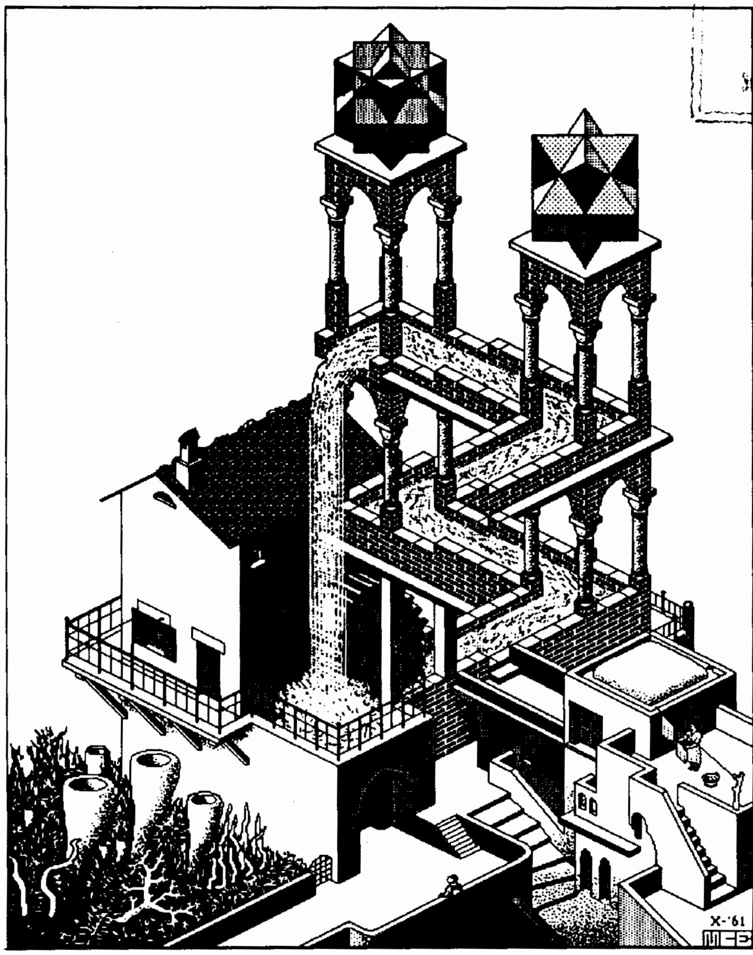
APUNTES DE MATEMATICAS APLICADAS I

- PARTE II -

IDALIA FLORES DE LA MOTA



Departamento de Ingeniería de Sistemas

División de Estudios de Posgrado   Facultad de Ingeniería, U. N. A. M.

APUNTES DE MATEMATICAS APLICADAS I

PARTE II

IDALIA FLORES DE LA MOTA

I N D I C E

CAPITULO 5

ESPACIOS VECTORIALES

5.1 INTRODUCCION

Hemos visto como se puede efectuar la eliminación para simplificar un sistema lineal $Ax = b$; afortunadamente, no sólo se facilita el cálculo del valor de x sino que también se responde a las cuestiones teóricas acerca de su existencia y su unicidad. El objetivo fundamental de ésta parte es adquirir además, un conocimiento diferente y mas profundo del problema.

Para ello necesitamos el concepto de espacio vectorial, introducimos la idea con los espacios mas importantes R, R^2, R^3,... El espacio R^2 se representa por el ya conocido plano x-y, donde las dos componentes del vector son las coordenadas x, y del punto correspondiente. También es familiar R^3 con las tres componentes que dan un punto en el espacio tridimensional y R es una recta. Lo importante en álgebra lineal es que la extensión a n dimensiones es directa.

En éstos espacios, como en todos los espacios vectoriales son posibles dos operaciones: podemos sumar dos vectores cualesquiera y podemos multiplicar vectores por escalares.

5.2 ESPACIOS VECTORIALES Y ALGEBRA EN R^N

Sean u y v vectores en R^n, donde $u = (u_1, u_2,..., u_n)$ y $v = (v_1, v_2,..., v_n)$

La suma u + v es el vector

$$u + v = (u_1 + v_1,\ u_2 + v_2, ...,\ u_n + v_n)$$

si r es un escalar, el múltiplo escalar ru se define

$$ru = (ru_1,\ ru_2,...,\ ru_n)$$

si u es un vector en R^n se define el negativo (o inverso aditivo) de u mediante -u y se define como:

$$- u = (-u_1,\ -u_2,...,-u_n)$$

la sustracción de vectores en R^n se define como $v - u = v + (-u)$ o en términos de las componentes:

$$\begin{aligned} v - u &= (v_1,\ v_2,...,v_n) + (-u_1,-u_2,...,-u_n) = \\ &= v_1 - u_1,\ v_2 - u_2,...,\ v_n - u_n \end{aligned}$$

Nota: Para sumar dos vectores, ambos deben estar contenidos en el mismo espacio R^n. La suma de vectores con diferente número de componentes no está definida.

Las propiedades mas importantes de la suma y el producto por un escalar en R^n se enumeran en el siguiente teorema

TEOREMA 5.1

Sean u, v y w en R^n y r, s escalares, entonces

a. $u + v = v + u$
b. $u + (v + w) = (u + v) + w$
c. $u + 0 = 0 + u = u$
d. $u + (-u) = 0$ es decir $u - u = 0$
e. $r(su) = (rs)u$
f. $r(u + v) = ru + rv$
g. $(r + s)u = ru + su$
h. $1u = u$

Este teorema permite manejar los vectores en R^n sin tener que expresarlos en términos de sus componentes, exactamente como se trabaja con números reales.

Para extender las nociones de distancia, norma y ángulo a R^n se comenzará con la siguiente generalización de producto punto en R^2 y en R^3.

DEFINICION 5.1 Si u y v son vectores en R^n entonces el producto interior euclideano se define como:

$$u \cdot v = u_1 v_1 + u_2 v_2 + \dots + u_n v_n$$

los vectores u y v se dicen ortogonales (o perpendiculares) si su producto interior es igual a cero: $u v = 0$

Ejemplo 5.1

Encuentre el producto interior euclideano de los vectores u y v en R^4 donde u y v son :

$$u = (-1, 3, 5, 7) \qquad v = (5, -4, 7, 0)$$

Solución:

$$u \cdot v = (-1)(5) + (3)(-4) + (5)(7) + (7)(0) = 18$$

Las propiedades básicas del producto interior en R^b se enumeran en el siguiente teorema.

TEOREMA 5.2

Si u, v y w son vectores en R^n y r es un escalar, entonces:

a. $u \cdot v = v \cdot u$
b. $(u + v) \cdot w = u \cdot w + v \cdot w$
c. $r(u\ v) = (ru)\ v = u\ (rv)$
d. $u \cdot u \geq 0$ y $u \cdot u = 0$ si y solo si $u = 0$

demostración del inciso d.

$u \cdot u = u_1^2 + u_2^2 + \ldots + u_n^2 \geq 0$ además la igualdad se cumple si y sólo si $u_1 = u_2 = \ldots = u_n$, es decir sólo si $u = 0$

DEFINICION 5.2 Norma y distancia en R^n.

Dados los vectores u y v en R^n se define la distancia entre los puntos u y v escrita d(u,v) como:

$$d(u, v) = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \ldots + (u_n - v_n)^2}$$

la norma o longitud del vector u, descrita como $\|u\|$ se define como

$$\|u\| = \sqrt{u \cdot u} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

donde $u\ u \geq 0$ y entonces la raíz cuadrada existe. Observe que

$$d(u, v) = \|u - v\|$$

Es importante tener una buena apreciación del concepto de norma, dada su relación con un cierto número de desarrollos del algebra lineal. Una primera relación de interés es la que se observa entre norma y distancia. Considere los puntos $P_1 = (a, b)$ y $P_2 = (c, d)$ en R^2

$$\|P_1\| = \sqrt{a^2 + b^2} \quad \text{y} \quad d(P_1, P_2) = \sqrt{(a-c)^2 + (b-d)^2}$$

así $\|P_1\|$ corresponde a la distancia del origen al punto P_1 y $d(P_1, P_2)$ corresponde a la distancia entre los puntos P_1 y P_2

Ejemplo 5.2

Dados los vectores $u = (1, 2, 0, 3)$ y $v = (0, 3, 5, 2)$ encuentre la norma de u, la norma de v y la distancia entre ellos.

