#91

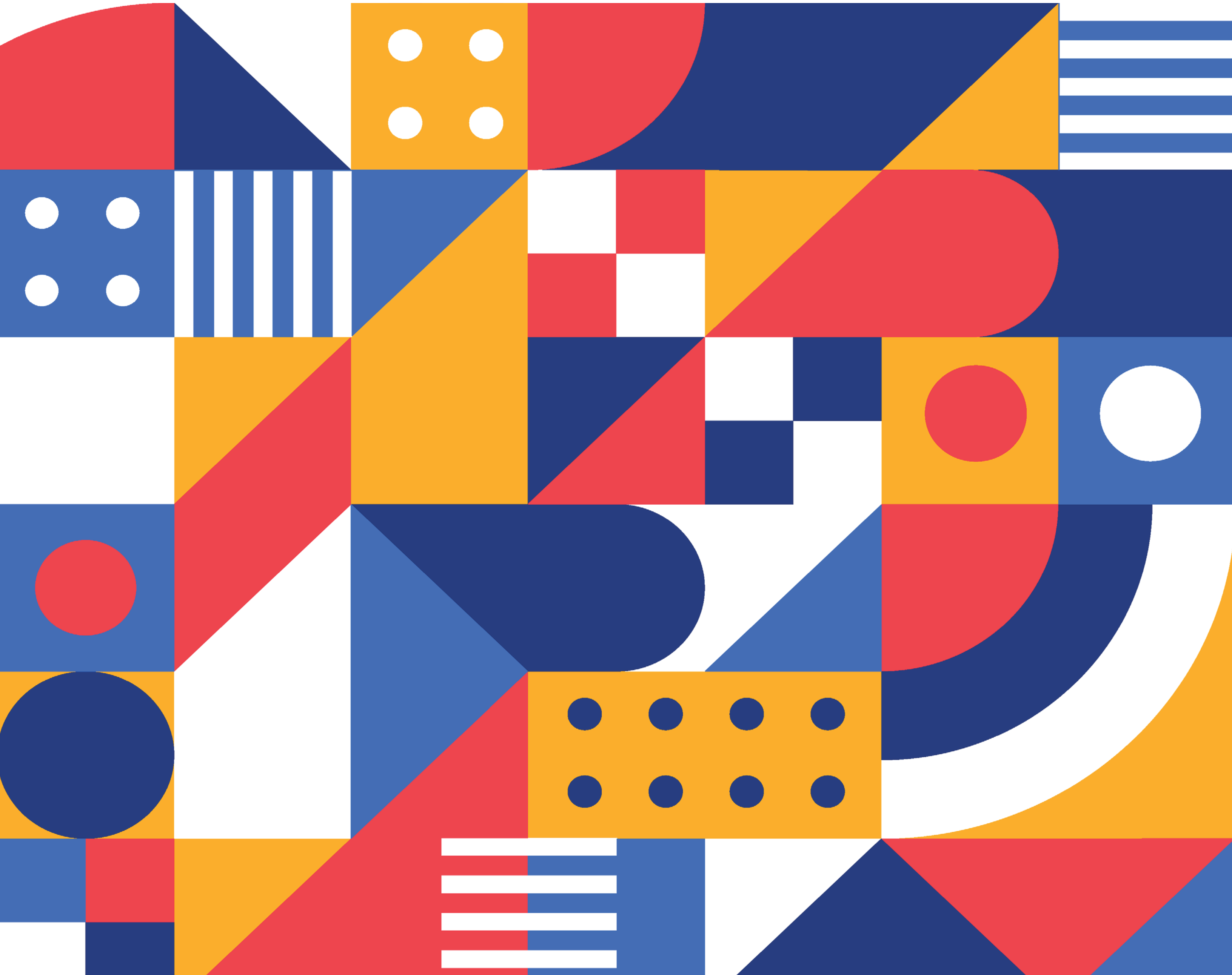
SERIE
MATEMÁTICAS

Pablo Barrera Sánchez y Franco Toledo de la Cruz

Álgebra lineal. Capitulo I

VÍNCULOS MATEMÁTICOS

AÑO
2007



Facultad de Ciencias

Vínculos matemáticos

Álgebra lineal, capítulo I

Pablo Barrera Sánchez*

Franco Toledo de la Cruz**

Departamento de Matemáticas

Facultad de Ciencias, UNAM

SEGUNDA EDICIÓN

Nº 91. 2011

* Profesores del Departamento de Matemáticas de la Facultad de Ciencias de la UNAM

** Profesor del Instituto de Astronomía de la UNAM

Impreso en la Coordinación de Servicios Editoriales de la Facultad de Ciencias, UNAM

CAPITULO I DEL TEXTO

DE

ALGEBRA LINEAL

POR:

PABLO BARRERA SÁNCHEZ.

FRANCO TOLEDO DE LA CRUZ.

FACULTAD DE CIENCIAS.

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO.

PROLOGO

El contenido del texto que presentamos es resultado de un curso impartido durante el segundo semestre de 1982. La recopilación fue hecha por la alumna *Elsa Campuzano V.*, quién posteriormente presentó el material como tema para obtener el título de Matemático.

Gracias al impulso decidido de numerosas personas y amigos en el Departamento de Matemáticas de la Facultad de Ciencias, especialmente del profesor *F. Javier Aurrecoechea F.*, hemos decidido hacer la revisión de esas notas, corregirlas y agregar los temas restantes para editar un texto de *Algebra Lineal*, de acuerdo a los programas de la licenciatura en la Facultad de Ciencias. En esta labor hemos contado con el apoyo invaluable de la alumna *Cynthia K. González G.*, quien tuvo a su cargo el trabajo de capturar y corregir las notas conforme se modificaban.

Queremos manifestar nuestro agradecimiento al Departamento de Matemáticas de la Facultad de Ciencias y al Instituto de Astronomía por el apoyo que nos han brindado.

Atentamente.

Dr. Pablo Barrera Sánchez.
Facultad de Ciencias.
U N A M

Fís. Franco Toledo de la Cruz
Instituto de Astronomía.
U N A M.

México D. F., octubre de 1991.

CAPITULO I

VECTORES Y MATRICES

INTRODUCCION.

En este capítulo introduciremos los conceptos de *vector* y de *matriz*. Gradualmente se irán definiendo las principales operaciones que se pueden realizar entre: vector-vector, vector-matriz y matriz-matriz. La presentación se hace de manera que la operación en una etapa, se generaliza en la etapa siguiente. Esta forma de presentar el material nos permitirá dar una interpretación dinámica de las operaciones mencionadas, en términos de acciones que se realizan sobre los componentes elementales.

Un aspecto de gran utilidad en el desarrollo es el de la notación que emplearemos y que especificamos a continuación:

Primero. Los escalares se denotarán con letras griegas minúsculas: $\alpha,\beta,\gamma,\delta,\mu,\varphi,\eta,\lambda,\ldots,\xi,\vartheta,\zeta$.

Segundo. Los vectores se denotarán con letras latinas minúsculas: $a,b,c,d,\ldots,x,y,z$. Sus componentes, por ser escalares, con griegas minúsculas y un subíndice: ξ_i denotará la i-ésima componente de un vector.

Tercero. Las matrices las denotaremos con letras latinas mayúsculas: $A,B,C,D,\ldots,X,Y,Z$. Sus elementos con griegas minúsculas y dos subíndices: ξ_{ij} indica un escalar elemento de una matriz.

Los vectores de partida sobre los cuales se basará el desarrollo de la teoría serán los vectores columna cuyas propiedades estudiaremos a continuación.

Vectores Columna y Transposicion de Vectores.

DEFINICION.

$$\text{Sea } \mathbb{R}^n = \left\{ \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} \;/\; \xi_i \in \mathbb{R};\ i = 1,\dots,n \right\}$$

entonces:

i) A los elementos de $\mathbb{R}^n$ los llamaremos *vectores columna de orden o dimensión* n.

ii) Los números $\xi_1,\dots,\xi_n$ se llamarán los *elementos o componentes* del vector.

En el desarrollo, cuando mencionemos simplemente un vector de órden n, estaremos refiriéndonos precisamente a un vector columna de órden n, y si n es fijo o se puede determinar del contexto, emplearemos únicamente la palabra vector.

El uso del término "*dimensión*", es por analogía con la geometría. Un vector de órden n se puede ver como un punto en el espacio n-dimensional.

Por ejemplo el vector:

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

quedaría representado así:

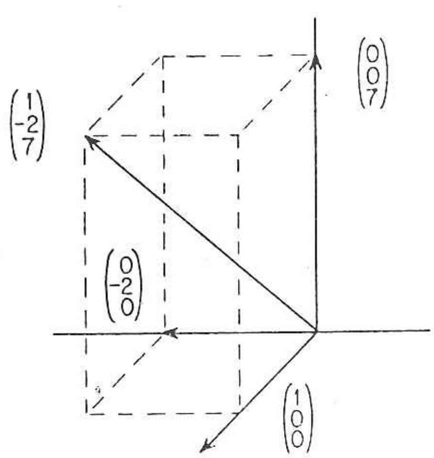


Figura 1

Notación:.

Para denotar un vector cuya i-ésima componente queremos especificar que es ξ_i, escribiremos $[\xi_i]$, es decir:

$$[\xi_i] = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

La estructura lineal de $\mathbb{R}^n$, viene dada por las operaciones básicas: *suma de dos vectores* y la *multiplicación de un vector por un escalar*. A continuación se definirán estas operaciones y se estudiarán sus propiedades.

SUMA DE DOS VECTORES.

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 $[\gamma_i] \in \mathbb{R}^n$ 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.

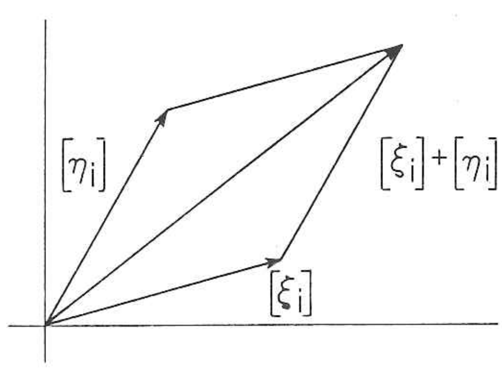


Figura 2

MULTIPLICACION DE UN VECTOR POR UN ESCALAR.

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 $[\beta_i] \in \mathbb{R}^n$ 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 $[\xi_i]$.

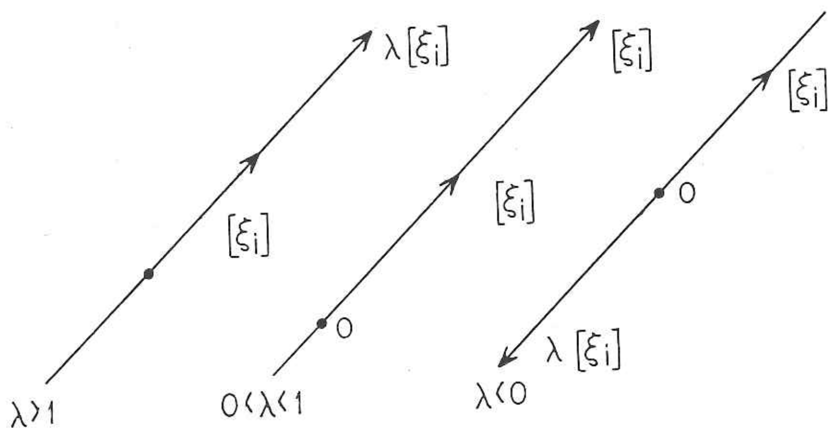


Figura 3

DEFINICION.

i) 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, sabremos si se trata del escalar cero o del vector cero.

ii) Dos vectores de órden n son iguales si y sólo si sus componentes correspondientes son iguales.

PROPIEDADES DE LOS VECTORES COLUMNA.

