} P_{m-2} \cdots L_3 \left(P_3 \bar{L}_2 P_3 \right) P_3 P_2 P_1$$

y si $P_3 \bar{L}_2 P_3 = \tilde{L}_3$ entonces

$$B = L_{m-1} P_{m-1} L_{m-2} P_{m-2} \cdots L_3 \tilde{L}_3 \quad P_3 P_2 P_1$$
$$= L_{m-1} P_{m-1} L_{m-2} \; P_{m-2} \cdots \bar{L}_3 \quad P_3 P_2 P_1 ,$$ donde $\bar{L}_3 = L_3 \tilde{L}_3$

Y en general si para alguna i tenemos que

$$B = L_{m-1} P_{m-1} L_{m-2} \; P_{m-2} \cdots L_i P_i \bar{L}_{i-1} \left(P_{i-1} P_{i-2} \cdots P_2 P_1 \right)$$

entonces, si $P_i \bar{L}_{i-1} P_i = \tilde{L}_i$ es de la forma (*), vamos a tener que

$$B = L_{m-1} P_{m-1} L_{m-2} P_{m-2} \cdots L_i \tilde{L}_i \left(P_i P_{i-1} \cdots P_2 P_1 \right)$$

es decir, que

$$B = L_{m-1} P_{m-1} \; L_{m-2} \; P_{m-2} \cdots \bar{L}_i \left(P_i P_{i-1} \cdots P_2 P_1 \right)$$

donde $\bar{L}_i = L_i \tilde{L}_i$.

Si para $i \leq 2 \leq m-1$ tenemos que

$$P_i \bar{L}_{i-1} P_i = \tilde{L}_i$$

(tomando $\bar{L}_1 = L_1$) entonces para $i = m-1$ tendremos que

$$\begin{aligned} B &= L_{m-1} P_{m-1} \bar{L}_{m-2} (P_{m-2} \cdots P_2 P_1) \\ &= L_{m-1} P_{m-1} L_{m-2} P_{m-1} (P_{m-1} P_{m-2} \ldots P_2 P_1) \\ &= L_{m-1} \tilde{L}_{m-1} (P_{m-1} P_{m-2} \ldots P_2 P_1) \\ &= \bar{L}_{m-1} (P_{m-1} P_{m-2} \cdots P_2 P_1) \end{aligned}$$

Es decir, B es de la forma

$$B = \tilde{L} P$$

donde $\tilde{L} = \bar{L}_{m-1}$ con $N(\tilde{L}) = \{0\}$

$P = P_{m-1} P_{m-2} \cdots P_2 P_1$

Veremos que existe una matriz L triangular inferior, con $N(L) = \{0\}$ tal que

$$PA = LU$$

Esta L existe ya que $N(\tilde{L}) = \{0\}$; y es tal que

$$L\tilde{L} = I_m$$

Veremos que, en efecto, para $2 \le i \le m-1$

$$P_i \bar{L}_{i-1} P_i = \tilde{L}_i$$

$$P_i \bar{L}_{i-1} P_i = \left[\begin{array}{c|c} I_{i-1} & 0 \\ \hline 0 & p^{(i)} \end{array}\right] \left[\begin{array}{c|c} \bar{L}^{(i-1)} & 0 \\ \hline \bar{M}^{(i-1)} & I_{m-(i-1)} \end{array}\right] \left[\begin{array}{c|c} I_{i-1} & 0 \\ \hline 0 & p^{(i)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} \bar{L}^{(i-1)} & 0 \\ \hline p^{(i)} \bar{M}^{(i-1)} & p^{(i)} \end{array}\right] \left[\begin{array}{c|c} I_{i-1} & 0 \\ \hline 0 & p^{(i)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} \bar{L}^{(i-1)} & 0 \\ \hline p^{(i)} \bar{M}^{(i-1)} & p^{(i)} p^{(i)} \end{array}\right] = \left[\begin{array}{c|c} \bar{L}^{(i-1)} & 0 \\ \hline p^{(i)} \bar{M}^{(i-1)} & I_{m-(i-1)} \end{array}\right]$$

$$= \begin{array}{cc} & \begin{array}{ccc} \qquad i-1 \qquad & 1 & \qquad m-i \qquad \end{array} \\ \begin{array}{c} i-1 \\ \\ \\ \\ m-i \end{array} & \left[\begin{array}{cccc|c:cccc} & & & & 0 & & & & \\ & \bar{L}^{(i-1)} & & & 0 & & 0 & & \\ & & & & \vdots & & & & \\ & & & & 0 & & & & \\ \hline * & * & \cdots & * & 1 & 0 & 0 & \cdots & 0 \\ \hdashline * & * & \cdots & * & 0 & & & & \\ * & * & \cdots & * & 0 & & I_{m-i} & & \\ \vdots & \vdots & & \vdots & \vdots & & & & \\ * & * & \cdots & * & 0 & & & & \end{array}\right] \end{array} = \begin{array}{cc} & \begin{array}{cc} i & m-i \end{array} \\ \begin{array}{c} i \\ m-i \end{array} & \left[\begin{array}{c|c} \tilde{L}^{(i)} & 0 \\ \hline \tilde{M}^{(i)} & I_{m-i} \end{array}\right] \end{array} = \tilde{L}_i$$

Además, como cada $\tilde{L}_i$ $2 \le i \le m-1$ es producto de matrices $m \times m$ de rango m $\to$ $\tilde{L}_i$ es de rango m y por lo tanto $N(\tilde{L}_i) = \{0\}$.

Hemos demostrado que dada $A_{m \times m}$ entonces existen una permutación P, una matriz triangular inferior L y una matriz triangular superior tal que

$$PA = LU \qquad N(L) = \{0\}.$$

Ejemplo

$$\text{Sea} \quad A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix}.$$

Encontrar P, L y U tal que $PA = LU$, $N(L) = \{0\}$.

Sea $P_1 = P_{12}$

$$P_1 A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix}$$

Sea $L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$

$$L_1 P_1 A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 \end{bmatrix} = A_1$$

Sea $P_2 = P_{23}$

$$P_2 A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & -1 & -1 \end{bmatrix}$$

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

$$L_2 P_2 A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -1 \end{bmatrix} = A_2$$

Sea $P_3 = P_{34}$

$$P_3 A_2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = A_3 = U$$

Por lo tanto

$$P = P_3 P_2 P_1 = P_{34} P_{23} P_{12}$$

$$= P_{34} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= P_{34} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

$$B = P_3 (L_2 P_2)(L_1 P_1) = P_3 L_2 (P_2 L_1 P_2)(P_2 P_1)$$

$$P_2 L_1 P_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} \cdot P_2 =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} = \tilde{L}_2$$

$$L_2 \tilde{L}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix} = \bar{L}_2$$

$$\therefore \quad B = P_3 \bar{L}_2 P_3 (P_3 P_2 P_1)$$

$$P_{\bar{3}} \bar{L}_2 P_3 = P_{\bar{3}4} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \tilde{L}_3 = \tilde{L}$$

Sea $L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ entonces

$$L\tilde{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

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

Por lo tanto

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Como seguramente se habrá notado, no obtuvimos un método para calcular la matriz L tal que $L\tilde{L} = I$. Esto se puede solucionar de la siguiente forma:

Cada una de las matrices triangulares que aparecen en el producto

$$B = L_{m-1} P_{m-1} \cdots L_2 P_2 L_1 P_1$$

es de la forma

$$L_i = \begin{array}{r|c|c|l} & i & & \\ i & I_i & 0 & \\ \hline & M^{(i)} & I_{m-i} & m-i \\ & & m-i & \end{array}$$

donde $M^{(i)}$ es una matriz $(m-i) \times i$

$$M^{(i)} = \begin{array}{cl} \left[\begin{matrix} 0 & 0 & \cdots & 0 & * \\ 0 & 0 & \cdots & 0 & * \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & * \end{matrix}\right] & m-i \\ i & \end{array}$$

$$= \ell_i e_i \quad ; \quad \ell_i \in \mathbb{R}^{m-i} \quad e_i \in \mathbb{R}^{i}$$

Entonces, para cada permutación P_k, con $k > i$,

$P_k L_i P_k$ es una matriz con la misma estructura que L_i :

si $k = i+q$, con $q \geq 1$

$$P_k = \begin{array}{c} \\ i+(q-1) \\ \\ \end{array} \begin{array}{c} i+(q-1) \\ \left[\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & p^{(k)} \end{array}\right]_{m-k} \\ m-[i+(q-1)] \end{array} = \begin{array}{c} \\ \left[\begin{array}{c|c} I_{i+(q-1)} & 0 \\ \hline 0 & p^{(k)} \end{array}\right]_{m-k} \\ m-k \end{array} = \begin{array}{c} \\ i \\ \\ \end{array} \begin{array}{c} i \\ \left[\begin{array}{c|c} I_i & 0 \\ \hline 0 & p'^{(k)} \end{array}\right]_{m-i} \\ m-i \end{array}$$

entonces

$$P_k L_i P_k = \begin{array}{c} \\ i \\ m-i \\ \\ \end{array} \begin{array}{c} i \\ \left[\begin{array}{c|c} I_i & 0 \\ \hline 0 & p'^{(k)} \end{array}\right] \\ m-i \end{array} \left[\begin{array}{c|c} I_i & 0 \\ \hline M^{(i)} & I_{m-i} \end{array}\right] \begin{array}{c} i \\ \left[\begin{array}{c|c} I_i & 0 \\ \hline 0 & p'^{(k)} \end{array}\right] \\ m-i \end{array} \begin{array}{c} \\ i \\ m-i \\ \\ \end{array}$$

$$= \left[\begin{array}{c|c} I_i & \\ \hline p'^{(k)} M^{(i)} & p'^{(k)} \end{array}\right] \left[\begin{array}{c|c} I_i & 0 \\ \hline 0 & p'^{(k)} \end{array}\right] = \left[\begin{array}{c|c} I_i & 0 \\ \hline p'^{(k)} M^{(i)} & p'^{(k)} p'^{(k)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_i & 0 \\ \hline p'^{(k)} M^{(i)} & I_{m-i} \end{array}\right]$$

Como $p'^{(k)} M^{(i)} = p'^{(k)} \ell_i e_i^t = (p'^{(k)} \ell_i) e_i^t = \ell_i' e_i^t$

$= M'^{(i)}$

$$P_k L_i P_k = \left[\begin{array}{c|c} I_i & 0 \\ \hline M'^{(i)} & I_{m-i} \end{array}\right] = \tilde{L}_i$$

Por lo tanto podemos factorizar a

$$B = L_{m-1}P_{m-1}\cdots L_2P_2L_1P_1$$

de la siguiente manera:

$$\begin{aligned}
B &= L_{m-1}(P_{m-1}L_{m-2}P_{m-1})(P_{m-1}P_{m-2}L_{m-3}P_{m-3}\cdots L_1P_1)\\
&= L_{m-1}\tilde{L}_{m-2}(P_{m-1}P_{m-2}L_{m-3}P_{m-2}P_{m-1})(P_{m-1}P_{m-2}P_{m-3}\cdot\\
&\quad \cdot L_{m-4}P_{m-4}L_{m-5}P_{m-5}\cdots P_2L_1P_1)\\
&= L_{m-1}\tilde{L}_{m-2}\tilde{L}_{m-3}(P_{m-1}P_{m-2}P_{m-3}L_{m-4}P_{m-3}P_{m-2}P_{m-1})\cdot\\
&\quad \cdot(P_{m-1}P_{m-2}P_{m-3}P_{m-4}L_{m-5}P_{m-5}\cdots L_2P_2L_1P_1)\\
&= L_{m-1}\tilde{L}_{m-2}\tilde{L}_{m-3}\tilde{L}_{m-4}\cdots(P_{m-1}P_{m-2}\cdots P_3L_2P_3\cdots P_{m-2}P_{m-1})\cdot\\
&\quad \cdot(P_{m-1}P_{m-2}\cdots P_3P_2L_1P_1)\\
&= L_{m-1}\tilde{L}_{m-2}\tilde{L}_{m-3}\cdots\tilde{L}_2(P_{m-1}P_{m-2}\cdots P_2L_1P_2\cdots P_{m-2}P_{m-1})\cdot\\
&\quad \cdot(P_{m-1}P_{m-2}\cdots P_2P_1)\\
&= L_{m-1}\tilde{L}_{m-2}\tilde{L}_{m-3}\cdots\tilde{L}_2\tilde{L}_1P
\end{aligned}$$

$$\therefore\quad B = \tilde{L}_{m-1}\tilde{L}_{m-2}\cdots\tilde{L}_2\tilde{L}_1P = \tilde{L}P \qquad (\tilde{L}_{m-1} = L_{m-1})$$

donde $P = P_{m-1}P_{m-2}\cdots P_2P_1$ y

$$\tilde{L}_i = P_{m-1}P_{m-2}\cdots P_{i+1}L_iP_{i+1}P_{i+2}\cdots P_{m-2}P_{m-1}$$

y donde cada matriz $\tilde{L}_i$ es de la forma

$$\tilde{L}_i = \left[\begin{array}{c|c} I_i & 0 \\ \hline \tilde{M}^{(i)} & I_{m-i} \end{array}\right], \quad \text{con} \quad \tilde{M}^{(i)} = \tilde{l}_i e_i^t \qquad \begin{array}{l} \tilde{l}_i \in \mathbb{R}^{m-i} \\ e_i \in \mathbb{R}^i \end{array}$$

$$1 \leq i \leq m-1$$

Sea $\bar{L}_i = \left[\begin{array}{c|c} I_i & 0 \\ \hline -\tilde{M}^{(i)} & I_{m-i} \end{array}\right]$

Por lo tanto

$$\bar{L}_i L_i = \left[\begin{array}{c|c} I_i & 0 \\ \hline -\tilde{M}^{(i)} & I_{m-i} \end{array}\right] \left[\begin{array}{c|c} I_i & 0 \\ \hline \tilde{M}^{(i)} & I_{m-i} \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_i & 0 \\ \hline -\tilde{M}^{(i)} + \tilde{M}^{(i)} & I_{m-i} \end{array}\right] = \left[\begin{array}{c|c} I_i & 0 \\ \hline 0 & I_{m-i} \end{array}\right] = I_m$$

Y por lo tanto si

$$\tilde{L}_{m-1} \tilde{L}_{m-2} \cdots \tilde{L}_2 \tilde{L}_1 = \tilde{L}$$

entonces

$$\bar{L}_1 \bar{L}_2 \cdots \bar{L}_{m-1} \tilde{L}_{m-1} \cdots \tilde{L}_2 L_1 = L \tilde{L} = I$$

y por lo tanto la matriz L tal que $BA = LU$ está dada por

$$L = \tilde{L}_1 \tilde{L}_2 \cdots \tilde{L}_{m-1}$$

■

En el ejemplo

$$L_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \quad L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix},$$

$$P_1 = P_{12}, \quad P_2 = P_{23}, \quad P_3 = P_{34}$$

$$\therefore \quad \tilde{L}_1 = P_3 P_2 L_1 P_2 P_3$$

$$P_2 L_1 P_2 = P_{23} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} = L_1'$$

$$P_3 L_1' P_3 = P_{34} \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \tilde{L}_1$$

$$\tilde{L}_2 = P_3 L_2 P_3$$

$$P_3 L_2 P_3 = P_{34} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \tilde{L}_2$$

$$\tilde{L}_3 = I_4 .$$

$$\therefore \quad L = \tilde{L}_1 \tilde{L}_2$$

$$\therefore \quad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\therefore L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Ahora, el tener la factorización $PA=LU$ de una matriz A de orden m nos es útil para resolver el sistema de ecuaciones

$$Ax=b$$

cuando $N(A)=\{0\}$, ya que en este caso si $PA=LU$ entonces $N(L)=N(U)=\{0\}$ y

$$\begin{aligned} Ax=b &\iff PAx=Pb \\ &\iff LUx=Pb \end{aligned}$$

Es decir, el sistema $Ax=b$ es equivalente al sistema

$$LUx=Pb$$

Si Hacemos $Ux=y$ entonces

$$LUx=Pb \iff \begin{cases} Ly=Pb \\ Ux=y \end{cases}$$

Estos dos sistemas son muy fáciles de resolver, ya que son triangulares.

Para el caso en que $N(A)\neq\{0\}$, utilizaremos la técnica para llevar a la matriz A a la forma escalonada

Ejemplo

Resolver el sistema $Ax=b$ donde

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix}, \quad b = \begin{bmatrix} 4 \\ 10 \\ 12 \\ 5 \end{bmatrix}$$

Por el ejemplo anterior $PA = LU$ donde

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Entonces tenemos que

$$Pb = \begin{bmatrix} 10 \\ 12 \\ 5 \\ 4 \end{bmatrix}$$

Primero tenemos que resolver el sistema

$$Ly = Pb$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \eta_3 \\ \eta_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 12 \\ 5 \\ 4 \end{bmatrix}$$

$$\Rightarrow \quad \begin{array}{ll} \eta_1 = 10 & \\ \eta_1 + \eta_2 = 12 & \therefore \quad \eta_2 = 2 \\ \eta_1 + \eta_2 + \eta_3 = 5 & \therefore \quad \eta_3 = 5 - 12 = -7 \\ \eta_4 = 4 & \end{array}$$

$$\therefore \quad y^t = [10 \quad 2 \quad -7 \quad 4]$$

Ahora hay que resolver el sistema $Ux = y$.

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ -7 \\ 4 \end{bmatrix}$$

$$\Rightarrow \quad \begin{array}{ll} \xi_4 = 4 & \\ -\xi_3 - \xi_4 = -7 & \Rightarrow \xi_3 = 7 - 4 = 3 \\ \xi_2 = 2 & \\ \xi_1 + \xi_2 + \xi_3 + \xi_4 = 10 & \\ \Rightarrow \xi_1 + 9 = 10 & \Rightarrow \xi_1 = 1 \end{array}$$

Por lo tanto la solución del sistema es

$$X = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

Supongamos que queremos resolver varios sistemas de ecuaciones

$$(*) \qquad \begin{array}{c} Ax = b_1 \\ Ax = b_2 \\ \vdots \\ Ax = b_k \end{array}$$

donde A es una matriz $m \times n$ y $b_i \in \mathbb{R}^m$,

y que nos interesa saber cuándo los k sistemas tienen solución.

Sea B la matriz $m \times k$ dada por

$$B = [b_1 | b_2 | \cdots | b_k]$$

Entonces es claro que los k sistemas tienen solución $\Longleftrightarrow$ $r([A|B]) = r(A)$.

Por lo tanto para resolver el sistema $(*)$ reducimos a la forma escalonada a la matriz $[A|B]$.

Notemos que si $(*)$ tiene solución, entonces existen $x_1, x_2, \dots, x_k \in \mathbb{R}^n$ tal que

$$
\begin{aligned}
Ax_1 &= b_1 \\
Ax_2 &= b_2 \\
&\vdots \\
Ax_k &= b_k
\end{aligned}
$$

es decir:

$$[Ax_1 \,|\, Ax_2 \,|\, \cdots \,|\, Ax_k] = [b_1 \,|\, b_2 \,|\, \cdots \,|\, b_k]$$

y por lo tanto existe una matriz $X_{n\times k}$ tal que

$$AX = B$$

es decir $R(B) \leq R(A)$

E inversamente, si existe una $X_{n\times k}$ tal que $AX = B \Rightarrow Ax_i = b_i \quad 1 \leq i \leq k$ $\Rightarrow$ el sistema $(*)$ tiene solución.

Es decir, el sistema (*) tiene solución

$\Leftrightarrow r([A|B]) = r(A)$

$\Leftrightarrow$ Existe $X_{n\times k}$ tal que $AX = B$.

Hasta ahora nos hemos ocupado en desarrollar las herramientas para resolver sistemas de ecuaciones del tipo

$$Ax = b$$

Supongamos que ahora nos interesa analizar

$$y^t A = c^t$$

Afortunadamente la teoría con que contamos nos sirve también para tratar este caso, ya que el sistema

$$y^t A = c^t$$

es equivalente al sistema

$$A^t y = c$$

Por lo tanto, para resolver el sistema (*) podemos aplicar nuestra técnica a la matriz

$$[A^t | c]$$

y todos los resultados que demostramos se aplican también a este caso.

En este capitulo hemos visto que el conjunto de soluciones del sistema

$$Ax - b$$

es una variedad lineal:

$$\{x \in \mathbb{R}^n : Ax = b\} = x_0 + N(A) = \mathcal{L}(b)$$

¿Será cierto que toda variedad lineal se puede obtener como el conjunto de soluciones de un sistema $Ax = b$?

Sea $\mathcal{L} = x_0 + \mathcal{S} \subseteq \mathbb{R}^n$ una variedad lineal. Nos interesa saber si siempre se pueden encontrar una matriz A $m \times n$ y un vector $b \in \mathbb{R}^m$ tal que

$$\mathcal{L} = \{x \in \mathbb{R}^n : Ax = b\} = \mathcal{L}(b)$$

En tal caso, tendriamos que

$$\mathcal{L} = x_0 + N(A) = x_0 + \mathcal{S}$$

Si encontramos una matriz A tal que $N(A) = \mathcal{S}$, y tomamos $b = Ax_0$, entonces tendriamos que, en efecto, $\mathcal{L} = \mathcal{L}(b)$.

Resolveremos el siguiente problema:

Dado $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$, encontrar una matriz A $m \times n$ tal que

$$\mathcal{S} = N(A).$$

Sea $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$, con $\dim(\mathcal{S}) = k < n$ ($\mathcal{S} = \mathbb{R}^n$ no nos interesa).

Sea $\{s_1, s_2, \ldots, s_k\}$ una base de $\mathcal{S}$ y sea $S = [s_1 | s_2 | \ldots | s_k]$, tenemos que encontrar una matriz $A_{m \times n}$ tal que

$$AS = 0$$

i.e:, tenemos que encontrar una base de $N(S^t)$.

Como $r(S) = r(S^t) = k$ entonces

$$\dim(N(S^t)) = n - k \quad :$$

$$N(S^t) = \{y \in \mathbb{R}^n : y^t S = 0\} \ ..$$

una base de $N(S^t)$ la podemos obtener resolviendo el sistema

$$S^t y = 0$$

sea $\{r_1, r_2, \ldots, r_{n-k}\} \subseteq \mathbb{R}^n$ una base de $N(S^t)$, y sean

$$A_1 = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_{n-k}^t \end{bmatrix} \qquad \text{y} \quad A_2 = \mathbb{Y} A_1$$

donde $\mathbb{Y}$ es cualquier matriz $[m-(n-k)] \times (n-k)$; es decir, A_2 consta de renglones que son combinaciones lineales de los de A_1.

$$\text{sea} \quad A = \begin{array}{c} n \\ \begin{bmatrix} A_1 \\ \\ A_2 \end{bmatrix} \end{array} \begin{array}{l} n-k \\ \\ m-(n-k) \end{array} \quad \text{entonces}$$

$$AS = \begin{array}{c} n \\ \begin{bmatrix} A_1 S \\ \\ \mathbb{Y} A_1 S \end{bmatrix} \end{array} \begin{array}{c} n-k \\ = \\ m-(n-k) \end{array} \begin{array}{c} n \\ \begin{bmatrix} 0 \\ \\ 0 \end{bmatrix} \end{array} = 0$$

Y por lo tanto $N(A) = R(S) = \mathcal{S}$, ya que, por construcción, como $r(A) = n-k \Rightarrow$
$\dim(N(A)) = n-(n-k) = k$
y como las columnas de S son k vectores independientes en $N(A) \Rightarrow$ forman una base de $N(A)$.

Observemos que cualquier matriz $\tilde{A}$ tal que $R(\tilde{A}^t) = R(A^t)$ es tal que

$$\tilde{A} S = 0$$

MATRICES INVERSAS

Dada $A_{m\times n}$ nos interesa saber bajo qué condiciones existen matrices $X_{n\times m}$ y $Y_{n\times m}$ tal que

$$AX = I_m$$
$$YA = I_n$$

RESULTADO

Sea A una matriz $m\times n$. Entonces son equivalentes:

(1) $r(A) = m \leq n$

(2) Los renglones de A son linealmente independientes.

(3) $N(A^t) = \{0^*\}$

(4) $R(A) = \mathbb{R}^m$

(5) $A : \mathbb{R}^n \to \mathbb{R}^m$ es sobre

(6) Existe una $X_{n\times m}$ tal que $AX = I_m$

Demostración:

(1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3) $\Leftrightarrow$ (4) obvio

(4) $\Rightarrow$ (5)

Si $R(A) = \mathbb{R}^m$ entonces dado $b \in \mathbb{R}^m$ existe $x \in \mathbb{R}^n$ tal que $Ax = b$ $\therefore$ $A : \mathbb{R}^n \to \mathbb{R}^m$ sobre.

(5) ⇒ (6)

Supongamos $A : \mathbb{R}^n \to \mathbb{R}^m$ sobre

$\Rightarrow \quad R(A) = \mathbb{R}^m$

$\Rightarrow \quad R(A) = R(I_m)$

$\Rightarrow$ Existe $\mathbb{X}_{n \times m}$ tal que $A\mathbb{X} = I_m$

(6) ⇒ (4)

Supongamos $A\mathbb{X} = I_m$ para $\mathbb{X}_{n \times m}$

$\Rightarrow \quad R(A) = R(I_m) = \mathbb{R}^m$

□

$$\begin{tikzcd}
\mathbb{R}^m \arrow[r, "\mathbb{X}"] & \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m
\end{tikzcd}$$

Observemos que $\mathbb{X}$ no es única, ya que, por ejemplo,

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \qquad \mathbb{X}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad \mathbb{X}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$$

son tales que $AX_1 = AX_2 = I_2$.

Es más, si $N(A) \neq \{0\}$, entonces sea $\{x_1^{(0)}, x_2^{(0)}, \ldots, x_{n-m}^{(0)}\}$ una base de $N(A)$ y sea $X_0 = [x_1^{(0)} | x_2^{(0)} | \cdots | x_{n-m}^{(0)}]$.

Si tenemos que $AX_1 = I_m$ donde $X_1 = [x_1^{(1)} | x_2^{(1)} | \cdots | x_m^{(1)}]$ sea $X_2 = [x_1^{(2)} | x_2^{(2)} | \cdots | x_m^{(2)}]$ donde

$$x_i^{(2)} = x_i^{(1)} + X_0 z_i \quad \text{para alguna } z_i \in \mathbb{R}^{n-m} \quad 1 \leq i \leq m$$

Es decir sea X_2 de la forma

$$X_2 = X_1 + X_0 Z \quad \text{para } Z_{(n-m) \times m}$$

Entonces

$$AX_2 = AX_1 + AX_0 Z = AX_1 = I_m$$

y por lo tanto

$$AX_1 = AX_2 = I_m.$$

RESULTADO

Sea A una matriz $m \times n$. Entonces son equivalentes:

(1) $r(A) = n \leq m$
(2) las columnas de A son independientes
(3) $N(A) = \{0\}$
(4) $R(A^t) = \mathbb{R}^{n*}$
(5) $A : \mathbb{R}^n \to \mathbb{R}^m$ es inyectiva
(6) Existe $Y_{n\times m}$ tal que $YA = I_n$

Demostración

$(1) \Leftrightarrow (2) \Leftrightarrow (3) \Leftrightarrow (4)$ obvio

$(4) \Rightarrow (6)$

$$R(A^t) = \mathbb{R}^{n*} \Rightarrow R(A^t) = R(I_n)$$
$$\Rightarrow \exists\ Y_{n\times m} \text{ tal que } YA = I_n$$

$(6) \Rightarrow (5)$

Supongamos $YA = I_n$ para alguna $Y_{n\times m}$

Entonces $Ax_1 = Ax_2 \Rightarrow YAx_1 = YAx_2 \Rightarrow I_n x_1 = I_n x_2$
$\Rightarrow x_1 = x_2$

$(5) \Rightarrow (3)$

$N(A) \neq \{0\} \Rightarrow Ax = 0$ para alguna $x \neq 0$
$\therefore\ Ax = A0 \quad \therefore\ A$ no inyectiva

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \arrow[r, "Y"] & \mathbb{R}^n \end{tikzcd}$$

Observemos que puede haber más de una matriz Y que tenga esta propiedad.

Si $N(A^t) \neq 0$, sea $\{y_1^t, y_2^t, \ldots, y_{m-n}^t\}$ una base de $R(A^t)$ y sea

$$Y_0 = \begin{bmatrix} y_1^t \\ y_2^t \\ \vdots \\ y_{m-n}^t \end{bmatrix}$$

Si $Y_1 A = I_n$ sea Y_2 tal que

$$\underset{n\times m}{Y_2} = Y_1 + \underset{n\times(m-n)}{Z}\ \underset{(m-n)\times m}{Y_0} \qquad \text{donde } Z \text{ es } n\times(n-m)$$

entonces $Y_2 A = Y_1 A + Z Y_0 A = Y_1 A = I_n$.

Si existe X tal que $AX = I_m$ decimos que X es una INVERSA POR LA DERECHA de A.

Si existe $X_{n \times m}$ tal que $XA = I_n$ entonces decimos que X es una INVERSA POR LA IZQUIERDA de A.

Para que una matriz pueda tener tanto inversa por la derecha como inversa por la izquierda es necesario que A sea una matriz cuadrada. Analizaremos el caso cuando A es cuadrada.

RESULTADO

Sea A una matriz de orden m. Supongamos que $AX = I_m$ para alguna matriz $X_{m \times m}$. Entonces X es única.

Demostración:

Supongamos $AX_1 = I_m \Rightarrow r(A) = m$

$\Rightarrow N(A) = \{0\}$

Si $AX_2 = I_m$ entonces

$$
\begin{aligned}
& AX_1 = AX_2 \\
\Rightarrow\ & A(X_1 - X_2) = 0 \\
\Rightarrow\ & X_1 - X_2 = 0 \qquad (N(A) = \{0\}) \\
\Rightarrow\ & X_1 = X_2
\end{aligned}
$$

RESULTADO

Sea A una matriz de orden m. Supongamos que existe alguna $X_{m\times m}$ tal que $XA = I_m$. Entonces X es única.

Demostración:

Supongamos $X_1 A = I_m \Rightarrow r(A) = m$

$\Rightarrow N(A^t) = \{0^t\}$

Si $X_2 A = I_m$ entonces

$$X_1 A = X_2 A$$
$$\Rightarrow (X_1 - X_2) A = 0$$
$$\Rightarrow X_1 - X_2 = 0 \qquad (N(A^t) = \{0^t\})$$
$$\Rightarrow X_1 = X_2$$

■

Por lo tanto tenemos que para matrices cuadradas de rango máximo entonces existe una única $X_{m\times m}$ tal que

$$AX = I_m$$

y existe una única $X_{m\times m}$ tal que

$$XA = I_m$$

¿Existe alguna relación entre estas dos matrices?

La respuesta a esta pregunta es afirmativa, ya que

$$
\begin{aligned}
AX = I_m &\Rightarrow AXA = A \\
&\Rightarrow A(XA - I) = 0 \\
&\Rightarrow XA - I = 0 \qquad (N(A) = \{0\}) \\
&\Rightarrow XA = I_m
\end{aligned}
$$

O sea que para matrices cuadradas tenemos que

(1) A tiene inversa por la derecha ⟺ A tiene inversa por la izquierda

(2) la inversa por la derecha de A es igual a la inversa por la izquierda de A.

Si existe $X_{m \times m}$ tal que $AX = I_m = XA$, entonces a X la llamamos INVERSA de A, y la denotaremos por A^{-1}.

RESULTADO

Sea A una matriz de orden m. Entonces son equivalentes

1) $r(A) = m$
2) $N(A) = \{0\}$
3) $R(A) = \mathbb{R}^m$
4) A tiene columnas independientes
5) A tiene renglones independientes

6) $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ es sobre

7) $A : \mathbb{R}^m \rightarrow \mathbb{R}^m$ es inyectiva

8) Existe una matriz A^{-1} de orden m tal que $AA^{-1} = A^{-1}A = I_m$

DEFINICION

Sea A una matriz cuadrada. Decimos que A es NO SINGULAR ó INVERTIBLE si A posee matriz inversa; en caso contrario decimos que A es SINGULAR.

PROPIEDADES

1) Si A es invertible entonces A^{-1} es única.

2) Si A es invertible entonces A^{-1} es invertible y $(A^{-1})^{-1} = A$

Demostración : $A(A^{-1}) = I_m \therefore (A^{-1})^{-1} = A$ ▪

3) Si A es no-singular entonces A^t es no-singular y

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

Demostración :

$$
\begin{aligned}
A^t (A^{-1})^t &= (A)^t (A^{-1})^t \\
&= ((A^{-1})(A))^t \\
&= I^t \\
&= I
\end{aligned}
$$

a la matriz $(A^{-1})^t$ la denotaremos por A^{-t}.

4) Si L es una matriz triangular inferior no-singular entonces L^{-1} es triangular inferior.

Demostración:

Sea B la matriz inversa de L. Entonces para $1 \leq i \leq m-1$ podemos dividir a L y B por bloques de la siguiente forma.

$$
L = \begin{array}{c|c|c|c}
 & i & & \\
i & L_1^{(i)} & 0 & \\
\hline
 & M^{(i)} & L_2^{(i)} & m-i \\
 & & m-i &
\end{array}
\qquad
B = \begin{array}{c|c|c|c}
 & i & & \\
i & B_{11}^{(i)} & B_{12}^{(i)} & \\
\hline
 & B_{21}^{(i)} & B_{22}^{(i)} & m-i \\
 & & m-i &
\end{array}
$$

donde $N(L_1^{(i)}) = \{0\}$.

Entonces, para $1 \leq i \leq m-1$ debemos de tener $B_{12}^{(i)} = 0$, ya que

$$LB = \left[\begin{array}{c|c} L_1^{(i)} & 0 \\ \hline M^{(i)} & L_2^{(i)} \end{array}\right] \left[\begin{array}{c|c} B_{11}^{(i)} & B_{12}^{(i)} \\ \hline B_{21}^{(i)} & B_{22}^{(i)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} L_1^{(i)} B_{11}^{(i)} & L_1^{(i)} B_{12}^{(i)} \\ \hline M^{(i)} B_{11}^{(i)} + L_2^{(i)} B_{21}^{(i)} & M^{(i)} B_{12}^{(i)} + L_2^{(i)} B_{22}^{(i)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} I_i & \\ \hline & I_{m-i} \end{array}\right]$$

$$\Rightarrow \quad L_1^{(i)} B_{12}^{(i)} = 0$$

$$\Rightarrow \quad B_{12}^{(i)} = 0 \qquad (\; N(L_1^{(i)}) = \{0\})$$

Por lo tanto para $1 \leq i \leq m-1$ $B_{12}^{(i)} = 0$

$\therefore$ B es triangular inferior.

5) Si U es una matriz triangular superior no-singular, entonces U^{-1} es triangular superior.

La demostración es análoga a la de matrices triangulares inferiores.

6) Sean A y B de orden m. Entonces

AB es no-singular $\Leftrightarrow$ A es no-singular y
B es no-singular

y $(AB)^{-1} = B^{-1}A^{-1}$

Demostración :

$\Leftarrow$ Si A y B son invertibles entonces

$$(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = Im$$

Por lo tanto AB es invertible y

$$(AB)^{-1} = B^{-1}A^{-1}$$

$\Rightarrow$ Supongamos AB es no-singular.
Entonces

$$A[B(AB)^{-1}] = (AB)(AB)^{-1} = Im$$

$\therefore$ A es invertible y $A^{-1} = B(AB)^{-1}$

analogamente

$$[(AB)^{-1}A]B = (AB)^{-1}(AB) = Im$$

$\therefore$ B es invertible y $B^{-1} = (AB)^{-1}A$

7) Sea $A = \begin{array}{ccc} & s & \\ s & \left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right] & m-s \\ & m-s & \end{array}$. Entonces

A no-singular $\iff$ A_1 es no-singular y
A_2 es no-singular

y $A^{-1} = \begin{array}{ccc} & s & \\ s & \left[\begin{array}{c|c} A_1^{-1} & 0 \\ \hline 0 & A_2^{-1} \end{array}\right] & m-s \\ & m-s & \end{array}$

Demostración :

$\Leftarrow$ Si A_1 y A_2 son invertibles entonces

$$\begin{array}{ccc} & s & \\ & \left[\begin{array}{c|c} A_1^{-1} & 0 \\ \hline 0 & A_2^{-1} \end{array}\right] & \\ m-s & m-s & \end{array} \begin{array}{ccc} & s & \\ s & \left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right] & m-s \\ & m-s & \end{array} = \begin{array}{cc} & s \\ s & \left[\begin{array}{c|c} I_s & 0 \\ \hline 0 & I_{m-s} \end{array}\right] \end{array}$$

$\therefore$ A invertible y $A^{-1} = \begin{array}{ccc} & s & \\ s & \left[\begin{array}{c|c} A_1^{-1} & 0 \\ \hline 0 & A_2^{-1} \end{array}\right] & m-s \\ & m-s & \end{array}$

$\Rightarrow$ Supongamos A es invertible.

Sea X la inversa de A .

Por lo tanto

$$\left[\begin{array}{c|c} A_1 & 0 \\ \hline 0 & A_2 \end{array}\right] \left[\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right] = I_m$$

$$\Rightarrow \quad A_1 X_{11} = I_s \quad \Rightarrow A_1 \text{ invertible}$$
$$A_2 X_{22} = I_{m-s} \Rightarrow A_2 \text{ invertible}$$

8) Sea $A = \begin{array}{cc} & \begin{array}{cc} s & m-s \end{array} \\ \begin{array}{c} s \\ m-s \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \end{array}$. Entonces

A es invertible $\Longleftrightarrow$ A_{11} es invertible y A_{22} es invertible

Demostración:

$\Leftarrow$ Supongamos A_{11} y A_{22} son invertibles

$\therefore \quad R(A_{11}) = \mathbb{R}^s \quad \therefore \quad R(A_{12}) \subseteq R(A_{11})$

Sea $A_{12} A_{22}^{-1} = \tilde{A}_{12}$

por lo tanto $R(\tilde{A}_{12}) \subseteq R(A_{12}) \subseteq R(A_{11})$

$\therefore$ existe una $X_{s \times (m-s)}$ tal que

$$A_{11}\,X = -\underset{s\times(m-s)}{\tilde{A}_{12}} = -A_{12}A_{22}^{-1}$$

y por lo tanto

$$\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \left[\begin{array}{c|c} A_{11}^{-1} & X \\ \hline 0 & A_{22}^{-1} \end{array}\right] = \left[\begin{array}{c|c} A_{11}A_{11}^{-1} & A_{11}X + A_{12}A_{22}^{-1} \\ \hline 0 & A_{22}A_{22}^{-1} \end{array}\right]$$

$$\left[\begin{array}{c|c} I_s & -\tilde{A}_{12} + \tilde{A}_{12} \\ \hline 0 & I_{m-s} \end{array}\right] = \left[\begin{array}{c|c} I_s & 0 \\ \hline 0 & I_{m-s} \end{array}\right]$$

$\therefore$ A es invertible.

$(\Rightarrow)$ Supongamos A invertible $\Rightarrow$ existe una matriz X tal que $AX = I_m$

$$\therefore \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline 0 & A_{22} \end{array}\right] \left[\begin{array}{c|c} X_{11} & X_{12} \\ \hline X_{21} & X_{22} \end{array}\right] = I_m$$

$$\Rightarrow \quad \begin{aligned} A_{22}X_{21} &= 0 \\ A_{22}X_{22} &= I_{m-s} \\ A_{11}X_{11} + A_{12}X_{21} &= I_s \end{aligned}$$

$$A_{22} X_{22} = I_{m-s} \Rightarrow A_{22} \text{ invertible}$$

$$\left.\begin{array}{l} A_{22} X_{21} = 0 \\ A_{22} \text{ invertible} \end{array}\right\} \Rightarrow X_{21} = 0$$

$$A_{11} X_{11} + A_{12} \cdot 0 = I_s \Rightarrow A_{11} X_{11} = I_s$$
$$\Rightarrow A_{11} \text{ invertible}$$

¿ Qué podemos decir cuando A es de la forma

$$A = \begin{array}{ccc} & s & \\ s & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & m-s \\ & m-s & \end{array} \quad ?$$

Primero analizaremos el caso cuando

$$A = \begin{array}{ccc} & m-1 & \\ m-1 & \left[\begin{array}{c|c} A_1 & a \\ \hline y^* & \alpha \end{array}\right] & 1 \\ & 1 & \end{array}$$

Observemos que $\alpha \neq 0$ y A_1 invertible $\Rightarrow$ A es invertible.

Por ejemplo :

$$A = \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ \hline 1 & 0 & 0 & 1 \end{array}\right] \quad \text{es singular}$$

ya que $a_4 = a_1 + a_2$.

$$\text{sea} \quad A = \begin{array}{c} \\ m-1 \\ \\ \end{array}\overset{m-1}{\left[\begin{array}{c|c} A_1 & a \\ \hline \gamma^t & \underset{1}{\alpha} \end{array}\right]}, \quad \text{con } A_1 \text{ invertible}$$

Veremos qué condiciones se deben tener para que A sea invertible.

Supongamos que A es invertible, y sea $A^{-1} = B$

$$\left[\begin{array}{c|c} A_1 & a \\ \hline \gamma^t & \alpha \end{array}\right] \left[\begin{array}{c|c} B_1 & b \\ \hline s^t & \beta \end{array}\right] = \left[\begin{array}{c|c} I_{m-1} & 0 \\ \hline 0 & 1 \end{array}\right]$$

$\Rightarrow$

$$\begin{aligned} A_1 B_1 + a s^t &= I_{m-1} \\ \gamma^t B_1 + \alpha s^t &= 0 \\ A_1 b + a\beta &= 0 \\ \gamma^t b + \alpha\beta &= 1 \end{aligned}$$

como A_1 es invertible $\Rightarrow$

$$A_1^{-1}A_1 b + A_1^{-1} a\beta = 0$$

$$\Rightarrow \quad b + A_1^{-1} a\beta = 0$$

$$\Rightarrow \quad \gamma^t b + \gamma^t A_1^{-1} a\beta = 0$$

pero tenemos que

$$\gamma^t b = 1 - \alpha\beta$$

$$\therefore \quad 1 - \alpha\beta + \gamma^t A_1^{-1} a\beta = 0$$

$$\therefore \quad \alpha\beta - \gamma^t A_1^{-1} a\beta = 1$$

$$\Rightarrow \quad \beta(\alpha - \gamma^t A_1^{-1} a) = 1$$

$$\Rightarrow \quad \alpha - \gamma^t A_1^{-1} a \neq 0$$

$\therefore$ A invertible $\Rightarrow$ $\alpha - \gamma^t A_1^{-1} a \neq 0$

Veremos también que es suficiente que $\alpha - \gamma^t A_1^{-1} a \neq 0$ para que A sea no-singular

supongamos $\alpha - \gamma^t A_1^{-1} a \neq 0$

Tenemos que encontrar una matriz B tal que $AB = I$

Sean $\beta = (\alpha - \gamma^t A_1^{-1} a)^{-1}$

$b = -A_1^{-1} a \beta$

La submatriz B_1 y el renglón s^t tienen que satisfacer

$$A_1 B_1 + a s^t = I_{m-1}$$
$$\gamma^t B_1 + \alpha s^t = 0$$

$$B_1 + A_1^{-1} a s^t = A_1^{-1}$$
$$\gamma^t B_1 + \alpha s^t = 0$$

$$\gamma^t B_1 + \gamma^t A_1^{-1} a s^t = \gamma^t A_1^{-1}$$
$$\gamma^t B_1 + \alpha s^t = 0$$

$$\therefore \quad (\alpha - \gamma^t A_1^{-1} a) s^t = -\gamma^t A_1^{-1}$$

y como $(\alpha - \gamma^t A_1^{-1} a) \neq 0$ podemos despejar s^t

$$\therefore \quad s^t = -(\alpha - \gamma^t A_1^{-1} a)^{-1} \gamma^t A_1^{-1}$$

y $\therefore \quad B_1 = A_1^{-1} - A_1^{-1} a s^t$

$$= A_1^{-1}(I - a s^t)$$
$$= A_1^{-1}(I + a(\alpha - \gamma^t A_1^{-1} a)^{-1} \gamma^t A_1^{-1})$$

$$\therefore \text{ si } \quad B = \left[\begin{array}{c|c} B_1 & b \\ \hline s^t & \beta \end{array}\right] \qquad \text{donde}$$

$$\beta = (\alpha - \gamma^t A_1^{-1} a)^{-1}$$

$$b = -A_1^{-1} a\,(\alpha - \gamma^t A_1^{-1} a)^{-1}$$

$$s^t = -(\alpha - \gamma^t A_1^{-1} a)^{-1} \gamma^t A_1^{-1}$$

$$B_1 = A_1^{-1}\left(I + a\cdot(\alpha - \gamma^t A_1^{-1} a)^{-1} \gamma^t A_1^{-1}\right)$$

es tal que

$$AB = I$$

$$\therefore \ (\alpha - \gamma^t A_1^{-1} a) \neq 0 \Rightarrow \exists\ B \text{ tal que } AB = I_n$$

$$\therefore \quad A \text{ invertible} \Leftrightarrow (\alpha - \gamma^t A_1^{-1} a) \neq 0$$

Hemos demostrado el siguiente

RESULTADO

Sea $$A = \begin{array}{cc} & \begin{array}{cc} n-1 & \quad \end{array} \\ n-1 & \left[\begin{array}{c|c} A_1 & a \\ \hline \gamma^t & \alpha \end{array}\right]_1 \\ & \qquad 1 \end{array},$$ con A_1 no-singular

Entonces

A es invertible $\Longleftrightarrow$ $\alpha - r^t A_1^{-1} a \neq 0$

. . . .

En el ejemplo anterior

$$\alpha - r^t A_1^{-1} a = 1 - \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 1 - 1 = 0$$

y $\therefore$ la matriz del ejemplo es singular.

Para matrices de la forma $A = \begin{array}{c@{}c} & \begin{array}{c|c} s & m-s \end{array} \\ \begin{array}{c} s \\ m-s \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$

con A_{11} invertible tenemos :

RESULTADO

Sea $A = \begin{array}{c@{}c} & \begin{array}{c|c} s & m-s \end{array} \\ \begin{array}{c} s \\ m-s \end{array} & \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \end{array}$ con A_{11} invertible.

Entonces

A es invertible $\Longleftrightarrow$ $(A_{22} - A_{21} A_{11}^{-1} A_{12})$ es invertible

Demostración :

$\Rightarrow$ Supongamos que A es invertible. Sea $A^{-1} = B$

$$\left[\begin{array}{c|c} \overset{s}{A_{11}} & A_{12} \\ \hline A_{21} & \underset{m-s}{A_{22}} \end{array}\right]\begin{matrix} s \\ m-s \end{matrix}\left[\begin{array}{c|c} \overset{s}{B_{11}} & B_{12} \\ \hline B_{21} & \underset{m-s}{B_{22}} \end{array}\right] = \left[\begin{array}{c|c} I_s & 0 \\ \hline 0 & I_{m-s} \end{array}\right]$$

$$\Rightarrow \left\{\begin{array}{l} A_{11}B_{11} + A_{12}B_{21} = I_s \\ A_{21}B_{11} + A_{22}B_{21} = 0 \\ \\ A_{11}B_{12} + A_{12}B_{22} = 0 \\ A_{21}B_{12} + A_{22}B_{22} = I_{m-s} \end{array}\right. \qquad (*)$$

$$\Rightarrow \begin{array}{l} B_{12} + A_{11}^{-1}A_{12}B_{22} = 0 \\ A_{21}B_{12} + A_{22}B_{22} = I_{m-s} \end{array}$$

$$\Rightarrow A_{21}(-A_{11}^{-1}A_{12}B_{22}) + A_{22}B_{22} = I_{m-s}$$

$$\Rightarrow (A_{22} - A_{21}A_{11}^{-1}A_{12})B_{22} = I_{m-s}$$

$\Rightarrow$ $A_{22} - A_{21}A_{11}^{-1}A_{12}$ es invertible

$\Leftarrow$ Supongamos que $A_{22} - A_{21}A_{11}^{-1}A_{12}$ es invertible. Tenemos que encontrar una B

$$B = \begin{matrix} s \\ \\ \end{matrix}\left[\begin{array}{c|c} \overset{s}{B_{11}} & B_{12} \\ \hline B_{21} & \underset{m-s}{B_{22}} \end{array}\right]\begin{matrix} \\ m-s \end{matrix} \qquad \text{tal que } AB = I_m$$

Para que tal B exista, el sistema (*) tiene que tener solución.

$$\begin{aligned} A_{11}B_{11} + A_{12}B_{21} &= I_s \\ A_{21}B_{11} + A_{22}B_{21} &= 0 \end{aligned}$$

$$\begin{aligned} B_{11} + A_{11}^{-1}A_{12}B_{21} &= A_{11}^{-1} \\ A_{21}B_{11} + A_{22}B_{21} &= 0 \end{aligned}$$

$$\Rightarrow \quad A_{21}(-A_{11}^{-1}A_{12}B_{21} + A_{11}^{-1}) + A_{22}B_{21} = 0$$

$$\Rightarrow \quad (A_{22} - A_{21}A_{11}^{-1}A_{12})B_{21} = -A_{11}^{-1}$$

como $A_{22}-A_{21}A_{11}^{-1}A_{12}$ es invertible, podemos despejar a B_{21}:

$$B_{21} = -(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1}A_{11}^{-1}$$

$$\therefore \quad B_{21} = -\left[A_{11}(A_{22}-A_{21}A_{11}^{-1}A_{12})\right]^{-1}$$

$$\begin{aligned} y \quad \therefore \quad B_{11} &= A_{11}^{-1} - A_{11}^{-1}A_{12}B_{21} \\ &= A_{11}^{-1}(I - A_{12}B_{21}) \end{aligned}$$

$$\therefore \quad B_{11} = A_{11}^{-1}\left[I - A_{12}\left(A_{11}[A_{22}-A_{21}A_{11}^{-1}A_{12}]\right)^{-1}\right]$$

$$\begin{aligned} \therefore \quad si \quad B_{22} &= (A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \\ B_{12} &= -A_{11}^{-1}A_{12}(A_{22}-A_{21}A_{11}^{-1}A_{12})^{-1} \end{aligned}$$

y

$$B_{21} = -\left[A_{11}(A_{22} - A_{21}A_{11}^{-1}A_{12})\right]^{-1}$$

$$B_{11} = A_{11}^{-1}\left[I - A_{12}\left[A_{11}(A_{22} - A_{21}A_{11}^{-1}A_{12})\right]^{-1}\right]$$

entonces B_{11}, B_{12}, B_{21} y B_{22} satisfacen el sistema $(*)$ y por lo tanto

$$B = \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right]$$

es tal que $AB = I$ y $\therefore$ A es invertible

$\therefore$ $A_{22} - A_{21}A_{11}^{-1}A_{12}$ invertible $\rightarrow$ A invertible

$\therefore$ A invertible $\Longleftrightarrow$ $A_{22} - A_{21}A_{11}^{-1}A_{12}$ invertible

$\square$

OBSERVACION

la suma de dos matrices no-singulares no necesariamente es invertible. Por ejemplo, la matriz I es invertible, pero $I - I = 0$ no lo es.

Nos interesa en particular saber cuando es invertible una matriz de la forma

$A + xy^t$ donde A es invertible.

Veremos qué pasa cuando $A + xy^t$ es singular:

supongamos que existe una $a \neq 0$ tal que

$$(A + xy^t)a = 0$$

$$\Rightarrow \quad Aa = -x\,(y^t a)$$

A no-singular $\Rightarrow$ $A\tilde{a} = -x$ para alguna $\tilde{a}$

$$\therefore \quad \tilde{a} = -A^{-1}x$$

$$\therefore \quad a = -A^{-1}x\,(y^t a) = \alpha\tilde{a}$$

$$\therefore \quad Aa = A(\alpha\tilde{a}) = -x\,(y^t a) = -x\,(y^t \alpha\tilde{a})$$

$$\therefore \quad \alpha A\hat{a} = -x\cdot\alpha = -x\,(y^t\alpha\,(-A^{-1}x))$$

$$\therefore \quad \alpha x\,(1 + y^t A^{-1}x) = 0$$

Como $\alpha \neq 0$ $(a = \alpha\tilde{a} \neq 0)$ y $x \neq 0$

$$\Rightarrow \quad 1 + y^t A^{-1}x = 0$$

$\therefore$ $A + xy^t$ singular $\Rightarrow$ $1 + y^t A^{-1}x = 0$

Ahora, supongamos $1 + y^t A^{-1}x = 0$. Entonces

$$(A + xy^t)A^{-1}x = x + x\,(y^t A^{-1}x) = x - x = 0 \quad .$$

Como $x \neq 0$ y A^{-1} es invertible $\Rightarrow$ $A^{-1}x \neq 0$

Por lo tanto si $1 + y^t A^{-1} x = 0$ entonces $(A + xy^t)$ es singular.

Por lo tanto

$$A + xy^t \text{ es singular} \iff 1 + y^t A^{-1} x = 0$$

Ahora, supongamos $1 + y^t A^{-1} x \neq 0$, trataremos de encontrar la matriz inversa de $A + xy^t$.