Solución:

$$\|u\| = \sqrt{(1)^2 + (2)^2 + (0)^2 + (3)^2} = \sqrt{14}$$

$$\|v\| = \sqrt{(0)^2 + (3)^2 + (5)^2 + (2)^2} = \sqrt{38}$$

$$d(u, v) = \sqrt{(1-0)^2 + (2-3)^2 + (0-5)^2 + (3-2)^2} = \sqrt{28}$$

Propiedades de la Magnitud o Norma

a) $\|v\|^2 = v \cdot v$
b) $\|rv\| = |r| \ \|v\|$

TEOREMA 5.3 DESIGUALDAD DE CAUCHY-SCHWARZ

Sean v y w vectores en R^n, entonces

$$|v \cdot w| \leq \|v\| \ \|w\|$$

demostración

Para escalares r y s cualesquiera se tiene

$$\|rv + sw\|^2 \geq 0 \qquad (5.1)$$

$\|rv + sw\|^2 = (rv + sw)\ (rv + sw) = r^2(vv) + 2rs(vw) + s^2(ww)$ (5.2)
haciendo r = ww y s = -(vw) se obtiene de las ecuaciones (5.1) y (5.2)

$$(ww)^2 (vv) - 2(ww)(vw)(vw) + (vw)^2 (ww) = (ww)^2 (vv) - (ww)(vw)^2 \geq 0$$

o $\qquad (ww)\ [(vv)(ww) - (vw)^2] \geq 0$

si (ww) = 0 entonces w = 0 y se tiene la igualdad con cero. Si (ww) $\neq$ 0 entonces ww > 0 y se tiene

$$(vv)(ww) - (vw)^2 \geq 0$$

$$(vv)(ww) \geq (vw)^2$$

o $$\|v\|^2 \ \|w\|^2 \geq (vw)^2$$

tomando raices cuadradas se tiene

$$\|v\| \ \|w\| \geq |v \cdot w|$$

Otra desigualdad importante que resulta fácilmente de la desigualdad de Cauchy- Schwarz es la desigualdad del triángulo: dados v y w wn R^n se tiene:

$$\|v + w\| \leq \|v\| + \|w\|$$

la figura siguiente indica el origen del nombre de desigualdad triangular

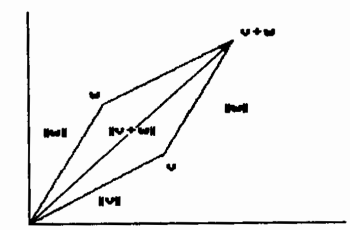


Figura 5.1

demostración

Usando las propiedades del producto punto así como la desigualdad de Cauchy- Schwarz tenemos

$$\begin{aligned}\|v + w\|^2 &= (v + w)(v + w)\\ &= v\cdot v + 2v\cdot w + w\cdot w\\ &\le v\cdot v + 2\,\|v\|\,\|w\| + w\cdot w\\ &= \|v\|^2 + 2\|v\|\,\|w\| + \|w\|^2\\ &= (\|v\| + \|w\|)^2\end{aligned}$$

y tomando raices cuadradas se tiene

$$\|v + w\| \le \|v\| + \|w\|$$

ya hemos visto algunas propiedades de los vectores en R^n, ahora procedemos a generalizarlas y definir un espacio vectorial como sigue:

DEFINICION 5.3 Un espacio vectorial real es un conjunto no vacio V de objetos llamados vectores junto con una regla para sumar dos vectores cualesquiera v y w para producir un vector $v+w \in V$ y una regla para multiplicar cualquier vector $v\in V$ por cualquier escalar $r\in R$ a fin de producir un vector $rv\in V$ llamado múltiplo escalar. Además se deben cumplir los siguientes axiomas, para todos los vectores u, v y w en V y para todos los escalares r y s

Propiedades de la suma

1. Si u y v son vectores en V entonces $u+v \in V$ (cerrado bajo la suma)
2. $u+v = v+u$ (ley conmutativa)
3. $u+(v+w) = (u+v)+w$ (ley asociativa)
4. Existe un vector $0\in V$ llamado vector cero tal que $0+u = u+0 = u$ para toda $u\in V$ (naturaleza del vector cero).
5. Para toda $u\in V$ existe $-u\in V$ tal que $u+(-u) = (-u)+u = 0$ ($-u$ como inverso aditivo de u)

Propiedades que incluyen el producto por un escalar

6. Si $r \epsilon R$ y $u \epsilon V$, entonces $ru \epsilon V$ (cerrado bajo el producto por un escalar)
7. $r(u+v) = ru + rv$ (ley distributiva)
8. $(r+s)u = ru + su$ (ley distributiva)
9. $r(su) = (rs)u$ (ley asociativa)
10. $1v = v$ (preservación de escala)

Ejemplo 5.3

El conjunto $V = R^n$ con las operaciones ordinarias de adición y multiplicación por escalares es un espacio vectorial.

Ejemplo 5.4

Sea V cualquier plano que pasa por el origen en R^3. Los puntos de V forman un espacio vectorial bajo las operaciones ordinarias para vectores en R^3. Por el ejemplo anterior se sabe que R^3 es un espacio vectorial bajo esas operaciones. Los axiomas (2) (3) (7) (8) ((9) y (10) se cumplen para todos los puntos de R^3 y por lo tanto para todos los puntos del plano V. Entonces solo es necesario demostrar que se cumplen los axiomas (1) (4) (5) y (6).

Solución:

Dado que el plano V pasa por el origen tiene una ecuación de la forma

$$ax + by + cz = 0 \tag{5.3}$$

por lo tanto, si $v= (v_1, v_2, v_3)$ y $w(w_1, w_2, w_3)$ son puntos en V entonces

$$av_1 + bv_2 + cv_3 = 0$$
$$aw_1 + bw_2 + cw_3 = 0$$

sumando éstas ecuaciones se tiene:

$$a(v_1 + w_1) + b(v_2 + w_2) + c(v_3 + w_3) = 0$$

ésta igualdad indica que las coordenadas del punto $v+w = (v_1 + w_1, v_2 + w_2, v_3 + w_3)$ satisfacen (5.3) y por lo tanto $v + w \epsilon V$, lo que muestra que el axioma 1 se cumple.

Si multiplicamos la ecuación $av_1 + bv_2 + cv_3 = 0$ por -1 resulta

$$a(-v_1) + b(-v_2) + c(-v_3) = 0$$

por lo tanto $-v \epsilon V$ y se cumple el axioma 5. La demostración de los axiomas 4 y 6 se dejan como ejercicio.

Ejemplo 5.5

Muestre que el conjunto P de todos los polinomios con coeficientes en R es un espacio vectorial, usando como suma de vectores y producto por un escalar la suma usual de polinomios y el producto usual de un polinomio por un escalar.

Solución: Sean p y q polinomios

$$p = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$$
$$q = b_0 + b_1x + b_2x^2 + \dots + b_mx^m$$

si $m \geq n$, recordemos que la suma de p y q está dada por:

$$p + q = (a_0 + b_0) + (a_1 + b_1)x + \dots + (a_n + b_n)x^n + \\ + b_{n+1}x_{n+1} + \dots + b_mx^m$$

se hace una definición similar si $m < n$. El producto de p por un escalar r está dado por:

$$rp = ra_0 + ra_1x + ra_2x^2 + \dots + ra_nx^n$$

se tiene el polinomio cero y -p y se cumplen los 8 axiomas restantes.

Ejemplo 5.6

Demuestre que el conjunto M de todas las matrices de mxn es un espacio vectorial, usando como suma de vectores y multiplicación por un escalar la suma y producto por un escalar usuales para las matrices.

Solución:
La suma de matrices de mxn y el producto de una matriz de mxn por un escalar produce de nuevo una matriz de mxn. Así M es cerrado bajo la suma de vectores y el producto por un escalar. Tomamos como vector cero en M la matriz cero usual, cuyas entradas son todas cero. Para cualquier matriz A en M consideremos -A como la matriz (-1)A.

Ejemplo 5.7

Este ejemplo muestra que no necesariamente la suma vectorial ha de estar relacionada con la suma ordinaria, y que el vector cero no necesariamente es el número real cero.

Sea $V = \{ x \mid x \in R, x > 0 \}$ la suma y producto escalar se definen como:

$$x \oplus y = xy$$

suma vectorial producto ordinario

y para $r \in R$, $x \in V$, se define

$$r \odot x = \underbrace{x^r}_{\text{potencia ordinaria}}$$

Muestre que V con éstas operaciones es un espacio vectorial.

Solución:

Como los elementos de V son $x \in R$ entonces se tiene

$$2 \oplus 3 = 2 \cdot 3 = 6$$
$$2 \odot 3 = 3^2 = 9$$
$$-2 \odot 3 = 3^{-2} = 1/9$$

entonces tenemos que demostrar que se satisfacen los axiomas:

1. Cerradura para la suma

Sean $x \in V$, $y \in V$ entonces $x \oplus y = xy$ como $x, y \in R^+$ y como el producto de dos reales positivos es otro real positivo $xy \in V$ entonces $x \oplus y \in V$

2. Cerradura para la multiplicación por un escalar. Sean $r \in R$ y $x \in V$ $r \odot x = x^r$ como $x \in R^+$ al elevarlo a cualquier potencia real se obtiene un número real positivo. En consecuencia $r \odot x \in V$.

