

20 FEB 1978

e. 21

celeste

distribución interna

albino bocaz

3227

APUNTES SOBRE MATRICES

Serie B, nº 25

CENTRO DE INVESTIGACIONES  
DE MATEMÁTICAS  
E INGENIERÍA

CENTRO LATINOAMERICANO DE DEMOGRAFIA

Sede: José M. Infante, 9. Casilla 91  
Teléfono, 495071. Santiago, (Chile)

Subsede: Facultad de Ciencias Económicas y Sociales,  
Ciudad Universitaria Rodrigo Facio.  
Casilla, 5249. San José (Costa Rica)

1. El concepto de matriz

Matriz es un arreglo rectangular o cuadrado de números encerrados por un par de paréntesis, tal como

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{bmatrix} \quad B = \begin{bmatrix} 3 & 2 & 5 \\ 2 & 5 & 8 \\ 5 & 8 & 15 \end{bmatrix}$$

con los cuales es posible realizar una serie de operaciones algebraicas.

Toda matriz rectangular está constituida por un número de (m) filas (o líneas) y un número (n) de columnas. En tal caso se dice que se tiene una matriz rectangular de orden mxn.

Cuando el número de líneas es igual al número de columnas, se tiene una matriz cuadrada de orden nxn.

Cada uno de los mxn números de una matriz rectangular (mxn) o de los nxn números de una matriz cuadrada (nxn), recibe el nombre de "elemento de la matriz".

En el caso de una matriz, en la que no se especifican los valores de cada uno de los elementos en forma numérica, se adopta la denotación  $(a_{ij})$  para el elemento ubicado en la intersección de la línea (i) con la columna (j). De esa manera para una matriz rectangular (mxn), se tendrá la representación

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix} = (a_{ij})_{mxn}$$

Cada línea de la matriz recibe el nombre de "vector línea" y cada columna de la matriz, el de "vector columna".

De ese modo una matriz rectangular (mxn) está formada por m vectores líneas o por n vectores columnas.

Si  $A_1, A_2, \dots, A_n$  son los (n) vectores columnas de una matriz  $A_{mxn}$ , se tendrá otra representación de la matriz en la forma

$$A = [A_1, A_2, \dots, A_j, \dots, A_n]$$

teniendo el vector de la columna j, las siguientes componentes

$$A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ a_{3j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

y siendo la longitud de este vector igual a

$$\sqrt{a_{1j}^2 + a_{2j}^2 + a_{3j}^2 + \dots + a_{mj}^2}$$

En una matriz  $A_{m \times n}$  puede suceder que  $(r)$  vectores columna sean linealmente independientes y que los otros  $(n-r)$  puedan deducirse como función lineal de los  $(r)$  primeros. En ese caso se dirá que la matriz  $A_{m \times n}$  tiene rango  $(r)$ .

De ese modo, la matriz

$$A = \begin{bmatrix} 3 & -1 & 2 & 1 & 0 & 0 \\ -2 & 4 & 0 & 0 & 1 & 0 \\ -4 & 3 & 8 & 0 & 0 & 1 \end{bmatrix} = [A_1, A_2, A_3, A_4, A_5, A_6]$$

es de rango 3 porque a base de los vectores unitarios  $A_4, A_5, A_6$  pueden expresarse los 3 primeros

$$\begin{aligned} A_1 &= 3A_4 - 2A_5 - 4A_6 \\ A_2 &= -A_4 + 4A_5 + 3A_6 \\ A_3 &= 2A_4 + 8A_6 \end{aligned}$$

Siempre en una matriz de rango  $r < n$ , es posible elegir arbitrariamente  $r$  vectores cualesquiera (base) y expresar los  $(n-r)$  vectores restantes en función de los primeros.

Así, en el caso anotado anteriormente podemos expresar los vectores  $A_4, A_5, A_6$  en función de los vectores  $A_1, A_2, A_3$ , pero evidentemente que la solución ahora no se hace en la forma directa como se hizo anteriormente.

$$\begin{aligned} \text{Resulta } 100A_4 &= 32A_1 + 16A_2 + 10A_3 \\ 100A_5 &= 14A_1 + 32A_2 - 5A_3 \\ 100A_6 &= -8A_1 - 4A_2 + 10A_3 \end{aligned}$$

Indudablemente que hay matrices  $(A)$  en que no vienen incluidos vectores unitarios como los indicados en la primera parte del ejemplo anterior, y de todas maneras se hace necesario buscar el rango de la matriz.

Así, por ejemplo la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ 3 & 4 & 1 & 2 \\ 2 & 1 & 2 & 1 \end{bmatrix} = [A_1, A_2, A_3, A_4]$$

es de rango 3, o sea que un vector cualquiera, p. ej.,  $A_4$  es función lineal de  $A_1, A_2, A_3$ .

En este caso, y siempre que se trate de matrices de orden reducido podemos escribir

$$A_4 = c_1 A_1 + c_2 A_2 + c_3 A_3$$

y determinar los coeficientes  $c_i$  a base de las ecuaciones

$$4 = c_1 + 2c_2 + 3c_3$$

$$1 = c_3$$

$$2 = 3c_1 + 4c_2 + c_3$$

$$1 = 2c_1 + c_2 + 2c_3$$

usando las 3 primeras ecuaciones, se tiene  $c_1 = -1; c_2 = 1; c_3 = 1$ , con lo cual

$$A_4 = -A_1 + A_2 + A_3$$

ya que la última ecuación se satisface igualmente con esos valores de  $c_i$ .

Nota 1: La determinación del rango de una matriz es asunto de bastante importancia, que se verá más adelante cuando se demuestre que siempre es posible descomponer una matriz (A) en la suma de matrices de rango 1.

Para el caso siguiente, de una matriz de 8x8

$$A = \begin{bmatrix} 0.58 & 0.63 & 0.46 & 0.51 & 0.35 & 0.18 & 0.45 & 0.24 \\ 0.63 & 0.81 & 0.36 & 0.54 & 0.45 & 0.00 & 0.27 & 0.00 \\ 0.46 & 0.36 & 0.52 & 0.42 & 0.20 & 0.36 & 0.60 & 0.48 \\ 0.51 & 0.54 & 0.42 & 0.45 & 0.30 & 0.18 & 0.42 & 0.24 \\ 0.35 & 0.45 & 0.20 & 0.30 & 0.25 & 0.00 & 0.15 & 0.00 \\ 0.18 & 0.00 & 0.36 & 0.18 & 0.00 & 0.36 & 0.48 & 0.48 \\ 0.45 & 0.27 & 0.60 & 0.42 & 0.15 & 0.48 & 0.73 & 0.64 \\ 0.24 & 0.00 & 0.48 & 0.24 & 0.00 & 0.48 & 0.64 & 0.64 \end{bmatrix}$$

se puede probar que esta matriz es de rango 2.

Nota 2: También puede presentarse el problema de dar una matriz, en que se encuentran indeterminados los términos de la diagonal principal y se desee qué valor deben tener esos elementos para que la matriz tenga un rango determinado.

Este último tipo de problema es uno de los problemas importantes del análisis factorial agravado por el hecho que los elementos  $a_{ij}$  para  $i \neq j$ , están sujetos a error de muestreo y debe buscarse estimaciones adecuadas de los elementos  $a_{ij}$ . En análisis factorial estos elementos reciben el nombre de "comunalidades".

También en una matriz, los vectores columna que la integran pueden presentar la particularidad de ser ortogonales, esto es, que la suma de los productos correspondientes sea nula. Así por ejemplo en la matriz

$$A = \begin{bmatrix} 2/\sqrt{14} & -4/\sqrt{42} & 1/\sqrt{3} \\ 1/\sqrt{14} & 5/\sqrt{42} & 1/\sqrt{3} \\ 3/\sqrt{14} & 1/\sqrt{42} & -1/\sqrt{3} \end{bmatrix} = [A_1, A_2, A_3]$$

se cumple esa condición. En este caso se dice que la matriz es "ortogonal".

Esta condición de ortogonalidad de la matriz tiene ventajas operacionales, como se verá más adelante en las aplicaciones numéricas.

## 2. Tipo de matrices

Aparte de los casos más generales de matrices mencionadas en 1), esto es, de matrices rectangulares y matrices cuadradas, existen diversos tipos de matrices que tienen alguna característica especial que las distingue de otra. Revisaremos las más frecuentemente encontradas, en un orden determinado.

### a) Matriz nula

es una matriz, en que cada elemento  $a_{ij}$  es igual a cero.

Cuando  $A$  es una matriz nula y no hay confusión en cuanto a su orden, se escribe  $A=0$  en lugar de un arreglo de  $m \times n$  ceros.

### b) Matriz identidad

es una matriz cuadrada  $A$  cuyos elementos son nulos, excepto en la diagonal que son iguales a 1.

De esa manera la matriz identidad  $I_3$  es la matriz

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

### c) Matriz escalar

si todos los elementos son nulos, excepto los elementos de la diagonal que son todos iguales a un escalar  $c$ , una matriz recibe el nombre de matriz escalar.