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

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. $(1)[\xi_i] = [\xi_i]$
6. $\lambda(\mu[\xi_i]) = (\lambda\mu)[\xi_i]$
7. $(\lambda + \mu)[\xi_i] = \lambda[\xi_i] + \mu[\xi_i]$
8. $\lambda([\xi_i] + [\eta_i]) = \lambda[\xi_i] + \lambda[\eta_i]$

Además de $\mathbb{R}^n$, otro conjunto elemental útil, que es necesario definir, es el conjunto de vectores renglón al que denotaremos por $\mathbb{R}^{n^*}$ debido a su relación directa con aquel.

DEFINICION.

$$\text{Sea } \mathbb{R}^{n^*} = \{\ [\xi_1\ \xi_2\ \cdots\ \xi_n]/\xi_i \in \mathbb{R};\ i = 1,2,\ldots,n\ \}$$

i) A los elementos de $\mathbb{R}^{n^*}$ los llamaremos *vectores renglón,de orden o dimensión* n.

ii) Los números $\xi_1,\ldots,\xi_n$ se llamarán *los elementos o componentes* del vector renglón.

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

Suma de dos Vectores Renglon.

DEFINICION.

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

La *suma* de los vectores $[\xi_1\ \xi_2 \cdots \xi_n]$ y $[\eta_1\ \eta_2 \cdots \eta_n]$ es el vector renglón $[\gamma_1\ \gamma_2 \cdots \gamma_n]$, denotado por $[\xi_1\ \xi_2 \cdots \xi_n] + [\eta_1\ \eta_2 \cdots \eta_n]$, dado por:

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

Multiplicacion de un Vector Renglon por un Escalar.

DEFINICION.

El *producto* de λ y $[\xi_1\ \xi_2 \cdots \xi_n]$ es el vector renglón $[\beta_1\ \beta_2 \cdots \beta_n]$, denotado por $\lambda[\xi_1\ \xi_2 \cdots \xi_n]$, dado por:

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

Propiedades de los Vectores Renglon.

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

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

2. $(\ [\xi_1\ \xi_2 \cdots \xi_n] + [\eta_1\ \eta_2 \cdots \eta_n]\ ) + [\zeta_1\ \zeta_2 \cdots \zeta_n] = [\xi_1\ \xi_2 \cdots \xi_n] +$
   $+ (\ [\eta_1\ \eta_2 \cdots \eta_n] + [\zeta_1\ \zeta_2 \cdots \zeta_n]\ )$

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

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

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. $(1)[\xi_1\ \xi_2 \ldots \xi_n] = [\xi_1\ \xi_2 \ldots \xi_n]$

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

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

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

Como se puede observar, existe una relación evidente entre $\mathbb{R}^n$ y $\mathbb{R}^{n^*}$, que es necesario precisar y estudiar con detenimiento. A continuación la establecemos mediante la operación de transposición.

TRANSPOSICION DE VECTORES.

DEFINICION.

Sea $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^{n^*}$

$$x \longmapsto x^t$$

tal que si $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$, entonces $x^t = [\xi_1 \ldots \xi_n]$.

A φ por tanto, la podemos ver como la función que transforma a la *columna* x en el *renglón* x^t, se lee "equis transpuesto". Es decir, $x^t = \varphi(x)$.

Es facil demostrar que φ es una función uno a uno y sobre, ya que:

1. Si $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$, $y = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \in \mathbb{R}^n$, entonces:

$$\varphi(x) = \varphi(y) \Rightarrow [\xi_1 \ldots \xi_n] = [\eta_1 \ldots \eta_n]$$

$$\Rightarrow \begin{matrix} \xi_1 = \eta_1 \\ \vdots \\ \xi_n = \eta_n \end{matrix} \quad \Rightarrow \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix} \Rightarrow x = y$$

$\therefore$ φ es inyectiva.

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

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

$\therefore$ φ es sobre.

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

$$\varphi^{-1}\left([\xi_1 \ldots \xi_n] \right) = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

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

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:

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

Entonces tendremos que:

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

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

PROPIEDADES DE LA TRANSPOSICION.

1. $(x^t)^t = x \qquad \forall\ x \in \mathbb{R}^n.$

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

Demostración:

$$x + y = \begin{bmatrix} \xi_1 + \eta_1 \\ \vdots \\ \xi_n + \eta_n \end{bmatrix}$$

$$\therefore\ (x + y)^t = [\xi_1 + \eta_1 \ldots\ \xi_n + \eta_n]$$

$$= [\xi_1 \ldots \xi_n] + [\eta_1 \ldots \eta_n]$$

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

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

Demostración:

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

$$\therefore\ (\alpha x)^t = [\alpha\xi_1 \ldots \alpha\xi_n]$$

$$= \alpha[\xi_1 \ldots \xi_n]$$

$$= \alpha x^t \ \blacksquare$$

Las operaciones definidas hasta ahora, tanto en $\mathbb{R}^n$ como en $\mathbb{R}^{n^*}$, permiten hacer un estudio de estos conjuntos como estructuras lineales que, como veremos más adelante, en particular nos servirán para definir planos y líneas rectas como subconjuntos que preservan las propiedades de aquellos. Existen otros dos conceptos muy importantes en relación con la interpretación geométrica de los vectores, los cuales nos son bastante familiares que dificilmente pensamos en los vectores sin tomarlos en cuenta, aunque no son relaciones tan sencillas como las lineales. Estos son los conceptos de norma de un vector y ángulo entre dos vectores, que a pesar de su naturaleza tan diferente, pueden definirse en términos de un concepto más básico (aunque no tan natural) que es el producto punto o producto escalar.

PRODUCTO PUNTO O PRODUCTO ESCALAR.

DEFINICION.

Sean $x, y \in \mathbb{R}^n$. El *producto punto o producto escalar* de x y y es el escalar denotado por $x*y$ dado por:

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

Observar que el producto punto de dos vectores es un número real.

Nota:

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

PROPIEDADES DEL PRODUCTO PUNTO.

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

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

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

   $x\cdot x = 0 \Leftrightarrow x = 0$

Demostración:

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

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

   $\therefore\ (\lambda x)\cdot y = \lambda(x\cdot y) = x\cdot(\lambda y)$ ■

3. Si $x_1 = [\xi_{i1}]$, $\quad x_2 = [\xi_{i2}]$; entonces

$$(x_1 + x_2)\cdot y = \sum_{i=1}^{n} (\xi_{i1} + \xi_{i2})\eta_i = \sum_{i=1}^{n} (\xi_{i1}\eta_i + \xi_{i2}\eta_i)$$

$$= \sum_{i=1}^{n} \xi_{i1}\eta_i + \sum_{i=1}^{n} \xi_{i2}\eta_i$$

$$= x_1\cdot y + x_2\cdot y$$ ■

4. $x\cdot(y_1 + y_2) = (y_1 + y_2)\cdot x$ $\quad$ (por 1)

   $= (y_1 x) + (y_2 x)$ $\quad$ (por 3)

   $= x\cdot y_1 + x\cdot y_2$ $\quad$ (por 1) ■

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

   como $\xi_i^2 \geq 0$ $\quad$ con $i = 1,2,\ldots,n$

   $\Rightarrow \xi_1^2 + \xi_2^2 + \ldots + \xi_n^2 \geq 0$

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

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

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

   $x = 0$ ■

Ejemplos

1. Si $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\Rightarrow$ $x \cdot y = (1)(1) + (0)(1) = 1$ ■

Observar, en el siguiente ejemplo, que:

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

2. Si $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $y = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ $\Rightarrow$ $x \cdot y = (1)(0) + (0)(1) = 0$ ■

Como se mencionó anteriormente, a partir del producto podemos definir la norma de los vectores. A continuación veremos como la definición que se obtiene refleja fielmente las técnicas y resultados que utilizamos cuando medimos segmentos lineales con una regla graduada.

DEFINICION DE NORMA EUCLIDIANA.

Sea $x \in \mathbb{R}^n$. Definimos la *norma Euclidiana* ó *longitud* de x, denotada por $\| x \|$, como:

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

PROPIEDADES DE NORMA.

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

Demostración::

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

$$\therefore \| \alpha x \| = | \alpha | \sqrt{\sum_{i=1}^{n} \xi_i^2} = | \alpha | \| x \|$$ ■

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

Demostración:

$$\| x - y \|^2 = \sum_{l=1}^{n} (\xi_l - \eta_l)^2 = \sum_{l=1}^{n} \xi_l^2 - \sum_{l=1}^{n} 2\xi_l\eta_l + \sum_{l=1}^{n} \eta_l^2$$

$$= \| x \|^2 + \| y \|^2 - 2\sum_{l=1}^{n} \xi_l\eta_l$$

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

3. *Teorema de Pitágoras*

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

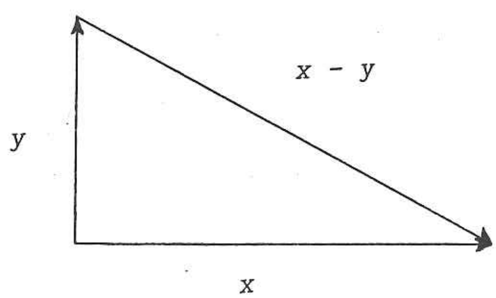


Figura 4

Demostración:

Por el resultado (2)

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

De manera análoga se puede ver que:

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

DEFINICION DE ORTOGONAL.

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

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

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

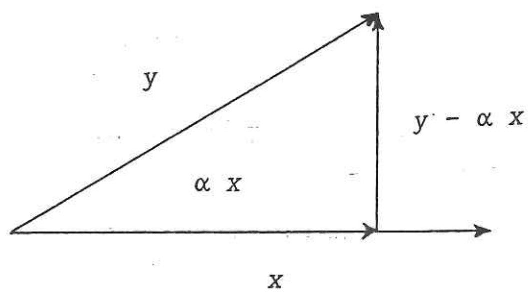


Figura 5

¡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:

$$y - \alpha x \text{ es ortogonal a } x \iff \alpha = \frac{x \cdot y}{\| x \|^2}$$

Demostración:

$$\begin{aligned}(y - \alpha x) \perp x &\iff x\cdot(y - \alpha x) = 0\\ &\iff x\cdot y + x\cdot(-\alpha x) = 0\\ &\iff x\cdot y - \alpha\,(x\cdot x) = 0\\ &\iff \alpha = \frac{x\cdot y}{x\cdot x} \quad \text{con } x \neq 0\\ &\iff \alpha = \frac{x\cdot y}{\| x \|^2}\end{aligned}$$

■

El resultado que se enuncia y demuestra a continuación es preliminar para la definición del ángulo entre dos vectores.

Desigualdad de Schwartz.

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

$$|x \cdot y| \leq \| x \| \| y \|$$

Demostración:

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

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

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

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

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

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

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

$$\Rightarrow |x \cdot y| \leq \| x \| \| y \| \quad \blacksquare$$

Ejemplos

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* mxn *sobre* $\mathbb{R}$ como un arreglo rectangular de mn números: $\alpha_{11},\alpha_{12},\ldots,\alpha_{mn}$, dispuestos en m renglones y n columnas. Los números que forman el arreglo se llaman *elementos* de la matriz.

Denotaremos a las matrices con letras mayúsculas.

Por lo tanto si A es una matriz mxn entonces A es de la forma:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \ldots & \alpha_{1n} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} \ldots & \alpha_{mn} \end{bmatrix}$$

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.

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

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

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

donde:

$$a_1 = \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{bmatrix}, \; a_2 = \begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{m2} \end{bmatrix}, \ldots, \; a_n = \begin{bmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{bmatrix}$$

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:

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

donde:

$$r_1^t = [\ \alpha_{11}\ \alpha_{12} \ldots \alpha_{1n}\ ]$$
$$r_2^t = [\ \alpha_{21}\ \alpha_{22} \ldots \alpha_{2n}\ ]$$
$$\vdots$$
$$r_m^t = [\ \alpha_{m1}\ \alpha_{m2} \ldots \alpha_{mn}\ ]$$

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

Al conjunto de todas las matrices mxn 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* $\begin{bmatrix} \otimes & \otimes & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes \\ \otimes & \otimes & \otimes & \otimes \end{bmatrix}$

2. *Columnas* $\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$

3. *Renglones* $\begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix}$

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 mxn son iguales si y solo si sus elementos correspondientes son iguales.

DEFINICION.

Una matriz A mxn 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 M_n.

