

INVERSA CONDICIONAL E INVERSA
GENERALIZADA DE UNA MATRIZ.
ESQUEMA GEOMETRICO *

MIGUEL A. MARMOLEJO L.
PROFESOR ASISTENTE.
DEPARTAMENTO DE MATEMATICA.
UNIVERSIDAD DEL VALLE- COLOMBIA.

1. PRELIMINARES

El propósito de este apartado es el de fijar alguna terminología y presentar algunas definiciones que servirán de referencia en el desarrollo del resto del trabajo.

1.1 Cada matriz $(A) = [a_{ij}]_{m \times n}$ tiene asociados cuatro espacios a saber : (a) $F(A)$, el espacio fila de A , (b) $N(A)$, el espacio nulo de A , (c) $C(A)$, el espacio columna de A , y (d) $N^*(A^t)$, el espacio nulo de A^t .

Además se sabe que:

(I) (a) $F(A)$ y $N(A)$ son ortogonales y (b) $R^n = F(A) \oplus N(A)$, donde $\oplus$ indica suma directa. Así, se tiene que:

$\dim F(A) + \dim N(A) = n$,
y cada $x \in R^n$ puede expresarse en forma única así:

$x = f + \alpha$,
donde $f \in F(A)$ y $\alpha \in N(A)$.

Así mismo,

(II) (a) $C(A)$ y $N^*(A^t)$ son ortogonales y (b) $R^m = C(A) \oplus N^*(A^t)$, donde $\oplus$ indica suma directa. Así, se tiene que:

$\dim C(A) + \dim N^*(A^t) = m$,
y cada $y \in R^m$ puede expresarse en forma única así:
 $y = c + \beta$,
donde $c \in C(A)$ y $\beta \in N^*(A^t)$.

1.2. El rango de una matriz A lo denotaremos por $\rho(A)$. Sobre el rango de una matriz mencionamos que:

(a) $\rho(A) = \rho(A^t A)$
 $= \dim F(A) = \dim C(A)$

1.3. Cada matriz $A = [a_{ij}]_{m \times n}$ tiene asociada la transformación lineal

$$\begin{aligned} A: \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ x &\rightarrow y = Ax \end{aligned}$$

Más aún; el conjunto imagen de A es $C(A)$ y el núcleo de A es $N(A)$.

En la figura 1 ilustraremos lo establecido hasta ahora.

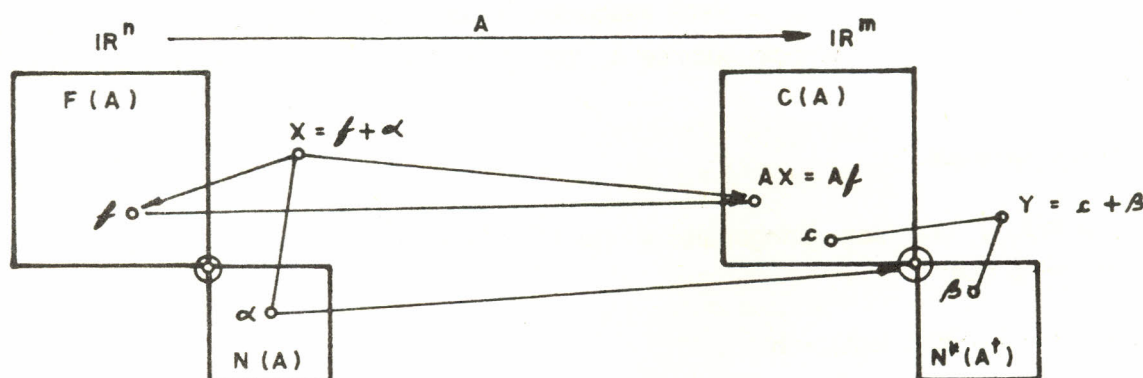


FIGURA 1

De otra parte, existe un método sencillo para calcular una base de cada uno de los cuatro espacios asociados a una matriz

$$A = [a_{ij}]_{m \times n}.$$

El método consiste en efectuar los pasos siguientes:

PASO 1. Forme la matriz $[A^t | I_n]$.