Así, por ejemplo, la matriz

$$A = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & c & 0 & 0 \\ 0 & 0 & c & 0 \\ 0 & 0 & 0 & c \end{bmatrix}$$

es la matriz escalar de orden 4

d) Matriz diagonal

es una matriz cuadrada en que todos los elementos son nulos, excepto los de la diagonal que son escalares de valor diferente.

Por ejemplo, la matriz

$$D = \begin{bmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{bmatrix}$$

es una matriz diagonal.

e) Matriz triangular

es una matriz cuadrada, en la que los elementos sobre la diagonal o bajo ella son nulos. De esa manera, hay dos tipos de matrices triangulares: las triangulares superiores y las triangulares inferiores.

Las matrices A y B siguientes

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 3 & 5 & 0 \\ 4 & 6 & 7 \end{bmatrix}$$

son de los tipos indicados.

Este tipo de matrices tiene la particularidad que es bastante sencillo calcular sus inversos y son las matrices a que se llega en la resolución de sistemas de ecuaciones simultáneas cuando se usa el método de Gauss-Jordan.

Por otra parte, como se verá más adelante, toda matriz no singular, puede expresarse como el producto de dos matrices triangulares: una inferior y la otra superior.

f) Matriz estocástica

es una matriz en la que los elementos  $a_{ij}$  son probabilidades y por lo tanto la suma de los elementos de cada línea es igual a 1.

Por ejemplo, la matriz

$$A = \begin{bmatrix} 1-a & 0 & a & 0 \\ b & 0 & 1-b & 0 \\ 0 & 1-c & 0 & c \\ 0 & d & 0 & 1-d \end{bmatrix}$$

es una matriz estocástica.

g) Matriz de permutación

es una matriz cuadrada en la que existe un elemento igual a 1 en cada línea y en cada columna.  
Las matrices

$$P_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad P_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad P_3 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad P_4 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$P_5 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

son matrices de permutación. La matriz identidad es un caso particular.

h) Matriz monomial

es una matriz cuadrada en la que en cada línea y en cada columna, hay un solo elemento no nulo.  
Las matrices

$$M_1 = \begin{bmatrix} 0 & 0 & 2 \\ -3 & 0 & 0 \\ 0 & 4 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 8 & 0 \\ 0 & 0 & 4 \\ -5 & 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0 & 2 \\ -3 & 0 \end{bmatrix}$$

son matrices monomiales.

i) Matriz escalonada

es una matriz, cuadrada o rectangular, en la que en cada línea el primer elemento no nulo que interviene es 1 y el número de ceros va en aumento a medida que pasamos de una fila a la siguiente.  
Por ejemplo, la matriz

$$A = \begin{bmatrix} 0 & 1 & d_{13} & d_{14} & d_{15} & d_{16} & d_{17} \\ 0 & 0 & 0 & 1 & d_{25} & d_{26} & d_{27} \\ 0 & 0 & 0 & 0 & 1 & d_{36} & d_{37} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

j) Matriz elemental

es una matriz que, aplicada a una matriz (A) realiza una "operación elemental" en esa matriz.



Las operaciones elementales que pueden realizarse en una matriz son:

- a) eliminar todos los elementos de la matriz, excepto los de la columna  $j$ . La matriz se denota por  $E_{jj}$  y es una matriz de rango  $1$ .
- b) eliminar todos los elementos de la matriz, excepto los de la línea  $i$ . La matriz se denota por  $E_{ii}$  y es una matriz de rango  $1$ .
- c) eliminar todos los elementos de la matriz, excepto los de la columna  $j$  pero trasladándolos a la columna  $i$ . La matriz es  $E_{ji}$  y es una matriz de rango  $1$ .
- d) eliminar todos los elementos de la matriz, excepto los de la línea  $i$  pero llevándolos a la línea  $j$ . La matriz es  $E_{ij}$  y es una matriz de rango  $1$ .
- e) intercambiar los elementos de las columnas  $i$  y  $j$ . La matriz la denotamos por  $H_{ij}$  y es una matriz de rango  $n$ .
- f) intercambiar los elementos de las líneas  $i$  y  $j$ . La matriz la denotamos por  $K_{ij}$  y es una matriz de rango  $n$ .
- g) multiplicar todos los elementos de la columna  $j$  por un escalar  $c$ . La matriz es  $K_j(c)$  y es una matriz de rango  $n$ .
- h) multiplicar todos los elementos de la línea  $i$  por un escalar  $c$ . La matriz es  $H_i(c)$  y es de rango  $n$ .
- i) agregar a la columna  $j$ ,  $c$  veces la columna  $i$ . La matriz es  $K_{ji}(c)$  y es una matriz de rango  $n$ .
- j) agregar a la línea  $i$ ,  $c$  veces la línea  $j$ . La matriz es  $H_{ij}(c)$  y es de rango  $n$ .

Para el caso de una matriz (A) de  $3 \times 3$ , algunas de estas matrices elementales, son las siguientes:

$$E_{22} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad E_{33} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad E_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$H_{12} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = K_{12} \quad H_{23}(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} = K_{32}(c)$$

$$H_3(c) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & c \end{bmatrix} = K_3(c)$$

Nota: Como se verá más adelante

$$H_{12} = E_{12} + E_{21} + E_{33}$$

$$H_{23}(c) = E_{12} + E_{22} + cE_{23} + E_{33}$$

$$H_3(o) = E_{12} + E_{22} + cE_{33}$$

k) Matriz simétrica

es una matriz cuadrada en la que el vector de la columna (i) tiene los mismos elementos que el vector de la línea (i).

La simetría de la matriz puede observarse trazando la diagonal principal y observando los elementos que quedan a ambos lados de esa diagonal. Si los elementos correspondientes son iguales, la matriz es simétrica. También puede saberse que una matriz es simétrica, si se la "traspone", esto es, si se escribe una matriz  $A'$  en la que las líneas son las columnas de la matriz primitiva A.

Por ejemplo, las matrices

$$A = \begin{bmatrix} 1 & 2 & 3 & 1 \\ 2 & 4 & 5 & 6 \\ 3 & 5 & 2 & 3 \\ 4 & 6 & 3 & 4 \end{bmatrix}$$

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

son matrices simétricas.

l) Matriz particionada

una matriz A puede subdivirse mediante líneas y columnas en una serie de submatrices. La matriz recibe en ese caso el nombre de "matriz particionada". La partición puede hacerse de diferentes modos buscando la manera que posteriormente permita trabajar más cómodamente con matrices de órdenes más reducidos. La matriz

$$A = \begin{array}{|c|c|c|c|c|c|} \hline 1 & 2 & 3 & 4 & 5 & 6 \\ \hline 2 & 3 & 4 & 5 & 6 & 7 \\ \hline 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 5 & 6 & 7 & 8 & 9 & 4 \\ \hline 6 & 7 & 8 & 9 & 4 & 1 \\ \hline \end{array} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{bmatrix}$$

es una matriz 6x6 que se la ha particionado en 4 submatrices, 2 vectores columna, 2 vectores línea y 1 escalar, mediante 2 líneas verticales y 2 líneas horizontales.

m) Matriz polinomial

es una matriz en la que los elementos  $(a_{ij})$  son funciones de un cierto parámetro  $\lambda$ . De esa manera la matriz

$$A(\lambda) = \begin{bmatrix} \lambda^2 + \lambda + 1 & \lambda^4 + 2\lambda^3 + 3\lambda^2 + 5 \\ \lambda^3 - 4 & \lambda^3 - 3\lambda^2 \end{bmatrix}$$

es una matriz polinomial de grado 2. Esta matriz puede expresarse en la forma

$$A(\lambda) = A_4 \lambda^4 + A_3 \lambda^3 + A_2 \lambda^2 + A_1 \lambda + A_0$$

siendo los coeficientes  $(A_i)$  matrices cuadradas  $2 \times 2$ .

n) Matriz singular

es una matriz cuadrada, cuyo determinante es nulo. La notación para determinante es  $|A|$  o  $\det A$ . Para que una matriz sea singular deberá haber al menos una línea o una columna de elementos que sean funciones lineales de los elementos de las otras líneas o columnas. Por ejemplo, la matriz

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 5 & 7 \end{bmatrix}$$

es singular porque la tercera línea es la suma de las dos primeras o la tercera columna es igual al doble de la segunda columna menos la primera columna. Cuando la matriz tiene todas sus líneas y columnas independientes linealmente, se dice que es una matriz no-singular.

o) Matriz idempotente

Cualquier matriz A puede elevarse a la potencia  $(k)$ , obteniéndose en tal caso la matriz  $A^k$ . (Más adelante se verá la importancia de la elevación a potencia de una matriz y la forma cómo se realiza).

Puede suceder que al elevar a la potencia  $(k+1)$  la matriz A, se obtenga la misma matriz A. En tal caso se dice que la matriz A es periódica de período k.

El período de referencia puede ser 1, en cuyo caso la matriz recibe el nombre de matriz idempotente.

También puede suceder que para un cierto  $(p)$  entero y positivo, se tenga que  $A^p = 0$ , en cuyo caso se dice que la matriz es nilpotente, de índice p.