3. Conmutatividad de la suma

$$x \oplus y = xy = yx = y \oplus x$$

por lo tanto $x \oplus y = y \oplus x$

4. Asociatividad

$$(x \oplus y) \oplus z = (xy) \oplus z = (xy)z = x(yz) = x \oplus (yz) = x \oplus (y \oplus z)$$

5. Idéntico aditivo

$x \oplus \theta = x$ por definición $x \oplus \theta = x\theta$ de donde $\theta = 1$, es decir el real positivo 1 es el idéntico aditivo o cero para este espacio

$$x \oplus 1 = x$$

6. Inverso aditivo. Para toda $x \in V$, $-x$ debe satisfacer $x \oplus (-x) = \theta$ pero $-x$ no es "menos x" es el símbolo para el inverso aditivo y θ no es el cero sino 1 entonces:

$$x \oplus (-x) = \theta$$

es equivalente a

$$x \cdot (-x) = 1$$

de donde $-x = 1/x$ como $x > 0$ entonces $1/x > 0$ por lo tanto $-x \epsilon V$

7. Ley asociativa para la multiplicación escalar. Sea $r \epsilon R$, $s \epsilon R$, $x \epsilon V$ entonces:

$$(rs) \odot x = x^{rs} \quad y \quad r \odot (s \odot x) = r \odot (x^s) = (x^s)^r$$

que por las leyes de los exponentes

$$(x^s)^r = x^{sr} = x^{rs} \text{ en consecuencia}$$

$$(rs) \odot x = r \odot (s \odot x)$$

8 y 9. Leyes distributivas. Sean $r \epsilon R$, $s \epsilon R$, $x \epsilon V$, $y \epsilon V$ entonces

$(r + s) \odot x = x^{r+s} = x^r \, x^s = x^r \oplus x^s = (r \odot x) \oplus (s \odot x)$

$r \odot (x \oplus y) = r \odot (xy) = (xy)^r = x^r \, y^r = x^r \oplus y^r = (r \odot x) \oplus (r \odot y)$

10. Sea $x \epsilon V$, entonces $1 \odot x = x^1 = x$

5.3 SUBESPACIOS

Hemos visto que R^n es un espacio vectorial en el que la suma de vectores consiste en sumar las componentes correspondientes de tales vectores y el producto por un escalar se obtiene multiplicando cada componente por el escalar. Cualquier subconjunto de R^n que con éstas mismas operaciones sea también un espacio vectorial, se llama subespacio de R^n .

Por ejemplo, el plano $x_1 x_2$ de R^3 formado por todos los vectores que tienen cero como tercera componente, como se puede ver en la figura 5.2

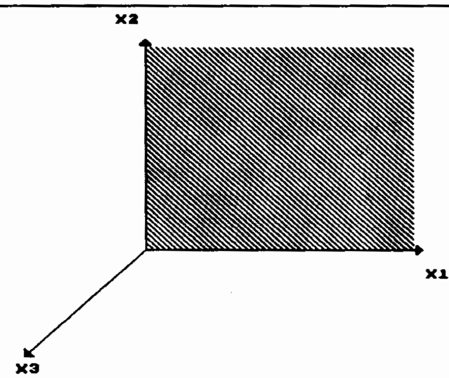


Figura 5.2

es un subespacio de R^3 sin embargo, el subconjunto $\{(m,n,p) \mid m,n,p \in Z\}$ de R^3 que consta de todos los vectores de R^3 con componentes enteras no es un subespacio, pues aunque éste subconjunto es cerrado bajo la suma de vectores, no es cerrado bajo el producto por un escalar, por ejemplo:

$0.5\ (1,\ 2,\ 5) = (0.5,\ 1,\ 2.5)$ no está en el subconjunto

<u>DEFINICION 5.4</u> Un subconjunto W de un espacio vectorial V es un subespacio de V si W es por si solo un espacio vectorial bajo las operaciones de suma de vectores y de producto por un escalar definidas en V.

Esto significa que el conjunto W debe ser cerrado bajo la suma y producto por un escalar. Las propiedades requeridas por un espacio vectorial se cumplen para W pues son válidas para todo V.

<u>TEOREMA 5.3 CRITERIO PARA UN SUBESPACIO</u>

Un subconjunto no vacío W de un espacio vectorial V es un subespacio de V si y solo si satisface las siguientes condiciones

a) Si u y v son vectores de W entonces u+v está en W
b) Si r es cualquier escalar de R y v está en W, entonces rv está en W

Las condiciones a) y b) indican que W es cerrado bajo la suma y bajo el producto por un escalar.

<u>demostración</u>.

-->) Si W es un subespacio de V, entonces satisface todos los axiomas de espacio vectorial, en particular, cumple con los axiomas 1 y 6 pero estos axiomas son las condiciones a y b.

<--) Suponemos que se cumplen a y b. Dado que estas condiciones son los axiomas 1 y 6 de espacio vectorial, solamente resta demostrar que W satisface los otros 8 axiomas. Los axiomas 2, 3, 7, 8, 9, 10 se satisfacen automáticamente ya que todos los vectores en V los satisfacen. Para completar la demostración sólo falta verificar que W satisface los axiomas 4 y 5.

4) Sea $u \in W$, por la condición b, $ru \in W$ para toda $r \in R$. Haciendo $r = -1$ se tiene $(-1)u \in W$ de donde $(-1)u + u \in W$ por a) y esto implica que $(-1)u + u = 0$ por lo que $0 \in W$

5) usando 4 consideramos $r = -1$ y $(-1)u = -u \in W$

Todo espacio vectorial V tiene al menos 2 subespacios. V es un subespacio de si mismo y el conjunto {0} que consiste únicamente del vector cero en V, este subespacio recibe el nombre de subespacio trivial. Los subespacios que no son ni {0} ni V se conocen como <u>subespacios propios</u>

El teorema 5.3 dice que para probar que W es un subespacio de V basta verificar que se cumplen las condiciones a y b.

Algunos ejemplos de subespacios son:

<u>Ejemplo 5.8</u>

Para todo espacio vectorial V, el subconjunto {0} que contiene solamente el vector cero, es un subespacio, puesto que:

$$0 + 0 = 0$$
$$r0 = 0 \text{ para toda } r \in R$$

y se le conoce como subespacio trivial.

<u>Ejemplo 5.9</u>

El conjunto M_{22}^0 de todas las matrices de 2x2 que tiene ceros en la diagonal principal es un subespacio del espacio vectorial M_{22} de todas las matrices de orden 2x2.

<u>Solución:</u>

Sean A y B dos matrices en M_{22}^0 como sigue:

$$A = \begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 0 & b_{12} \\ b_{21} & 0 \end{bmatrix}$$

entonces rA y A+B son iguales a:

$$rA = \begin{bmatrix} 0 & ra_{12} \\ ra_{21} & 0 \end{bmatrix} \qquad A+B = \begin{bmatrix} 0 & a_{12} + b_{12} \\ a_{12} + b_{21} & 0 \end{bmatrix}$$

que como se ve tienen ceros en la diagonal principal, entonces pertenecen a M_{22}^0. Por lo tanto M_{22}^0 es un subespacio de M_{22}.

<u>Ejemplo 5.10</u>

Sea Ax = 0 un sistema homogéneo de m ecuaciones lineales con n incógnitas y sea W el conjunto de todos los vectores solución del sistema y sean u y v dos vectores en W. Se demostrará que W es un subespacio de R^n

Solución:

Se dice que un vector $v \epsilon R^n$ es un vector solución del sistema si $x_1 = v_1$, $x_2 = v_2, \ldots, x_n = v_n$ es una solución del sistema.

Para mostrar que es cerrado bajo la suma y bajo el producto por un escalar, es necesario demostrar que u+v y ru son vectores solución.

dado que u y v son vectores solución se tiene que:

$$Au = 0 \quad y \quad Av = 0$$

y de aquí $\quad A(u + v) = Au + Av = 0 + 0 = 0$

entonces $u + v$ satisface $Ax = 0$ lo que implica $u + v \epsilon W$
además

$$A(ru) = r(Au) = r0 = 0$$

se puede concluir que ru satisface $Ax = 0$ entonces $ru \epsilon W$, por lo tanto W es un subespacio de R^n

Ejemplo 5.11

Sea $H = \{ (x, y) \mid y = 2x \quad x,y \epsilon R \}$ esto es, H consiste de los vectores en R^2 que están sobre una recta que pasa por el origen. Verificaremos que H es un subespacio de R^2

Solución:

Sean $\quad a = (x_1, y_1) \epsilon H$ entonces $y_1 = 2x_1$
$\qquad b = (x_2, y_2) \epsilon H$ entonces $y_2 = 2x_2$

$$a+b = (x_1 + x_2, y_1 + y_2) = (x_1 + x_2, 2x_1 + 2x_2) = (x_1 + x_2, 2(x_1 + x_2))$$

de aquí se concluye que $a+b \epsilon H$

$$ra = r(x_1, y_1) = (rx_1, ry_1) = (rx_1, r2x_1) = (rx_1, 2(rx_1))$$

de aquí se concluye que $ra \epsilon H$ y por lo tanto H es un subespacio de R^2

Ejemplo 5.12

Considere todos los vectores en R^2 cuyas componentes sean positivas o cero. Si el espacio vectorial V es R^2 con las operaciones usuales de suma y producto para vectores entonces este subconjunto es el primer cuadrante, las coordenadas satisfacen $x \geq 0$, $y \geq 0$.