PASO 2. Efectúe operaciones elementales en las fila de la matriz anterior hasta conseguir una forma escalonada de la matriz A^t . Al final de este paso se obtiene una matriz que podemos describir por bloques así: ($r = \rho(A)$).

$$\left[\begin{array}{c|c} E_{r \times m} & P_{r \times n} \\ \hline O_{n-r \times m} & P_{n-r \times n} \end{array} \right]$$

Las filas de la matriz $E_{r \times m}$ conforman una base para $C(A)$; el conjunto imagen de A , y las filas de la matriz $P_{n-r \times n}$ conforman una base para $N(A)$; el núcleo de A . Al efectuar el paso 2 a la matriz

$[A]_m$ se obtienen sendas bases para $C(A^\dagger) = F(A)$ y $N^*(A^\dagger)$.

1.4 Para cada subespacio U de R^k existe un subespacio W de R^k tal que $R^k = U \oplus W$. Donde $\oplus$ indica suma directa. Un tal subespacio W es llamado un complemento de U .

1.5. (i) Se dice que x_0 es una solución mínima cuadrada del sistema de ecuaciones $Ax = y$, si:
 $\|Ax_0 - y\| \leq \|Ax - y\|$ para todo x .

(ii) Se dice que x_0 es una mejor solución aproximada del sistema de ecuaciones $Ax = y$, si:

$$(a) \quad \|Ax_0 - y\| \leq \|Ax - y\| \text{ para todo } x, y$$

$$(b) \quad \text{Si } x^* \text{ es tal que } \|Ax_0 - y\| = \|Ax^* - y\|, \\ \text{entonces } \|x_0\| \leq \|x^*\|.$$

2. INVERSA CONDICIONAL E INVERSA GENERALIZADA DE UNA MATRIZ: DEFINICIONES

2.1. **Definición:** Sea A una matriz $m \times n$. Una matriz $A^c_{n \times m}$ es una inversa condicional (c-inversa) de A , si $AA^cA = A$.

2.2. **Definición:** Sea A una matriz $m \times n$. Una matriz $A^+_{n \times m}$ es una inversa generalizada (g-inversa) de A , si:

$$(i) \quad AA^+A = A$$

$$(ii) \quad A^+AA^+ = A^+$$

$$(iii) \quad AA^+ \text{ y } A^+A \text{ son simétricas.}$$

Obsérvese que una g-inversa de A es una c-inversa de A . A continuación veremos una manera de conseguir una c-inversa de una matriz.

Sean: A una matriz $m \times n$, W un complemento de $N(A)$ y U un complemento de $C(A)$. Así; $R^n = W \oplus N(A)$ y $R^m = C(A) \oplus U$, por lo que cada $x \in R^n$ y cada $y \in R^m$ pueden expresarse en forma única así: $x = w + \alpha$ y $y = c + u$, donde w, α, c y u pertenecen a $W, N(A), C(A)$ y U respectivamente. Ahora; si definimos $A^c: R^m \rightarrow R^n$ de manera que $A^c[Aw] = w$ para cada $Aw \in C(A)$, entonces se tiene que para cada $x \in R^n$: $AA^cAx = AA^cA(w + \alpha) = AA^cAw = Aw = Ax$.

De aquí que $AA^cA = A$. lo cual indica que A^c es una c-inversa de la matriz A .

De otra parte; dado que W y U son arbitrarios y dado que hay libertad para definir $A^c(u)$ para cada $u \in U$, se concluye que en general una matriz puede tener más de una c-inversa. En la figura

2 ilustramos esto.

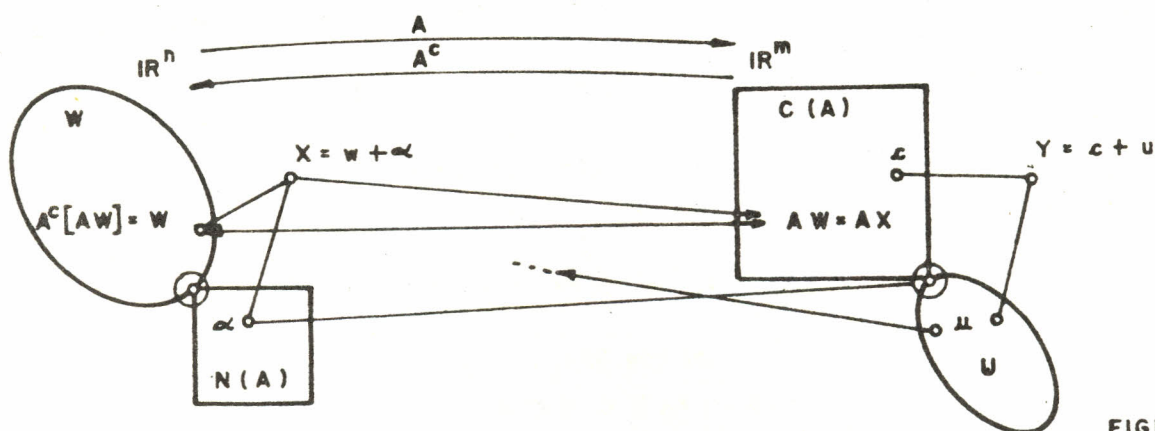


FIGURA 2

3. C-INVERSA DE UNA MATRIZ: PROPIEDADES

En este apartado veremos:

- (a) Algunas propiedades generales de una c-inversa de A.
- (b) Algunas propiedades que posee una c-inversa de A, según:
 - (i) La manera de definir $A^c(u)$ para cada $u \in U$.
 - (ii) El rango de A.
 - (iii) La escogencia de W y de U.

1. Al definir $A^c: R^m \rightarrow R^n$ de manera $A^c[Aw] = w$ para cada $Aw \in C(A)$, se tiene que: (ver figura número 2)

- (1) $AA^cA = A$
- (2) AA^c y A^cA son matrices idempotentes,
- (3) $C(A) = C(AA^c)$, por lo que $\rho(A) = \rho(AA^c)$,
- (4) $C(A^cA) = W$, por lo que $\rho(A) = \rho(A^cA)$,
- (5) $W \subseteq C(A^c)$, por lo que $\rho(A) \leq \rho(A^c)$,
- (6) $Ax = y$ tiene solución sii $AA^cy = y$, y
- (7) Si $Ax = y$ tiene solución, su solución general está dada por: $x = A^cy + \alpha$, $\alpha \in N(A)$.

$$x = A^C y + (I - A^C A) h, h \in R^n.$$

II. Al definir $A^C: R^m \rightarrow R^n$ de manera que:

$$A^C [Aw] = w, Aw \in C(A), y$$

$$A^C(u) = 0, u \in U$$

Se tiene que: (ver figura número 2)

$$(1)' AA^C A = A \text{ y } A^C AA^C = A^C,$$

$$(2), (3), (4),$$

$$(5)' W = C(A^C), \text{ por lo que } \rho(A) = \rho(A^C),$$

$$(6) \text{ y } (7).$$

III. Si $\rho(A) = m < n$, entonces $U = \{0\}$. De aquí que al definir $A^C: R^m \rightarrow R^n$ de manera que $A^C [Aw] = w$ para cada $Aw \in C(A)$, se tiene que: (ver figura número 3)

$$(1)', (2), (3), (4), (5)',$$

$$(6)' Ax = y \text{ siempre tiene solución,}$$

$$(7)' Ax = y \text{ siempre tiene solución y su solución general está dada por:}$$

$$x = A^C y + (I - A^C A)h, h \in R^n, y$$

$$(8) AA^C = I$$

IV. Si $\rho(A) = n < m$, entonces $N(A) = \{0\}$, por lo que al definir $A^C: R^m \rightarrow R^n$ de manera que $A^C [Aw] = w$ para cada $Aw \in C(A)$, se tiene que: (ver figura número 4)

$$(1)', (2), (3), (4), (5)', (6),$$

$$(7)' Si Ax = y \text{ tiene solución, ésta es única y está dada por } x = A^C y, y$$

$$(8)' A^C A = I$$

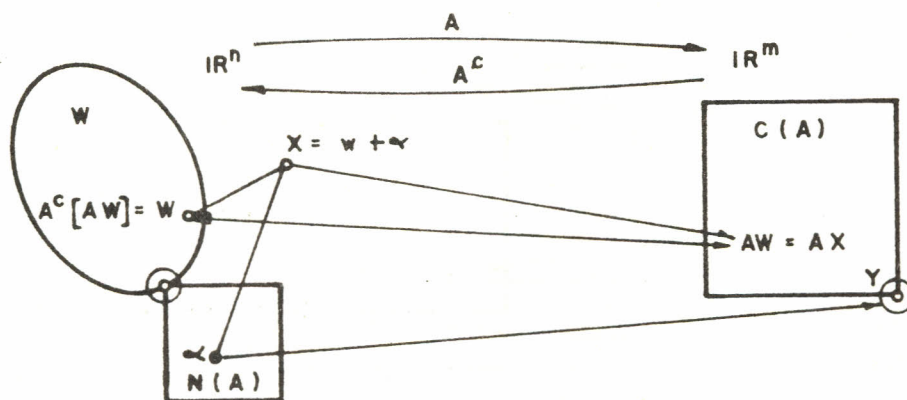


FIGURA 3

V. Si $r(A) = n = m$, entonces $N(A) = \{0\} = U$.

De aquí que necesariamente $A^C = A^{-1}$

VI. Si $U = N^*(A^t)$ y se define $A^C: R^m \rightarrow R^n$ de manera que:

$$A^C [Aw] = w, \quad Aw \in C(A), \text{ y}$$

$$A^C u = \alpha, \quad u \in U, \quad \alpha \in N(A)$$

Entonces se tiene que: (ver figura número 5).