Finalmente puede suceder que la matriz A, elevada al cuadrado conduzca a la matriz identidad, en cuyo caso la matriz es una matriz involutoria. La matriz involutoria es una matriz que es igual a su inversa.

Como ejemplo de las matrices citadas anteriormente, se tiene

$$A = \begin{bmatrix} 1 & -2 & -6 \\ -3 & 2 & 9 \\ 2 & 0 & -3 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -3 & -4 \\ -1 & 3 & 4 \\ 1 & -3 & -4 \end{bmatrix}$$
$$C = \begin{bmatrix} 2 & -3 & -5 \\ -1 & 4 & 5 \\ 1 & -3 & -4 \end{bmatrix} \quad D = \begin{bmatrix} 4 & 3 & 3 \\ -1 & 0 & -1 \\ -4 & -4 & -3 \end{bmatrix}$$

siendo A una matriz periódica de período 2, la matriz B es nilpotente, la matriz C es idempotente y la matriz D es involutoria.

p) Matriz de variancia-covariancia

es una matriz cuadrada en la que los elementos ( $a_{ij}$ ) son las covariancias entre las variables  $x_i$  y  $x_j$ .

Así, por ejemplo, para el caso de 4 variables aleatorias, se tendrá la siguiente matriz simétrica

$$V = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_{44} \end{bmatrix}$$

q) Matriz Jacobiana

Si se tiene  $n$  variables  $x_i$  y se reemplazan por  $(n)$  variables  $y_i$  ligadas linealmente, la matriz cuadrada  $n \times n$ , en la que los elementos ( $a_{ij}$ ) son las derivadas parciales de primer orden  $\frac{\partial x_i}{\partial y_j}$  se denomina matriz Jacobiana.

Por ejemplo, si se tiene las relaciones lineales

$$x_1 = y_1 + 2y_2 + 4y_3$$

$$x_2 = y_2 + 4y_3$$

$$x_3 = y_3$$

que permiten cambiar las variables  $x_i$  por las variables  $y_i$ , la matriz Jacobiana será

$$J = \begin{bmatrix} 1 & 2 & 4 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{bmatrix}$$

El valor absoluto del determinante de esta matriz se denomina "Jacobiano de la transformación". Para el caso recién indicado este Jacobiano es igual a 1.

r) Matriz Hessiana

Si se tiene una función  $F$  de  $(n)$  variables  $(x_i)$ , la matriz cuyo elemento  $(a_{ij})$  es igual a

$$\left( \frac{\partial^2 F}{\partial x_i \partial x_j} \right)$$

recibe el nombre de matriz Hessiana. El valor del determinante de esta matriz, para los valores que anulan

las primeras derivadas  $\left( \frac{\partial^2 F}{\partial x_i \partial x_j} \right)$  nos permitirá ver

si la función  $F$  pasa por un máximo o por un mínimo. Por ejemplo, si la función  $F$  es la siguiente

$$F = x_1^2 - 2x_1 + x_2^2 - 6x_2 + 16$$

la matriz Hessiana es igual a

$$H = \begin{bmatrix} \frac{\partial^2 F}{\partial x_1^2} & \frac{\partial^2 F}{\partial x_1 \partial x_2} \\ \frac{\partial^2 F}{\partial x_2 \partial x_1} & \frac{\partial^2 F}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

s) Matriz Grammiana

es la matriz que resulta de multiplicar una matriz  $(A)$  por su traspuesta  $(A')$ . Si al realizarse el producto, la matriz Grammiana se reduce a una matriz identidad, se dice que la matriz  $A$  es ortogonal.

Así, por ejemplo, la matriz A

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{bmatrix}$$

tiene una matriz Grammiana igual a

$$B = I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

### 3. Operaciones en matrices

#### Suma de matrices

Si se tiene las matrices  $A = (a_{ij})$ ;  $B = (b_{ij})$ , de  $m \times n$  elementos, su suma (diferencia),  $A \pm B$  es una matriz  $C$  de  $m \times n$  elementos, en que cada elemento de  $C$  es la suma (diferencia) de los correspondientes elementos de  $A$  y  $B$ .

En símbolos, si

$$A = (a_{ij})_{m \times n} ; B = (b_{ij})_{m \times n}$$

entonces

$$A \pm B = C = (c_{ij})_{m \times n} = (a_{ij} \pm b_{ij})_{m \times n}$$

Así, por ejemplo, si se tiene las matrices

$$A = \begin{bmatrix} 283.9 & 165.6 \\ 333.3 & 126.6 \\ 247.6 & 81.1 \\ 170.7 & 55.3 \\ 89.5 & 26.4 \\ 37.4 & 10.4 \end{bmatrix} \quad B = \begin{bmatrix} 198.1 & 22.3 \\ 151.6 & 12.1 \\ 117.2 & 8.5 \\ 97.1 & 7.7 \\ 63.3 & 5.1 \\ 35.0 & 3.4 \end{bmatrix}$$

que representan la distribución de la PEA -Chile, 1960- en las edades 15-24; 25-34, .... 65 y más, en el área urbana y rural respectivamente, la población económicamente activa por grupos de edad solamente es la matriz  $(A + B)$

$$C = \begin{bmatrix} 482.0 & 187.9 \\ 484.9 & 138.7 \\ 364.8 & 89.6 \\ 267.8 & 63.0 \\ 152.8 & 31.5 \\ 72.4 & 13.8 \end{bmatrix}$$

Caso especial

Si al hacer la diferencia (A - B) se obtiene la matriz nula 0, entonces  $a_{ij} - b_{ij} = 0$  ó  $a_{ij} = b_{ij}$ . En ese caso, se dice que las matrices A y B son iguales.

Si se acepta que 3 matrices A, B, C son del mismo orden, o sea, matrices m x n, puede establecerse que

- a)  $A + B = B + A$ ; ley conmutativa
- b)  $A + (B + C) = (A + B) + C$ ; ley asociativa
- c)  $k(A + B) = kA + kB = (A + B)k$ , siendo k un escalar.
- d) existe una matriz D tal que  $A + D = B$
- e)  $\text{Traza}(A + B) = \text{traza } A + \text{traza } B$ .

Multiplicación de matrices

Supongamos que disponemos de 2 tablas estadísticas de doble entrada en las que se indica la distribución de la población según edad y actividad (tabla 1) y según actividad y migración (tabla 2).

Tabla 1

		Actividad				
Edades		1	2	3	4	Total
1	$a_{11} y_1$	$a_{12} y_2$	$a_{13} y_3$	$a_{14} y_4$		$x_1$
2	$a_{21} y_1$	$a_{22} y_2$	$a_{23} y_3$	$a_{24} y_4$		$x_2$
3	$a_{31} y_1$	$a_{32} y_2$	$a_{33} y_3$	$a_{34} y_4$		$x_3$
4	$a_{41} y_1$	$a_{42} y_2$	$a_{43} y_3$	$a_{44} y_4$		$x_4$
5	$a_{51} y_1$	$a_{52} y_2$	$a_{53} y_3$	$a_{54} y_4$		$x_5$
6	$a_{61} y_1$	$a_{62} y_2$	$a_{63} y_3$	$a_{64} y_4$		$x_6$
Total		$y_1$	$y_2$	$y_3$	$y_4$	

Tabla 2

Actividad	Migración		Total
	1	2	
1	$b_{11} z_1$	$b_{12} z_2$	$y_1$
2	$b_{21} z_1$	$b_{22} z_2$	$y_2$
3	$b_{31} z_1$	$b_{32} z_2$	$y_3$
4	$b_{41} z_1$	$b_{42} z_2$	$y_4$
Total	$z_1$	$z_2$	

Si denotamos por  $x$ ,  $y$ ,  $z$  los siguientes vectores columnas

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \quad z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$

y por  $A$  y  $B$  las matrices

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$$

se puede pensar, en resumir las tablas 1 y 2, en la forma matricial

$$\begin{aligned} Ay &= x \quad (1) \\ Bz &= y \quad (2) \end{aligned}$$

por lo cual tendríamos que aceptar alguna regla para multiplicar las matrices  $A$  y  $B$  por los vectores  $y$ ,  $z$ , respectivamente.

Esta regla sería: Cada línea de la matriz se multiplica por la columna que forma el vector y se consigue un nuevo vector columna con el mismo número de líneas de la matriz.