Este subconjunto sin embargo no es un subespacio ya que aunque contenga al cero y sea cerrado bajo la suma, no es cerrado bajo el producto ya que si $r = -1$ se tiene $r(x, y) = -1(x, y) = (-x, -y)$ que se encuentra en el tercer cuadrante. Si se incluye el tercer cuadrante además del primero, no es cerrado bajo la suma ya que:

$(1, 2) + (-2, -1) = (-1, 1)$ que está en el segundo cuadrante.

Por lo tanto el menor subespacio que contiene al primer cuadrante es todo el espacio R^2

Nota 1: No todo espacio vectorial tiene subespacios propios.

Ejemplo 5.13

Sea W un subespacio de R. Si $W \neq \{0\}$ entonces W contiene un número real $\alpha \neq 0$ por el axioma 6 existe $1/\alpha$ tal que

$$(1/\alpha)\ \alpha = 1 \in W$$

y

$$1\ \alpha = \alpha \in W \text{ para toda } \alpha \in R$$

así si W no es el subespacio trivial entonces $W = R$, esto es, R no tiene subespacios propios.

Se puede pensar que cualquier recta o plano de R^3 es un subespacio, sin embargo, para que esto sea cierto, este subespacio de R^3 debe contener al cero. Geométricamente significa que esta recta o plano debe pasar por el origen.

5.4 COMBINACIONES LINEALES

Como ya se decía, un plano o una recta en R^3 es un subespacio sólo si pasa por el origen. Un plano que pasa por el origen está totalmente determinado por dos vectores cualesquiera v_1 y v_2 distintos de cero, y no paralelos que están en el, como se ve en la figura 5.3

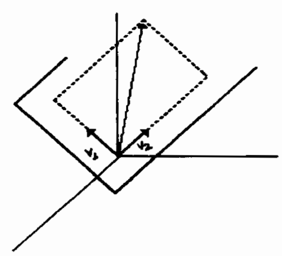


Figura 5.3

la figura muestra que cualquier vector de este plano tiene la forma

$$r_1 v_1 + r_2 v_2$$

ésta expresión es una combinación lineal de los vectores v_1 y v_2

<u>DEFINICION 5.5</u> Se dice que un vector w es un combinación lineal de los vectores $v_1, v_2, \ldots, v_n$ si es posible expresarlo en la forma:

$$w = r_1 v_1 + r_2 v_2 + \ldots + r_n v_n$$

donde $r_1, r_2, \ldots, r_n$ son escalares.

<u>Ejemplo 5.14</u>

Sean los vectores u = (1, 2, -1) y v = (6, 4, 2) en R^3. Demuestre que w = (9, 2, 7) es una combinación lineal de u y v.

<u>Solución:</u>

Para que w sea una combinación lineal de u y v, deben existir escalares r_1 y r_2 tales que

$$w = r_1 u + r_2 v$$

es decir

$$(9, 2, 7) = r_1 (1, 2, -1) + r_2 (6, 4, 2)$$

que equivale a

$$(9, 2, 7) = (r_1 + 6r_2, 2r_1 + 4r_2, -r_1 + 2r_2)$$

igualando las componentes correspondientes resulta:

$$\begin{aligned} r_1 + 6r_2 &= 9 \\ 2r_1 + 4r_2 &= 2 \\ -r_1 + 2r_2 &= 7 \end{aligned}$$

resolviendo éste sistema se obtiene $r_1 = -3$ y $r_2 = 2$ de donde

$$w = -3u + 2v$$

en general definimos:

<u>DEFINICION 5.6</u> Sea { $v_1, v_2, \ldots, v_n$ } un conjunto de vectores en el espacio vectorial V. El espacio generado por { $v_1, v_2, \ldots, v_n$ } es el conjunto de todas las combinaciones posibles de $v_1, v_2, \ldots, v_n$ y se denota

$$\text{gen}(v_1, v_2, \ldots, v_n) = \{r_1 v_1 + r_2 v_2 + \ldots + r_n v_n \mid r_i \in R \ i=1,2,\ldots,n\}$$

Ejemplo 5.15

Para el ejemplo del plano en R^3 tendríamos que el conjunto de todas las combinaciones lineales de v_1 y v_2 sería

$$gen(v_1,\ v_2) = (\ r_1 v_1 + r_2 v_2 \mid r_1,\ r_2 \in R\ )$$

y geométricamente

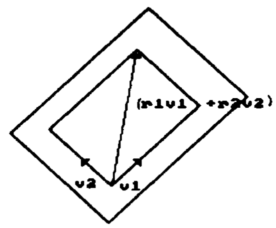


Figura 5.4

además $gen(v_1, v_2, \ldots, v_n) \subset V$ es un subespacio de V

TEOREMA 5.4

Si $S = \{\ v_1,\ v_2, \ldots, v_n\ \} \subset V$ siendo V un espacio vectorial, entonces gen S es un subespacio de V

demostración:

Sean x, y $\in$ gen S, entonces x, y son combinaciones lineales

$$x = r_1 v_1 + \ldots + r_n v_n$$
$$y = s_1 v_1 + \ldots + s_n v_n$$

entonces

$$x + y = (r_1 + s_1)v_1 + \ldots + (r_n + s_n)v_n$$

y

$$kx = (kr_1)v_1 + \ldots + (kr_n)v_n$$

que también se encuentra en gen S. Por lo tanto gen S es un subespacio de V. Se le llama subespacio de V generado por S.

Ejemplo 5.16

Determine si $v_1 = (1, 1, 2)$, $v_2 = (1, 0, 1)$ y $v_3 = (2, 1, 3)$ generan a R^3

Solución:

Es necesario determinar si un vector arbitrario $b = (b_1, b_2, b_3)$ en R^3 se puede expresar como una combinación lineal de v_1, v_2 y v_3 esto es:

$$b = r_1 v_1 + r_2 v_2 + r_3 v_3$$

en términos de sus componentes resulta:

$$(b_1, b_2, b_3) = (r_1 + r_2 + 2r_3, r_1 + r_3, 2r_1 + r_2 + 3r_3)$$

o bien

$$\begin{aligned} r_1 + r_2 + 2r_3 &= b_1 \\ r_1 + \quad + r_3 &= b_2 \\ 2r_1 + r_2 + 3r_3 &= b_3 \end{aligned}$$

entonces el problema se reduce a determinar si éste sistema es consistente para todos los valores de b_1, b_2 y b_3. Esto es equivalente a que la matriz de los coeficientes sea invertible.

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

o lo que es lo mismo det $A \neq 0$, como det $A = 0$, A no es invertible y por lo tanto v_1, v_2 y v_3 no generan a R^3.

Ejemplo 5.17

Describa geométricamente el subespacio de R^4 generado por el vector $v = (2, 3, -1, 4)$

Solución:

El subespacio gen(v) es una recta de R^4 que contiene al origen, cada punto de esta recta es un múltiplo escalar rv de v, esto es, cada punto tiene la forma (x_1, x_2, x_3, x_4) donde $x_1 = 2r$, $x_2 = 3r$, $x_3 = -r$ y $x_4 = 4r$ para algún escalar r.

Nota 2: Que un determinado vector u se pueda expresar como una combinación lineal de otro u otros vectores v_i se puede ver también, como el hecho de que ese vector u pertenezca al espacio generado por los vectores v_i.

Ejemplo 5.18

¿Es posible escribir (3, -1, 4) como combinación lineal de (0, 1, 1), (1, -1, 0) y (3, -5, -2)? o lo que es lo mismo ¿está (3, -1, 4) en gen {(1, -1, 0), (0, 1, 1), (3, -5, -2)}?

Solución:

En cualquier caso se tiene:

$$(3,\ -1,\ 4) = r_1\ (1,\ -1,\ 0) + r_2\ (0,\ 1,\ 1) + r_3\ (3,\ -5,\ -2)$$

o equivalentemente:

$$\begin{aligned} 3 &= r_1 + \qquad + 3r_3 \\ -1 &= -r_1 + r_2 - 5r_3 \\ 4 &= \qquad\ \ r_2 - 2r_3 \end{aligned}$$

escribiendo en forma matricial y efectuando eliminación gaussiana se tiene:

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & 3 \\ 0 & 1 & -2 & 2 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

entonces no existe solución y (3, -1, 4) no se puede escribir como combinación lineal de los vectores dados.

5.5 INDEPENDENCIA LINEAL

Hemos visto que un conjunto de vectores gen(S) genera un espacio vectorial V si S C V y cada vector en V es una combinación lineal de v_1, $v_2, \ldots, v_r$. Los conjuntos de generadores son muy útiles ya que en muchos casos es posible estudiar un espacio vectorial V si primero se trabaja con los vectores que pertenecen a algún conjunto de generadores S y después se aplican los resultados a la totalidad de V. El problema de encontrar los conjuntos mas pequeños de generadores para un espacio vectorial depende del concepto de independencia lineal.