Supongamos

$$(A + xy^t)(A^{-1} - x_1 y_1^t) = I$$

$$\Rightarrow \quad I - A x_1 y_1^t + xy^t A^{-1} - xy^t x_1 y_1^t = I$$

$$\Rightarrow \quad A x_1 y_1^t - xy^t A^{-1} + xy^t x y_1^t = 0$$

sea $\quad A x_1 = x$

$$\therefore \quad x(y_1^t - y^t A^{-1} + y^t x_1 y_1^t) = 0$$

$$\Rightarrow x(y^t A^{-1} - (1 + y^t x_1) y_1^t) = 0$$

$$\Rightarrow x(y^t A^{-1} - (1 + y^t A^{-1} x) y_1^t) = 0$$

Este producto es de la forma ab^t

$$ab^t = 0 \Rightarrow a=0 \text{ ó } b=0$$

$\therefore$ como $x \neq 0 \Rightarrow$

$$y^t A^{-1} - (1+y^t A^{-1}x) y_1^t = 0$$

$$\Rightarrow \quad y_1^t = \frac{y^t A^{-1}}{1+y^t A^{-1}x} \qquad (1+y^t A^{-1}x \neq 0)$$

$$\therefore \quad A^{-1} - x_1 y_1^t = A^{-1} - (A^{-1}x) \frac{y^t A^{-1}}{1+y^t A^{-1}x}$$

$$\therefore \quad A^{-1} - x_1 y_1^t = A^{-1} - \frac{A^{-1}xy^t A^{-1}}{1+y^t A^{-1}x}$$

$\therefore$ La inversa de $A+xy^t$ es

$$A^{-1} - \frac{A^{-1}xy^t A^{-1}}{1+y^t A^{-1}x}$$

Hemos demostrado :

RESULTADO

Sea A no-singular. Entonces

$(A+xy^t)$ es no-singular $\Leftrightarrow$ $1+y^t A^{-1}x \neq 0$, y en el caso de que $A+xy^t$ sea invertible

$$(A+xy^t)^{-1} = A^{-1} - \frac{A^{-1}xy^t A^{-1}}{1+y^t A^{-1}x}$$

SUMA DE SUBESPACIOS

Supongamos que tenemos una matriz $A_{m \times n}$

$$A = [A_1 \mid A_2]$$

donde $A_1 = [a_1^{(1)} \mid a_2^{(1)} \mid \cdots \mid a_{n_1}^{(1)}]$ es $m \times n_1$

$A_2 = [a_1^{(2)} \mid a_2^{(2)} \mid \cdots \mid a_{n_2}^{(2)}]$ es $m \times n_2$

y $n = n_1 + n_2$.

Entonces

$$b \in R(A) \iff b = Ax \qquad x \in IR^n$$

$$\iff b = [A_1 \mid A_2] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{matrix} n_1 \\ n_2 \end{matrix}$$

$$\iff b = A_1 x_1 + A_2 x_2$$

$$\iff b = b_1 + b_2 \quad \text{con} \quad \begin{matrix} b_1 \in R(A_1) \\ b_2 \in R(A_2) \end{matrix}$$

Si $\mathcal{S} = R(A)$, $\mathcal{S}_1 = R(A_1)$ y $\mathcal{S}_2 = R(A_2)$ entonces tenemos que

$$s \in \mathcal{S} \iff s = s_1 + s_2 \quad , \quad \begin{matrix} s_1 \in \mathcal{S}_1 \\ s_2 \in \mathcal{S}_2 \end{matrix}$$

DEFINICION

Sean $\mathcal{S}_1, \mathcal{S}_2 \leq \mathbb{R}^m$ subespacios de $\mathbb{R}^m$. La SUMA de $\mathcal{S}_1$ y $\mathcal{S}_2$, denotada por $\mathcal{S}_1 + \mathcal{S}_2$ ó por $[\mathcal{S}_1, \mathcal{S}_2]$ es

$$\mathcal{S}_1 + \mathcal{S}_2 = \{ x \in \mathbb{R}^m : x = s_1 + s_2 \; ; \; s_1 \in \mathcal{S}_1 ; s_2 \in \mathcal{S}_2 \}$$

Es fácil ver que $\mathcal{S}_1 + \mathcal{S}_2$ es un subespacio de $\mathbb{R}^m$.

OBSERVACION :

La representación $x = s_1 + s_2$ de un vector $x \in \mathcal{S}_1 + \mathcal{S}_2$ no tiene porqué ser única. Por ejemplo, si

$$A_1 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$\mathcal{S}_1 = R(A_1)$ $\qquad \mathcal{S}_2 = R(A_2)$ $\qquad$ entonces

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$$= s_1 + s_2 = s_1' + s_2'$$

PROPIEDADES

Sean $\mathcal{S}_1, \mathcal{S}_2$: subespacios de $\mathbb{R}^m$. Entonces

1) $[\mathcal{S}_1, \mathcal{S}_2] = [\mathcal{S}_2, \mathcal{S}_1]$

2) $\mathcal{S}_1 \subseteq [\mathcal{S}_1, \mathcal{S}_2]$
$\mathcal{S}_2 \subseteq [\mathcal{S}_1, \mathcal{S}_2]$

3) $[\mathcal{S}_1, \mathcal{S}_2] \neq \mathcal{S}_1 \Leftrightarrow \mathcal{S}_2 \not\subset \mathcal{S}_1$
$[\mathcal{S}_1, \mathcal{S}_2] \neq \mathcal{S}_2 \Leftrightarrow \mathcal{S}_1 \not\subset \mathcal{S}_2$

Demostración de (3):

$[\mathcal{S}_1, \mathcal{S}_2] = \mathcal{S}_1 \Leftrightarrow \mathcal{S}_2 \subset \mathcal{S}_1$

$[\Rightarrow$ Supongamos $[\mathcal{S}_1, \mathcal{S}_2] = \mathcal{S}_1$

$y \in \mathcal{S}_2 \Rightarrow y \in [\mathcal{S}_1, \mathcal{S}_2] \Rightarrow y \in \mathcal{S}_1$

$\therefore \mathcal{S}_2 \subseteq \mathcal{S}_1$

$\Leftarrow]$ Supongamos $\mathcal{S}_2 \subset \mathcal{S}_1$

Sea $x \in [\mathcal{S}_1, \mathcal{S}_2] \quad \therefore \quad x = s_1 + s_2$ $\quad s_1 \in \mathcal{S}_1$
$s_2 \in \mathcal{S}_2 \subseteq \mathcal{S}_1$

$\therefore x \in \mathcal{S}_1 \quad \therefore \quad [\mathcal{S}_1, \mathcal{S}_2] \subseteq \mathcal{S}_1$

$\therefore [\mathcal{S}_1, \mathcal{S}_2] = \mathcal{S}_1$ ∎

Observar la analogía entre la suma de subespacios $[\mathcal{S}_1, \mathcal{S}_2]$ y el subespacio generado por dos vectores $[a_1, a_2]$.

El concepto de suma de dos subespacios se puede extender a n subespacios:

Si $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k$ son subespacios de $\mathbb{R}^m$, entonces

$$\mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_k = \left\{ x \in \mathbb{R}^m : x = \sum_{i=1}^{k} s_i \ , \ s_i \in \mathcal{S}_i \right\}.$$

Si tenemos que $\mathcal{S} = R(A)$, donde

$$A = [\overset{n_1}{A_1} \mid \overset{n_2}{A_2}]$$

entonces

$$\mathcal{S} = R(A_1) + R(A_2) = \mathcal{S}_1 + \mathcal{S}_2$$

observemos que

$$\dim(\mathcal{S}) = r(A) \leq r(A_1) + r(A_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2)$$

o sea que en general vamos a tener que

$$\dim(\mathcal{S}_1 + \mathcal{S}_2) \leq \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2)$$

¿ Cuándo tendremos que

$$\dim(\mathcal{S}_1+\mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) \ ?$$

RESULTADO

Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ subespacios de $\mathbb{R}^m$. Entonces

$$\dim(\mathcal{S}_1+\mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$$

Demostración:

Sean $\{a_1, a_2, \dots, a_k\}$ una base de $\mathcal{S}_1 \cap \mathcal{S}_2$

$\{a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_m\}$ una base de $\mathcal{S}_1$

y $\{a_1, a_2, \dots, a_k, c_1, c_2, \dots, c_n\}$ una base de $\mathcal{S}_2$

entonces

$\{a_1, a_2, \dots, a_k, b_1, b_2, \dots, b_m, c_1, c_2, \dots, c_n\}$ es una base de $\mathcal{S}_1 + \mathcal{S}_2$:

Claramente $\{a_1, \dots, a_k, b_1, \dots, b_m, c_1, \dots, c_n\}$ es un conjunto generador de $\mathcal{S}_1 + \mathcal{S}_2$, veremos que también es un conjunto linealmente independiente.

Sea $A = [a_1|a_2|\cdots|a_k]$

$B = [b_1|b_2|\cdots|b_m]$

$C = [c_1|c_2|\cdots|c_n]$

por lo tanto

$$\mathcal{S}_1 = R([A|B]) \quad ; \quad N([A|B]) = \{0\}$$

$$\mathcal{S}_2 = R([A|C]) \quad ; \quad N([A|C]) = \{0\}$$

$$\mathcal{S}_1 + \mathcal{S}_2 = R([A|B|C]) \quad ; \quad \mathcal{S}_1 \cap \mathcal{S}_2 = R(A)$$

Demostraremos que $N([A|B|C]) = \{0\}$

suponqamos $X = \begin{array}{c|l} \begin{bmatrix} x_1 \\ \hline x_2 \\ \hline x_3 \end{bmatrix} & \begin{array}{l} k \\ m \\ n \end{array} \end{array}$ es tal que

$$[A|B|C]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$$

$$\Rightarrow \quad Ax_1 + Bx_2 + Cx_3 = 0$$

$$\Rightarrow \quad -Cx_3 = [A|B]\begin{array}{c l} \begin{bmatrix} x_1 \\ \hline x_2 \end{bmatrix} & \begin{array}{l} k \\ m \end{array} \end{array}$$

$$\therefore \; -Cx_3 \in R([A|C]) \text{ y } -Cx_3 \in R([A|B])$$

$$\therefore \; -Cx_3 \in \mathcal{S}_1 \cap \mathcal{S}_2 = R(A)$$

$$\therefore \; -Cx_3 = A\tilde{x}_1$$

$$\Rightarrow [A|C]\begin{bmatrix}\tilde{x}_1 \\ x_3\end{bmatrix} = 0 \Rightarrow \tilde{x}_1 = 0 \text{ y } x_3 = 0$$

$$(N([A|C]) = \{0\})$$

$$\Rightarrow Ax_1 + Bx_2 + Cx_3 = Ax_1 + Bx_2 = 0$$

$$\Rightarrow [A|B]\begin{bmatrix}x_1 \\ x_2\end{bmatrix} = 0$$

$$\Rightarrow x_1 = 0, \; x_2 = 0 \qquad (N([A|B]) = \{0\})$$

$$\Rightarrow x = 0$$

$$\therefore \; N([A|B|C]) = \{0\}$$

$\therefore$ $\{a_1, \ldots, a_k, b_1, \ldots, b_m, c_1, \ldots, c_n\}$ es una base de $\mathcal{S}_1 + \mathcal{S}_2$.

Por lo tanto

$$\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) = (k+m) + (k+n)$$

$$= k + (k+m+n)$$

$$= \dim(\mathcal{S}_1 \cap \mathcal{S}_2) + \dim(\mathcal{S}_1 + \mathcal{S}_2)$$

$$\therefore \quad \dim(\mathcal{S}_1 + \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$$

∎

Daremos una segunda demostración de este resultado:

Si $\{a_1^{(1)}, a_2^{(1)}, \ldots, a_{n_1}^{(1)}\}$ es una base de $\mathcal{S}_1$

$\{a_1^{(2)}, a_2^{(2)}, \ldots, a_{n_2}^{(2)}\}$ es una base de $\mathcal{S}_2$ y

$A_1 = [a_1^{(1)} | a_2^{(1)} | \cdots | a_{n_1}^{(1)}]$, $\quad A_2 = [a_1^{(2)} | a_2^{(2)} | \cdots | a_{n_2}^{(2)}]$

entonces $\mathcal{S}_1 + \mathcal{S}_2 = R(A)$, donde $A = [A_1 | A_2]$

Por lo tanto

$$\dim(R(A)) = \dim(\mathcal{S}_1 + \mathcal{S}_2)$$

$\dim(N(A)) = \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$ (ésto ya lo probamos)

$$\begin{aligned}\therefore \dim(R(A)) + \dim(N(A)) &= n_1 + n_2 \\ &= \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) \\ &= \dim(\mathcal{S}_1 + \mathcal{S}_2) + \dim(\mathcal{S}_1 \cap \mathcal{S}_2)\end{aligned}$$

$$\therefore \quad \dim(\mathcal{S}_1 + \mathcal{S}_2) + \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2)$$

∎

DEFINICION

Sean $\mathcal{S}_1, \mathcal{S}_2$ subespacios de $\mathbb{R}^m$. Entonces

1) $\{\mathcal{S}_1, \mathcal{S}_2\}$ es INDEPENDIENTE si $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0\}$
2) $\{\mathcal{S}_1, \mathcal{S}_2\}$ es DEPENDIENTE si $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \{0\}$
3) $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$ es la SUMA DIRECTA de $\mathcal{S}_1$ y $\mathcal{S}_2$ si $\mathcal{S}_1 \cap \mathcal{S}_2 = \{0\}$, y la denotamos por $\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2$

Sean $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k$ subespacios de $\mathbb{R}^m$. Entonces

1) $\{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k\}$ es INDEPENDIENTE si $\forall i$, $1 \leq i \leq k$

$$\mathcal{S}_i \cap \{\mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_{i-1} + \mathcal{S}_{i+1} + \cdots + \mathcal{S}_k\} = \{0\}$$

2) $\{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k\}$ es DEPENDIENTE si existe alguna i, $1 \leq i \leq k$ tal que

$$\mathcal{S}_i \cap \{\mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_{i-1} + \mathcal{S}_{i+1} + \cdots + \mathcal{S}_k\} \neq \{0\}.$$

3) $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_k$ es la SUMA DIRECTA de $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k$ si $\{\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_k\}$ es independiente. En tal caso escribimos

$$\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_k$$

Observar que los conceptos de dependencia e independencia para subespacios son análogos a las nociones de dependencia e independencia para vectores.

RESULTADO

$\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_k \Leftrightarrow$ cada vector $s \in \mathcal{S}$ tiene una única representación $s = s_1 + s_2 + \cdots + s_k$, $s_i \in \mathcal{S}_i$.

Demostración:

$\Rightarrow]$ Supongamos que para alguna $s \in \mathcal{S}$

$$s = s_1 + s_2 + \cdots + s_k = a_1 + a_2 + \cdots + a_k$$

donde $s_i, a_i \in \mathcal{S}_i$ $1 \leq i \leq k$ y donde

$s_j \neq a_j$ para alguna j .

$$\Rightarrow \quad s_j - a_j = -\left[(s_1 - a_1) + (s_2 - a_2) + \cdots + (s_{j-1} - a_{j-1}) \right.$$
$$\left. + (s_{j+1} - a_{j+1}) + \cdots + (s_k - a_k) \right]$$

$$\neq 0$$

$$\Rightarrow \quad s_j - a_j \in \mathcal{S}_j \cap (\mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_{j-1} + \mathcal{S}_{j+1} + \cdots + \mathcal{S}_k)$$

$\Rightarrow$ $\mathcal{S}$ no es la suma directa de $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k$.

$\Leftarrow]$ Supongamos que $\mathcal{S}$ no es la suma directa de $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k$

$$\therefore \quad \mathcal{S}_j \cap (\mathcal{S}_1 + \mathcal{S}_2 + \cdots + \mathcal{S}_{j-1} + \mathcal{S}_{j+1} + \cdots + \mathcal{S}_k) \neq \{0\}$$

sea $s \in \mathcal{S}_j \cap (\mathcal{S}_1 + \cdots + \mathcal{S}_{j-1} + \mathcal{S}_{j+1} + \cdots + \mathcal{S}_k)$, $s \neq 0$

$$\therefore \quad s = s_j = s_1 + s_2 + \cdots + s_{j-1} + s_{j+1} + \cdots + s_k ,$$
con $s_i \in \mathcal{S}_i \quad 1 \le i \le k$

$$\therefore \quad s = s_1 + s_2 + \cdots + s_{j-1} + 0 + s_{j+1} + \cdots + s_k$$
$$= 0 + 0 + \cdots + 0 + s_j + 0 + \cdots + 0$$

Como $s \neq 0 \Rightarrow s_i \neq 0$ para alguna i,
$i \in \{1, 2, \ldots, j-1, j+1, \ldots, k\}$

y por lo tanto el vector s tiene dos representaciones de la forma indicada.

■

RESULTADO

$$\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_k \Leftrightarrow$$
$$\dim(\mathcal{S}) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) + \cdots + \dim(\mathcal{S}_k).$$

Demostración:

Sea $\{ s_1^{(i)}, s_2^{(i)}, \ldots, s_{r_i}^{(i)} \}$ una base de $\mathcal{S}_i$,
$1 \le i \le k$

y sea $A_i = \left[s_1^{(i)} \,|\, s_2^{(i)} \,|\, \cdots \,|\, s_{r_i}^{(i)} \right] \qquad 1 \le i \le k$

Por lo tanto

$\mathcal{S} = R(A)$ donde

$$A = \left[A_1 \,|\, A_2 \,|\, \cdots \,|\, A_k \right]$$

$\Leftarrow$ Supongamos $\mathcal{S}$ no es la suma directa de $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_k$

por lo tanto existe un vector $s \in \mathcal{S}$ tal que

$$s = Ax_1 = Ax_2 \qquad \text{con } x_1 \neq x_2$$

$$\Rightarrow A(x_1 - x_2) = 0 \qquad x_1 \neq x_2$$

$$\Rightarrow N(A) \neq \{0\}$$

$$\Rightarrow r(A) < r_1 + r_2 + \cdots + r_k$$

$$\Rightarrow \dim(\mathcal{S}) < \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) + \cdots + \dim(\mathcal{S}_k)$$

$\therefore$ $\dim(\mathcal{S}) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) + \cdots + \dim(\mathcal{S}_k) \Rightarrow$
$\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_k$

$\Rightarrow$ Supongamos que

$$\dim(\mathcal{S}) < \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) + \cdots + \dim(\mathcal{S}_k)$$

$$\Rightarrow r(A) < r_1 + r_2 + \cdots + r_k$$

$$\Rightarrow N(A) \neq \{0\}$$

$\Rightarrow$ $\exists\ s \neq 0$ tal que $As = 0 = A0$

$\Rightarrow$ el vector 0 tiene dos representaciones de la forma $s_1 + s_2 + \cdots + s_k$

con $s_i \in \mathcal{S}_i$

$\Rightarrow$ $\mathcal{S}$ no es la suma directa de $\mathcal{S}_1, \mathcal{S}_2, \ldots, \mathcal{S}_K$

$$\therefore \quad \mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_K \Rightarrow \dim(\mathcal{S}) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) + \cdots + \dim(\mathcal{S}_K)$$

■

RESULTADO

$$\mathcal{S} = \mathcal{S}_1 \oplus \mathcal{S}_2 \oplus \cdots \oplus \mathcal{S}_K \iff$$

$$\{s_1^{(1)}, s_2^{(1)}, \ldots, s_{r_1}^{(1)}\} \cup \{s_1^{(2)}, s_2^{(2)}, \ldots, s_{r_2}^{(2)}\} \cup \cdots \cup \{s_1^{(K)}, s_2^{(K)}, \ldots, s_{r_K}^{(K)}\}$$

es una base de $\mathcal{S}$.

RESULTADO

Sea $\{s_1, s_2, \ldots, s_K\}$ una base de $\mathcal{S}$. Entonces

$$\mathcal{S} = [s_1] \oplus [s_2] \oplus \cdots \oplus [s_K] .$$

DEFINICION

Sean $\mathcal{S}_1$ y $\mathcal{S}_2$ subespacios de $\mathbb{R}^m$. Decimos que $\mathcal{S}_1$ y $\mathcal{S}_2$ son ORTOGONALES si todo vector s_1 de $\mathcal{S}_1$ es ortogonal a todo vector s_2 de $\mathcal{S}_2$:

$$s_1^t s_2 = 0 \qquad \forall\, s_1 \in \mathcal{S}_1 \;,\; s_2 \in \mathcal{S}_2$$

DEFINICION

Sea $\mathcal{S}$ un subespacio de $\mathbb{R}^m$. El COMPLEMENTO ORTOGONAL de $\mathcal{S}$, denotado por $\mathcal{S}^\perp$ es:

$$\begin{aligned}\mathcal{S}^\perp &= \{x \in \mathbb{R}^m : x \perp s \quad \forall s \in \mathcal{S}\}\\ &= \{x \in \mathbb{R}^m : x^t s = 0 \quad \forall s \in \mathcal{S}\}\end{aligned}$$

RESULTADOS

Sea $\mathcal{S}$ un subespacio de $\mathbb{R}^m$. Entonces

1) $\mathcal{S}^\perp$ es un subespacio de $\mathbb{R}^m$.

Demostración:

$$\begin{aligned} x \in \mathcal{S}^\perp &\Rightarrow \lambda (x^t s) = (\lambda x)^t s = 0 \quad \forall s \in \mathcal{S}\\ &\Rightarrow \lambda x \in \mathcal{S}^\perp\end{aligned}$$

$$\begin{aligned} x_1, x_2 \in \mathcal{S}^\perp &\Rightarrow (x_1 + x_2)^t s = x_1^t s + x_2^t s = 0 \quad \forall s \in \mathcal{S}\\ &\Rightarrow x_1 + x_2 \in \mathcal{S}^\perp\end{aligned}$$

$\square$

2) Sea $\{s_1, s_2, \ldots, s_k\}$ un conjunto de generadores de $\mathcal{S}$. Entonces

$$\mathcal{S}^\perp = \{x \in \mathbb{R}^m : x \perp s_i \,,\; 1 \leq i \leq k\}$$

Demostración:

Sea $\tilde{\mathcal{J}} = \{x \in \mathbb{R}^m : x \perp s_i \ , \ 1 \leq i \leq k\}$

y $S = [s_1|s_2|\cdots|s_k]$, entonces

$\mathcal{J}^\perp \subseteq \tilde{\mathcal{J}}$ obviamente. Además:

$$x \in \tilde{\mathcal{J}} \Rightarrow x^t s_i = 0 \quad 1 \leq i \leq k$$

$$\Rightarrow x^t [s_1|s_2|\cdots|s_k] = 0$$

$$\Rightarrow x^t S = 0^t$$

$$\Rightarrow x^t (S\tilde{x}) = 0 \quad \forall \tilde{x} \in \mathbb{R}^k$$

$\Rightarrow x \in \mathcal{J}^\perp$ (ya que las columnas de S generan $\mathcal{J}$).

$$\therefore \tilde{\mathcal{J}} \subseteq \mathcal{J}^\perp \quad \therefore \tilde{\mathcal{J}} = \mathcal{J}^\perp \quad \blacksquare$$

3) Sea $\{s_1, s_2, \ldots, s_k\}$ un conjunto de generadores de $\mathcal{J}$, es decir, $\mathcal{J} = R(S)$, donde

$$S^t = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_k^t \end{bmatrix}$$

Entonces

$$R(S)^{\perp} = \mathcal{S}^{\perp} = \{ x \in \mathbb{R}^m : S^t x = 0 \} = N(S^t).$$

Demostración :

$$\begin{aligned} \mathcal{S}^{\perp} &= \{ x \in \mathbb{R}^m : s_i^t x = 0 , \ 1 \leq i \leq k \} \\ &= \{ x \in \mathbb{R}^m : S^t x = 0 \} \\ &= N(S^t) \end{aligned}$$

■

4) $\mathcal{S}^{\perp} \cap \mathcal{S} = \{0\}$

Demostración :

Si existe $x \neq 0$ tal que $x \in \mathcal{S}^{\perp} \cap \mathcal{S}$

$\Rightarrow \quad x^t s = 0 \quad \forall s \in \mathcal{S} \qquad (x \in \mathcal{S}^{\perp})$

$\Rightarrow \quad x^t x = 0 \qquad (x \in \mathcal{S}) \quad !$

Esto es imposible, ya que $x \neq 0$

$\therefore \quad \mathcal{S}^{\perp} \cap \mathcal{S} = \{0\}$

5) $\dim(\mathcal{S}^{\perp}) = m - \dim(\mathcal{S})$

Demostración

Sea S^t la matriz del resultado (3), entonces

$$\dim(\mathcal{S}^{\perp}) = \dim(N(S^t)) = m-k = m - \dim(\mathcal{S}).$$

6) $\mathbb{R}^m = \mathcal{S} \oplus \mathcal{S}^{\perp}$

Demostración:

Por el resultado (5)

$$\dim(\mathcal{S}^{\perp}) + \dim(\mathcal{S}) = m$$

Por el resultado (4)

$$\mathcal{S}^{\perp} \cap \mathcal{S} = \{0\}$$

$$\therefore \quad \dim(\mathcal{S}^{\perp}) + \dim(\mathcal{S}) = m = \dim(\mathcal{S}^{\perp} + \mathcal{S})$$

$$\therefore \quad \mathbb{R}^m = \mathcal{S} \oplus \mathcal{S}^{\perp}$$

7) $(\mathcal{S}^{\perp})^{\perp} = \mathcal{S}$

Demostración:

Sea $s \in \mathcal{S}$ $\quad \therefore \quad s = Sx$

$\therefore \quad s^t y = x^t S^t y = 0 \quad \forall \, y \in N(S^t) = \mathcal{S}^\perp$

$\therefore \quad s \in (\mathcal{S}^\perp)^\perp \quad \therefore \quad \mathcal{S} \subseteq (\mathcal{S}^\perp)^\perp$

Por el resultado anterior,

$$\mathbb{R}^m = \mathcal{S} \oplus \mathcal{S}^\perp$$

$$\mathbb{R}^m = \mathcal{S}^\perp \oplus (\mathcal{S}^\perp)^\perp$$

$\therefore \quad \dim(\mathcal{S}) = \dim((\mathcal{S}^\perp)^\perp)$

y como $\mathcal{S} \subseteq ((\mathcal{S}^\perp)^\perp) \Rightarrow \mathcal{S} = (\mathcal{S}^\perp)^\perp$ ∎

8) $R(A) = N(A^t)^\perp$.

9) $R(A^t) = N(A)^\perp$.

10) $\mathbb{R}^m = R(A) \oplus N(A^t)$.

11) $\mathbb{R}^n = N(A) \oplus R(A^t)$.

INTERSECCION DE VARIEDADES LINEALES

Supongamos que tenemos dos variedades lineales

$$\mathcal{L}_1 = x_1 + \mathcal{S}_1 \subseteq \mathbb{R}^n$$

$$\mathcal{L}_2 = x_2 + \mathcal{S}_2 \subseteq \mathbb{R}^n$$

donde $\mathcal{S}_1 \neq \mathbb{R}^n$ y $\mathcal{S}_2 \neq \mathbb{R}^n$.

¿ Cuando tendremos que $\mathcal{L}_1 \cap \mathcal{L}_2 \neq \emptyset$?

Para que $\mathcal{L}_1 \cap \mathcal{L}_2 \neq \emptyset$ se necesita que existan $s_1 \in \mathcal{S}_1$ y $s_2 \in \mathcal{S}_2$ tales que

$$x_1 + s_1 = x_2 + s_2$$

Pero $x_1 + s_1 = x_2 + s_2 \iff x_1 - x_2 = -s_1 + s_2$

$$\iff x_1 - x_2 \in \mathcal{S}_1 + \mathcal{S}_2$$

Por lo tanto, las variedades lineales determinadas por $\mathcal{S}_1$ y $\mathcal{S}_2$ siempre se intersecan si $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$

¿ Bajo qué condiciones tenemos que $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$?

Lo que hay que observar primero es que si $\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) < n$ entonces es imposible que $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$, ya que en este caso

$$\dim(\mathcal{S}_1 + \mathcal{S}_2) \leq \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) < n$$

Por lo tanto es necesario que

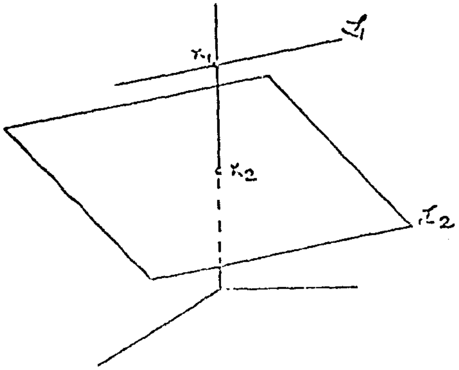
$$\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) \geq n.$$

Sin embargo esta condición no es suficiente, ya que si, por ejemplo

$$\mathcal{S}_1 = [a_1] \subseteq \mathbb{R}^3 \qquad \mathcal{S}_2 = [a_1, a_2] \subseteq \mathbb{R}^3$$

$$\mathcal{L}_1 = x_1 + \mathcal{S}_1 \ , \quad \mathcal{L}_2 = x_2 + \mathcal{S}_2 , \qquad \text{con } x_1 \neq x_2$$

entonces $\mathcal{L}_1 \cap \mathcal{L}_2 = \emptyset$ y $\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) = 3$



Sin embargo, tampoco basta con tener estas dos condiciones para que $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$, como se puede ver en el sig. ejemplo.

$$\text{Sea} \quad A_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

y sean $\mathcal{S}_1 = R(A_1)$, $\mathcal{S}_2 = R(A_2)$ subespacios de $\mathbb{R}^5$

$$\therefore \quad \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) = 6 > 5$$

$$\mathcal{S}_1 \not\subset \mathcal{S}_2 \quad y \quad \mathcal{S}_2 \not\subset \mathcal{S}_1$$

Sin embargo $\dim(\mathcal{S}_1 + \mathcal{S}_2) = 4$

Esto se debe a que $\dim(\mathcal{S}_1 \cap \mathcal{S}_2) = 2$.

O sea, no es suficiente con pedir que dos subespacios $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{R}^n$ no sean paralelos y que la suma de sus dimensiones sea mayor que el espacio que los contiene, para tener que $\mathcal{S}_1 + \mathcal{S}_2 = \mathbb{R}^n$ .. hay que exigir además que $\mathcal{S}_1 \cap \mathcal{S}_2$ sea pequeña.

RESULTADO

Sean $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathbb{R}^n$ subespacios de $\mathbb{R}^n$ tales que

$$\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) = n + k \,, \quad k \geq 0$$

Entonces

$$\mathbb{R}^n = \mathcal{S}_1 + \mathcal{S}_2 \iff \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = k$$

Demostración

$$\begin{aligned} \dim(\mathcal{S}_1 + \mathcal{S}_2) &= \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) \\ &= n + k - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) \end{aligned}$$

Por tanto

$$\begin{aligned} \dim(\mathcal{S}_1 + \mathcal{S}_2) = n &\iff n + k - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = n \\ &\iff \dim(\mathcal{S}_1 \cap \mathcal{S}_2) = k \end{aligned}$$

RESULTADO

Sean $\mathcal{L}_1, \mathcal{L}_2 \subseteq \mathbb{R}^n$ dos variedades lineales de $\mathbb{R}^n$, cuyos subespacios asociados son $\mathcal{S}_1$ y $\mathcal{S}_2$ respectivamente. Entonces $\mathcal{L}_1 \cap \mathcal{L}_2 \neq \emptyset$ si $\dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) = n + \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$.

CAPITULO IV

ESPACIOS VECTORIALES

ESPACIOS VECTORIALES

DEFINICION

Un ESPACIO VECTORIAL REAL consta de

a) Un conjunto $\mathbb{V} \neq \emptyset$

b) Dos operaciones $+ : \mathbb{V} \times \mathbb{V} \to \mathbb{V}$ y $\cdot : \mathbb{R} \times \mathbb{V} \to \mathbb{V}$
$+ (v_1, v_2) = v_1 + v_2$
$\cdot (\alpha, v) = \alpha \cdot v = \alpha v$

con las siguientes propiedades :

1) $v_1 + v_2 = v_2 + v_1$
2) $v_1 + (v_2 + v_3) = (v_1 + v_2) + v_3$
3) Existe un vector $0 \in \mathbb{V}$ tal que $0 + v = v \quad \forall v \in \mathbb{V}$
4) Para todo vector $v \in \mathbb{V}$ existe un único vector $-v \in \mathbb{V}$ tal que $v + (-v) = 0$

5) $1 \cdot v = v \quad \forall v \in \mathbb{V}$
6) $(\alpha\beta) v = \alpha (\beta v) \quad \forall v \in \mathbb{V}, \ \alpha, \beta \in \mathbb{R}$
7) $\alpha (v_1 + v_2) = \alpha v_1 + \alpha v_2$
8) $(\alpha + \beta) v = \alpha v + \beta v$

a los elementos de $\mathbb{V}$ los llamamos VECTORES .

EJEMPLOS DE ESPACIOS VECTORIALES

1) $\mathbb{R}^m$, con las operaciones de suma de vectores y multiplicación por un escalar como las definimos.

2) $\{f:\mathbb{R}\to\mathbb{R} \ : \ f \text{ es polinomio}\}$ donde

$$(f+g)(t) = f(t)+g(t)$$
$$(\alpha f)(t) = \alpha[f(t)]$$

3) $\{f:\mathbb{R}\to\mathbb{R} : f \text{ es polinomio de grado} \leq k\}$ (con las mismas operaciones que en (2))

4) $M(m,n)$ con las operaciones de suma de matrices y multiplicación de un escalar por una matriz como las definimos.

5) Un cuadrado mágico 3×3 es una matriz de orden 3 $A=[\alpha_{ij}]$ tal que

a) la suma de las columnas es siempre la misma

b) la suma de los renglones es siempre la misma

c) la suma de las diagonales es siempre la misma.

d) las sumas de a, b y c son la misma cantidad.

Es decir, si $A_{3\times3}$ es un cuadrado mágico, entonces existe un real $s = s(A)$ tal que

$$\alpha_{i1} + \alpha_{i2} + \alpha_{i3} = s(A) \qquad i = 1,2,3$$

$$\alpha_{1j} + \alpha_{2j} + \alpha_{3j} = s(A) \qquad j = 1,2,3$$

$$\alpha_{11} + \alpha_{22} + \alpha_{33} = s(A)$$

$$\alpha_{13} + \alpha_{22} + \alpha_{31} = s(A)$$

El conjunto de los cuadrados mágicos 3x3 es un espacio vectorial:

Sean A y B cuadrados mágicos de orden 3. Entonces

$$\gamma A = \begin{bmatrix} \gamma\alpha_{11} & \gamma\alpha_{12} & \gamma\alpha_{13} \\ \gamma\alpha_{21} & \gamma\alpha_{22} & \gamma\alpha_{23} \\ \gamma\alpha_{31} & \gamma\alpha_{32} & \gamma\alpha_{33} \end{bmatrix}$$

$$\therefore \quad \gamma\alpha_{i1} + \gamma\alpha_{i2} + \gamma\alpha_{i3} = \gamma S(A) \qquad 1 \leq i \leq 3$$

$$\gamma\alpha_{1j} + \gamma\alpha_{2j} + \gamma\alpha_{3j} = \gamma S(A) \qquad 1 \leq j \leq 3$$

$$\gamma\alpha_{11} + \gamma\alpha_{22} + \gamma\alpha_{33} = \gamma S(A)$$

$$\gamma\alpha_{13} + \gamma\alpha_{22} + \gamma\alpha_{31} = \gamma S(A)$$

Por lo tanto γA es un cuadrado mágico, y $S(\gamma A) = \gamma(S(A))$.

$$A+B = \begin{bmatrix} \alpha_{11}+\beta_{11} & \alpha_{12}+\beta_{12} & \alpha_{13}+\beta_{13} \\ \alpha_{21}+\beta_{21} & \alpha_{22}+\beta_{22} & \alpha_{23}+\beta_{23} \\ \alpha_{31}+\beta_{31} & \alpha_{32}+\beta_{32} & \alpha_{33}+\beta_{33} \end{bmatrix}$$

$$\therefore \quad \sum_{k=1}^{3}(\alpha_{ik}+\beta_{ik}) = \sum_{k=1}^{3}\alpha_{ik} + \sum_{k=1}^{3}\beta_{ik} = S(A)+S(B), \quad 1 \leq i \leq 3$$

$$\sum_{k=1}^{3}(\alpha_{kj}+\beta_{kj}) = \sum_{k=1}^{3}\alpha_{kj} + \sum_{k=1}^{3}\beta_{kj} = S(A)+S(B), \qquad 1 \leq j \leq 3$$

$$\sum_{k=1}^{3}(\alpha_{kk}+\beta_{kk}) = \sum_{k=1}^{3}\alpha_{kk} + \sum_{k=1}^{3}\beta_{kk} = S(A)+S(B)$$

$$\sum_{k=1}^{3}(\alpha_{k,4-k}+\beta_{k,4-k}) = \sum_{k=1}^{3}\alpha_{k,4-k} + \sum_{k=1}^{3}\beta_{k,4-k} = S(A)+S(B)$$

Por lo tanto A+B es un cuadrado mágico, y $s(A+B) = s(A) + s(B)$.

Las propiedades de espacio vectorial se heredan de las matrices.

6) Un cuadrado mágico de orden n es una matriz de orden n $A=[\alpha_{ij}]$ tal que tiene asociado un real $s = s(A)$ y

a) La suma de las columnas de A vale $s(A)$:

$$\sum_{k=1}^{n} \alpha_{kj} = s(A) \quad , \quad 1 \leq j \leq n$$

b) La suma de los renglones de A vale $s(A)$:

$$\sum_{k=1}^{n} \alpha_{ik} = s(A), \qquad 1 \leq k \leq n$$

c) La suma de cada diagonal vale $s(A)$:

$$\sum_{k=1}^{n} \alpha_{1,n+1-k} = s(A)$$

$$\sum_{k=1}^{n} \alpha_{kk} = s(A).$$

El conjunto de los cuadrados mágicos $n \times n$ es un espacio vectorial.

7) $V = \{ \{s_k\} : a_1 s_k + a_2 s_{k-1} = 0 \}$, donde

$\{s_k\}$ es una sucesión de números reales y

$$\alpha \{s_k\} = \{\alpha s_k\}$$
$$\{s_k\} + \{z_k\} = \{s_k + z_k\}$$

Si $\{s_k\}, \{z_k\} \in V$ entonces

$$a_1 (s_k + z_k) + a_2 (s_{k-1} + z_{k-1}) =$$
$$= (a_1 s_k + a_2 s_{k-1}) + (a_1 z_k + a_2 z_{k-1}) = 0$$

$$\therefore \{s_k\} + \{z_k\} \in V$$

$$a_1 (\alpha s_k) + a_2 (\alpha s_{k-1}) = \alpha (a_1 s_k + a_2 s_{k-1}) = 0$$

$$\therefore \alpha \{s_k\} \in V .$$

Los propiedades de espacio vectorial son cons cuencia de los propiedades de los números reale
Por ejemplo:

Si $\{s_k\}$ y $\{z_k\} \in V$, entonces $\{s_k\} + \{z_k\} \in V$ y $\alpha \{s_k\}$, $\alpha \{z_k\} \in V$ y por lo tanto:

$$\alpha\left[\{s_k\}+\{z_k\}\right] = \alpha\cdot\{s_k+z_k\} = \{\alpha(s_k+z_k)\}$$

$$= \{\alpha s_k+\alpha z_k\} = \{\alpha s_k\}+\{\alpha z_k\}$$

$$= \alpha\{s_k\}+\alpha\{z_k\}$$

$$\therefore\ \alpha\left[\{s_k\}+\{z_k\}\right] = \alpha\{s_k\}+\alpha\{z_k\}.$$

8) $\mathbb{V} = \left\{ \{s_k\} : a_0 s_k + a_1 s_{k-1} + \cdots + a_r s_{k-r} = 0 \right\}$

9) $\mathbb{V} = \left\{ y : \mathbb{R}\to\mathbb{R} \ : \ y \text{ es diferenciable dos veces y } a_0 y''(\tau) + a_1 y'(\tau) + a_2 y(\tau) = 0 \right\}$

donde $(y_1+y_2)(\tau) = y_1(\tau)+y_2(\tau)$

$(\alpha y)(\tau) = \alpha\left[y(\tau)\right]$

Si $y_1, y_2 \in \mathbb{V}$, entonces y_1+y_2 es diferenciable dos veces y

$$a_0(y_1+y_2)''(\tau) + a_1(y_1+y_2)'(\tau) + a_2(y_1+y_2)(\tau) =$$

$$= a_0(y_1''+y_2'')(\tau) + a_1(y_1'+y_2')(\tau) + a_2(y_1+y_2)(\tau)$$

$$= a_0 y_1''(\tau) + a_1 y_1'(\tau) + a_2 y_1(\tau) +$$
$$+ a_0 y_2''(\tau) + a_1 y_2''(\tau) + a_2 y_2''(\tau)$$

$= 0$

Por lo tanto $y_1 + y_2 \in V$.

si $y \in V$ entonces (αy) es diferenciable dos veces y

$$a_0 (\alpha y)''(t) + a_1 (\alpha y)'(t) + a_2 (\alpha y)(t) =$$

$$= \alpha a_0 y''(t) + \alpha a_1 y'(t) + \alpha a_2 y(t) = 0$$

$\therefore \; \alpha y \in V.$

Las propiedades de espacio vectorial se cumplen, ya que al aplicar una función se obtienen números reales.

10) $V = \{ f : \mathbb{R} \to \mathbb{R} :$ f es contínua y periódica, de período $2\pi \}$

con $(f+g)(t) = f(t) + g(t)$
$(\alpha f)(t) = \alpha f(t)$.

si $f, g \in V$ entonces

$f+g$ es contínua, y

$(f+g)(t+2\pi) = f(t+2\pi) + g(t+2\pi) = f(t) + g(t)$

$\therefore$ $f+g$ es periódica de período 2π

$\therefore$ $f+g \in V$

$$(\alpha f)(\tau+2\pi) = \alpha f(\tau+2\pi) = \alpha f(\tau)$$

$\therefore$ (αf) es continua y periódica de período 2π
$\therefore$ $(\alpha f) \in V$.

ii) sea $I = [\tau_1, \tau_n] \subseteq \mathbb{R}$ y sea

$P: \quad \tau_1 < \tau_2 < \dots < \tau_n$ una partición de I

sean $I_1 = [\tau_1, \tau_2]$

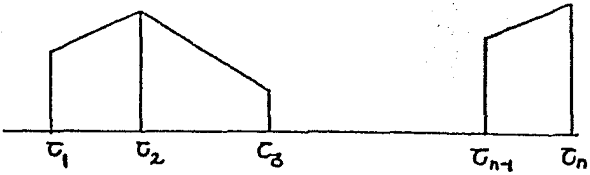
$I_2 = [\tau_2, \tau_3]$

$\vdots$

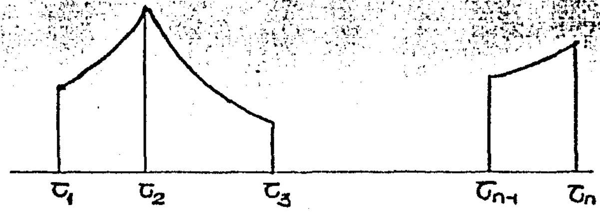
$I_{n-1} = [\tau_{n-1}, \tau_n]$

Sean

A) $\mathcal{S}_1 = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado ≤ 1 y f continua en $I \}$



B) $\mathcal{S}_2 = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado ≤ 2 $(1 \leq i \leq n-1)$ y f es continua en I



C) $\mathcal{S}_m = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$ $(1 \leq i \leq n-1)$ y f es continua en $I \}$

Entonces $\mathcal{S}_1, \mathcal{S}_2, \dots, \mathcal{S}_m$ son espacios vectoriales.

Ya que si $f_1, f_2 \in \mathcal{S}_k$, $1 \leq k < m$ entonces

$\therefore$ $(f_1+f_2)|_{I_i}$ es un polinomio de grado $\leq k$ $(1\leq i\leq n-1)$

Y (f_1+f_2) es continua en I.

$\therefore$ $f_1+f_2 \in \mathcal{S}_k$,

y si $\alpha \in \mathbb{R}$, $f \in \mathcal{S}_k$ entonces

$(\alpha f)|_{I_i}$ es un polinomio de grado $\leq k$ $(1\leq i\leq n-1)$

y (αf) es continua en I y $\therefore$ $(\alpha f) \in \mathcal{S}_k$

$\therefore$ $\mathcal{S}_k$, $k \in \mathbb{N}$, es un espacio vectorial (es fácil checarlo).

Sean

D) $\mathcal{S}_{m,1} = \{f: I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$ $(1\leq i\leq n-1)$ y f es C^1 en $I\}$ $m\geq 1$.

$\mathcal{S}_{m,2} = \{f: I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$ $(1\leq i\leq n-1)$ y f es C^2 en I,$\}$ con $m\geq 2$

$\vdots$

$\mathcal{S}_{m,k} = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$ $(1 \leq i \leq n-1)$ y f es C^k en I, $m \geq k \}$

Observar que

$\mathcal{S}_{m,m} = \{ f : I \to \mathbb{R} : f$ es un polinomio de grado $\leq$

$\mathcal{S}_{m,k}$ es un espacio vectorial :

Si $f_1, f_2 \in \mathcal{S}_{m,k}$ entonces

$(f_1+f_2)|_{I_i}$ es un polinomio de grado $\leq m$ y

(f_1+f_2) es C^k en I.

Si $f \in \mathcal{S}_{m,k}$, $\alpha \in \mathbb{R}$ entonces

$(\alpha f)|_{I_i}$ es un polinomio de grado $\leq m$ y

(αf) es C^k en I.

$\therefore$ (f_1+f_2) y (αf) $\in$ $\mathcal{S}_{m,k}$.

Con esto ya se puede ver fácilmente que $\mathcal{S}_{m,k}$ es un espacio vectorial.

DEFINICION

Sea V un espacio vectorial (real). Un SUBESPACIO de V es un subconjunto $W \subseteq V$ no-vacío tal que

1) $w_1 + w_2 \in W \quad \forall\, w_1, w_2 \in W$

2) $\alpha w \in W \quad \forall\, w \in W$

Ejemplos

1) V es un subespacio de V

2) $\{0\}$ es un subespacio de V, llamado SUBESPACIO NULO de V

3) Sea $V = \{ f : \mathbb{R} \to \mathbb{R} \mid f$ es un polinomio $\}$,

$W = \{ f : \mathbb{R} \to \mathbb{R} \mid f$ es polinomio de grado $\leq K \}$,

entonces W es un subespacio de V.

4) Sean $V = \{ f : \mathbb{R} \to \mathbb{R} : f$ es continua $\}$,

$W = \{ f : \mathbb{R} \to \mathbb{R} : f$ continua y periódica, de periodo $2\pi \}$.

entonces $W \subseteq V$

RESULTADO

Sean W_1 y W_2 subespacios de V. Entonces $W_1 \cap W_2$ es un subespacio de V.

La demostración es análoga a la que dimos para $V = \mathbb{R}^n$.

DEFINICION

Sean $v_1, v_2, \dots, v_n \in V$. El SUBESPACIO GENERADO por $v_1, v_2, \dots, v_n$, denotado por $[v_1, v_2, \dots, v_n]$ está dado por:

$$[v_1, v_2, \dots, v_n] = \left\{ v \in V : v = \sum_{i=1}^{n} \alpha_i v_i \ , \ \alpha_i \in \mathbb{R} \right\}$$

Si $W = [v_1, v_2, \dots, v_n]$, decimos que $\{v_1, v_2, \dots, v_n\}$ es UN CONJUNTO DE GENERADORES de W.

PROPIEDADES

1) Si $x_1, x_2, \dots, x_k \in [v_1, v_2, \dots, v_n] \Rightarrow$

$$[x_1, x_2, \dots, x_k] \subseteq [v_1, v_2, \dots, v_n]$$

Demostración:

$$\text{si} \quad x_1 = \sum_{i=1}^{n} E_{i1} v_i \;,\; x_2 = \sum_{i=1}^{n} E_{i2} v_i, \ldots, x_n = \sum_{i=1}^{n} E_{in} v_i$$

entonces si $y \in [x_1, x_2, \ldots, x_k] \Rightarrow$

$$y = \sum_{j=1}^{k} \alpha_j x_j = \sum_{j=1}^{k} \alpha_j \left(\sum_{i=1}^{n} E_{ij} v_i \right)$$

$$= \sum_{i=1}^{n} \left(\sum_{j=1}^{k} \alpha_j E_{ij} \right) v_i \in [v_1, v_2, \ldots, v_n]$$

$$\therefore \quad [x_1, x_2, \ldots, x_n] \subseteq [v_1, v_2, \ldots, v_n]$$

2) $[v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n] \neq [v_1, v_2, \ldots, v_n] \Leftrightarrow$

$$v_i \notin [v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n]$$

3) $[v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n] = [v_1, v_2, \ldots, v_n] \Leftrightarrow$

$$v_i \in [v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n]$$

DEFINICION

Sean $v_1, v_2, \dots, v_n \in V$. Decimos que $\{v_1, v_2, \dots, v_n\}$ es LINEALMENTE DEPENDIENTE si uno de los vectores se puede obtener como una combinación lineal de los $n-1$ restantes. Es decir, si existe $i \in \{1, 2, \dots, n\}$ tal que

$$v_i \in [v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n]$$

Decimos que $\{v_1, v_2, \dots, v_n\}$ es LINEALMENTE INDEPENDIENTE si ninguno de los vectores v_i está en el subespacio generado por los $n-1$ vectores restantes. Es decir, si

$$v_i \notin [v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n] \quad \forall i, \ 1 \leq i \leq n$$

RESULTADO

Sea $\{v_1, v_2, \dots, v_n\} \subseteq V$. Entonces

1) $\{v_1, v_2, \dots, v_n\}$ es linealmente dependiente $\Leftrightarrow$

el vector cero se puede expresar como una combinación lineal

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0$$

donde al menos una de las α_i's es diferente de cero.

2) $\{v_1, v_2, \dots, v_n\}$ es linealmente independiente $\Leftrightarrow$

la única combinación lineal de $v_1, v_2, \dots, v_n$ que nos da el vector cero es la que tiene todos sus coeficientes iguales a cero.
$\left(\sum_{i=1}^{n} \alpha_i v_i = 0 \Rightarrow \alpha_i = 0 \;, 1 \le i \le n \right)$.

Demostración:

(1) $\Rightarrow$ Supongamos $v_i \in [v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n]$

$$\Rightarrow v_i = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \cdots + v_n$$

$$\Rightarrow 0 = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_{i-1} v_{i-1} - v_i + \alpha_{i+1} v_{i+1} + \cdots + v_n$$

$\Rightarrow \quad 0 = \sum_{j=1}^{n} \alpha_j v_j$, donde $\alpha_i = -1 \neq 0$

$\Leftarrow$] Supongamos $\sum_{j=1}^{n} \alpha_j v_j = 0$ con $\alpha_i \neq 0$

$$\Rightarrow \sum_{j=1}^{n} \left(\frac{\alpha_j}{\alpha_i}\right) v_j = 0$$

$$\Rightarrow \sum_{j=1}^{i-1} \left(\frac{\alpha_j}{\alpha_i}\right) v_j + v_i + \sum_{j=i+1}^{n} \left(\frac{\alpha_j}{\alpha_i}\right) v_j = 0$$

$$\Rightarrow v_i = \sum_{j=1}^{i-1} \left(-\frac{\alpha_j}{\alpha_i}\right) v_j + \sum_{j=i+1}^{n} \left(-\frac{\alpha_j}{\alpha_i}\right) v_j$$

$$\Rightarrow v_i \in [v_1, v_2, \dots, v_{i-1}, v_{i+1}, \dots, v_n]$$

■

(2) Tenemos que

$\{v_1, v_2, \dots, v_n\}$ no dependiente $\Leftrightarrow$ $\{v_1, v_2, \dots, v_n\}$ independiente

Por lo tanto la demostración para (1) y (2) es la misma

DEFINICION

Decimos que el espacio vectorial V es FINITAMENTE GENERADO si tiene un número finito de generadores, es decir, si existen $v_1, v_2, \dots, v_n \in V$ tal que

$$V = [v_1, v_2, \dots, v_n]$$

Ejemplos

1) $\mathbb{R}^m$
$\{f : \mathbb{R} \to \mathbb{R} \mid f$ es polinomio de grado $\leq k\}$
$M(m,n)$
$\{A \in M_n :$ A es cuadrado mágico $\}$
$\{\{s_k\} : a_0 s_k + a_1 s_{k-1} + \cdots + a_r s_{k-r} = 0\}$
$\{y : I \to \mathbb{R} : a_0 y''(t) + a_1 y'(t) + a_2 y(t) = 0\}$
$\mathcal{S}_m$
$\mathcal{S}_{m,k}$

son finitamente generados

2) $\{f : \mathbb{R} \to \mathbb{R} : f$ es un polinomio $\}$
$\{f : \mathbb{R} \to \mathbb{R} : f$ es continua y periódica de período $2\pi\}$

no son finitamente generados.

Nosotros trabajaremos únicamente con espacios finitamente generados.

RESULTADO

Sea V un espacio vectorial finitamente generado, tal que

$$V = [v_1, v_2, \ldots, v_m] = [w_1, w_2, \ldots, w_n] \quad \text{con} \quad m < n.$$

Entonces $\{w_1, w_2, \ldots, w_n\}$ es linealmente dependiente.

Demostración :

Supongamos $\{\omega_1, \omega_2, \ldots, \omega_m\}$ es independiente.

Veremos que $V = [w_1, w_2, \ldots, w_m]$.

(si $\{\omega_1, \omega_2, \ldots, \omega_m\}$ es dependiente entonces $\{w_1, w_2, \ldots, w_n\}$ es dependiente y no hay nada que demostrar).

Como $\{\omega_1, \omega_2, \ldots, \omega_m\}$ es independiente $\Rightarrow$ $\omega_1 \neq 0$

$$\therefore \quad \omega_1 = \lambda_1^{(1)} v_1 + \lambda_2^{(1)} v_2 + \cdots + \lambda_m^{(1)} v_m$$

Podemos suponer s.p.g. $\lambda_1^{(1)} \neq 0$ (si $\lambda_1^{(1)} = 0$ entonces renombramos las v's)

entonces $v_1 \in [w_1, v_2, \ldots, v_m]$ y por lo tanto $V \subseteq [w_1, v_2, \ldots, v_m] \subseteq V$

$$\therefore \quad V = [w_1, v_2, \ldots, v_m]$$

En un segundo paso, sabemos que $w_2 \notin [w_1]$

$\therefore \quad w_2 = \lambda_1^{(2)} w_1 + \lambda_2^{(2)} v_2 + \cdots + \lambda_m^{(2)} v_m \neq 0$ donde al menos una $\lambda_i^{(2)} \neq 0$ para $2 \leq i \leq m$.- ésto es resultado de que si

$$\lambda_1^{(2)} \neq 0 \quad \text{pero} \quad \lambda_i^{(2)} = 0 \quad \forall\, i,\ 2 \leq i \leq m$$

entonces $w_2 = \lambda_1^{(2)} w_1 \neq 0$! contradice la hipótesis de que $w_2 \notin [w_1]$

$\therefore$ alguna $\lambda_i^{(2)} \neq 0$ para $2 \leq i \leq m$.