A base de las tablas 1 y 2 podemos deducir una tabla 3, del siguiente tipo

Tabla 3

Edades	Migración		Total
	1	2	
1	$c_{11} z_1$	$c_{12} z_2$	$x_1$
2	$c_{21} z_1$	$c_{22} z_2$	$x_2$
3	$c_{31} z_1$	$c_{32} z_2$	$x_3$
4	$c_{41} z_1$	$c_{42} z_2$	$x_4$
5	$c_{51} z_1$	$c_{52} z_2$	$x_5$
6	$c_{61} z_1$	$c_{62} z_2$	$x_6$
	$z_1$	$z_2$	

que podría, también, resumirse en la relación

$$Cz = x \quad (3)$$

Reemplazando en la relación (1) el valor (y) tomado de la relación (2) se tiene:

$$Ay = x$$

$$A(Bz) = x$$

con lo cual se llegaría a que

$$C = AB$$

o sea,

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \\ c_{51} & c_{52} \\ c_{61} & c_{62} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix}$$

pero de tabla 1 tenemos que

$$x_1 = a_{11} y_1 + a_{12} y_2 + a_{13} y_3 + a_{14} y_4$$

y reemplazando los valores ( $y_i$ ) de tabla 2, se tiene

$$x_1 = a_{11} (b_{11} z_1 + b_{12} z_2) + a_{12} (b_{21} z_1 + b_{22} z_2) +$$

$$+ a_{13} (b_{31} z_1 + b_{32} z_2) + a_{14} (b_{41} z_1 + b_{42} z_2)$$

$$x_1 = (a_{11} b_{11} + a_{12} b_{21} + a_{13} b_{31} + a_{14} b_{41}) z_1 +$$

$$+ (a_{11} b_{12} + a_{12} b_{22} + a_{13} b_{32} + a_{14} b_{42}) z_2$$

$$x_1 = c_{11} z_1 + c_{12} z_2$$

notando por lo tanto que  $c_{11}$  se consigue multiplicando la primera línea de A por la primera columna de B. De la misma manera  $c_{12}$ , es el resultado de la multiplicación de la primera línea de A y de la segunda columna de B. De allí se deduce que para poder multiplicar 2 matrices es necesario que el número de columnas de la primera matriz sea igual al número de líneas de la segunda matriz. De ese modo una matriz  $A_{m \times p}$  puede multiplicarse por una matriz  $B_{p \times n}$ , teniéndose entonces que:

Por producto AB en ese orden de la matriz  $A = (a_{ij})_{m \times p}$  y de la matriz

$B = (b_{ij})_{p \times n}$  se entiende la matriz  $C = (c_{ij})_{m \times n}$ , en que

$$c_{ij} = a_{i1} b_{1j} + a_{i2} b_{2j} + \dots + a_{ip} b_{pj} = \sum_{k=1}^p a_{ik} b_{kj}$$

$$\begin{pmatrix} i = 1, 2, \dots, m \\ j = 1, 2, \dots, n \end{pmatrix}$$

Por ejemplo, si se tiene las matrices

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

$$B = \begin{bmatrix} 2 & 5 & -7 & 1 \\ 5 & -1 & 5 & -2 \\ -7 & 5 & 11 & 10 \\ 1 & -2 & 10 & 5 \end{bmatrix}$$

su producto  $\underline{AB}$  es igual a

$$C = \begin{bmatrix} 18 & 0 & 0 & 0 \\ 0 & 18 & 0 & 0 \\ 0 & 0 & 18 & 0 \\ 0 & 0 & 0 & 18 \end{bmatrix} = 18 I_4$$

Si se tiene 3 matrices A, B, C que pueden sumarse y multiplicarse, se cumple que

- a)  $A(B + C) = AB + AC$ ; 1.<sup>a</sup> ley distributiva
- b)  $(A + B)C = AC + BC$ ; 2.<sup>a</sup> ley distributiva
- c)  $A(BC) = (AB)C$  ; ley asociativa

teniéndose además

- d)  $AB \neq BA$
- e)  $\det(AB) = (\det A)(\det B)$
- f)  $AB = 0$ , no implica necesariamente  $A = 0$  ó  $B = 0$
- g)  $AB = AC$ , no implica necesariamente  $B = C$

Producto de matrices particionadas

Podemos subdividir la matriz (A) de la siguiente manera

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \\ a_{51} & a_{52} & a_{53} & a_{54} \\ a_{61} & a_{62} & a_{63} & a_{64} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \\ A_{31} & A_{32} \end{bmatrix}$$

y la matriz B, en la forma

$$B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \\ b_{41} & b_{42} \end{bmatrix} = \begin{bmatrix} B_{11} \\ B_{21} \end{bmatrix}$$

el producto (AB) es igual a

$$AB = \begin{bmatrix} A_{11} B_{11} + A_{12} B_{21} \\ A_{21} B_{11} + A_{22} B_{21} \\ A_{31} B_{11} + A_{32} B_{21} \end{bmatrix}$$

y en general si se tiene dos matrices (A) y (B) particionadas en la forma

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

el producto (AB) de ellas es igual a

$$AB = \begin{bmatrix} A_{11} B_{11} + A_{12} B_{21} & A_{11} B_{12} + A_{12} B_{22} \\ A_{21} B_{11} + A_{22} B_{21} & A_{21} B_{12} + A_{22} B_{22} \end{bmatrix}$$

cumpléndose la misma regla de multiplicación de dos matrices en la que no se ha realizado la partición.

Ejemplo 1: Calcular el producto (AB) de las matrices particionadas, así

$$\begin{aligned} A_{11} &= \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} & A_{12} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} & B_{11} &= \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} & B_{12} &= \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ A_{21} &= \begin{bmatrix} 1 & 0 \end{bmatrix} & A_{22} &= \begin{bmatrix} 1 \end{bmatrix} & B_{21} &= \begin{bmatrix} 2 & 3 & 1 \end{bmatrix} & B_{22} &= \begin{bmatrix} 2 \end{bmatrix} \end{aligned}$$

Solución

$$\begin{aligned} A_{11} B_{11} &= \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 3 \\ 7 & 5 & 5 \end{bmatrix} & A_{12} B_{21} &= 0 & A_{11} B_{12} &= 0 \\ & A_{12} B_{22} &= 0 & A_{22} B_{22} &= 0 \\ A_{21} B_{11} &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} & A_{21} B_{11} + A_{22} B_{21} &= (3, 4, 2) \\ A_{22} B_{21} &= (1) (2 \ 3 \ 1) = (2 \ 3 \ 1) & A_{22} B_{22} &= 2 \end{aligned}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 3 & 2 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 1 & 0 \\ 2 & 3 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 & 3 & 0 \\ 7 & 5 & 5 & 0 \\ 3 & 4 & 2 & 2 \end{bmatrix}$$

Ejemplo 2: Determinar el cuadrado de la matriz

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ a & b & -1 & 0 \\ c & d & 0 & -1 \end{bmatrix}$$

Solución: Esta matriz puede escribirse en la forma particio-  
nada

$$A = \begin{bmatrix} I_2 & 0 \\ A_{21} - I_2 \end{bmatrix}$$

y por lo tanto el cuadrado vale

$$\begin{bmatrix} I_2 & 0 \\ A_{21} - I_2 \end{bmatrix} \begin{bmatrix} I_2 & 0 \\ A_{21} - I_2 \end{bmatrix} = \begin{bmatrix} I_2^2 + 0 & 0 \\ 0 & I_2^2 + 0 \end{bmatrix} = \begin{bmatrix} I_2 & 0 \\ 0 & I_2 \end{bmatrix} = I_4$$

Problema. Si las matrices (A) y (B) han sido particionadas en la forma

$$A = \left[ \begin{array}{cc|cc} 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 0 \\ \hline 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{array} \right] \quad B = \left[ \begin{array}{cc|cc} 1 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 2 & 2 \end{array} \right]$$

determinar los productos (AB) y (BA).

Casos especiales de  
multiplicación de  
matrices

Aunque se ha dado la regla general para multiplicar dos matrices (A) y (B) y se ha deducido también de allí la regla para multiplicar matrices particionadas, es conveniente indicar una serie de casos de multiplicación, en que las matrices que intervienen como factores son matrices especiales y aún pueden reducirse al producto de vectores.

Consideraremos los 8 casos siguientes:

- Caso 1: Multiplicación de una matriz (A) por la matriz identidad (I)
- Caso 2: Multiplicación de una matriz por un vector columna
- Caso 3: Multiplicación de un vector línea por una matriz (A)
- Caso 4: Multiplicación de un vector línea por un vector columna
- Caso 5: Multiplicación de un vector columna por un vector línea
- Caso 6: Multiplicación de una matriz (A) por su matriz inversa
- Caso 7: Multiplicación de una matriz (A) por una matriz diagonal
- Caso 8: Multiplicación de una matriz (A) por una matriz elemental

Caso 1. Multiplicación de una matriz (A) por la matriz identidad (I)

En este caso, se obtiene la matriz inicial A, ya se pre-multiplicue o postmultiplique por la matriz identidad.

En símbolos  $IA = AI = A$

Caso 2. Multiplicación de una matriz (A) por un vector columna (x)

En este caso, se obtiene como resultado otro vector columna, cuyo número de elementos es igual al número de líneas de la matriz (A).

En símbolos

$$Ax = y$$

siendo  $A_{m \times n}$ ,  $x_{n \times 1}$ ,  $y_{m \times 1}$

Nota: Si el vector (x) es el vector unitario ( $e_i$ ), el vector (y) resultante es el vector columna  $A_i$  de la matriz A.

Caso 3. Multiplicación de un vector línea (x') por una matriz (A).

En este caso el resultado es un vector línea y'.

En símbolos  $x' A = y'$

siendo  $A_{m \times n}$ ,  $x'_{1 \times m}$ ,  $y'_{1 \times n}$

Nota: Si el vector  $x'$  es el vector unitario  $e'_i$ , el vector resultante es el vector línea  $A'_i$  de la línea (i) de la matriz (A).