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:

SUMA DE MATRICES.

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. Podriamos haber definido A + B por columnas ó por renglones.

1.a. Es decir si $A = [a_1|a_2|\ldots|a_n]$, $B = [b_1|b_2|\ldots|b_n]$, entonces

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

ya que si $C = [c_1 | c_2 | \ldots | c_n]$ entonces:

$$c_j = \begin{bmatrix} \gamma_{1j} \\ \vdots \\ \gamma_{mj} \end{bmatrix} = \begin{bmatrix} \gamma_{1j} + \beta_{1j} \\ \vdots \\ \gamma_{mj} + \beta_{mj} \end{bmatrix} = \begin{bmatrix} \gamma_{1j} \\ \vdots \\ \gamma_{mj} \end{bmatrix} + \begin{bmatrix} \beta_{1j} \\ \vdots \\ \beta_{mj} \end{bmatrix}$$

$$= a_j + b_j \quad \text{con} \quad 1 \leq j \leq n.$$

1.b. Analogamente si

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

podríamos haber definido $\quad A + B = \begin{bmatrix} r_1^t + s_1^t \\ \vdots \\ r_m^t + s_m^t \end{bmatrix}$

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.

Producto de una Matriz por un Escalar.

DEFINICION.

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

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

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

Observaciones.

Podriamos haber definido a λA por columnas o por renglones:

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

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

Ejemplos

1. *Vectores Columna*

Si A es una matriz mxl, entonces A es de la forma:

$$A = \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{bmatrix}$$

que es exactamente la estructura de los miembros de $\mathbb{R}^n$.
No distinguiremos entre una matriz $A \in M_{(m,n)}$ y un vector $a \in \mathbb{R}^n$.

2. *Vectores Renglón*

Si A es una matriz lxn, entonces A es de la forma:

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

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

3. *Escalares*

Si A es una matriz 1 x 1 entonces $A = [\alpha_{11}]$, $\alpha_{11} \in \mathbb{R}$.

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

Notación:

Si $A \in M_{(m,n)}$ escribiremos A_{mxn}.

PROPIEDADES DE MATRICES.

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

PRODUCTO MATRICIAL.

Los escalares que intervienen en el sistema se pueden agrupar en tres estructuras diferentes de la siguiente manera:

a) Los coeficientes α_{ij} del sistema en la matriz A_{mxn};

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

b) Las incógnitas ξ_i en el vector x;

$$x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

c) Los valores β_j en el vector b;

$$b = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}$$

Con estos elementos, es usual expresar el sistema anterior como una expresión de la forma Ax = b, para lo cual hemos de definir el producto de una matriz por un vector apropiadamente. El primer paso para lograr este objetivo es reconocer que si en el sistema tomamos al vector b como la columna cuyas componentes son los valores β_j, el miembro izquierdo de cada ecuación también es componente de un vector al cual denotaremos como Ax, es decir

$$Ax = \begin{bmatrix} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n \\ \vdots \qquad \vdots \qquad \vdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} = b$$

Esta expresión nos conduce a dar la siguiente definición, en donde es muy importante observar que el número de columnas de la matriz de coeficientes es igual al número de incógnitas, en tanto que el número de renglones coincide con el número de ecuaciones.

DEFINICION.

$$\text{Sea } A_{mxn} = \begin{bmatrix} \alpha_{11} & \alpha_{12} \ldots & \alpha_{1n} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} \ldots & \alpha_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$$

el producto Ax es un vector de dimensión m tal que

$$Ax = \begin{bmatrix} \alpha_{11}\xi_1 + \alpha_{12}\xi_2 + \ldots + \alpha_{1n}\xi_n \\ \vdots \qquad \vdots \qquad \vdots \\ \alpha_{m1}\xi_1 + \alpha_{m2}\xi_2 + \ldots + \alpha_{mn}\xi_n \end{bmatrix}$$

Como primera consecuencia de esta definición tenemos, que como no hay ninguna restricción sobre el valor de m, este puede tomar el valor m = 1. En ese caso, la matriz A tiene la forma

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

y por lo tanto tenemos de hecho el producto de un renglón por una columna en donde el resultado sería un vector con una sola componente, esto es, un escalar. De esta forma, introducimos la siguiente definición.

DEFINICION.

Sean $r^t \in \mathbb{R}^{n^*}$ y $x \in \mathbb{R}^n$. Entonces el producto $r^t x$ es el escalar dado por

$$r^t x = \alpha_1 \xi_1 + \ldots + \alpha_n \xi_n$$

con $r^t = [\alpha_1 \; \alpha_2 \; \ldots \; \alpha_n]$ y $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$

Si comparamos esta definición con la de producto interior o producto punto, vemos que nos da el mismo resultado, pero que la expresión $r^t x$ es una consecuencia del producto, más general, Ax. Por esta razón, el producto punto $r \cdot x$ lo denotaremos como

$$r \cdot x = r^t x$$

Una vez hecha la definición Ax, y continuando con nuestra línea de presentación, el siguiente paso consiste en buscar las formas de interpretación posibles que se le puede dar, dependiendo de las diferentes maneras de ver a una matriz.

1) Si consideramos a A por sus renglones, tenemos

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

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

Fijándonos ahora en el producto Ax, resulta que las componentes del vector son precisamente el producto del renglón correspondiente con la columna x, esto es

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

2) Veamos ahora qué sucede si A la tomamos como un arreglo de columnas, o sea

$$A = [a_1 | a_2 | \dots | a_n].$$

En este caso, fijémonos en el vector Ax como un todo, es decir sin tomar separadamente sus componentes. Como Ax es el vector columna en $\mathbb{R}$, podemos usar las operaciones definidas en $\mathbb{R}$ para obtener que

$$Ax = \xi_1 \begin{bmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m1} \end{bmatrix} + \xi_2 \begin{bmatrix} \alpha_{12} \\ \vdots \\ \alpha_{m2} \end{bmatrix} + \dots + \xi_n \begin{bmatrix} \alpha_{1n} \\ \vdots \\ \alpha_{mn} \end{bmatrix}$$

donde las columnas del miembro derecho son precísamente las columnas de A. Así

$$Ax = [a_1 | a_2 | \dots | a_n] \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix} = \xi_1 a_1 + \xi_2 a_2 + \dots + \xi_n a_n$$

Observación.

Para poder realizar las operaciones necesarias, no existe ninguna restricción sobre m (número de renglones de A), pero es necesario que la dimensión del vector x y la de las columnas de A sean iguales para que los productos $r_i^t x$ puedan efectuarse en un caso y la multiplicación de cada ξ_i por a_i sea posible en el otro.

Ahora, si nos fijamos en que al multiplicar una matriz mxn por un vector nx1 obtenemos un vector mx1, podemos ver a la multiplicación por A_{mxn} como una función que a cada $x \in \mathbb{R}^n$ le asocia una $y \in \mathbb{R}^m$:

$$\begin{aligned} A_{mxn} : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ x &\longmapsto y = Ax \end{aligned}$$

Por ejemplo, si

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

$$A(x) = Ax = (1)\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + (0)\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\therefore \qquad \begin{bmatrix} 1 \\ 0 \end{bmatrix} \xmapsto{A} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

PROPIEDADES DEL PRODUCTO MATRIZ-VECTOR.

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

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

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

Demostración:

(Las demostraciones que se presentan a continuación, se basan en la primera interpretación. Se deja para ejercicios las demostraciones en que se usará la segunda).

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

$$= \begin{bmatrix} r_1^t x_1 \\ \vdots \\ r_m^t x_1 \end{bmatrix} + \begin{bmatrix} r_1^t x_2 \\ \vdots \\ r_m^t x_2 \end{bmatrix} = Ax_1 + Ax_2$$ ■

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

3. Si $B = \begin{bmatrix} s_1^t \\ \vdots \\ s_m^t \end{bmatrix}$ entonces

$$(A + B)x = \begin{bmatrix} r_1^t + s_1^t\ x \\ \vdots \\ r_m^t + s_m^t\ x \end{bmatrix} = \begin{bmatrix} r_1^t x + s_1^t x \\ \vdots \\ r_m^t x + s_m^t x \end{bmatrix}$$

$$= \begin{bmatrix} r_1^t x \\ \vdots \\ r_m^t x \end{bmatrix} + \begin{bmatrix} s_1^t x \\ \vdots \\ s_m^t x \end{bmatrix} = Ax + Bx$$ ■

Más adelante, cuando tratemos la solución del sistema:

$$Ax = b$$

una de las primeras cuestiones que enfrentaremos será conocer si existirá una x que la satisfaga. De acuerdo a lo que hemos discutido, esto es cierto si existen números reales $\xi_1,\ldots,\xi_n$ tales que b se pueda expresar como:

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

si esto es posible, una solución es $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}$

de aqui se infiere la importancia que tiene entender bien la interpretación de la multiplicación de una matriz por un vector.

MULTIPLICACION DE UN RENGLON POR UNA MATRIZ.

Analogamente a como definimos el producto matriz por columna $A_{mxn}\, x_{nx1} = b_{mx1}$ definiremos el producto de un renglón por una matriz:

$$y^t_{1xm}\ A_{mxn}$$

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 1xn dado por:

$$y^tA = [y^ta_1\ y^ta_2\ \dots\ y^ta_n]$$

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

DEFINICION 2.

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

$$y^tA = \eta_1 r^t_1 + \eta_2 r^t_2 + \dots + \eta_m r^t_m$$

donde

$r^t_i = [\alpha_{i1}\ \alpha_{i2} \dots \alpha_{in}]$ $\quad 1 \le i \le m$ es el i-ésimo renglón de A.

Como en el producto anterior, estas dos definiciones son equivalentes, ya que si:

$d^t_1 = \eta_1 r^t_1 + \eta_2 r^t_2 + \dots + \eta_m r^t_m$ y

$d^t_2 = [y^ta_1\ y^ta_2\ \dots\ y^ta_n]$ entonces

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

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

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

Observación.

Ahora no hay ninguna restricción sobre n (número de renglones de A), pero es necesario que la dimensión del renglón y^t sea m (número de renglones de A) para que los productos $y^t a_j$ estén definidos y los factores de $\eta_i r_i^t$ se correspondan uno con otro.

Tambien podemos ver la multiplicación de y^t_{1xm} por A_{mxn} como una función que actúa sobre y^t, y que le asocia un $d^t \in \mathbb{R}^{n^*}$.

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

por ejemplo, si

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

$$\begin{aligned} y^t A &= (1)[1\ 2] + (0)[0\ 1] + (-1)[1\ 0] \\ &= [1\ 2] + [-1\ 0] \\ &= [0\ 2] \end{aligned}$$

$$\therefore \qquad [1\ 0\ -1] \longmapsto [0\ 2]$$

PROPIEDADES DEL PRODUCTO RENGLON-MATRIZ.

1. $(y_1 + y_2)^t A = y_1^t A + y_2^t A \qquad \forall\ A_{mxn},\ y_1, y_2 \in \mathbb{R}^m$

2. $(\lambda y)^t A = \lambda(y^t A) \qquad \forall\ A_{mxn},\ y \in \mathbb{R}^m,\ \lambda \in \mathbb{R}$

3. $y^t(A + B) = y^tA + y^tB \qquad \forall\ A_{mxn}\ ,\ B_{mxn}\ , y \in \mathbb{R}^m$

Demostración:

1. $(y_1 + y_2)^tA = [(y_1 + y_2)^ta_1\ (y_1 + y_2)^ta_2\ \dots\ (y_1 + y_2)^ta_n]$

$= [y_1^ta_1 + y_2^ta_1\ \ y_1^ta_2 + y_2^ta_2\ \dots\ y_1^ta_n + y_2^ta_n]$

$= [y_1^ta_1\ y_1^ta_2\ \dots\ y_1^ta_n] + [y_2^ta_1\ y_2^ta_2\ \dots\ y_2^ta_n]$

$= y_1^tA + y_2^tA$ ■

2. $(\lambda y)^tA = (\lambda y^t)A$

$= [(\lambda y^t)a_1\ (\lambda y^t)a_2\ \dots\ (\lambda y^t)a_n]$

$= [\lambda(y^ta_1)\ \lambda(y^ta_2)\ \dots\ \lambda(y^ta_n)]$

$= \lambda[y^ta_1\ y^ta_2\ \dots\ y^ta_n]$

$= \lambda(y^tA)$ ■

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

$= [y^ta_1 + y^tb_1\ \ y^ta_2 + y^tb_2\ \dots\ y^ta_n + y^tb_n]$

$= [y^ta_1\ y^ta_2\ \dots\ y^ta_n] + [y^tb_1\ y^tb_2\ \dots\ y^tb_n]$

$= y^tA + y^tB$ ■

Observación Importante.