Supondremos s.p.g. $\lambda_2^{(2)} \neq 0$

$\Rightarrow \quad v_2 \in [w_1, w_2, v_3, \ldots, v_m]$

$$\therefore \quad V = [w_1, v_2, v_3, \ldots, v_m] \subseteq [w_1, w_2, v_3, \ldots, v_m] \subseteq V$$

$$\therefore \quad V = [w_1, w_2, v_3, \ldots, v_m]$$

En el k-ésimo paso $1 \leq k \leq m$ si

$$V = [w_1, w_2, \dots, w_{k-1}, v_k, v_{k+1}, \dots, v_m]$$ entonces

vamos a tener que

$$V = [w_1, w_2, \dots, w_k, v_{k+1}, \dots, v_m]$$

ya que como $\{w_1, \dots, w_m\}$ es independiente
$\Rightarrow$ $w_k \notin [w_1, w_2, \dots, w_{k-1}]$ $(y \therefore w_k \neq 0)$

Como $\omega_k \in V \Rightarrow$

$$\omega_k = \lambda_1^{(k)} w_1 + \lambda_2^{(k)} w_2 + \dots + \lambda_{k-1}^{(k)} w_{k-1} + \lambda_k^{(k)} v_k + \lambda_{k+1}^{(k)} v_{k+1} + \dots + \lambda_m^{(k)} v_m .$$

si $\lambda_k^{(k)} = \lambda_{k+1}^{(k)} = \dots = \lambda_m^{(k)} = 0$ entonces

vamos a tener que

$$w_k = \lambda_1^{(k)} w_1 + \lambda_2^{(k)} w_2 + \dots + \lambda_{k-1}^{(k)} w_{k-1} \neq 0$$

$$\Rightarrow w_k \in [w_1, w_2, \dots, w_{k-1}] \; !$$

Por lo tanto al menos una $\lambda_i^{(k)}$ es diferente de cero, $k \leq i \leq m$. Vamos a suponer, sin pérdida de generalidad, que $\lambda_k^{(k)} \neq 0$.

Por lo tanto

$$\omega_k = \lambda_1^{(k)} w_1 + \lambda_2^{(k)} w_2 + \cdots + \lambda_{k-1}^{(k)} w_{k-1} + \lambda_k^{(k)} v_k + \cdots + \lambda_m^{(k)} v_m$$

donde $\lambda_k^{(k)} \neq 0$

$$\therefore \quad v_k \in [w_1, w_2, \ldots, w_k, v_{k+1}, \ldots, v_m]$$

$$\therefore V = [w_1, w_2, \ldots, w_{k-1}, v_k, v_{k+1}, \ldots, v_m] \subseteq$$

$$\subseteq [w_1, w_2, \ldots, w_{k-1}, w_k, v_{k+1}, \ldots, v_m] \subseteq V$$

$$\therefore \quad V = [w_1, w_2, \ldots, w_k, v_{k+1}, \ldots, v_m]$$

Por lo tanto, en el m-ésimo paso obtenemos que

$$V = [w_1, w_2, \ldots, w_m] = [w_1, w_2, \ldots, w_m, w_{m+1}, \ldots, w_n]$$

Como $n > m$ entonces $w_{m+1}, \ldots, w_n \in [w_1, w_2, \ldots, \omega_n]$

$\therefore$ $\{w_1, w_2, \ldots, w_n\}$ es linealmente dependiente

$\blacksquare$

DEFINICION

Sea V un espacio vectorial. Decimos que $\{v_1, v_2, \ldots, v_n\} \subseteq V$ es un CONJUNTO MINIMO DE GENERADORES de V si

1) $V = [v_1, v_2, \ldots, v_n]$

2) $V \neq [v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n] \quad \forall i, \; 1 \leq i \leq n.$

Es decir, si $\{v_1, v_2, \ldots, v_n\}$ genera todo V, pero al quitar un vector a $\{v_1, v_2, \ldots, v_n\}$ entonces ya no generamos todo V.

PROPIEDAD

Todo conjunto mínimo de generadores de V es linealmente independiente.

Demostración:

Sea $\{v_1, v_2, \ldots, v_n\}$ un conjunto mínimo de generadores de V.

Supongamos $\{v_1, v_2, \ldots, v_n\}$ es linealmente dependiente. Por lo tanto existe una i, $1 \leq i \leq n$, tal que

$$v_i \in [v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n]$$

$\Rightarrow$ $[v_1, v_2, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n] = [v_1, v_2, \ldots, v_n]$!

$\Rightarrow$ $[v_1, v_2, \ldots, v_n]$ no es un conjunto mínimo de generadores

$\therefore$ $\{v_1, v_2, \ldots, v_n\}$ es linealmente independiente. ■

RESULTADO

Sea V un espacio vectorial finitamente generado. Si $\{v_1, v_2, \ldots, v_n\}$ y $\{w_1, w_2, \ldots, w_r\}$ son dos conjuntos mínimos de generadores de V entonces $n = r$.

Demostración:

$r > n$ no es posible, ya que $r > n \Rightarrow$ $\{w_1, w_2, \ldots, w_r\}$ es dependiente.

$\therefore$ $r \leq n$.

$n > r$ no es posible, ya que $n > r \Rightarrow$ $\{v_1, v_2, \ldots, v_n\}$ es dependiente.

$\therefore$ $n \leq r$.

$\therefore$ $r = n$

□

DEFINICION

La DIMENSION de un espacio vectorial $\mathcal{V}$ es la cardinalidad de uno de sus conjuntos mínimos de generadores.

DEFINICION

UNA BASE de un espacio vectorial $\mathcal{V}$ (finitamente generado) es un conjunto de vectores $\{v_1, v_2, \dots, v_n\}$ con las siguientes propiedades:

1) $\mathcal{V} = [v_1, v_2, \dots, v_n]$ (genera a $\mathcal{V}$)

2) $\{v_1, v_2, \dots, v_n\}$ es linealmente independiente.

Observar que $\{v_1, v_2, \dots, v_n\}$ es una base de $\mathcal{V} \Leftrightarrow \{v_1, v_2, \dots, v_n\}$ es un conjunto mínimo de generadores de $\mathcal{V}$.

Ejemplos

1) Sea $\mathcal{V} = \{f : \mathbb{R} \to \mathbb{R} \mid f$ es polinomio de grado $\leq k$

entonces $\dim \mathcal{V} = k+1$, y una base de $\mathcal{V}$ es $\{1, x, x^2, \dots, x^k\}$.

2) Sea $V = M(m,n)$. Entonces $\dim V = mn$ y una base de V es

$$\{E_{11}, E_{12}, \ldots, E_{1n}, E_{21}, E_{22}, \ldots, E_{2n}, \ldots, E_{m1}, E_{m2}, \ldots, E_{mn}\}$$

donde $E_{ij} = e_i e_j^t$; $e_i \in \mathbb{R}^m$, $e_j \in \mathbb{R}^n$; $1 \leq i \leq m$
$1 \leq j \leq n$.

3) Sea $V = \{A \in M_3 : A \text{ es cuadrado mágico}\}$.

Veremos que 3 valores determinan un cuadrado mágico:

$$A = \begin{bmatrix} \beta & \alpha_{12} & \delta \\ \alpha_{21} & \gamma & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix}$$

$$\begin{matrix} \beta + \alpha_{12} + \delta = S \\ \alpha_{12} + \gamma + \alpha_{32} = S \end{matrix} \Rightarrow \alpha_{32} = \beta - \gamma + \delta$$

$$\begin{matrix} \beta + \alpha_{21} + \alpha_{31} = S \\ \delta + \gamma + \alpha_{31} = S \end{matrix} \Rightarrow \alpha_{21} = -\beta + \gamma + \delta$$

$$\begin{matrix} \delta + \alpha_{23} + \alpha_{33} = S \\ \beta + \gamma + \alpha_{33} = S \end{matrix} \Rightarrow \alpha_{23} = \beta + \gamma - \delta$$

$$\therefore \quad S = \alpha_{21} + \gamma + \alpha_{23} = -\beta + \gamma + \delta + \gamma + \beta + \gamma - \delta$$
$$= 3\gamma$$

$$\therefore \quad \beta + \alpha_{12} + \delta = s$$

$$\therefore \quad \alpha_{12} = s - \beta - \delta = 3\gamma - \beta - \delta$$

$$\gamma_{31} = s - \gamma - \delta = 2\gamma - \delta$$

$$\alpha_{33} = s - \beta - \gamma = 2\gamma - \beta$$

$$\therefore \quad A = \begin{bmatrix} \beta & 3\gamma - \beta - \delta & \delta \\ -\beta + \gamma + \delta & \gamma & \beta + \gamma - \delta \\ 2\gamma - \delta & \beta - \gamma + \delta & 2\gamma - \beta \end{bmatrix}$$

$$\therefore \quad A = \beta \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 3 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 2 \end{bmatrix} + \delta \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix}$$

Por lo tanto $\dim V = 3$ y una base de V es

$$\left\{ \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 3 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 2 \end{bmatrix}, \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} \right\} =$$

$= \{ A_1, A_2, A_3 \}$. $\{ A_1, A_2, A_3 \}$ es una base de

V ya que $\{A_1, A_2, A_3\}$ genera V y además es linealmente independiente.

Checaremos que $\{A_1, A_2, A_3\}$ es independiente.

Supongamos $\beta A_1 + \gamma A_2 + \delta A_3 = 0$, entonces

$$\beta \begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} + \gamma \begin{bmatrix} 0 & 3 & 0 \\ 1 & 1 & 1 \\ 2 & -1 & 2 \end{bmatrix} + \delta \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} = 0 .$$

$$\Rightarrow \begin{bmatrix} \beta & 3\gamma-\beta-\delta & \delta \\ -\beta+\gamma+\delta & \gamma & \beta+\gamma-\delta \\ 2\gamma-\delta & \beta-\gamma+\delta & 2\gamma-\beta \end{bmatrix} = 0$$

$$\Rightarrow \quad \beta = \gamma = \delta = 0$$

$\therefore$ $\{A_1, A_2, A_3\}$ es linealmente independiente

y por lo tanto

$\{A_1, A_2, A_3\}$ es una base de V.

4) Sea $V = \{ \{s_k\} : a_1 s_k + a_2 s_{k-1} = 0 \}$.

Entonces, cada sucesión en V queda determinada por un número :

ya que si $a_1, a_2 \neq 0$ entonces

$$s_k = \left(-\frac{a_2}{a_1}\right) s_{k-1}$$

Por lo tanto $\dim(V) = 1$, y una base de V está dada por:

$$\left\{ \left\{ 1, \left(-\frac{a_2}{a_1}\right), \left(-\frac{a_2}{a_1}\right)^2, \left(-\frac{a_2}{a_1}\right)^3, \dots, \left(-\frac{a_2}{a_1}\right)^k, \dots \right\} \right\}$$

5) Sea $V = \{ \{s_k\} : a_0 s_k + a_1 s_{k-1} + \cdots + a_r s_{k-r} = 0 \}$

sup. $a_0, a_1, \dots, a_r \neq 0$. entonces

$$s_k = -\frac{a_1}{a_0} s_{k-1} - \frac{a_2}{a_0} s_{k-2} - \cdots - \frac{a_r}{a_0} s_{k-r}$$

∴ Cada sucesión en V está determinada por los primeros r elementos de la sucesión, y por lo tanto $\dim V = r$.

Como se podrá observar en los ejemplos, el número de parámetros nos da $\dim(V)$.

6) Sea $\mathbb{V} = \{ y : \mathbb{R} \to \mathbb{R} : y$ es diferenciable dos veces y $a_0 y''(\tau) + a_1 y'(\tau) + a_2 y(\tau) = 0 \}$

Veremos que $\dim(\mathbb{V}) = 2$.

Supongamos que tenemos la ecuación diferencial

$$a_0 y''(\tau) + a_1 y'(\tau) + a_2 y(\tau) = 0 \quad \cdots \quad (*)$$

donde $a_0, a_1, a_2 \in \mathbb{R}$.

Buscaremos soluciones de $(*)$ de la forma $y = e^{r\tau}$.

Supongamos $y = e^{r\tau}$ es solución de $(*)$, entonces

$$a_0 (e^{r\tau})'' + a_1 (e^{r\tau})' + a_2 (e^{r\tau}) = 0$$

$$\Rightarrow \quad e^{r\tau} (a_0 r^2 + a_1 r + a_2) = 0$$

$$\Rightarrow \quad a_0 r^2 + a_1 r + a_2 = 0 \quad \cdots \quad (**)$$

Por lo tanto $e^{r\tau}$ es solución de $(*)$ $\Leftrightarrow$ r es solución de $(**)$.

Supongamos que $(**)$ tiene raíces reales r_1, r_2, con $r_1 \neq r_2$.

Entonces $e^{r_1\tau}$ y $e^{r_2\tau}$ son soluciones de (*).

Veremos que $\{e^{r_1\tau}, e^{r_2\tau}\}$ es una base del espacio de soluciones de (*):

$\{e^{r_1\tau}, e^{r_2\tau}\}$ es independiente, ya que si

$$\alpha e^{r_1\tau} + \beta e^{r_2\tau} = 0$$

$$\Rightarrow \quad (\alpha e^{r_1\tau} + \beta e^{r_2\tau})' = 0$$

$$\Rightarrow \quad \alpha r_1 e^{r_1\tau} + \beta r_2 e^{r_2\tau} = 0$$

$$\Rightarrow \begin{cases} \alpha e^{r_1\tau} + \beta e^{r_2\tau} = 0 \\ \alpha r_1 e^{r_1\tau} + \beta r_2 e^{r_2\tau} = 0 \end{cases} \qquad \forall \tau$$

$$\Rightarrow \begin{cases} \alpha + \beta = 0 \\ \alpha r_1 + \beta r_2 = 0 \end{cases}$$

Si $r_1 = 0$ ó $r_2 = 0$ $\Rightarrow$ $\alpha = \beta = 0$

Si $r_1 \neq 0$ y $r_2 \neq 0$ entonces tenemos que

$$\begin{cases} \alpha + \beta = 0 \\ \beta(r_1 - r_2) = 0 \end{cases}$$

Como $r_1 - r_2 \neq 0$ $\Rightarrow$ $\beta = 0$ $\Rightarrow$ $\alpha = 0$

$\therefore$ $\alpha = \beta = 0$ $\therefore$ $\{e^{r_1\tau}, e^{r_2\tau}\}$ es independiente.

Ahora veremos que $\{e^{r_1\tau}, e^{r_2\tau}\}$ genera al espacio de soluciones de $(*)$.

Sea $y(\tau)$ cualquier solución de $(*)$.

Supongamos $y(\tau_0) = y_0$ $y'(\tau_0) = y'_0$, y

supongamos $y(\tau) = \alpha_1 e^{r_1\tau} + \alpha_2 e^{r_2\tau}$, entonces

$$\begin{cases} \alpha_1 e^{r_1\tau_0} + \alpha_2 e^{r_2\tau_0} = y_0 \\ \alpha_1 r_1 e^{r_1\tau_0} + \alpha_2 r_2 e^{r_2\tau_0} = y'_0 \end{cases}$$

$\Rightarrow$ $\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}$ es solución del sistema

$$Aa = \begin{bmatrix} e^{r_1\tau_0} & e^{r_2\tau_0} \\ r_1 e^{r_1\tau_0} & r_2 e^{r_2\tau_0} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} y_0 \\ y'_0 \end{bmatrix}$$

Por otro lado
Este sistema tiene solución, ya que como $r_1 \neq r_2$, entonces la matriz A es no-singular. (y además, la solución es única)

Supongamos que $\varphi(\tau) = \alpha_1 e^{r_1\tau} + \alpha_2 e^{r_2\tau}$

$\therefore$ $\varphi(\tau)$ es solución de $(*)$, ya que $\varphi(\tau)$ es una combinación lineal de $e^{r_1\tau}$ y $e^{r_2\tau}$

Por otro lado $\varphi(\tau_0) = y_0$ y
$\varphi'(\tau_0) = y_0'$

Pero teníamos que $y(\tau)$ era solución de $(*)$ y
$y(\tau_0) = y_0$
$y'(\tau_0) = y_0'$

En los cursos de ecuaciones diferenciales se demuestra que existe una única solución $y(\tau)$ de la ecuación $(*)$ tal que $y(\tau_0) = y_0$, $y'(\tau_0) = y_0'$. Por lo tanto

$$\varphi(\tau) = y(\tau)$$

$\therefore$ $y(\tau) = \alpha_1 e^{r_1\tau} + \alpha_2 e^{r_2\tau}$

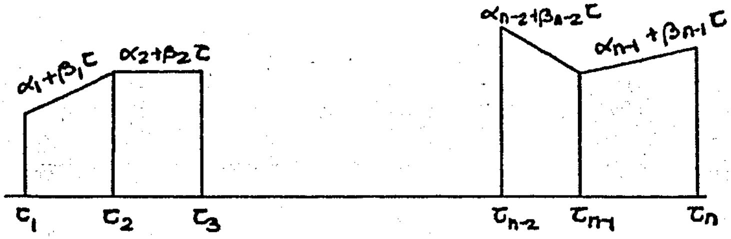
$\therefore$ $\{e^{r_1\tau}, e^{r_2\tau}\}$ genera al espacio de soluciones de $(*)$.

$\therefore$ $\{e^{r_1\tau}, e^{r_2\tau}\}$ es una base del espacio de soluciones de $(*)$, y por lo tanto $\dim[V] = 2$

7) Calcularemos la dimensión de

$\mathcal{S}_1 = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado ≤ 1

y f es continua en $I \}$



En cada intervalo I_i la función es de la forma

$$\alpha_i + \beta_i \tau$$

Por lo tanto, para que la función sea continua en I necesitamos que:

$$\alpha_1 + \beta_1 \tau_2 = \alpha_2 \beta_2 \tau_2$$
$$\alpha_2 + \beta_2 \tau_3 = \alpha_3 + \beta_3 \tau_3$$
$$\vdots$$
$$\alpha_{n-3} + \beta_{n-3} \tau_{n-2} = \alpha_{n-2} + \beta_{n-2} \tau_{n-2}$$
$$\alpha_{n-2} + \beta_{n-2} \tau_{n-1} = \alpha_{n-1} + \beta_{n-1} \tau_{n-1}$$

Es decir, $\alpha_1, \beta_1, \alpha_2, \beta_2, \ldots, \alpha_{n-1}, \beta_{n-1}$ tienen que ser tales que

$$Ax = 0, \quad \text{donde}$$

$$A = \begin{bmatrix} 1 & \tau_2 & -1 & -\tau_2 & & & & & & & \\ 0 & 0 & 1 & \tau_3 & -1 & -\tau_3 & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & \tau_{n-1} & -1 & -\tau_{n-1} \end{bmatrix}$$

$$\text{y} \quad x^t = \begin{bmatrix} \alpha_1 & \beta_1 & \alpha_2 & \beta_2 & \cdots & \alpha_{n-1} & \beta_{n-1} \end{bmatrix}$$

Por lo tanto una función $f \in \mathcal{S}_1 \Leftrightarrow$

$$f|_{I_i} = f_i = \alpha_i + \beta_i \tau \qquad (1 \leq i \leq n-1)$$

$$\Leftrightarrow \quad Ax = 0$$

Ahora, es fácil ver que existe una correspondencia biunívoca entre $\mathcal{S}_1$ y $N(A)$, y por lo tanto $\dim(N(A)) = \dim(\mathcal{S}_1)$.

A es una matriz $(n-2) \times [2(n-1)]$, de rango $r(A) = n-2$, $\therefore$ $\dim(\mathcal{S}_1) = \dim(N(A)) = 2n-2-(n-2)$

$\therefore$ $\dim(\mathcal{S}_1) = n =$ # de ptos. de la partición

8) Calcularemos la dimensión de

$$\mathcal{S}_m = \{ f : I \to \mathbb{R} : f|_{I_i} \text{ es un polinomio de grado} \leq m \text{ y } f \text{ es continua en } I \}$$

En cada intervalo I_i la función es de la forma

$$\alpha_0^{(i)} + \alpha_1^{(i)} \tau + \alpha_2^{(i)} \tau^2 + \cdots + \alpha_m^{(i)} \tau^m$$

Para que f sea continua en I se deben de cumplir las siguientes condiciones.

$$\alpha_0^{(1)} + \alpha_1^{(1)} \tau_2 + \alpha_2^{(1)} \tau_2^2 + \cdots + \alpha_m^{(1)} \tau_2^m = \alpha_0^{(2)} + \alpha_1^{(2)} \tau + \cdots + \alpha_m^{(2)} \tau_2^m$$

$$\alpha_0^{(2)} + \alpha_1^{(2)} \tau_5 + \alpha_2^{(2)} \tau_3^2 + \cdots + \alpha_m^{(2)} \tau_3^m = \alpha_0^{(3)} + \alpha_1^{(3)} \tau_3 + \alpha_2^{(3)} \tau_3^2 + \cdots + \alpha_m^{(3)} \tau_3^m$$

$$\vdots \qquad\qquad\qquad\qquad \vdots \qquad\qquad\qquad\qquad \vdots$$

$$\alpha_0^{(n-2)} + \alpha_1^{(n-2)} \tau_{n-1} + \alpha_2^{(n-2)} \tau_{n-1}^2 + \cdots + \alpha_m^{(n-2)} \tau_{n-1}^m = \alpha_0^{(n-1)} + \alpha_1^{(n-1)} \tau_{n-1} + \alpha_2^{(n-1)} \tau_{n-1}^{(2)} + \cdots + \alpha_m^{(n-1)} \tau_{n-1}^m$$

Al igual que en el ejemplo (7), estas $n-2$ condiciones los podemos escribir en lenguaje matricial

$$A \cdot X = 0$$

donde $x^t = [x_1^t \mid x_2^t \mid \dots \mid x_{n-1}^t]$ con

$$
\begin{aligned}
x_1^t &= \left[\alpha_0^{(1)} \quad \alpha_1^{(1)} \quad \alpha_2^{(1)} \quad \dots \quad \alpha_m^{(1)}\right] \\
x_2^t &= \left[\alpha_0^{(2)} \quad \alpha_1^{(2)} \quad \alpha_2^{(2)} \quad \dots \quad \alpha_m^{(2)}\right] \\
&\ \ \vdots \\
x_{n-1}^t &= \left[\alpha_0^{(n-1)} \quad \alpha_1^{(n-1)} \quad \alpha_2^{(n-1)} \quad \dots \quad \alpha_m^{(n-1)}\right]
\end{aligned}
$$

y donde $A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_{n-2}^t \end{bmatrix}$ con

$$
r_1^t = \left[\begin{array}{c|c|c} m+1 & m+1 & (m+1)(n-3) \\ t_2^t & -t_2^t & 0 \end{array}\right]
$$

$$
r_2^t = \left[\begin{array}{c|c|c|c} m+1 & m+1 & m+1 & (m+1)(n-4) \\ 0 & t_3^t & -t_3^t & 0 \end{array}\right]
$$

$$
r_3^t = \left[\begin{array}{c|c|c|c|c} m+1 & m+1 & m+1 & m+1 & (m+1)(n-5) \\ 0 & 0 & t_4^t & -t_4^t & 0 \end{array}\right]
$$

$$
\vdots \qquad\qquad \vdots
$$

$$
r_{n-2}^t = \left[\begin{array}{c|c|c|c|c|c|c|c} m+1 & m+1 & m+1 & m+1 & m+1 & \cdots & m+1 & m+1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & t_{n-1}^t & -t_{n-1}^t \end{array}\right]
$$

donde $t_i^t = \left[1 \quad \tau_i \quad \tau_i^2 \quad \cdots \quad \tau_i^m\right]$

y $\dim(N(A)) = \dim(\mathcal{S}_m)$.

Como se puede ver, los renglones de A son independientes $\therefore$ $r(A) = n-2$. Como A es una matriz $(n-2) \times (m+1)(n-1)$

$$\therefore \dim(N(A)) = (m+1)(n-1) - (n-2)$$
$$= m(n-1) + 1$$

$$\therefore \dim(\mathcal{S}_m) = m(n-1) + 1.$$

9) Calcularemos la dimensión de

$$\mathcal{S}_{m,1} = \{ f : I \to \mathbb{R} : f|_{I_i} \text{ es un polinomio de grado} \leq m \text{ y } f \text{ es } C^1 \text{ en } I \}.$$

En cada intervalo I_i la función es de la forma

$$\alpha_0^{(i)} + \alpha_1^{(i)} \tau + \alpha_2^{(i)} \tau^2 + \cdots + \alpha_m^{(i)} \tau^m$$

Para que f sea continua se tienen que cumplir las condiciones que impusimos en el ejemplo (8).

La derivada de f está dada por:

$$f'(\tau) = \alpha_1^{(i)} + 2\alpha_2^{(i)} \tau + 3\alpha_3^{(i)} \tau^2 + \cdots + m\alpha_m^{(i)} \tau^{m-1}, \text{ si } \tau \in I_i$$

Para que $f'(\tau)$ sea continua en I se necesitan cumplir:

$$\alpha_1^{(1)} + 2\alpha_2^{(1)}\tau_2 + 3\alpha_3^{(1)}\tau_2^2 + \cdots + m\alpha_m^{(1)}\tau_2^{m-1} = \alpha_1^{(2)} + 2\alpha_2^{(2)}\tau_2 + 3\alpha_3^{(2)}\tau_2^2 + \cdots + m\alpha_m^{(2)}\tau_2^{m-}$$

$$\alpha_1^{(2)} + 2\alpha_2^{(2)}\tau_3 + 3\alpha_3^{(2)}\tau_3^2 + \cdots + m\alpha_m^{(2)}\tau_3^{m-1} = \alpha_1^{(3)} + 2\alpha_2^{(3)}\tau_3 + 3\alpha_3^{(3)}\tau_3^2 + \cdots + m\alpha_m^{(3)}\tau_3^{m-}$$

$$\vdots \qquad\qquad \vdots$$

$$\alpha_1^{(n-2)} + 2\alpha_2^{(n-2)}\tau_{n-1} + 3\alpha_3^{(n-2)}\tau_{n-1}^2 + \cdots + m\alpha_m^{(n-2)}\tau_{n-1}^{m-1} = \alpha_1^{(n-1)} + 2\alpha_2^{(n-1)}\tau_{n-1} + 3\alpha_3^{(n-1)}\tau_{n-1}^2 + \cdots + m\alpha_m^{(n-1)}\tau_{n-1}^{m-}$$

Sean

$$s_1^t = \left[\begin{array}{c|c|c} m+1 & m+1 & (m+1)(n-3) \\ z_2^t & -z_2^t & 0 \end{array}\right]$$

$$s_2^t = \left[\begin{array}{c|c|c|c} m+1 & m+1 & m+1 & (m+1)(n-4) \\ 0 & z_3^t & -z_3^t & 0 \end{array}\right]$$

$$s_3^t = \left[\begin{array}{c|c|c|c|c} m+1 & m+1 & m+1 & m+1 & (m+1)(n-5) \\ 0 & 0 & z_4 & -z_4^t & 0 \end{array}\right]$$

$$\vdots$$

$$s_{n-2}^t = \left[\begin{array}{c|c|c|c|c|c|c|c} m+1 & m+1 & m+1 & m+1 & m+1 & & m+1 & m+1 \\ 0 & 0 & 0 & 0 & 0 & \cdots & z_{n-1}^t & -z_{n-1}^t \end{array}\right.$$

donde $z_i^t = [0 \;\; 1 \;\; 2\tau_i \;\; 3\tau_i^2 \;\; \cdots \;\; m\tau_i^{m-1}]$

Los condiciones para que la función f y su primera derivada f' sean continuas se pueden escribir en notación matricial de la siguiente manera:

$Ax = 0$

donde $x \in \mathbb{R}^{(m+1)(n-1)}$ es como en el ejemplo (8)
y

$$A = \begin{bmatrix} r_1^t \\ s_1^t \\ r_2^t \\ s_2^t \\ \vdots \\ r_{n-2}^t \\ s_{n-2}^t \end{bmatrix}$$

Entonces $\dim(\mathcal{S}_{m,1}) = \dim(N(A))$

Como se puede ver facilmente, A es una matriz escalonada con renglones independientes.

Por lo tanto $\dim(N(A)) = (m+1)(n-1) - 2(n-2)$
$= m(n-1) - (n-2) + 1$

Por lo tanto $\dim(\mathcal{S}_{m,1}) = m(n-1) - (n-2) + 1$

10) Calcularemos la dimensión de

$\mathcal{S}_{m,k} = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$, y f es C^k en $I \}$

Siguiendo un procedimiento análogo al usado en los ejemplos (7) - (9)

$$\dim(\mathcal{S}_{m,k}) = \#\text{ variables} - \#\text{ condiciones}$$

El número de variables es el mismo que en los ejemplos (7) y (9) : $(m+1)(n-1)$, ya que tenemos $(n-1)$ intervalos, y en cada intervalo la función está determinada por $(m+1)$ parámetros.

La continuidad de $f, f', f^{(2)}, \ldots, f^{(k)}$ está dada por $(n-2)(k+1)$ condiciones, que determinan una matriz escalonada con renglones independientes.

Por lo tanto

$$\begin{aligned}\dim(\mathcal{S}_{m,k}) &= (m+1)(n-1) - (n-2)(k+1) \\ &= m(n-1) + n-1 - (n-2) - k(n-2) \\ &= m(n-1) - k(n-2) + 1\end{aligned}$$

$$\therefore \quad \dim(\mathcal{S}_{m,k}) = m(n-1) - k(n-2) + 1$$

$\square$

DEFINICION

Sean V y W dos espacios vectoriales (reales). Una TRANSFORMACION LINEAL de V en W es una función $\tilde{T}: V \to W$ tal que

$$\tilde{T}(\alpha x + y) = \alpha \tilde{T}(x) + \tilde{T}(y)$$

para todo $x, y \in V$, $\alpha \in \mathbb{R}$.

PROPIEDADES

Sea $\tilde{T}: V \to W$ una transformación lineal. Entonces

1) Si $S \subseteq V$ es un subespacio $\Rightarrow \tilde{T}(S) \subseteq W$ es un subespacio.

Demostración:

Supongamos $S \subseteq V$ es un subespacio de V

Sean $\tilde{T}(s_1), \tilde{T}(s_2) \in \tilde{T}(S)$

$\therefore \quad \tilde{T}(s_1) + \tilde{T}(s_2) = \tilde{T}(s_1 + s_2) \qquad (\tilde{T} \text{ lineal})$

$\therefore \quad \tilde{T}(s_1) + \tilde{T}(s_2) \in \tilde{T}(S) \qquad (S \text{ subespacio})$

y $\alpha \tilde{T}(s) = \tilde{T}(\alpha s) \in \tilde{T}(S)$

Por lo tanto $\tilde{T}(\mathcal{S}) \subseteq W$ es subespacio de W.

2) Si $\tilde{\mathfrak{F}} \subseteq W$ es subespacio $\Rightarrow$

$\tilde{T}^{-1}(\tilde{\mathfrak{F}}) = \{v \in V : \tilde{T}v \in \tilde{\mathfrak{F}}\}$ es un subespacio.

Demostración:

Sean $v_1, v_2 \in \tilde{T}^{-1}(\tilde{\mathfrak{F}}) \Rightarrow \tilde{T}(v_1), \tilde{T}(v_2) \in \tilde{\mathfrak{F}}$

$\therefore \tilde{T}(v_1) + \tilde{T}(v_2) \in \tilde{\mathfrak{F}}$ $\quad$ ($\tilde{\mathfrak{F}}$ subespacio)

$\therefore \tilde{T}(v_1 + v_2) \in \tilde{\mathfrak{F}}$ $\quad$ ($\tilde{T}$ lineal)

$\therefore v_1 + v_2 \in \tilde{T}^{-1}(\tilde{\mathfrak{F}})$

Sea $v \in \tilde{T}^{-1}(\tilde{\mathfrak{F}}) \Rightarrow \tilde{T}(v) \in \tilde{\mathfrak{F}}$

$\therefore \alpha \tilde{T}(v) \in \tilde{\mathfrak{F}}$ $\quad$ ($\tilde{\mathfrak{F}}$ subespacio)

$\therefore \tilde{T}(\alpha v) \in \tilde{\mathfrak{F}}$ $\quad$ ($\tilde{T}$ lineal)

$\therefore \alpha v \in \tilde{T}^{-1}(\tilde{\mathfrak{F}})$

Por lo tanto $\tilde{T}^{-1}(\tilde{\mathfrak{F}})$ es un subespacio de V.

3) Si $\mathcal{S} \subseteq V$ es un subespacio de V entonces $\dim(\tilde{T}(\mathcal{S})) \leq \dim(\mathcal{S})$.

Demostración :

Sea $\{\tilde{T}(s_1), \tilde{T}(s_2), \dots, \tilde{T}(s_K)\}$ una base de $\tilde{T}(\mathcal{S})$.

Supongamos $\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_K s_K = 0$

$\Rightarrow \tilde{T}(\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_k s_k) = 0$

$\Rightarrow \alpha_1 \tilde{T}(s_1) + \alpha_2 \tilde{T}(s_2) + \cdots + \alpha_k \tilde{T}(s_k) = 0$

$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$

$\Rightarrow \{s_1, s_2, \dots, s_K\} \subseteq \mathcal{S}$ es linealmente independiente.

$\therefore \quad \dim(\mathcal{S}) \geq K = \dim(\tilde{T}(\mathcal{S}))$

$\therefore \quad \dim(\tilde{T}(\mathcal{S})) \leq \dim(\mathcal{S})$

$\square$

4) Si $\tilde{T}$ es inyectiva entonces $\dim(\mathcal{S}) = \dim(\tilde{T}(\mathcal{S}))$, si $\mathcal{S} \subseteq V$ subespacio.

Demostración

Supongamos $\{s_1, s_2, \dots, s_K\}$ es una base de $\mathcal{S}$, entonces

$$\{\tilde{T}(s_1), \tilde{T}(s_2), \ldots, \tilde{T}(s_k)\} \text{ es una base de } \tilde{T}(\mathcal{S}),$$

ya que si

$$\alpha_1 \tilde{T}(s_1) + \alpha_2 \tilde{T}(s_2) + \cdots + \alpha_k \tilde{T}(s_k) = 0$$

$$\Rightarrow \tilde{T}(\alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_k s_k) = 0$$

$$\Rightarrow \alpha_1 s_1 + \alpha_2 s_2 + \cdots + \alpha_k s_k = 0 \quad (\tilde{T} \text{ 1-1})$$

$$\Rightarrow \alpha_1 = \alpha_2 = \cdots = \alpha_k = 0 \quad (\{s_1, s_2, \ldots, s_k\} \text{ independiente})$$

$\therefore$ $\{\tilde{T}(s_1), \tilde{T}(s_2), \ldots, \tilde{T}(s_k)\}$ es linealmente independiente.

$$\therefore \quad \dim(\tilde{T}(\mathcal{S})) \geq k = \dim(\mathcal{S})$$

pero, por el inciso anterior

$$\dim(\tilde{T}(\mathcal{S})) \leq \dim(\mathcal{S})$$

$$\therefore \quad \dim(\tilde{T}(\mathcal{S})) = \dim(\mathcal{S}) \quad \square$$

En particular, para $\mathcal{S} = \mathbb{V}$, vamos a tener

$$\dim(\mathbb{V}) = \dim(\tilde{T}(\mathbb{V})) \leq \dim(\mathbb{W})$$

cuando $\tilde{T}$ es inyectiva.

$\blacksquare$

5) Si $\tilde{T}$ es sobre entonces $\dim(\mathcal{W}) \leq \dim(\mathcal{V})$

Demostración:

Si $\tilde{T}$ es sobre $\Rightarrow T(\mathcal{V}) = \mathcal{W}$

Por lo tanto, aplicando el resultado (3) con $\mathcal{S} = \mathcal{V}$:

$$\dim(\tilde{T}(\mathcal{V})) = \dim(\mathcal{W}) \leq \dim(\mathcal{V})$$

$\blacksquare$

DEFINICION

Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales, y sea $\tilde{T}: \mathbb{V} \to \mathbb{W}$ una transformación lineal. Entonces el NUCLEO ó KERNEL de $\tilde{T}$, denotado por $N(\tilde{T})$ ó por $\operatorname{Ker}(\tilde{T})$, es

$$N(\tilde{T}) = \{ v \in \mathbb{V} : \tilde{T}(v) = 0 \in \mathbb{W} \}$$

RESULTADO

Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales, y sea $\tilde{T}: \mathbb{V} \to \mathbb{W}$ lineal. Entonces $N(\tilde{T})$ es un subespacio de $\mathbb{V}$.

Demostración :

Si $v \in N(\tilde{T})$ entonces $\tilde{T}(\alpha v) = \alpha \tilde{T}(v) = 0$ $\forall \alpha \in \mathbb{R}$ $\therefore$ $\alpha v \in N(\tilde{T})$

Si $v_1, v_2 \in N(\tilde{T})$ entonces

$$\tilde{T}(v_1 + v_2) = \tilde{T}(v_1) + \tilde{T}(v_2) = 0$$

$\therefore$ $v_1 + v_2 \in N(\tilde{T})$

$\therefore$ $N(\tilde{T}) \subseteq \mathbb{V}$ es un subespacio de $\mathbb{V}$

DEFINICION

Sean V y W dos espacios vectoriales. Una transformación $\tilde{T}: V \to W$ es un ISOMORFISMO de V sobre W si

1) $\tilde{T}$ es lineal
2) $\tilde{T}$ es biyectiva.

Si existe un isomorfismo $\tilde{T}: V \to W$ entonces decimos que V y W son ISOMORFOS, y escribimos $V \cong W$.

RESULTADO

Sea V un espacio vectorial tal que $\dim(V) = m$. Entonces $V \cong \mathbb{R}^m$.

Demostración

Sea $\{v_1, v_2, \ldots, v_m\}$ una base de V.

Por lo tanto si $v \in V \Rightarrow v = \sum_{i=1}^{m} \alpha_i v_i$

Sea $\tilde{T}: V \to \mathbb{R}^m$ dada por :

$$\tilde{T}(V) = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$

Entonces $\tilde{T}$ es un isomorfismo :

$\tilde{T}$ está bien definida ya que la expresión

$$V = \sum_{i=1}^{m} \alpha_i v_i$$

para los vectores de V es única.

$\tilde{T}$ es sobre, ya que si $y = [\eta_i] \in \mathbb{R}^m$

$$\Rightarrow \quad y = \tilde{T}(x), \quad \text{donde} \quad x = \sum_{i=1}^{m} \eta_i v_i \quad .$$

$\tilde{T}$ es inyectiva, ya que si $\tilde{T}(x_1) = \tilde{T}(x_2)$ entonces

$$\text{si} \quad x_1 = \sum_{i=1}^{m} \xi_i^{(1)} v_i \qquad x_2 = \sum_{i=1}^{m} \xi_i^{(2)} v_i$$

y por lo tanto $\xi_i^{(1)} = \xi_i^{(2)} \qquad 1 \leq i \leq m$

$$\therefore \quad x_1 = x_2$$

$\tilde{T}$ es lineal :

$$\tilde{T}(\alpha x_1 + x_2) = \tilde{T}\left(\sum_{i=1}^{m} \alpha \xi_i^{(1)} v_i + \sum_{i=1}^{m} \xi_i^{(2)} v_i \right)$$

$$= \tilde{T}\left(\sum_{i=1}^{m} (\alpha \xi_i^{(1)} + \xi_i^{(2)}) v_i \right)$$

$$= \left[\alpha \xi_i^{(1)} + \xi_i^{(2)} \right]$$

$$= \left[\alpha \xi_i^{(1)} \right] + \left[\xi_i^{(2)} \right]$$

$$= \alpha \left[\xi_i^{(1)} \right] + \left[\xi_i^{(2)} \right]$$

$$= \alpha \tilde{T}(x_1) + \tilde{T}(x_2)$$

Recordar que $[\xi_i] = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_m \end{bmatrix}$

$\therefore \quad \tilde{T} : V \to \mathbb{R}^m$ es un isomorfismo

$\therefore \quad V \cong \mathbb{R}^m$

∎

Sean V y W espacios vectoriales reales.

Sea $L(V,W) = \{\tilde{T} : V \to W \;\because\; \tilde{T}$ es lineal $\}$.

Si $\tilde{T}, \tilde{U} \in L(V,W)$ entonces definimos

$\tilde{T}+\tilde{U}$ y $\alpha\tilde{T}$ como las transformaciones dadas por

$$(\tilde{T}+\tilde{U})(v) = \tilde{T}(v) + \tilde{U}(v)$$

$$(\alpha\tilde{T})(v) = \alpha\tilde{T}(v)$$

Por lo tanto $\tilde{T}+\tilde{U}$ y $\alpha\tilde{T}$ son transformaciones de V en W. Veremos que $\tilde{T}+\tilde{U}$ y $\alpha\tilde{T}$ son transformaciones lineales de V en W.

$$\begin{aligned}(\tilde{T}+\tilde{U})(\alpha v_1+v_2) &= \tilde{T}(\alpha v_1+v_2) + \tilde{U}(\alpha v_1+v_2)\\ &= \alpha\tilde{T}(v_1)+\tilde{T}(v_2) + \alpha\tilde{U}(v_1)+\tilde{U}(v_2)\\ &= \alpha(\tilde{T}(v_1)+\tilde{U}(v_1)) + \tilde{T}(v_2)+\tilde{U}(v_2)\\ &= \alpha(\tilde{T}+\tilde{U})(v_1) + (\tilde{T}+\tilde{U})v_2\end{aligned}$$

$\therefore$ $\tilde{T}+\tilde{U}$ es lineal $\therefore$ $\tilde{T}+\tilde{U} \in L(V,W)$.

$$\begin{aligned}(\alpha\tilde{T})(\beta v_1+v_2) &= \alpha\tilde{T}(\beta v_1) + \alpha\tilde{T}(v_2)\\ &= \beta(\alpha\tilde{T})(v_1) + (\alpha\tilde{T})(v_2)\end{aligned}$$

$\therefore$ $\alpha\tilde{T}$ es lineal y $\therefore$ $\alpha\tilde{T} \in L(V,W)$.

Se puede checar que L(V,W), con las operaciones de suma de transformaciones y producto de una transformación por un escalar, es un espacio vectorial.

DEFINICION

Sea V un espacio vectorial de dimensión m. Un SISTEMA DE COORDENADAS en V es un isomorfismo de V sobre $\mathbb{R}^m$.

OBSERVACION

Supongamos que V es un espacio vectorial de dimensión m. Entonces una vez escogida una base ordenada $\{v_1, v_2, \dots, v_m\}$ de V, un sistema de coordenadas en V queda unívocamente determinado por la correspondencia

$$v_i \mapsto e_i \qquad 1 \leq i \leq m$$

donde $\{e_1, e_2, \dots, e_n\}$ es la base canónica de $\mathbb{R}^m$.

Bajo esta correspondencia

$$v = \sum_{i=1}^{n} \xi_i v_i \longmapsto x = [\xi_i]$$

ESCOGER UNA BASE ORDENADA EN V ⇒ ESCOGER UN SISTEMA DE COORDENADAS.

DEFINICION

Sea $\mathbb{V}$ un espacio vectorial de dimensión m. Sea $\mathcal{B}=\{v_1,v_2,\dots,v_m\}$ una base de $\mathbb{V}$. Por lo tanto el sistema de coordenadas determinado por la base es la función $c:\mathbb{V}\to\mathbb{R}^m$ dada por

$$c(v)=[\xi_i] \qquad \text{si} \qquad v=\sum_{i=1}^{m}\xi_i v_i$$

Decimos que $\xi_1,\xi_2,\dots,\xi_m$ son las COORDENADAS DEL VECTOR v CON RESPECTO A LA BASE ORDENADA

Escribiremos $v=[\xi_i]_{\mathcal{B}}$ para indicar que las coordenadas del vector v con respecto a la base $\mathcal{B}$ son $\xi_1,\xi_2,\dots,\xi_m$. A veces identificaremos al vector $v\in\mathbb{V}$ con el vector de coordenadas de v, $x=[\xi_i]\in\mathbb{R}^m$.

RESULTADO

Sean $\mathbb{V}$ y $\mathbb{W}$ espacios vectoriales de dimensión n y m respectivamente. Entonces $L(\mathbb{V},\mathbb{W})$ es isomorfo a $M(m,n)$.

Demostración:

Daremos un isomorfismo de $L(\mathbb{V},\mathbb{W})$ en $M(m,n)$: Por el resultado anterior $\mathbb{V}\cong\mathbb{R}^n$ y $\mathbb{W}\cong\mathbb{R}^m$.

Sean $c_I : V \to \mathbb{R}^n$ y $S_I : W \to \mathbb{R}^m$ sistemas de coordenadas en V y W respectivamente.

Sea $\tilde{T} \in L(V,W)$. Entonces en el diagrama observamos que podemos definir una $T : \mathbb{R}^n \to \mathbb{R}^m$ por medio de $T = S_I \circ \tilde{T} \circ c_I^{-1}$.

$$\begin{tikzcd}
V \arrow[r, "\tilde{T}"] \arrow[d, "c_I"'] & W \arrow[d, "S_I"] \\
\mathbb{R}^n \arrow[r, dashed, "T"'] & \mathbb{R}^m
\end{tikzcd}$$

Sea $\mathcal{F} : L(V,W) \to M(m,n)$ dada por $\mathcal{F}(\tilde{T}) = S_I \circ \tilde{T} \circ c_I^{-1} = T \in M(m,n)$. Entonces

$\mathcal{F}$ es biyectiva:

Si $\tilde{T}_1$ y $\tilde{T}_2 \in L(V,W)$ entonces

$$\mathcal{F}(\tilde{T}_1) = \mathcal{F}(\tilde{T}_2) \Rightarrow S_I \circ \tilde{T}_1 \circ c_I^{-1} = S_I \circ \tilde{T}_2 \circ c_I^{-1} \Rightarrow \tilde{T}_1 = \tilde{T}_2$$

(ya que S_I y c_I son biyectivas)

$\mathcal{F}$ es sobre:

Sea $T \in M(m,n)$, y sea $\tilde{T} \in L(V,W)$ dada por $\tilde{T} = S_I^{-1} \circ T \circ c_I$ $\therefore$ $\mathcal{F}(\tilde{T}) = S_I \circ (S_I^{-1} \circ T \circ c_I) \circ c_I^{-1} = T$.

$\mathcal{F}$ es lineal:

Si $\tilde{T}_1, \tilde{T}_2 \in L(V,W)$ y $\alpha \in \mathbb{R}$, entonces :

$$
\begin{aligned}
F(\alpha\tilde{T}_1+\tilde{T}_2) &= S_I\circ(\alpha\tilde{T}_1+\tilde{T}_2)\circ C_I^{-1}\\
&= S_I\circ\alpha\tilde{T}_1\circ C_I^{-1} + S_I\circ\tilde{T}_2\circ C_I^{-1}\\
&= \alpha\,(S_I\circ\tilde{T}_1\circ C_I^{-1}) + S_I\circ\tilde{T}_2\circ C_I^{-1}\\
&= \alpha\,F(\tilde{T}_1) + F(\tilde{T}_2).
\end{aligned}
$$

$\therefore\ F : L(V,W) \to M(m,n)$ es un isomorfismo

$\therefore\ L(V,W) \cong M(m,n)$ ■

Este resultado es muy importante, ya que si tenemos dos espacios vectoriales de dimensión finita, entonces cada transformación lineal $\tilde{T} : V \to W$ nos determina una matriz $T : \mathbb{R}^n \to \mathbb{R}^m$, y reciprocamente se tenemos una matriz $T : \mathbb{R}^n \to \mathbb{R}^m$, ésta nos determina unívocamente una transformación lineal $\tilde{T} : V \to W$.

OBSERVACION

La matriz asociada a una transformación lineal $\tilde{T} : V \to W$ depende de los sistemas de coordenadas que hayamos elegido.

Supongamos ahora que tenemos una transformación lineal $\tilde{T} : \mathbb{V} \to \mathbb{W}$.

¿ cómo calculamos una matriz $T : \mathbb{R}^n \to \mathbb{R}^m$ asociada a $\tilde{T}$?

Elijamos sistemas de coordenadas en $\mathbb{V}$ y $\mathbb{W}$:

sea $\mathcal{B} = \{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{V}$ y

sea $\mathcal{C} = \{w_1, w_2, \ldots, w_m\}$ una base de $\mathbb{W}$.

Sea $v \in \mathbb{V}$. Por lo tanto $v = \sum_{j=1}^{n} \xi_j v_j$

$$\therefore \tilde{T}(v) = \tilde{T}\left(\sum_{j=1}^{n} \xi_j v_j\right)$$

$$\therefore \tilde{T}(v) = \sum_{j=1}^{n} \xi_j \tilde{T}(v_j) = w \in \mathbb{W}$$

Por otro lado, $\tilde{T}(v_j) \in \mathbb{W} \quad 1 \leq j \leq n$,

Por lo tanto

$$\begin{aligned}
\tilde{T}(v_1) &= \tau_{11} w_1 + \tau_{21} w_2 + \cdots + \tau_{m1} w_m \\
\tilde{T}(v_2) &= \tau_{12} w_1 + \tau_{22} w_2 + \cdots + \tau_{m2} w_m \\
&\vdots \qquad\qquad\qquad\qquad \vdots \\
\tilde{T}(v_n) &= \tau_{1n} w_1 + \tau_{2n} w_2 + \cdots + \tau_{mn} w_m
\end{aligned}$$

Por lo tanto

$$\tilde{T}(v) = \sum_{j=1}^{n} \xi_j \tilde{T}(v_j)$$

$$= \sum_{j=1}^{n} \xi_j \left(\sum_{i=1}^{m} \tau_{ij} w_i \right)$$

$$= \sum_{i=1}^{m} \left(\sum_{j=1}^{n} \xi_j \tau_{ij} \right) w_i = \sum_{i=1}^{n} \eta_i w_i$$

De aquí podemos ver que si el vector de coordenadas de v en la base $\mathcal{B}$ es $[\xi_i]$, entonces el vector de coordenadas de $\tilde{T}v$ en la base $\mathcal{C}$ de W es $[\eta_i] = [\sum_{j=1}^{n} \tau_{ij} \xi_j]$.

Es decir:

$$\begin{bmatrix} \tau_{11} & \tau_{12} & \cdots & \tau_{1n} \\ \tau_{21} & \tau_{22} & \cdots & \tau_{2n} \\ \vdots & & & \\ \tau_{m1} & \tau_{m2} & \cdots & \tau_{mn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_m \end{bmatrix}$$

Por lo tanto la matriz asociada a la transformación $\tilde{T}$ es $T = [\tau_{ij}]$

Ahora, el vector de coordenadas de v_i con respecto a la base ordenada $\{v_1, v_2, \ldots, v_n\}$ de V es $e_i \in \mathbb{R}^n$, y por lo tanto el vector de coordenadas de $\tilde{T}(v_i)$ con respecto a la base $\{w_1, w_2, \ldots, w_m\}$ de W estará dado por Te_i.- la i-ésima columna de la matriz T.

Por lo tanto, la i-ésima columna de la matriz T se obtiene aplicando la transformación $\tilde{T}$ al i-ésimo vector de la base de V, y después expresando $\tilde{T}(v_i)$ en términos de la base de W:

$$Te_i = S_I \cdot \tilde{T}(v_i)$$

Por lo tanto, la matriz asociada a la transformación $\tilde{T}$ está dada por:

$$T = \left[S_I \circ \tilde{T}(v_1) \mid S_I \cdot \tilde{T}(v_2) \mid \cdots \mid S_I \circ \tilde{T}(v_n) \right]$$

Esta discusión nos permite concluir que TODA TRANSFORMACION LINEAL QUEDA DETERMINADA POR SU EFECTO EN UNA BASE.

CAMBIOS DE BASE

Sea V un espacio vectorial de dimensión n. Sean $\mathcal{B}_I$ y $\mathcal{B}_{II}$ dos bases de V, y sean C_I y C_{II} los sistemas de coordenadas asociados a $\mathcal{B}_I$ y $\mathcal{B}_{II}$ respectivamente.

Si las dos bases (ordenadas) son diferentes, entonces las coordenadas de un vector $v \in \mathcal{V}$ con respecto a $\mathcal{B}_I$ serán diferentes de las coordenadas de v con respecto a $\mathcal{B}_{II}$.

$$\begin{tikzcd} \mathcal{V} \arrow[r, dash, "\tilde{I}"] \arrow[d, "c_I"'] & \mathcal{V} \arrow[d, "c_{II}"] \\ \mathbb{R}^n \arrow[r, dash, "A"'] & \mathbb{R}^n \end{tikzcd}$$

Supongamos que c_I es el sistema de coordenadas "viejo" y c_{II} es el sistema de coordenadas "nuevo".

Sea $A : \mathbb{R}^n \to \mathbb{R}^n$ la matriz asociada a la transformación identidad $\tilde{I} : \mathcal{V} \to \mathcal{V}$:

Como se indica en el diagrama anterior, tenemos que

$$A = c_{II} \circ c_I^{-1}$$

Entonces, si $x \in \mathbb{R}^n$ es el vector de coordenadas del vector $v \in \mathcal{V}$ en el sistema c_I, Ax es el vector de coordenadas de v en el sistema c_{II}.

La i-ésima columna de A se obtiene expresando al i-ésimo vector de la base $\mathcal{B}_I$ en términos de los vectores de la base $\mathcal{B}_{II}$.

Observemos además que si

$$x = C_{I}(v) \quad y \quad y = C_{II}(v)$$

entonces $v = C_{I}^{-1}(x)$ y por lo tanto

$$y = C_{II}(C_{I}^{-1}(x)) = (C_{II} \circ C_{I}^{-1})(x) = Ax$$

$$\therefore \quad x = (C_{I} \circ C_{II}^{-1})\, y = A^{-1} x .$$

Ahora, supongamos que $\tilde{T} : V \to W$ es una transformación lineal, y que deseamos efectuar un cambio de base en V y otro en W.