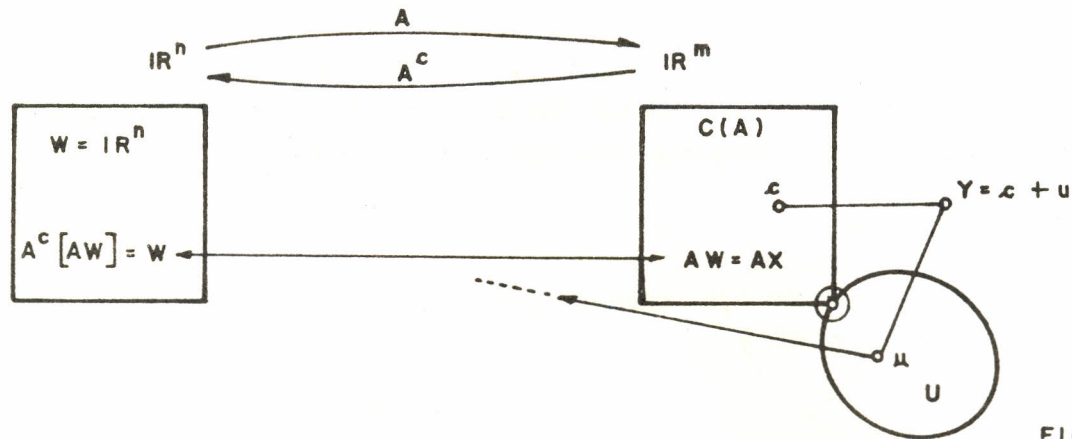


FIGURA 4

(1), (2), (3), (4), (5), (6), (7),

(9) AA^C es la proyección ortogonal de $\mathbb{R}^m$ sobre el espacio columna de A . Por esto, AA^C es simétrica e idempotente. De aquí que $x_0 = A^C y$ es una solución mínima cuadrada de $Ax = y$, pues: (ver figura número 6)

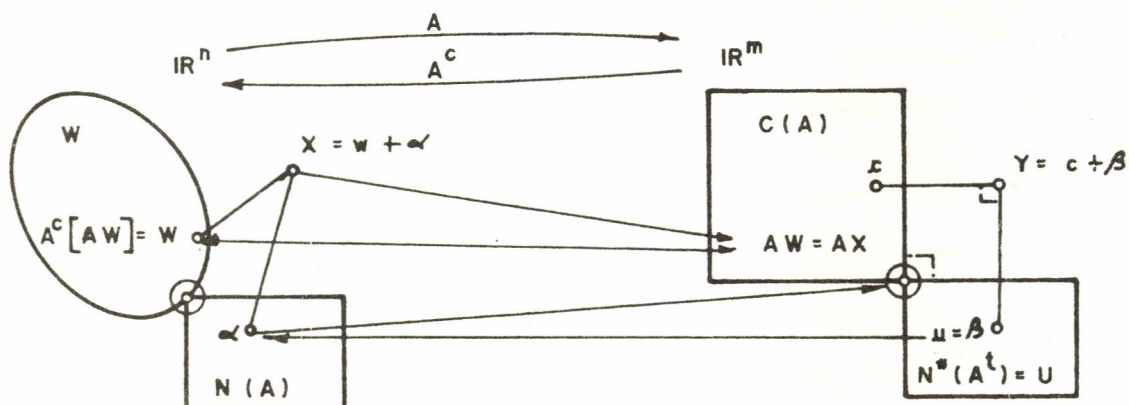


FIGURA 5

$\| AA^C y - y \| = \| \text{Proyección ortogonal de } y \text{ sobre } C(A) - y \| \leq \| Ax - y \|$ para todo x .

Además. como: $AA^C y = A[A^C y + \alpha]$, $\alpha \in N(A)$, se concluye que en general $Ax = y$ puede tener más de una solución mínima cuadrada.