Caso 4. Multiplicación de un vector línea por un vector columna.

En este caso el producto es un escalar. De ese modo si

$$x' = (x_1, x_2, \dots, x_n) \quad y' = (y_1, y_2, \dots, y_n)$$

los productos ( $x' y$ ) e ( $y' x$ ) son iguales al mismo escalar.

$$x_1 y_1 + x_2 y_2 + x_3 y_3 + \dots + x_n y_n$$

Casos especiales:

a) El producto escalar de los vectores

$$x' = (x_1, x_2, x_3, \dots, x_n)' \quad e' = (1, 1, 1, \dots, 1)'$$

$$\text{es } x_1 + x_2 + x_3 + \dots + x_n = \sum x_i = n\bar{x}$$

b) El producto escalar del vector (x) por sí mismo, es

$$x' x = x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2 = \sum x_i^2$$

c) El producto escalar del vector (x) por el vector (Ax) es

$$x' Ax = \sum_i \sum_j a_{ij} x_i x_j$$

d) El producto escalar del vector (e) por el vector (A) es  $e'Ae = \sum_i \sum_j a_{ij}$

Caso 5. Multiplicación de un vector columna (y) por un vector línea (x').

El resultado es una matriz. Si el vector columna (y) tiene (m) elementos y el vector línea tiene (n) elementos, el producto es una matriz  $A_{m \times n}$ .

En símbolos

$$y_{m \times 1} x'_{1 \times n} = A_{m \times n}$$

Veamos un caso sencillo

$$y' = (y_1, y_2, y_3, y_4) \quad x' = (x_1, x_2, x_3)$$

las dos matrices (producto de vectores) que pueden formarse son

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} = \begin{bmatrix} y_1 x_1 & y_1 x_2 & y_1 x_3 \\ y_2 x_1 & y_2 x_2 & y_2 x_3 \\ y_3 x_1 & y_3 x_2 & y_3 x_3 \\ y_4 x_1 & y_4 x_2 & y_4 x_3 \end{bmatrix}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} y_1 & y_2 & y_3 & y_4 \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & x_1 y_3 & x_1 y_4 \\ x_2 y_1 & x_2 y_2 & x_2 y_3 & x_2 y_4 \\ x_3 y_1 & x_3 y_2 & x_3 y_3 & x_3 y_4 \end{bmatrix}$$

Nota: Si el vector  $y = e_i$  y el vector  $x = e_j$ , se tendrá

$$= y x' = e_i e_j' = \begin{bmatrix} 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 0 & \dots & 0 \end{bmatrix} = E_{ij}$$

resultando la relación de importancia

$$e_i e_j' = E_{ij}$$

que permite expresar una matriz elemental ( $E_{ij}$ ) en función de "vectores unitarios

$e_i$  y  $e_j$  correspondientes

Caso 6. Multiplicación de una matriz (A) por su recíproco (A<sup>-1</sup>).

En este caso el producto (post o premultiplicación) es la matriz identidad (I).

En símbolos.

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

Caso 7. Multiplicación de una matriz (A) por una matriz diagonal (D).

Se pueden presentar dos casos: premultiplicación y postmultiplicación. Cuando la matriz (A) se pre-multiplica por una matriz diagonal de elementos (d<sub>1</sub>, d<sub>2</sub>, ... d<sub>n</sub>)

cada línea de la matriz queda multiplicada por

$$d_1, d_2, \dots, d_n.$$

Así, por ejemplo:

$$\begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} d_1 a_{11} & d_1 a_{12} \\ d_2 a_{21} & d_2 a_{22} \end{bmatrix}$$

En el caso de post-multiplicación, cada columna queda amplificada por d, d, ... d.

Así por ejemplo

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} d_1 a_{11} & d_2 a_{12} \\ d_1 a_{21} & d_2 a_{22} \end{bmatrix}$$

Caso 8. Multiplicación de una matriz (A) por una matriz elemental

Las matrices elementales son las matrices E, H y K indicadas al señalar el tipo j) de matrices.

Consideraremos primeramente la multiplicación por la matriz E<sub>ij</sub> que es una matriz nxn, con todos los elementos

o excepto el elemento de la celda (ij) que es igual a 1.

Para el caso de post-multiplicación se tiene

$$A E_{ij} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1i} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2i} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{j1} & a_{j2} & \dots & a_{ji} & \dots & a_{jn} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{ni} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & \dots & a_{1i} & \dots & 0 \\ 0 & 0 & \dots & a_{2i} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{ni} & \dots & 0 \end{bmatrix}$$

esto es, una matriz con todos los elementos nulos, excepto los de la columna (j) que son iguales a los elementos del vector de la columna (i) de la matriz (A).







reemplazando en la relación anterior, obtenemos

$$XY' = \sum_1^r (Xe_i) (e_i'Y') = \sum_1^r x_i y_i'$$

siendo  $(x_i)$  el  $i$ ésimo vector columna de la matriz  $(X)$  e  $(y_i')$  el  $i$ ésimo vector línea de la matriz  $(Y')$ .

Ahora bien, si la matriz  $(A)$  es igual al producto de las matrices  $X$  e  $Y'$ , entonces podremos escribir

$$A = \sum_1^r x_i y_i'$$

y como cada matriz  $(x_i y_i')$  es una matriz de rango  $\underline{1}$ , la matriz  $(A)$  ha quedado descompuesta en  $(r)$  matrices de rango  $\underline{1}$ .

#### Otros tipos de operaciones

Aparte de la suma, resta y multiplicación de matrices es conveniente destacar otros tipos de operaciones, que se realizan con frecuencia en las matrices.

- a) Equivalencia: Mediante operaciones elementales es posible pasar de una matriz  $(A)$  a otra matriz  $(B)$ , que presente alguna particularidad importante. Se dice en tal caso que las matrices  $A$  y  $B$  son equivalentes. Así, por ejemplo, mediante "operaciones línea" puede llevarse una matriz  $(A)$  a una matriz triangular. Estas operaciones línea son las que se usan en el método de eliminación de incógnitas cuando se resuelve un sistema de ecuaciones simultáneas.

Veamos la siguiente matriz

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 3 & 4 & 3 \\ 3 & 3 & 4 \end{bmatrix}$$

podemos agregar a las líneas 2 y 3 la primera línea multiplicada por  $(-3)$  o sea, la operación

$$(I - 3E_{21} - 3E_{31}) A$$

lo que nos da

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

que es una matriz triangular.

Una vez que se ha obtenido la matriz triangular, esta matriz triangular puede llevarse a una matriz diagonal realizando operaciones columna.

En el caso recién indicado, aplicando la operación

$$I - E_{12} - E_{13}$$

que equivale a restarle a las columnas 2 y 3, la columna 1, se tiene la matriz identidad I.

Resumiendo el caso numérico indicado se tiene

$$(I - 3E_{21} - 3E_{31}) A (I - E_{12} - E_{13}) = I$$

que podrá escribirse en la forma

$$A = (I - 3E_{21} - 3E_{31})^{-1} (I - E_{12} - E_{13})^{-1}$$

quedando expresada la matriz (A) por el producto de 2 matrices.

Nota: Luego veremos cómo se determina los inversos de matrices de la forma

$$I + c_1 E_{ij} + c_2 E_{kl} + c_3 E_{mn} + \dots$$

esto es, de una matriz que es igual a la suma de una matriz identidad y diversas matrices elementales.

b) Congruencia: Dos matrices cuadradas (A) y (B), se dice que son congruentes, cuando es posible encontrar una matriz (P) tal que

$$B = P^t A P$$

La congruencia es un caso especial de equivalencia, pero ahora las matrices que multiplican (pre y post) a la matriz A, una es la traspuesta de la otra.

La matriz (B) resultante debe tener alguna forma especial, para poder deducir de allí conclusiones de interés.

La matriz B deberá por tal razón ser una matriz diagonal o bien una matriz identidad, en cuyo caso la matriz (A) deberá ser una matriz simétrica.

Veamos un ejemplo con la matriz (A) siguiente

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

Aplicando la operación línea

$$(I - 2E_{21} + E_{31} - 2E_{41})$$

que equivale a hacer 0 todos los elementos de la primera columna excepto el primero, se tiene

$$(I - 2E_{21} + E_{31} - 2E_{41}) A = \begin{bmatrix} 1 & 2 & -1 & 2 \\ 0 & -3 & 4 & -6 \\ 0 & 4 & 0 & 0 \\ 0 & -6 & 0 & -3 \end{bmatrix}$$

y por la simetría de la matriz (A) podemos aplicar la operación columna de igual tipo indicado, o sea

$$B = (I - 2E_{21} + E_{31} - 2E_{41}) A (I - 2E_{12} + E_{13} - 2E_{14}) =$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 4 & -6 \\ 0 & 4 & 0 & 0 \\ 0 & -6 & 0 & -3 \end{bmatrix}$$

pudiendo verse, que la matriz resultante sigue simétrica.