Por definición 2

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

es decir $d^t = y^tA$ *si y solo si* d^t *es una combinación de los renglones de* A.

Si nos planteamos el problema:

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

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

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

MULTIPLICACION DE MATRICES.

Supongamos que tenemos dos matrices A_{mxn} y B_{nxk}.
¿Cómo definimos el producto AB de modo que respete los productos que hemos definido anteriormente?

Si consideramos a B como un arreglo de k columnas;

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

el producto se podría indicar como:

$$AB = A[b_1 | b_2 | \ldots | b_k]$$

Para que AB coincida en Ax cuando B sea de dimensión. nx1. La multiplicación AB se define como:

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

Por otro lado, considerando a A como un arreglo de m renglones:

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

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

La multiplicación de un renglón por una matriz y el producto AB cuando A sea de dimensión 1xn, nos conduce a definir.

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

Tambien podremos definir AB como se hace usualmente, elemento a elemento:

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

Estas tres formas, aparentemente diferentes, de definir el producto AB son equivalentes, ya que:

$$[Ab_1 | Ab_2 | \ldots | Ab_k] = \begin{bmatrix} r_1^t b_1 & r_1^t b_2 & \cdots & r_1^t b_k \\ \vdots & \vdots & & \vdots \\ r_m^t b_1 & r_m^t b_2 & \cdots & r_m^t b_k \end{bmatrix}$$

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

Con esta aclaración, nos permitimos hacer la siguiente definición.

DEFINICION.

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

1. $AB = [Ab_1 | Ab_2 | \ldots | Ab_k]$

2. $AB = \begin{bmatrix} r^t_1 B \\ \vdots \\ r^t_m B \end{bmatrix}$

3. $AB = \begin{bmatrix} r^t_1 b_1 & r^t_1 b_2 & \ldots & r^t_1 b_k \\ \vdots & \vdots & & \vdots \\ r^t_m b_1 & r^t_m b_2 & \ldots & r^t_m b_k \end{bmatrix}$

es decir $\gamma_{ij} = \gamma^t_i b_j \qquad 1 \leq i \leq m,\ 1 \leq j \leq k$

Usaremos cualquiera de las tres expresiones que sea más adecuada al problema a resolver.

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).

$$A_{2x3} \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}$$

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.

$$\begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix} B_{3x2} = \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}$$

3. Cada entrada de AB es el producto de un renglón y una columna:
   γ_{ij} = (i-ésimo renglón de A)·(j-ésima columna de B).

$$\begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix} \begin{bmatrix} * & * & * \\ * & * & * \\ * & * & * \end{bmatrix} = \begin{bmatrix} * & * & * \\ * & * & * \end{bmatrix}$$

Observación.

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

Propiedades del Producto Matriz-Matriz.

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 las siguientes razones:

Si A es mxn y B es nxk, con $k \neq m$, entonces AB esta definido pero BA no lo está.

Si A es de mxn y B es nxm, con $m \neq n$, entonces tanto AB como BA están definidos, pero no pueden ser iguales ya que AB es mxm y BA es nxn, $n \neq m$.

Si tanto A como B son nxn, entonces AB y BA son de orden n, pero aún en este caso la conmutatividad puede fallar, como en el siguiente ejemplo:

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

Interpretacion del Producto Matriz-Matriz.

Nos podriamos preguntar para que 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 $AB = [Ab_1 | Ab_2 | \ldots | Ab_k]$

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

$$Ab_j = a_1\beta_{1j} + a_2\beta_{2j} + \ldots + a_n\beta_{nj} \quad \text{con} \quad 1 \leq j \leq k.$$

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:

$$B_{nxk}: M_{(m,n)} \longrightarrow M_{(m,k)}$$

$$A_{mxn} \xmapsto{B_{nxk}} (AB)_{mxk}$$

La utilidad de esta interpretación reside en que a veces necesitaremos realizar algunas operaciones sobre las columnas de A, porque desearemos transformarla en otra con la cual sea más facil trabajar. Esto lo logramos escogiendo adecuadamente una matriz B_{nxn} y multiplicandola por la derecha de A (posmultiplicando).

2. Por otro lado, tenemos que

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

Habiamos visto que $r_i^t B$ es una combinación de los renglones de B

$$r_i^t B = \alpha_{i1} s_1^t + \alpha_{i2} s_2^t + \ldots + \alpha_{in} s_n^t \quad \text{con} \quad 1 \leq i \leq m.$$

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

Es decir, *los renglones de* AB *son combinaciones de los renglones de* B.

Podemos pensar en esta ocasión que A est! operando sobre B:

$$A_{m\times n} : M_{(n,k)} \longrightarrow M_{(m,k)}$$

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

A veces necesitaremos realizar operaciones sobre los renglones de una matriz, por ejemplo de la matriz B, o 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 $A_{m\times n} = [a_1 | a_2 | \ldots | a_n]$.

1a. Queremos que $AB = [\lambda_1 a_1 | \lambda_2 a_2 | \ldots | \lambda_n a_n]$. Esto es, si queremos multiplicar a las columnas de A por los escalares λ_i, debe multiplicarse a A por B a la derecha, donde B es:

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

ya que
$$\begin{aligned} Ab_1 &= \lambda_1 a_1 + 0a_2 + \ldots + 0a_n \\ Ab_1 &= 0a_1 + \lambda_2 a_1 + \ldots + 0a_n \\ &\vdots \\ Ab_n &= 0a_1 + 0a_2 + \ldots + \lambda_n a_n \end{aligned}$$

1b. Queremos $AB = [a_2 | a_1 | a_3 | \ldots | a_n]$, es decir, si queremos cambiar la primera columna de A en la segunda, entonces:

$$B = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

ya que $Ab_1 = 0a_1 + 1a_2 + 0a_3 + \dots + 0a_n$

$Ab_2 = 1a_1 + 0a_2 + 0a_3 + \dots + 0a_n$

$Ab_j = 0a_1 + \dots + 1a_j + \dots + 0a_n$

con $3 \le j \le n$.

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

2a. Dar A tal que:

$$AB = \begin{bmatrix} s_1^t \\ \lambda_1 s_1^t + \lambda_2 s_2^t \\ s_3^t \\ \vdots \\ s_n^t \end{bmatrix} \Rightarrow A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_1 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

2b. Dar A tal que

$$AB = \begin{bmatrix} s_2^t \\ s_1^t \\ s_3^t \\ \vdots \\ s_n^t \end{bmatrix} \Rightarrow A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

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 renglón de B)+(j-ésimo renglón de B)

$$A = \left[\begin{array}{ccccc} 1 & & & & \\ & \ddots & & & \\ \cdots & 1 & \cdots & 0 & \\ & \vdots & & \vdots & \\ \cdots & \lambda & \cdots & 1 & \\ & & & & \ddots \\ & & & & 1 \end{array}\right] \begin{array}{c} \\ \\ i \\ \\ j \\ \\ \\ \end{array}$$

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

Entonces si queremos transformar a B en otra matriz que deja invariables los renglones uno y tres, y al segundo le sume un múltiplo del primero, premultiplicamos con una matriz de la forma

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

En particular, podemos escoger λ de manera que:

$$\lambda[1\ 2\ 1] + [2\ 0\ 1] = [0\ *\ *],$$

es decir, para que en el primer elemento del segundo renglón sea nulo.

Si escogemos $\lambda = -2$, entonces

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

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

Más adelante veremos que necesitaremos efectuar este tipo de operación, en el proceso conocido como triangulación de una matriz.

Casos Especiales del Producto de Matrices.

1. *Producto de renglón por columna*

Sea $x^t \in M_{(1,n)}$, $y \in M_{(n,1)}$. Entonces $x^t y \in M_{(1,1)} \sim \mathbb{R}$

O sea, que el producto escalar es un caso particular del producto de matrices; $x^t y = x \cdot y$.

2. *Producto de columna por renglón*

Sean $a \in \mathbb{R}^m$, $b \in \mathbb{R}^n$. Entonces el producto ab^t está definido y ab^t es una matriz mxn. Este tipo de producto es muy importante.

Observar que:

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

Es decir, *cada renglón de* ab^t *es un múltiplo de* b^t.

Por otra parte

$$ab^t = a[\beta_1 \ \beta_2 \dots \beta_n] = [\beta_1 a \ \beta_2 a \ \dots \ \beta_n a]$$

Es decir, *cada columna de* ab^t *es un múltiplo del vector* a.

Ejemplo.

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

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

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

$$e_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{l} \\ \\ \\ \text{lugar } i \\ \\ \\ \\ \end{array}$$

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

Realicemos el producto.

$$e_ie_j^k = \begin{array}{c} \\ \\ i \\ \\ \\ \end{array}\begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{array}{c} \qquad\quad j \\ \left[\, 0 \;\dots\; 1 \;\dots\; 0 \,\right] \end{array} = \begin{array}{c} \begin{array}{ccccccc} & & & j & & & \end{array} \\ \begin{bmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \end{array} \begin{array}{c} \\ \\ \\ i \\ \\ \\ \end{array}$$

Asi, $e_ie_j^t$ es la matriz mxn 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 $e_ie_j^t$ por

$$E_{ij} = e_ie_j^t$$

Expresaremos a continuación la matriz A, como suma de matrices elementales de 3 formas distintas:

1. *Escribiremos a A como columnas*

$$A_{mxn} = [a_1 | a_2 | \dots | a_n]$$

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

Ahora

$$[a_1|0|\ldots|0] = [1a_1|0a_2|\ldots|0a_n] = a_1e_1^t$$

$$[0|a_2|\ldots|0] = [0a_1|1a_2|\ldots|0a_n] = a_2e_2^t$$

$$\vdots$$

$$[0|0|\ldots|a_n] = [0a_1|0a_2|\ldots|1a_n] = a_ne_n^t$$

$$\therefore\ A = a_1e_1^t + a_2e_2^t + \ldots + a_ne_n^t$$

$$\therefore\ A = \sum_{j=1}^{n} a_je_j^t$$

Nota: Como $a_j \in \mathbb{R}^m$ y $e_j^t \in \mathbb{R}^{n^*}$, $a_je_j^t \in M_{(m,n)}$

2. *Si escribimos a A como renglones*

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

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

Pero

$$\begin{bmatrix} r_1^t \\ 0 \\ \vdots \\ 0 \end{bmatrix} = e_1r_1^t\ ; \quad \begin{bmatrix} 0 \\ r_2^t \\ \vdots \\ 0 \end{bmatrix} = e_2r_2^t\ ; \ \ldots\ ; \quad \begin{bmatrix} 0 \\ 0 \\ \vdots \\ r_m^t \end{bmatrix} = e_mr_m^t$$

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

$$e_i r_i^t \in M_{(m,n)}$$

$$\therefore \qquad A = \sum_{i=1}^{n} e_i r_i^t$$

3. *Finalmente consideremos a A por sus entradas*

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

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

$$\dots + \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ \alpha_{m1} & 0 & \dots & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \alpha_{m2} & \dots & 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_{mn} \end{bmatrix} =$$

$$= \alpha_{11}E_{11} + \alpha_{12}E_{12} + \dots + \alpha_{1n}E_{1n} +$$

$$+ \alpha_{21}E_{21} + \alpha_{22}E_{22} + \dots + \alpha_{2n}E_{2n} + \dots$$

$$\dots + \alpha_{m1}E_{m1} + \alpha_{m2}E_{m2} + \dots + \alpha_{mn}E_{mn}$$

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

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

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

Observación.