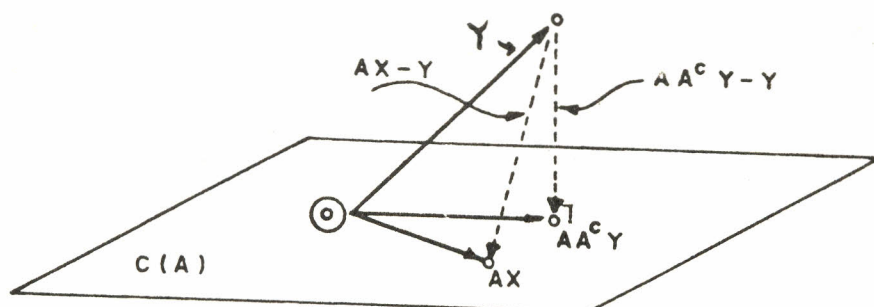


FIGURA 6

VII. Si $W = F(A)$, al definir $A^C: R^m \rightarrow R^n$ de manera que: $A^C[Aw] = w$ para cada $Aw \in C(A)$, se tiene que: (ver figura número 7).

(1), (2), (3), (4), (5), (6), (7),

(10) $A^C A$ es la proyección ortogonal de R^n sobre el espacio fila de A . Por esto, $A^C A$ es simétrica e idempotente. De aquí que $w = A^C[Aw]$ es el vector de menor longitud entre aquellos vectores x tales que $Ax = Aw$. En efecto; si

$Ax = Aw$, entonces $x = w + \alpha$, $\alpha \in N(A)$, por lo que:

$$\|x\|^2 = \|w + \alpha\|^2 = \|w\|^2 + \|\alpha\|^2 \geq \|w\|^2$$

VIII. Si $U = N^*(A^t)$ y $W = F(A)$, al definir $A^C: R^m \rightarrow R^n$ de manera que:

$$A^C[Aw] = w, \quad Aw \in C(A)$$

$$A^C[u] = \theta, \quad u \in U,$$

Entonces se tiene que: (ver figura número 8)

(1)', (2), (3), (4), (5)', (6), (7), (8), (8)', (9) y (10).

Además; (9) y (10) expresan que $x_0 = A^C y$ es una mejor solución aproximada de $Ax = y$.

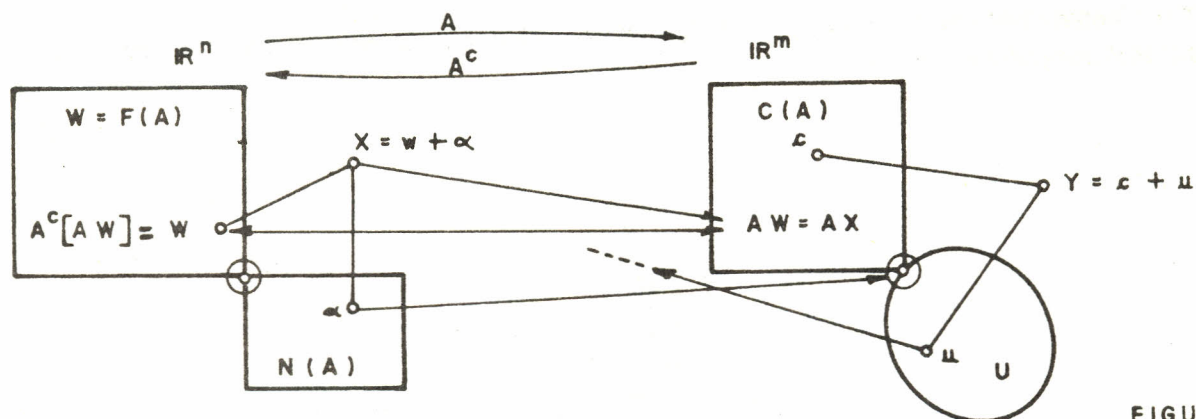


FIGURA 7

De otro lado; en primer lugar, note que A^c definida así es una g-inversa de A , y en segundo lugar; dado que $N^*(A^t)$ es el único complemento ortogonal de $C(A)$ y que $F(A)$ es el único complemento ortogonal de $N(A)$, se concluye que A^c así definida es única. En consecuencia: toda matriz A tiene una única g-inversa. Además; $Ax = y$ siempre tiene una única mejor solución aproximada, a saber $x_0 = A^+y$.