Aplicamos ahora la operación línea

$$(I + 4/3 E_{32} - 2 E_{42})$$

con lo cual

$$(I + \frac{4}{3} E_{32} - 2E_{42}) B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 4 & -6 \\ 0 & 0 & 16/3 & -8 \\ 0 & 0 & -8 & 9 \end{bmatrix}$$

y multiplicando la operación columna

$$(I + 4/3 E_{23} - 2 E_{24})$$

a la matriz anterior, se tiene

$$C = (I + \frac{4}{3} E_{32} - 2E_{42}) B (I + \frac{4}{3} E_{23} - 2E_{24}) =$$

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 16/3 & -8 \\ 0 & 0 & -8 & 9 \end{bmatrix}$$

que nuevamente es una matriz simétrica.

Trabajando ahora con las líneas 3 y 4 podemos aplicar la operación línea

$$I + 3/2 E_{42}$$

con lo cual tenemos

$$(I + \frac{3}{2} E_{42}) C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 16/3 & -8 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

y finalmente con la operación columna del mismo tipo

$$(I + \frac{3}{2} E_{42}) C (I + \frac{3}{2} E_{24}) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 16/3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix}$$

que es una matriz diagonal.

El proceso puede resumirse ahora multiplicando las matrices "operaciones línea" o las "operaciones columna". Se tiene

$$\begin{aligned} P' &= (I + \frac{3}{2} E_{42}) (I + \frac{4}{3} E_{32} - 2E_{42}) (I - 2E_{21} + E_{31} - 2E_{41}) \\ &= I - 2E_{21} - \frac{5}{3} E_{31} - E_{41} + \frac{4}{3} E_{32} - \frac{1}{2} E_{42} \end{aligned}$$

$$\begin{aligned} P &= (I - 2E_{12} + E_{13} - 2E_{14}) (I + \frac{4}{3} E_{23} - 2E_{24}) (I + \frac{3}{2} E_{24}) \\ &= I - 2E_{12} - \frac{5}{3} E_{13} - E_{14} + \frac{4}{3} E_{23} - \frac{1}{2} E_{24} \end{aligned}$$

o sea

$$P' A P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -3 & 0 & 0 \\ 0 & 0 & 16/3 & 0 \\ 0 & 0 & 0 & -3 \end{bmatrix} \text{ siendo } P = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -2 & 1 & 0 & 0 \\ -5/3 & 4/3 & 1 & 0 \\ -1 & -\frac{1}{2} & 0 & 1 \end{bmatrix}$$

c) Similitud: Se dice que dos matrices A y B son similares si existe una matriz no-singular P con la que se cumple:

$$B = P^{-1} A P$$

Siempre será de interés encontrar la matriz (P), que nos conduzca a una matriz triangular B. Esta matriz triangular (como se verá más adelante) tiene en la diagonal números que son las raíces características de la matriz (A).

Además se tiene que  $B^k = P^{-1} A^k P$

lo que nos permitirá elevar a potencias fraccionarias, por ejemplo, una matriz (A), siendo éste un problema importante en Demografía.

- d) Trasposición de una matriz: Trasponer una matriz es cambiar de posición los vectores columna que forman la matriz y dejarlos como vectores línea.

Así, por ejemplo, trasponer la matriz

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

es dejarla en la forma

$$A' = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{bmatrix}$$

La notación para matriz traspuesta es  $A'$ .

Puede probarse que

- la traspuesta de una suma de matrices es la suma de las traspuestas, esto es,

$$(A + B)' = A' + B'$$

- la traspuesta de un producto de matrices es el producto de sus traspuestas en orden inverso, esto es,

$$(AB)' = B' A'$$

- e) Ortogonalización de una matriz: Se dice que una matriz cuadrada (P) es ortogonal, cuando  $P P' = I$ .

Si se dispone de una matriz cuadrada (A) es posible transformarla en una matriz ortogonal (P), premultiplicando la matriz (A) dada por la matriz (I + L) siendo (L) una matriz estrictamente triangular inferior, o sea

$$L = \begin{bmatrix} 0 & 0 & 0 \dots 0 \\ l_{21} & 0 & 0 \dots 0 \\ l_{31} & l_{32} & 0 \dots 0 \\ \dots & \dots & \dots \\ l_{n1} & l_{n2} & l_{n3} \dots 0 \end{bmatrix}$$

El método de ortogonalización recibe el nombre de "método de ortogonalización de Gram-Schmidt".

De acuerdo a la proposición anterior tenemos la relación

$$(I + L) A = P$$

y si denotamos por  $(I + L)_i$  y  $P_i$  los vectores de la línea (i) de las matrices  $(I + L)$  y  $P$ , se tendrá

$$(I + L)_i A = P_i$$

que es equivalente a la relación

$$P_i^i = l_{i1} A_1^i + l_{i2} A_2^i + l_{i3} A_3^i + \dots + A_i^i$$

siendo  $A_1, A_2, \dots, A_i, \dots, A_n$  los vectores línea de la matriz (A).

Para  $i = 1$ , tenemos

$$P_1^i = A_1^i$$

Para  $i = 2$ , en cambio,

$$P_2^i = l_{21} A_1^i + A_2^i$$

o sea

$$P_2^i = l_{21} P_1^i + A_2^i$$

pero dado que  $P_1$  y  $P_2$  son ortogonales, podemos escribir la ecuación

$$P_2^i P_1^i = 0 = l_{21} (P_1^i P_1^i) + A_2^i P_1^i$$

con lo cual

$$l_{21} = - A_2^i P_1^i / P_1^i P_1^i$$

Para  $i = 3$  tenemos

$$P_3^i = l_{31} A_1^i + l_{32} A_2^i + A_3^i$$

pero dado que  $A_2$  es función de  $P_1$  y  $P_2$ , podemos decir que  $P_3$  es función lineal de esos dos vectores y escribir

$$P_3^i = \lambda_{31} P_1^i + \lambda_{32} P_2^i + A_3^i$$



Como  $(P_3)$  es ortogonal con  $P_1$  y  $P_2$ , tenemos

$$P_3^T P_1 = \lambda_{31} P_1^T P_1 + A_3^T P_1 = 0$$

$$P_3^T P_2 = \lambda_{32} P_2^T P_2 + A_3^T P_2 = 0$$

o sea

$$\boxed{\begin{aligned} \lambda_{31} &= -A_3^T P_1 / P_1^T P_1 \\ \lambda_{32} &= -A_3^T P_2 / P_2^T P_2 \end{aligned}}$$

con lo cual

$$\begin{aligned} P_3 &= \begin{bmatrix} \lambda_{31} & \lambda_{32} & 1 \end{bmatrix} \begin{bmatrix} P_1^T \\ P_2^T \\ A_3^T \end{bmatrix} = \begin{bmatrix} \lambda_{31} & \lambda_{32} & 1 \end{bmatrix} \begin{bmatrix} A_1^T \\ \lambda_{21} A_1^T + A_2^T \\ A_3^T \end{bmatrix} \\ &= \begin{bmatrix} \lambda_{31} & \lambda_{32} & 1 \end{bmatrix} \left[ \begin{pmatrix} 1 \\ \lambda_{21} \\ 0 \end{pmatrix} A_1^T + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} A_2^T + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} A_3^T \right] \end{aligned}$$

teniendo entonces como tercera línea para la matriz  $(L)$

$$\begin{bmatrix} \lambda_{31} & \lambda_{32} \end{bmatrix} = \begin{bmatrix} \lambda_{31} & \lambda_{32} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \lambda_{21} & 1 \end{bmatrix}$$

Para  $P_4$  podemos escribir, de la misma manera

$$P_4^T = \lambda_{41} P_1^T + \lambda_{42} P_2^T + \lambda_{43} P_3^T + A_4^T$$

y como  $(P_4)$  debe ser ortogonal con los (3) primeros vectores  $(P_i)$  ya determinados, tenemos las ecuaciones

$$P_4^T P_1 = 0 = \lambda_{41} P_1^T P_1 + A_4^T P_1$$

$$P_4^T P_2 = 0 = \lambda_{42} P_2^T P_2 + A_4^T P_2$$

$$P_4^T P_3 = 0 = \lambda_{43} P_3^T P_3 + A_4^T P_3$$

relaciones que nos determinan los valores de los escalares  $\lambda_{41}, \lambda_{42}$  y  $\lambda_{43}$ .

La cuarta línea de la matriz (L) será entonces

$$\begin{bmatrix} l_{41} & l_{42} & l_{43} \end{bmatrix} = \begin{bmatrix} \lambda_{41} & \lambda_{42} & \lambda_{43} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \lambda_{21} & 1 & 0 \\ \lambda_{31} & \lambda_{32} & 1 \end{bmatrix}$$

Podemos continuar el proceso y tener, en general,

$$P_i^i = \lambda_{i1} P_1 + \lambda_{i2} P_2 + \dots + \lambda_{i(i-1)} P_{i-1} + A_i^i$$

y dado que  $(P_i)$  es ortogonal con los  $(i-1)$  vectores  $(P_h)$  anteriores

$$P_i^i P_h = 0 = \lambda_{ih} P_h^i P_h + A_i^i P_h$$

o 
$$\lambda_{ih} = -A_i^i P_h / P_h^i P_h$$

y tener para la fila (i) de la matriz L, la expresión

$$\begin{bmatrix} l_{i1} & l_{i2} & \dots & l_{i,i-1} \end{bmatrix} = \begin{bmatrix} \lambda_{i1} & \lambda_{i2} & \dots & \lambda_{i,i-1} \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ \lambda_{21} & 1 & & 0 \\ \lambda_{31} & \lambda_{32} & & 1 \\ \dots & \dots & \dots & \dots \\ \lambda_{i-1,1} & \lambda_{i-1,2} & \dots & 1 \end{bmatrix}$$

que es equivalente a la multiplicación de la i-ésima línea de la matriz  $\Delta$  por la matriz de las anteriores.