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

Como consecuencia de estas tres consideraciones, podemos obtener una serie de resultados de acuerdo a nuestros propósitos. Como una muestra, expresaremos el producto de dos matrices como una suma de matrices:

Sea $A_{mxn} = [a_1 | a_2 | \ldots | a_n]$

$$B_{mxk} = \begin{bmatrix} s_1^t \\ s_2^t \\ \vdots \\ s_n^t \end{bmatrix}$$

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

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

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

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

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

TRANSPOSICION DE MATRICES.

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

$$\underset{n\times 1}{x} \longmapsto \underset{1\times n}{x^t}$$

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

$$\underset{1\times n}{x^t} \longmapsto \underset{n\times 1}{x}$$

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, si queremos que la operación sea una generalización de la ya discutida.

DEFINICION.

Sea $A \in M_{(m,n)}$. Entonces, la *matriz transpuesta* de A, denotada por A^t, es la matriz nxm que se obtiene de A intercambiando sus renglones y columnas:

$$A = \begin{bmatrix} \alpha_{11} \alpha_{12} \dots \alpha_{1n} \\ \alpha_{21} \alpha_{22} \dots \alpha_{2n} \\ \vdots \quad \ddots \quad \vdots \\ \alpha_{m1} \alpha_{m2} \dots \alpha_{mn} \end{bmatrix} \Rightarrow A^t = \begin{bmatrix} \alpha_{11} \alpha_{21} \dots \alpha_{m1} \\ \alpha_{12} \alpha_{22} \dots \alpha_{m2} \\ \vdots \quad \ddots \quad \vdots \\ \alpha_{1n} \alpha_{2n} \dots \alpha_{mn} \end{bmatrix}$$

Tenemos entonces tres maneras de determinar la transpuesta de A

1. Si $A = [a_1 | a_2 | \ldots | a_n]$ entonces $A^t = \begin{bmatrix} a_1^t \\ a_2^t \\ \vdots \\ a_n^t \end{bmatrix}$

2. Si $A = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_m^t \end{bmatrix}$ entonces $A^t = [r_1 | r_2 | \ldots | r_m]$

3. Si $A = [\alpha_{ij}]$, $1 \le j \le n$; $1 \le i \le m$, entonces $A^t = [\alpha_{ji}]$

Ejemplos

1. $x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix} \Rightarrow x^t = [\xi_1 \ \xi_2 \ \cdots \ \xi_n]$

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

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

Las matrices que cumplen con $A^t = A$ ó $A^t = -A$ nos serán muy útiles posteriormente. Por ésta razón y por su estructura las clasificamos como sigue.

DEFINICION.

Sea A_{nxn}. Decimos que A es *simétrica* si $A^t = A$; decimos que A es *anti-simétrica* si $A^t = -A$.

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

PROPIEDADES DE LA TRANSPOSICION.

1. $(A^t)^t = A$ $\qquad \forall\ A_{mxn}$.

2. $(A + B)^t = A^t + B^t$ $\qquad \forall\ A, B \in M_{(m,n)}$.

3. $(Ax)^t = x^t A^t$ $\qquad \forall\ A_{mxn},\ x \in \mathbb{R}^n$.

4. $(y^t A)^t = A^t y$ $\qquad \forall\ A_{mxn},\ y \in \mathbb{R}^m$.

5. $(AB)^t = B^t A^t$ $\qquad \forall\ A_{mxn}, B_{nxk}$.

6. $(A^t A)^t = A^t (A^t)^t = A^t A$ $\qquad \forall\ A_{mxn}$.

Demostración:

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

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

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

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

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

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

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

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

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

$$\therefore (Ax)^t = [r_1^t x \quad r_2^t x \quad \ldots \quad r_m^t x]$$

$$= [x^t r_1 \quad x^t r_2 \quad \ldots \quad x^t r_m] = x^t A^t$$

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

4. $(y^t A)^t = [y^t a_1 \quad y^t a_2 \quad \ldots \quad y^t a_n]$

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

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

5. $AB = [Ab_1 | Ab_2 | \ldots | Ab_k]$

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

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

MATRICES ESPECIALES

La presencia de elementos $\alpha_{ij} = 0$ en una matriz generalmente simplifica su análisis y los cálculos en que interviene dicha matriz.

Algunas matrices con determinada distribución de ceros son tan importantes que merecen nombres especiales, dependiendo de la forma en que quede la distribución de sus elementos.

MATRICES DIAGONALES.

DEFINICION.

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 todos los elementos fuera de la diagonal son nulos.

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

A veces denotaremos a una matriz diagonal por:

$$D = \text{diag}(\delta_i)$$

donde δ_i es el i-ésimo elemento diagonal, ó por:

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

¿ 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:

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

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 $D = \text{diag}(\delta_i)$, con $\delta_i = \delta \quad \forall i = 1,\dots,n$, entonces $DA = AD$ $\forall A_{nxn}$.

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

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

se llama *matriz identidad de orden n.*

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

Observar que:

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

donde $e_i \in \mathbb{R}^n \qquad 1 \le i \le n.$

PROPIEDAD.

Sea A de mxn. Entonces $I_m A = A I_n = A$

Demostración:

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

$$AI_n = [a_1|a_2|\dots|a_n]\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = [a_1|a_2|\dots|a_n] \quad \blacksquare$$

Permutaciones Elementales.

DEFINICION 1.

Una matriz P_{ij} de orden n se dice que es una *permutación elemental,* si se obtiene a partir de la identidad intercambiando la columna i-ésima con la j-ésima, o sea que es de la forma:

$$P_{ij} = [e_1|e_2|\dots|e_{i-1}|e_j|e_{i+2}|\dots|e_{j-1}|e_i|e_{j+1}|\dots|e_n]$$

donde $i < j$.

Es decir

$$P_{ij} = \begin{array}{c} \begin{array}{ccccc} \quad & & i & & j \end{array} \\ \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & \dots & 0 & \dots & 1 & \\ & & \vdots & & \vdots & \\ & \dots & 1 & \dots & 0 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \end{array}$$

Observese que P_{ij} es tambien la matriz que se obtiene a partir de la matriz identidad intercambiando el renglón i-ésimo con el j-ésimo. Podemos, por lo tanto, definir una matriz de permutación en función de sus renglones.

DEFINICION 2.

Una matriz P_{ij} de orden n se dice que es una *permutación elemental* si es de la forma

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

Por ejemplo

$$
P_{23} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}
$$

Propiedades de las Permutaciones Elementales.

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

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

Demostración:

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

2. Sea $P_{ij} = [e_1 | \ldots | e_j | \ldots | e_i | \ldots | e_n]$

$$\Rightarrow (P_{ij})^2 = P_{ij}[e_1 | \ldots | e_j | \ldots | e_i | \ldots | e_n]$$
$$= [P_{ij}e_1 | P_{ij}e_2 | \ldots | P_{ij}e_j | \ldots | P_{ij}e_i | \ldots | P_{ij}e_n]$$
$$= [e_1 | \ldots | e_i | \ldots | e_j | \ldots | e_n] = I$$

■

La propiedad número 2, que se puede escribir como $P_{ij}^2 = P_{ij}(P_{ij}I) = I$ tiene una interpretación operativa. Nos está diciendo que si intercambiamos dos renglones de la matriz identidad y despues los volvemos a intercambiar, entonces al final tendremos a la matriz original, la identidad. Esto tambien se puede pensar por columnas: si intercambiamos dos columnas de la identidad y despues 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 $y^t \in \mathbb{R}^n$?

$$\begin{aligned} y^t P_{ij} &= y^t[e_1|\dots|e_j|\dots|e_i|\dots|e_n] \\ &= [y^t e_1|\dots|y^t e_j|\dots|y^t e_i|\dots|y^t e_n] \\ &= [\eta_1|\dots|\eta_j|\dots|\eta_i|\dots|\eta_n] \end{aligned}$$

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

Ejemplo.

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

Ahora, sea A una matriz nxk. ¿ Qué efecto produce sobre A la premultiplicación por P_{ij}?

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

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:

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

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

$$\begin{aligned} P_{ij} &= A\ [e_1|\ldots|e_j|\ldots|e_i|\ldots|e_n] \\ &= [Ae_1|\ldots|Ae_j|\ldots|Ae_i|\ldots|Ae_n] \\ &= [a_1|\ldots|a_j|\ldots|a_i|\ldots|a_n] \end{aligned}$$

Ejemplo:

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

MATRICES DE PERMUTACION.

DEFINICION.

Una matriz cuadrada P de orden n, es *matriz de permutación* si sus columnas se obtienen permutando las columnas de la matriz identidad:

$$P = [p_1|p_2|\ldots|p_n \quad = \quad e_{\sigma(1)}|e_{\sigma(2)}|\ldots|e_{\sigma(n)}]$$

donde σ es una permutación de los elementos del conjunto $\{1,\ldots,n\}$.

PROPIEDADES.

1. Si $P = [e_{\sigma(1)} | e_{\sigma(2)} | \ldots | e_{\sigma(n)}]$, entonces $P = \begin{bmatrix} e^t_{\sigma^{-1}(1)} \\ e^t_{\sigma^{-1}(2)} \\ \vdots \\ e^t_{\sigma^{-1}(n)} \end{bmatrix}$

donde σ^{-1} es la permutación inversa de σ.

2. Si $P = [e_{\sigma(1)} | e_{\sigma(2)} | \ldots | e_{\sigma(n)}]$ entonces $AP = [a_{\sigma(1)} | a_{\sigma(2)} | \ldots | a_{\sigma(n)}]$

y $PA = \begin{bmatrix} r^t_{\sigma^{-1}(1)} \\ r^t_{\sigma^{-1}(2)} \\ \vdots \\ r^t_{\sigma^{-1}(n)} \end{bmatrix}$

MATRICES TRAPEZOIDALES.

DEFINICION.

1. Sea A una matriz mxn. decimos que A es *trapezoidal superior (inferior)* si $i \quad j \Rightarrow \alpha_{ij} = 0$.

2. Si una matriz A es trapezoidal superior (inferior) y además es cuadrada se dice que es *triangular superior (inferior)*. Utilizaremos la letra U para denotar exclusivamente matrices triangulares superiores y la letra L para triangulares inferiores. La distinción entre triangulares de un mismo tipo se hará mediante el uso de subíndices.

Las matrices trapezoidales superiores tienen una de las siguientes formas:

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

$(m > n)$

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

$(m < n)$

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

$(m = n)$

Las matrices trapezoidales inferiores tienen una de las siguientes formas:

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

$(m > n)$

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

$(m < n)$

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

$(m = n)$

A veces indicaremos una matriz por su forma para referirnos a ella. Por ejemplo:

$$
U = \begin{bmatrix} \quad \end{bmatrix} \qquad L = \begin{bmatrix} \quad \end{bmatrix}
$$

1. Si L es triangular inferior, entonces

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

Demostración:

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

2. Si U es triangular superior, entonces

$$U = [\mu_{11} e_1 | \mu_{12} e_1 + \mu_{22} e_2 | \ldots | \mu_{1n} e_1 + \mu_{2n} e_2 + \ldots + \mu_{nn} e_n]$$

Demostración:

$$U = [e_1 | e_2 | \ldots | e_n] \begin{bmatrix} \mu_{11} & \mu_{12} & \ldots & \mu_{1n} \\ & \mu_{22} & \ldots & \mu_{2n} \\ & & \ddots & \vdots \\ & & & \mu_{nn} \end{bmatrix}$$

$$= [\mu_{11} e_1 | \mu_{12} e_1 + \mu_{22} e_2 | \ldots | \mu_{1n} e_1 + \mu_{2n} e_2 + \ldots + \mu_{nn} e_n] \blacksquare$$

3. El producto de dos matrices triangulares inferiores es triangular inferior.

Demostración:

Supongamos que

$$L_1 = \begin{bmatrix} \lambda_{11} & & \\ \lambda_{21} & \lambda_{22} & \\ \vdots & \vdots & \ddots \\ \lambda_{n1} & \lambda_{n2} \dots & \lambda_{nn} \end{bmatrix} \qquad L_2 = \begin{bmatrix} r_1^t \\ r_2^t \\ \vdots \\ r_n^t \end{bmatrix}$$

entonces

$$L_1L_2 = \begin{bmatrix} \lambda_{11} r_1^t \\ \lambda_{21} r_1^t + \lambda_{22} r_2^t \\ \vdots \\ \lambda_{n1} r_1^t + \lambda_{n2} r_2^t + \dots + \lambda_{nn} r_n^t \end{bmatrix} = \begin{bmatrix} z_1^t \\ z_2^t \\ \vdots \\ z_n^t \end{bmatrix}$$

Si L_2 es triangular inferior entonces

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

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

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

$$\vdots$$

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

por lo tanto

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

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

$$= (\lambda_{21}\alpha_{11} + \lambda_{22}\alpha_{21}) e_1^t + (\lambda_{22}\alpha_{22}) e_2^t$$

$$\vdots$$

$$z_n^t = \lambda_{n1}(\alpha_{11} e_1^t) + \lambda_{n2}(\alpha_{21} e_1^t + \alpha_{22} e_2^t) + \dots$$

$$\dots + \lambda_{nn}(\alpha_{n1} e_1^t + \alpha_{n2} e_2^t + \alpha_{n3} e_3^t + \dots + \alpha_{nn} e_n^t)$$

$$= (\lambda_{n1}\alpha_{11} + \lambda_{n2}\alpha_{21} + \dots + \lambda_{nn}\alpha_{n1}) e_1^t + (\lambda_{n2}\alpha_{22} + \dots + \lambda_{nn}\alpha_{n2}) e_2^t + \dots$$

$$\dots + (\lambda_{nn}\alpha_{nn}) e_n^t$$

$$\therefore z_n^t = \zeta_{11} e_1^t$$
$$z_n^t = \zeta_{21} e_1^t + \zeta_{22} e_2^t$$
$$\vdots$$
$$z_n^t = \zeta_{n1} e_1^t + \zeta_{n2} e_2^t + \dots + \zeta_{nn} e_n^t$$

$\therefore L_1 L_2$ es triangular inferior. ■

4. El producto de dos matrices triangulares superiores es triangular superior.

Demostración:

Supongamos

$$U_1 = \begin{bmatrix} \mu_{11} & \mu_{12} & \dots & \mu_{1n} \\ & \mu_{22} & \dots & \mu_{2n} \\ & & \ddots & \vdots \\ & & & \mu_{nn} \end{bmatrix} \qquad U_2 = [a_1 | a_2 | \dots | a_n]$$

Entonces

$$U_2 U_1 = [\mu_{11} a_1 | \mu_{12} a_1 + \mu_{22} a_2 | \dots | \mu_{1n} a_1 + \mu_{2n} a_2 + \dots + \mu_{nn} a_n]$$

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

Si A es triangular superior, entonces

$$a_n = \alpha_{11} e_1$$
$$a_n = \alpha_{12} e_1 + \alpha_{22} e_2$$
$$\vdots$$
$$a_n = \alpha_{1n} e_1 + \alpha_{2n} e_2 + \dots + \alpha_{nn} e_n$$

por lo tanto

$$b_1 = \mu_{11}(\alpha_{11}e_1) = (\mu_{11}\alpha_{11})e_1$$

$$= \beta_{11}e_1$$

$$b_2 = \mu_{12}(\alpha_{11}e_1) + \mu_{22}(\alpha_{12}e_1 + \alpha_{22}e_2)$$

$$= (\mu_{12}\alpha_{11} + \mu_{22}\alpha_{12})e_1 + (\mu_{22}\alpha_{22})e_2$$

$$= \beta_{12}e_1 + \beta_{22}e_2$$

$$\vdots$$

$$b_n = \mu_{1n}(\alpha_{11}e_1) + \mu_{2n}(\alpha_{12}e_1 + \alpha_{22}e_2) + \dots$$

$$\dots + \mu_{nn}(\alpha_{1n}e_1 + \alpha_{2n}e_2 + \alpha_{n3}e_3 + \dots + \alpha_{nn}e_n)$$

$$= (\mu_{1n}\alpha_{11} + \mu_{2n}\alpha_{12} + \dots + \mu_{nn}\alpha_{1n})e_1 + (\mu_{2n}\alpha_{22} + \dots + \mu_{nn}\alpha_{2n})e_2 + \dots$$

$$\dots + (\mu_{nn}\alpha_{nn})e_n$$

$$= \beta_{1n}e_1 + \beta_{2n}e_2 + \dots + \beta_{nn}e_n$$

$\therefore$ U_2U_1 es triangular superior. ■

Matrices por Bloques.

Hemos visto que una matriz puede considerarse como un arreglo de escalares, vectores columna o vectores renglón. Es natural que ahora nos preguntemos si también es factible ver a una matriz como un arreglo de submatrices. Veremos que ésto no solamente es posible sino que además nos será de utilidad para construir algoritmos prácticos de manipulación con matrices.

Las ideas esenciales surgen de la manipulación de sistemas de ecuaciones que pueden manipularse en pequeños subsistemas. Por ejemplo, supongamos que tenemos el sistema de ecuaciones:

$$\left.\begin{aligned} 1\,\xi_1 + 1\,\xi_2 + 1\,\xi_3 + 1\,\xi_4 &= 1 \\ 0\,\xi_1 + 1\,\xi_2 + 1\,\xi_3 + 1\,\xi_4 &= 0 \\ 0\,\xi_1 + 0\,\xi_2 + 1\,\xi_3 + 0\,\xi_4 &= 1 \\ 0\,\xi_1 + 0\,\xi_2 + 1\,\xi_3 + 1\,\xi_4 &= 1 \end{aligned}\right\} \qquad (1)$$

que se puede escribir como Ax = b, con

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix} \quad \text{y } b = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix} \qquad (2)$$

Si observamos las dos últimas ecuaciones del sistema, vemos que forman un sistema de dos ecuaciones en las incógnitas ξ_3 y ξ_4, que se puede resolver independientemente de las otras ecuacines. Una vez determinado ξ_3 y ξ_4, sustituímos en las otras dos ecuaciones sus valores para quedarnos nuevamente con dos ecuaciones en las incógnitas ξ_1 y ξ_2, que resolvemos para tener la solución del sistema original. El tratamiento de sistemas con este tipo de estructuras es equivalente a considerar una matriz como un arreglo de submatrices, como veremos a continuación.

Si del sistema anterior separamos las dos últimas ecuaciones, nos quedan dos subsistemas;

$$\begin{aligned} 1\ \xi_1 + 1\ \xi_2 + 1\ \xi_3 + 1\ \xi_4 &= 1 \\ 0\ \xi_1 + 1\ \xi_2 + 1\ \xi_3 + 1\ \xi_4 &= 0 \end{aligned} \tag{3}$$

$$\begin{aligned} 0\ \xi_1 + 0\ \xi_2 + 1\ \xi_3 + 0\ \xi_4 &= 1 \\ 0\ \xi_1 + 0\ \xi_2 + 1\ \xi_3 + 1\ \xi_4 &= 1 \end{aligned} \tag{4}$$

El sistema (3) se puede escribir como

$$A_1 x = b_1$$

$$\text{con } A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{5}$$

mientras que el sistema (4) se puede expresar como

$$A_2 x = b_2$$

donde

$$A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad x = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \\ \xi_4 \end{bmatrix}, \quad b_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{6}$$

Si comparamos (5) y (6) con (2) vemos que esto es equivalente a separar a la matriz A en dos matrices A_1 y A_2, en tanto que b se ha dividido en dos vectores b_1 y b_2.

A continuación estudiaremos las operaciones entre matrices en términos de bloques y como punto de partida trataremos las que involucran a los vectores.

DEFINICION.

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

$$\text{Sean } x_1 = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_r \end{bmatrix} \text{ y } x_2 = \begin{bmatrix} \xi_{r+1} \\ \xi_{r+2} \\ \vdots \\ \xi_n \end{bmatrix} \text{ donde } 1 \le r \le n.$$

Si escribimos $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, 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 una forma se llama *partir por bloques.*

Ejemplo.

$$\text{Sea } x = \begin{bmatrix} 1 \\ 0 \\ 2 \\ 1 \\ 3 \end{bmatrix}. \text{ Si } x = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad x = \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \text{ entonces}$$

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

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

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 ambas 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

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

Además $x^t y = [x_1^t \; x_2^t] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = x_1^t y_1 + x_2^t y_2$

Ejemplo:

$$\Big[[1 \;\, 2][3 \;\, 2 \;\, 1]\Big] \begin{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} = [1 \;\, 2] \begin{bmatrix} 0 \\ 1 \end{bmatrix} + [3 \;\, 2 \;\, 1] \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Esto nos hace ver que si x y y están divididos por bloques, al sumar x con y sus particiones se comportan como si fueran escalares. 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 vectores, y en algunas aplicaciones más sofisticadas se utiliza. Sin embargo, el objeto de la partición por bloque es la simplificación, mientras menos dividamos a un vector, tanto mejor.

Consideremos ahora la partición de matrices.

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

$$A_1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} \cdots & \alpha_{1n} \\ \vdots & \vdots & \vdots \\ \alpha_{s1} & \alpha_{s2} \cdots & \alpha_{sn} \end{bmatrix}$$ la matriz de los primeros s renglones de A.

$$A_2 = \begin{bmatrix} \alpha_{s+1,1} & \alpha_{s+1,2} \cdots & \alpha_{s+1,n} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} \cdots & \alpha_{mn} \end{bmatrix}$$ $1 \le s \le m$ los últimos m-s renglones de A.

Entonces podemos escribir

$A = \begin{bmatrix} A_1 \\ \hline A_2 \end{bmatrix}$ y decimos que A_1 y A_2 son *submatrices* ó *bloques* de A.

Cuando sea necesario, indicaremos las dimensiones de A de la siguiente manera:

$$A_{mxn} = \begin{array}{c} \begin{bmatrix} A_1 \\ \hline A_2 \end{bmatrix} \begin{array}{l} s \\ m-s \end{array} \\ n \end{array}$$

Es decir, A está dividida en dos bloques:

A_1 de dimensión sxn

y A_2 de dimensión $(m - s)xn$

Sean $A = \begin{array}{c} \begin{bmatrix} A_1 \\ \hline A_2 \end{bmatrix} \begin{array}{l} s \\ m-s \end{array} \\ n \end{array}$ y $B = \begin{array}{c} \begin{bmatrix} B_1 \\ \hline B_2 \end{bmatrix} \begin{array}{l} s \\ m-s \end{array} \\ n \end{array}$

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

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

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

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

El producto Ax, se puede expresar en términos de los bloques de A como

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

Si $y^t \in \mathbb{R}^{m^*}$ y queremos expresar el producto $y^t A$ con $A = \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right]$, se requerirá dividir a y^t de la siguiente manera.

$$y^t = [\ \underset{s}{y_1^t} \mid \underset{m-s}{y_2^t}\ ]$$

entonces

$$y^t A = y^t \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] = [\ \underset{s}{y_1^t} \mid \underset{m-s}{y_2^t}\ ] \left[\begin{array}{c} A_1 \\ \hline A_2 \end{array}\right] \begin{array}{l} s \\ m-s \end{array} = \underset{1\,xn}{y_1^t A_1} + \underset{1\,xn}{y_2^t A_2}$$

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

Un ejemplo de matriz por bloques sería

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

También podríamos dividir a A por medio de las submatrices

$$A_1 = \begin{bmatrix} \alpha_{11} & \alpha_{12} \cdots & \alpha_{1r} \\ \vdots & \vdots & \vdots \\ \alpha_{m1} & \alpha_{m2} \cdots & \alpha_{mr} \end{bmatrix}$$ matriz con las primeras r columnas de A.

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

donde $1 \leq r \leq n$.

En este caso escribimos

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

Si pensamos en el sistema $Ax = b$, esta partición de A es equivalente a separar a la incógnita x, en subgrupos.