Ejemplo 5.19

Determine si gen {(1, -1, 0), (0, 1, 1), (3, -5, -2)} genera todo R^3

Solución:

Sea $(x_1,\ x_2,\ x_3)$ un vector cualquiera en R^3, entonces queremos saber si es posible escribir

$$(x_1,\ x_2,\ x_3) = r_1\ (1,\ -1,\ 0) + r_2\ (0,\ 1,\ 1) + r_3\ (3,\ -5,\ -2)$$

que es equivalente a

$$\left[\begin{array}{ccc|c} 1 & 0 & 3 & x_1 \\ -1 & 1 & -5 & x_2 \\ 0 & 1 & -2 & x_3 \end{array}\right]$$

la cual se reduce a

$$\left[\begin{array}{ccc:l} 1 & 0 & 3 & x_1 \\ 0 & 1 & -2 & x_1 + x_2 \\ 0 & 0 & 0 & x_3 - x_2 - x_1 \end{array}\right]$$

hay una solución sólo si $x_3 - x_2 - x_1 = 0$, pero esto impone una restricción a (x_1, x_2, x_3) y por lo tanto la primera de las ecuaciones no se puede resolver para un vector arbitrario (x_1, x_2, x_3), esto significa que este conjunto de vectores no genera R^3 lo que no implica que no se pueda generar otro espacio con éste conjunto, entonces cabe pedir: describa gen(S) donde

$$S = \{(1, -1, 0), (0, 1, 1), (3, -5, -2)\}$$

Suponga que (x_1, x_2, x_3) está en gen(S) entonces la ecuación

$$(x_1, x_2, x_3) = r_1 (1, -1, 0) + r_2 (0, 1, 1) + r_3 (3, -5, -2)$$

debe tener solución, pero se tiene que $x_3 - x_2 - x_1 = 0$ por lo cual gen(S) = $\{ (x_1, x_2, x_3) \mid x_3 = x_1 + x_2\}$ Es decir, el espacio generado por S consta de todos los vectores cuya tercera componente es la suma de las dos primeras, así por ejemplo (1, 3, 5) no esta en gen(S) y (1, 3, 4) ϵ gen(S).

Por ejemplo R^2 esta generado por S = { (1, 0), (0, 1)} pero también puede generarse por el conjunto mayor S' = { (1, 0), (0, 1), (1, 1)}. R^2 y funciones sobre R^2 se pueden analizar a través de los conjuntos generadores, así entonces por economía, habría que encontrar los conjuntos generadores mas pequeños. Para lograr esto se requiere del concepto de <u>independencia lineal</u>

<u>DEFINICION 5.7</u> Un conjunto $\{v_1, v_2, \ldots, v_n\}$ de vectores en un espacio vectorial V se llama linealmente independiente si

$$r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = 0$$

tiene como única solución $r_1 = r_2 = \ldots = r_n = 0$, es decir la solución trivial.

<u>DEFINICION 5.8</u> Un conjunto $\{v_1, v_2, \ldots, v_n\}$ de vectores en un espacio vectorial V se llama linealmente dependiente si

$$r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = 0$$

tiene una solución no trivial, es decir, existen escalares $r_1, r_2, \ldots, r_n$ no todos cero tales que el sistema descrito tiene solución

<u>Ejemplo 5.20</u>

En cada uno de los espacios vectoriales especificados determine si los vectores listados son linealmente independientes o dependientes

a) $V = R^2$ con $S = \{ x_1 = (2, 1), x_2 = (6, 3), x_3 = (0, 1)\}$

geométricamente se tiene:

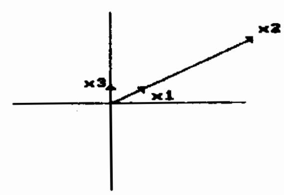


Figura 5.5

Note que

$$-3x_1 + x_2 + 0x_3 = 0$$

de donde x_1, x_2, x_3 son linealmente dependientes

otra forma sería

$$r_1\, x_1 + r_2\, x_2 + r_3\, x_3 = 0$$

o

$$r_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} 6 \\ 3 \end{bmatrix} + r_3 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

y en forma matricial

$$\begin{bmatrix} 2 & 6 & 0 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

de donde

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

lo que implica $r_3 = 0$, $r_2 = t$, $r_1 = -3t$ $\quad t \in R$
por lo tanto x_1, x_2, x_3 son linealmente dependientes.

b) Si $V = R^3$ con

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad x_2 = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \qquad x_3 = \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}$$

entonces

$$r_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + r_2 \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} + r_3 \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

o bien

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ 1 & 1 & 4 & 0 \end{array}\right]$$

como se tiene una matriz triangular inferior es invertible y además la solución es la trivial.

$$r_1 = r_2 = r_3 = 0$$

por lo cual x_1, x_2, y x_3 son linealmente independientes.

c) Si $V = M_{2x2}$ con

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad M_2 = \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} \qquad M_3 = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}$$

entonces escribimos

$$r_1 M_1 + r_2 M_2 + r_3 M_3 = 0$$

esto es

$$r_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + r_2 \begin{bmatrix} 4 & 0 \\ 0 & 0 \end{bmatrix} + r_3 \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

o bien

$$\begin{bmatrix} r_1 + 4r_2 & r_3 \\ r_3 & r_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

de aquí

$$r_1 + 4r_2 = 0$$
$$r_3 = 0$$

lo que implica que $r_1 = -4r_2$ si $r_2 = t$ entonces $r_1 = -4t$ $t \epsilon R$ por lo tanto M_1, M_2, M_3 son linealmente dependientes.

Ejemplo 5.21

Determine si el conjunto formado por los vectores $v_1 = (1, -2, 3)$, $v_2 = (5, 6, -1)$ y $v_3 = (3, 2, 1)$ es linealmente dependiente o independiente.

Solución:

En términos de los componentes, la ecuación vectorial

$$r_1 v_1 + r_2 v_2 + r_3 v_3 = 0$$

se transforma en

$$r_1 (1, -2, 3) + r_2 (5, 6, -1) + r_3 (3, 2, 1) = (0, 0, 0)$$

o el sistema:

$$\begin{aligned} r_1 + 5r_2 + 3r_3 &= 0 \\ -2r_1 + 6r_2 + 2r_3 &= 0 \\ 3r_1 - r_2 + r_3 &= 0 \end{aligned}$$

resolviendo éste sistema se tiene

$$r_1 = -\tfrac{1}{2} t \qquad r_2 = -\tfrac{1}{2} t \qquad r_3 = t$$

por lo que el sistema tiene soluciones no triviales y en consecuencia v_1, v_2 y v_3 son linealmente dependientes.

Nota 3: El término "linealmente dependiente" sugiere que los vectores "dependen" unos de otros de alguna manera. Para ver ésta dependencia, suponga que $v = \{v_1, v_2, \ldots, v_r\}$ es un conjunto linealmente dependiente por lo tanto, la ecuación

$$r_1 v_1 + r_2 v_2 + \ldots + r_n v_n = 0$$

tiene otra solución además de $r_1 = r_2 = \ldots = r_r = 0$. Supongamos que $r_1 \neq 0$ entonces multiplicando esta ecuación por $1/r_1$ y resolviendo para v_1 se obtiene:

$$v_1 = \left(-\frac{r_2}{r_1}\right) v_2 + \ldots + \left(-\frac{r_r}{r_1}\right) v_r$$

por lo que v_1 se puede expresar como una combinación lineal de los vectores restantes $v_2, \ldots, v_r$. Por lo que podemos concluir que: un conjunto de vectores es linealmente dependiente si y solo si al menos uno de los vectores es una combinación lineal de los restantes.

Ejemplo 5.22

Dos vectores v_1 y v_2 forman un conjunto l.d. si y solo si uno de los vectores es un múltiplo escalar del otro. Para ver esto, suponga que $S = \{v_1, v_2\}$ es linealmente dependiente. Dado que la ecuación vectorial

$$r_1 v_1 + r_2 v_2 = 0$$

tiene una solución además de la trivial $r_1 = r_2 = 0$, ésta ecuación se puede expresar como:

$$v_1 = -\frac{r_2}{r_1} v_2 \qquad o \qquad v_2 = -\frac{r_1}{r_1} v_1$$

estas ecuaciones indican que v_1 es múltiplo escalar de v_2 y viceversa. Por consecuencia, dos vectores en R^2 o en R^3 son l.d. si y solo si pertenecen a la misma recta que pasa por el origen

para R^3 decimos que si v_1, v_2 y v_3 son tres vectores en R^3 entonces el conjunto $S = \{v_1, v_2, v_3\}$ es l.d. si y sólo si los tres vectores pertenecen al mismo plano que pasa por el origen.