De esa manera se tiene que

$$L = \Delta(I + \Delta),$$

o sea que

$$I + L = I + \Delta + \Delta^2$$

Ejemplo 1

Ortogonalizar la matriz

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

Los vectores líneas son los siguientes

$$A_1^i = [1, 1, 1]$$

$$A_2^i = [1, -2, 1]$$

$$A_3^i = [1, 2, 3]$$

De allí que

$$P_1^i = A_1^i = [1, 1, 1]$$

El segundo vector  $P_2^i$  es igual a

$$P_2^i = \lambda_{21} P_1^i + A_2^i$$

siendo

$$\lambda_{21} = -\frac{A_2^i P_1^i}{P_1^i P_1^i} = 0$$

con lo cual

$$P_2^i = A_2^i = [1, -2, 1]$$

Para el vector  $P_3^i$  tenemos igualmente

$$P_3^i = \lambda_{31} P_1^i + \lambda_{32} P_2^i + A_3^i$$

siendo  $\lambda_{31} = -\frac{A_3^i P_1^i}{P_1^i P_1^i} = -\frac{6}{3} = -2$ ;  $\lambda_{32} = -\frac{A_3^i P_2^i}{P_2^i P_2^i} = 0$

con lo cual

$$P_3^i = -2(1, 1, 1) + (1, 2, 3) = (-1, 0, 1)$$

y de allí que la matriz  $(I + L)$  es igual a

$$I + L = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$$

y la matriz ortogonal  $(P)$  es igual a

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

f) Derivación: Consideraremos los dos tipos siguientes de derivación

- derivada de una matriz con respecto a un escalar
- derivada de un escalar con respecto a una matriz

Derivada de una matriz con respecto a un escalar

La derivada de una matriz con respecto a un escalar es sencillamente la matriz que resulta de derivar cada elemento de la matriz con respecto a ese escalar.

En símbolos, si  $(Y)$  es una matriz de elementos  $(y_{ij})$ , entonces

$$\frac{\partial A}{\partial x_{ij}} = \left( \frac{\partial y_{ij}}{\partial x_{ij}} \right)$$

Ejemplo 1: Derivar la matriz polinomial

$$A(t) = \begin{bmatrix} t^2 + t & t^2 \\ t + 1 & 1 \end{bmatrix}$$

con respecto al escalar "t".

Solución: De acuerdo a la definición, se tiene

$$\frac{dA}{dt} = \begin{bmatrix} 2t + 1 & 2t \\ 1 & 0 \end{bmatrix}$$

Ejemplo 2: Derivar la matriz

$$Y = \begin{bmatrix} x & 2x^3 & 3/x^4 \\ e^x & \text{sen}x & \log_e x \end{bmatrix}$$

con respecto a  $(x)$ .

Solución: Se tiene

$$\frac{dY}{dx} = \begin{bmatrix} 1 & 6x^2 & -3/4x^5 \\ e^x & \cos x & 1/x \end{bmatrix}$$

Derivada de un escalar con respecto a una matriz

En este caso se supone que el escalar es función de la matriz con respecto a la que va a derivar. Por ejemplo, se tiene la matriz  $Y = A X$  siendo  $A$  y  $X$  matrices de elementos  $(a_{ij})$  y  $(x_{ij})$  respectivamente y se quiere calcular la derivada:

$$\frac{\partial y_{ij}}{\partial X}$$

ya que

$$y_{ij} = \sum_1^n a_{ik} x_{kj}$$

entonces

$$\left( \frac{\partial y_{ij}}{\partial X} \right)$$

es una matriz de elementos

$$\left( \frac{\partial y_{ij}}{\partial x_{ij}} \right)$$

que más abajo veremos

Caso particular: El escalar es función de un vector columna  $(x)$  y se desea la derivada de ese escalar con respecto a ese vector. En ese caso, el resultado es un vector columna cuyos elementos son las derivadas parciales del escalar (función) con respecto a cada una de las variables  $(x_i)$  que componen el vector.

En símbolos

$$\frac{\partial F}{\partial X} = \begin{bmatrix} \frac{\partial F}{\partial x_1} \\ \frac{\partial F}{\partial x_2} \\ \vdots \\ \frac{\partial F}{\partial x_n} \end{bmatrix}$$

Pasaremos a considerar ahora una serie de derivadas de funciones escalares y matrices especiales.

Caso 1: Derivada de una constante

Consideraremos la matriz (A) de elementos constantes ( $a_{ij}$ ). Tenemos, para la derivada de la matriz (A) con respecto a un escalar variable ( $x_{ij}$ ):

$$\frac{\partial A_j}{\partial x_{ij}} = 0$$

siendo  $0$  la matriz nula.

Para la derivada del escalar ( $a_{ij}$ ) con respecto a la matriz (X) de elementos ( $x_{ij}$ ), se tiene

$$\frac{\partial a_{ij}}{\partial X} = 0$$

Caso 2: Derivada del escalar  $F = x' Ay = \sum \sum a_{ij} x_i y_j$ , respecto a la matriz A.

Ya que

$$A = \sum_i \sum_j a_{ij} E_{ij}$$

$$F = x' \left( \sum_i \sum_j a_{ij} E_{ij} \right) y$$

$$\frac{\partial F}{\partial a_{ij}} = x' E_{ij} y = (x' e_i) (e_j' y) = x_i y_j$$

$$\frac{\partial (x' Ay)}{\partial A} = (x_i y_j) = \underline{x y'}$$

Caso 3: Derivada del escalar  $F = x' Ax = \sum \sum a_{ii} x_i^2 + \sum_i \sum_{j \neq i} a_{ij} x_i x_j$

Es un caso particular del anterior, y por lo tanto

$$\frac{\partial (x' Ax)}{\partial A} = x x'$$

Caso 4: Derivada del escalar  $F = x' Ax$  siendo (A) una matriz simétrica

Ya que

$$F = \sum \sum a_{ii} x_i^2 + \sum \sum_{i \neq j} a_{ij} x_i x_j$$

$$\frac{\partial F}{\partial a_{ii}} = x_i^2 \quad \frac{\partial F}{\partial a_{ij}} = x_i x_j + x_j x_i = 2x_i x_j$$

y de allí que

$$\frac{\partial F}{\partial \lambda} = \begin{bmatrix} x_1^2 & 2x_1 x_2 & 2x_1 x_3 & \dots & 2x_1 x_n \\ 2x_2 x_1 & x_2^2 & 2x_2 x_3 & \dots & 2x_2 x_n \\ \dots & \dots & \dots & \dots & \dots \\ 2x_n x_1 & 2x_n x_2 & 2x_n x_3 & \dots & x_n^2 \end{bmatrix} = 2xx' - D(x'x)$$

siendo D una matriz diagonal de elementos  $(x_1^2, x_2^2, \dots, x_n^2)$

Caso 5: Derivada de una matriz (X) con respecto a ella misma.

La matriz tiene los elementos  $(x_{ij})$  y por lo tanto

$$\frac{\partial X}{\partial x_{ij}} = \begin{matrix} & & & & (j) \\ & & & & \begin{bmatrix} 0 & 0 \dots 0 \dots 0 \\ 0 & 0 \dots 0 \dots 0 \\ (i) & 0 & 0 \dots 1 \dots 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 \dots 0 \dots 0 \end{bmatrix} \\ & & & & \end{matrix} = E_{ij}$$

o sea, una matriz del mismo orden que la matriz (A) con todos los elementos nulos excepto el de la celda (ij) que es igual a 1. Es, por lo tanto, la matriz elemental  $E_{ij}$ .

La otra derivada

$$\frac{\partial x_{ij}}{\partial X} = \begin{matrix} & & & & (j) \\ & & & & \begin{bmatrix} 0 & 0 \dots 0 \dots 0 \\ 0 & 0 \dots 0 \dots 0 \\ (i) & 0 & 0 \dots 1 \dots 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 \dots 0 \dots 0 \end{bmatrix} \\ & & & & \end{matrix} = E_{ij}$$

Caso 6: Derivada de  $Y = U + V - W$

siendo  $(y_{ij}); (u_{ij}); (v_{ij}); (w_{ij})$  los elementos de las matrices.