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

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

Entonces

$$Ax = m\ \underset{\ \ r \quad n-r}{[A_1 | A_2]} \begin{array}{c} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{array} \begin{array}{l} r \\ n-r \end{array} = A_1x_1 + A_2x_2$$

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

$$y^t[A_1 | A_2] = [y^tA_1 | y^tA_2]$$

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

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

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

Entonces

1. Si $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \begin{array}{l} r \\ n-r \end{array}$

$$\Rightarrow Ax = \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left[\begin{array}{c} A_{11}x_1 + A_{12}x_2 \\ \hline A_{21}x_1 + A_{22}x_2 \end{array}\right]$$

2. Si $y^t = \begin{array}{c|c} y_1^t & y_2^t \\ s & m-s \end{array}$ i.e. $y^t = [\, \underset{s}{y_1^t} \mid \underset{m-s}{y_2^t} \,]$

$$y^tA = [\, y_1^t \mid y_2^t \,]\left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] = [y_1^tA_{11} + y_2^tA_{21} \quad y_1^tA_{12} + y_2^tA_{22}]$$

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

$$\text{Si} \quad A_{mxn} = \begin{array}{cc} \left[\begin{array}{c|c} A_{11} & A_{12} \\ \hline A_{21} & A_{22} \end{array}\right] & \begin{array}{c} s \\ m-s \end{array} \\ \begin{array}{cc} r & n-r \end{array} & \end{array} \qquad B_{nxk} = \begin{array}{cc} \left[\begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array}\right] & \begin{array}{c} r \\ n-r \end{array} \\ \begin{array}{cc} k & k-t \end{array} & \end{array}$$

entonces

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

DEFINICION.

Decimos que un conjunto de matrices está dividido *conformemente con respecto a una operación* si las operaciones a realizar entre los bloques correspondientes están definidas.

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

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

Las siguientes matrices están divididas conformemente con respecto a la multiplicación pero 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.

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

Demostración:

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

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

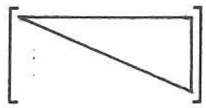
$$= \begin{array}{cl} \left[\begin{array}{c|c} A^t_{11} & A^t_{21} \\ \hline A^t_{12} & A^t_{22} \end{array}\right] & \begin{array}{l} r \\ n-r \end{array} \\ \begin{array}{cc} s & m-s \end{array} & \end{array} \qquad \blacksquare$$

Por medio del siguiente problema haremos 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.

Triangulacion de Matrices.

PROBLEMA:

Dada una matriz A_{nxn} encontrar una matriz L_{nxn} tal que

$LA =$  $= U$

donde L no debe ser trivial, es decir, no debe tener renglones con puros ceros. Como el producto se hace a la izquierda de A, este problema es equivalente a determinar combinaciones de los renglones de A, que sean los renglones de U.

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 L_1 tal que deje inveriante el primer renglón de A y anule la primera componente de los renglones restantes, es decir el resultado del producto L_1A debe tener la forma

$$L_1A = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \dots & * \end{bmatrix} = A_1$$

En el segundo paso buscaríamos una matriz L_2 tal que deje invariantes los dos primeros renglones de A_1 y anule la segunda componente de los restantes

$$L_2A_1 = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} = A_2$$

En general, en el k-ésimo paso buscaríamos una matriz L_k tal que

$$L_k A_{k-1} = \begin{bmatrix} * & * & \ldots & * & * & \ldots & * \\ 0 & * & \ldots & * & * & \ldots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & * & * & \ldots & * \\ 0 & 0 & \ldots & 0 & * & \ldots & * \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 & 0 & \ldots & * \end{bmatrix} = A_k .$$

Al terminar el paso $n - 1$, el proceso acaba, con A_{n-1} una matriz triangular superior, que por lo tanto, denotaremos como U. Así, tenemos que

$$L_1 A = A_1$$

$$L_2 A_1 = A_2$$

$$\vdots$$

$$L_k A_{k-1} = A_k = U$$

Si sustituímos A_1 en la segunda ecuación, A_2 en la tercera y así sucesivamente, hasta el paso n-1, obtenemos

$$(L_{n-1} L_{n-2} \ldots L_2 L_1) A = U$$

y si $L = L_{n-1} L_{n-2} \ldots L_2 L_1$ tenemos resuelto el problema de determinar L tal que

$$LA = U$$

Analicemos ahora el procedimiento anterior, en términos de matrices por bloques.

En el primer paso tenemos

$$A_1 = L_1 A = \begin{array}{c} 1 \\ n-1 \end{array} \left[\begin{array}{c|ccc} * & * & \cdots & * \\ \hline 0 & * & \cdots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \cdots & * \end{array}\right] = \left[\begin{array}{c|c} * & * \; \cdots \; * \\ \hline 0 & \overline{A}_1 \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array}$$

$$\begin{array}{cc} 1 & n-1 \end{array} \qquad\qquad \begin{array}{cc} 1 & n-1 \end{array}$$

En el segundo paso, para multiplicar L_2 por A_1 dividimos a L_2 por bloques

$$L_2 = \left[\begin{array}{c|c} L_{11} & L_{12} \\ \hline L_{21} & L_{22} \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array}$$

$$\begin{array}{cc} 1 & n-1 \end{array}$$

Por tanto

$$A_2 = L_2 A_1 = \begin{array}{c} 1 \\ n-1 \end{array} \left[\begin{array}{c|c} L_{11} & L_{12} \\ \hline L_{21} & L_{22} \end{array}\right] \left[\begin{array}{c|c} * & * \; \cdots \; * \\ \hline 0 & \overline{A}_1 \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array} = \left[\begin{array}{c|cccc} * & * & * & \cdots & * \\ \hline 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \cdots & * \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array}$$

$$\begin{array}{cc} 1 & n-1 \end{array} \qquad \begin{array}{cc} 1 & n-1 \end{array} \qquad \begin{array}{cc} 1 & n-1 \end{array}$$

Como queremos que el primer renglón de A_1 no se altere, nos conviene poner $L_{11} = I_1$ y $L_{12} = 0_{1x(n-1)}$. L_{21} queda forzado a ser $L_{21} = 0_{(n-1)x1}$ ya que el primer renglón no interviene en esta etapa. Si llamamos a $L_{22} = \overline{L}_2$ entonces

$$L_2 = \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & \overline{L}_2 \end{array}\right] \begin{array}{c} 1 \\ n-1 \end{array}$$

$$\begin{array}{cc} 1 & n-1 \end{array}$$

$$\therefore\ L_2A_1 = \begin{array}{c} 1 \\ n-1 \\ \\ \end{array}\!\!\begin{array}{c} \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & \overline{L}_2 \end{array}\right] \\ \begin{array}{cc} 1 & n-1 \end{array} \end{array} \begin{array}{c} \left[\begin{array}{c|c} * & * \ \ldots \ * \\ \hline 0 & \overline{A}_1 \end{array}\right] \\ \begin{array}{cc} 1 & n-1 \end{array} \end{array}\!\!\begin{array}{c} 1 \\ n-1 \\ \\ \end{array} = \begin{array}{c} \left[\begin{array}{c|c} * & * \ \ldots \ * \\ \hline 0 & \overline{L}_2\overline{A}_1 \end{array}\right] \\ \begin{array}{cc} 1 & n-1 \end{array} \end{array}\!\!\begin{array}{c} 1 \\ n-1 \\ \\ \end{array}$$

De aquí vemos que $\overline{L}_2$ debe ser tal que

$$\overline{L}_2\overline{A}_1 = \begin{array}{c} \left[\begin{array}{ccccc} * & * & * & \ldots & * \\ 0 & * & * & \ldots & * \\ 0 & * & * & \ldots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \ldots & * \end{array}\right] \\ n-1 \end{array} n-1 = \begin{array}{c} \left[\begin{array}{c|c} * & * \ \ldots \ * \\ \hline 0 & \overline{A}_1 \end{array}\right] \\ \begin{array}{cc} 1 & n-1 \end{array} \end{array}\!\!\begin{array}{c} 1 \\ n-1 \\ \\ \end{array}$$

Es decir, para determinar $\overline{L}_2$ tenemos que resolver otra vez el problema de hacer ceros abajo de un elemento como en el primer paso, pero en este caso las matrices con las que estamos trabajando son de una dimensión menor.

Veremos ahora qué forma tiene la matriz L_3.

$$\text{Sea}\quad L_3 = \begin{array}{c} 2 \\ n-2 \\ \\ \end{array}\!\!\begin{array}{c} \left[\begin{array}{c|c} L_{11} & L_{12} \\ \hline L_{21} & L_{22} \end{array}\right] \\ \begin{array}{cc} 2 & n-2 \end{array} \end{array}$$

Por lo tanto

$$L_3 = \begin{array}{c} 2 \\ n-2 \\ \\ \end{array}\!\!\begin{array}{c} \left[\begin{array}{c|c} L_{11} & L_{12} \\ \hline L_{21} & L_{22} \end{array}\right] \\ \begin{array}{cc} 2 & n-2 \end{array} \end{array} \begin{array}{c} \left[\begin{array}{c|c} & \\ \hline 0 & \overline{A}_2 \end{array}\right] \\ \begin{array}{cc} 2 & n-2 \end{array} \end{array}\!\!\begin{array}{c} 2 \\ n-2 \\ \\ \end{array} = \begin{array}{c} \left[\begin{array}{c|c} & \\ \hline 0 & \begin{array}{cccc} * & * & \ldots & * \\ 0 & * & \ldots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \ldots & * \end{array} \end{array}\right] \\ \begin{array}{cc} 2 & n-2 \end{array} \end{array}\!\!\begin{array}{c} 2 \\ n-2 \\ \\ \end{array}$$

Para que los dos primeros renglones de A2 no se modifiquen necesitamos que $L_{11} = I_2$ y que $L_{12} = 0_{2x(n-2)}$.

Además, L_{21} está forzado a ser $0_{(n-2)x2}$ para que el primer renglón y el segundo no intervengan en este paso.

$$\therefore\ L_3 = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & \bar{L}_3 \end{array}\right]$$

donde

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

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

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

y tal que

$$L_k A_{k-1} = \begin{array}{cc} \left[\begin{array}{c|c} \diagdown & \square \\ \hline 0 & \bar{A}_k \end{array}\right] & \begin{array}{l} k \\ n-k \end{array} \\ \begin{array}{cc} k & n-k \end{array} & \end{array} = A_k$$

$$\bar{L}_1\bar{A}_{k-1} = \left[\begin{array}{c|ccc} * & * & \dots & * \\ \hline 0 & * & \dots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \dots & * \end{array}\right] = \begin{array}{cc} \left[\begin{array}{c|c} * & * \ \dots \ * \\ \hline 0 & \bar{A}_k \end{array}\right] & \begin{array}{l} 1 \\ n-k \end{array} \\ \begin{array}{cc} 1 & n-k \end{array} & \end{array}$$

Hemos visto que el problema de encontrar una matriz L tal que $LA = [◹]$ se reduce a encontrar matrices $\overline{L}_k$ tal que

$$\overline{L}_k \overline{A}_{k-1} = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \dots & * \end{bmatrix}$$

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

Una vez determinadas las matrices $\overline{L}_k$ construímos

$$L_k = \left[\begin{array}{c|c} I_{k-1} & 0 \\ \hline 0 & \overline{L}_k \end{array}\right]$$

y la L que buscamos será $L = L_{n-1} L_{n-2} \dots L_2 L_1$ y tendremos

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

Nos queda ahora el siguiente

PROBLEMA:

Dada una matriz A_{mxn} encontrar una matriz L_{mxm} tal que

$$LA = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \dots & * \end{bmatrix}$$

$$\text{Sea} \quad \overline{L} = \begin{bmatrix} \overline{s}_1^t \\ \overline{s}_2^t \\ \vdots \\ \overline{s}_m^t \end{bmatrix}$$

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

$$\overline{L}\,\overline{A} = \begin{bmatrix} * & * & * & \ldots & * \\ 0 & * & * & \ldots & * \\ 0 & * & * & \ldots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \ldots & * \end{bmatrix}$$