4. G - INVERSA DE UNA MATRIZ : PROPIEDADES Y CALCULO ALGEBRAICO

En este apartado veremos algunas propiedades que posee la g-inversa de una matriz y una manera de calcularla. Empezaremos reescribiendo lo establecido en 3. VIII.

I. Para toda matriz $A_{m \times n}$ existe una única matriz $A^+_{n \times m}$, la g-inversa de A , tal que:

- (1) $AA^+A = A$ y $A^+AA^+ = A^+$,
- (2) AA^+ y A^+A son simétricas e idempotentes,
- (3) $\rho(A) = \rho(A^+) = \rho(AA^+) = \rho(A^+A)$,
- (4) $N(A^+) = N^*(A^t)$, $C(A^+) = F(A)$,
 $F(A^+) = C(A)$,
- (5) $Ax = y$ tiene solución si $AA^+y = y$,

(6) Si $Ax = y$ tiene solución, su solución general ésta dada por:

$$x = A^+y + (I - A^+A)h, \quad h \in R^n.$$

(7) AA^+ es la proyección ortogonal de R^m sobre $C(A)$ y A^+A es la proyección ortogonal de R^n sobre $F(A)$, por lo que; $Ax = y$ siempre tiene una única mejor solución aproximada, a saber: $x_0 = A^+y$.

II. Si $A = O_{m \times n}$, entonces $A^+ = O_{n \times m}$.

III. Si $\rho(A) = m < n$, entonces AA^t es inversible. De otro lado, como A^+A es la proyección ortogonal de $\mathbb{R}^n$ sobre $F(A)$. Entonces:

$$A^+AA^t = A^t,$$

de donde;

$$A^+ = A^t[AA^t]^{-1}$$

IV. Si $\rho(A) = n < m$, entonces A^tA es inversible. Puesto que AA^+ es simétrica y puesto que:

$$AA^+A = A,$$

Entonces:

$$A^tAA^+ = A^t,$$

de donde;

$$A^+ = [A^tA]^{-1} A^t$$

V. Si $\rho(A) = n = m$, entonces $A^+ = A^{-1}$.

VI. Sean B una matriz $m \times r$ y C una matriz $r \times n$. Si $F(B) = C(C)$, entonces

$$(BC)^+ = C^+B^+$$

(ver figura número 9).

En particular:

(i) Para cualquier matriz A :

$$(AA^t)^+ = (A^t)^+ A^+ \text{ y}$$

$$(A^tA)^+ = A^+ (A^t)^+.$$

ii) Si B es una matriz $m \times r$ y C es una matriz $r \times n$ y si $\rho(B) = \rho = \rho(C)$, entonces $(BC)^+ = C^+B^+$. Más explícitamente, según III y IV, $(BC)^+ = C^t[CC^t]^{-1} [B^tB]^{-1} B^t$.

VII. Si A es simétrica, entonces A^+ es simétrica. Más aún: (ver figura número 10)

(i) $Ax = 0$ sii $A^+x = 0$

(ii) $Ax = \lambda x$ sii $A^+x = \lambda^{-1}x$; $\lambda \neq 0$

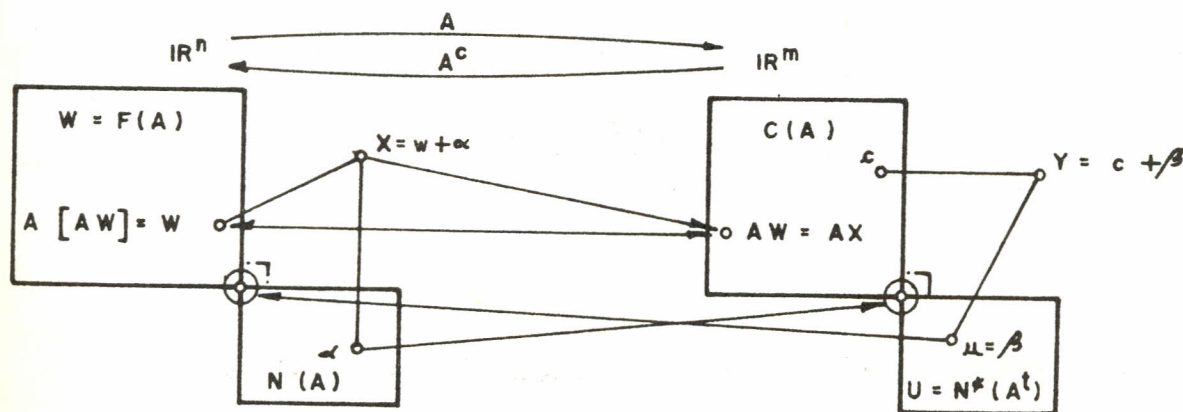


FIGURA 8

Por último veremos un método para calcular la g-inversa de una matriz $A = [a_{ij}]_{m \times n}$. El método consiste en efectuar los pasos siguientes:

PASO 1: Forme la matriz $[A|I_m]$.