Sea $T_{I} : \mathbb{R}^n \to \mathbb{R}^m$ la matriz asociada a $\tilde{T}$ con respecto a los sistemas C_{I}, S_{I}, y sea $T_{II} : \mathbb{R}^n \to \mathbb{R}^m$ la matriz asociada a $\tilde{T}$ con respecto a los sistemas C_{II}, S_{II}, donde

$C_{I}, C_{II} : V \to \mathbb{R}^n$

$S_{I}, S_{II} : W \to \mathbb{R}^m$ son sistemas de coordenadas.

Consideremos el siguiente diagrama:

$$\begin{tikzcd}
V \arrow[r, "\tilde{I}_V"] \arrow[d, "C_{II}"'] & V \arrow[r, "\tilde{T}"] \arrow[d, "C_{I}"'] & W \arrow[r, "\tilde{I}_W"] \arrow[d, "S_{I}"'] & W \arrow[d, "S_{II}"'] \\
\mathbb{R}^n \arrow[r, "(C_{I} \cdot C_{II}^{-1})"'] \arrow[rrr, bend right, "T_{II}"'] & \mathbb{R}^n \arrow[r, "T_{I}"'] & \mathbb{R}^m \arrow[r, "(S_{II} \circ S_{I}^{-1})"'] & \mathbb{R}^m
\end{tikzcd}$$

Obsérvese que el diagrama es conmutativo. Lo que queremos decir con esta frase es que la función que se obtiene al "ir" de V a $\mathbb{R}^n$ por la parte superior del diagrama

$$\begin{tikzcd}
V \arrow[r, "\tilde{I}_V"] \arrow[d, "C_{II}"'] & V \arrow[d, "C_I"] \\
\mathbb{R}^n \arrow[r, "A"'] & \mathbb{R}^n
\end{tikzcd}$$

es la misma que la que se obtiene al "ir" por la parte inferior del diagrama, esto es

$$A \circ C_{II} = C_I \circ \tilde{I}_V .$$

como se puede checar ya que $A = C_I \circ C_{II}^{-1}$,

y similarmente con las otras partes del diagrama.

Por lo tanto viendo la parte superior del diagrama de un extremo al otro tenemos:

$$T_{II} = (C_I \cdot C_{II}^{-1})\, T_I\, (S_{II} \circ S_I^{-1})$$

y despejando

$$T_I = (C_{II} \circ C_I^{-1})\, T_{II}\, (S_I \circ S_{II}^{-1})$$

Si $A = C_I \circ C_{II}^{-1}$ y $B = S_I \circ S_{II}^{-1}$
entonces

$$T_{II} = A\, T_I\, B^{-1}$$

$$T_I = A^{-1}\, T_{II}\, B$$

DEFINICIÓN

Sean T_1, T_2 dos matrices $m \times n$. Decimos que T_1 y T_2 son EQUIVALENTES si existen dos matrices A $m \times m$, B $n \times n$ no-singulares tales que

$$T_1 = A^{-1} T_2 B$$

Como acabamos de ver, las matrices equivalentes representan la misma transformación en distintas bases de $\mathbb{R}^n$ y $\mathbb{R}^m$.

DEFINICION

Sean T_1, T_2 dos matrices de orden n. Decimos que T_1 es SIMILAR a T_2 si existe una matriz A de orden n no-singular tal que

$$T_1 = A^{-1} T_2 A .$$

Si T_1 es similar a T_2, entonces T_1 y T_2 representan la misma transformación en bases distintas de $\mathbb{R}^n$.

DEFINICION

Sea $\mathbb{V}$ un espacio vectorial, y sean $\mathbb{V}_1, \mathbb{V}_2 \subseteq \mathbb{V}$. La SUMA de $\mathbb{V}_1$ y $\mathbb{V}_2$, denotada por $\mathbb{V}_1 + \mathbb{V}_2$, está dada por

$$\mathbb{V}_1 + \mathbb{V}_2 = \{ v \in \mathbb{V} : v = v_1 + v_2 \ , \ v_1 \in \mathbb{V}_1, \ v_2 \in \mathbb{V}_2 \}$$

Se puede ver fácilmente que si $\mathbb{V}_1$ y $\mathbb{V}_2$ son subespacios de $\mathbb{V}$, entonces $\mathbb{V}_1 + \mathbb{V}_2$ es subespacio de $\mathbb{V}$.

DEFINICION

Sea $\mathbb{V}$ un espacio vectorial, y $\mathbb{V}_1, \mathbb{V}_2 \subseteq \mathbb{V}$ subespacios de $\mathbb{V}$.

Decimos que $\mathbb{W} = \mathbb{V}_1 + \mathbb{V}_2$ es la SUMA DIRECTA de $\mathbb{V}_1$ y $\mathbb{V}_2$ si $\mathbb{V}_1 \cap \mathbb{V}_2 = \{0\}$.

En este caso escribimos

$$\mathbb{W} = \mathbb{V}_1 \oplus \mathbb{V}_2 .$$

RESULTADO

Sean V un espacio vectorial y V_1, V_2 subespacios de V. Entonces

$$V = V_1 \oplus V_2 \iff \dim(V) = \dim(V_1) + \dim(V_2).$$

La demostración es análoga a la que dimos para $V = \mathbb{R}^n$.

Describiremos algunas de las transformaciones $T: \mathbb{R}^n \to \mathbb{R}^n$.

1) TRANSFORMACIONES DE SIMILITUD (ELONGACIONES).

(a) Sea $T = \lambda I_n$, $\lambda \neq 0$. Entonces

$$Tx = \lambda x \quad \forall\, x \in \mathbb{R}^n$$

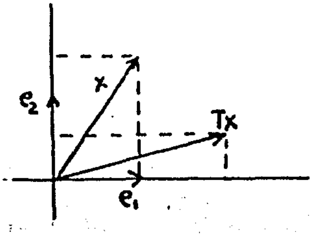
Si $|\lambda| > 1$ entonces T produce el efecto de alargar los vectores.
Si $|\lambda| < 1$ entonces T produce el efecto de acortar los vectores.

(b) Sea $\{v_1, v_2, \ldots, v_n\}$ una base de $\mathbb{R}^n$, y supongamos que en este sistema de coordenadas tenemos una transformación $T: \mathbb{R}^n \to \mathbb{R}^n$ dada por

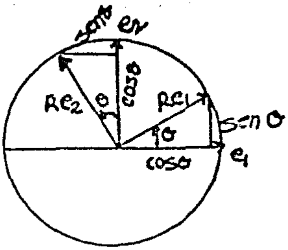
$$T = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

Entonces $Tv_1 = \lambda_1 v_1$, $Tv_2 = \lambda_2 v_2, \ldots$, $Tv_n = \lambda_n v_n$
O sea, que al aplicar T, acortamos a los vectores en unas direcciones y los

alargamos en otras direcciones.



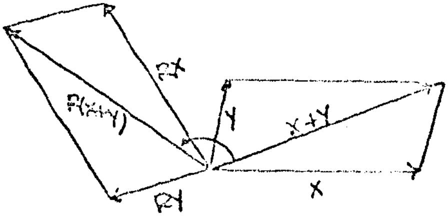
2) ROTACIONES



Si queremos efectuar una rotación $R: \mathbb{R}^2 \to \mathbb{R}^2$, entonces

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \longmapsto \begin{bmatrix} \cos\theta \\ \operatorname{sen}\theta \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \end{bmatrix} \longmapsto \begin{bmatrix} -\operatorname{sen}\theta \\ \cos\theta \end{bmatrix}$$

$\therefore \quad R = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$ representa una rotación de θ grados (en el sentido contrario a las manecillas del reloj).



3) PROYECCIONES

Supongamos $\mathbb{R}^n = \mathcal{S}_1 \oplus \mathcal{S}_2$. Por lo tanto todo vector $x \in \mathbb{R}^n$ es de la forma

$$x = x_1 + x_2 \qquad \text{con} \quad x_1 \in \mathcal{S}_1, \quad x_2 \in \mathcal{S}_2 ,$$

y para cada $x \in \mathbb{R}^n$ ésta expresión es única.

Definimos $P_1, P_2 : \mathbb{R}^n \to \mathbb{R}^n$ como sigue:

$$P_1 x = x_1$$
$$P_2 x = x_2$$

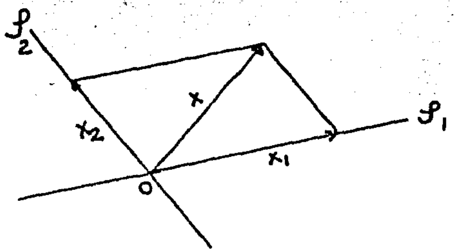
Decimos que x_1 es la proyección de x sobre $\mathcal{S}_1$, paralela a $\mathcal{S}_2$; y que x_2 es la proyección de x sobre $\mathcal{S}_2$, paralela a $\mathcal{S}_1$.

Observemos que P_1 y P_2 tienen la propiedad de que $P_1^2 = P_1$, $P_2^2 = P_2$

$$(P_1^2 x = P_1(P_1 x) = P_1 x_1 = x_1 = P_1 x \quad \therefore \quad P_1^2 = P_1) .$$

Además $P_1 + P_2 = I_n$, ya que

$$(P_1 + P_2) x = P_1 x + P_2 x = x_1 + x_2 = x .$$



Si $\mathbb{R}^n = \mathcal{S} \oplus \mathcal{S}^\perp$ entonces estamos en el caso de proyecciones ortogonales.

Veremos que si P es una proyección ortogonal, entonces $P^t = P$.

Observemos que podíamos haber definido una proyección $P : \mathbb{R}^n \to \mathbb{R}^n$ como una matriz P tal que $P^2 = P$.

Ahora, si P es una proyección, entonces P^t también lo es, ya que

si $P^2 = P \Rightarrow (P^t)^2 = P^t$.

Supongamos que $\mathbb{R}^n = \mathcal{J} \oplus \mathcal{J}^\perp$ y que P es la proyección sobre $\mathcal{J}$ paralela a $\mathcal{J}^\perp$ (P es la proyección ortogonal sobre $\mathcal{J}$).

Entonces $\mathcal{J} = R(P)$, ya que:

si $x_1 \in \mathcal{J}$ entonces $Px_1 = x_1 \quad \therefore x_1 \in R(P)$
$\therefore \quad \mathcal{J} \subseteq R(P)$

por otro lado $Px \in \mathcal{J} \quad \forall x \in \mathbb{R}^n$ y por lo tanto $R(P) \subseteq \mathcal{J}$

$$\therefore \quad R(P) = \mathcal{J}$$

Además $N(P) = \mathcal{J}^\perp$

$\therefore \quad \mathcal{J}^\perp = N(P) = R(P^t)^\perp$ (por lo visto en la última parte del cap 3)
$\therefore \quad \mathcal{J} = R(P^t)$

Por lo tanto, si P es una proyección ortogonal entonces $R(P^t) = R(P)$

Por lo tanto para cualquier $x \in \mathbb{R}^n$

$$Px \in R(P^t)$$

$$\therefore \quad P^t(Px) = Px$$

$$\therefore \quad P^tP = P$$

Transponiendo esta última igualdad obtenemos que

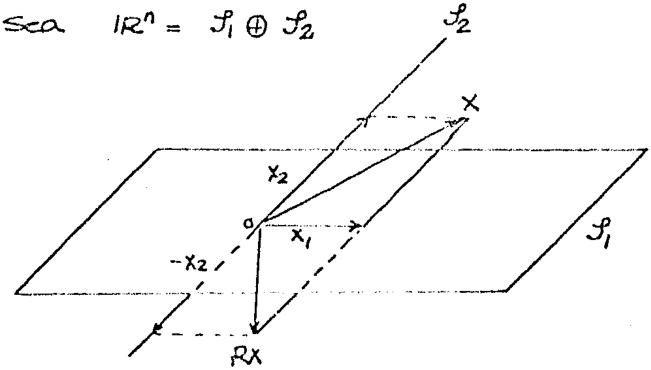
$$P^tP = P^t$$

Por lo tanto $P^tP = P = P^t$

Es decir, si P es una proyección ortogonal, entonces $P^t = P$.

4) REFLEXIONES

Sea $\mathbb{R}^n = \mathcal{S}_1 \oplus \mathcal{S}_2$



Sean $P_1, P_2 : \mathbb{R}^n \to \mathbb{R}^n$ las proyecciones sobre $\mathcal{S}_1$ y $\mathcal{S}_2$ respectivamente:

si $x = x_1 + x_2 \in \mathbb{R}^n$, con $x_1 \in \mathcal{S}_1$, $x_2 \in \mathcal{S}_2$ entonces

$$P_1 x = x_1$$
$$P_2 x = x_2$$

La reflexión de x sobre $\mathcal{S}_1$, paralela a $\mathcal{S}_2$, está dada por la matriz $R : \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$Rx = P_1 x - P_2 x = x_1 - x_2$$

Es decir:

$$Rx = (P_1 - P_2)\, x$$

Esto lo podemos reescribir, usando la

igualdad $I = P_1 + P_2$

$$R = P_1 - P_2 = P_1 - (I - P_1) = 2P_1 - I$$

Por lo tanto la reflexión sobre $\mathcal{S} = \mathcal{S}_1$ (paralela a $\mathcal{S}_2$) está dada por la matriz

$$R = 2P_1 - I$$

donde P_1 es la proyección sobre $\mathcal{S}$.

Observemos que si R es la reflexión sobre $\mathcal{S}$, entonces

$$R^2 = 4P_1^2 - 4P_1 + I = 4P_1 - 4P_1 + I = I$$

$$\therefore \quad R^2 = I$$

Por lo tanto tenemos que

R es la reflexión sobre $\mathcal{S}$ (paralela a $\mathcal{S}_2$) $\Leftrightarrow$ $P_1 = \frac{1}{2}(I+R)$ es la proyección sobre $\mathcal{S}$ (paralela a $\mathcal{S}_2$).

Observemos que podríamos haber definido una reflexión como una matriz $R : \mathbb{R}^n \to \mathbb{R}^n$ tal que $R^2 = I_n$, ya que si $R^2 = I \Rightarrow$

$P_1 = \frac{1}{2}(I+R)$ es una proyección, ya que

$$
\begin{aligned}
P_1^2 &= \frac{1}{4}(I^2 + 2R + R^2) \\
&= \frac{1}{4}(I + 2R + I) \\
&= \frac{1}{2}(I+R) = P_1
\end{aligned}
$$

$\therefore$ P_1 es una proyección.

EJEMPLO EN $\mathbb{R}^2$:

Sea $\{a, c\}$ una base de $\mathbb{R}^2$,

Sea $b \in \mathbb{R}^2$ tal que $c^t b = 0$, $b^t a = 1$

Entonces $P : \mathbb{R}^2 \to \mathbb{R}^2$ dada por

$P = a b^t$ es la proyección sobre $[a]$ paralela a $[c]$:

Ya que si $x = \alpha a + \gamma c \in \mathbb{R}^2$ entonces

$$
\begin{aligned}
Px &= ab^t(\alpha a + \gamma c) \\
&= ab^t(\alpha a) + ab^t(\gamma c) \\
&= \alpha a(b^t a) + \gamma a(b^t c) \\
&= \alpha a .
\end{aligned}
$$

ESPACIO DUAL

DEFINICION

Sea V un espacio vectorial. Una transformación lineal $\ell : V \to \mathbb{R}$ se llama FUNCIONAL LINEAL.

Ejemplo

Sea $V = \{ A \in M_3 : A$ es cuadrado mágico $\}$

Sea $\ell : V \to \mathbb{R}$ definida como sigue:

Si A es un cuadrado mágico, entonces

$\ell(A) = s$, donde $s = \alpha_{11} + \alpha_{12} + \alpha_{13} = s(A)$.

$\therefore$ Como vimos anteriormente, si $A, B \in V$ entonces

$$\ell(A+B) = \ell(A) + \ell(B)$$
$$\ell(\alpha A) = \alpha \ell(A)$$

Por lo tanto $\ell : V \to \mathbb{R}$ es un funcional lineal.

DEFINICION

Sea $\mathbb{V}$ un espacio vectorial. Al espacio vectorial

$$L(\mathbb{V}, \mathbb{R}) = \{ \ell : \mathbb{V} \to \mathbb{R} : \ell \text{ es lineal} \}$$

lo llamamos ESPACIO DUAL de $\mathbb{V}$, y lo denotaremos por $\mathbb{V}^*$:

$$\mathbb{V}^* = \{ \ell : \mathbb{V} \to \mathbb{R} : \ell \text{ es lineal} \}$$

Por ejemplo, si $\mathbb{V} = \mathbb{R}^n$, entonces

$$\begin{aligned} \mathbb{R}^{n*} &= \{ \ell : \mathbb{R}^n \to \mathbb{R} : \ell \text{ es lineal} \} \\ &= M(1,n) \\ &= \{ c^t : c \in \mathbb{R}^n \} \end{aligned}$$

RESULTADO

Sea $\mathbb{V}$ un espacio vectorial de dimensión n. Entonces $\mathbb{V}^*$ es isomorfo a $\mathbb{R}^{n*}$.

Demostración :

$$\begin{tikzcd}
\mathbb{V} \arrow[r, "\ell"] \arrow[d, "S_I"'] & \mathbb{R} \arrow[d, "I"] \\
\mathbb{R}^n \arrow[r, "c^t"'] & \mathbb{R}
\end{tikzcd}$$

Por un resultado anterior $L(\mathbb{V}, \mathbb{R}) \cong M(1,n)$

$$\therefore \quad \mathbb{V}^* \cong \mathbb{R}^{n*}$$

DEFINICION

Sea $\mathbb{V}$ un espacio vectorial y $\mathcal{L} \subseteq \mathbb{V}$ un subconjunto no-vacío de $\mathbb{V}$. Decimos que $\mathcal{L}$ es una VARIEDAD LINEAL si $\mathcal{L}$ es de la forma

$$\mathcal{L} = v_0 + \mathcal{S} = \{ v \in \mathbb{V} : v = v_0 + s \ , \ s \in \mathcal{S} \}$$

donde $\mathcal{S} \subseteq \mathbb{V}$ es un subespacio de $\mathbb{V}$, y $v_0 \in \mathcal{L}$.

RESULTADO

Sea $\mathbb{V}$ un espacio vectorial y $\ell : \mathbb{V} \to \mathbb{R}$, $\ell \in \mathbb{V}^*$ un funcional lineal. Entonces

$$\{ v \in \mathbb{V} : \ell(v) = \alpha \}$$

es una variedad lineal.

Demostración :

Supongamos $\{v \in \mathbb{V} : \ell(v) = \alpha\} \neq \phi$.

Sea $v_0 \in \{v \in \mathbb{V} : \ell(v) = \alpha\}$. Por lo tanto

$$\begin{aligned} \ell(v) = \alpha &\Leftrightarrow \ell(v) - \ell(v_0) = 0 \\ &\Leftrightarrow \ell(v - v_0) = 0 \qquad (\ell \text{ lineal}) \\ &\Leftrightarrow v - v_0 \in N(\ell) \\ &\Leftrightarrow v \in v_0 + N(\ell) \end{aligned}$$

$\therefore \{v \in \mathbb{V} : \ell(v) = \alpha\} = v_0 + N(\ell)$

$\therefore \{v \in \mathbb{V} : \ell(v) = \alpha\}$ es una variedad lineal. ∎

RESULTADO

Sean $\mathbb{V}$ un espacio vectorial, y $\ell_1, \ell_2, \dots, \ell_K \in \mathbb{V}^*$ funcionales lineales. Sean

$$\begin{aligned} \mathcal{L}_1 &= \{v \in \mathbb{V} : \ell_1(v) = \alpha_1\} \\ \mathcal{L}_2 &= \{v \in \mathbb{V} : \ell_2(v) = \alpha_2\} \\ &\vdots \qquad\qquad\qquad \vdots \\ \mathcal{L}_K &= \{v \in \mathbb{V} : \ell_K(v) = \alpha_K\} \end{aligned}$$

variedades lineales en $\mathbb{V}$. Supongamos que $\mathcal{L} = \mathcal{L}_1 \cap \mathcal{L}_2 \cap \dots \cap \mathcal{L}_K \neq \phi$. Entonces

$$\mathcal{L} = \{ v \in V : l_i(v) = \alpha_i \quad , \quad 1 \leq i \leq k \}$$

es una variedad lineal.

Demostración:

Como $\mathcal{L}_1 \cap \mathcal{L}_2 \cap \cdots \cap \mathcal{L}_k \neq \phi$ podemos escoger un vector $v_0 \in \mathcal{L} = \mathcal{L}_1 \cap \mathcal{L}_2 \cap \cdots \cap \mathcal{L}_k$

Por lo tanto $l_i(v_0) = \alpha_i \quad , \quad 1 \leq i \leq k$ y

$$\begin{aligned} v \in \mathcal{L} \quad & \Longleftrightarrow \quad l_i(v) - l_i(v_0) = 0 \qquad 1 \leq i \leq k \\ & \Longleftrightarrow \quad l_i(v - v_0) = 0 \qquad 1 \leq i \leq k \\ & \Longleftrightarrow \quad v - v_0 \in N(l_i) \qquad 1 \leq i \leq k \\ & \Longleftrightarrow \quad v - v_0 \in \bigcap_{i=1}^{k} N(l_i) \\ & \Longleftrightarrow \quad v \in v_0 + \bigcap_{i=1}^{k} N(l_i) \end{aligned}$$

Como la intersección (finita) de subespacios es un subespacio, entonces $\bigcap_{1}^{k} N(l_i)$ es un subespacio de $\mathbb{R}^n$.

Por lo tanto la intersección (finita) de variedades lineales es una variedad lineal.

∎

RESULTADO

Sea V un espacio vectorial de dimensión n. Sea $\{v_1, v_2, \ldots, v_n\}$ una base de V, y sea $l_i : V \to \mathbb{R}$ la función dada por

$$l_i(v) = \xi_i \ , \quad \text{donde} \quad v = \sum_{j=1}^{n} \xi_j v_j \in V \ ; \ 1 \leq i \leq n$$

Entonces

$l_i : V \to \mathbb{R}$ es un funcional lineal, $1 \leq i \leq n$

y $\{l_1, l_2, \ldots, l_n\}$ es una base de V^*

Demostración:

Para cada i, $1 \leq i \leq n$, l_i está bien definida ya que la expresión

$$v = \sum_{j=1}^{n} \xi_j v_j \qquad \text{para cada vector } v \in V$$

es única.

Sean $v, w \in V$ y $\alpha \in \mathbb{R}$. Por lo tanto

$$v = \sum_{j=1}^{n} \xi_j v_j \qquad w = \sum_{j=1}^{n} \eta_j v_j$$

y tenemos que

$$\begin{aligned} l_i(\alpha v+w) &= l_i\left(\alpha \sum_{j=1}^{n} \xi_j v_j + \sum_{j=1}^{n} \eta_j v_j\right) \\ &= l_i\left(\sum_{j=1}^{n} (\alpha \xi_j + \eta_j)\right) \\ &= \alpha \xi_i + \eta_i \\ &= \alpha\, l_i(v) + l_i(w) \end{aligned}$$

$\therefore$ l_i es un funcional lineal, $1 \leq i \leq n$.

$\therefore$ $l_i \in V^*$, $1 \leq i \leq n$.

Veremos que $\{l_1, l_2, \ldots, l_n\}$ es una base de V^*. Basta con ver que $\{l_1, l_2, \ldots, l_n\}$ es independiente:

Supongamos

$$\alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n = 0 \quad \in V^*$$

$$\begin{aligned} \Rightarrow (\alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n)(v_1) &= 0 \in \mathbb{R} \\ (\alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n)(v_2) &= 0 \in \mathbb{R} \\ \vdots \qquad\qquad & \quad \vdots \\ (\alpha_1 l_1 + \alpha_2 l_2 + \cdots + \alpha_n l_n)(v_n) &= 0 \in \mathbb{R} \end{aligned}$$

$$\Rightarrow \quad \begin{array}{l} \alpha_1 = 0 \\ \alpha_2 = 0 \\ \vdots \\ \alpha_n = 0 \end{array}$$

$\therefore \quad \{l_1, l_2, \dots, l_n\}$ es independiente

$\therefore \quad \{l_1, l_2, \dots, l_n\}$ es una base de V^*.

Observemos que

$l_i : V \to \mathbb{R}$ es la función que a cada $v \in V$ le asocia su i-ésima coordenada en la base dada.

A $l_1, l_2, \dots, l_n$ las llamamos FUNCIONES COORDENADAS de la base dada.

∎

RESULTADO

Sea $\mathcal{L} = v_0 + \mathcal{S} \subseteq \mathbb{V}$ una variedad lineal.

Supongamos $\dim(\mathbb{V}) = n$ y $\dim(\mathcal{S}) = k$.

Entonces $\mathcal{L}$ es de la forma

$$\mathcal{L} = v_0 + \bigcap_{i=k+1}^{n} N(\ell_i)$$

donde $\ell_i : \mathbb{V} \to \mathbb{R}$ son funcionales lineales $1 \leq i \leq n$

Demostración:

Sea $\mathcal{L} = v_0 + \mathcal{S}$ una variedad lineal en $\mathbb{V}$.

Sea $\{s_1, s_2, \dots, s_k\}$ una base de $\mathcal{S}$ y sea $\{s_1, s_2, \dots, s_k, s_{k+1}, s_{k+2}, \dots, s_n\}$ una base de $\mathbb{V}$.

Sean $\ell_i : \mathbb{V} \to \mathbb{R}$ las funciones coordenadas en la base dada

$$\ell_i(v) = \xi_i \qquad \text{si} \quad v = \sum_{j=1}^{n} \xi_j s_j \in \mathbb{V}.$$

Por lo tanto ℓ_i es un funcional lineal, $1 \leq i \leq n$, y si $v = \sum_{j=1}^{n} \xi_j s_j \in V$ entonces

$$v \in \mathcal{S} \Leftrightarrow \xi_i = 0 \quad k+1 \leq i \leq n$$

$$\Leftrightarrow l_i(v) = 0 \quad k+1 \leq i \leq n$$

$$\Leftrightarrow v \in N(l_i) \quad k+1 \leq i \leq n$$

$$\Leftrightarrow v \in \bigcap_{i=k+1}^{n} N(l_i)$$

Por lo tanto $\mathcal{S} = \bigcap_{i=k+1}^{n} N(l_i)$

$$\therefore \quad \mathcal{L} = v_0 + \bigcap_{i=k+1}^{n} N(l_i) \quad \blacksquare$$

CAPITULO VI

FUNCIONES CUADRATICAS
Y
FORMAS BILINEALES

FUNCIONES CUADRATICAS Y FORMAS BILINEALES

INTRODUCCION

Supongamos que tenemos un vector $x = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} \in \mathbb{R}^2$ y una función cuadrática en x

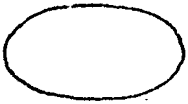
$$g(x) = \gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2 + \delta_1 \xi_1 + \delta_2 \xi_2 + \gamma$$

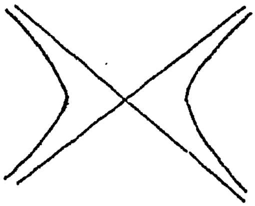
Nos interesa estudiar las curvas de nivel de g:

$$n(g, cte) = \{ x \in \mathbb{R}^2 : g(x) = cte \}$$

Por nuestra experiencia en geometría analítica, sabemos que las curvas de nivel pueden tener una de las siguientes formas:

elipse





hipérbola

parábola

Observemos que, si excluimos a las parábolas, el tipo de curva asociado a la función q depende exclusivamente de los términos cuadráticos

$$\gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2$$

lo cual nos sugiere estudiar con más detenimiento a las "formas cuadráticas", que son funciones cuadráticas "homogéneas":

Sea

$$\Phi_2 = \left\{ \varphi : \mathbb{R}^2 \to \mathbb{R} \;:\; \varphi(x) = \gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2 \;;\; x^t = [\xi_1 \; \xi_2] \right\}$$

Entonces :

1) Φ_2 es un espacio vectorial, ya que si

$\varphi, \varphi' \in \Phi_2$ y $\alpha \in \mathbb{R}$:

$$(\varphi + \varphi')(x) = (\gamma_1 + \gamma_1')\xi_1^2 + (\gamma_2 + \gamma_2')\xi_2^2 + (\beta + \beta')(\xi_1 \xi_2)$$

$$(\alpha \varphi)(x) = (\alpha \gamma_1)\xi_1^2 + (\alpha \gamma_2)\xi_2^2 + (\alpha \beta)(\xi_1 \xi_2)$$

donde
$$\varphi(x) = \gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2$$
$$\varphi'(x) = \gamma_1' \xi_1^2 + \gamma_2' \xi_2^2 + \beta' \xi_1 \xi_2$$

$\therefore \; \varphi + \varphi', \; \alpha \varphi \in \Phi_2$.

2) Una base de Q_2 es $\{\xi_1^2, \xi_2^2, \xi_1\xi_2\}$:

Es claro que $\{\xi_1^2, \xi_2^2, \xi_1\xi_2\}$ genera Q_2, veremos que también es linealmente independiente :

$$\text{si} \quad \gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2 = 0$$

$$\left.\begin{array}{l} x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \Rightarrow \gamma_1 = 0 \\ x = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Rightarrow \gamma_2 = 0 \end{array}\right\} \Rightarrow \beta \xi_1 \xi_2 = 0$$

$$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Rightarrow \beta = 0$$

∎

Lo que queremos ahora es investigar las curvas de nivel de los elementos de Q_2 .- sin embargo ésta base que acabamos de dar no es la más adecuada para identificar de qué tipo curva de nivel se trata. El procedimiento usual es efectuar un cambio de coordenadas lineal en $\mathbb{R}^2$, es decir, de la forma

$$\tilde{x} = Ux$$

con U no-singular.

Bajo este cambio de base

$$n(\varphi, cte) = \{ x \in \mathbb{R}^2 : \varphi(x) = cte \}$$

se transforma en el conjunto

$$\{ \tilde{x} \in \mathbb{R}^2 : \tilde{x} = Ux \ , \ x \in n(\varphi, cte) \} .$$

Este conjunto, como se verá más adelante, se puede escribir de la forma

$$n(\tilde{\varphi}, cte) = \{ \tilde{x} \in \mathbb{R}^2 : \tilde{\varphi}(\tilde{x}) = cte \}$$

donde $\tilde{\varphi}(\tilde{x})$ se obtiene de $\varphi(x)$ por medio de U.

Lo que necesitamos es encontrar una representación de φ que nos haga fácil obtener $\tilde{\varphi}$, tal que para $\tilde{\varphi}$ sea fácil decir cómo son sus curvas de nivel. Veremos que ésto se puede lograr con matrices de la forma que indicaremos más adelante.

Es natural que al cuestionarnos sobre las propiedades de los elementos de Φ_2, nos preguntemos si son transformaciones lineales.

¿ $\varphi(\alpha x) = \alpha \varphi(x)$?

¿ $\varphi(x+y) = \varphi(x) + \varphi(y)$?

Veremos que ninguna de estas dos propiedades se cumple :

$$\varphi(\alpha x) = \gamma_1(\alpha\xi_1)^2 + \gamma_2(\alpha\xi_2)^2 + \beta(\alpha\xi_1)(\alpha\xi_2) = \alpha^2\varphi(x)$$
$$\varphi(x+y) = \gamma_1(\xi_1+\eta_1)^2 + \gamma_2(\xi_2+\eta_2)^2 + \beta(\xi_1+\eta_1)(\xi_2+\eta_2)$$

Por lo tanto

$$\varphi(\alpha x) = \alpha^2\varphi(x)$$
$$\varphi(x+y) = \varphi(x) + \varphi(y) + b(x,y)$$

donde $b(x,y) = 2\gamma_1\xi_1\eta_1 + 2\gamma_2\xi_2\eta_2 + \beta(\xi_1\eta_2 + \xi_2\eta_1)$

Observemos que

$$\begin{aligned} b(x,y) &= (2\gamma_1\xi_1 + \beta\xi_2)\eta_1 + (\beta\xi_1 + 2\gamma_2\xi_2)\eta_2 \\ &= \begin{bmatrix} 2\gamma_1\xi_1 + \beta\xi_2 & \beta\xi_1 + 2\gamma_2\xi_2 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \\ &= \begin{bmatrix} \xi_1 & \xi_2 \end{bmatrix} \begin{bmatrix} 2\gamma_1 & \beta \\ \beta & 2\gamma_2 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = x^t B y \end{aligned}$$

También se puede notar que

$$b(x,y) = b(y,x)$$

y que $b(\alpha x_1 + x_2, y) = \alpha b(x_1,y) + b(x_2,y)$

además

$$\varphi(2x) = 4\varphi(x) = \varphi(x) + \varphi(x) + b(x,x)$$

II.6

$$\therefore \quad \varphi(x) = \frac{1}{2}\, b(x,x)$$

Por lo tanto si $\varphi(x)$ es una forma cuadrática, entonces

1) $\varphi(x+y) = \varphi(x) + \varphi(y) + b(x,y)$

donde $b(x,y) = x^t B y$, con $B = B^t$

2) $\varphi(\alpha x) = \alpha^2 \varphi(x)$

3) $\varphi(x) = \frac{1}{2}\, b(x,x)$, donde b es como en (1)

Observar que la igualdad

$$\varphi(x) = \frac{1}{2} x^t B x$$

nos proporciona una manera de asignar una matriz a la función φ.

También podemos ver que podemos estudiar las formas cuadráticas a través de las "formas bilineales":

$$B_2 = \left\{ b : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : b(x,y) = \sum_{i=1}^{2} \sum_{j=1}^{2} \beta_{ij}\, \xi_i \eta_j \right\}$$

Propiedades de B_2

1) B_2 es un espacio vectorial

2) Si $b(x,y) \in B_2$ entonces

b es lineal en x (tomando a y constante)
y b es lineal en y (tomando a x constante)

3) B_2 es isomorfo a $M_2 = M(2,2)$

La correspondencia $b \longmapsto B$, donde $b(x,y) = x^t By$, es un isomorfismo entre B_2 y M_2.

4) B_2 no es isomorfo a Q_2, ya que por el inciso anterior $\dim(B_2) = 4$, pero $\dim(Q_2) = 3$.

5) Recordemos que si $q(x) \in Q_2$ entonces

$$q(x+y) = q(x) + q(y) + b(x,y)$$

donde $b(x,y) = x^t By$, con $B = B^t$, y tal que $b(x,y) = b(y,x)$

Esto nos sugiere que quizás si

$$S_2 = \{ B \in M_2 : B = B^t \}$$

y

$$\tilde{S}_2 = \{ b \in B_2 \ : \ b(x,y) = b(y,x) \}$$

entonces S_2 es isomorfo a $\tilde{S}_2$.

En efecto, $S_2 \cong \tilde{S}_2$, ya que la correspondencia

$b \xrightarrow{T} B$, donde $b \in \tilde{S}_2$ está dada

por $b(x,y) = x^t B y$, es un isomorfismo.

Es fácil ver que T es lineal, y además

$$\begin{aligned} b(x,y) = b(y,x) &\Leftrightarrow x^t B y = y^t B x & \forall x,y \in \mathbb{R}_2 \\ &\Leftrightarrow y^t B^t x = y^t B x & \forall x,y \in \mathbb{R}_2 \\ &\Leftrightarrow B = B^t . \end{aligned}$$

6) φ_2 es isomorfo a S_2:

Si $\varphi(x) = \gamma_1 \xi_1^2 + \gamma_2 \xi_2^2 + \beta \xi_1 \xi_2$

Sea $B = \begin{bmatrix} \gamma_1 & \beta/2 \\ \beta/2 & \gamma_2 \end{bmatrix}$

(o sea que B es la matriz tal que $\varphi(x+y) = \varphi(x) + \varphi(y) + x^t B y$).

Se puede checar fácilmente que la correspondencia $B \xrightarrow{T} \varphi$ es un isomorfismo de φ_2 a S_2.

7) Sea $T : M_2 \to \varphi_2$ la transformación dada por:

Si $A \in M_2$ entonces

$T(A) = \varphi$ es la forma cuadrática $\in \varphi_2$ tal que $\varphi(x) = x^t A x$.

Se puede checar fácilmente que T es lineal y sobre, y que

$$N(T) = \{ A \in M_2 : A^t = -A \}$$

8) sea $A_2 = \{ A \in M_2 : A^t = -A \}$

(Si $A \in A_2$ decimos que A es anti-simétrica).

Entonces $M_2 = S_2 \oplus A_2$ ya que

$S_2 \cap A_2 = \{0\}$ y toda matriz $A = [\alpha_{ij}]$

de M_2 se puede escribir de la forma

$$A = \frac{1}{2} \begin{bmatrix} 2\alpha_{11} & \alpha_{12} + \alpha_{21} \\ \alpha_{12} + \alpha_{21} & 2\alpha_{22} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 0 & \alpha_{12} - \alpha_{21} \\ \alpha_{21} - \alpha_{12} & 0 \end{bmatrix}$$

$= A_S + A_{S^-}$, con $A_S \in S_2$ y $A_{S^-} \in A_2$.

FUNCIONES CUADRATICAS Y FORMAS BILINEALES EN $\mathbb{R}^n$

DEFINICION

Sea $X = [\xi_i] \in \mathbb{R}^n$. Una FORMA CUADRATICA HOMOGENEA en n variables $\xi_1, \xi_2, \dots, \xi_n$ es un polinomio cuadrático homogéneo :

$$\varphi(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \xi_i \xi_j$$

Denotaremos al conjunto de las formas cuadráticas por

$$Q_n = \{ \varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \ : \ \varphi(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \xi_i \xi_j \}$$

RESULTADO

Q_n es un espacio vectorial de dimensión $\frac{n(n+1)}{2}$.

. Demostración :

Únicamente veremos que si $\varphi, \varphi' \in Q_n$ y $\alpha \in \mathbb{R}$ entonces $\varphi + \varphi'$ y $\alpha\varphi \in Q_n$, y daremos una base de Q_n.

Si $\varphi, \varphi' \in Q_n$, $\alpha \in \mathbb{R}$:

$$(\varphi+\varphi')(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} (\gamma_{ij} + \gamma'_{ij}) E_i E_j \quad \therefore (\varphi+\varphi') \in Q_n$$

$$(\alpha\varphi)(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} (\alpha \gamma_{ij}) E_i E_j \quad \therefore (\alpha \varphi) \in Q_n .$$

$\{ E_1^2, E_1E_2, \ldots, E_1E_n, E_2^2, E_2E_3, \ldots, E_2E_n, E_3^2, \ldots, E_3E_n, \ldots E_n^2 \}$ es una base de Q_n :

Ya que si $\varphi(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} E_i E_j$

entonces

$$\varphi(x) = \gamma_{11} E_1^2 + (\gamma_{12}+\gamma_{21}) E_1E_2 + (\gamma_{13}+\gamma_{31}) E_1 E_3 + \cdots + (\gamma_{1n}+\gamma_{n1}) E_1E_n$$

$$+ \gamma_{22} E_2^2 + (\gamma_{23}+\gamma_{32}) E_2E_3 + \cdots + (\gamma_{2n}+\gamma_{n2}) E_2 E_n$$

$$+ \cdots + \gamma_{nn}^2$$

$$= \sum_{i\leq j}^{n} \sum_{j=1}^{n} \alpha_{ij} E_i E_j$$

$\therefore \{E_iE_j\}$, $1 \leq i \leq j$ $1 \leq j \leq n$ genera Q_n.

Ahora, supongamos

$$\varphi(x) = \sum_{i \leq j}^{n} \sum_{j=1}^{n} \gamma_{ij} \Xi_i \Xi_j = 0 \qquad \forall x$$

Entonces

$$\varphi(e_1) = 0 \Rightarrow \gamma_{11} = 0$$

$$\varphi(e_2) = 0 \Rightarrow \gamma_{22} = 0$$

$$\vdots$$

$$\varphi(e_n) = 0 \Rightarrow \gamma_{nn} = 0$$

$$\therefore \quad \varphi(x) = \sum_{i<j}^{n} \sum_{j=1}^{n} \gamma_{ij} \Xi_i \Xi_j$$ y por lo tanto

$$\varphi(e_i + e_j) = \gamma_{ij} \quad \text{si} \quad i<j$$

$$\therefore \quad i<j \, , \quad \varphi(e_i + e_j) = 0 \Rightarrow \gamma_{ij} = 0$$

Por lo tanto $\displaystyle\sum_{i \leq j}^{n} \sum_{j=1}^{n} \gamma_{ij} \Xi_i \Xi_j = 0$

$$\Rightarrow \gamma_{ij} = 0 \qquad 1 \leq i \leq j \, , \quad 1 \leq j \leq n$$

$\therefore \quad \{\Xi_i \Xi_j\} \quad 1 \leq i \leq j \, , \ 1 \leq j \leq n$ es independiente.

y por lo tanto es una base de Q_n.

y por lo tanto

$$\dim(Q_n) = 1+2+\cdots+n = \frac{n(n+1)}{2}$$

$\blacksquare$

PROPIEDADES DE LAS FORMAS CUADRATICAS HOMOGENEAS

1) $\varphi(\alpha X) = \alpha^2 \varphi(X) \qquad \forall \varphi \in Q_n, \; X \in \mathbb{R}^n$

2) $\varphi(X+Y) = \varphi(X) + \varphi(Y) + 2\, b(X,Y)$ donde

$b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ es tal que

i) $b(X,Y) = b(Y,X)$
ii) $b(X,Y)$ es lineal en X
$b(X,Y)$ es lineal en Y.

y donde $\varphi(X) = b(X,X)$.

Demostración :

1) $$\varphi(\alpha X) = \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}(\alpha\xi_i)(\alpha\xi_j) = \alpha^2 \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}\xi_i\xi_j$$ $\blacksquare$

2) $$\varphi(X+Y) = \sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}(\xi_i+\eta_i)(\xi_j+\eta_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} (\xi_i \xi_j + \eta_i \eta_j) + \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} (\xi_i \eta_j + \xi_j \eta_i)$$

$\therefore \quad \varphi(X+Y) = \varphi(X) + \varphi(Y) + 2\, b(X,Y)$ donde

$$b(X,Y) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} (\xi_i \eta_j + \xi_j \eta_i)$$

Además

$$b(X,Y) = \frac{1}{2} \Big(\varphi(X+Y) - \varphi(X) - \varphi(Y) \Big)$$

de donde se ve que $b(X,Y) = b(Y,X)$, y

$b(X,Y)$ es lineal en X y en Y :

$$b(\alpha X_1 + X_2, Y) = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \left[(\alpha \xi_i^{(1)} + \xi_i^{(2)}) \eta_j + (\alpha \xi_j^{(1)} + \xi_j^{(2)}) \eta_i \right]$$

$$= \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \left[\alpha (\xi_i^{(1)} \eta_j + \xi_j^{(1)} \eta_i) + \xi_i^{(2)} \eta_j + \xi_j^{(2)} \eta_i \right]$$

$$= \alpha \cdot \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \left[\xi_i^{(1)} \eta_j + \xi_j^{(1)} \eta_i \right] +$$

$$+ \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} (\xi_i^{(2)} \eta_j + \xi_j^{(2)} \eta_i)$$

$$= \alpha\, b(X_1, Y) + b(X_2, Y) .$$

Por la simetría, $b(x,y)$ también es lineal en y.

Por último

$$b(x,x) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}\,(\xi_i\xi_j + \xi_i\xi_j)$$

$$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}\,(2\xi_i\xi_j) = \sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij}\,\xi_i\xi_j$$

Por lo tanto

$$\varphi(x) = b(x,x) \quad \square$$

Observar que de paso hemos demostrado que si $\varphi(x)$ es una forma cuadrática homogénea, entonces

$$b(x,y) = \frac{1}{2}(\varphi(x+y) - \varphi(x) - \varphi(y))$$

es tal que

$b(x,y) = b(y,x)$
$b(x,y)$ es lineal en x
$b(x,y)$ es lineal en y.

Esto lo usaremos más adelante.

$b(x,y)$ se puede ver como una función

$$b : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$$

Estudiaremos funciones $b : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$, donde m y n pueden ser diferentes.

DEFINICION

Una FUNCION BILINEAL EN $\mathbb{R}^m \times \mathbb{R}^n$ es una función

$$F : \mathbb{R}^m \times \mathbb{R}^n \longrightarrow \mathbb{R}$$

tal que $x, x_1, x_2 \in \mathbb{R}^m$, $y, y_1, y_2 \in \mathbb{R}^n$, $\alpha, \beta \in \mathbb{R}$ se cumplen las propiedades :

1) $F(\alpha x_1 + x_2, y) = \alpha F(x_1, y) + F(x_2, y)$

2) $F(x, \beta y_1 + y_2) = \beta F(x, y_1) + F(x, y_2)$

Es decir, decimos que $F(x,y)$ es bilineal si F es lineal en x (para y fijo) y en y (para x fijo).

Veremos que si escogemos sistemas de coordenadas en $\mathbb{R}^m$ y en $\mathbb{R}^n$, entonces podemos encontrar una expresión general ó forma para el valor de $F(x,y)$.

RESULTADO

Supongamos que $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ es una función bilineal. Entonces

$$F(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} F(e_i, e_j)\, \xi_i\, \eta_j$$

donde $x = [\xi_i] \in \mathbb{R}^m$, $y = [\eta_j] \in \mathbb{R}^n$, $e_i \in \mathbb{R}^m$, $e_j \in \mathbb{R}^n$.

Demostración :

Supongamos $x = \sum_{i=1}^{m} \xi_i e_i$ $\quad y = \sum_{j=1}^{n} \eta_j e_j$

$\Rightarrow$

$$F(x,y) = F\Big(\sum_{i=1}^{m} \xi_i e_i \, , \, \sum_{j=1}^{n} \eta_j e_j \Big)$$

$$= \sum_{i=1}^{n} \xi_i\, F\Big(e_i, \sum_{j=1}^{n} \eta_j e_j\Big)$$

$$= \sum_{i=1}^{n} \xi_i \sum_{j=1}^{n} \eta_j\, F(e_i, e_j)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} F(e_i, e_j)\, \xi_i\, \eta_j$$

DEFINICION

Una FORMA BILINEAL HOMOGENEA en las variables $\xi_1, \xi_2, \dots, \xi_m, \eta_1, \eta_2, \dots, \eta_n$ es un polinomio

$$b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \, \xi_i \eta_j \quad ; \quad \begin{array}{l} x = [\xi_i] \in \mathbb{R}^m \\ y = [\eta_j] \in \mathbb{R}^n . \end{array}$$

Por lo tanto si $F : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ es una función bilineal en $\mathbb{R}^m \times \mathbb{R}^n$ entonces

$F(x,y)$ es una forma bilineal.

RESULTADO

Sea $b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \, \xi_i \eta_j$ una forma bilineal. Entonces

$$b(x,y) = x^t B y, \quad \text{donde} \quad B = [\beta_{ij}]$$

Demostración:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \, \xi_i \eta_j = \sum_{i=1}^{m} \xi_i \left(\sum_{j=1}^{n} \beta_{ij} \eta_j \right)$$

$$= \sum_{i=1}^{m} \xi_i \cdot \left[\beta_{i1} \eta_1 + \beta_{i2} \eta_2 + \cdots + \beta_{in} \eta_n \right]$$

$$= \left(\sum_{i=1}^{m} \xi_i \left[\beta_{i1} \;\; \beta_{i2} \cdots \beta_{in} \right] \right) \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$$

$$= \left[\xi_1 \;\; \xi_2 \cdots \xi_m \right] \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2n} \\ & & & \\ \beta_{m1} & \beta_{m2} & \cdots & \beta_{mn} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}$$

$$= X^t B Y \quad \square$$

Hay que observar que la forma bilineal $b(X,Y)$ está completamente determinada por la matriz B.

Observar también que la entrada β_{ij} de la matriz B es el coeficiente de $\xi_i \eta_j$ en la forma $b(X,Y)$.

RESULTADO

Sea $b(X,Y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j$ una forma bilineal.

Entonces $b(X+Y, X+Y) = b(X,X) + b(Y,Y) + b(X,Y) + b(Y,X)$.

Demostración:

$$b(X+Y, X+Y) = b(X, X+Y) + b(Y, X+Y)$$
$$= b(X,X) + b(X,Y) + b(Y,X) + b(Y,Y) \quad \square$$

RESULTADO

sea $B(m,n) = \{ b : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \ : \ b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j \}$

sea $(b+b')(x,y) = b(x,y) + b'(x,y)$
$(\alpha b)(x,y) = \alpha \, b(x,y)$

Entonces $B(m,n)$, con éstas operaciones, es un espacio vectorial, y su dimensión es mn.

Demostración:

Veremos únicamente que $\{\xi_i \eta_j\}$ $1 \le i \le m$, $1 \le j \le n$ es una base de $B(m,n)$.

Que $\{\xi_i \eta_j\}$ genera $B(m,n)$ es obvio. Además, si

$b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j = 0$, sean $x = e_i \in \mathbb{R}^m$
$y = e_j \in \mathbb{R}^n$

y entonces

$$\beta_{ij} = b(e_i, e_j) = 0 \qquad 1 \le i \le m \ , \ 1 \le j \le n$$

Por lo tanto $\{\xi_i \eta_j\}$ es independiente y genera $B(m,n)$

Por lo tanto $\{\xi_i \eta_j\}$ es una base de $B(m,n)$ y $\dim(B(m,n)) = mn$.

COROLARIO

$B(m,n)$ es isomorfo a $M(m,n)$

Demostración:

Sea $T : B(m,n) \to M(m,n)$ la transformación lineal dada por

$$T(\xi_i \eta_j) = e_i e_j^t \qquad \text{donde} \quad \begin{array}{l} e_i \in \mathbb{R}^m \\ e_j \in \mathbb{R}^n \end{array}$$

Entonces T es un isomorfismo entre $B(m,n)$ y $M(m,n)$.

Observemos que si

$$b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j \in B(m,n)$$

entonces $b(x,y) = x^t B y$ y

$$T(b) = B$$

▪

Denotaremos a $B(n,n)$ por B_n, i.e.

$$B_n = \left\{ b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \ : \ b(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \beta_{ij} \xi_i \eta_j \right\}$$

COROLARIO

B_n es isomorfo a M_n.

Recordemos que

$$Q_n = \left\{ \varphi : \mathbb{R}^n \longrightarrow \mathbb{R} \quad : \quad \varphi(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} \xi_i \xi_j \right\}$$

Como la dimensión de B_n es n^2 y la dimensión de $Q_n = \frac{n(n+1)}{2}$, entonces B_n y Q_n no pueden ser isomorfos. (y por lo tanto M_n y Q_n tampoco pueden ser isomorfos).

Sin embargo, encontraremos un subespacio de B_n, y uno de M_n, que sí sean isomorfos a Q_n.

Definimos una transformación

$$\tilde{T} : B_n \to Q_n$$

como sigue:

si $b \in B_n$ entonces $\tilde{T}(b) = \varphi \in Q_n$

donde

$$\tilde{T}(b)(x) = b(x,x) = \varphi(x)$$

$\tilde{T}$ es lineal, ya que si $b_1, b_2 \in B_n$, entonces

$$\tilde{T}(\alpha b_1 + b_2)(X) = (\alpha b_1 + b_2)(X,X)$$

$$= \alpha b_1(X,X) + b_2(X,X)$$

$$= \alpha \tilde{T}(b_1)(X) + \tilde{T}(b_2)(X)$$

$$\therefore \tilde{T}(\alpha b_1 + b_2) = \alpha \tilde{T}(b_1) + \tilde{T}(b_2) .$$

$\tilde{T}$ es una función sobre, ya que si tomamos cualquier $\varphi \in Q_n$, entonces sea $b \in B_n$ tal que

$$b(X,Y) = \frac{\varphi(X+Y) - \varphi(X) - \varphi(Y)}{2}$$

[Sabemos que $b \in B_n$ ya que anteriormente demostramos que si $\varphi \in Q_n$, entonces la función b definida arriba pertenece a B_n]

Entonces $b(X,X) = \varphi(X)$ y por lo tanto

$$\tilde{T}(b) = \varphi$$

Sin embargo $\tilde{T}$ no es una función inyectiva, ya que $\dim(B_n) > \dim(Q_n)$.

¿ Cuál es el kernel de $\tilde{T}$?

$$N(\tilde{T}) = \{ b \in B_n : \tilde{T}(b) = 0 \in Q_n \}$$

$$= \{ b \in B_n : b(x,x) = 0 \}$$

Supongamos $b \in N(\tilde{T}) \Rightarrow b(x,x) = 0 \quad \forall x \in \mathbb{R}^n$. Pero, como vimos anteriormente

$$b(x+y, x+y) = b(x,x) + b(y,y) + b(x,y) + b(y,x) = 0$$

$$\Rightarrow b(x,y) + b(y,x) = 0 \qquad \forall x,y \in \mathbb{R}^n$$

Inversamente, si $b(x,y) = -b(y,x) \quad \forall x,y \in \mathbb{R}^n$ entonces $b(x,x) = -b(x,x)$ y por lo tanto $b(x,x) = 0 \quad \forall x \in \mathbb{R}^n \quad \therefore \quad b \in N(\tilde{T})$

Por lo tanto

$$N(\tilde{T}) = \{ b \in B_n : b(x,y) = -b(y,x) \}$$

DEFINICION

Sea $b \in B_n$. Decimos que :

1) b es SIMETRICA si $b(x,y) = b(y,x) \quad \forall x,y \in \mathbb{R}^n$

2) b es ANTISIMETRICA si $b(x,y) = -b(y,x) \quad \forall x,y \in \mathbb{R}^n$

sea $\tilde{S}_n = \{ b \in B_n : b$ es simétrica $\}$

$\tilde{A}_n = \{ b \in B_n : b$ es antisimétrica $\}$

RESULTADO

$$B_n = \tilde{S}_n \oplus \tilde{A}_n$$

Demostración

Sea $b \in B_n$. Entonces b se puede escribir de la forma

$$b(x,y) = \frac{b(x,y) + b(y,x)}{2} + \frac{b(x,y) - b(y,x)}{2}$$

$\therefore \; b(x,y) = b_s(x,y) + b_{s^-}(x,y)$ donde

$$b_s \in \tilde{S}_n \quad y \quad b_{s^-} \in \tilde{A}_n .$$

además $\tilde{S}_n \cap \tilde{A}_n = \{0\}$, ya que si existe una $b \in B_n$ tal que $b \in \tilde{S}_n \cap \tilde{A}_n$ entonces

$$b(x,y) = b(y,x)$$
$$-b(x,y) = b(y,x)$$

$\therefore \; b(x,y) = -b(x,y) \quad \forall \, x,y$

$\Rightarrow \; b = 0$

CODOLARIO

$\tilde{S}_n$ es isomorfo a Q_n.