La derivada de Y con respecto al escalar  $(x_{ij})$  es igual a

la matriz de elementos

$$\left( \frac{\partial u_{mn}}{\partial x_{ij}} + \frac{\partial v_{mn}}{\partial x_{ij}} - \frac{\partial w_{mn}}{\partial x_{ij}} \right)$$

y la derivada del escalar  $y_{ij}$  con respecto a la matriz (X) es la matriz

$$\frac{\partial y_{ij}}{\partial X} = \frac{\partial u_{ij}}{\partial X} + \frac{\partial v_{ij}}{\partial X} - \frac{\partial u_{ij}}{\partial X}$$

Caso 7: Derivada de  $Y = UV$

Se tiene en este caso

$$\frac{\partial Y}{\partial x_{ij}} = \frac{\partial U}{\partial x_{ij}} V + U \frac{\partial V}{\partial x_{ij}}$$

$$y_{ij} = \sum_{k=1}^n u_{ik} v_{kj}$$

y para la 2.<sup>a</sup> derivada, dado que

$$\frac{\partial y_{ij}}{\partial X} = \sum_{k=1}^n \left( \frac{\partial u_{ik}}{\partial X} v_{kj} + u_{ik} \frac{\partial v_{kj}}{\partial X} \right)$$

Caso 8: Derivada de  $Y = AX$

siendo (A) una matriz de elementos independientes de las  $(x_{ij})$ .

La derivada de la matriz (Y) con respecto al escalar  $x_{ij}$

es

$$\frac{\partial Y}{\partial x_{ij}} = A E_{ij}$$

notando que

$$\frac{\partial Y}{\partial x_{ij}} = A \frac{\partial X}{\partial x_{ij}}$$

Por otra parte, la derivada del escalar  $y_{ij}$  con respecto a la matriz (X) es igual a

$$\boxed{\frac{\partial y_{ij}}{\partial X} = A' E_{ij}}$$



notando que

$$\frac{\partial y_{ij}}{\partial X} = \sum_1^n a_{ik} \frac{\partial x_{kj}}{\partial X} = \sum_1^n a_{ik} (E_{kj})$$

Caso 9: Derivada de  $Y = X X$

La derivada de la matriz (Y) con respecto a  $(x_{ij})$  es igual

$$\frac{\partial Y}{\partial x_{ij}} = \frac{\partial X}{\partial x_{ij}} X + X \frac{\partial X}{\partial x_{ij}} = E_{ij} X + X E_{ij}$$

La derivada del escalar  $(y_{ij})$  con respecto a la matriz (X) es igual a

$$\begin{aligned} \frac{\partial y_{ij}}{\partial X} &= \sum_1^n \left( \frac{\partial x_{ik}}{\partial X} \right) x_{kj} + \sum_1^n x_{ik} \left( \frac{\partial x_{kj}}{\partial X} \right) \\ &= \sum_1^n E_{ik} x_{kj} + \sum_1^n x_{ik} (E_{kj}) \\ &= e_i \left( \sum_1^n e_k' x_{kj} \right) + \left( \sum_1^n x_{ik} e_k \right) e_j' = e_i (x_{1j}, x_{2j}, \dots, x_{nj}) \end{aligned}$$

$$+ \begin{pmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{pmatrix} e_j'$$

$$= e_i' (X e_j)' + (e_i' X) e_j' = (e_i' e_j') X' + X' (e_i e_j') =$$

$$= E_{ij} X' + X' E_{ij}$$

Caso 10: Derivada de  $Y = X' X$

La derivada de la matriz (Y) con respecto a  $(x_{ij})$  es igual a

$$\frac{\partial Y}{\partial x_{ij}} = \frac{\partial X'}{\partial x_{ij}} X + X' \frac{\partial X}{\partial x_{ij}} = E_{ij}' X + X' E_{ij}$$

La derivada del escalar  $(y_{ij})$  con respecto a la matriz  $(X)$  es igual a

$$\begin{aligned} \frac{\partial y_{ij}}{\partial X} &= \sum_1^n \frac{\partial x_{ik}}{\partial X} x_{kj} + \sum_1^n x_{ik} \frac{\partial x_{kj}}{\partial X} \\ &= \sum_1^n E'_{ik} x_{kj} + \sum_1^n x_{ik} E_{kj} = X E'_{ij} + X' E_{ij} \end{aligned}$$

Problema: Demuestre que si  $Y = X X' X$

$$\frac{\partial Y}{\partial x_{ij}} = E_{ij} X' X + X E'_{ij} X + X X' E_{ij}$$

$$\frac{\partial y_{ij}}{\partial X} = E_{ij} X' X + X E'_{ij} X + X X' E_{ij}$$

Caso 11: Derivada de  $X^{-1}$

Partiendo de la definición de inverso, se tiene

$$I = X X^{-1} = X R$$

siendo  $R$  una matriz de elementos  $(r_{ij})$ .

La derivada de la matriz  $I$  con respecto a un escalar  $x_{ij}$  será igual a

$$\begin{aligned} 0 &= X \frac{\partial R}{\partial x_{ij}} + \frac{\partial X}{\partial x_{ij}} R \\ &= X \frac{\partial R}{\partial x_{ij}} + E_{ij} R \end{aligned}$$

con lo cual

$$\frac{\partial R}{\partial x_{ij}} = -R E_{ij} R$$

La derivada de un escalar  $r_{ij}$  de la matriz inversa  $(R)$  está dada por

$$\frac{\partial r_{ij}}{\partial X} = \sum_1^n x_{ik} \frac{\partial r_{kj}}{\partial X} + \sum_1^n \frac{\partial x_{ik}}{\partial X} R = 0$$

$$= \sum_1^n x_{ik} \frac{\partial^r_{kj}}{\partial X} + \sum_1^n E_{ij} R$$

$$= \sum_1^n x_{ik} \frac{\partial^r_{kj}}{\partial X} + E_{ij} R' = 0$$

Observando esta última ecuación puede decirse que

$$\frac{\partial^r_{kj}}{\partial X} = -Z E_{kj} R'$$

siendo  $Z$  una matriz, función de la matriz  $R$ , y reemplazando en la ecuación anterior tenemos

$$-\sum_1^n x_{ik} (Z E_{kj} R') + E_{ij} R' = 0$$

$$-Z (\sum x_{ik} e_k) e_j' R' + E_{ij} R' = 0$$

$$-Z X' E_{ij} R' + E_{ij} R' = 0$$

con lo cual

$$-Z X' + I = 0$$

o sea

$$Z = (X')^{-1} = R'$$

y por lo tanto

$$\boxed{\frac{\partial^n_{ij}}{\partial X} = -R' E_{ij} R'}$$

Problema: Demuestre que si  $Y = X^n$

$$\frac{\partial X^n}{\partial x_{ij}} = \sum_{l=0}^{n-1} X^l E_{ij} X^{n-l-1}$$

$$\frac{\partial y_{ij}}{\partial X} = \sum_{s=0}^{n-1} (X')^s E_{ij} (X')^{n-s-1}$$

Ya que hemos visto una serie de derivadas de escalares con respecto a una matriz y de una matriz con respecto a un escalar, consideraremos -por conveniencia- en forma separada el caso de derivación de un escalar con respecto a un vector columna (x) de componentes  $(x_1, x_2, \dots, x_n)$  o a un vector columna (y) de componentes  $(y_1, y_2, \dots, y_n)$ .

Caso 1: Derivada de la función  $F = x'y = y'x = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$   
La derivada de esta función con respecto al vector (x) es igual a

$$\frac{\partial F}{\partial x} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = y$$

De la misma manera la derivada de esa función (escalar) con respecto al vector (y) será igual a

$$\frac{\partial F}{\partial y} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x$$

teniéndose por lo tanto

$$\frac{\partial(x'y)}{\partial x} = \frac{\partial(y'x)}{\partial x} = y \quad \frac{\partial(x'y)}{\partial y} = \frac{\partial(y'x)}{\partial y} = x$$

Caso 2: Derivada de la función cuadrática simple

$$F = x'x = x_1^2 + x_2^2 + \dots + x_n^2$$

Derivando con respecto al vector (x) se encuentra

$$\frac{\partial F}{\partial x} = \begin{pmatrix} 2x_1 \\ 2x_2 \\ \vdots \\ 2x_n \end{pmatrix} = 2x$$

Caso 3: Derivada de la función  $F = x'Ax = \sum_i a_{ii} x_i^2 + \sum_{i \neq j} a_{ij} x_i x_j$  con  $A=A'$

Ya que

$$\frac{\partial F}{\partial x_i} = 2a_{ii} x_i + \sum_{j \neq i} (a_{ij} x_i + a_{ji} x_j) =$$

$$= 2a_{ii} x_i + 2 \sum_{j \neq i} a_{ij} x_j$$

se tiene

$$\frac{\partial(x' Ax)}{\partial x} = 2Ax$$

Caso 4: Derivada de la función  $F = (x'y)^2$  con respecto al vector  $(y)$

Ya que

$$F = (x'y) (x'y) = (y'x) (x'y) = y' Xy$$

siendo  $X$  una matriz simétrica, o sea el caso anterior. De allí que

$$\frac{\partial(x'y)^2}{\partial y} = 2xy = 2xx'y = \underline{2(x'y)x}$$