PASO 2: Efectúe operaciones elementales en las filas de la anterior matriz hasta conseguir una forma escalonada de la matriz A. Al final de este paso se obtiene una matriz que podemos describir por bloques así: ($r = \rho(A)$).

$$\left[\begin{array}{c|c} E_{rxn} & P_{rxm} \\ \hline O_{m-rxn} & P_{m-rxm} \end{array} \right]$$

PASO 3: Forme la matriz

$$\left[\begin{array}{c|c} E_{rxn} & A^t \\ \hline P_{m-rxm} & O_{m-rxn} \end{array} \right]$$

Paso 4: Lleve la matriz anterior a su forma escalonada reducida. Al final de este paso se obtiene la matriz $[I_m | (A^+)^t]$

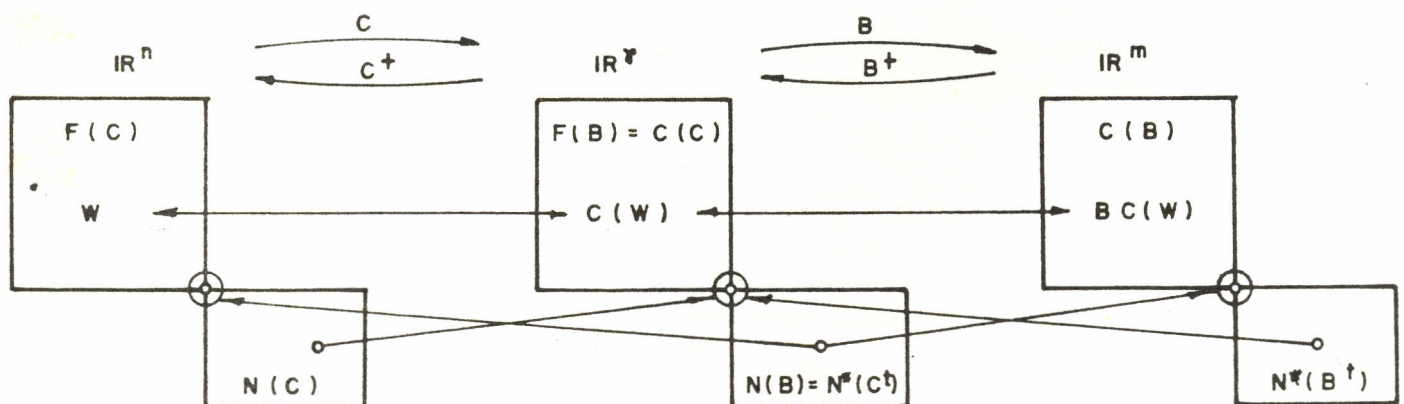


FIGURA 9

5. APLICACION: AJUSTE DE DATOS POR MINIMOS CUADRADOS

Como es sabido; el sistema de ecuaciones $Ax = y$:

(i) Tiene infinitas soluciones ó (ii) Tiene solución única ó (iii) No tiene solución.

En el trabajo experimental generalmente se da la opción (iii). En este caso nos preguntamos si existe una "solución aproximada", para una definición conveniente de solución aproximada.

El método de los mínimos cuadrados busca una solución mínima cuadrada. Es decir, busca x_0 tal que $\|Ax_0 - y\| \leq \|Ax - y\|$ para todo x .

Como hemos visto, $Ax = y$ tiene siempre al menos una solución mínima cuadrada, a saber: $x = A^+y$, la mejor solución aproximada de $Ax = y$.

Más aún; si $A = [a_{ij}]_{m \times n}$ tiene rango n , se desprende de lo establecido en 3. VI (9) y 3. IV que $Ax = y$ tiene sólo una solución mínima cuadrada.