Demostración:

La transformación $\tilde{T} : B_n \to Q_n$ dada por $\tilde{T}(b)(x) = b(x,x)$ es lineal y sobre y $N(\tilde{T}) = \tilde{A}_n$.

Por otro lado $B_n = \tilde{S}_n \oplus \tilde{A}_n$

$\therefore$ $\tilde{T}|\tilde{S}_n$ es un isomorfismo de $\tilde{S}_n$ en Q_n ▪

Hemos visto que
$$M_n \cong B_n$$
$$\tilde{S}_n \cong Q_n$$

Es natural esperar que si

$$S_n = \{ A \in M_n : A^* = A \}$$

$$A_n = \{ A \in M_n : A^* = -A \}$$

entonces
$$S_n \cong \tilde{S}_n \cong Q_n$$
$$A_n \cong \tilde{A}_n$$

RESULTADO

$\tilde{S}_n$ es isomorfo a S_n y
$\tilde{A}_n$ es isomorfo a A_n.

Demostración:

Habíamos visto que la transformación

$T: B_n \to M_n$ dada por

$T(b) = B$, donde $b(X,Y) = \sum_{i=1}^{m} \sum_{J=1}^{n} \beta_{ij} \xi_i \eta_j$

y $B = [\beta_{ij}]$

es un isomorfismo entre B_n y M_n.

Si $b \in \tilde{S}_n$ entonces $b(e_i, e_j) = b(e_j, e_i)$
y por lo tanto $\beta_{ij} = \beta_{ji}$ $\therefore$ $B = B^t$

$\therefore$ $T(\tilde{S}_n) = S_n$

Si $b \in \tilde{A}_n$ entonces $-b(e_i, e_j) = b(e_j, e_i)$
$\therefore$ $-\beta_{ij} = \beta_{ji}$ $\therefore$ $-B = B^t$

$\therefore$ $T(\tilde{A}_n) = A_n$

Por lo tanto $S_n \cong \tilde{S}_n$ y $A_n \cong \tilde{A}_n$.

RESULTADO

$$M_n = S_n \oplus \dot{A}_n$$

Demostración

Por un lado tenemos que $B_n = \tilde{S}_n \oplus \tilde{A}_n$, y por otro B_n es isomorfo a M_n.

Por lo tanto, si $T : B_n \to M_n$ es un isomorfismo de B_n en M_n, entonces

$$T(B_n) = T(\tilde{S}_n) \oplus T(\tilde{A}_n)$$

$$\therefore \quad M_n = S_n \oplus \dot{A}_n \quad \blacksquare$$

RESULTADO

S_n es isomorfo a Q_n

Demostración

$$\tilde{S}_n \cong Q_n, \quad \tilde{S}_n \cong S_n \quad \Rightarrow \quad S_n \cong Q_n \quad \blacksquare$$

Por el resultado anterior, una forma cuadrática siempre la podemos escribir de la forma :

$$\varphi(x) = x^t A x \qquad \text{con} \quad A^t = A$$

Veremos que esta representación es adecuada para un cambio de coordenadas en $\mathbb{R}^n$:

Sea $\mathcal{N}(\varphi, cte) = \{ x \in \mathbb{R}^n : \varphi(x) = cte \}$ una curva de nivel de φ.

Si $\tilde{x} = Sx$, donde S es no-singular, representa un cambio de coordenadas en $\mathbb{R}^n$, entonces

$$S(\mathcal{N}(\varphi, cte)) = \{ \tilde{x} \in \mathbb{R}^n : \tilde{x} = Sx \quad \text{con} \quad x \in \mathcal{N}(\varphi, cte) \}$$

$$= \{ \tilde{x} \in \mathbb{R}^n : \tilde{x} = Sx \;, \ni \; \varphi(x) = cte \}$$

$$= \{ \tilde{x} \in \mathbb{R}^n : \tilde{x} = Sx \;; \; x^t A x = cte \}$$

$$= \{ \tilde{x} \in \mathbb{R}^n : (S^{-1}\tilde{x})^t A (S^{-1}\tilde{x}) = cte \}$$

$$= \{ \tilde{x} \in \mathbb{R}^n : \tilde{x}^t S^{-t} A S^{-1} \tilde{x} = cte \}$$

$$= \{ \tilde{x} \in \mathbb{R}^n : \tilde{x}^t \tilde{A} \tilde{x} = cte , \tilde{A} = S^{-t} A S^{-1} \}$$

$$= \mathcal{N}(\tilde{\varphi}, cte)$$

o sea, que al efectuar un cambio de coordenadas en $\mathbb{R}^n$, entonces la forma cuadrática $\varphi(x)$ se transforma en la forma cuadrática : $\tilde{\varphi}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x}$, donde $\tilde{A} = S^{-t} A S^{-1}$, y las curvas de nivel de $\mathcal{G}$ se transforman en curvas de nivel de $\tilde{\mathcal{G}}$.

Este mismo análisis se puede hacer para formas bilineales :

$$\begin{tikzcd}
\mathbb{R}^m \times \mathbb{R}^n \arrow[r, "b"] \arrow[d, "S"', shift right=4ex] \arrow[d, "C", shift left=4ex] & \mathbb{R} \arrow[d, "I"] \\
\mathbb{R}^m \times \mathbb{R}^n \arrow[r, "\tilde{b}"'] & \mathbb{R}
\end{tikzcd}$$

Si $b : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ es una forma bilineal :

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

y $\tilde{x} = Sx$, $\tilde{y} = Cy$ son cambios de coordenadas en $\mathbb{R}^m$ y $\mathbb{R}^n$ respectivamente (y por lo tanto S y C son no-singulares) entonces

$$b(x,y) = x^t B y = (S^{-1}\tilde{x})^t B (C^{-1}\tilde{y}) =$$

$$= \tilde{x}^t S^{-t} B C^{-1} \tilde{y} = \tilde{x}^t \tilde{B} \tilde{y} \quad \text{con } \tilde{B} = S^{-t} B C^{-1}$$

$$= \tilde{b}(\tilde{x}, \tilde{y}).$$

Por lo tanto, si b y $\tilde{b}$ representan la misma forma bilineal, entonces

$$\tilde{B} = S^{-t} B C^{-1} \quad , \text{ con } S \text{ y } C \text{ no-singulares.}$$

DEFINICION

Sean $b(x,y) = x^t B y$ y $\tilde{b}(\tilde{x},\tilde{y}) = \tilde{x}^t \tilde{B} \tilde{y}$ dos formas bilineales.

Decimos que b y $\tilde{b}$ son EQUIVALENTES si representan a la misma función bilineal con respecto a determinadas bases de $\mathbb{R}^m$ y $\mathbb{R}^n$, es decir, si existen $S \in M_m$ y $C \in M_n$ ambas no-singulares tales que

$$b(x,y) = \tilde{b}(Sx, Cy)$$

Escribimos $b(x,y) \sim \tilde{b}(\tilde{x},\tilde{y})$

Sean $B, \tilde{B} \in M(m,n)$. Decimos que B y $\tilde{B}$ son EQUIVALENTES si existen $S \in M_m$ y $C \in M_n$ ambas no-singulares tales que

$$B = S \tilde{B} C$$

Sean $\varphi(X) = X^t A X$ y $\tilde{\varphi}(\tilde{X}) = \tilde{X}^t \tilde{A} X$ (con $A = A^t$, $\tilde{A} = \tilde{A}^t$) dos formas cuadráticas.

Decimos que φ y $\tilde{\varphi}$ son CONGRUENTES si existe una matriz no-singular S tal que

$$\varphi(X) = \tilde{\varphi}(SX)$$

es decir, si φ y $\tilde{\varphi}$ representan la misma función con respecto a dos bases de $\mathbb{R}^n$.

Sean A y $\tilde{A} \in M_n$. Decimos que A es CONGRUENTE a $\tilde{A}$ si existe $S \in M_n$ no-singular tal que

$$A = S^t \tilde{A} S$$

Escribimos $\varphi(X) \sim \tilde{\varphi}(\tilde{X})$, $A \sim \tilde{A}$

A continuación veremos cómo transformar una forma cuadrática en otra forma cuadrática, para la cual sea más fácil decir cómo son sus curvas de nivel, mediante un cambio de coordenadas.

Sin embargo, antes de proseguir, veremos el concepto de función bilineal de $V \times W$ en $\mathbb{R}$, donde V y W son cualesquiera espacios vectoriales, y también veremos que las podemos estudiar por medio de formas bilineales en $\mathbb{R}^m \times \mathbb{R}^n$

DEFINICION

Sean V y W espacios vectoriales. Una FUNCION BILINEAL en $V \times W$ es una función

$F: V \times W \to \mathbb{R}$ tal que para todo $v, v_1, v_2 \in V$, $w, w_1, w_2 \in W$, $\alpha, \beta \in \mathbb{R}$ se tiene:

1) $F(\alpha v_1 + v_2, w) = \alpha F(v_1, w) + F(v_2, w)$

2) $F(v, \beta w_1 + w_2) = \beta F(v, w_1) + F(v, w_2)$

Es decir, F es bilineal si F es lineal en v (para w fijo) y en w (para v fijo).

Las funciones bilineales en $V \times W$ se pueden estudiar por medio de las formas bilineales en $\mathbb{R}^m \times \mathbb{R}^n$ (donde $\dim(V) = m$ y $\dim(W) = n$).

Esto se debe a que como $V \times W$ es isomorfo a $\mathbb{R}^m \times \mathbb{R}^n$, entonces $\tilde{L}(V \times W, \mathbb{R})$ es isomorfo a $\tilde{L}(\mathbb{R}^m \times \mathbb{R}^n, \mathbb{R})$,

donde $\tilde{L}(V \times W, \mathbb{R}) = \{F: V \times W \to \mathbb{R} : F \text{ bilineal}\}$.

Si tomamos coordenadas en V y en W entonces vamos a tener que:

$$\begin{tikzcd}
\mathbb{V} \arrow[d, "S_I"'] & \times & \mathbb{W} \arrow[d, "C_I"] \arrow[r, "F"] & \mathbb{R} \arrow[d, "I"] \\
\mathbb{R}^m & \times & \mathbb{R}^n \arrow[r, "b"'] & \mathbb{R}
\end{tikzcd}$$

Si $v \in \mathbb{V}$, $w \in \mathbb{W}$, y

$$S_I : \mathbb{V} \to \mathbb{R}^m$$
$$C_I : \mathbb{W} \to \mathbb{R}^n$$

son sistemas de coordenadas en $\mathbb{V}$ y $\mathbb{W}$ respectivamente, entonces

$$F(v,w) = b(S_I(v), C_I(w)).$$

$$= S_I(v)^t \, B \, C_I(w)$$

$$= X^t B Y$$

donde $X = S_I(v) \in \mathbb{R}^m$
$Y = C_I(w) \in \mathbb{R}^n$

es decir, donde $X \in \mathbb{R}^m$ es el vector de coordenadas de $v \in \mathbb{V}$, y $Y \in \mathbb{R}^n$ es el vector de coordenadas de $w \in \mathbb{W}$,

Observemos que si $\{v_1, v_2, \ldots, v_m\}$ es la base de $\mathbb{V}$ que determina S_I, y $\{w_1, w_2, \ldots, w_n\}$ es la base de W que determina C_I, entonces

$$F(v_i, w_j) = e_i^t B e_j = \beta_{ij} \qquad e_i \in \mathbb{R}^m, \; e_j \in \mathbb{R}^n$$
$$1 \leq i \leq m, \quad 1 \leq j \leq n \, .$$

Por lo tanto, si tenemos una función bilineal $F: V \times W \to \mathbb{R}$, entonces

$$F(v,w) = x^t B y$$

donde B es la matriz dada por

$$\beta_{ij} = F(v_i, w_j)$$

con $\{v_1, v_2, \ldots, v_m\}$ una base de V,

$\{w_1, w_2, \ldots, w_n\}$ una base de W

y donde $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$ son los vectores de coordenadas de $v \in V$ y $w \in W$ con respecto a los bases en cuestión.

EJEMPLOS

Sea $I = [\tau_1, \tau_r] \subseteq \mathbb{R}$

P: $\tau_1 < \tau_2 < \cdots < \tau_n$ una partición de I

y sean

$$I_1 = [\tau_1, \tau_2], \quad I_2 = [\tau_2, \tau_3], \ldots, \quad I_{r-1} = [\tau_{r-1}, \tau_r]$$

Sea $\mathcal{S}_{m,k} = \{ f : I \to \mathbb{R} : f|_{I_i}$ es un polinomio de grado $\leq m$ y f es C^k en $I \}$

Entonces

1) La función $\mathcal{F} : \mathcal{S}_{m,m} \times \mathcal{S}_{n,n} \to \mathbb{R}$ dada por

$$\mathcal{F}(p,q) = \sum_{i=1}^{r} p(\tau_i)\, q(\tau_i) \quad , \text{ donde } \begin{array}{l} p \in \mathcal{S}_{m,m} \\ q \in \mathcal{S}_{n,n} \end{array}$$

es bilineal.

(Recordemos que

$\mathcal{S}_{m,m} = \{ f : I \to \mathbb{R} \mid f$ es un polinomio grado $\leq m \}$)

Veremos que $\mathcal{F}$ es lineal en la primera componente:

si $p, p_1 \in \mathcal{S}_{m,m}$ entonces

$$\mathcal{F}(\alpha p + p_1, q) = \sum_{i=1}^{r} (\alpha p + p_1)(\tau_i)\, q(\tau_i)$$

$$= \sum_{i=1}^{r} [\alpha p(\tau_i) + p_1(\tau_i)]\, q(\tau_i)$$

$$= \sum_{i=1}^{r} \alpha p(\tau_i) q(\tau_i) + \sum_{i=1}^{r} p_1(\tau_i) q(\tau_i)$$

$$= \alpha F(p,q) + F(p_1, q) .$$

Por lo tanto F es lineal en p. Analogamente se ve que F es lineal en q.

Una base de $\mathcal{P}_{m,m}$ es $\{1, \tau, \tau^2, \ldots, \tau^m\}$, y una base de $\mathcal{P}_{n,n}$ es $\{1, \tau, \tau^2, \ldots, \tau^n\}$

Por lo tanto si B es la matriz asociada a F en estas bases, entonces

$$\beta_{ij} = F(\tau^{i-1}, \tau^{j-1}) = \sum_{k=1}^{r} (\tau_k)^{i-1} (\tau_k)^{j-1} = \sum_{k=1}^{r} \tau_k^{i+j-2}$$

$$\therefore \quad B = \begin{bmatrix} 1 & \sum_{k=1}^{r} \tau_k & \sum_{k=1}^{r} \tau_k^2 & \cdots & \sum_{k=1}^{r} \tau_k^{n-1} \\ \sum_{k=1}^{r} \tau_k & \sum_{k=1}^{r} \tau_k^2 & \sum_{k=1}^{r} \tau_k^3 & \cdots & \sum_{k=1}^{r} \tau_k^{n} \\ \sum_{k=1}^{r} \tau_k^2 & \sum_{k=1}^{r} \tau_k^3 & \sum_{k=1}^{r} \tau_k^4 & \cdots & \sum_{k=1}^{r} \tau_k^{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \sum_{k=1}^{r} \tau_k^{m-1} & \sum_{k=1}^{r} \tau_k^{m} & \sum_{k=1}^{r} \tau_k^{m+1} & \cdots & \sum_{k=1}^{r} \tau_k^{m+n-2} \end{bmatrix}$$

2) La función $F: \mathcal{S}_{m,2} \times \mathcal{S}_{n,2} \to \mathbb{R}$ dada por

$$F(p,q) = \int_{\tau_1}^{\tau_r} p''(\sigma)\, q''(\sigma)\, d\sigma$$

donde $p \in \mathcal{S}_{m,2}$ y $q \in \mathcal{S}_{n,2}$ es bilineal :

Si p_1 y $p_2 \in \mathcal{S}_{m,2}$ y $q \in \mathcal{S}_{n,2}$ y $\alpha \in \mathbb{R}$, entonces

$$F(\alpha p_1 + p_2, q) = \int_{\tau_1}^{\tau_r} (\alpha p_1 + p_2)''(\sigma)\, q''(\sigma)\, d\sigma$$

$$= \int_{\tau_1}^{\tau_r} (\alpha p_1'' + p_2'')(\sigma)\ q''(\sigma)\, d\sigma$$

$$= \int_{\tau_1}^{\tau_r} \alpha p_1''(\sigma)\, q''(\sigma)\, d\sigma + \int_{\tau_1}^{\tau_r} p_2''(\sigma)\, q''(\sigma)\, d\sigma$$

$$= \alpha\, F(p_1, q) + F(p_2, q)$$

Por lo tanto F es lineal en p, y de manera análoga se ve que F es lineal en q.

■

3) Sea $A = [\alpha_{ij}] \in M_n$. la TRAZA de A es

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

Entonces la función $F: M_n \times M_n \to \mathbb{R}$ dada por

$$F(A,B) = tr(ASB)$$

donde $A, S, B \in M_n$ es bilineal.

Primero observemos que si

$$A = \begin{bmatrix} r_1{}^t \\ r_2{}^t \\ \vdots \\ r_n{}^t \end{bmatrix} \qquad B = [b_1 | b_2 | \cdots | b_n]$$

entonces $tr(AB) = \sum_{i=1}^{n} r_i{}^t b_i$

Por lo tanto, si $A, C \in M_n$; $C = \begin{bmatrix} z_1{}^t \\ z_2{}^t \\ \vdots \\ z_n{}^t \end{bmatrix}$

entonces

$$tr[(\alpha A + C)(SB)] = \sum_{i=1}^{n} (\alpha r_i{}^t + z_i{}^t)(Sb_i)$$

$$= \sum_{i=1}^{n} (\alpha r_i{}^t S b_i) + \sum_{i=1}^{n} z_i{}^t S b_i$$

$$= \alpha \sum_{i=1}^{n} r_i^t (Sb_i) + \sum_{i=1}^{n} z_i^t (Sb_i)$$

$$= \alpha \, tr(ASB) + tr(CSB)$$

$$= \alpha F(A,B) + F(C,B)$$

Por lo tanto, F es lineal en A. De manera análoga se puede ver que F es lineal en B.

Por lo tanto F es bilineal.

Observemos además que $tr(A) = tr(A^t)$, y que por lo tanto

$$F(A,B) = tr(ASB)$$

$$= tr(B^t S^t A^t)$$

$\therefore$ si $S = S^t$

entonces $F(A,B) = tr(B^t S A^t)$

$$= F(B^t, A^t)$$

DEFINICION

Sea $\mathcal{V}$ un espacio vectorial de dimensión n, y sea $S_I : \mathcal{V} \to \mathbb{R}$ un sistema de coordenadas en $\mathcal{V}$

Decimos que $\tilde{\varphi} : \mathcal{V} \to \mathbb{R}$ es una FUNCION CUADRATICA en $\mathcal{V}$ si existe una forma cuadrática $\varphi : \mathbb{R}^n \to \mathbb{R}$ tal que

$$\tilde{\varphi}(v) = (\varphi \circ S_I)(v)$$

Es decir, si el siguiente diagrama es conmutativo

$$\begin{tikzcd}
\mathcal{V} \arrow[r, "\tilde{\varphi}"] \arrow[d, "S_I"'] & \mathbb{R} \arrow[d, "I"] \\
\mathbb{R}^n \arrow[r, "\varphi"'] & \mathbb{R}
\end{tikzcd}$$

Por lo tanto $\tilde{\varphi} : \mathcal{V} \to \mathbb{R}$ es una función cuadrática en $\mathcal{V}$ si

$$\tilde{\varphi}(v) = x^t A x \qquad , \; A = A^t$$

donde $x \in \mathbb{R}^n$ es el vector de coordenadas de $v \in \mathcal{V}$.

Toda la teoría que desarrollamos para formas cuadráticas en $\mathbb{R}^n$, y formas bilineales en $\mathbb{R}^m \times \mathbb{R}^n$ se aplica en estos casos.

Por ejemplo, decimos que $F: \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ es SIMETRICA si $F(v,w) = F(w,v)$ $\forall\, v,w \in \mathbb{V}$, y si

$\tilde{\varphi}: \mathbb{V} \to \mathbb{R}$ está definida de la siguiente forma:

$\tilde{\varphi}(v) = F(v,v)$ donde $F: \mathbb{V} \times \mathbb{V} \to \mathbb{R}$ es bilineal y simétrica, entonces $\tilde{\varphi}$ es una función cuadrática en $\mathbb{V}$.

EJEMPLOS

1) Sea $F: \mathcal{P}_{n,n} \times \mathcal{P}_{n,n} \to \mathbb{R}$ dada por

$$F(p,q) = \sum_{i=1}^{r} p(\tau_i)\, q(\tau_i) \qquad \text{donde} \quad p,q \in \mathcal{P}_{n,n}$$

Por lo tanto F es bilineal y simétrica

Por lo tanto la función $\tilde{\varphi}: \mathcal{P}_{n,n} \to \mathbb{R}$ dada por

$$\tilde{\varphi}(p) = F(p,p) = \sum_{i=1}^{r} p(\tau_i)^2$$

es una función cuadrática en $\mathcal{P}_{n,n}$

2) Sea $\mathcal{F} : \mathcal{S}_{n,2} \times \mathcal{S}_{n,2} \to \mathbb{R}$ dada por

$$\mathcal{F}(p,q) = \int_{\tau_1}^{\tau_r} p''(\sigma)\, q''(\sigma)\, d\sigma \quad , \quad p,q \in \mathcal{S}_{n,2}$$

Entonces $\mathcal{F}$ es bilineal y simétrica.

Por lo tanto si $\tilde{\varphi} : \mathcal{S}_{n,2} \to \mathbb{R}$ está dada por

$$\tilde{\varphi}(p) = \mathcal{F}(p,p) = \int_{\tau_1}^{\tau_r} (p''(\sigma))^2\, d\sigma$$

es una función cuadrática en $\mathcal{S}_{n,2}$.

3) Sea $\mathcal{F} : S_n \times S_n \to \mathbb{R}$ dada por

$$\mathcal{F}(A,B) = tr(ASB) \qquad \text{donde} \quad A,S,B \in S_n .$$

Entonces $\mathcal{F}$ es bilineal y $\mathcal{F}$ es simétrica, ya que $\mathcal{F}(A,B) = \mathcal{F}(B^t, A^t) = \mathcal{F}(B,A)$ si $B = B^t$ y $A = A^t$.

Por lo tanto la función $\tilde{\varphi} : S_n \to \mathbb{R}$ dada por

$$\tilde{\varphi}(A) = tr(ASA) \quad , \text{ donde } \quad S = S^t$$

es una función cuadrática en S_n.

REDUCCION DE FORMAS CUADRATICAS A FORMAS CANONICAS.

Las formas cuadráticas para las cuales es sencillo decir cómo son sus curvas[1] de nivel son aquellas que se reducen a una suma de cuadrados:

$$\varphi(x) = \sum_{i=1}^{n} \alpha_i \xi_i^2 = x^t A x,$$

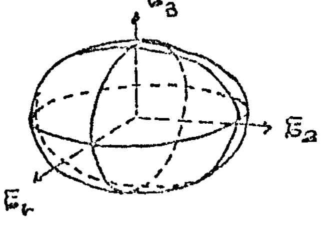
donde A es una matriz diagonal.

Veremos algunos ejemplos de superficies de nivel de formas cuadráticas diagonales para el caso $n=3$

1) $\varphi(x) = \delta_1 \xi_1^2 + \delta_2 \xi_2^2 + \delta_3 \xi_3^2$, con $\delta_i > 0$

$\mathcal{N}(\varphi,\gamma) = \{ x \in \mathbb{R}^3 : \varphi(x) = \gamma, \gamma > 0 \}$ es un

ellpsoide.

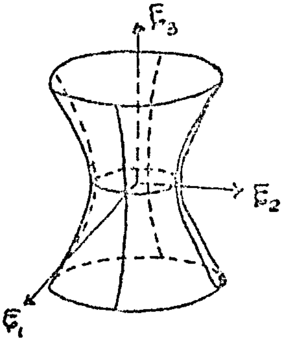


---

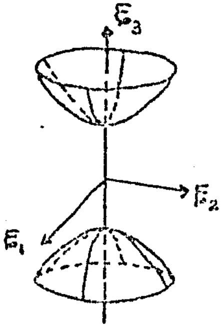
[1] Nosotros nos referimos a las "curvas" de nivel, aunque el objeto geométrico del cual estemos hablando no siempre sea una curva en $\mathbb{R}^n$. Por ejemplo, para $n=3$ tenemos superficies, y para el caso general "hipersuperficies"

2) Sea $\varphi(x) = \delta_1 \xi_1^2 + \delta_2 \xi_2^2 - \delta_3 \xi_3^2$, $\delta_i > 0$

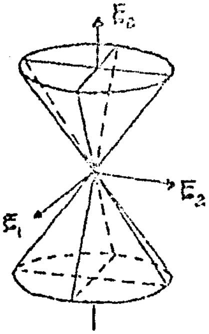
(a) $\pi(\varphi, \gamma)$ con $\gamma > 0$ es un hiperboloide de una hoja



(b) $\pi(\varphi, \gamma)$ con $\gamma < 0$ es un hiperboloide de dos hojas

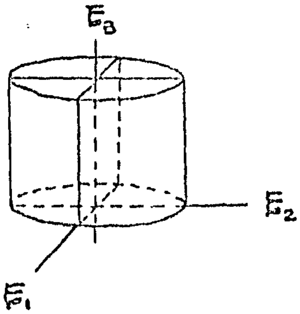


(c) $\mathcal{H}(\varphi, 0)$ es un cono



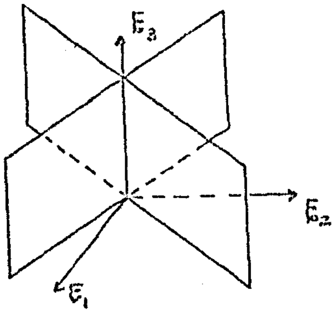
2) sea $\varphi(x) = \delta_1 \xi_1^2 + \delta_2 \xi_2^2$, $\delta_1, \delta_2 > 0$

$\mathcal{H}(\varphi, r)$, $r > 0$ es un cilindro elíptico

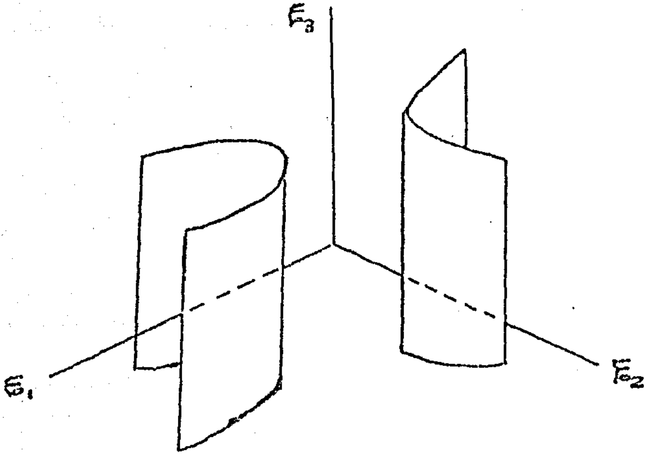


3) Sea $\varphi(x) = \delta_1 E_1^2 - \delta_2 E_2^2$, $\delta_1, \delta_2 > 0$

(a) $\mathcal{Z}(\varphi, 0)$ son dos planos que se cruzan .

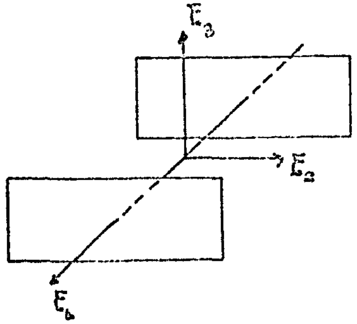


(b) $\mathcal{Z}(\varphi, \gamma)$, $\gamma \neq 0$ es un cilindro hiperbólico .



4) Sea $\varphi(x) = \delta_1 \xi_1^2$, $\delta_1 > 0$

(a) $\pi(\varphi,\gamma)$, $\gamma > 0$ son dos planos paralelos



(b) $\pi(\varphi,0)$ es el plano $\xi_1 = 0$.

Decimos que una forma cuadrática $\varphi : \mathbb{R}^n \to \mathbb{R}$ está en la forma canónica si

$$\varphi(X) = \xi_1^2 + \xi_2^2 + \cdots + \xi_p^2 - \xi_{p+1}^2 - \cdots - \xi_{p+q}^2$$

con $p+q \leq n$.

Por lo tanto $\varphi(X)$ es una forma canónica si

$$\varphi(X) = X^t A X \;, \quad \text{con} \quad A = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right] \;, \quad p+q \leq n \,.$$

Veremos que toda forma cuadrática homogénea se puede llevar a una forma canónica, es decir, toda forma cuadrática se puede reducir a una suma de a lo más n cuadrados, usando la técnica de completar cuadrados.

Antes de ver el caso general lo ilustraremos para el caso $n=2$:

sea $\varphi(X) = \alpha_{11}\xi_1^2 + 2\alpha_{12}\xi_1\xi_2 + \alpha_{22}\xi_2^2$

(a)

Si $\alpha_{11} > 0$, entonces

$$\varphi(X) = (\sqrt{\alpha_{11}}\,\xi_1)^2 + 2\alpha_{12}\xi_1\xi_2 + \left(\frac{\alpha_{12}}{\sqrt{\alpha_{11}}}\xi_2\right)^2 - \left(\frac{\alpha_{12}}{\sqrt{\alpha_{11}}}\xi_2\right)^2 + \alpha_{22}^2\xi_2^2$$

$$\therefore \varphi(X) = \left(\sqrt{\alpha_{11}}\,\xi_1 + \frac{\alpha_{12}}{\sqrt{\alpha_{11}}}\xi_2\right)^2 \pm \left(\sqrt{\left|\alpha_{22} - \frac{\alpha_{12}^2}{\alpha_{11}}\right|}\;\xi_2\right)^2$$

donde el signo ± se escoge según $\alpha_{22} - \dfrac{\alpha_{12}^2}{\alpha_{11}}$ sea positivo ó negativo.

sea $\tilde{\xi}_1 = \sqrt{\alpha_{11}}\,\xi_1 + \dfrac{\alpha_{12}}{\sqrt{\alpha_{11}}}\xi_2$

$$\tilde{\xi}_2 = \left(\left| \alpha_{22} - \frac{\alpha_{12}^2}{\alpha_{11}} \right| \right)^{1/2} \xi_2$$

Entonces $\varphi(X) = \tilde{\xi}_1^2 \pm \tilde{\xi}_2^2 = \tilde{\varphi}(\tilde{X})$

(b) si $\alpha_{11} < 0$ entonces

$$\varphi(X) = -\left[|\alpha_{11}| \xi_1^2 - 2\alpha_{12} \xi_1 \xi_2 \right] + \alpha_{22} \xi_2^2$$

$$= -\left(\sqrt{|\alpha_{11}|}\, \xi_1 - \frac{\alpha_{12}}{\sqrt{|\alpha_{11}|}} \xi_2 \right)^2 + \left(\alpha_{22} + \frac{\alpha_{12}^2}{|\alpha_{11}|} \right) \xi_2^2$$

$$= -\left(\sqrt{|\alpha_{11}|}\, \xi_1 - \frac{\alpha_{12}}{\sqrt{|\alpha_{11}|}} \xi_2 \right)^2 \pm \left(\sqrt{\left| \alpha_{22} + \frac{\alpha_{12}^2}{|\alpha_{11}|} \right|}\, \xi_2 \right)^2,$$

$\therefore$ $\varphi(X) = -\tilde{\xi}_1^2 \pm \tilde{\xi}_2^2$

donde $\tilde{\xi}_1 = \sqrt{|\alpha_{11}|}\, \xi_1 - \dfrac{\alpha_{12}}{\sqrt{|\alpha_{11}|}} \xi_2$

$$\tilde{\xi}_2 = \sqrt{\left| \alpha_{22} + \frac{\alpha_{12}^2}{|\alpha_{11}|} \right|}\, \xi_2$$

(c) Si $\alpha_{11}=0$ y $\alpha_{22}\neq 0$, entonces

$$\varphi(x)=\bar{\alpha}_{11}\bar{\xi}_1^2+2\bar{\alpha}_{12}\bar{\xi}_1\bar{\xi}_2=\bar{\varphi}(\bar{x})$$

donde $\bar{\alpha}_{11}=\alpha_{22}\neq 0$; $\bar{\xi}_1=\xi_2$
$\bar{\alpha}_{12}=\alpha_{12}$; $\bar{\xi}_2=\xi_1$

Por lo tanto, como $\bar{\alpha}_{11}\neq 0$,

$\varphi(x)=\bar{\varphi}(\bar{x})=\tilde{\varphi}(\tilde{x})$, donde $\tilde{\varphi}(\tilde{x})$ se obtiene a partir de $\bar{\varphi}(\bar{x})$ completando cuadrados, usando la técnica del inciso (a) ó (b) según $\bar{\alpha}_{11}$ sea positivo ó negativo

(d) Si $\alpha_{11}=0$ y $\alpha_{22}=0$; $\alpha_{12}\neq 0$, entonces

$$\varphi(x)=2\alpha_{12}\xi_1\xi_2$$

Sean $\bar{\xi}_1=\xi_1+\xi_2$

$\bar{\xi}_2=\xi_1-\xi_2$

Por lo tanto $\bar{\xi}_1^2-\bar{\xi}_2^2=4\xi_1\xi_2$

$$\therefore\ \varphi(x)=\frac{1}{2}\alpha_{12}\bar{\xi}_1^2-\frac{1}{2}\alpha_{12}\bar{\xi}_2^2$$

$$=\pm\left(\sqrt{\tfrac{1}{2}|\alpha_{12}|}\ \bar{\xi}_1\right)^2\mp\left(\sqrt{\tfrac{1}{2}|\alpha_{12}|}\ \bar{\xi}_2\right)^2$$

$$\therefore \quad \varphi(x) = \pm \tilde{\xi}_1^2 \mp \tilde{\xi}_2^2$$

donde

$$\tilde{\xi}_1 = \sqrt{\tfrac{1}{2}|\alpha_{12}|}\ \bar{\xi}_1 \ , \quad \tilde{\xi}_2 = \sqrt{\tfrac{1}{2}|\alpha_{12}|}\ \bar{\xi}_2$$

Por lo tanto, cualquier forma cuadrática $\varphi : \mathbb{R}^2 \to \mathbb{R}$, la podemos reducir a una suma de a lo más 2 cuadrados.

Ahora veremos el caso $\varphi : \mathbb{R}^n \to \mathbb{R}$.

Nota:

En lo que sigue escribiremos

$$\varphi(x) = \varphi(\xi_1, \xi_2, \dots, \xi_n)$$

para resaltar que φ depende de las variables $\xi_1, \xi_2, \dots, \xi_n$.

LEMA 1.

Sea $\varphi(x)$ una forma cuadrática en las variables $\xi_1, \xi_2, \dots, \xi_n$:

$$\varphi(x) = \varphi(\xi_1, \xi_2, \dots, \xi_n) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} \xi_i \xi_j$$

donde $\alpha_{11} \neq 0$

Entonces existe un cambio de coordenadas en $\mathbb{R}^n$

$$\bar{X} = UX$$

tal que

$$\varphi(X) = \pm \bar{\xi}_1^2 + \bar{\varphi}(\bar{\xi}_2, \bar{\xi}_3, \ldots, \bar{\xi}_n)$$

Demostración :

$$\text{Sea} \quad \varphi(X) = X^t A X = \varphi(\xi_1, \xi_2, \cdots, \xi_n)$$

$$\begin{aligned} \therefore \quad \varphi(X) = \ & \alpha_{11}\xi_1^2 + 2\alpha_{12}\xi_1\xi_2 + \cdots + 2\alpha_{1n}\xi_1\xi_n + \\ & + \alpha_{22}\xi_2^2 + 2\alpha_{23}\xi_2\xi_3 + \cdots + 2\alpha_{2n}\xi_2\xi_n + \\ & + \cdots + \alpha_{nn}\xi_n^2 \end{aligned}$$

Agrupemos los términos en los que aparece ξ_1 para completar un cuadrado perfecto :

$$\alpha_{11}\xi_1^2 + 2\alpha_{12}\xi_1\xi_2 + \cdots + 2\alpha_{1n}\xi_1\xi_n =$$

$$= \frac{\alpha_{11}}{|\alpha_{11}|}\left[|\alpha_{11}|\,\xi_1^2 + 2\frac{|\alpha_{11}|}{\alpha_{11}}\left(\alpha_{12}\xi_2 + \alpha_{13}\xi_3 + \cdots + \alpha_{1n}\xi_n\right)\xi_1 \right]$$

$$= \frac{\alpha_{11}}{|\alpha_{11}|}\left[\left(\sqrt{|\alpha_{11}|}\,\xi_1\right)^2 + 2\frac{|\alpha_{11}|}{\alpha_{11}}\,\xi_1\left(\alpha_{12}\xi_2 + \cdots + \alpha_{1n}\xi_n\right)\right.$$

$$+\left(\frac{1}{\sqrt{|\alpha_{11}|}}\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)\right)^2$$

$$\left.-\left(\frac{1}{\sqrt{|\alpha_{11}|}}\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)\right)^2\right]$$

$$= \frac{\alpha_{11}}{|\alpha_{11}|}\left[\left(\sqrt{|\alpha_{11}|}\,\xi_1\right) + \left(\frac{|\alpha_{11}|}{\alpha_{11}}\right)\left(\frac{1}{\sqrt{|\alpha_{11}|}}\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)\right)\right]^2$$

$$-\frac{\alpha_{11}}{|\alpha_{11}|}\left(\frac{1}{|\alpha_{11}|}\right)\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)^2$$

$$= \frac{\alpha_{11}}{|\alpha_{11}|}\left[\left(\sqrt{|\alpha_{11}|}\,\xi_1\right) + \left(\frac{|\alpha_{11}|}{\alpha_{11}}\right)\frac{1}{\sqrt{|\alpha_{11}|}}\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)\right]^2$$

$$-\frac{1}{\alpha_{11}}\left(\alpha_{12}\xi_2+\cdots+\alpha_{1n}\xi_n\right)^2$$

Por lo tanto

$$\varphi(x) = \frac{\alpha_{11}}{|\alpha_{11}|}\left[\sqrt{|\alpha_{11}|}\,\xi_1 + \left(\frac{|\alpha_{11}|}{\alpha_{11}}\right)\frac{1}{\sqrt{|\alpha_{11}|}}\sum_{i=2}^{n}\alpha_{1i}\xi_i\right]^2 \quad +$$

$$+ \ \alpha_{22}\xi_2^2 + \alpha_{33}\xi_3^2 + \cdots + \alpha_{nn}\xi_n^2 \ +$$

$$+ 2\left(\sum_{i=3}^{n}\alpha_{2i}\xi_2\xi_i + \sum_{i=4}^{n}\alpha_{3i}\xi_3\xi_i + \cdots + \sum_{i=n}^{n}\alpha_{n-1,i}\xi_{n-1}\xi_i\right)$$

$$- \frac{1}{\alpha_{11}}\left(\sum_{i=2}^{n}\alpha_{1i}\xi_i\right)^2$$

Sea $U = \begin{bmatrix} \sqrt{|\alpha_{11}|} & \pm\frac{\alpha_{12}}{\sqrt{|\alpha_{11}|}} & \pm\frac{\alpha_{13}}{\sqrt{|\alpha_{11}|}} & \cdots & \pm\frac{\alpha_{1n}}{\sqrt{|\alpha_{11}|}} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$

Como $\alpha_{11} \neq 0 \Rightarrow U$ es no-singular

Sea $\bar{X} = UX$ un cambio de coordenadas en $\mathbb{R}^n$.

Por lo tanto

$$\varphi(X) = \pm\bar{\xi}_1^2 + \bar{\varphi}(\bar{\xi}_2, \bar{\xi}_3, \ldots, \bar{\xi}_n)$$

$\square$

LEMA 2.

Sea $\varphi(x) = \varphi(\xi_1, \xi_2, \dots, \xi_n)$ una forma cuadrática en las variables $\xi_1, \xi_2, \dots, \xi_n$:

$$\varphi(x) = x^t A x \quad , \quad A = A^t$$

tal que $\varphi \neq 0$; $\alpha_{11} = \alpha_{22} = \dots = \alpha_{nn} = 0$ y tal que φ sí depende de la variable ξ_1

Entonces existe un cambio de coordenadas en $\mathbb{R}^n$, $\hat{x} = Sx$, tal que

$$\varphi(x) = \hat{\varphi}(\hat{x}) = \hat{x}^t \hat{A} \hat{x} \quad , \quad \text{donde } \hat{\alpha}_{11} \neq 0$$

Demostración :

Como $\varphi \neq 0$ y φ sí depende de la variable ξ_1 $\Rightarrow$ $\alpha_{1i} \neq 0$ para alguna $i > 1$

Sean $\hat{\xi}_1 = \xi_1 + \xi_i$

$\hat{\xi}_i = \xi_1 - \xi_i$

$\hat{\xi}_k = \xi_k$ para $k \neq 1, i$

Por lo tanto

$$\hat{\xi}_1^2 - \hat{\xi}_i^2 = 4\xi_1\xi_i$$

sea $\hat{x} = Sx$ donde

$$S = \begin{array}{cc} & \begin{array}{cc} i & \end{array} \\ \begin{array}{c} i \\ \\ \end{array} & \left[\begin{array}{c|c} S_1 & 0 \\ \hline 0 & I_{n-i} \end{array}\right] n-i \\ & \begin{array}{cc} & n-i \end{array} \end{array}$$

$$y \; S_1 = \begin{array}{cc} \begin{array}{ccccc} & & & i & \end{array} & \\ \left[\begin{array}{ccccc} 1 & 0 & \cdots & 0 & 1 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 1 & 0 & \cdots & 0 & -1 \end{array}\right] & i \end{array}$$

$\therefore$ S es no-singular .

Como $\alpha_{11} = \alpha_{22} = \cdots = \alpha_{nn} = 0$ entonces

$$\varphi(x) = 2\left[\sum_{j=2}^{n} \alpha_{1j}\xi_1\xi_j + \sum_{j=3}^{n} \alpha_{2j}\xi_2\xi_j + \cdots + \sum_{j=n}^{n} \alpha_{n-1,j}\xi_{n-1}\xi_j \right]$$

Veremos qué forma tienen los productos $\xi_k\xi_\ell$ en el nuevo sistema de coordenadas:

$$\xi_1\xi_i = \frac{\hat{\xi}_1^2}{4} - \frac{\hat{\xi}_i^2}{4}$$

$$\xi_1\xi_k = \left(\frac{\hat{\xi}_1 + \hat{\xi}_i}{2}\right)\hat{\xi}_k \qquad \text{si } k \neq 1$$

$$\xi_\ell\xi_k = \hat{\xi}_\ell\hat{\xi}_k \qquad \text{si } \ell,k \neq 1,i$$

$$\xi_i\xi_k = \left(\frac{\hat{\xi}_1 - \hat{\xi}_i}{2}\right)\hat{\xi}_k$$

Por lo tanto, al efectuar el cambio de coordenadas, el termino $\hat{\xi}_1^2$ unicamente es producido por el producto $\xi_1\xi_i$, y por lo tanto no puede ser cancelado.

Por lo tanto $\varphi(x) = \hat{\varphi}(\hat{x}) = \hat{x}^*\hat{A}x$, donde

$$\hat{\alpha}_{11} = \frac{1}{4} \neq 0$$

RESULTADO

Sea $\varphi(x) = \varphi(\xi_1, \xi_2, \dots, \xi_n)$ una forma cuadrática en las variables $\xi_1, \xi_2, \dots, \xi_n$. Entonces

$\varphi(x) \sim \tilde{\varphi}(\tilde{x})$ donde

$$\tilde{\varphi}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \cdots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \tilde{\xi}_{p+2}^2 - \cdots - \tilde{\xi}_{p+q}^2$$

con $p+q \leq n$.

Es decir, toda forma cuadrática se puede reducir a una suma de a lo más n cuadrados.

Demostración:

La demostración es por inducción sobre el número de variables.

Para $n=1$ se cumple el resultado:

Si $\varphi(X) = \alpha_{11}\xi_1^2 \Rightarrow \varphi(X) = \pm\left(\sqrt{|\alpha_{11}|}\ \xi_1\right)^2 = \tilde{\varphi}(\tilde{X})$

Supongamos cierto el resultado para formas cuadráticas que dependan de $n-1$ variables.

Sea $\varphi(X) = \varphi(\xi_1, \xi_2, \dots, \xi_n) = X^t A X$

Hay que considerar dos casos:

I) al menos algún α_{ii} es diferente de cero.

Sea $X' = P_{1i}X$. Entonces

$\varphi(X) \sim \varphi'(X') = (X')^t A'(X')$ donde $\alpha'_{11} = \alpha_{ii} \neq 0$.

Por lo tanto, por el lema 1, existe un cambio de coordenadas tal que

$$\varphi'(X') = \pm\bar{\xi}_1^2 + \bar{\varphi}(\bar{\xi}_2, \bar{\xi}_3, \dots, \bar{\xi}_n)$$

Como la forma cuadrática $\bar{\varphi}$ depende de $n-1$ variables, podemos aplicar la hipótesis de inducción, para obtener que

$$\varphi'(X') \sim \varphi''(X'') = \pm\xi_1''^2 \pm \xi_2''^2 \pm \dots \pm \xi_{p+q}''^2$$

donde $p+q \leq n$.

Mediante un cambio de coordenadas, dado por una permutación, podemos ordenar los cuadrados de φ'' de manera que siempre aparezcan primero los que no están afectados por el signo $-$.

Por lo tanto

$$\varphi(x) \sim \varphi'(x') \sim \varphi''(x'') \sim \tilde{\varphi}(\tilde{x}) \quad , \text{ donde}$$

$$\tilde{\varphi}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \cdots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \cdots - \tilde{\xi}_{p+q}^2$$

II) Supongamos $\alpha_{11} = \alpha_{22} = \cdots = \alpha_{nn} = 0$

A) Si $\alpha_{1i} = \alpha_{i1} = 0$ para i, $1 \leq i \leq n$, entonces eso significa que la variable ξ_1 no aparece en la forma cuadrática:

$$\varphi(x) = \varphi(\xi_1, \xi_2, \ldots, \xi_n) = \varphi(0, \xi_2, \ldots, \xi_n)$$

y por lo tanto φ depende únicamente de $n-1$ variables.

Mediante un cambio de coordenadas, dado por $x' = P_{1n} x$, tenemos que

$$\varphi(x) \sim \varphi'(x') \quad , \quad \text{con}$$

$$\varphi'(x') = \varphi'(\xi_1', \xi_2', \ldots, \xi_n') = \varphi'(\xi_1', \xi_2', \ldots, \xi_{n-1}', 0)$$

φ' depende únicamente de $n-1$ variables y por lo tanto podemos aplicar la hipótesis de inducción para obtener

$$\varphi(X) \sim \varphi'(X') \sim \bar{\varphi}(\bar{X}) \quad , \quad \text{donde}$$

$$\bar{\varphi}(\bar{X}) = \pm \bar{\xi}_1^2 \pm \bar{\xi}_2^2 \pm \cdots \pm \bar{\xi}_{p+q}^2$$

Mediante un cambio de coordenadas, dado por una permutación, podemos ordenar los cuadrados de manera que primero estén los que no están afectados por el signo $-$:

$$\varphi(X) \sim \bar{\varphi}(\bar{X}) \sim \tilde{\varphi}(\tilde{X}) \, .$$

$$\tilde{c}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \cdots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \cdots - \tilde{\xi}_{p+q}^2$$

B) Supongamos que existe alguna $i > 1$ tal que $\alpha_{ii} \neq 0$.

Entonces, por el lema 2, mediante un cambio de coordenadas, podemos obtener una forma $\hat{\varphi}(\hat{X})$ equivalente a $\varphi(X)$, tal que

$$\hat{\varphi}(\hat{X}) = \hat{X}^t \hat{A} \hat{X} \quad , \quad \text{con } \hat{\alpha}_{11} \neq 0 \, .$$

Por lo tanto $\hat{\varphi}$ se reduce al caso I, y

$$\varphi(X) \sim \hat{\varphi}(\hat{X}) \sim \tilde{\varphi}(\tilde{X}) \quad , \quad \text{donde}$$

$$\tilde{\varphi}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \dots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \dots - \tilde{\xi}_{p+q}^2$$

con $p+q \leq n$.

Por lo tanto, hemos demostrado que toda forma cuadrática $\varphi(x) = x^t A x$ es equivalente a una forma cuadrática $\tilde{\varphi}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x}$, donde

$$\tilde{A} = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right] \qquad p+q \leq n$$

Es decir, existe un cambio de coordenadas $\tilde{x} = Sx$ tal que

$$\varphi(x) = \tilde{\varphi}(Sx) \qquad y$$

$$A = S^t \tilde{A} S \quad ; \quad \tilde{A} = S^{-t} A S^{-1}$$

Daremos otra demostración de este resultado, tratando de encontrar la matriz S por un procedimiento diferente al de completar cuadrados.

Primero observemos que si $\tilde{\varphi}(\tilde{x}) = \tilde{x}^t \tilde{A} x$ donde $\tilde{A}$ es diagonal, es equivalente a $\varphi(x) = x^t A x$, entonces vamos a tener que

$$\tilde{\varphi}(\tilde{e}_i+\tilde{e}_j) = \varphi(\tilde{e}_i) + \varphi(\tilde{e}_j) + 2\,\tilde{e}_i^T\tilde{A}\,\tilde{e}_j$$

donde $\tilde{\varphi}(\tilde{e}_i) = \tilde{\alpha}_{ii}$ y

$$\tilde{e}_i^t\,\tilde{A}\,\tilde{e}_j = \tilde{\alpha}_{ij} \quad , \quad i \neq j$$

y como $\tilde{A}$ es diagonal

$$\tilde{\varphi}(\tilde{e}_i+\tilde{e}_j) = \tilde{\varphi}(\tilde{e}_i) + \tilde{\varphi}(\tilde{e}_j) , \quad i \neq j$$

Como $\varphi(x)$ es equivalente a $\tilde{\varphi}(\tilde{x})$, entonces

$$\tilde{x} = Sx$$

para alguna S no-singular.

Suponqamos $\tilde{e}_i = S s_i$

$\tilde{e}_j = S s_j$ donde $i \neq j$

por lo tanto

$$\tilde{\varphi}(\tilde{e}_i) + \tilde{\varphi}(\tilde{e}_j) = \varphi(s_i) + \varphi(s_j) = \tilde{\varphi}(\tilde{e}_i+\tilde{e}_j) =$$
$$= \varphi(s_i+s_j) \quad , \quad \text{para } i \neq j$$

Esto nos indica que para reducir una forma cuadrática a una suma de cuadrados tenemos que encontrar $s_1, s_2, \ldots, s_n \in \mathbb{R}^n$ tales que

$$\varphi(s_i+s_j) = \varphi(s_i)+\varphi(s_j) \quad \text{si } i\neq j$$

es decir, tales que $s_i^t A s_j^t = 0$ si $i \neq j$.

Consideremos el caso $n=3$.

Sea $s_1 \in \mathbb{R}^3$ $\quad s_1 \neq 0$

$$\therefore \quad \varphi(s_1+s_2) = \varphi(s_1)+\varphi(s_2)+2s_1^t A s_2$$

Por lo tanto queremos encontrar un vector $s_2 \in \mathbb{R}^3$ tal que

$$s_1^t A s_2 = 0$$

Observemos que $s_1^t A$ es una matriz 1×3, y queremos encontrar un $s_2 \in \mathbb{R}^3$ tal que $s_2 \in N(s_1^t A)$, $s_2 \neq 0$.

Pero $\dim(N(s_1^t A)) = 2$, y siempre podemos encontrar el vector s_2 que buscamos.

De manera similar, nos interesa encontrar un $s_3 \in \mathbb{R}^3$ tal que

$$s_1^t A s_3 = 0 \quad \text{y} \quad s_2^t A s_3 = 0$$

es decir, queremos un $s_3 \neq 0$ tal que

$$s_3 \in N((s_1^t A)) \quad \text{y} \quad s_3 \in N(s_2^t A)$$

Tal s_3 siempre existe, ya que

$$\dim\left(N(s_1^t A) \cap N(s_2^t A)\right) > 0 .$$

RESULTADO

Sea A una matriz simétrica. Entonces $\tilde{A} = S^{t} A S^{-1}$ para alguna matriz S no-singular, donde $\tilde{A}$ es diagonal.

Demostración:

La demostración es por inducción sobre n, donde $A \in M_n$.

Si $A = [\alpha_{11}]$, tomemos $S^{-1} = [1/\sqrt{|\alpha_{11}|}]$ (si $\alpha_{11} \neq 0$)

$$\therefore \quad \tilde{A} = [\pm 1] = S^{-t} A S^{-1} .$$

Por lo tanto para $n=1$ vale.

Supongamos cierto el resultado para $n-1$.

Sea $\varphi(x) = x^t A x$, $A = A^t$, $A \in M_n$.

El caso $A=0$ no nos interesa, por tanto supondremos $A \neq 0$ y $\therefore$ $\varphi \neq 0$.

Sea $s_1 \neq 0$, $s_1 \in \mathbb{R}^n$, tal que

$\varphi(s_1) \neq 0$, tal s_1 siempre existe ya que $\varphi \neq 0$.

Queremos obtener $n-1$ vectores $s_2, s_3, \ldots, s_n \in \mathbb{R}^n$ tal que

$$\varphi(s_1 + s_j) = \varphi(s_1) + \varphi(s_j) + 2\, s_1^t A s_j = \varphi(s_1) + \varphi(s_j)$$

es decir, tal que $s_1^t A s_j = 0$. si $j \neq 1$

Sea $b = A^t s_1 = A s_1 \in \mathbb{R}^n$, y sea

$$V(s_1) = \{ x \in \mathbb{R}^n : s_1^t A x = 0 \}$$

$$\therefore \; V(s_1) = \{ x \in \mathbb{R}^n : b^t x = 0 \}$$

Por lo. tanto nos interesa encontrar $n-1$ vectores independientes $s_2, s_3, \ldots, s_n$ tal que

$$s_i \in N(b^t)$$

Como $s_1 \neq 0$ y $\varphi(s_1) = s_1^t A s_1 \neq 0$
$\Rightarrow$ $b = A s_1 \neq 0$ y por lo tanto la dimensión de $N(b^t)$ es $n-1$.