TEOREMA 5.5

Sea $V = \{ v_1, v_2, \ldots, v_r \}$ un conjunto de vectores en R^n. Si $r > n$, entonces V es linealmente dependiente.

demostración:

Suponga que

$$\begin{aligned} v_1 &= (v_{11}, v_{12}, \ldots, v_{1n}) \\ v_2 &= (v_{21}, v_{22}, \ldots, v_{2n}) \\ &\cdots\cdots\cdots\cdots\cdots \\ v_r &= (v_{r1}, v_{r2}, \ldots, v_{rn}) \end{aligned}$$

considere la ecuación

$$r_1 v_1 + r_2 v_2 + \ldots + r_r v_r = 0$$

expresando ambos lados de la ecuación en términos de sus componentes se tiene que:

$$\begin{aligned} v_{11} r_1 + v_{21} r_2 + \ldots + v_{r1} r_r &= 0 \\ v_{12} r_1 + v_{22} r_2 + \ldots + v_{r2} r_r &= 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ v_{1n} r_1 + v_{2n} r_2 + \ldots + v_{rn} r_r &= 0 \end{aligned}$$

este es un sistema homogéneo de n ecuaciones con r incógnitas. Dado que $r > n$ entonces por el resultado que dice, que si un sistema homogéneo de ecuaciones lineales tiene mas incógnitas que ecuaciones, siempre tiene un número infinito de soluciones, lo que implica que el sistema tiene soluciones no triviales y por lo tanto S es un conjunto l.d.

5.6 BASES Y DIMENSION

El conjunto de vectores $S = \{(1, 1), (1, -1)\}$ genera a R^2. Es decir, cualquier vector de R^2 es una combinación lineal de (1, 1), (1,-1). El conjunto de vectores

$$T = \{ (1, 1), (1, -1), (1, 0)\}$$

también genera a R^2. Los conjuntos S y T difieren en que S es linealmente independiente mientras que T es linealmente dependiente. Esto hace que haya una diferencia al expresar un vector como una combinación lineal de los vectores de cada conjunto. Por ejemplo, al expresar (2, 4) en términos de los vectores en S, se tiene como única posibilidad

$$(2, 4) = 3(1, 1) - 1(1, -1)$$

sin embargo, en términos de los vectores de T, hay varias posibilidades:

$$(2, 4) = 3(1, 1) - 1(1, -1) + 0(1, 0)$$
$$(2, 4) = 0(1, 1) - 4(1, -1) + 6(1, 0)$$
$$(2, 4) = 4(1, 1) + 0(1, -1) - 2(1, 0)$$

o en general

$$(2, 4) = (k+4)(1, 1) + k(1, -1) + (-2-2k)(1, 0)$$

La conclusión es: si un conjunto S de vectores genera V y S es linealmente dependiente, entonces la representación de un vector x en términos de los vectores en S no es única. Si se desea unicidad, el conjunto generador también deberá ser linealmente independiente. Un conjunto con esas características recibe el nombre de <u>base</u> para V. Las bases se emplean en teoría de códigos entre otras áreas.

<u>DEFINICION 5.9</u> Si V es un espacio vectorial y $S = \{v_1, v_2, \ldots, v_r\}$ es un conjunto finito de vectores en V, entonces se dice que S es una base de V si

a) S es linealmente independiente
b) S genera a V

Ejemplo 5.23

Sean $v_1 = (1, 2)$ y $v_2 = (3, -1)$. Demuestre que el conjunto S = { v_1, v_2 } es una base de R^2

Solución:

Se tiene que demostrar que S es linealmente independiente y que genera a R^2

Para demostrar que S genera a R^2, se debe mostrar que un vector arbitrario $b = (b_1, b_2)$ se puede expresar como una combinación lineal de v_1 y v_2, esto es

$$r_1 (1, 2) + r_2 (3, -1) = (b_1, b_2) \qquad (5.4)$$

tiene una solución única para cualquier $b = (b_1, b_2)$

Y para demostrar que v_1 y v_2 son linealmente independientes se debe demostrar que

$$r_1 (1, 2) + r_2 (3, -1) = (0, 0) \qquad (5.5)$$

tiene solamente la solución $r_1 = r_2 = 0$. Las ecuaciones (5.4) y (5.5) escritas en términos de sus componentes son iguales a:

$$\begin{aligned} r_1 + 3r_2 &= b_1 \\ 2r_1 - r_2 &= b_2 \end{aligned}$$

y

$$\begin{aligned} r_1 + 3r_2 &= 0 \\ 2r_1 - r_2 &= 0 \end{aligned}$$

en forma matricial se tiene:

$$\left[\begin{array}{cc:c} 1 & 3 & b_1 \\ 2 & -1 & b_2 \end{array}\right] \qquad \left[\begin{array}{cc:c} 1 & 3 & 0 \\ 2 & -1 & 0 \end{array}\right]$$

lo que se puede resolver escribiendo

$$\left[\begin{array}{cc:c:c} 1 & 3 & b_1 & 0 \\ 2 & -1 & b_2 & 0 \end{array}\right]$$

usando eliminación gaussiana se tiene:

la independencia lineal se comprueba con esta parte

$$\left[\begin{array}{cc|c|c} 1 & 3 & b_1 & 0 \\ 2 & -1 & \dfrac{-b_2 + 2b_1}{7} & 0 \end{array}\right]$$

El espacio generado se comprueba con esta parte

se tiene entonces que para demostrar que S genera se tiene:

$$r_2 = \frac{2b_1 - b_2}{7} \qquad r_1 = \frac{b_1 + 3b_2}{7}$$

por lo cual dando valores arbitrarios a b_1 y b_2 se obtienen los valores para r_1 y r_2 que multiplicados por v_1 y v_2 pueden expresar como combinación lineal a b_1 y b_2

Para la independencia lineal se tiene que

$$\begin{aligned} r_1 + 3r_2 &= 0 \\ r_2 &= 0 \end{aligned}$$

lo que implica $r_1 = r_2 = 0$ por lo tanto S es una base de R^2

<u>TEOREMA 5.6</u>

Sea $S = \{ v_1, v_2, \ldots, v_n \}$ una base para el espacio vectorial V, y sea $v \in V$. Los coeficientes en la representación

$$v = r_1 v_1 + \ldots + r_n v_n$$

son únicos.

<u>demostración:</u>

Supóngase que se tienen las dos representaciones

$$\begin{aligned} v &= a_1 v_1 + \ldots + a_n v_n \\ v &= b_1 v_1 + \ldots + b_n v_n \end{aligned}$$

para v; se demostrará que los coeficientes son iguales. Para ello, formamos v + (-v) que es igual a cero y se combinan los términos para obtener

$$0 = (a_1 - b_1)v_1 + \ldots + (a_n - b_n)v_n$$

como S es una base, es un conjunto l.i. por lo tanto

$$a_1 - b_1 = a_2 - b_2 = \ldots = a_n - b_n = 0$$

lo que implica

$$a_1 = b_1, \quad a_2 = b_2, \ldots, a_n = b_n$$

Ejemplo 5.24

Sean $v_1 = (1, 2, 1)$, $v_2 = (2, 9, 0)$ y $v_3 = (3, 3, 4)$. Demuestre que el conjunto $S = \{v_1, v_2, v_3\}$ es una base de R^3

Solución:

Entonces se tiene que demostrar que para cualquier vector $b = (b_1, b_2, b_3)$ la ecuación

$$r_1 v_1 + r_2 v_2 + \ldots + r_3 v_3 = b$$

siempre tiene solución, y que la única solución de

$$r_1 v_1 + r_2 v_2 + \ldots + r_3 v_3 = 0$$

es la trivial, esto es $r_1 = r_2 = r_3 = 0$

De estos dos sistemas podemos concluir, que la matriz de coeficientes asociada

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{bmatrix}$$

es invertible, esto es que existe solución única para $Ar = 0$ y para $Ar = b$. Que A es invertible se puede verificar con su determinante, si $\det A \neq 0$ entonces A es invertible.

$$\det A = \begin{vmatrix} 1 & 2 & 3 \\ 2 & 9 & 3 \\ 1 & 0 & 4 \end{vmatrix} = -1$$

por lo tanto S es una base de R^3

TEOREMA 5.7

Sean $v_1, v_2, \ldots, v_n$ vectores en R^n las siguientes condiciones son equivalentes

1. Los vectores son linealmente independientes
2. Los vectores generan todo R^n
3. La matriz A que tiene estos vectores como vectores columna es invertible.

Ejemplo 5.25

Sea el conjunto S = { M_1, M_2, M_3} tal que

$$M_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad M_2 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad M_4 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

demuestre que S es una base del espacio vectorial M_{22} de las matrices de 2x2

Solución

Para comprobar que S genera a M_{22}, observe que, dada una matriz arbitraria

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

se puede expresar como:

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

$$= a\, M_1 + b\, M_2 + c\, M_3 + d\, M_4$$

es decir, una combinación lineal de M_1, M_2, M_3, y M_4

Para verificar que S es l.i. hacemos

$$a\, M_1 + b\, M_2 + c\, M_3 + d\, M_4 = 0$$

o bien

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

entonces

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

por lo tanto $a = b = c = d = 0$ y S es linealmente independiente.