Tenemos que

$$\overline{L}\,\overline{A} = \begin{bmatrix} \overline{s}_1^t \\ \overline{s}_2^t \\ \vdots \\ \overline{s}_m^t \end{bmatrix} \overline{A} = \begin{bmatrix} \overline{s}_1^t\,\overline{A} \\ \overline{s}_2^t\,\overline{A} \\ \vdots \\ \overline{s}_m^t\,\overline{A} \end{bmatrix} = \overline{L} \begin{bmatrix} \overline{r}_1^t \\ \overline{r}_2^t \\ \vdots \\ \overline{r}_m^t \end{bmatrix}$$

Ahora

$$\overline{s}_1^t\overline{A} = \overline{\lambda}_{11}\overline{r}_1^t + \overline{\lambda}_{12}\overline{r}_2^t + \ldots + \overline{\lambda}_{1m}\overline{r}_m^t$$

Como el primer renglón queda invariante, nos conviene escoger

$$\overline{\lambda}_{11} = 1$$

$$\overline{\lambda}_{12} = \overline{\lambda}_{13} = \ldots = \overline{\lambda}_{1m} = 0$$

$$\therefore \ \overline{s}_1^t = [1 \ 0 \ \ldots \ 0]$$

Para el segundo renglón tenemos

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

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

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

$$\therefore\ \bar{\lambda}_{21}\bar{r}_1^t = [\bar{\lambda}_{12}\bar{\alpha}_{11}\ \ \bar{\lambda}_{21}\bar{\alpha}_{12}\ \ldots\ \bar{\lambda}_{21}\bar{\alpha}_{1n}]$$

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

$$\bar{\lambda}_{21}\bar{r}_1^t + \bar{\lambda}_{22}\bar{r}_2^t = [\bar{\lambda}_{21}\bar{\alpha}_{11} + \bar{\lambda}_{22}\bar{\alpha}_{21}\ldots\ \bar{\lambda}_{21}\bar{\alpha}_{1n} + \bar{\lambda}_{22}\bar{\alpha}_{2n}]$$

De aquí vemos que si podemos hacer

$$\bar{\lambda}_{21}\bar{\alpha}_{11} + \bar{\lambda}_{22}\bar{\alpha}_{21} = 0$$

entonces podemos hacer

$$\bar{s}_2^t\bar{A} = \bar{\lambda}_{12}\bar{r}_1^t + \bar{\lambda}_{22}\bar{r}_2^t$$

tomando

$$\bar{\lambda}_{12} = \bar{\lambda}_{13} = \ldots = \bar{\lambda}_{1m} = 0$$

En efecto, podemos escoger $\bar{\lambda}_{21}$ y $\bar{\lambda}_{22}$

tales que

$$\bar{\lambda}_{21}\bar{\alpha}_{11} + \bar{\lambda}_{22}\bar{\alpha}_{21} = 0$$

ya que

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

y si $\bar{\alpha}_{11} \neq 0$ entonces

$$\bar{\lambda}_{21} = -\ \bar{\lambda}_{22}(\bar{\alpha}_{21}/\ \bar{\alpha}_{11})$$

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

$$\therefore\ \bar{\lambda}_{21} = -\ (\bar{\alpha}_{21}/\ \bar{\alpha}_{11})$$

$$\therefore \; \bar{s}_2^t = [\bar{\lambda}_{21} \; \bar{\lambda}_{22} \; \dots \; \bar{\lambda}_{2m}]$$

$$\therefore \; \bar{s}_2^t = [-\,(\bar{\alpha}_{21}/\,\bar{\alpha}_{11}) \;\; 1 \; 0 \; \dots \; 0]$$

similarmente se puede ver que

$$\bar{s}_3^t = [-\,(\bar{\alpha}_{31}/\,\bar{\alpha}_{11}) \;\; 0 \;\; 1 \; \dots \; 0]$$

$$\vdots$$

$$\bar{s}_m^t = [-\,(\bar{\alpha}_{m1}/\,\bar{\alpha}_{11}) \;\; 0 \;\; 0 \; \dots \; 1]$$

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

$$\bar{L} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\,(\bar{\alpha}_{21}/\,\bar{\alpha}_{11}) & 1 & 0 & \dots & 0 \\ -\,(\bar{\alpha}_{31}/\,\bar{\alpha}_{11}) & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\,(\bar{\alpha}_{m1}/\,\bar{\alpha}_{11}) & 0 & 0 & \dots & 1 \end{bmatrix}$$

Observar que $\bar{L}$ es una matriz triangular inferior.

Como se puede ver, el procedimiento es valido si $\bar{\alpha}_{11} \neq 0$. Más adelante veremos como proceder si no se cumple esta condición.

Ejemplo:

Llevar a la forma triangular a la matriz

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

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

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

■

Analicemos ahora con cuidado la estructura de las matrices $L_1,\dots,L_{n-1}$.

Por lo que acabamos de ver, $\overline{L}_1$ es de la forma.

$$\overline{L}_1 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_{21} & 1 & 0 & \dots & 0 \\ \lambda_{31} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n1} & 0 & 0 & \dots & 1 \end{bmatrix}$$

donde

$$\lambda_{i1} = -\,(\alpha_{i1} / \alpha_{11}) \qquad i = 2,\dots,n$$

Resulta entonces que la matriz L_1 difiere de la identidad sólo en la primera columna y λ_{i1} se escoge para que el producto nos dé:

$$L_1 A = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \dots & * \end{bmatrix} = A_1 = \begin{bmatrix} \alpha^1_{11} & \alpha^1_{12} & \dots & \alpha^1_{1n} \\ 0 & \alpha^1_{22} & \dots & \alpha^1_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & \alpha^1_{n2} & \dots & \alpha^1_{nn} \end{bmatrix}$$

En el segundo paso tendremos que

$$L_2 = \left[\begin{array}{c|c} I_1 & 0 \\ \hline 0 & \overline{L}_2 \end{array}\right]$$

donde $$\overline{L}_2 = \begin{bmatrix} \alpha^1_{22} & \alpha^1_{23} & \dots & \alpha^1_{2n} \\ \alpha^1_{32} & \alpha^1_{33} & \dots & \alpha^1_{3n} \\ \vdots & \vdots & & \vdots \\ \alpha^1_{n2} & \alpha^1_{n3} & \dots & \alpha^1_{nn} \end{bmatrix} = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & * & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & * & * & \dots & * \end{bmatrix}$$

por lo tanto $\overline{L}_2$ es de la forma

$$\overline{L}_2 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_{32} & 1 & 0 & \dots & 0 \\ \lambda_{42} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n2} & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$\lambda_{i2} = -(\alpha^1_{i2} / \alpha^1_{22}) \qquad i = 3,\dots,n$$

que al sustituir en L_2 nos queda

$$L_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_{32} & 1 & 0 & \dots & 0 \\ 0 & \lambda_{42} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & \lambda_{n2} & 0 & 0 & \dots & 1 \end{bmatrix}$$

Ahora la matriz L_2 difiere de la identidad en la segunda columna para que el producto haga cero la segunda componente.

$$L_2 A_1 = \begin{bmatrix} * & * & * & \dots & * \\ 0 & * & * & \dots & * \\ 0 & 0 & * & \dots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \dots & * \end{bmatrix} = A_2 = \begin{bmatrix} \alpha^2_{11} & \alpha^2_{12} & \alpha^2_{13} & \dots & \alpha^2_{1n} \\ 0 & \alpha^2_{22} & \alpha^2_{23} & \dots & \alpha^2_{2n} \\ 0 & 0 & \alpha^2_{33} & \dots & \alpha^2_{3n} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \alpha^2_{n3} & \dots & \alpha^2_{nn} \end{bmatrix}$$

Para el tercer paso

$$L_3 = \left[\begin{array}{c|c} I_2 & 0 \\ \hline 0 & \overline{L}_3 \end{array}\right]$$

con

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

Por lo tanto $\overline{L}_3$ es de la forma

$$\overline{L}_3 = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \lambda_{43} & 1 & 0 & \dots & 0 \\ \lambda_{53} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda_{n3} & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$L_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & \lambda_{43} & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda_{53} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \lambda_{n3} & 0 & 0 & \dots & 1 \end{bmatrix}$$

Nuevamente L_3 difiere de I en la tercera columna.
En general, en el k-ésimo paso vamos a tener

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

Por lo tanto tenemos

$$L_{n-1}L_{n-2}\ \ldots\ L_1A = [\nabla] = U$$

donde cada una de las L's es una matriz triangular inferior que difieren de la identidad en una sola columna.

El resultado anterior se puede expresar adecuadamente para obtener uno de los resultados más importantes de la teoría de matrices conocido como *la factorización* A = LU de la matriz A. Este se obtiene si notamos que cada una de las operaciones efectuada por cada matriz L_i se puede invertir, es decir, podemos encontrar una L_i^{-1} tal que su efecto sea el de deshacer el efecto de L_i.

Empecemos por L_1:

Como

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

si quisieramos de A_1 recuperar A, lo único que tenemos que hacer es:

1) Restar al 2^o 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

Regresando nuevamente al producto L_1A, vemos que las entradas λ_{i1} de L_1, son los múltiplos de r_1^t que se suma a los demás renglones para obtener A_1. Podemos concluir, por tanto, que L_1^{-1} debe tener la misma forma que L_1, solo que cambiando los signos de $\lambda_{21},\lambda_{31},\ldots,\lambda_{n1}$, es decir:

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

Podemos confirmar lo anterior si observamos que

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

$$= \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ -\lambda_{21} + \lambda_{43} & 1 & 0 & \dots & 0 \\ -\lambda_{31} + \lambda_{53} & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ -\lambda_{n1} + \lambda_{n3} & 0 & 0 & \dots & 1 \end{bmatrix}$$

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

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

$$L_2 = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & \lambda_{32} & 1 & & & \\ 0 & \lambda_{42} & 0 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & \lambda_{n2} & 0 & 0 & \dots & 1 \end{bmatrix}, \quad L_2^{-1} = \begin{bmatrix} 1 & & & & & \\ 0 & 1 & & & & \\ 0 & -\lambda_{32} & 1 & & & \\ 0 & -\lambda_{42} & 0 & 1 & & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & -\lambda_{n2} & 0 & 0 & \dots & 1 \end{bmatrix}$$

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

etcétera...

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

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, multiplicando sucesivamente por

$$L_{n-1}^{-1},\ L_{n-2}^{-1},\ \dots,\ L_1^{-1}$$

$$(L_{n-1}^{-1}L_{n-1})\, L_{n-2} \dots L_1A = I_n\, L_{n-2} \dots L_1A = L_{n-1}^{-1}U$$

$$(L_{n-2}^{-1}L_{n-2})\, L_{n-3} \dots L_1A = I_n\, L_{n-3} \dots L_1A = L_{n-2}^{-1}\, L_{n-1}^{-1}U$$

$$\vdots \qquad\qquad \vdots$$

$$(L_1^{-1}L_1)\; A = I_nA = L_1^{-1}\, L_2 \dots L_{n-2}^{-1}\, L_{n-1}^{-1}\; U$$

nos queda

$$A = L_1^{-1}\, L_2^{-1} \dots L_{n-2}^{-1}\, L_{n-1}^{-1}\; U$$

Como cada una de las matrices L_i es triangular inferior, tenemos que $L_1^{-1}\, L_2^{-1} \dots L_{n-2}^{-1}\, L_{n-1}^{-1}$ es triangular inferior.

Por lo tanto a la matriz A la podemos factorizar como A = LU donde L es triangular inferior y U triangular superior.

En la práctica, la factorización

$$A = LU$$

de una matriz A, presenta algunos problemas cuando se utiliza en la solución de sistemas, en especial cuando alguna de las α_{ii} se anula o tiene un valor peque~o comparado con los de la misma columna. Esta situación ha planteado la necesidad de buscar alguna variante de ésta factorización que evite los problemas en que puede incurrir. El resultado de esta busqueda se conoce como la factorización

$$PA = LU$$

donde P es una matriz de permutación que resulta del intercambio de algunos renglones durante el proceso.

Para terminar este capítulo, mencionaremos que el uso eficiente de las computadoras en la resolución de sistemas de ecuaciones, ha planteado tal necesidad de buscar factorizaciones adecuadas al problema y a la estructura de la matriz del sistema, que actualmente ya se han desarrollado varios tipos más, como la Factorización de Cholesky, Factorización QR.