Por lo tanto podemos encontrar $n-1$ vectores independientes tales que

$$s_1^t A s_j = 0 \qquad 2 \leq j \leq n$$

si tomamos $\{s_2, s_3, \ldots, s_n\}$ una base de $N(b^t)$.

Entonces en este caso vamos a tener que $\{s_1, s_2, \ldots, s_n\}$ es una base de $\mathbb{R}^n$. Sea $S^{-1} = [s_1 | s_2 | \cdots | s_n]$.

Por lo tanto, $x = S^{-1} \bar{x}$, donde

$x \in \mathbb{R}^n$ está expresado en la base canónica $\{e_1, e_2, \ldots, e_n\}$ de $\mathbb{R}^n$ y $\bar{x} \in \mathbb{R}^n$ está expresado en la base $\{s_1, s_2, \ldots, s_n\}$ de $\mathbb{R}^n$.

Sea $\phi(s_1) = \delta_1 = s_1^t A s_1$.

Por lo tanto, en la base $\{s_1, s_2, \ldots, s_n\}$ la matriz asociada a ϕ es

$$\tilde{A}_1 = S^{-t} A S^{-1} = \left[\begin{array}{c|c} \delta_1 & 0 \\ \hline 0 & A_1^{(1)} \end{array}\right]_{n-1} \quad \text{con } A_1^{(1)} \text{ simétrica}$$

$$\hphantom{\tilde{A}_1 = S^{-t} A S^{-1} = \quad\quad} n-1$$

y por lo tanto $A \sim A_1$.

Como $A_1^{(1)}$ es simétrica de orden $n-1$, podemos aplicar la hipótesis de inducción:

Existe una matriz C_1^{-1} de orden $n-1$, no singular, tal que

$$\bar{C}_1^{-t} A_1^{(1)} C_1^{-1} = \begin{array}{cc} \begin{bmatrix} \delta_2 & & & \\ & \delta_3 & & \\ & & \ddots & \\ & & & \delta_n \end{bmatrix} & n-1 \\ n-1 & \end{array}$$

Sea C_1 la inversa de C_1^{-1}, y sea

$$C = \begin{array}{cc} & \begin{array}{cc} 1 & n-1 \end{array} \\ \begin{array}{c} 1 \\ n-1 \end{array} & \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & C_1 \end{array}\right] \end{array}$$

por lo tanto C es no-singular, y

$\tilde{x} = (CS)x = C(\bar{x})$ representa un cambio de coordenadas en $\mathbb{R}^n$.

Por lo tanto $q(x) \sim \tilde{q}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x}$

Por lo tanto $A \sim \tilde{A}$.

Pero $x = (CS)^{-1} \tilde{x}$ $\therefore$ $x = S^{-1} C^{-1} \tilde{x}$

$$\therefore \quad q(x) = (S^{-1}C^{-1}\tilde{x})^t A (S^{-1}C^{-1}\tilde{x})$$

$$= \tilde{x}^t (C^{-t} S^{-t} A S^{-1} C^{-1}) \tilde{x}$$

Por lo tanto $\tilde{A} = C^{-t} S^{-t} A S^{-1} C^{-1}$.

Veremos que $\tilde{A}$ es diagonal :

$$S^{-t} A S^{-1} = \begin{array}{cc} \begin{array}{c} 1 \\ n-1 \end{array} & \left[\begin{array}{c|c} \delta_1 & 0 \\ \hline 0 & A_1^{(1)} \end{array}\right] \\ & \begin{array}{cc} 1 & n-1 \end{array} \end{array} = \tilde{A}_1$$

$$\therefore \quad \tilde{A} = C^{-t} \tilde{A}_1 C^{-1}$$

$$= \begin{array}{cc} \begin{array}{cc} 1 & n-1 \end{array} & \\ \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & C_1^{-t} \end{array}\right] & \begin{array}{c} 1 \\ n-1 \end{array} \end{array} \begin{array}{cc} \begin{array}{cc} 1 & n-1 \end{array} & \\ \left[\begin{array}{c|c} \delta_1 & 0 \\ \hline 0 & A_1^{(1)} \end{array}\right] & \begin{array}{c} 1 \\ n-1 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & C_1^{-1} \end{array}\right] \end{array}$$

$$= \begin{array}{cc} \begin{array}{cc} 1 & n-1 \end{array} & \\ \left[\begin{array}{c|c} \delta_1 & 0 \\ \hline 0 & C_1^{-t} A_1^{(1)} \end{array}\right] & \begin{array}{c} 1 \\ n-1 \end{array} \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & C_1^{-1} \end{array}\right] \end{array} = \begin{array}{cc} & \begin{array}{cc} 1 & n-1 \end{array} \\ \begin{array}{c} 1 \\ n-1 \end{array} & \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & C_1^{-t} A_1^{(1)} C_1^{-1} \end{array}\right] \end{array}$$

$$= \begin{array}{cc} n & \\ \left[\begin{array}{cccc} \delta_1 & & & \\ & \delta_2 & & \\ & & \ddots & \\ & & & \delta_n \end{array}\right] & n . \end{array}$$

$\therefore$ Toda matriz simétrica es congruente a una matriz diagonal.

□

Ejemplo :

Sea $A = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 5 & -2 \\ 3 & -2 & -8 \end{bmatrix}$; $\varphi(x) = x^t A x$.

Sea $s_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\therefore$ $\varphi(s_1) = 5 \neq 0$.

$$s_1^t A = \begin{bmatrix} 2 & 5 & -2 \end{bmatrix} = b^t$$

$$\therefore \quad x \in N(b^t) \Leftrightarrow 2\xi_1 + 5\xi_2 - 2\xi_3 = 0$$

$$\xi_2 = 2, \ \xi_3 = 0 \Rightarrow \xi_1 = -5$$

$$\xi_2 = 0, \ \xi_3 = 1 \Rightarrow \xi_1 = 1$$

Por lo tanto si $s_2 = \begin{bmatrix} -5 \\ 2 \\ 0 \end{bmatrix}$ $s_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$

Entonces $\{ s_1, s_2, s_3 \}$ es una base de $\mathbb{R}^3$

Por lo tanto :

$\varphi(s_1) = 5$
$\varphi(s_2) = -20$
$\varphi(s_3) = -2$

$s_1^t A s_2 = s_1^t A s_3 = 0$

$s_2^t A s_3 = -15$

Por lo tanto, en la base $\{s_1, s_2, s_3\}$ la matriz asociada a φ es

$$\tilde{A}_1 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & -20 & -15 \\ 0 & -15 & -2 \end{bmatrix}$$

y $\tilde{A}_1 = S^{-t} A S^{-1}$, donde

$$S^{-1} = \begin{bmatrix} 0 & -5 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

RESULTADO

Sea $\varphi(x) = x^t A x$ una forma cuadrática tal que $\varphi(x) \sim \tilde{\varphi}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x}$, donde

$$\tilde{A} = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right]$$

Entonces $r(A) = p+q$.

Demostración :

Como $\varphi(x) \sim \tilde{\varphi}(\tilde{x}) \Rightarrow$ existe S no-singular tal que

$$A = S^t \tilde{A} S$$

Como el rango de una matriz se conserva al multiplicar por matrices invertibles,

$$r(A) = r(\hat{A}) = p+q.$$

□

El siguiente resultado nos indica que la forma canónica que se obtiene es independiente de cómo se haya procedido para obtenerla.

TEOREMA DE LA INERCIA

Sea $\varphi(x) = x^t A x$ $(A = A^t)$ una forma cuadrática. Si existen S_1 y S_2 no-singulares tales que

$$S_1^{-t} A S_1^{-1} = \left[\begin{array}{c|c|c} I_{p_1} & & \\ \hline & -I_{q_1} & \\ \hline & & 0 \end{array}\right] = \tilde{A}$$

$$S_2^{-t} A S_2^{-1} = \left[\begin{array}{c|c|c} I_{p_2} & & \\ \hline & -I_{q_2} & \\ \hline & & 0 \end{array}\right] = \bar{A}$$

Entonces $p_1 = p_2$ y $q_1 = q_2$

Es decir, si una forma cuadrática se reduce a una suma de cuadrados en dos bases distintas de $\mathbb{R}^n$, entonces el número de cuadrados "positivos" y el número de cuadrados "negativos" es el mismo en los dos casos.

Demostración :

Sea $x \in \mathbb{R}^n$. Supongamos que $p_1 > p_2$, y que

$$\tilde{x} = S_1 x$$

$$\bar{x} = S_2 x$$

por lo tanto $\varphi(x) = \tilde{\varphi}_1(S_1 x) = \tilde{\varphi}_1(\tilde{x})$

$$\varphi(x) = \bar{\varphi}_2(S_2 x) = \bar{\varphi}_2(\bar{x})$$

Supongamos

$$S_1^{-1} = [s_1^{(1)} | s_2^{(1)} | \cdots | s_n^{(1)}]$$

$$S_2^{-1} = [s_1^{(2)} | s_2^{(2)} | \cdots | s_n^{(2)}]$$

Sea $\mathcal{S}_1 \subseteq \mathbb{R}^n$ el subespacio generado por las primeras p_1 columnas de S_1^{-1}:

$$\mathcal{S}_1 = [s_1^{(1)}, s_2^{(1)}, \ldots, s_{p_1}^{(1)}] \quad ,$$

y sea $\mathcal{S}_2 \subseteq \mathbb{R}^n$ el subespacio generado por las últimas $n-p_2$ columnas de S_2^{-1} :

$$\mathcal{S}_2 = [s_{p_2+1}^{(2)}, s_{p_2+2}^{(2)}, \ldots, s_n^{(2)}]$$

entonces $\mathcal{S}_1 \cap \mathcal{S}_2 \neq \{0\}$, ya que

$$\dim(\mathcal{S}_1 + \mathcal{S}_2) = \dim(\mathcal{S}_1) + \dim(\mathcal{S}_2) - \dim(\mathcal{S}_1 \cap \mathcal{S}_2) \leq n$$

$$\therefore \quad p_1 + (n - p_2) \leq n + \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$$

$$\therefore \quad 0 < p_1 - p_2 \leq \dim(\mathcal{S}_1 \cap \mathcal{S}_2)$$

$\therefore \quad \dim(\mathcal{S}_1 \cap \mathcal{S}_2) > 0 \qquad$ (ya que $p_1 > p_2$)

Sea $\quad x \neq 0 \quad$ tal que $\quad x \in \mathcal{S}_1 \cap \mathcal{S}_2$

Por lo tanto

$$x = \xi_1^{(1)} s_1^{(1)} + \xi_2^{(1)} s_2^{(1)} + \cdots + \xi_{p_1}^{(1)} s_{p_1}^{(1)} \qquad y$$

$$x = \xi_{p_2+1}^{(2)} s_{p_2+1}^{(2)} + \xi_{p_2+2}^{(2)} s_{p_2+2}^{(2)} + \cdots + \xi_n^{(2)} s_n^{(2)}$$

Sean

$$\tilde{x} = \begin{bmatrix} \xi_1^{(1)} \\ \vdots \\ \xi_{p_1}^{(1)} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \bar{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \xi_{p_2+1} \\ \vdots \\ \xi_n \end{bmatrix}$$

por lo tanto $\quad x = S_1^{-1} \tilde{x} = S_2^{-1} \bar{x} \quad$ y

$$\varphi(x) = \tilde{\varphi}(\tilde{x}) = \bar{\varphi}(\bar{x})$$

pero $\quad \tilde{\varphi}(\tilde{x}) = \tilde{x}^* \tilde{A} \tilde{x} = \sum_{i=1}^{p_1} (\xi_i^{(1)})^2 > 0 \qquad (\tilde{x} \neq 0)$

y $\quad \bar{\varphi}(\bar{x}) = \bar{x}^* \bar{A} \bar{x} = \sum_{i=p_3}^{n} -(\xi_i^{(2)})^2 \leq 0 \qquad (\bar{x} \neq 0)$

Observar que puede pasar que $\bar{\varphi}(\bar{x}) = 0$ a pesar de que $\bar{x}$ sea diferente de 0. Por ejemplo, esto sucede cuando $\bar{x} = e_n \in \mathbb{R}^n$ y $p_2 + q_2 < n$.

Por lo tanto $\varphi(x) > 0$ y $\varphi(x) \leq 0$!

Esto es una contradicción

$$\therefore \quad S_1 \cap S_2 = \{0\} \quad \text{y} \quad p_1 \leq p_2 .$$

De manera análoga se prueba que $p_2 \leq p_1$, y por lo tanto que $p_1 = p_2$.

De forma similar se demuestra que $q_1 = q_2$ ∎

DEFINICION .

Sea $\varphi(x) = x^t A x$ una forma cuadrática. Si $x^t A x > 0$ para todo $x \neq 0$ decimos que:

la forma cuadrática $\varphi(x)$ es POSITIVA DEFINIDA y que la matriz A es POSITIVA DEFINIDA; escribiremos $A > 0$.
(Observar que en esta definición estamos suponiendo que $A = A^t$).

RESULTADO

Sea $A > 0$ de orden n. Entonces $r(A) = n$.

Demostración :

Si $x \neq 0 \Rightarrow Ax \neq 0$ ya que $x^t(Ax) > 0 \quad \forall x \neq 0$

■

RESULTADO

Sea $A > 0$ de orden n, si $A \sim \tilde{A}$, donde

$$\tilde{A} = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right]$$

entonces $p = n$ y $q = 0$, es decir $\tilde{A} = I_n$

Demostración :

Como $A > 0$, por el resultado anterior, $r(A) = p + q = n$.

Suponqamos $q \neq 0$.

Sea $\varphi(x) = x^t A x$, por lo tanto $\varphi(x) > 0$ $\forall \, x \neq 0$.

Supongamos $\tilde{A} = S^{-t} A S^{-1}$, con S no-singular.

Por lo tanto $\tilde{e}_{p+1} = Sx$ para alguna $x \in \mathbb{R}^n$, $x \neq 0$.

$$\therefore \quad \varphi(x) = \tilde{\varphi}(\tilde{e}_{p+1}) = -1 < 0 \qquad (q \neq 0)$$

$\therefore$ $\varphi(x) > 0$ ya que $A>0$, y por otro lado $\varphi(x) < 0$!

Esto es una contradicción, por lo tanto $q=0$ y $p=n$

∎

RESULTADO

Sea $A \in M_n$ una matriz simétrica. Entonces $A > 0 \iff$ existe $S \in M_n$ no-singular tal que $A = S^tS$.

Demostración:

$\Leftarrow]$ Supongamos $A = S^tS$, con S invertible.

Entonces
$$\begin{aligned} x^tAx &= (xS^t)(Sx) \\ &= (Sx)^t(Sx) > 0 \quad \text{si } x \neq 0 \end{aligned}$$

Por lo tanto $A > 0$.

$\Rightarrow]$ Supongamos $A>0$. Por lo tanto, por el resultado anterior $A \sim I_n$ $\therefore$ existe

S no-singular tal que $A = S^t I_n S = S^t S$ ▫

RESULTADO

Sea $S \in M(m,n)$ una matriz con columnas independientes. Entonces $S^t S$ es positiva definida.

Demostración:

Sea $A = S^t S$. Como $N(S) = \{0\} \Rightarrow r(S) = r(S^t) = n$. Por lo tanto $r(A) = n$ y

$$x^t A x = x^t S^t S x = (Sx)^t (Sx) > 0 \quad \text{si} \quad x \neq 0.$$

▫

RESULTADO

Sea
$$A = \begin{array}{c|c} & \begin{array}{cc} k & n-k \end{array} \\ \begin{array}{c} k \\ n-k \end{array} & \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] \end{array}.$$
Entonces

$$A > 0 \iff A_{11} > 0 \ \text{y} \ A_{22} > 0.$$

Demostración:

$\Rightarrow$ Supongamos $A > 0$. Sea $x_1 \in \mathbb{R}^k$, $x_1 \neq 0$ y sea

$$x = \begin{bmatrix} x_1 \\ \hline 0 \end{bmatrix} \begin{array}{l} k \\ n-k \end{array} \in \mathbb{R}^n$$

Por lo tanto $x \neq 0$ y $x^t A x > 0$.

Pero
$$x^t A x = \begin{array}{cc} k & n-k \\ \left[\; x_1^t \right. & \left. 0 \;\right] \end{array} \begin{array}{r} \\ k \\ n-k \end{array} \begin{array}{c} \begin{array}{cc} k & \quad n-k \end{array} \\ \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{22} \end{array}\right] \end{array} \left[\begin{array}{c} x_1 \\ 0 \end{array}\right] \begin{array}{l} k \\ n-k \end{array}$$

$$= x_1^t A_{11} x_1$$

$\therefore \quad x_1^t A_{11} x_1 > 0$

$\therefore \quad A_{11} > 0$

análogamente se ve que $A_{22} > 0$.

$\Leftarrow]$ Supongamos que $A_{11} > 0$ y $A_{22} > 0$.

Sea $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^n, \quad x \neq 0$

$\therefore \quad x_1 \neq 0$ ó $x_2 \neq 0$ y

$$x^t A x = x_1^t A_{11} x_1 + x_2^t A_{22} x_2 > 0$$

■

RESULTADO

Supongamos $A \sim \tilde{A}$. Entonces $A > 0 \Leftrightarrow \tilde{A} > 0$.

Demostración :

Supongamos $A = S^t \tilde{A} S$ con S no-singular.

$\boxed{\Rightarrow}$ Si $A > 0$. Sea $\tilde{x} \in \mathbb{R}^n$, $\tilde{x} \neq 0$

$\therefore$ $\tilde{x} = Sx$ para alguna $x \neq 0$

$\therefore$ $x^t A x = \tilde{x}^t \tilde{A} \tilde{x} > 0$ $\therefore$ $\tilde{A} > 0$

$\boxed{\Leftarrow}$ Si $\tilde{A} > 0$, sea $x \in \mathbb{R}^n$, $x \neq 0$

$\therefore$ $x^t A x = (Sx)^t A \, (Sx) > 0$

□

RESULTADO

Sea $A > 0$ una matriz de orden n. Entonces existe $U \in M_n$ triangular superior no-singular tal que

$$A = U^t U$$

Demostración por Inducción sobre n :

Para $n=1$: si $A = [\alpha_{11}]$, entonces

$$A = [\sqrt{\alpha_{11}}][\sqrt{\alpha_{11}}] \qquad (A>0 \Rightarrow \alpha_{11}>0).$$

Supongamos cierto el resultado para matrices positivas definidas de orden $n-1$.

Sea $A>0$ de orden n. Por lo tanto $\varphi(X) = X^t A X$ es positiva definida y $\alpha_{11} = \varphi(e_1) > 0$.

Trataremos de encontrar una matriz U_1 de la forma

$$U_1 = \begin{array}{c|c|c|l} & 1 & n-1 & \\ 1 & \mu_1 & z_1^t & \\ n-1 & 0 & I_{n-1} & n-1 \end{array}$$

tal que $\qquad A = U_1^t \tilde{A}_1 U_1 \quad$ y $\quad U_1$ sea no-singular.

donde $\tilde{A}_1$ es de la forma

$$\tilde{A}_1 = \begin{array}{c|c|l} 1 & n-1 & \\ I_1 & 0 & 1 \\ 0 & A_1^{(1)} & n-1 \end{array}$$

y tal que $A_1^{(1)}$ (y por lo tanto A_1 también) sea positiva definida.

Resolveremos $U_1^t \tilde{A}_1 U_1 = A$:

$$\begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \mu_1 & 0 \\ \hline z_1 & I_{n-1} \end{array}\right] \end{array} \begin{array}{c} 1 \\ n-1 \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & A_1^{(1)} \end{array}\right] \end{array} \begin{array}{c} 1 \\ n-1 \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \mu_1 & z_1^t \\ \hline 0 & I_{n-1} \end{array}\right] \end{array} \begin{array}{c} 1 \\ = \\ n-1 \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \alpha_{11} & a^t \\ \hline a & A_1 \end{array}\right] \end{array} = A$$

$$\Rightarrow \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \mu_1 & 0 \\ \hline z_1 & A_1^{(1)} \end{array}\right] \end{array} \begin{array}{c} 1 \\ n-1 \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \mu_1 & z_1^t \\ \hline 0 & I_{n-1} \end{array}\right] \end{array} = \begin{array}{c} 1 \\ n-1 \end{array} \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \mu_1^2 & \mu_1 z_1^t \\ \hline \mu_1 z_1 & z_1 z_1^t + A_1^{(1)} \end{array}\right] \end{array} = A$$

$$\Rightarrow \begin{aligned} \mu_1^2 &= \alpha_{11} \\ \mu_1 z_1 &= a \\ z_1 z_1^t + A_1^{(1)} &= A_1 \end{aligned}$$

$$\therefore \quad \mu_1 = \sqrt{\alpha_{11}} \quad ; \quad z_1 = \frac{a}{\sqrt{\alpha_{11}}} \quad Y$$

$$A_1^{(1)} = A_1 - \frac{a}{\sqrt{\alpha_{11}}} \frac{a^t}{\sqrt{\alpha_{11}}} = A_1 - \frac{1}{\alpha_{11}} a a^t$$

Por lo tanto

$$U_1 = \begin{array}{c} \begin{array}{cc} 1 & n-1 \end{array} \\ \left[\begin{array}{c|c} \sqrt{\alpha_{11}} & (1/\sqrt{\alpha_{11}})\, a^t \\ \hline 0 & I_{n-1} \end{array}\right] \end{array} \begin{array}{c} 1 \\ n-1 \end{array}$$

;

es no-singular y por lo tanto A es congruente a $\tilde{A}_1$.

Como $A > 0 \Rightarrow \tilde{A}_1 > 0$ y por lo tanto, por un resultado anterior, $A_1^{(1)} > 0$.

Por lo tanto, podemos aplicar la hipótesis de inducción a $A_1^{(1)}$ para obtener que

$$A = U^* U.$$

□

Observar que si aplicamos inductivamente este procedimiento, obtenemos que

$$\begin{aligned} A &= U_1^* U_2^* \cdots U_n^* U_n \cdots U_2 U_1 \\ &= (U_n \cdots U_2 U_1)^* (U_n \cdots U_2 U_1) \\ &= U^* U, \end{aligned}$$

donde cada U_i es de la forma

$$U_i = \begin{array}{c|c|cc|} & i-1 & \multicolumn{2}{c}{n-(i-1)} \\ i-1 & I_{i-1} & & \\ \cline{2-4} 1 & & \multicolumn{2}{c|}{u_i} \\ \cline{3-4} n-i & 0 & 0 & I_{n-i} \\ & i-1 & 1 & n-i \end{array}$$

.

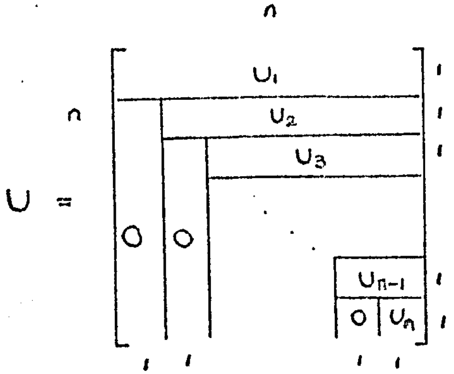
La matriz U es muy fácil de construir, ya que

$$
U_{i+1}U_i = \begin{array}{c}
 \\ i \\ \\ n-i \\ \\
\end{array}
\left[\begin{array}{c|cc}
i & \multicolumn{2}{c}{n-i} \\
I_i & \multicolumn{2}{c}{0} \\ \hline
 & \multicolumn{2}{c}{U_{i+1}} \\ \cline{2-3}
0 & \multicolumn{1}{c|}{0} & I_{n-(i+1)} \\
i & 1 & n-(i+1)
\end{array}\right]
\begin{array}{c}
 \\ \\ 1 \\ \\ \\
\end{array}
\left[\begin{array}{c|cc}
i-1 & \multicolumn{2}{c}{n-(i-1)} \\
I_{i-1} & \multicolumn{2}{c}{0} \\ \hline
 & \multicolumn{2}{c}{U_i} \\ \cline{2-3}
 & \multicolumn{1}{c|}{0} & I_{n-i} \\
\end{array}\right]
\begin{array}{c}
 \\ i-1 \\ 1 \\ n-i \\
\end{array}
$$

$$
= \begin{array}{c}
 \\ i-1 \\ \\ \\ n-(i-1) \\ \\
\end{array}
\left[\begin{array}{c|ccc}
i-1 & \multicolumn{3}{c}{n-(i-1)} \\
I_{i-1} & \multicolumn{3}{c}{0} \\ \hline
 & 1 & 0 \;\; 0 & \cdots \;\; 0 \\ \cline{2-4}
 & \multicolumn{1}{c|}{0} & \multicolumn{2}{c}{U_{i+1}} \\ \cline{2-4}
0 & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & I_{n-(i+1)} \\
i-1 & 1 & 1 & n-(i+1)
\end{array}\right]
\begin{array}{c}
 \\ i-1 \\ 1 \\ 1 \\ \\ \\
\end{array}
\left[\begin{array}{c|ccc}
i-1 & \multicolumn{3}{c}{n-(i-1)} \\
I_{i-1} & \multicolumn{3}{c}{0} \\ \hline
 & \multicolumn{3}{c}{U_i} \\ \cline{2-4}
 & \multicolumn{1}{c|}{0} & 1 & 0 \cdots 0 \\ \cline{2-4}
0 & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & I_{n-(i+1)} \\
i-1 & 1 & 1 & n-(i+1)
\end{array}\right]
\begin{array}{c}
 \\ i-1 \\ 1 \\ 1 \\ n-(i+1) \\ \\
\end{array}
$$

$$
= \begin{array}{c}
 \\ i-1 \\ \\ \\ n-(i-1) \\ \\
\end{array}
\left[\begin{array}{c|ccc}
i-1 & \multicolumn{3}{c}{n-(i-1)} \\
I_{i-1} & \multicolumn{3}{c}{0} \\ \hline
 & \multicolumn{3}{c}{U_i} \\ \cline{2-4}
 & \multicolumn{1}{c|}{} & \multicolumn{2}{c}{U_{i+1}} \\ \cline{3-4}
0 & \multicolumn{1}{c|}{0} & \multicolumn{1}{c|}{0} & I_{n-(i+1)} \\
i-1 & 1 & 1 & n-
\end{array}\right]
$$

Lo cual nos permite concluir que



DEFINICION

Sea $\varphi(x) = x^t A x$ una forma cuadrática.
Si $x^t A x < 0$ para toda $x \neq 0$ decimos que :

la forma cuadrática $\varphi(x)$ es NEGATIVA DEFINIDA y que la matriz A es NEGATIVA DEFINIDA.

Escribimos $A < 0$.

Nota: estamos suponiendo $A = A^t$.

RESULTADO

$$A<0 \iff (-A)>0$$

Demostración

$$\begin{aligned} A<0 &\iff x^t A x < 0 \quad \forall x \neq 0 \\ &\iff -(x^t A x) > 0 \quad \forall x \neq 0 \\ &\iff x^t(-A)x > 0 \quad \forall x \neq 0 \\ &\iff -A > 0 \end{aligned}$$

■

COROLARIO

1) Si $A \in M_n$ y $A<0 \Rightarrow r(A)=n$.

2) Si $A \in M_n$ y $A<0 \Rightarrow q=n, p=0 \, (A \sim -I_n)$

3) $A<0 \iff$ existe S no-singular tal que $A=-S^tS$

4) Si $S \in M(m,n)$ tiene columnas independientes $\Rightarrow -S^tS$ es negativa definida.

5) Sea $A = \left[\begin{array}{c|c} A_{11} & 0 \\ \hline 0 & A_{12} \end{array}\right]$. Entonces

$$A<0 \iff A_{11}<0 \text{ y } A_{22}<0$$

6) Supongamos $A \sim \tilde{A}$. Entonces $A<0 \iff \tilde{A}<0$

7) Sea $A<0$ de orden n. Entonces existe $U \in M_n$ triangular superior no-singular tal que

$$A = -U^t U$$

$\blacksquare$

DEFINICION

Sea $A = A^t$.
Decimos que:

1) A es SEMI-DEFINIDA POSITIVA si $x^t A x \geq 0 \ \forall x \in \mathbb{R}^n$.
Escribimos $A \geq 0$.

2) A es SEMI-DEFINIDA NEGATIVA si $x^t A x \leq 0 \ \forall x \in \mathbb{R}^n$
Escribimos $A \leq 0$

3) A es INDEFINIDA si existen $x_1, x_2 \in \mathbb{R}^n$ tales que $x_1^t A x_1 > 0$ y $x_2^t A x_2 < 0$

FUNCIONES CUADRATICAS NO-HOMOGENEAS

Sea $f(x) = x^t A x + b^t x + \gamma$ una función cuadrática no-homogénea.

Decimos que $f(x)$ está en la forma canónica si

$$f(x) = \xi_1^2 + \xi_2^2 + \cdots + \xi_p^2 - \xi_{p+1}^2 - \cdots - \xi_{p+q}^2 + \gamma \quad \text{ó}$$

$$f(x) = \xi_1^2 + \xi_2^2 + \cdots + \xi_p^2 - \xi_{p+1}^2 - \cdots - \xi_{p+q}^2 + \xi_{p+q+1} + \gamma$$

Veremos cómo son las superficies de nivel para

$$f(x) = \delta_1 \xi_1^2 \pm \delta_2 \xi_2^2 + \xi_3 \quad , \quad \text{con} \quad \delta_1, \delta_2 > 0$$

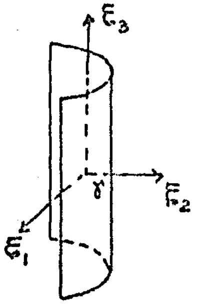
El caso cuando

$$f(x) = \delta_1 \xi_1^2 \pm \delta_2 \xi_2^2 \pm \delta_3 \xi_3^2$$

ya lo analizamos al estudiar las formas cuadráticas homogéneas.

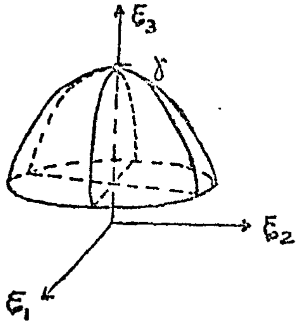
1) $f(x) = \delta_1 \xi_1^2 + \xi_2$

$\mathcal{R}(f, \gamma)$ es un cilindro parabólico



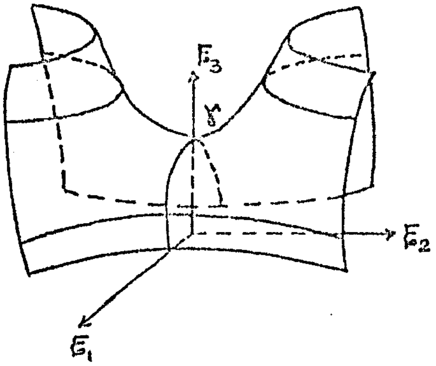
2) $f(x) = \delta_1 \xi_1^2 + \delta_2 \xi_2^2 + \xi_3$

$\mathcal{R}(f, \gamma)$ es un paraboloide elíptico.



3) $f(x) = \delta_1 E_1^2 - \delta_2 E_2^2 + E_3$

$\mathcal{H}(f, \gamma)$ es un paraboloide hiperbólico



Veremos que toda función cuadrática no-homogénea se puede llevar a la forma canónica, mediante un cambio de coordenadas y una traslación.

LEMA 1.

Sea $f(x) = x^t A x + b^t x + \gamma$, entonces existe un cambio de coordenadas tal que

$$f(x) \sim \bar{f}(\bar{x}) = \bar{x}^t \bar{A} \bar{x} + \bar{b}^t \bar{x} + \bar{\gamma} \quad , \quad \text{donde}$$

$$\bar{A} = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right]$$

Demostración:

Sea S no-singular tal que $\bar{A} = S^{-t} A S^{-1}$.
Sea $\bar{x} = Sx$, entonces

$$f(x) = (S^{-1}\bar{x})^t A (S^{-1}\bar{x}) + b^t (S^{-1}\bar{x}) + \gamma$$

$$= \bar{x}^t (S^{-t} A S^{-1}) \bar{x} + (b^t S^{-1}) \bar{x} + \gamma$$

$$= \bar{x}^t \bar{A} \bar{x} + (S^{-t} b)^t \bar{x} + \bar{\gamma} \quad , \quad \bar{\gamma} = \gamma$$

$$= \bar{x}^t \bar{A} \bar{x} + \bar{b}^t \bar{x} + \bar{\gamma} \quad , \quad \text{donde} \quad \bar{b} = S^{-t} b$$

∎

LEMA 2.

Supongamos $\bar{f}(\bar{x}) = \bar{x}^t \bar{A} \bar{x} + \bar{b}^t \bar{x} + \bar{\gamma}$ donde

$$\bar{A} = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right]$$

Entonces podemos llevar a f a una de las formas canónicas.

Demostración:

Sea $$\bar{f}(\bar{x}) = \bar{\xi}_1^2 + \bar{\xi}_2^2 + \cdots + \bar{\xi}_p^2 - \bar{\xi}_{p+1}^2 - \cdots - \bar{\xi}_{p+q}^2 + $$
$$+ \bar{\beta}_1 \bar{\xi}_1 + \bar{\beta}_2 \bar{\xi}_2 + \cdots + \bar{\beta}_n \bar{\xi}_n + \bar{\gamma}.$$

caso I: $\bar{f}(\bar{x}) = \bar{x}^t \bar{A} \bar{x} + \bar{b}^t \bar{x} + \bar{\gamma}$

si $\bar{\beta}_{p+q+1} = \bar{\beta}_{p+q+2} = \cdots = \bar{\beta}_n = 0$

entonces podemos usar la técnica de completar cuadrados:

$$\bar{f}(\bar{x}) = \left[\bar{\xi}_1^2 + \bar{\beta}_1 \bar{\xi}_1 + \left(\frac{\bar{\beta}_1}{2}\right)^2\right] + \cdots + \left[\bar{\xi}_p^2 + \bar{\beta}_2 \bar{\xi}_2 + \left(\frac{\bar{\beta}_p}{2}\right)^2\right] - \sum_{i=1}^{p}\left(\frac{\bar{\beta}_i}{2}\right)^2$$
$$- \left[\bar{\xi}_{p+1}^2 - \bar{\beta}_{p+1}\bar{\xi}_{p+1} + \left(\frac{\bar{\beta}_{p+1}}{2}\right)^2\right] - \cdots - \left[\bar{\xi}_{p+q}^2 - \bar{\beta}_{p+q}\bar{\xi}_{p+q} + \left(\frac{\bar{\beta}_{p+q}}{2}\right)^2\right] + \sum_{i=p+1}^{p+q}\left(\frac{\bar{\beta}_i}{2}\right)^2$$
$$+ \gamma$$

$$= \left(\bar{\xi}_1 + \frac{\bar{\beta}_1}{2}\right)^2 + \cdots + \left(\bar{\xi}_p + \frac{\bar{\beta}_p}{2}\right)^2 - \left(\bar{\xi}_{p+1} - \frac{\bar{\beta}_{p+1}}{2}\right)^2 - \cdots$$

$$- \left(\bar{\xi}_{p+q} - \frac{\bar{\beta}_{p+q}}{2}\right)^2 - \sum_{i=1}^{p}\left(\frac{\bar{\beta}_i}{2}\right)^2 + \sum_{i=p+1}^{p+q}\left(\frac{\bar{\beta}_i}{2}\right)^2 + \bar{\gamma}$$

sea $\tilde{x} = \bar{x} + x_0$ donde

$$x_0 = \frac{1}{2}\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \\ -\beta_{p+1} \\ \vdots \\ -\beta_{p+q} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad ; \quad \text{sea} \quad \tilde{\gamma} = \bar{\gamma} + \sum_{i=p+1}^{p+q}\left(\frac{\bar{\beta}_i}{2}\right)^2 - \sum_{i=1}^{p}\left(\frac{\bar{\beta}_i}{2}\right)^2 \ ;$$

$$\therefore \quad \bar{f}(\bar{x}) = \tilde{\xi}_1^2 + \cdots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \cdots - \tilde{\xi}_{p+q}^2 + \tilde{\gamma} = \tilde{f}(\tilde{x})$$

$$= \tilde{f}(\bar{x} + x_0)$$

y por lo tanto las curvas de nivel de $\bar{f}$ y $\tilde{f}$ son las mismas, pero trasladadas.

caso II : $\bar{f}(\bar{x}) = \bar{x}^t \bar{A} \bar{x} + \bar{b}^t \bar{x} + \gamma$:

Supongamos $\bar{\beta}_k \neq 0$, con $k > p+q$.

Sea $\tilde{\xi}_i = \bar{\xi}_i$ si $i \neq p+q+1 , k$

$\tilde{\xi}_{p+q+1} = \bar{b}^t \bar{x}$

$\tilde{\xi}_k = \bar{\xi}_{p+q+1}$

Por lo tanto $\tilde{x} = S \bar{x}$, donde S es la matriz

$$S = \begin{array}{c|c|cccccccccc|c}
 & p+q & & & & & n-(p+q) & & & & & & \\
\hline
p+q & I_{p+q} & & & & & 0 & & & & & & \\
\hline
1 & \bar{\beta}_1 \cdots \bar{\beta}_{p+q} & \beta_{p+q+1} & & \cdots & & & \bar{\beta}_k & & \cdots & & \beta_n & \leftarrow p+q+1 \\
\hline
 & & 0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \\
 & & 0 & 0 & 1 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \\
 & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \\
 & & 0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0 & \\
 & 0 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & \leftarrow k \\
 & & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 & \\
n-(p+q+1) & & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 & \\
 & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots & \\
 & & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 & \\
\hline
 & & \uparrow & & & & & \uparrow & & & & & \\
 & & p+q+1 & & & & & k & & & & &
\end{array}$$

Como $\tilde\beta_k \neq 0 \Rightarrow S$ es no-singular y por lo tanto $\tilde x = S\bar x$ representa un cambio de coordenadas en $\mathbb{R}^n$

$\therefore$

$$\bar f(\bar x) = \bar x^t \bar A \bar x + \bar b^t \bar x + \bar\gamma = \tilde x^t \tilde A \tilde x + \tilde\xi_{p+q+1} + \tilde\gamma = \tilde f(S\bar x),$$

donde $\tilde A = \bar A$ y $\tilde\gamma = \bar\gamma$ □

RESULTADO

Sea $f(x) = x^t A x + b^t x + \gamma$. Entonces f se puede llevar a una de las formas canónicas.

Demostración:

Por el lema 1

$$f(x) \sim \bar f(\bar x) = \bar x^t \bar A \bar x + \bar b^t \bar x + \bar\gamma, \quad \text{donde}$$

$$\bar A = \left[\begin{array}{c|c|c} I_p & & \\ \hline & -I_q & \\ \hline & & 0 \end{array}\right]$$

Por el lema 2

$$\bar f(\bar x) = \tilde f(\bar x + x_0) = \tilde f(\tilde x) = \tilde x^t \tilde A \tilde x + \tilde\gamma, \text{ con } \tilde A = \bar A$$

ó $$\bar f(\bar x) = \tilde f(S\bar x) = \tilde f(\tilde x) = \tilde x^t \tilde A \tilde x + \tilde\xi_{p+q+1} + \tilde\gamma, \text{ con } \tilde A = \bar A$$

□

Analizaremos las funciones cuadráticas canónicas :

Caso I : $f(x) = x^t A x + b^t x + \gamma$, con $A > 0$

RESULTADO

Sea $f(x) = x^t A x + b^t x + \gamma$, con $A > 0$. Entonces $f(x)$ tiene mínimo.

Demostración:

Si $A > 0 \Rightarrow A = U^t U$

$$\therefore\ f(x) = x^t (U^t U) x + b^t (U^{-1} U) x + \gamma$$

$$= (x^t U^t)(Ux) + (b^t U^{-1})(Ux) + \gamma$$

$$= (Ux)^t (Ux) + (U^{-t} b)^t (Ux) + \gamma$$

$$= \left(Ux + \frac{U^{-t}b}{2}\right)^t \left(Ux + \frac{U^{-t}b}{2}\right) + \gamma - \left(\frac{U^{-t}b}{2}\right)^t \left(\frac{U^{-t}b}{2}\right)$$

$$= \left[U\left(x + \frac{U^{-1}U^{-t}b}{2}\right)\right]^t \left[U\left(x + \frac{U^{-1}U^{-t}b}{2}\right)\right] + \gamma - \frac{b^t U^{-1} U^{-t} b}{4}$$

$$= \left[\left(x + \frac{U^{-1}U^{-t}b}{2}\right)^t U^t\right] \left[U\left(x + \frac{U^{-1}U^{-t}b}{2}\right)\right] + \gamma - \frac{b^t U^{-1} U^{-t} b}{4}$$

$$\therefore f(x) = \left[x + \frac{U^{-1}U^{-t}b}{2}\right]^{t}(U^{t}U)\left[x + \frac{U^{-1}U^{-t}b}{2}\right] + \gamma - \frac{b^{t}U^{-1}U^{-t}b}{4}$$

Pero teníamos que $A = U^{t}U$ y por lo tanto $A^{-1} = U^{-1}U^{-t}$

$$\Rightarrow \quad f(x) = \left(x + \frac{A^{-1}b}{2}\right)^{t} A \left(x + \frac{A^{-1}b}{2}\right) + \gamma - \frac{b^{t}A^{-1}b}{4}$$

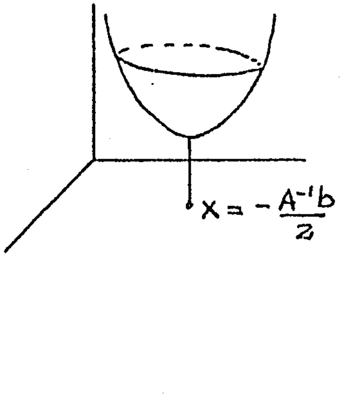
Como A es positiva definida, $f(x)$ tiene mínimo cuando

$$x = -\frac{A^{-1}b}{2}$$

Por ejemplo, para $n=2$ la gráfica de f

$$\mathcal{G}(f) = \left\{ \left[\begin{matrix} x \\ E_3 \end{matrix}\right]_1^2 \in \mathbb{R}^3 \;:\; f(x) = E_3 \right\}$$

tiene la forma:



Caso II : $f(x) = x^t A x + b^t x + \gamma$, con $A < 0$.

RESULTADO

Sea $f(x) = x^t A x + b^t x + \gamma$, con $A < 0$. Entonces $f(x)$ tiene máximo.

Demostración :

De manera análoga a como se procedió en el resultado anterior, se puede ver que

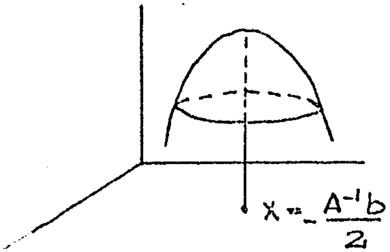
$$f(x) = \left(x + \frac{A^{-1}b}{2}\right)^t A \left(x + \frac{A^{-1}b}{2}\right) + \gamma - \frac{b^t A^{-1} b}{4}$$

y por lo tanto, como $A < 0$ $f(x)$ tiene máximo cuando

$$x = -\frac{A^{-1}b}{2}$$

□

Por ejemplo, para $n = 2$ la gráfica de f tiene la forma :



caso III : $f(x) = x^t A x + b^t x + \gamma$ con A indefinida y $r(A) = n$

RESULTADO

Sea $f(x) = x^t A x + b^t x + \gamma$, con A indefinida , $r(A) = n$. Entonces $f(x)$ no tiene ni máximo ni mínimo.

Demostración :

Como $r(A) = n \Rightarrow$

$f(x) \sim \tilde{f}(\tilde{x})$, donde $\tilde{f}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x} + \tilde{\gamma}$

Esto nos sugiere que quizás podamos escribir a $f(x)$ como

$$f(x) = (x - x_0)^t A (x - x_0) + \tilde{\gamma} \quad :$$

Supongamos $f(x) = (x - x_0)^t A \ (x - x_0) + \tilde{\gamma}$

$$\Rightarrow \quad f(x) = (x^t - x_0^t) A (x - x_0) + \tilde{\gamma}$$

$$= x^t A x - 2 x_0^t A x + x_0^t A x_0 + \tilde{\gamma}$$

$$\Rightarrow \quad \begin{aligned} b^t &= -2 x_0^t A \\ \gamma &= \tilde{\gamma} + x_0^t A x_0 \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} b &= -2 A^t x_0 = -2 A x_0 \qquad (A = A^t) \\ \gamma &= \tilde{\gamma} + x_0^t A x_0 \end{aligned}$$

Como $r(A)=n \Rightarrow A$ es invertible y por lo tanto

$$x_0 = -\frac{A^{-1}b}{2} \quad \text{y}$$

$$\tilde{\gamma} = \gamma - \left(-\frac{A^{-1}b}{2}\right)^t A \left(-\frac{A^{-1}b}{2}\right) = \gamma - \left(\frac{b^t A^{-t}}{2}\right)\left(\frac{AA^{-1}b}{2}\right)$$

$$= \gamma - \frac{b^t A^{-1} b}{2\cdot 2} \qquad (A = A^t \Rightarrow (A^{-1})^t = A^{-1})$$

$$\therefore \quad f(x) = \left(x + \frac{A^{-1}b}{2}\right)^t A \left(x + \frac{A^{-1}b}{2}\right) + \gamma - \frac{b^t A^{-1} b}{2^2}$$

Como A es indefinida, entonces $A \sim \tilde{A}$,

donde $\tilde{A} = \left[\begin{array}{c|c} I_p & 0 \\ \hline 0 & -I_q \end{array}\right]$ $\qquad p+q=n$.

y $A = S^t \tilde{A} S$. Si $\tilde{x} = S\left(x + \frac{A^{-1}b}{2}\right)$ entonces

$$f(x) = \left(x + \frac{A^{-1}b}{2}\right)^t S^t \tilde{A} S \left(x + \frac{A^{-1}b}{2}\right) + \gamma - \frac{b^t A^{-1} b}{4}$$

$$= \left[S\left(x + \frac{A^{-1}b}{2}\right)\right] \tilde{A} \left[S\left(x + \frac{A^{-1}b}{2}\right)\right] + \gamma - \frac{b^t A^{-1} b}{4}$$

$$\therefore \quad f(x) = \tilde{f}(\tilde{x}) = \tilde{x}^t \tilde{A} \tilde{x} + \tilde{\gamma}$$

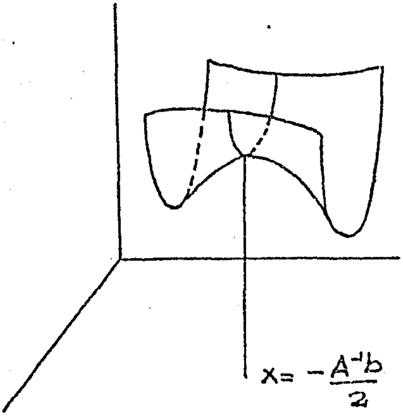
$$= \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \cdots + \tilde{\xi}_p^2 - \tilde{\xi}_{p+1}^2 - \cdots - \tilde{\xi}_n^2 + \tilde{\gamma}$$

de donde se puede concluir que los valores de la función pueden ser arbitrariamente grandes, tanto positivamente como negativamente.

El punto $-\frac{A^{-1}b}{2}$ se llama "punto silla".

□

Por ejemplo, para $n=2$, la gráfica de f tiene la forma:



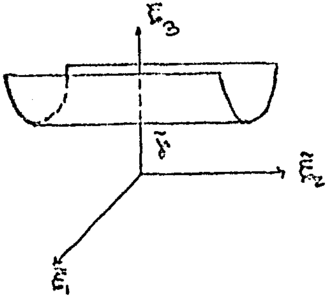
Caso IV :

$f(x) = x^t A x + b^t x + \gamma$, con A semi-definida positiva.

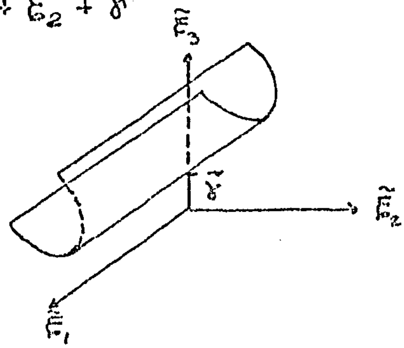
Entonces puede suceder que $f(x)$ tenga mínimo - y que el conjunto de puntos donde f alcanza su mínimo sea de cardinalidad infinita - ó puede pasar que f no tenga mínimo.

Para $n=2$ las formas canónicas, y sus respectivas gráficas son:

(a) $\tilde{f}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\gamma}$



(b) $\tilde{f}(\tilde{x}) = \tilde{\xi}_1^2 + \tilde{\xi}_2 + \tilde{\gamma}$



Caso V :

$f(x) = x^t A x + b^t x + \gamma$, con A semi-definida negativa .

Es análogo al caso IV .

CAPITULO VI

ESPACIOS CON PRODUCTO INTERIOR

ESPACIOS CON PRODUCTO INTERIOR

INTRODUCCION

En el primer capítulo definimos la norma euclideana de $x \in \mathbb{R}^n$ como

$$\|x\| = \sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2} = \sqrt{x^t x}$$

y vimos que el producto $x^t y$ nos servía para definir el ángulo formado por los vectores x y y.

Supongamos que $S: \mathbb{R}^n \to \mathbb{R}^n$ es una matriz no-singular que nos da un cambio de coordenadas en $\mathbb{R}^n$ de x a $\tilde{x}$; es decir

$$\tilde{x} = Sx$$

entonces

$$\|\tilde{x}\|^2 = \tilde{x}^t \tilde{x} = (Sx)^t (Sx) = x^t (S^t S) x$$

Observemos que, en general, $\|\tilde{x}\| \neq \|x\|$, pero sin embargo

$$\|\tilde{x}\|^2 = \tilde{\xi}_1^2 + \tilde{\xi}_2^2 + \cdots + \tilde{\xi}_n^2 .$$

es decir, la norma euclideana de $\mathbb{R}^n$ en el sistema de coordenadas S "no se ve" como la norma euclideana desde el sistema I_n

Es muy conveniente poder manejar la norma de $\tilde{x}$ en términos de x. Así podemos pensar que

$$\|\tilde{x}\|^2 = \|x\|_M^2 = x^t M x \quad , \text{ donde } M = S^t S > 0.$$

Angulos :

En $\mathbb{R}^n$ tenemos que el ángulo entre x y y está dado por:

$$\theta = \text{arc cos}\left(\frac{x^t y}{\|x\| \, \|y\|}\right)$$

Fue posible introducir este concepto gracias a la desigualdad de Schwartz :

$$|x^t y| \leq \|x\| \, \|y\|$$

Veremos que la noción de ángulo se puede extender a espacios vectoriales en los cuales se tenga un producto escalar similar a $x^t y$.

Por ejemplo, en $\mathbb{R}^n$, donde la norma está dada por

$$\|x\|_M^2 = x^t M x \quad , \text{ con } M > 0 .$$

DEFINICION

Una NORMA en un espacio vectorial real $\mathbb{V}$ es una función

$$\| \cdot \| : \mathbb{V} \to \mathbb{R}$$

con las siguientes propiedades:

1) $\|v\| \geq 0 \quad \forall\, v \in \mathbb{V} \quad ; \quad \|v\| = 0 \Leftrightarrow v = 0$

2) $\|\lambda v\| = |\lambda| \cdot \|v\| \qquad \lambda \in \mathbb{R}, \ v \in \mathbb{V}$

3) $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\| \qquad \forall\, v_1, v_2 \in \mathbb{V}$.

EJEMPLOS DE NORMAS EN $\mathbb{R}^n$

i) $\|x\|_2 = \sqrt{\xi_1^2 + \xi_2^2 + \cdots + \xi_n^2}$

Demostración:

(1) Claramente se cumple.

(2) $\|\lambda x\|_2^2 = \sum_{i=1}^{n} (\lambda \xi_i)^2 = \lambda^2 \sum_{i=1}^{n} \xi_i^2 = \lambda^2 \|x\|_2^2$

$\therefore \|\lambda x\|_2 = |\lambda| \cdot \|x\|_2$

(3) $\|x+y\|_2^2 = \sum_{i=1}^{n} (\xi_i + \eta_i)^2$

$$= \sum_{i=1}^{n} \xi_i^2 + 2 \sum_{i=1}^{n} \xi_i \eta_i + \sum_{i=1}^{n} \eta_i^2$$

$$\leq \|x\|_2^2 + 2 \left| \sum_{i=1}^{n} \xi_i \eta_i \right| + \|y\|_2^2$$

$$= \|x\|_2^2 + 2 |x^t y| + \|y\|_2^2$$

$$\leq \|x\|_2^2 + 2 \|x\|_2 \|y\|_2 + \|y\|_2^2$$

$$= (\|x\|_2 + \|y\|_2)^2$$

$\therefore \|x+y\|_2 \leq \|x\|_2 + \|y\|_2$

$\therefore$ $\|x\|_2$ es una norma en $\mathbb{R}^n$

■

2) $\|x\|_1 = |\xi_1| + |\xi_2| + \cdots + |\xi_n|$

Demostración

(1) $\sum_{i=1}^{n} |\xi_i| \geq 0 \quad \forall\, x \in \mathbb{R}^n \quad ; \quad \sum_{i=1}^{n} |\xi_i| = 0 \iff x = 0$

(2) $\sum_{i=1}^{n} |\lambda \xi_i| = \sum_{i=1}^{n} (|\lambda|\,|\xi_i|) = |\lambda| \cdot \sum_{i=1}^{n} |\xi_i| = |\lambda|\ \|x\|_1$

(3) $$\|x+y\|_1 = \sum_{i=1}^{n} |\xi_i + \eta_i| \leq \sum_{i=1}^{n} (|\xi_i| + |\eta_i|)$$

$$= \sum_{i=1}^{n} |\xi_i| + \sum_{i=1}^{n} |\eta_i| = \|x\|_1 + \|y\|_1$$

$\therefore \quad \|x\|_1$ es una norma en $\mathbb{R}^n$ ▫

3) $\|x\|_\infty = \max \{ |\xi_1|, |\xi_2|, \ldots, |\xi_n| \}$

(1) Claramente $\|x\|_\infty \geq 0 \quad \forall\, x \in \mathbb{R}^n$

$\|x\|_\infty = 0 \iff x = 0$

(2) $$\max \{ |\lambda \xi_1|, |\lambda \xi_2|, \ldots, |\lambda \xi_n| \} =$$

$$= \max \{ |\lambda|\,|\xi_1|, |\lambda|\,|\xi_2|, \ldots, |\lambda|\,|\xi_n| \}$$

$$= |\lambda| \max \{ |\xi_1|, |\xi_2|, \ldots, |\xi_n| \}$$

$$= |\lambda|\ \|x\|_\infty .$$

$$= \|\lambda x\|_\infty .$$

$$(3) \quad \|x+y\|_\infty = \max\{|\xi_1+\eta_1|, |\xi_2+\eta_2|, \dots, |\xi_n+\eta_n|\}$$

$$\leq \max\{|\xi_1|+|\eta_1|, |\xi_2|+|\eta_2|, \dots |\xi_n|+|\eta_n|\}$$

$$= \max\{|\xi_1|, |\xi_2|, \dots, |\xi_n|\} +$$

$$+ \max\{|\eta_1|, |\eta_2|, \dots, |\eta_n|\}$$

$$= \|x\|_\infty + \|y\|_\infty$$

$\therefore$ $\|x\|_\infty$ es una norma en $\mathbb{R}^n$ ◻

4) $\|x\|_M = \sqrt{x^t M x}$, donde $M > 0$.