Ejemplo 5.26

Muestre que el conjunto $S = \{ (1, 2), (3, -1), (1, 0)\}$ no es una base de R^2.

Solución:

El conjunto S es l.d. porque:

$$(1, 2) + 2(3, -1) - 7(1, 0) = (0, 0)$$

por lo cual S no es una base de R^2

TEOREMA 5.8 REDUCCION DE UN CONJUNTO GENERADOR A UNA BASE

Si $S = \{v_1, \ldots, v_m\}$ es un conjunto de vectores no nulos que genera a un subespacio W de un espacio vectorial V, entonces algún subconjunto de S es una base para W. Dicha base se puede encontrar excluyendo de S aquellos vectores que son combinaciones lineales de predecesores.

demostración:

Si S es un conjunto linealmente independiente, entonces por definición S es una base para W.

Si S es linealmente dependiente, uno de los vectores puede escribirse como una combinación lineal de los otros. Supóngase que v_m es dicho vector. Se afirma que $S' = \{ v_1, \ldots, v_{m-1} \}$ aún genera a W. Para comprobar ésto, sea x un elemento de W con

$$x = c_1 v_1 + \ldots + c_{m-1} v_{m-1} + c_m v_m$$

pero $v_m = d_1 v_1 + \ldots + d_{m-1} v_{m-1}$ por lo que ésta expresión se puede sustituir en la combinación lineal anterior para obtener

$$x = c_1 v_1 + \ldots + c_{m-1} v_{m-1} + c_m (d_1 v_1 + \ldots + d_{m-1} v_{m-1}) =$$
$$= (c_1 + c_m d_1) v_1 + \ldots + (c_{m-1} + c_m d_{m-1}) v_{m-1}$$

por lo tanto S' genera a W. Si S' es linealmente independiente S' es una base para W. Si S' es l.d. uno de los vectores en S' es una combinación lineal de los demás, y se procede a efectuar nuevamente el mismo procedimiento. En ésta forma debe llegarse al fin a un conjunto linealmente independiente que genera W, y que por lo tanto sea una base de W.

Se llega asi al siguiente resultado fundamental :

Cualquier espacio vectorial finito, generado por un conjunto de vectores no nulos, tiene por lo menos una base.

Ejemplo 5.27

Considere el subespacio W de R^3 generado por el conjunto de vectores S = {(2, 1, 3), (-1, 2, 0), (1, 8, 6)} encuentre una base de W

Solución:

Comenzamos por examinar el segundo vector, como (-1, 2, 0) no es múltiplo de (2, 1, 3) entonces (-1, 2, 0) se queda en la lista. Procedemos con el vector (1, 8, 6) y verificamos si puede expresarse en la forma

$$(1, 8, 6) = r_1(2, 1, 3) + r_2(-1, 2, 0)$$

para escalares r_1, r_2, este sistema de ecuaciones se puede escribir en forma matricial como

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right]$$

para encontrar la solución en r_1 y r_2 procedemos a efectuar eliminación gaussiana:

$$\left[\begin{array}{cc|c} 2 & -1 & 1 \\ 1 & 2 & 8 \\ 3 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 2 & -1 & 1 \\ 3 & 0 & 6 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & -5 & -15 \\ 0 & -6 & -18 \end{array}\right] \quad \left[\begin{array}{cc|c} 1 & 2 & 8 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{array}\right]$$

Obtenemos la solución $r_2 = 3$ y $r_1 = 2$ y concluimos que (1, 8, 6) está en el conjunto S' = { (2, 1, 3), (-1, 2, 0)} donde S' es una base de W.

Tal como se ha presentado hasta ahora, la técnica de exclusión parece potencialmente laboriosa. Por ejemplo, para aplicarla a una lista $v_1, v_2, \ldots, v_6$ parece que debemos probar si tienen soluciones para cinco ecuaciones vectoriales

$$v_2 = r_1 v_1, \qquad v_3 = r_1 v_1 + r_2 v_2$$

y asi sucesivamente.

En el espacio vectorial R^n, cada una de éstas ecuaciones conducirá a un sistema lineal, según se ilustró en el ejemplo anterior. En realidad se pueden resolver cinco sistemas lineales al mismo tiempo, reduciendo una sola matriz, según se enuncia en el recuadro. Explicaremos la validez de esta técnica en el siguiente ejemplo.

Técnicas de exclusión en R^n

La eliminación de los vectores distintos de cero que son combinación lineal de los predecesores en una lista $v_1, v_2, \ldots, v_k$ se puede lograr como sigue:

1. Formar la matriz A con v_j como el j-ésimo vector colum a.

2. Reducir A a la forma escalonada por filas como en el método de Gauss.

3. Los vectores de las columnas de A que dan lugar a pivotes (distintos de cero) deben permanecer y generarán todo el espacio columna de A. Esto es, se excluyen de la lista aquellos vectores de ls columnas de A que no proporcionen pivotes distintos de cero

Ejemplo 5.28

Sea W el subespacio de R^5 generado por:

$v_1 = (1, -1, 0, 2, 1)$ $\quad$ $v_2 = (2, 1, -2, 0, 0)$
$v_3 = (0, -3, 2, 4, 2)$ $\quad$ $v_4 = (3, 3, -4, -2, -1)$
$v_5 = (2, 4, 1, 0, 1)$ $\quad$ $v_6 = (5, 7, -3, -2, 0)$

use la técnica de exclusión para intenter acortar la lista v_i $i = 1, \ldots, 6$ de generadores de W.

Solución:

Se reduce la matriz que tiene v_j como j-ésimo vector columna y se tiene una forma escalonada.

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ -1 & 1 & -3 & 3 & 4 & 7 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 2 & 0 & 4 & -2 & 0 & -2 \\ 1 & 0 & 2 & -1 & 1 & 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 0 & 3 & 2 & 5 \\ 0 & 3 & -3 & 6 & 6 & 12 \\ 0 & -2 & 2 & -4 & 1 & -3 \\ 0 & -4 & 4 & -8 & -4 & -12 \\ 0 & -2 & 2 & -4 & -1 & -5 \end{bmatrix}$$

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

De la última matriz vemos que el subespacio W está generado por las columnas que contienen pivotes distintos de cero y corresponden a v_1, v_2 y v_5 en la matriz original. Se concluye que:

$$W = \text{gen}\ \{v_1, v_2, v_5\}$$

<u>DEFINICION 5.10</u> Un espacio vectorial V tiene dimensión finita si hay un subconjunto finito de vectores de V, $\{v_1, v_2, \ldots, v_n\}$ que forman una base de V (la dimensión es igual al número de vectores en una base de V). Si no existe tal subconjunto, V tiene <u>dimensión infinita</u> (además se convendrá en que el espacio trivial tiene dimensión finita, aún cuando no tiene conjuntos l.i. y por consecuencia, no tiene ninguna base)

La dimensión de un espacio vectorial V se denota por **dim V**

<u>TEOREMA 5.9</u>

1. Si $S = \{v_1, \ldots, v_m\}$ es un conjunto de n vectores linealmente independientes en un espacio V de n dimensiones, entonces S es una base de V.

2. Si $S = \{v_1, \ldots, v_m\}$ es un conjunto de n vectores que generan un espacio V de n dimensiones, entonces S es una base de V.

3. Si $S = \{v_1, \ldots, v_m\}$ es un conjunto linealmente independiente en un espacio V de n dimensiones y $r < n$, entonces a S se le pueden agregar vectores hasta formar una base de V, es decir, existen vectores $v_{r+1}, \ldots, v_n$ tales que $\{v_1, v_2, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ es una base de V.

Este teorema dice que si se conoce la dimensión n de un espacio vectorial V de dimensión finita, entonces los incisos 1 y 2 simplifican el problema de determinar si un conjunto S de n vectores es una base. Para mostrar que S es una base, basta con mostrar que S genera a V, o que S es linealmente independiente. Es claro que si no se conoce la dimensión de V se deben probar ambas cosas. La utilidad del inciso 3 se verá mas adelante.

TEOREMA 5.10

Si $S = \{v_1, \ldots, v_n\}$ es una base de V, entonces

a) Cualquier conjunto de m vectores con $m > n$ es linealmente dependiente, y por lo tanto no es una base para V.

b) Cualquier conjunto de m vectores con $m < n$ no puede generar a V y por lo tanto, no es una base de V.