I. Adaptación, por mínimos cuadrados, de una línea recta. Sea una línea recta:

$$y = a + bx$$

que se quiere adaptar a los puntos: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Si los puntos fuesen colineales, los coeficientes desconocidos a y b satisfacerían la ecuación de la recta, es decir, se tendría:

$$y_i = a + bx_i, \quad i = 1, 2, \dots, n.$$

Las anteriores igualdades pueden escribirse así:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = Ax$$

Si los puntos no son colineales, de lo que se trata es de encontrar una solución mínima cuadrada de $Ax = y$.

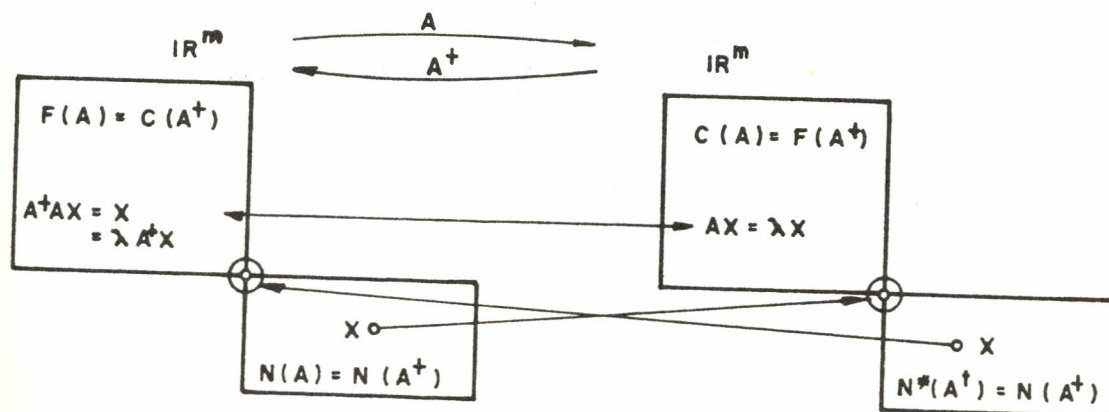


FIGURA 10

Ejemplo. Adaptemos, por mínimos cuadrados, los puntos (0,1), (1,3), (2,4) y (3,4) a una línea recta $y = a + bx$.

Lo que debemos hacer es encontrar una solución mínima cuadrada de

$$Ax = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 4 \\ 4 \end{bmatrix} = y$$

Una tal solución está dada por $x_0 = A^+y$. Ahora bien, como $r(A) = 2$, entonces ésta es la única solución mínima cuadrada de $Ax = y$. Por lo tanto los coeficientes buscados están dados por:

$$x_0 = \begin{bmatrix} a \\ b \end{bmatrix} = A^+y = [A^tA]^{-1} A^ty = \begin{bmatrix} 1.5 \\ 1 \end{bmatrix}$$

II. Adaptación, por mínimos cuadrados, de un polinomio.

Para adaptar, por mínimos cuadrados, un polinomio de grado $\leq m$:

$$y = a_0 + a_1x + a_2x^2 + \dots + a_mx^m$$

a un conjunto de n puntos: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, se busca una solución mínima cuadrada de

$$Ax = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^m \\ 1 & x_2 & x_2^2 & \dots & x_2^m \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^m \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = y$$

Ejemplo. Encontremos un polinomio de grado ≤ 2 que mejor se adapte, por mínimos cuadrados, a los puntos: (0.1,-0.18), (0.2,0.31), (0.3,1.03), (0.4,2.48) y (0.5,3.73).

Debemos encontrar una solución mínima cuadrada de

$$Ax = \begin{bmatrix} 1 & 0.1 & 0.01 \\ 1 & 0.2 & 0.04 \\ 1 & 0.3 & 0.09 \\ 1 & 0.4 & 0.16 \\ 1 & 0.5 & 0.25 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} -0.18 \\ 0.31 \\ 1.03 \\ 2.48 \\ 3.73 \end{bmatrix} = y$$

Puesto que $r(A) = 3$, existe una única solución mínima cuadrada de $Ax = y$, a saber:

$$x_0 = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = A^+ y = \begin{bmatrix} -0.40 \\ 0.35 \\ 16.10 \end{bmatrix}$$

BIBLIOGRAFIA:

- [1] GRAYBILL, F.A.: *Introduction to matrices with applications in statistics*. Wadsworth Publishing Company Inc. 1971.
- [2] RORRES, C. y ANTON H.: *Aplicaciones de algebra lineal*. Limusa, 1979.
- [3] STRANG, G.: *Algebra lineal y sus aplicaciones*. Fondo Educativo Interamericano, S.A., 1980.