Después veremos que $\|x\|_M$ es una norma en $\mathbb{R}^n$

Hay muchas funciones que tienen las mismas propiedades que el producto escalar $x^t y$:

Si $F(x,y) = x^t y$; $F: \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,

entonces F tiene las siguientes propiedades :

i) $F(x,y) = F(y,x)$ $\forall\, x,y \in \mathbb{R}^n$

ii) $F(\lambda x, y) = \lambda F(x,y)$ $\forall\, x,y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

iii) $F(x_1+x_2, y) = F(x_1,y) + F(x_2,y)$ $\forall\, x_1,x_2, y \in \mathbb{R}$

iv) $F(x,x) > 0$ si $x \neq 0$, $F(0,0) = 0$.

Por ejemplo

$$F(X,Y) = X^t M Y \quad \text{con} \quad M>0$$

ESPACIOS EUCLIDEANOS

Empezaremos por definir en un espacio vectorial $\mathbb{V}$ una función similar al producto escalar en $\mathbb{R}^n$, que llamaremos producto interior en $\mathbb{V}$.

DEFINICION

Un PRODUCTO INTERIOR en $\mathbb{V}$ es una función

$$F : \mathbb{V} \times \mathbb{V} \to \mathbb{R}$$

con las siguientes propiedades:

1) F es bilineal
2) F es simétrica
3) $F(V,V)$ es positiva definida

Escribiremos $F(V,W) = \langle V,W \rangle$, si F es un producto interior en $\mathbb{V}$.

RESULTADO

Sea V un espacio vectorial real de dimensión n, y $F : V \times V \rightarrow \mathbb{R}$ un producto interior en V.

Sea $\{v_1, v_2, \ldots, v_n\}$ una base de V.

Si $v = \sum_{i=1}^{n} \xi_i v_i$ y $w = \sum_{i=1}^{n} \eta_i v_i$ son dos vectores en V entonces

$$F(v,w) = x^t M y \qquad \text{donde}$$

$x = [\xi_i] \in \mathbb{R}^n$ es el vector de coordenadas de v en la base dada

$y = [\eta_i] \in \mathbb{R}^n$ es el vector de coordenadas de w en la base dada

$M = [\mu_{ij}] \in M_n$ está dada por

$$\mu_{ij} = F(v_i, v_j) = \langle v_i, v_j \rangle ; \; 1 \leq i,j \leq n .$$

DEFINICION :

Un ESPACIO EUCLIDEANO es un espacio vectorial en el que está definido un producto interior.
Escribiremos $(V, \langle \cdot \rangle)$

PROPIEDADES DEL PRODUCTO INTERIOR

1) $\langle v-w, v-w\rangle = \langle v,v\rangle + \langle w,w\rangle - 2\langle v,w\rangle$

Demostración:

$$\begin{aligned}\langle v-w, v-w\rangle &= \langle v, v-w\rangle + \langle -w, v-w\rangle\\ &= \langle v,v\rangle + \langle v,-w\rangle + \langle -w,v\rangle + \langle -w,-w\rangle\\ &= \langle v,v\rangle - \langle v,w\rangle - \langle w,v\rangle + \langle w,w\rangle\\ &= \langle v,v\rangle - 2\langle v,w\rangle + \langle w,w\rangle \quad \blacksquare\end{aligned}$$

2) Si $\langle v,w\rangle = 0$ entonces

$$\langle v-w, v-w\rangle = \langle v,v\rangle + \langle w,w\rangle$$

Demostración.

Es consecuencia inmediata del resultado anterior $\blacksquare$

DEFINICION

Sean $v,w \in (V, \langle\cdot\rangle)$. Diremos que v es ORTOGONAL a w si $\langle v,w\rangle = 0$. En tal caso escribiremos

$$v \perp w \ .$$

RESULTADO

Sean $v, \omega \in (\mathbb{V}, <,>)$, $\omega, v \neq 0$. Entonces

$$(\omega - \alpha v) \perp v \iff \alpha = \frac{\langle v, w\rangle}{\langle v, v\rangle}$$

Demostración :

$$\begin{aligned} \omega - \alpha v \perp v &\iff \langle v, \omega - \alpha v\rangle = 0 \\ &\iff \langle v, \omega\rangle - \alpha \langle v, v\rangle = 0 \\ &\iff \alpha = \frac{\langle v, w\rangle}{\langle v, v\rangle} \end{aligned}$$

▪

Por lo tanto , dados dos vectores independientes , siempre podemos encontrar un vector que sea ortogonal a uno de ellos .

DESIGUALDAD DE SCHWARTZ

Sea $(\mathbb{V}, <, >)$ un espacio euclideano , y sean $v, \omega \in \mathbb{V}$. Entonces

$$\langle v, \omega\rangle^2 \leq \langle v, v\rangle \langle \omega, \omega\rangle$$

Demostración

Sea $z = \alpha v$.

Por lo tanto

$$\langle \omega,\omega \rangle = \langle (\omega-z)+z , (\omega-z)+z \rangle$$

Y si $\omega - z \perp z$ entonces

$$\langle \omega,\omega \rangle = \langle \omega-z,\omega-z \rangle + \langle z,z \rangle$$

$$\therefore \quad \langle z,z \rangle \leq \langle \omega,\omega \rangle$$

Por el resultado anterior $\omega - z \perp z \Leftrightarrow$

$z = \alpha v$ con $\alpha = \dfrac{\langle v,\omega \rangle}{\langle v,v \rangle}$, por lo tanto

$$\therefore \langle z,z \rangle = \langle \alpha v, \alpha v \rangle = \alpha^2 \langle v,v \rangle = \frac{\langle v,\omega \rangle^2}{\langle v,v \rangle^2} \langle v,v \rangle$$

$$\therefore \langle z,z \rangle = \frac{\langle v,\omega \rangle^2}{\langle v,v \rangle} \leq \langle w,w \rangle$$

$$\therefore \quad \langle v,\omega \rangle^2 \leq \langle v,v \rangle \langle w,w \rangle \quad \blacksquare$$

DEFINICION

Sea $(V, \langle , \rangle)$ un espacio euclideano. Definimos la siguiente función $\| \cdot \| : V \rightarrow \mathbb{R}$, tal que

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

RESULTADO

$\| \cdot \| : V \to \mathbb{R}$; $\|v\| = \sqrt{\langle v,v \rangle}$ es una norma en V.

Demostración

(1) $\langle v,v \rangle > 0$ si $v \neq 0$; $\langle v,v \rangle = 0 \Leftrightarrow v = 0$

$\therefore$ $\|v\| > 0$ si $v \neq 0$ y $\|v\| = 0 \Leftrightarrow v = 0$.

(2) $\| \alpha v \| = \sqrt{\langle \alpha v, \alpha v \rangle} = \sqrt{\alpha^2 \langle v,v \rangle} = |\alpha| \sqrt{\langle v,v \rangle}$

$= |\alpha| \, \|v\|$.

(3)
$$\begin{aligned} \|v+w\|^2 &= \langle v+w, v+w \rangle \\ &= \langle v,v \rangle + 2 \langle v,w \rangle + \langle w,w \rangle \\ &\leq \langle v,v \rangle + 2 \sqrt{\langle v,v \rangle} \sqrt{\langle w,w \rangle} + \langle w,w \rangle \\ &= \|v\|^2 + 2 \|v\| \, \|w\| + \|w\|^2 \\ &= (\|v\| + \|w\|)^2 \end{aligned}$$

$\therefore$ $\|v+w\| \leq \|v\| + \|w\|$

■

COROLARIO

Sea $\mathbb{R}^n$ con producto interior $\langle x,y \rangle = x^t M y$, con $M > 0$. Entonces

$\|x\|_M = \sqrt{x^t M x}$ es una norma en $\mathbb{R}^n$

■

RESULTADO

Sean $v,w \in (V,<.>)$, $v,w \neq 0$. Si $v \perp w$ entonces $\{v,w\}$ es linealmente independiente.

Demostración:

Suponqamos $v \perp w$, y sea

$$\alpha v + \beta w = 0 .$$

$\Rightarrow$ $\langle v, \alpha v + \beta w \rangle = 0$
$\Rightarrow$ $\alpha \langle v,v \rangle + \beta \langle v,w \rangle = 0$
$\Rightarrow$ $\alpha \langle v,v \rangle = 0$, ya que $\langle v,w \rangle = 0$
$\Rightarrow$ $\alpha = 0$, ya que $v \neq 0$

Por otro lado, de manera análoga tenemos:

$$\langle w, \alpha v + \beta w \rangle = 0 \Rightarrow \beta = 0$$

$\therefore$ $\alpha v + \beta w = 0 \Rightarrow \alpha = \beta = 0$

$\therefore$ $\{v,w\}$ linealmente independiente. ∎

DEFINICION

Sea $(V,<,>)$ un espacio euclideano, y sea $\{v_1, v_2, \dots, v_k\} \subseteq V$ tal que
$\langle v_i, v_j \rangle = 0$ si $i \neq j$ $1 \leq i,j \leq k$
$\langle v_i, v_i \rangle \neq 0$ para $1 \leq i \leq k$

Entonces decimos que $\{v_1, v_2, \dots, v_k\}$ es un conjunto ORTOGONAL.

El resultado anterior se puede generalizar de la siguiente manera:

RESULTADO

Sea $(V, <, >)$ un espacio euclideano. Si $\{v_1, v_2, \dots, v_k\}$ es ortogonal entonces $\{v_1, v_2, \dots, v_k\}$ es linealmente independiente.

Demostración:

Supongamos

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k = 0$$

$$\Rightarrow \langle v_i, \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_k v_k \rangle = 0 \qquad 1 \leq i \leq k$$

$$\Rightarrow \alpha_1 \langle v_i, v_1 \rangle + \alpha_2 \langle v_i, v_2 \rangle + \cdots + \alpha_k \langle v_i, v_k \rangle = 0 \ ; \quad 1 \leq i \leq k$$

y como $\langle v_i, v_j \rangle = 0$ si $i \neq j$ entonces tenemos que

$$\alpha_i \langle v_i, v_i \rangle = 0 \qquad \forall i \ , \quad 1 \leq i \leq k$$

$$\Rightarrow \quad \alpha_i = 0 \quad , \text{ya que } \langle v_i, v_i \rangle \neq 0$$

$\Rightarrow \{v_1, v_2, \dots, v_k\}$ es independiente ▫

Veremos que el tener un conjunto de generadores ortogonal de un subespacio nos puede ser de gran utilidad, como se muestra a continuación:

Ejemplo

Sea $\mathbb{R}^n$ con producto interior $\langle x,y\rangle = x^t y$

Resolver el sistema $AX = b$, donde

$A = [a_1|a_2|\cdots|a_n]$ y $\{a_1, a_2, \ldots, a_n\}$ es ortogonal.

Por el resultado anterior, $\{a_1, a_2, \cdots, a_n\}$ es linealmente independiente y por lo tanto $R(A) = \mathbb{R}^n$.

Por lo tanto $Ax = b$ tiene solución única, y

$x = [\xi_i]$, donde $\xi_i = \dfrac{a_i^t b}{a_i^t a_i}$ es la solución del sistema.

Esto se puede ver si observamos que

$$A^t A = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{bmatrix} [a_1|a_2|\cdots|a_n]$$

$$= \begin{bmatrix} a_1{}^t a_1 & & \\ & a_2{}^t a_2 & \\ & & & a_n{}^t a_n \end{bmatrix} = D \,, \quad D \text{ diagonal.}$$

$$\therefore \quad A^t A x = A^t y$$

$$\therefore \quad Dx = A^t y \qquad y \therefore \qquad x = D^{-1} A^t y$$

$$\therefore \quad \xi_i = \frac{1}{a_1{}^t a_1} \left(a_1{}^t y\right)$$

Por lo tanto, si la matriz del sistema $Ax=b$ tiene columnas ortogonales, entonces éste es muy fácil de resolver.

Es natural que nos preguntemos:

Dado $(V, \langle , \rangle)$, ¿ siempre existen bases ortogonales de V?

Para $V = \mathbb{R}^n$ la respuesta a esta pregunta es afirmativa:

RESULTADO

Sea $\mathbb{R}^n$ con producto interno $\langle x,y\rangle = x^t M y$, $M>0$. Entonces siempre podemos construir una base ortogonal de $\mathbb{R}^n$.

Demostración :

Sea $x_1 \in \mathbb{R}^n$, $x_1 \neq 0$.

Sea $M^t x_1 = Mx_1 = r_1$, y sea

$$\mathcal{S}_1 = \{ y \in \mathbb{R}^n : x_1^t M y = 0 \} = \{ y \in \mathbb{R}^n : (Mx_1)^t y = 0 \}$$

$$\therefore \ \mathcal{S}_1 = \{ y \in \mathbb{R}^n : r_1^t y = 0 \ ; \ r_1 = Mx_1 \}$$

Como $M>0$ y $x_1 \neq 0 \Rightarrow Mx_1 \neq 0$ y por lo tanto $\dim (N(r_1^t)) = \dim(\mathcal{S}_1) = n-1$

Por lo tanto podemos tomar $x_2 \in \mathcal{S}_1$, $x_2 \neq 0$, tal que $\{x_1, x_2\}$ es ortogonal.

Sea $\mathcal{S}_2 = \{ y \in \mathbb{R}^n : r_2^t y = 0 \ ; \ r_2 = Mx_2 \}$

Por lo tanto $\dim(\mathcal{S}_2) = n-1$ (ésto se ve de igual manera que para $\mathcal{S}_1$).

Entonces

$$\mathcal{S}_1 \cap \mathcal{S}_2 = \left\{ y \in \mathbb{R}^n : \begin{bmatrix} r_1^t \\ r_2^t \end{bmatrix} y = 0 \right\}$$

Como $x_1 \perp x_2 \Rightarrow \{Mx_1, Mx_2\} = \{r_1, r_2\}$ es independiente y por lo tanto $\dim(\mathcal{S}_1 \cap \mathcal{S}_2) = n-2$.

Por lo tanto podemos tomar $x_3 \in [\mathcal{S}_1 \cap \mathcal{S}_2]$, $x_3 \neq 0$, y $\{x_1, x_2, x_3\}$ es ortogonal.

En general vamos a tener en el k-ésimo, $1 \leq k \leq n-1$, que si

$$\{x_1, x_2, \ldots, x_k\} \text{ es ortogonal}$$

y si $\mathcal{S}_i = \{y \in \mathbb{R}^n : (Mx_i)^t y = 0\}$

$$= \{y \in \mathbb{R}^n : r_i^t y = 0 ; \ Mx_i = r_i\}, \ 1 \leq i \leq k,$$

entonces $\dim(\mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_k) = n-k$, ya que

$\mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_k = N(A_k)$ donde A_k es la matriz $k \times n$ dada por

$$A_k = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_k^t \end{bmatrix}$$

Como $M > 0$ y $\{x_1, x_2, \ldots, x_k\}$ es ortogonal $\Rightarrow$ $\{Mx_1, Mx_2, \ldots, Mx_k\} = \{r_1, r_2, \ldots, r_k\}$ es independiente y por lo tanto $\dim(N(A_k)) = n-k$.

Y por lo tanto, para $k \leq n-1$ podemos tomar un vector $x_{k+1} \in \mathcal{S}_1 \cap \mathcal{S}_2 \cap \cdots \cap \mathcal{S}_k$, $x_{k+1} \neq 0$, y entonces

$$\{x_1, x_2, \ldots, x_k, x_{k+1}\}$$

será un conjunto ortogonal en $\mathbb{R}^n$.

Para $k = n-1$ obtenemos un conjunto ortogonal $\{x_1, x_2, \ldots, x_n\} \subseteq \mathbb{R}^n$.

Por lo tanto $\{x_1, x_2, \ldots, x_n\}$ es una base ortogonal de $\mathbb{R}^n$

$\square$

COROLARIO

Sea $(V, \langle\, ,\rangle)$ un espacio euclideano de dimensión n. Entonces podemos encontrar n vectores $\omega_1, \omega_2, \ldots, \omega_n \in V$ tal que $\{\omega_1, \omega_2, \ldots, \omega_n\}$ sea una base ortogonal de V.

Demostración :

Sea $C_I : V \to \mathbb{R}^n$ un sistema de coordenadas en V.

Por lo tanto $\langle v, \omega\rangle = x^t M y$, donde

$$M = [\mu_{ij}] \; ; \quad \text{con} \quad \mu_{ij} = \langle C_I^{-1}(e_i), C_I^{-1}(e_j)\rangle$$
$$= \langle v_i, v_j\rangle, \quad \text{donde}$$

$$v_i = C_I^{-1}(e_i) \quad , \quad 1 \le i \le n.$$

y donde $$v = \sum_{i=1}^{n} \xi_i v_i = \sum_{i=1}^{n} \xi_i \left[C_I^{-1}(e_i) \right]$$

$$w = \sum_{i=1}^{n} \eta_i v_i = \sum_{i=1}^{n} \eta_i \left[C_I^{-1}(e_i) \right]$$

Por lo tanto $\langle v,w \rangle = 0 \iff x^t M y = 0$, es decir,

$$v \perp w \iff x \perp y \quad .$$

Por lo tanto, si $\{x_1, x_2, \dots, x_n\}$ es una base ortogonal de $\mathbb{R}^n \Rightarrow$

$$\{ C_I^{-1}(x_1),\ C_I^{-1}(x_2), \dots ,\ C_I^{-1}(x_n) \}$$

es una base ortogonal de $\mathcal{V}$.

$\square$

Veremos que para cualquier subespacio $\mathcal{S} \subseteq \mathbb{R}^n$ siempre podemos encontrar una base ortogonal de $\mathcal{S}$:

RESULTADO

Sea $\mathbb{R}^n$ con producto interno $\langle x,y \rangle = x^t M y$, $M > 0$. Si $\mathcal{S} \subseteq \mathbb{R}^n$ es un subespacio de $\mathbb{R}^n$, entonces siempre podemos una base ortogonal de $\mathcal{S}$.

Demostración :

El procedimiento que utilizamos para construir la base ortogonal de $\mathcal{S}$ es esencialmente el mismo que utilizamos cuando $\mathcal{S} = \mathbb{R}^n$:

Sea $s_1 \in \mathcal{S}$, $s_1 \neq 0$, y sean

$$\mathcal{S}_1 = \{ y \in \mathbb{R}^n : (Ms_1)^t y = 0 \}$$

$$\tilde{\mathcal{S}}_1 = \{ y \in \mathcal{S} : (Ms_1)^t y = 0 \} = \mathcal{S}_1 \cap \mathcal{S} .$$

Entonces $\dim(\mathcal{S}_1 \cap \mathcal{S}) = \dim(\tilde{\mathcal{S}}_1) = k-1$, ya que : $\mathcal{S}_1 + \mathcal{S} = \mathbb{R}^n$ y por lo tanto

$$\dim(\mathcal{S}_1 + \mathcal{S}) + \dim(\mathcal{S}_1 \cap \mathcal{S}) = \dim(\mathcal{S}_1) + \dim(\mathcal{S})$$

$$\Rightarrow \quad n + \dim(\mathcal{S}_1 \cap \mathcal{S}) = (n-1) + k$$

$$\Rightarrow \quad \dim(\mathcal{S}_1 \cap \mathcal{S}) = k-1$$

Por lo tanto podemos tomar un vector $s_2 \in \tilde{\mathcal{S}}_1$, $s_2 \neq 0$ y por lo tanto $\{s_1, s_2\} \subseteq \mathcal{S}$ es ortogonal.

En general, para $1 \leq i \leq k-1$ vamos a tener que , si

$$\tilde{\mathcal{S}}_i = \{ y \in \mathcal{S} : (Ms_i)^t y = 0 \} \quad , \qquad \text{donde}$$

$$\{s_1, s_2, \ldots, s_i\} \subseteq \mathcal{S} \quad \text{es ortogonal}$$

entonces $\dim(\tilde{\mathcal{S}}_1 \cap \tilde{\mathcal{S}}_2 \cap \cdots \cap \tilde{\mathcal{S}}_i) = k-i$ y

por lo tanto podemos tomar

$$s_{i+1} \in \tilde{\mathcal{S}}_1 \cap \tilde{\mathcal{S}}_2 \cap \cdots \cap \tilde{\mathcal{S}}_i \quad , \quad s_{i+1} \neq 0$$

y por lo tanto

$$\{s_1, s_2, \ldots, s_i, s_{i+1}\} \quad \text{es ortogonal.}$$

Al cabo del paso $k-1$ obtenemos un conjunto ortogonal

$$\{s_1, s_2, \ldots, s_k\} \subseteq \mathcal{S}$$

y por lo tanto

$$\{s_1, s_2, \ldots, s_k\} \quad \text{es una base ortogonal de } \mathcal{S}.$$

∎

COROLARIO

Sea $(V, \langle , \rangle)$ un espacio euclideano. Si $W \subseteq V$ es un subespacio de V, entonces podemos construir una base ortogonal de W

∎

PROCEDIMIENTOS EFICIENTES PARA CONSTRUIR BASES ORTOGONALES

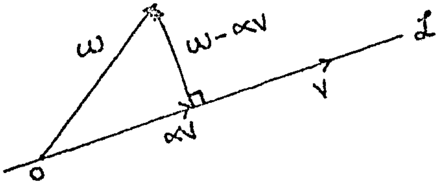
Hemos visto hasta ahora que el tener una base ortogonal de un subespacio $W \subseteq (V, < >)$ puede simplificar determinadas operaciones que tengamos que realizar con los vectores de W, y que siempre podemos encontrar una base ortogonal de cualquier subespacio de V. Trataremos de encontrar procedimientos eficientes para construir bases ortogonales de un subespacio dado.

INTRODUCCION

La distancia de un punto a una recta se puede ver como un método para construir un vector ortogonal a una recta.

Suponqamos $(V, <,>)$ es un espacio euclideano.

Sea $\mathcal{L} \subseteq V$ una recta que pasa por el origen : $\mathcal{L} = [v]$



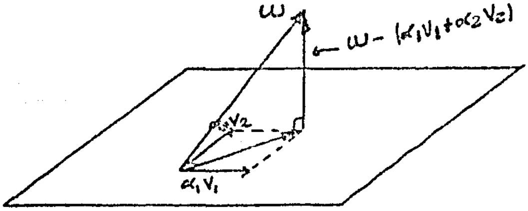
Sea $w \in \mathcal{V}$ un vector que no está en la recta $\mathcal{L}$, entonces, como vimos anteriormente,

$$w - \alpha v \perp v \iff \alpha = \frac{\langle v,w\rangle}{\langle v,v\rangle}$$

, y el vector $\left(\frac{\langle v,w\rangle}{\langle v,v\rangle}\right)v$ es el punto más cercano de $\mathcal{L}$ al punto w.

Por lo tanto, contamos con un procedimiento para construir un vector ortogonal a otro.

¿ Cómo calculamos un vector ortogonal a dos vectores ortogonales dados ?



Es decir, sean $v_1, v_2 \in \mathcal{V}$ tal que $\{v_1, v_2\}$ es ortogonal, y sea $w \in \mathcal{V}$ tal que $w \notin [v_1, v_2]$.

Buscaremos un vector de la forma

$$v_3 = w - (\alpha_1 v_1 + \alpha_2 v_2)$$

, tal que

$V_1 \perp V_3$ y $V_2 \perp V_3$

Por lo tanto necesitamos que

$$\langle V_1, \omega - \alpha_1 V_1 - \alpha_2 V_2 \rangle = 0$$
$$\langle V_2, \omega - \alpha_1 V_1 - \alpha_2 V_2 \rangle = 0$$

$$\therefore \quad \langle V_1, \omega \rangle - \alpha_1 \langle V_1, V_1 \rangle - \alpha_2 \langle V_1, V_2 \rangle = 0$$
$$\langle V_2, \omega \rangle - \alpha_1 \langle V_1, V_2 \rangle - \alpha_2 \langle V_2, V_2 \rangle = 0$$

$$\Rightarrow \quad \langle V_1, \omega \rangle - \alpha_1 \langle V_1, V_1 \rangle = 0$$
$$\langle V_2, \omega \rangle - \alpha_2 \langle V_2, V_2 \rangle = 0 \qquad (\langle V_1, V_2 \rangle = 0)$$

$$\Rightarrow \quad \alpha_1 = \frac{\langle V_1, \omega \rangle}{\langle V_1, V_1 \rangle} \quad , \quad \alpha_2 = \frac{\langle V_2, \omega \rangle}{\langle V_2, V_2 \rangle}$$

Por lo tanto, si $V_1 \perp V_2$ y $\omega \notin [V_1, V_2]$, entonces

$$V_3 = \omega - \frac{\langle V_1, \omega \rangle}{\langle V_1, V_1 \rangle} V_1 - \frac{\langle V_2, \omega \rangle}{\langle V_2, V_2 \rangle} V_2$$

es ortogonal a V_1 y a V_2.

Observemos que:

$\frac{\langle V_1, \omega \rangle}{\langle V_1, V_1 \rangle}$ es el punto más cercano sobre $[V_1]$ a ω, y

que $\frac{\langle V_2, W \rangle}{\langle V_2, V_2 \rangle}$ es el punto más cercano sobre $[V_2]$ a ω.

Ejemplo

Sea $(V, \langle , \rangle)$ un espacio euclideano. Dado $\{w_1, w_2, w_3\}$ linealmente independiente, construir $\{v_1, v_2, v_3\}$ ortogonal.

Este problema prácticamente ya está resuelto, si procedemos a construir v_1, v_2 y v_3 siguiendo la misma idea que en los ejemplos anteriores:

Sean $v_1 = w_1$

$$v_2 = w_2 - \frac{\langle v_1, w_2 \rangle}{\langle v_1, v_1 \rangle} v_1$$

$$v_3 = w_3 - \frac{\langle v_1, w_3 \rangle}{\langle v_1, v_1 \rangle} v_1 - \frac{\langle v_2, w_3 \rangle}{\langle v_2, v_2 \rangle} v_2$$

Entonces $\{v_1, v_2, v_3\}$ es ortogonal.

Por lo tanto, a partir de 3 vectores independientes siempre podemos construir 3 vectores ortogonales.

Más adelante veremos que este método se puede extender a conjuntos independientes de n vectores, y que por lo tanto nos proporciona un procedimiento para construir bases ortogonales.

El problema de construir bases ortogonales de subespacios está ligado al siguiente

PROBLEMA

Sea IR^n con producto interior $\langle x,z\rangle = x^t M z$, $M>0$, y sea $\mathcal{S} \subseteq IR^n$ un subespacio de IR^n.

Dado $y \in IR^n$, encontrar $\tilde{y} \in \mathcal{S}$ tal que

$$\| y - \tilde{y} \|^2 = \min_{s \in \mathcal{S}} \| y - s \|^2$$

Es decir, dado un subespacio $\mathcal{S} \subseteq IR^n$ y un punto $y \in IR^n$, queremos saber si existe un punto $\tilde{y}$ de $\mathcal{S}$ que sea el más cercano a y, y en caso afirmativo, desarrollar la técnica para encontrarlo.

Este problema constituye la médula de este capítulo.

Veremos que siempre existe un punto $\tilde{y} \in \mathcal{S}$ tal que la distancia de y a $\tilde{y}$ es mínima, y que este punto es único.

RESULTADO

Sea $\mathbb{R}^n$ con producto interior $\langle x,z\rangle = x^t M z$, $M>0$, y sea $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$. Dado $y \in \mathbb{R}^n$ siempre existe un único vector $\hat{y} \in \mathcal{S}$ tal que

$$\| y-\hat{y} \|^2 = \min_{s\in\mathcal{S}} \| y-s \|^2$$

Demostración:

Sea $f: \mathcal{S} \to \mathbb{R}$ definida por

$$f(s) = \| y-s \|^2 \qquad \text{si} \quad s \in \mathcal{S}.$$

es decir, si $s\in\mathcal{S}$ entonces
$f(s) =$ (distancia de y a s)^2.

Veremos que f tiene mínimo:

Sea $\{s_1, s_2, \dots, s_k\}$ una base de $\mathcal{S}$, $k\leq n$, y sea

$$S = [s_1 | s_2 | \cdots | s_k]$$

Por lo tanto, si $s\in\mathcal{S} \Rightarrow s = Sx$ para alguna $x\in\mathbb{R}^k$

$$\therefore \quad f(s) = \| y - Sx \|^2 = (y-Sx)^t M (y-Sx)$$
$$= (y^t - (Sx)^t)(My - MSx)$$

$$= y^t M y - 2y^t M S x + (Sx)^t M (Sx)$$

$$= x^t (S^t M S) x - 2 (S^t M y)^t x + \|y\|^2$$

$$\therefore \quad f(s) = x^t A x + b^t x + \gamma$$

La matriz $A = S^t M S$ es positiva definida, ya que $M > 0$ y S tiene columnas independientes.

Por lo tanto

$$\varphi(x) = x^t A x > 0$$

y por lo tanto $f(s)$ tiene mínimo, y éste es único.

■

COROLARIO

Sea $(V, <, >)$ un espacio euclideano, y $W \subseteq V$ un subespacio de V. Sea $v \in V$, entonces existe un único $\tilde{v} \in W$ tal que

$$\|v - \tilde{v}\|^2 = \min_{w \in W} \|v - w\|^2$$

■

PROYECCION ORTOGONAL

Sea $(V, <,>)$ un espacio euclideano, y sea $\mathcal{S} \subseteq V$ un subespacio de V. Definimos

$P_{\mathcal{S}} : V \to \mathcal{S}$ como sigue

$P_{\mathcal{S}}(\omega) = \tilde{\omega} \in \mathcal{S}$, donde

$$\| \omega - \tilde{\omega} \|^2 = \min_{s \in \mathcal{S}} \| \omega - s \|^2$$

Por el resultado anterior esta función está bien definida, y además

i) $P_{\mathcal{S}}(s) = s \quad \forall s \in \mathcal{S}$

ii) $P_{\mathcal{S}}(P_{\mathcal{S}}(\omega)) = P_{\mathcal{S}}(\omega) \quad \forall \omega \in \mathbb{R}^n$, es decir,

$P_{\mathcal{S}}^2 = P_{\mathcal{S}}$

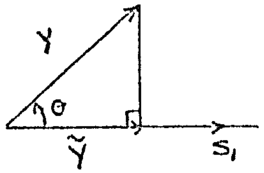
iii) $\| \omega - P_{\mathcal{S}}(\omega) \|^2 \leq \| \omega - s \|^2 \quad \forall s \in \mathcal{S}$

Como veremos más adelante, la propiedad (iii) de esta función implica que $\omega - P_{\mathcal{S}}(\omega)$ es ortogonal a $s \ \forall s \in \mathcal{S}$.

$P_{\mathcal{S}}$ se llama la PROYECCION ORTOGONAL SOBRE $\mathcal{S}$.

EJEMPLO

Veamos qué forma tiene $P_{\mathcal{S}}$ cuando $\dim(\mathcal{S})=1$.



Supongamos $\mathcal{S} = [s_1] \subseteq \mathbb{R}^n$.

Entonces $\tilde{y} = (\|y\| \cos\theta)\dfrac{s_1}{\|s_1\|} = \|y\| \dfrac{\langle s_1, y\rangle}{\|s_1\|\,\|y\|}\,\dfrac{s_1}{\|s_1\|}$.

$$\therefore \quad \tilde{y} = \frac{\langle s_1, y\rangle}{\|s_1\|^2} s_1 = \frac{\langle s_1, y\rangle}{\langle s_1, s_1\rangle} s_1 = P_{[s_1]}(y)$$

(a) si $\langle x, z\rangle = x^t z$ entonces

$$P_{\mathcal{S}}(y) = \frac{s_1^t y}{s_1^t s_1}\cdot s_1 = \frac{s_1 s_1^t}{s_1^t s_1} y$$

$$\therefore \quad P_{[s_1]} = \frac{s_1 s_1^t}{s_1^t s_1}$$

(b) si $\langle x, z\rangle = x^t M z$ entonces

$$P_{[s_1]}(y) = \frac{s_1^t M y}{s_1^t M s_1} s_1 = s_1 \cdot \frac{s_1^t M y}{s_1^t M s_1}$$

$$= \frac{s_1 s_1^t M}{s_1^t M s_1} y = \frac{s_1 (Ms_1)^t}{(Ms_1)^t s_1} y \qquad (M>0)$$

$$\therefore \; P_{[s_1]} = \frac{s_1 (Ms_1)^t}{(Ms_1)^t s_1}$$

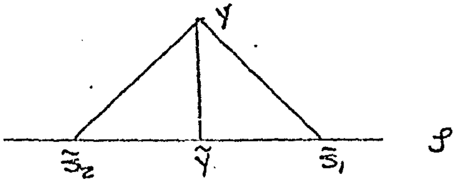
Veremos que si $\tilde{y} \in \mathcal{S}$ es tal que

$$y - \tilde{y} \perp s \quad \forall s \in \mathcal{S} \quad , \quad \text{entonces}$$

$\tilde{y} = P_{\mathcal{S}}(y)$.

Este resultado se puede entender mejor si nos fijamos en la geometría involucrada:

Supongamos $\mathcal{S} \subseteq \mathbb{R}^n$ es un subespacio de dimensión 1:



Si $y \notin \mathcal{S}$ y $\tilde{s}_1$ es cualquier punto sobre $\mathcal{S}$ tal que $y - \tilde{s}_1$ no es ortogonal a $\mathcal{S}$, entonces $\|y - \tilde{s}_1\|^2 > \min_{s \in \mathcal{S}} \|y - s\|^2$, ya que si formamos el triángulo isósceles $\tilde{s}_1 y \tilde{s}_2$ tal que $\|y - \tilde{s}_1\| = \|y - \tilde{s}_2\|$, entonces, como en un triángulo isósceles la perpendicular es "más chica" que los dos lados iguales, si $\tilde{y}$ es tal que $y-\tilde{y} \perp s$ para todo $s \in \mathcal{S}$, entonces

$$\| y-\tilde{y} \|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2 \quad y \text{ por lo tanto } \tilde{y} = P_{\mathcal{S}}(y).$$

Este mismo razonamiento se puede extender a subespacios de dimensión mayor. Por ejemplo, si $\dim(\mathcal{S}) = 2$, entonces para que $\tilde{y}$ sea tal que

$\| y-\tilde{y} \|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2$ necesitamos que $y-\tilde{y}$ sea ortogonal a cualquier recta en $\mathcal{S}$ que pase por $\tilde{y}$.- y por lo tanto $y-\tilde{y}$ debe ser ortogonal a cualquier vector $s \in \mathcal{S}$.

PROPIEDAD FUNDAMENTAL

Sea $\mathbb{R}^n$ con producto interior $\langle , \rangle$ y $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$. Sea $y \in \mathbb{R}^n$ y $\tilde{y} \in \mathcal{S}$. Entonces

$$\| y-\tilde{y} \|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2 \iff y-\tilde{y} \perp s \quad \forall\, s \in \mathcal{S}$$

Es decir, si $y-\tilde{y}$ es ortogonal a todo vector $s \in \mathcal{S}$, entonces $\tilde{y}$ es el punto más cercano a y de los puntos de $\mathcal{S}$.

Demostración:

Sea $f(s) = \|y-s\|^2$ para $s \in \mathcal{S}$.

Por lo tanto $f(s)$ tiene mínimo, y éste es único.

Veremos que $\tilde{y} \in \mathcal{S}$ tal que $y-\tilde{y} \perp s \ \forall s \in \mathcal{S}$ es el mínimo:

$$\|y-s\|^2 = \|(y-\tilde{y}) + (\tilde{y}-s)\|^2 \qquad \forall s \in \mathcal{S}$$

Como $\tilde{y}, s \in \mathcal{S} \Rightarrow \tilde{y}-s \in \mathcal{S}$, y por lo tanto, si $y-\tilde{y} \perp s \ \forall s \in \mathcal{S} \Rightarrow (y-\tilde{y}) \perp (\tilde{y}-s)$

$$\therefore \quad \|y-s\|^2 = \|y-\tilde{y}\|^2 + \|\tilde{y}-s\|^2 \qquad \forall s \in \mathcal{S}$$

$$\therefore \quad \|y-\tilde{y}\|^2 \leq \|y-s\|^2 \qquad \forall s \in \mathcal{S}$$

$$\therefore \quad \tilde{y} = P_{\mathcal{S}}(y)$$

Por lo tanto, por el resultado que demostramos sobre la existencia y unicidad de $P_{\mathcal{S}}(y) = \tilde{y}$:

$$y-\tilde{y} \perp s \ \forall s \in \mathcal{S} \Leftrightarrow \|y-\tilde{y}\|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2$$

$\blacksquare$

COROLARIO

Sea $\mathbb{R}^n$ con producto interior $<,>$ y $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de dimensión k.

Sea $\{s_1, s_2, \dots, s_{k'}\}$, $k' \geq k$, un conjunto de generadores de $\mathcal{S}$.

Sea $y \in \mathbb{R}^n$, $\tilde{y} \in \mathcal{S}$. Entonces

$$\|y-\tilde{y}\|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2 \Leftrightarrow y-\tilde{y} \perp s \quad \forall s \in \mathcal{S}$$

$$\Leftrightarrow y-\tilde{y} \perp s_i \quad 1 \leq i \leq k'$$

∎

COROLARIO

Sea $(\mathbb{V}, <,>)$ un espacio euclideano y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de dimensión k.

Sea $\{s_1, s_2, \dots, s_{k'}\}$, $k' \geq k$, un conjunto de generadores de $\mathcal{S}$.

Sea $\omega \in \mathbb{V}$, $\tilde{\omega} \in \mathcal{S}$. Entonces

$$\|\omega-\tilde{\omega}\|^2 = \min_{s \in \mathcal{S}} \|\omega-s\|^2 \Leftrightarrow \omega-\tilde{\omega} \perp s \quad \forall s \in \mathcal{S}$$

$$\Leftrightarrow \omega-\tilde{\omega} \perp s_i, \quad 1 \leq i \leq k'$$

∎

ECUACIONES NORMALES

Sea $(\mathbb{V}, <,>)$ un espacio euclideano y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de $\mathbb{V}$ de dimensión k.

Sea $\{s_1, s_2, \ldots, s_k\}$ una base de $\mathcal{S}$.

Si $\omega \in \mathbb{V}$ entonces, por el corolario anterior

$$\|\omega - \tilde{\omega}\|^2 = \min_{s \in \mathcal{S}} \|\omega - s\|^2 \iff \omega - \tilde{\omega} \perp s_i, \quad 1 \leq i \leq k$$

$$\iff \langle \omega - \tilde{\omega}, s_i \rangle = 0, \quad 1 \leq i \leq k$$

$$\iff \langle \omega, s_i \rangle = \langle \tilde{\omega}, s_i \rangle, \quad 1 \leq i \leq k$$

$$\iff \langle s_i, \tilde{\omega} \rangle = \langle s_i, W \rangle, \quad 1 \leq i \leq k$$

El sistema de ecuaciones

$$\boxed{\langle s_i, \tilde{\omega} \rangle = \langle s_i, \omega \rangle \qquad 1 \leq i \leq k}$$

se llama ECUACIONES NORMALES.

Si $\mathbb{V} = \mathbb{R}^n$, entonces, si $y \in \mathbb{V}$

$$\|y - \tilde{y}\|^2 = \min_{s \in \mathcal{S}} \|y - s\|^2 \iff \langle s_i, \tilde{y} \rangle = \langle s_i, y \rangle$$

$$1 \leq i \leq k$$

Como $\hat{y} \in \mathcal{S} \Rightarrow \hat{y} = Sx$ para alguna $x \in \mathbb{R}^k$, donde

$$S = [s_1 | s_2 | \cdots | s_k]$$

por lo tanto las ecuaciones normales se pueden escribir de la siguiente manera:

$$\langle s_i , Sx \rangle = \langle s_i , y \rangle \qquad 1 \leq i \leq k$$

Escribiremos las ecuaciones normales en el lenguaje matricial:

A) Supongamos $\langle x, z \rangle = x^t z$. Entonces

$$\langle s_i , Sx \rangle = \langle s_i , y \rangle \qquad 1 \leq i \leq k$$

se transforma en :

$$s_i^t \, Sx = s_i^t \, y \qquad 1 \leq i \leq k$$

$$\therefore \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_k^t \end{bmatrix} Sx = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_k^t \end{bmatrix} y$$

$$\therefore \boxed{S^t S \, x = S^t y}$$

Como ya sabemos, esta ecuación debe de tener solución, lo cual se comprueba observando que como S tiene columnas independientes entonces $S^tS > 0$, y por lo tanto el sistema tiene solución única.

B) Supongamos $\langle x, z\rangle = x^t M z$, $M>0$. Entonces las ecuaciones normales:

$$\langle s_i, Sx\rangle = \langle s_i, y\rangle \qquad 1 \leq i \leq k$$

se convierten en:

$$s_i^t M (Sx) = s_i^t M y \qquad 1 \leq i \leq k$$

$$\therefore \begin{array}{c} s_1^t M (Sx) = s_1^t M y \\ s_2^t M (Sx) = s_2^t M y \\ \vdots \qquad\qquad \vdots \\ s_k^t M (Sx) = s_k^t M y \end{array}$$

$$\begin{bmatrix} s_1^t M \\ s_2^t M \\ \vdots \\ s_k^t M \end{bmatrix} (Sx) = \begin{bmatrix} s_1^t M \\ s_2^t M \\ \vdots \\ s_k^t M \end{bmatrix} y$$

$$\therefore \quad (S^t M)(Sx) = (S^t M) y$$

Por lo tanto, en el caso más general de producto interior en $\mathbb{R}^n$, las ecuaciones normales se pueden escribir de la forma

$$\boxed{(MS)^t(Sx) = (MS)^t y}$$

Al igual que para $M = I_n$, podemos checar que este sistema de ecuaciones siempre tiene solución, y ésta es única. Esto se ve fácilmente si observamos que la matriz del sistema es positiva definida:

$$(MS)^t S = S^t M S$$

$$\therefore \quad x^t (S^t M S) x = (Sx)^t M (Sx) > 0 \quad \text{si } x \neq 0,$$

ya que $M > 0$ y S tiene columnas independientes.

Intuimos que este sistema debe de ser muy fácil de resolver si $\{s_1, s_2, \dots, s_k\}$ es una base ortogonal de $\mathcal{S}$:

TEOREMA DE LA PROYECCION

Sea $(V, \langle , \rangle)$ un espacio euclideano y $\mathcal{S} \subseteq V$ un subespacio de V de dimensión k. Sea $P_\mathcal{S} : V \to \mathcal{S}$ la proyección ortogonal sobre $\mathcal{S}$. Entonces

$$P_\mathcal{S} = P_{[s_1]} + P_{[s_2]} + \cdots + P_{[s_k]} \qquad , \text{ donde}$$

$\{s_1, s_2, \ldots, s_k\}$ es una base ortogonal de $\mathcal{S}$.

Demostración:

Sea $\omega \in V$, entonces $P_{\mathcal{S}}(\omega) = \tilde{\omega} \iff$

$$\langle s_i, \tilde{\omega} \rangle = \langle s_i, \omega \rangle \qquad 1 \leq i \leq k$$

donde $\{s_1, s_2, \ldots, s_k\}$ es una base de $\mathcal{S}$.

Como $\tilde{\omega} \in \mathcal{S} \Rightarrow \tilde{\omega} = \xi_1 s_1 + \xi_2 s_2 + \cdots + \xi_k s_k$

$$= \sum_{j=1}^{k} \xi_j s_j$$

Por lo tanto

$$P_{\mathcal{S}}(\omega) = \tilde{\omega} \iff \langle s_i, \sum_{j=1}^{k} \xi_j s_j \rangle = \langle s_i, \omega \rangle \qquad 1 \leq i \leq k$$

$$\iff \sum_{j=1}^{k} \xi_j \langle s_i, s_j \rangle = \langle s_i, \omega \rangle, \qquad 1 \leq i \leq k$$

Si $\{s_1, s_2, \ldots, s_k\}$ es ortogonal, entonces $\langle s_i, s_j \rangle = 0$ para $i \neq j$, y por lo tanto

$$P_{\mathcal{S}}(\omega) = \tilde{\omega} \iff \xi_i \langle s_i, s_i \rangle = \langle s_i, \omega \rangle, \quad 1 \leq i \leq k$$

$$\iff \xi_i = \frac{\langle s_i, \omega \rangle}{\langle s_i, s_i \rangle}$$

Por lo tanto

$$P_{\mathcal{S}}(w) = \frac{\langle s_1, w\rangle}{\langle s_1, s_1\rangle} s_1 + \frac{\langle s_2, w\rangle}{\langle s_2, s_2\rangle} s_2 + \cdots + \frac{\langle s_k, w\rangle}{\langle s_k, s_k\rangle} s_k$$

Pero, como habíamos visto anteriormente

$$P_{[s_i]}(w) = \frac{\langle s_i, w\rangle}{\langle s_i, s_i\rangle} s_i, \qquad 1 \leq i \leq k$$

$$\therefore \quad P_{\mathcal{S}}(w) = P_{[s_1]}(w) + P_{[s_2]}(w) + \cdots + P_{[s_k]}(w)$$

Supongamos que $V = \mathbb{R}^n$ y $\langle x, z\rangle = x^t M z$, entonces la matriz asociada al sistema de ecuaciones normales es diagonal, cuando $\{s_1, s_2, \dots, s_k\}$ es una base ortogonal de $\mathcal{S}$:

En este caso, el sistema de ecuaciones normales está dado por:

$$(MS)^t (Sx) = (MS)^t y \qquad \text{(ver pgs. 36-38)}$$

Si las columnas de S son ortogonales entonces

$$s_i^t M s_j = 0 \qquad i \neq j$$

es decir, $\quad (Ms_i)^t s_j = 0 \qquad i \neq j$

Por lo tanto

$$MS = M\left[s_1|s_2|\cdots|s_k\right] = \left[Ms_1\,|\,Ms_2|\cdots|Ms_k\right]$$

$$\therefore (MS)^t = \begin{bmatrix} (Ms_1)^t \\ (Ms_2)^t \\ \vdots \\ (Ms_k)^t \end{bmatrix}$$

$$\therefore (MS)^t S = \begin{bmatrix} (Ms_1)^t \\ (Ms_2)^t \\ \vdots \\ (Ms_k)^t \end{bmatrix} \left[s_1|s_2|\cdots|s_k\right]$$

$$= \begin{bmatrix} (Ms_1)^t s_1 & (Ms_1)^t s_2 & \cdots & (Ms_1)^t s_k \\ (Ms_2)^t s_1 & (Ms_2)^t s_2 & \cdots & (Ms_2)^t s_k \\ \vdots & \vdots & & \vdots \\ (Ms_k)^t s_1 & (Ms_k)^t s_2 & \cdots & (Ms_k)^t s_k \end{bmatrix}$$

$$= \begin{bmatrix} (Ms_1)^t s_1 & & & \\ & (Ms_2)^t s_2 & & \\ & & \ddots & \\ & & & (Ms_k)^t s_k \end{bmatrix}$$

$$= D$$

En este caso, $P_{\mathcal{S}} : \mathbb{R}^n \to \mathcal{S}$ estará dado por una matriz. Veamos cómo es esta matriz:

Por el teorema de la proyección:

$$P_{\mathcal{S}} = P_{[s_1]} + P_{[s_2]} + \cdots + P_{[s_k]}$$

Pero $\quad P_{[s_i]} = \dfrac{s_i (Ms_i)^t}{(Ms_i)^t s_i} \qquad 1 \leq i \leq k \qquad$ (ver pag. VI.32)

Por lo tanto

$$P_{\mathcal{S}} = \frac{s_1 (Ms_1)^t}{(Ms_1)^t s_1} + \frac{s_2 (Ms_2)^t}{(Ms_2)^t s_2} + \cdots + \frac{s_k (Ms_k)^t}{(Ms_k)^t s_k}$$

Observemos que si $\langle , \rangle$ es el producto interno usual de $\mathbb{R}^n$, es decir $\langle x, z \rangle = x^t z = x^t M z$, con $M = I_n$, entonces

$$P_{\mathcal{S}} = \frac{s_1 s_1^t}{s_1^t s_1} + \frac{s_2 s_2^t}{s_2^t s_2} + \cdots + \frac{s_k s_k^t}{s_k^t s_k}$$

■

ALGORITMO DE GRAM-SCHMIDT

El algoritmo de Gram-Schmidt nos proporciona un método para resolver el siguiente

PROBLEMA

Sea $(V,<,>)$ un espacio euclideano y $\mathcal{S} \subseteq V$ un subespacio de V de dimensión k.

Dada $\{s_1,s_2,\ldots,s_k\} \subseteq \mathcal{S}$ una base de $\mathcal{S}$, construir una base ortogonal de $\mathcal{S}$ a partir de $\{s_1,s_2,\ldots,s_k\}$.

SOLUCION

1) Sea $q_1 = s_1$

2) Para $i=2,3,\ldots,k$ hacemos lo siguiente :

Calcular $\tilde{s}_i$ tal que $\tilde{s}_i$ es el punto más cercano a s_i del subespacio generado por $q_1,q_2,\ldots,q_{i-1}$. Es decir, $\tilde{s}_i$ es la proyección ortogonal de s_i sobre $[q_1,q_2,\ldots,q_{i-1}]$.

Por lo tanto $s_i - \tilde{s}_i \perp s \quad \forall s \in [q_1,\ldots,q_{i-1}]$

Sea $q_i = s_i - \tilde{s}_i$.

al terminar, en el k-ésimo paso, obtenemos un conjunto de k vectores

$$\{q_1, q_2, \ldots, q_k\} \subseteq V$$

Veremos que $\{q_1, q_2, \ldots, q_k\}$ es una base ortogonal de $\mathcal{S}$:

Es claro que $\{q_1, q_2, \ldots, q_k\}$ es ortogonal, ya que, por construcción, en cada paso

$$q_i \perp q_1, q_2, \ldots, q_{i-1} \qquad , \quad 2 \leq i \leq k$$

y por lo tanto $\{q_1, q_2, \ldots, q_i\}$. es ortogonal, para $1 \leq i \leq k$.

Observemos además, que en el primer paso hacemos

$$q_1 = s_1$$

En el segundo paso hacemos

$$q_2 = s_2 - P_{12}\, q_1 \qquad , \text{ donde } P_{12} = \frac{\langle q_1, s_2 \rangle}{\langle q_1, q_1 \rangle}$$

Y en general, en el i-ésimo paso, $2 \leq i \leq k$ hacemos

$$q_i = s_i - P_{1i}\, q_1 - P_{2i}\, q_2 - \cdots - P_{i-1,i}\, q_{i-1} \; ; \qquad \text{donde}$$

$$p_{1i} = \frac{\langle q_1, s_i \rangle}{\langle q_1, q_1 \rangle}, \quad p_{2i} = \frac{\langle q_2, s_i \rangle}{\langle q_2, q_2 \rangle}, \quad \ldots, \quad p_{i-1,i} = \frac{\langle q_{i-1}, s_i \rangle}{\langle q_{i-1}, q_{i-1} \rangle}$$

Es decir, tenemos que

$$(*) \quad \begin{cases} q_1 = s_1 \\ q_2 = s_2 - p_{12}\, q_1 \\ \quad \vdots \qquad\qquad \vdots \\ q_k = s_k - p_{1k}\, q_1 - p_{2k}\, q_2 - \cdots - p_{k-1,k}\, q_{k-1} \end{cases}$$

y por lo tanto

$$\begin{aligned} q_1 &\in [s_1] \\ q_2 &\in [s_2, q_1] \subseteq [s_2, s_1] \\ q_3 &\in [s_3, q_1, q_2] \subseteq [s_1, s_2, s_3] \\ &\;\vdots \qquad\qquad\qquad \vdots \\ q_k &\in [s_k, q_1, q_2, \ldots, q_{k-1}] \subseteq [s_1, s_2, \ldots, s_k] \end{aligned}$$

Por lo tanto $[q_1, q_2, \ldots, q_k] \subseteq [s_1, s_2, \ldots, s_k]$.

Por lo tanto $[q_1, q_2, \ldots, q_k] \subseteq \mathcal{S}$, y como la dimensión de $\mathcal{S}$ es k, y $\{q_1, q_2, \ldots, q_k\}$ es independiente $\Rightarrow$

$\{q_1, q_2, \ldots, q_k\}$ es una base ortogonal de $\mathcal{S}$.

$\blacksquare$

El algoritmo de Gram-Schmidt nos resuelve por completo los siguientes problemas

PROBLEMA 1.

Sea $(\mathbb{V}, <,>)$ un espacio euclideano. Construir una base ortogonal de $\mathbb{V}$ a partir de cualquier base de $\mathbb{V}$.

Solución:

Es inmediata, si aplicamos el procedimiento de Gram-Schmidt a $\mathcal{S} = \mathbb{V}$.

PROBLEMA 2.

Sea $(\mathbb{V}, <,>)$ un espacio euclideano, y $\mathcal{S} \subseteq \mathbb{V}$ un subespacio de $\mathbb{V}$. Dado $\omega \in \mathbb{V}$ encontrar la proyección ortogonal de ω sobre $\mathcal{S}$. Es decir, encontrar $\tilde{\omega} \in \mathcal{S}$ tal que

$$\|\omega - \tilde{\omega}\|^2 = \min_{s \in \mathcal{S}} \|\omega - s\|^2 .$$

Solución:

Sea $\{s_1, s_2, \ldots, s_k\}$ una base de $\mathcal{S}$, y sea $s_{k+1} = \omega$.

Si construimos un conjunto ortogonal $\{q_1, q_2, \dots, q_k, q_{k+1}\}$ a partir de $\{s_1, s_2, \dots, s_k, s_{k+1}\}$, entonces

$$q_{k+1} = s_{k+1} - P_{[q_1, q_2, \dots, q_k]}(s_{k+1}) = s_{k+1} - \tilde{s}_{k+1}$$

$$= w - P_{\mathcal{S}}(w) = w - \tilde{s}_{k+1}$$

y por lo tanto $P_{\mathcal{S}}(w) = \tilde{s}_{k+1}$ (ver pg. VI.43)

Sin embargo, éstas no son las únicas consecuencias que se pueden obtener al aplicar el algoritmo de Gram-Schmidt a un conjunto independiente de vectores.

Supongamos ahora que $V = \mathbb{R}^n$ y que obtuvimos una base ortogonal $\{q_1, q_2, \dots, q_k\}$ de $\mathcal{S} \subseteq \mathbb{R}^n$ aplicando el procedimiento de Gram-Schmidt a la base $\{s_1, s_2, \dots, s_k\}$ de $\mathcal{S}$.

Entonces (ver pag. VI.45)

$$q_1 = s_1$$
$$q_2 = s_2 - \rho_{12} q_1$$
$$\vdots$$
$$q_k = s_k - \rho_{1k} q_1 - \rho_{2k} q_2 - \dots - \rho_{k-1,k} q_{k-1}$$

donde:

$$p_{12} = \frac{\langle q_1, s_2\rangle}{\langle q_1, q_1\rangle}$$

$$p_{13} = \frac{\langle q_1, s_3\rangle}{\langle q_1, q_1\rangle}, \quad p_{23} = \frac{\langle q_2, s_3\rangle}{\langle q_2, q_2\rangle}$$

$$\vdots \qquad\qquad \vdots$$

$$p_{1k} = \frac{\langle q_1, s_k\rangle}{\langle q_1, q_1\rangle}, \quad p_{2k} = \frac{\langle q_2, s_k\rangle}{\langle q_2, q_2\rangle}, \quad \ldots, \quad p_{k-1,k} = \frac{\langle q_{k-1}, s_k\rangle}{\langle q_{k-1}, q_{k-1}\rangle}$$

Este sistema de ecuaciones lo podemos reescribir de la forma:

$$\begin{aligned} s_1 &= q_1 \\ s_2 &= p_{12} q_1 + q_2 \\ s_3 &= p_{13} q_1 + p_{23} q_2 + q_3 \\ &\vdots \\ s_k &= p_{1k} q_1 + p_{2k} q_2 + p_{3k} q_3 + \cdots + p_{k-1,k} q_{k-1} + q_k \end{aligned}$$

Y como $s_i, q_i \in \mathbb{R}^n$ para $1 \leq i \leq k$, entonces lo podemos escribir en forma matricial:

$$[s_1 | s_2 | \cdots | s_k] = [q_1 | q_2 | \cdots | q_k] \begin{bmatrix} 1 & p_{12} & p_{13} & \cdots & p_{1,k-1} & p_{1k} \\ 0 & 1 & p_{23} & \cdots & p_{2,k-1} & p_{2k} \\ 0 & 0 & 1 & \cdots & p_{3,k-1} & p_{3,k} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & p_{k-1,k} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}$$

Por lo tanto si tenemos una base $\{s_1, s_2, \dots, s_k\}$ de $\mathcal{S} \subseteq \mathbb{R}^n$, entonces a partir de ésta podemos construir una base ortogonal $\{q_1, q_2, \dots, q_k\}$ de $\mathcal{S}$, y éstas dos bases están relacionadas de la siguiente manera:

$$\underset{n\times k}{S} = \underset{n\times k}{Q}\,\underset{k\times k}{R}$$

donde $S = [s_1 | s_2 | \cdots | s_k]$ y $Q = [q_1 | q_2 | \cdots | q_k]$

y R es una matriz triangular superior no-singular, de orden k.

Es decir, la matriz que da el cambio de coordenadas de las s's a las q's es R, ya que si

$$s = Sx \in \mathcal{S}$$

$$\Rightarrow \qquad s = Sx = Q(Rx)$$

RESULTADO

Sea S $n\times k$ con columnas independientes, entonces existe Q $n\times k$ con columnas ortogonales y $R_{k\times k}$ triangular superior no-singular tal que

$$S = QR$$

Demostración:

Aplicar el algoritmo de Gram-Schmidt a las columnas de S.

Volvamos al punto central de que nos hemos estado ocupando en este capítulo:

Sea $\mathbb{R}^n$ con producto interior $\langle x, z\rangle = x^t M z$, $M > 0$, y sea $\mathcal{S} \in \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$ de dimensión k. Dado $y \in \mathbb{R}^n$ encontrar $\tilde{y} \in \mathcal{S}$, donde $\tilde{y}$ es la proyección ortogonal de y sobre $\mathcal{S}$.

Supongamos que $\mathcal{S} = [s_1, s_2, \ldots, s_k]$ y sea

$\underset{n\times k}{S} = [s_1 | s_2 | \ldots | s_k]$.

Si tenemos la factorización $S = QR$ de S, donde $Q_{n\times k}$ tiene columnas ortogonales y $\underset{k\times k}{R}$ es triangular superior no-singular, entonces el problema ya está prácticamente resuelto:

Las ecuaciones normales están dadas por

$$(MS)^t (Sx) = (MS)^t y$$

por lo tanto $[M(QR)]^t (QR)x = [M(QR)]^t y$

$$\Rightarrow \quad R^t (Q^t MQ) Rx = R^t (Q^t M) y$$

Como R es invertible $\Rightarrow R^t$ es invertible y por lo tanto

$$(R^{-t} R^t)(Q^t MQ) Rx = (R^{-t} R^t)(Q^t M) y$$

$$\Rightarrow \quad (Q^t MQ) Rx = Q^t M y$$

Como las columnas de Q son ortogonales $\Rightarrow$

$$q_i^t M q_j = 0 \quad \text{si } i \neq j$$

$$q_i^t M q_i > 0$$

y por lo tanto la matriz $Q^t MQ$ es diagonal y positiva definida, y por lo tanto es invertible

$$\therefore \quad Rx = (Q^t MQ)^{-1} (Q^t M) y$$

$$\therefore \quad \tilde{y} = Sx = QRx = Q(Q^t MQ)^{-1} (Q^t M) y = P_S(y)$$

Por lo tanto $P_S = Q(Q^t MQ)^{-1}(Q^t M)$
$= Q D^{-1} Q^t M$

donde $D = Q^t MQ$

Entonces tendremos que :

$$Q^t M Q = \begin{bmatrix} q_1^t M q_1 & & & \\ & q_2^t M q_2 & & \\ & & \ddots & \\ & & & q_k^t M q_k \end{bmatrix}$$

$$= \begin{bmatrix} \|q_1\|^2 & & & \\ & \|q_2\|^2 & & \\ & & \ddots & \\ & & & \|q_k\|^2 \end{bmatrix}$$

y por lo tanto $D^{-1} = \operatorname{diag}\left(1/\delta_i \right)$, donde $\delta_i = \|q_i\|^2$.

Por lo tanto $P_\rho = Q D^{-1} Q^t M =$

$$= [q_1 | q_2 | \cdots | q_k] \begin{bmatrix} \|q_1\|^{-2} & & \\ & \|q_2\|^{-2} & \\ & & \|q_k\|^{-2} \end{bmatrix} \begin{bmatrix} (Mq_1)^t \\ (Mq_2)^t \\ \vdots \\ (Mq_k)^t \end{bmatrix}$$

$$= \left[\|q_1\|^{-2} q_1 \,|\, \|q_2\|^{-2} q_2 \,|\, \cdots \,|\, \|q_k\|^{-2} q_k \right] \begin{bmatrix} (Mq_1)^t \\ (Mq_2)^t \\ \vdots \\ (Mq_k)^t \end{bmatrix}$$

$$\therefore \quad P_\rho = \frac{q_1 (Mq_1)^t}{\|q_1\|^2} + \frac{q_2 (Mq_2)^t}{\|q_2\|^2} + \cdots + \frac{q_k (Mq_k)^t}{\|q_k\|^2}$$

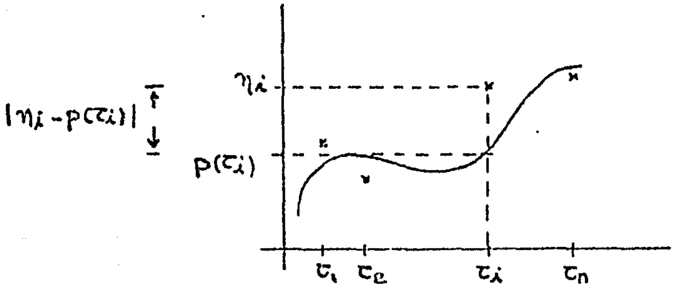
Confirmando el resultado que habíamos obtenido antes (el teorema de la proyección).

EJEMPLOS

(I) Dado un conjunto de n datos experimentales,

$$(\tau_i, \eta_i) \qquad 1 \leq i \leq n$$

encontrar el polinomio de grado $\leq k-1$ que mejor aproxime los datos.



Supongamos $\quad p(\tau) = \xi_1 + \xi_2 \tau + \xi_3 \tau^2 + \cdots + \xi_k \tau^{k-1}$

Necesitamos encontrar $\xi_1, \xi_2, \dots, \xi_k \in \mathbb{R}$ tal que

$$\sum_{i=1}^{n} (\eta_i - p(\tau_i))^2 \quad \text{sea} \quad \text{mínimo} .$$

Sean
$$X = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix} \in \mathbb{R}^k \quad , \quad Y = \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} \in \mathbb{R}^n$$

Entonces $p(\tau)$ se puede escribir de la forma

$$p(\tau) = \begin{bmatrix} 1 & \tau & \tau^2 & \dots & \tau^{k-1} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_k \end{bmatrix}$$

Por lo tanto si hacemos

$$z_i^t = \begin{bmatrix} 1 & \tau_i & \tau_i^2 & \dots & \tau_i^{k-1} \end{bmatrix} \quad , \quad 1 \leq i \leq n$$

entonces $p(\tau_i) = z_i^t X \quad , \quad 1 \leq i \leq n$

Por lo tanto queremos minimizar

$$f(X) = \sum_{i=1}^{n} (\eta_i - p(\tau_i))^2 = \sum_{i=1}^{n} (\eta_i - z_i^t X)^2$$

$$= \left\| \begin{bmatrix} \eta_1 - z_1^t x \\ \eta_2 - z_2^t x \\ \vdots \\ \eta_n - z_n^t x \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix} - \begin{bmatrix} z_1^t \\ z_2^t \\ \vdots \\ z_n^t \end{bmatrix} x \right\|^2$$

$$= \| y - Sx \|^2$$ donde

$$S = \begin{bmatrix} z_1^t \\ z_2^t \\ \vdots \\ z_n^t \end{bmatrix} = \begin{bmatrix} 1 & \tau_1 & \tau_1^2 & \cdots & \tau_1^{k-1} \\ 1 & \tau_2 & \tau_2^2 & \cdots & \tau_2^{k-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \tau_n & \tau_n^2 & \cdots & \tau_n^{k-1} \end{bmatrix}$$

Hemos traducido el problema al lenguaje matricial :

Encontrar el polinomio que mejor se aproxime a los datos equivale a encontrar el punto de R(S) que sea el más cercano a y.

(II) Sea $C[\alpha,\beta] = \left\{ f : [\alpha,\beta] \to \mathbb{R} \text{ tal que } f \text{ es continua} \right\}$

Entonces $C[\alpha,\beta]$ es un espacio vectorial que no es finitamente generado, ya que se pueden encontrar conjuntos independientes de funciones continuas tan grandes como se quiera. Por ejemplo:

$$\{1, \tau, \tau^2, \dots, \tau^n, \dots\} \subseteq C[\alpha, \beta]$$

es independiente.

Definamos un producto interior $\langle , \rangle_\omega : C[\alpha,\beta] \rightarrow \mathbb{R}$ como sigue:

Si $f, g \in C[\alpha, \beta]$ entonces

$$\langle f, g \rangle_\omega = \int_\alpha^\beta \omega(\tau)\, f(\tau)\, g(\tau)\, d\tau \quad , \text{ donde}$$

$\omega \in C[\alpha,\beta]$ y $\omega(\tau) > 0 \quad \forall\, \tau \in [\alpha, \beta]$.

Veremos que esta función efectivamente define un producto interior en $C[\alpha,\beta]$:

(1) $$\langle f, f \rangle_W = \int_\alpha^\beta \omega(\tau)\, f^2(\tau) \geq 0 \qquad \left(\begin{matrix} \omega(\tau) f^2(\tau) \geq 0 \\ \forall\, \tau \end{matrix} \right)$$

$$\begin{aligned} \langle f, f \rangle_\omega = 0 &\iff \omega(\tau)\, f^2(\tau) = 0 \quad \forall\, \tau \in [\alpha,\beta] \\ &\iff f^2(\tau) = 0 \quad \forall\, \tau \in [\alpha,\beta] \\ &\iff f = 0 \end{aligned}$$

$$(2) \quad \langle f,g\rangle_\omega = \int_\alpha^\beta \omega(\tau) f(\tau) g(\tau)\, d\tau = \int_\alpha^\beta \omega(\tau) g(\tau) f(\tau)\, d\tau$$

$$= \langle g,f\rangle_\omega$$

$$(3) \quad \langle \lambda f_1 + f_2, g\rangle_\omega = \int_\alpha^\beta \omega(\tau)\, [\lambda f_1 + f_2](\tau)\, g(\tau)\, d\tau$$

$$= \lambda \int_\alpha^\beta \omega(\tau) f_1(\tau) g(\tau)\, d\tau + \int_\alpha^\beta \omega(\tau) f_2(\tau) g(\tau)\, d\tau$$

$$= \lambda \langle f_1,g\rangle_\omega + \langle f_2,g\rangle_\omega$$

Por lo tanto $\langle f,g\rangle_\omega$ es un producto interior en $C[\alpha,\beta]$.

Por lo tanto la función $\|\cdot\|_\omega : C[\alpha,\beta] \to \mathbb{R}$ dada por

$$\|f\|_\omega = \sqrt{\langle f,f\rangle_\omega}$$

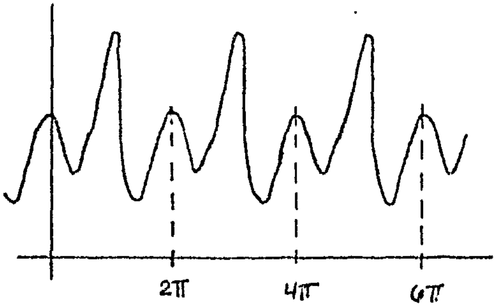
es una norma en $C[\alpha,\beta]$.

Solamente nos interesan subespacios de dimensión finita de $C[\alpha,\beta]$.

(III) Series de Fourier

Sea $f = f_e$ una función obtenida por métodos experimentales tal que esperamos que f_e sea periódica de período 2π, es decir

$$f_e(\tau) = f_e(\tau + 2\pi)$$



Un fenómeno periódico puede representarse como superposición de funciones elementales :

$$1, \text{sen}\,\tau, \cos\tau, \text{sen}\,2\tau, \cos 2\tau, \ldots$$

Veremos cómo aproximar la función fe usando un número finito de tales funciones.

Sea V el subespacio de funciones periódicas, y sea $W \subseteq V$

$$W = [1, \text{sen}\,\tau, \cos\tau, \text{sen}\,2\tau, \cos 2\tau, \ldots, \text{sen}\,n\tau, \cos n\tau]$$

El conjunto $\{1, \operatorname{sen} \tau, \cos \tau, \ldots, \operatorname{sen} n\tau, \cos n\tau\}$ es ortogonal en $[0,2\pi]$; ya que

$$\langle \operatorname{sen} k\tau, \cos l\tau \rangle = \int_0^{2\pi} \operatorname{sen} k\tau \, \cos k\tau \, d\tau = 0 \quad \forall k, l$$

$$\langle \operatorname{sen} k\tau, \operatorname{sen} l\tau \rangle = \int_0^{2\pi} \operatorname{sen} k\tau \, \operatorname{sen} l\tau \, d\tau = 0 \quad \text{si } k \neq l$$

$$\langle \cos k\tau, \cos l\tau \rangle = \int_0^{2\pi} \cos k\tau \, \operatorname{sen} l\tau \, d\tau = 0 \quad \text{si } k \neq l$$

$$\langle 1, \operatorname{sen} k\tau \rangle = \int_0^{2\pi} \operatorname{sen} k\tau \, d\tau = 0 \quad \forall k$$

$$\langle 1, \cos k\tau \rangle = \int_0^{2\pi} \cos k\tau \, d\tau = 0 \quad \forall k$$

Por lo tanto $\{1, \operatorname{sen} \tau, \cos \tau, \ldots, \operatorname{sen} n\tau, \cos n\tau\}$ es una base de W.

Encontraremos la mejor aproximación $\hat{f} \in W$ de nuestra función, es decir, calcularemos $\hat{f} \in W$ tal que

$$\| f_e - \hat{f} \|^2 = \min_{p \in W} \| f_e - p \|^2$$

Supongamos $\hat{f}(\tau) = \alpha_0 + \sum_{k=1}^{n} (\alpha_k \cos k\tau + \beta_k \operatorname{sen} k\tau)$

Por lo tanto, necesitamos que $f_e(t) - \hat{f}(t)$ sea ortogonal a $p(t)$, donde $p(t)$ es cualquier función en W.

Como $\{1, \operatorname{sen} t, \cos t, \ldots, \operatorname{sen} nt, \cos nt\}$ es una base ortogonal de W, entonces

$$\alpha_0 = \frac{\langle 1, f_e \rangle}{\langle 1, 1 \rangle} = \frac{1}{2\pi} \int_0^{2\pi} f_e(t)\,dt$$

$$\alpha_k = \frac{\langle \cos kt, f_e(t) \rangle}{\langle \cos kt, \cos kt \rangle} = \frac{1}{\pi} \int_0^{2\pi} f_e(t) \cos kt \, dt$$

$$\beta_k = \frac{\langle \operatorname{sen} kt, f_e(t) \rangle}{\langle \operatorname{sen} kt, \operatorname{sen} kt \rangle} = \frac{1}{\pi} \int_0^{2\pi} f_e(t) \operatorname{sen} kt \, dt$$

Y por lo tanto la función $\hat{f} \in W$ que mejor aproxima a la función f_e es

$$\hat{f}(t) = \alpha_0 + \sum_{k=1}^{n} (\alpha_k \cos kt + \beta_k \operatorname{sen} kt)$$

donde $\alpha_0, \alpha_k, \beta_k$ $(1 \leq k \leq n)$ son como arriba.

PROCEDIMIENTO DE GRAM-SCHMIDT MODIFICADO

Nosotros resolvimos el problema de encontrar el punto más cercano de un subespacio a un vector dado, y obtuvimos que toda matriz con columnas independientes se puede factorizar de la forma

$$S = QR$$

donde Q tiene columnas ortogonales y R es triangular superior, con 1's en la diagonal.

Sin embargo, podríamos haber empezado por preguntarnos si toda matriz S, con $N(S) = \{0\}$, se puede factorizar de tal manera, sin considerar el aspecto geométrico del problema, sino únicamente el aspecto técnico:

Dada $S_{n \times k}$ con columnas independientes encontrar $Q_{n \times k}$ con columnas ortogonales y $R_{k \times k}$ triangular superior, con 1's en la diagonal, tal que $S = QR$.

El algoritmo de Gram-Schmidt nos proporciona un método para calcular la descomposición $S = QR$ de una matriz S con columnas independientes. Veamos más detalladamente qué significa ésto en términos de las matrices involucradas.

Aplicar el algoritmo de Gram-Schmidt a las columnas de S equivale a resolver el sistema $S = QR$ para Q y R.

Si $S = [s_1 | s_2 | \cdots | s_k]$, $Q = [q_1 | q_2 | \cdots | q_k]$ y

$$R = \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1k} \\ & 1 & \rho_{23} & \cdots & \rho_{2k} \\ & & 1 & \cdots & \rho_{3k} \\ & & & \ddots & \vdots \\ & & & & 1 \end{bmatrix}$$

Entonces el sistema se puede escribir de la siguiente manera (ver pag VI.45):

$$(*) \quad \begin{aligned} q_1 &= s_1 \\ q_2 &= s_2 - \rho_{12} q_1 \\ q_3 &= s_3 - \rho_{13} q_1 - \rho_{23} q_2 \\ &\vdots \\ q_k &= s_k - \rho_{1k} q_1 - \rho_{2k} q_2 - \cdots - \rho_{k-1,k} q_{k-1} \end{aligned}$$

En cada paso del algoritmo calculamos una columna de Q y una columna de R: en el i-ésimo paso del procedimiento de Gram-Schmidt calculamos q_i y la i-ésima columna de la matriz R.

Observemos que al resolver el sistema $(*)$ según el algoritmo de Gram-Schmidt, entonces en el i-ésimo paso:

para $2 \leq i \leq k$ tenemos que restar un múltiplo de q_1 a s_i

para $3 \leq i \leq k$ tenemos que restar un múltiplo de q_2 a $(s_i - P_{1i} q_1)$

para $4 \leq i \leq k$ tenemos que restar un múltiplo de q_3 a $(s_i - P_{1i} q_1 - P_{2i} q_2)$

$\vdots$

para $i = k$ tenemos que restar un múltiplo de q_{k-1} a $(s_k - P_{1k} q_1 - P_{2k} q_2 - \cdots - P_{k-2,k} q_{k-2})$

Esto nos sugiere resolver el sistema $S = QR$ por medio del siguiente algoritmo:

En el primer paso hacemos

$$
\begin{aligned}
q_1 &= s_1 = s_1^{(0)} \\
s_2^{(1)} &= s_2 - P_{12} q_1 \\
s_3^{(1)} &= s_3 - P_{13} q_1 \\
&\vdots \\
s_k^{(1)} &= s_k - P_{1k} q_1
\end{aligned}
$$

donde $P_{1i} = \dfrac{\langle q_1, s_i \rangle}{\langle q_1, q_1 \rangle}$, $2 \leq i \leq k$

Por lo tanto, en el primer paso calculamos el primer renglón de R y $(k-1)$ vectores $s_2^{(1)}, s_3^{(1)}, \ldots, s_k^{(1)}$ tal que $s_i^{(1)} \perp q_1$ para $2 \leq i \leq k$.

En el segundo paso hacemos

$$q_2 = s_2^{(1)}$$

$$\begin{aligned} s_3^{(2)} &= s_3^{(1)} - \rho_{23}\, q_2 \\ s_4^{(2)} &= s_4^{(1)} - \rho_{24}\, q_2 \\ &\ \vdots \qquad\quad \vdots \\ s_k^{(2)} &= s_k^{(1)} - \rho_{2k}\, q_2 \end{aligned}$$

donde $\rho_{2i} = \dfrac{\langle q_2, s_i^{(1)} \rangle}{\langle q_2, q_2 \rangle}$, $3 \leq i \leq k$

Es decir, en el segundo paso calculamos el segundo renglón de R y $(k-2)$ vectores $s_3^{(2)}$, $s_4^{(2)}, \ldots, s_k^{(2)}$ tales que $s_i^{(2)} \perp q_2$ para $3 \leq i \leq k$.

Observemos que además $s_i^{(2)} \perp q_1$, $3 \leq i \leq k$, ya que como $q_1 \perp q_2$ y $s_i^{(1)} \perp q_1$, $3 \leq i \leq k$ $\Rightarrow$ $s_i^{(2)} \perp q_1$.

En el tercer paso hacemos

$$q_3 = s_3^{(2)}$$

$$\begin{aligned} s_4^{(3)} &= s_4^{(2)} - \rho_{34}\, q_3 \\ s_5^{(3)} &= s_5^{(2)} - \rho_{35}\, q_3 \\ &\vdots \\ s_k^{(3)} &= s_k^{(2)} - \rho_{3k}\, q_3 \end{aligned}$$

donde $\rho_{3j} = \dfrac{\langle q_3, s_j^{(2)} \rangle}{\langle q_3, q_3 \rangle}$, $4 \le j \le k$

Es decir, calculamos el tercer renglón de R y $(k-3)$ vectores $s_4^{(3)}, s_5^{(3)}, \ldots, s_k^{(3)}$ tales que

$$s_j^{(3)} \perp q_1, \quad s_j^{(3)} \perp q_2 \quad y \quad s_j^{(3)} \perp q_3, \quad 4 \le j \le k.$$

En general, en el i-ésimo paso, $1 \le i \le k$, hacemos

$$q_i = s_i^{(i-1)}$$

$$\begin{aligned} s_{i+1}^{(i)} &= s_{i+1}^{(i-1)} - \rho_{i,i+1}\, q_i \\ s_{i+2}^{(i)} &= s_{i+2}^{(i-1)} - \rho_{i,i+2}\, q_i \\ &\vdots \\ s_k^{(i)} &= s_k^{(i-1)} - \rho_{i,k}\, q_i \end{aligned}$$

donde $\rho_{ij} = \dfrac{\langle q_i, s_j^{(i-1)} \rangle}{\langle q_i, q_i \rangle}$ $\quad i+1 \le j \le k$

Es decir calculamos el i-ésimo renglón de R y $(k-i)$ vectores

$s_{i+1}^{(i)}, s_{i+2}^{(i)}, \dots, s_k^{(i)}$ tales que cada uno

de ellos es ortogonal a $q_1, q_2, \dots, q_i$

En el k-ésimo paso obtenemos que $\mathcal{S} = R(S) = R(Q)$ donde

$Q = [q_1 | q_2 | \cdots | q_k]$ tiene columnas ortogonales.

Este algoritmo se llama PROCEDIMIENTO DE GRAM-SCHMIDT MODIFICADO.

En la práctica se ha encontrado que numéricamente este método es más efectivo, y además tiene la propiedad de que en cada paso descompone a $\mathcal{S}$ en la suma de un subespacio y su complemento ortogonal en $\mathcal{S}$:

Después del i-ésimo paso tenemos que

$$\mathcal{S} = [q_1, q_2, \dots, q_i] \oplus [s_{i+1}^{(i)}, s_{i+2}^{(i)}, \dots, s_k^{(i)}]$$

$$\therefore \mathcal{S} = \mathcal{S}_i \oplus \tilde{\mathcal{S}}_i^{\perp}, \quad \text{donde} \quad \tilde{S}_i = S_i^{\perp} \cap \mathcal{S}$$

$$\mathcal{S}_i = [q_1, q_2, \dots, q_i].$$

▫

VI.68

TRANSFORMACIONES ORTOGONALES

Sea $\mathbb{R}^n$ con producto interior $\langle x,y\rangle = x^t y$. Nos interesa estudiar las transformaciones lineales $A : \mathbb{R}^n \to \mathbb{R}^n$ tales que

$$\| Ax \| = \| x \| \qquad \forall\, x \in \mathbb{R}^n$$

donde $\| x \| = \sqrt{\langle x, x\rangle}$.

Es decir, nos interesan las transformaciones lineales que preservan distancias y ángulos - y que por lo tanto no deforman los objetos geométricos.

Veamos qué propiedades tienen las matrices $A_{m\times n}$ tales que $\| Ax \| = \| x \| \quad \forall\, x \in \mathbb{R}^n$

$$\| \underset{m\times n}{A} x \| = \| x \| \iff (Ax)^t (Ax) = x^t x$$

$$\iff x^t (\underset{n\times m}{A^t} \underset{m\times n}{A})\, x = x^t x$$

$$\iff A^t A = I_n$$

Por lo tanto

$$\underset{\forall x \in \mathbb{R}^n}{\| \underset{m\times n}{A} x \| = \| x \|} \iff A^t A = I_n$$

Por lo tanto

$$I_n = \begin{bmatrix} a_1{}^t \\ a_2{}^t \\ \vdots \\ a_n{}^t \end{bmatrix} \begin{bmatrix} a_1 | a_2 | \cdots | a_n \end{bmatrix} = \begin{bmatrix} a_1{}^t a_1 & a_1{}^t a_2 & a_1{}^t a_n \\ a_2{}^t a_1 & a_2{}^t a_2 & a_2{}^t a_n \\ a_n{}^t a_1 & a_n{}^t a_2 & a_n{}^t a_n \end{bmatrix}$$

$$\Rightarrow \quad a_i{}^t a_j = 0 \quad \text{si } i \neq j$$

$$a_i{}^t a_i = 1$$

Hemos demostrado el siguiente

RESULTADO

Sea A una matriz $m \times n$ tal que $\|Ax\| = \|x\|$ $\forall x \in \mathbb{R}^n$, donde $\|x\|^2 = x^t x$. Entonces

1) las columnas de A son ortogonales
2) los renglones de A son ortogonales
3) los renglones y las columnas de A tienen norma = 1.

DEFINICION

Sea A una matriz de orden n. Decimos que A es ORTOGONAL si $A^t A = I_n$.

Por lo tanto las matrices $A: \mathbb{R}^n \to \mathbb{R}^n$ ortogonales son las que preservan distancias.

Ejemplos de transformaciones ortogonales :

1) Las rotaciones planas
2) Las reflexiones

PROPIEDADES

Sea $A_{n \times n}$ una matriz ortogonal . Entonces

1) A^t es ortogonal
2) A es invertible y $A^{-1} = A^t$
3) Si $B_{n \times n}$ es ortogonal , entonces AB es ortogonal .

Demostración :

(1),(2) $A^tA = I_n \Rightarrow A^{-1} = A^t$ y

$$A^tA = AA^t = (A^t)^t A^t = I_n \quad \therefore \quad A^t \text{ ortogonal}$$

(3) $(AB)^t (AB) = B^t (A^tA) B = I_n$.

DEFINICION

Sea $(\mathbb{V},<,>)$ un espacio euclideano. Decimos que $\{v_1,v_2,\dots,v_k\}\subseteq\mathbb{V}$ es ORTONORMAL si

(1) $\{v_1,v_2,\dots,v_k\}$ es ortogonal

(2) $\|v_i\|=1, \qquad 1\leq i\leq k$

RESULTADO

Sea $A=[a_1|a_2|\cdots|a_n] = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_n^t \end{bmatrix}$ de orden n.

Entonces

A es ortogonal $\Leftrightarrow$ $\{a_1,a_2,\dots,a_n\}$ es ortonormal
$\Leftrightarrow$ $\{r_1,r_2,\dots,r_n\}$ es ortonormal

▫

RESULTADO

Sea $\mathbb{R}^n$ con producto interno $\langle x,y\rangle = x^t y$.
Sea $Q=[q_1|q_2|\cdots|q_n]$ con columnas ortogonales.

Entonces $$\tilde{Q}=\left[\frac{q_1}{\|q_1\|}\;\middle|\;\frac{q_2}{\|q_2\|}\;\middle|\cdots\middle|\;\frac{q_n}{\|q_n\|}\right]=[\tilde{q}_1|\tilde{q}_2|\cdots|\tilde{q}_n]$$

es ortogonal.

Demostración:

$$\tilde{q}_i^t \tilde{q}_j = \frac{1}{\|q_i\|} \frac{1}{\|q_j\|} q_i^t q_j = 0 \quad \text{si } i \neq j$$

$$\tilde{q}_i^t q_i = \frac{1}{\|q_i\|^2} q_i^t q_i = \frac{1}{\|q_i\|^2} (\|q_i\|^2) = 1 .$$

□

Observemos que si Q y $\tilde{Q}$ son como en el resultado anterior, y

$$D = \begin{bmatrix} \|q_1\| & & & \\ & \|q_2\| & & \\ & & \ddots & \\ & & & \|q_n\| \end{bmatrix}$$

entonces $\tilde{Q} = Q D^{-1}$

REFLEXIONES ORTOGONALES

Las reflexiones ortogonales forman una clase importante de matrices ortogonales, y las estudiaremos con más detalle.

Sea $\mathbb{R}^n$ con producto interior $\langle x,y\rangle = x^t y$.
Sea $\mathcal{S} \subseteq \mathbb{R}^n$ un subespacio de $\mathbb{R}^n$

$$\therefore \quad \mathbb{R}^n = \mathcal{S} \oplus \mathcal{S}^{\perp}$$

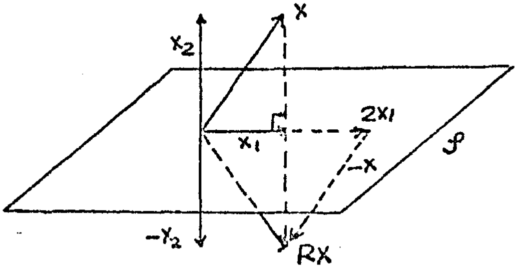
$\therefore$ si $x \in \mathbb{R}^n \Rightarrow x = x_1 + x_2$ con $x_1 \in \mathcal{S}$, $x_2 \in \mathcal{S}^{\perp}$

$$\therefore \quad x = P_{\mathcal{S}}(x) + P_{\mathcal{S}^{\perp}}(x)$$

Por lo tanto la reflexión ortogonal sobre $\mathcal{S}$ está dada por

$$R_{\mathcal{S}}(x) = P_{\mathcal{S}}(x) - P_{\mathcal{S}^{\perp}}(x) = P_{\mathcal{S}}(x) - (x - P_{\mathcal{S}}(x))$$

$$\therefore \quad R_{\mathcal{S}}(x) = (2P_{\mathcal{S}} - I)\,x$$



- Cuando $\dim(\mathcal{S}) = 1$ decimos que $R_{\mathcal{S}}$ es una REFLEXION ELEMENTAL

Veamos qué forma tiene una reflexión elemental.

Supongamos $\mathcal{S} = [z]$

Entonces la proyección ortogonal sobre $\mathcal{S}$ está dada por

$$P_{[z]} = \frac{zz^t}{z^tz}$$

y por lo tanto $R_{[z]} = 2\frac{zz^t}{z^tz} - I$.

En los cursos de análisis numérico se ve que las reflexiones son muy estables numéricamente.

USO DE LAS TRANSFORMACIONES ORTOGONALES EN EL PROBLEMA DE CUADRADOS MINIMOS

Volvamos a considerar el problema de:

Dado $y \in \mathbb{R}^n$ y $\mathcal{S} \subseteq \mathbb{R}^n$, encontrar $\tilde{y} \in \mathcal{S}$ tal que

$$\|y-\tilde{y}\|^2 = \min_{s \in \mathcal{S}} \|y-s\|^2 .$$

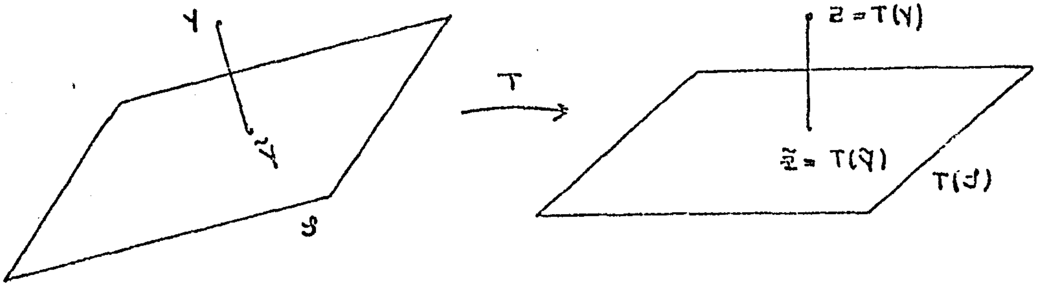
La solución de éste la conseguimos construyendo una base ortogonal $\{q_1, q_2, \ldots, q_k\}$ de $\mathcal{S}$ a partir de la base $\{s_1, s_2, \ldots, s_k\}$ de $\mathcal{S}$, y obteniendo la factorización $S = QR$ de la matriz S, donde

$$S = [s_1 | s_2 | \cdots | s_k]$$
$$Q = [q_1 | q_2 | \cdots | q_k] .$$

Ahora que conocemos las transformaciones ortogonales, intuimos que éstas pueden sernos de utilidad para resolver el problema de cuadrados mínimos.

Observemos que si $T : \mathbb{R}^n \to \mathbb{R}^n$ es una transformación ortogonal entonces el problema de encontrar el punto $\tilde{y} \in \mathcal{S}$, que sea el más cercano a $y \in \mathbb{R}^n$ de todos los vectores en $\mathcal{S}$, se convierte en encontrar el punto $\tilde{z} = T(\tilde{y}) \in T(\mathcal{S})$ que sea el más cercano a $z = T(y)$;

ya que $\| y - \tilde{y} \| = \| T(y - \tilde{y}) \| = \| T(y) - T(\tilde{y}) \| = \| z - \tilde{z} \|$



Encontrar el vector $\tilde{z} \in T(\mathcal{S})$ tal que

$$\|y-\tilde{y}\| = \|z-\tilde{z}\|$$

es especialmente fácil si la transformación T es tal que $T(\mathcal{S})$ es uno de los "subespacios coordenados", es decir, si

$$T(\mathcal{S}) = \left\{ s' \in \mathbb{R}^n : \; s' = \begin{bmatrix} z' \\ 0 \end{bmatrix} \begin{matrix} k \\ n-k \end{matrix} \right\} \qquad (*)$$

Hallaremos una matriz T con esta propiedad.

Supongamos que $\{\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_k\}$ es una base ortonormal de $\mathcal{S}$, entonces la podemos completar a una base ortonormal de $\mathbb{R}^n$:

$$\{\tilde{q}_1, \tilde{q}_2, \ldots, \tilde{q}_k, \tilde{q}_{k+1}, \ldots, \tilde{q}_n\}.$$

Para que $T(\mathcal{S})$ sea como en $(*)$ necesitamos que

$$T\tilde{q}_1 = e_1 \; , \; T\tilde{q}_2 = e_2, \ldots, \; T\tilde{q}_k = e_k$$

ya que si $T\tilde{q}_i = e_i$ $\quad 1 \leq i \leq k$ entonces,

como $\mathcal{S} = R(S) = R([\tilde{q}_1 | \tilde{q}_2 | \cdots | \tilde{q}_k]) \Rightarrow$

$$T(\mathcal{S}) = \mathcal{R}\left(T\,[\tilde{q}_1|\tilde{q}_2|\cdots|\tilde{q}_k]\right) = [e_1, e_2, \ldots, e_k] .$$

Por lo tanto, sea $T : \mathbb{R}^n \to \mathbb{R}^n$ dado por

$$T\tilde{q}_i = e_i \quad , \; 1 \leq i \leq n$$

Sea $\tilde{Q} = [\tilde{q}_1|\tilde{q}_2|\cdots|\tilde{q}_n]$

Por lo tanto tenemos que

$$T\tilde{Q} = T\,[\tilde{q}_1|\tilde{q}_2|\cdots|\tilde{q}_n] = [e_1|e_2|\cdots|e_n] = I_n$$

$$\therefore \quad T\tilde{Q} = I_n$$

$\therefore \; T = \tilde{Q}^t$ ya que, como las columnas de $\tilde{Q}$ son ortonormales $\Rightarrow$ $\tilde{Q}$ es ortogonal.

Sea
$$z = \tilde{Q}^t y = \begin{array}{c} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \end{array}\begin{array}{l} k \\ n-k \end{array}$$

Por lo tanto

$$\min_{x \in \mathbb{R}^k} \| y - Sx \|^2 = \min_{x \in \mathbb{R}^k} \| \tilde{Q}^t (y - Sx) \|^2$$

$$= \min_{x \in \mathbb{R}^k} \| \tilde{Q}^t y - \tilde{Q}^t S x \|^2$$

$$= \min_{x \in \mathbb{R}^k} \| z - \tilde{Q}^t S x \|^2$$

Pero $R(\tilde{Q}^t S) = [e_1, e_2, \dots, e_k]$

$$= \left\{ s' \in \mathbb{R}^n : \; s' = \left[\begin{array}{c} z' \\ 0 \end{array}\right]\begin{array}{l} k \\ n-k \end{array}, \; z' \in \mathbb{R}^k \right\}$$

$$\Rightarrow \quad \tilde{Q}^t S = \overset{k}{\left[\begin{array}{c} N \\ \hline O \end{array}\right]}\begin{array}{l} k \\ n-k \end{array} \quad \text{para alguna } N_{k \times k} \text{ no-singular.}$$

(Esto es consecuencia de que si r_i^t es el i-ésimo renglón de $\tilde{Q}^t S$, entonces $i \geq k+1 \Rightarrow r_i^t x = 0 \quad \forall\, x \in \mathbb{R}^k \Rightarrow r_i^t = 0^t$).

Por lo tanto

$$\min_{x \in \mathbb{R}^k} \| y - Sx \|^2 = \min_{x \in \mathbb{R}^k} \left\| \begin{array}{c} k \\ n-k \end{array}\left[\begin{array}{c} z_1 \\ z_2 \end{array}\right] - \left[\begin{array}{c} N \\ 0 \end{array}\right] x \right\|^2$$

$$= \min_{x \in \mathbb{R}^k} \left\| \begin{array}{c} k \\ n-k \end{array}\left[\begin{array}{c} z_1 - Nx \\ z_2 \end{array}\right] \right\|^2$$

$$= \min_{x \in \mathbb{R}^k} \left(\| z_1 - Nx \|^2 + \| z_2 \|^2 \right)$$

Concluimos que si para alguna $x \in \mathbb{R}^k$ tenemos que $z_1 - Nx = 0$, entonces esa x nos da el valor mínimo.

Como N es no-singular, entonces el sistema

$$N x = z_1$$

siempre tiene solución, y por lo tanto

$$\tilde{y} = Sx \quad , \quad \text{donde} \quad x = N^{-1} z_1 .$$

Por lo tanto, para poder resolver nuestro problema necesitamos una matriz ortogonal $\tilde{Q}$ tal que

$$\tilde{Q}^t S = \begin{array}{c} k \\ \left[\begin{array}{c} N \\ \hline 0 \end{array}\right] \end{array} \begin{array}{l} k \\ n-k \end{array} \quad \text{con } N_{k \times k} \text{ no-singular.}$$

o sea que necesitamos encontrar una factorización de S:

$$\underset{n \times k}{S} = \underset{n \times n}{\tilde{Q}} \begin{array}{c} k \\ \left[\begin{array}{c} N \\ 0 \end{array}\right] \end{array} \begin{array}{l} k \\ n-k \end{array}$$

con $\tilde{Q}$ ortogonal y $N_{k \times k}$ no-singular.

Recordemos que ya tenemos la descomposición

$$\underset{n \times k}{S} = \underset{n \times k}{Q} R_{k \times k}$$

donde R es triangular superior no-singular y

$Q_{n\times k}$ tiene columnas ortogonales

Podemos pasar de esta descomposición a la otra completando a una base de $\mathbb{R}^n$ y después normalizando ; y entonces obtenemos

$$S = \begin{array}{c} \\ n \end{array}\!\!\begin{array}{c|c} k & n-k \\ \Big[\, Q D_1^{-1} & Q_2 D_2^{-1} \,\Big] \end{array} \begin{bmatrix} D_1 R \\ \hline O \end{bmatrix}\begin{array}{l} k \\ n-k \end{array}$$

$$= \underset{n\times n}{\tilde{Q}} \overset{k}{\begin{bmatrix} N \\ \hline O \end{bmatrix}}\begin{array}{l} k \\ n-k \end{array}$$

donde $Q_2 = [q_{k+1} | q_{k+2} | \cdots | q_n]$

$$S = QR = [q_1 | q_2 | \cdots | q_k] R$$

con $\{q_1, q_2, \dots, q_k, q_{k+1}, \dots, q_n\}$ una base ortogonal de $\mathbb{R}^n$ y

$$D_1 = \begin{bmatrix} \|q_1\| & & & \\ & \|q_2\| & & \\ & & \ddots & \\ & & & \|q_k\| \end{bmatrix}$$

$$D_2 = \begin{bmatrix} \|q_{k+1}\| & & & \\ & \|q_{k+2}\| & & \\ & & \ddots & \\ & & & \|q_n\| \end{bmatrix}$$

O sea que no solamente podemos lograr que N sea no-singular, sino que también sea triangular superior - lo cual hace que la solución del sistema $Nx = z_1$ sea muy fácil de calcular.

Se podrá pensar que no se ganó nada, salvo complicarse la vida, ya que este procedimiento representa más trabajo que el que se efectúa con el procedimiento Gram-Schmidt.

Esto es cierto, pero lo que nos interesa es la solución del problema de la proyección ortogonal teniendo la factorización $S = \tilde{Q}\tilde{R}$, que resumimos a continuación, ya que después procederemos a calcular $\tilde{Q}$ y $\tilde{R}$ de manera distinta.

PROBLEMA

Sea $\mathbb{R}^n$ con producto interior $\langle x,z\rangle = x^t z$. Dado $y \in \mathbb{R}^n$ y $\mathcal{S} \in \mathbb{R}^n$ un subespacio encontrar $\tilde{y} \in \mathcal{S}$ tal que

$$\| y - \tilde{y} \|^2 = \min_{s \in \mathcal{S}} \| y - s \|^2 = \min_{x \in \mathbb{R}^k} \| y - Sx \|$$

donde $S = [s_1 | s_2 | \ldots | s_k]$, $\mathcal{S} = R(S)$ y $\dim(\mathcal{S}) = k$.

SOLUCION

1) Calcular la descomposición

$$S = \tilde{Q}\tilde{R}$$

donde $\tilde{Q}_{n \times n}$ es ortogonal y

$$\tilde{R}_{n \times k} = \begin{array}{cc} k & \\ \left[\begin{array}{c} N \\ \hline 0 \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \end{array}$$

con N triangular superior no-singular

2) Sea $$\tilde{Q}^t y = z = \left[\begin{array}{c} z_1 \\ \hline z_2 \end{array}\right] \begin{array}{l} k \\ n-k \end{array}$$

3) Resolver $N x = z_1$, es decir sea $x = N^{-1} z_1$

4) $\hat{y} = S x = S N^{-1} z_1$

■

Daremos ahora un método muy eficiente para calcular la descomposición

$$S = \tilde{Q}\tilde{R}$$

de una matriz S con columnas independientes.

ALGORITMO PARA CALCULAR LA FACTORIZACION $S=\tilde{Q}\tilde{R}$

Sea $S_{n\times k}$ con columnas independientes. Entonces si existen $\tilde{Q}$ y $\tilde{R}$ tal que

$$S = \tilde{Q}\tilde{R}$$

con $\tilde{Q}$ ortogonal de orden n y $\tilde{R}$ $n\times k$ de la forma

$$\tilde{R} = \begin{array}{cl} k & \\ \left[\begin{array}{c} N \\ \hline 0 \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \end{array} \quad ; \quad N = [\nabla] \text{ no-singular}$$

entonces $\tilde{Q}^t S = \tilde{R}$.

Como observamos anteriormente, las reflexiones ortogonales son muy estables numéricamente.- trataremos de expresar a la matriz $\tilde{Q}^t$ como producto de reflexiones elementales:

$$\tilde{Q}^t = R_k R_{k-1} \cdots R_2 R_1$$

Supongamos $S = [s_1|s_2|\cdots|s_k] = [\sigma_{ij}]$.

En el primer paso queremos encontrar una reflexión elemental R_1 tal que

$$R_1 S = \begin{bmatrix} \circledast & * & \cdots & * \\ 0 & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{bmatrix} = \left[\begin{array}{c|ccc} \circledast & * & \cdots & * \\ \hline & & & \\ 0 & & S_1 & \\ & & & \end{array}\right]$$

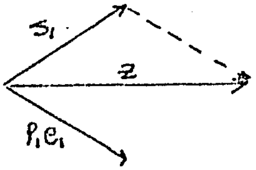
Como $R_1 S = [R_1 s_1 \mid R_1 s_2 \mid \cdots \mid R_1 s_k]$

necesitamos que $R_1 s_1 = \rho_1 e_1$ donde

$$\rho_1 = \begin{cases} \|s_1\| & \text{si } \sigma_{11} > 0 \\ -\|s_1\| & \text{si } \sigma_{11} < 0 \end{cases}.$$

$(R_1 s_1 = (\pm\|s_1\|) e_1 \Rightarrow \|R s_1\| = \|s_1\|)$.

Sea $z_1 = s_1 + \rho_1 e_1$



Veremos que $R_1 = R_{[z]}$ es tal que $R_1 s_1 = \rho_1 e_1$:

$$z_1^t z_1 = (s_1 + \rho_1 e_1)^t (s_1 + \rho_1 e_1)$$

$$= s_1^t s_1 + 2\rho_1 e_1^t s_1 + \rho_1^2$$

$$= s_1^t s_1 + 2\rho_1 e_1^t s_1 + \|s_1\|^2$$

$$= 2 s_1^t s_1 + 2\rho_1 e_1^t s_1$$

$$= 2(s_1^t s_1 + \rho_1 e_1^t s_1).$$

Por lo tanto

$$R_1(s_1) = R_{[z_1]}(s_1) = 2\left(\frac{z_1 z_1^t}{z_1^t z_1} - I\right) s_1$$

$$= 2z_1 \frac{z_1^t s_1}{z_1^t z_1} - s_1$$

$$= 2z_1 \frac{(s_1 + \rho e_1)^t s_1}{z_1^t z_1} - s_1$$

$$= 2z_1 \frac{(s_1^t s_1 + \rho e_1^t s_1)}{2(s_1^t s_1 + \rho e_1^t s_1)} - s_1$$

$$= z_1 - s_1 = s_1 + \rho_1 e_1 - s_1$$

$\therefore$ $R_1(s_1) = R_{[z_1]}$

y por lo tanto

$$R_1 S = R_{[2,1]} S = \begin{array}{cc} \begin{array}{cc} 1 & k-1 \end{array} & \\ \left[\begin{array}{c|c} \otimes & * \cdots * \\ \hline 0 & \bar{S}_1 \end{array}\right] & \begin{array}{c} 1 \\ n-1 \end{array} \end{array} = S_1$$

Observemos que $$z = \begin{bmatrix} \sigma_{11} \pm \|s_1\| \\ \sigma_{21} \\ \vdots \\ \sigma_{n1} \end{bmatrix}$$

es decir, z es "casi igual" a la primera columna de S -- solamente difieren en la primera componente.

Además, basta con conocer el vector z para determinar R_1.

En el segundo paso aplicamos este mismo procedimiento a $\bar{S}_1$ para encontrar $\bar{R}_2$ $(n-1) \times (n-1)$ tal que

$$\bar{R}_2 \bar{S}_1 = \begin{array}{cc} & \begin{array}{cc} 1 & k-1 \end{array} \\ \begin{array}{c} 1 \\ n-2 \end{array} & \left[\begin{array}{c|c} \otimes & * \cdots * \\ \hline 0 & \bar{S}_2 \end{array}\right] \end{array}$$

$$\therefore \quad S_1 \quad R_2 = \begin{array}{c|c|c} 1 & n-1 & \\ \hline I_1 & 0 & 1 \\ \hline 0 & \bar{R}_2 & n-1 \end{array}$$

$$\text{Entonces} \qquad R_2 S_1 = \begin{array}{c|c|c} & 1 & k-1 \\ \hline 1 & \circledast & * * \cdots * \\ \hline n-1 & 0 & \bar{R}_2 \bar{S}_1 \end{array}$$

$$= \begin{array}{c|cc|c} & \multicolumn{2}{c|}{2} & k-2 \\ \hline 2 & \circledast & * & * \cdots * \\ & & \circledast & * \cdots * \\ \hline n-2 & \multicolumn{2}{c|}{0} & \bar{S}_2 \end{array} = S_2$$

Observar que como $\bar{S}_1$ es de dimensión menor que S, entonces calcular $\bar{R}_2$, y por lo tanto R_2 es más simple que calcular la R_1

En el tercer paso aplicamos el procedimiento a $\bar{S}_2$, para obtener una $\bar{R}_3$ tal que

$$R_3 S_2 = \begin{array}{c|c|c} & 2 & n-2 \\ \hline 2 & I_2 & \\ \hline n-2 & & \bar{R}_3 \end{array} S_2 \quad = \begin{array}{c|c|c} & 3 & k-3 \\ \hline 3 & \diagdown & \Box \\ \hline n-3 & 0 & \bar{S}_3 \end{array} = S_3$$

Podemos continuar este proceso, hasta obtener

$$R_k \cdots R_2 R_1 S = \tilde{R} = \begin{array}{cc} k & \\ \left[\begin{array}{c} \diagdown \\ \hline 0 \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \end{array}$$

Sea $\tilde{Q}^t = R_k \cdots R_2 R_1$

$\therefore \quad S = \tilde{Q} \tilde{R} = \widehat{R}_1^t \widetilde{R}_2^t \cdots \widetilde{R}_k^t$.

BIBLIOGRAFIA

1. G. Birkhoff and S. MacLane
   A SURVEY OF MODERN ALGEBRA
   Macmillan, New York 1977

2. C.G. Broyden
   BASIC MATRICES
   Macmillan, London 1975

3. G.W. Stewart
   INTRODUCTION TO MATRIX COMPUTATIONS
   Academic Press, New York 1973

4. G. Strang
   LINEAR ALGEBRA AND ITS APPLICATIONS
   Academic Press, New York 1980

5. V.V. Voevodin
   ALGEBRA LINEAL
   Editorial Mir, Moscú 1982

6. L.I. Golovina
   ALGEBRA LINEAL Y ALGUNAS DE SUS APLICACIONES
   Editorial Mir, Moscú 1980