Este teorema simplemente dice que el número de vectores en una base es único.

demostración:

a) Sea $T = \{w_1, w_2, \ldots, w_n, w_{n+1}\}$. Es decir, T contiene exactamente n+1 vectores. Se demostrará que T no puede ser una base probando que T es linealmente dependiente. Para ello considérse

$$c_1 w_1 + c_2 w_2 + \ldots + c_n w_n + c_{n+1} w_{n+1} = 0 \qquad (5.6)$$

pero cada w_k puede escribirse como:

$$w_k = a_{1k} v_1 + a_{2k} v_2 + \ldots + a_{nk} v_n \qquad k = 1, 2, \ldots, n+1$$

ya que S genera a V. Sustituyendo esto en la ecuación (5.6) se tiene:

$$\sum_{k=1}^{n+1} c_k w_k = \sum_{k=1}^{n+1} c_k \sum_{j=1}^{n} a_{jk} v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n+1} a_{jk} c_k \right) v_j = 0$$

desarrollando se tiene:

$$\begin{aligned} a_{11} c_1 v_1 + a_{12} c_2 v_2 + \ldots + a_{1,n+1} c_{n+1} v_n &= 0 \\ a_{21} c_1 v_1 + a_{22} c_2 v_2 + \ldots + a_{2,n+1} c_{n+1} v_n &= 0 \\ a_{n1} c_1 v_1 + a_{n2} c_2 v_2 + \ldots + a_{n,n+1} c_{n+1} v_n &= 0 \end{aligned}$$

Observe que éste sistema homogéneo tiene menos ecuaciones que incógnitas, por lo que existe entonces solución no trivial para $c_1, c_2, \ldots, c_{n+1}$. Esto quiere decir que T es l.d. Ahora suponga que T contiene mas de n+1 vectores. Sea T' un subconjunto de n+1 vectores, tal como se acaba de demostrar, debe ser l.d. Como $T' \subset T$, el conjunto T contiene un subconjunto l.d. y es entonces un conjunto l.d.

b) Suponga que T contiene n-1 vectores y que genera a V. Entonces por el teorema 5.8 T debe contener una base S^* para V. Si S^* contiene r vectores, entonces deberá ser $r \le n-1$. Como S^* es una base y S tiene r+1 o mas vectores, se infiere de la parte a) que S es l.d. Lo que contradice que S sea una base.

TEOREMA 5.11

Cualesquiera dos bases de un espacio vectorial de dimensión finita tienen el mismo número de vectores

demostración:

Sean $S = \{v_1,\dots,v_n\}$ y $S' = \{w_1, w_2,\dots,w_m\}$ dos bases de un espacio vectorial V de dimensión finita. Dado que S es una base y S' es un conjunto l.i. por el teorema 5.9 se afirma que $m \le n$. De igual manera puesto que S' es una base y S es un conjunto l.i. se tiene $n \le m$ lo que implica que $m = n$.

Ejemplo 5.28 (base canónica)

Muestre que R^n tiene la base canónica $E = \{e_1, e_2,\dots,e_n\}$ donde $e_j = (0, 0,\dots, \underset{\text{j-ésima componente}}{1}, 0,\dots,0)$

Solución:
Para el espacio generado considere cualquier $x = (x_1, x_2,\dots, x_n) \in R^n$ y observe que:

$$\begin{aligned} x &= x_1 (1, 0,\dots,0) + x_2 (0, 1,\dots,0) + \dots + x_n (0,\dots,0, 1) \\ &= x_1 e_1 + x_2 e_2 + \dots + x_n e_n \end{aligned}$$

para el caso de la independencia lineal considere

$$c_1 e_1 + c_2 e_2 + \dots + c_n e_n = (0, 0,\dots,0)$$

de donde se ve que $c_1 = c_2 = \dots = 0$. Por lo tanto, E es una base de R^n y $\dim R^n = n$

esto es razonable ya que se tiene:

R - recta	objeto	unidimensional
R^2 - plano	"	bidimensional
R^3 - espacio	"	tridimensional

Ejemplo 5.29

Demuestre que M_{23} tiene dimensión 6

Solución:

Una base que es la canónica de M_{23} es la siguiente

$$
S = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right.
$$
$$
\left. \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right\}
$$

por lo que dim M_{23} = número de vectores en $S = 6$

Ahora que ya sabemos que es una base de un espacio vectorial, podemos plantear el segundo problema fundamental del álgebra lineal.

EL PROBLEMA DE LA BASE

Sea V un espacio vectorial. El problema de la base puede presentarse en una de las dos formas siguientes:

Problema 1. Construir una base para V. Tomando los vectores de V.

Problema 2. Dado un conjunto S de vectores en V, construir una base para V añadiendo o bien eliminando, algunos (pero no todos) los vectores de S, o ambas cosas.

se puede resolver esto planteando lo siguiente:

Problema 1. Si puede escogerse un conjunto de vectores de V que genere V entonces eliminando vectores dependientes se obtendrá una base para V.

Problema 2. Si el conjunto dado S genera V, se procede como en el problema 1. Si no, se aumenta S añadiendo mas vectores hasta que se obtenga un conjunto generador y l.i.

Ejemplo 5.30 (2o.Problema de la base)

Sea S = {(1, 0, 3), (2, 1, 4), encuentre una base T para R^3 que contenga a S.

Como R^3 tiene dimensión 3, entonces T debe contener 3 vectores. El conjunto S, como está, ya es l.i. por lo que sólo hace falta añadir un vector a S, de tal forma que este nuevo vector sea l.i. con los dos ya existentes.

Solución 1: (prueba y error)

Como $\{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$ es una base para R^3 entonces verificamos la independencia lineal de cada uno de éstos vectores con los de S.

Esto es, si existen r_1, r_2 y r_3 tales que

$$r_1 (1, 0, 3) + r_2 (2, 1, 4) + r_3 (1, 0, 0) = (0, 0, 0)$$

lo que implica que

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 3 & 4 & 0 & 0 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array}\right]$$

asi los vectores son linealmente independientes y forman una base $T = \{(1, 0, 3), (2, 1, 4), (1, 0, 0)\}$

Solución 2:

Otra forma de resolverlo es la siguiente:

$$\begin{aligned} \text{gen}\,\{(1, 0, 3), (2, 1, 4)\} &= \{ x \mid x = a (1, 0, 3) + b (2, 1, 4)\} \\ &= \{ x \mid x = (a + 2b, b, 3a + 4b)\} \end{aligned}$$

por lo tanto, hay que procurar que el nuevo vector no sea de la forma $(a + 2b, b, 3a + 4b)$. Para ello se supone que el nuevo vector es (x_1, x_2, x_3) y se hace que la ecuación

$$(a + 2b, b, 3a + 4b) = (x_1, x_2, x_3)$$

no tenga solución para a y b. Igualando las componentes se obtienen las ecuaciones cuya representación matricial es

$$\left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 3 & 4 & x_3 \end{array}\right]$$

lo que se reduce a

$$\left[\begin{array}{cc|c} 1 & 2 & x_1 \\ 0 & 1 & x_2 \\ 0 & 0 & x_3 - 3x_1 + 2x_2 \end{array}\right]$$

por lo tanto si x_1, x_2 y x_3 se eligen de tal manera que $x_3 - 3x_1 + 2x_2 \neq 0$ se tendrá un vector que no está en gen $\{(1, 0, 3), (2, 1, 4)\}$, en tal caso $x = (0, 1, 0)$ es una elección aceptable y

$$T = \{ (1, 0, 3), (2, 1, 4), (0, 1, 0)\}$$

es una base de R^3.

Solución 3:

Si el tercer vector (x_1, x_2, x_3) hiciera que $\{(1, 0, 3), (2, 1, 4), (x_1, x_2, x_3)\}$ fuera linealmente dependiente entonces

$$\det \begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 4 \\ x_1 & x_2 & x_3 \end{bmatrix} \neq 0$$

calculando el determinante se tiene

$$x_3 + 2x_2 - 3x_1 \neq 0$$

que es la misma condición obtenida mediante la solución 2. Algunos terceros vectores posibles son (1, 0, 0), (0, 1, 0), (0, 1, 1) o (1, 1, 0) en realidad hay un número infinito de posibilidades.

5.7 RESUMEN

1. Un espacio vectorial es un conjunto no vacío V de objetos llamados vectores junto con reglas para sumar dos vectores cualesquiera v y w de V y para multiplicar culaquier vector v de V y cualquier ecalar r de R. Se requiere que V sea cerrado bajo esta suma de vectores y multiplicación por un escalar, de modo que v+w y rv estén en V. Es más, los axiomas siguientes se deben satisfacer para todos los vectores u, v y w de V y todos los escalares r y s de R.

2. Para todos los vectores v y w de R^n, tenemos

Desigualdad de Schwarz: $|v \cdot w| \leq \|v\| \, \|w\|$

Desigualdad triangular : $\|v + w\| \leq \|v\| + \|w\|$

3. Un subconjunto S de un espacio vectorial V es un subespacio de V si, y solo si, es no vacío y satisface las dos propiedades de cierre:

$v + w \in S$ para todos los vectores v y $w \in S$

$rv \in S$ para todos los vectores $v \in S$ y escalares $r \in R$

4. El espacio generado por $\{ v_1, v_2, \ldots, v_n \}$ es el conjunto de todas las combinaciones posibles de $v_1, v_2, \ldots, v_n$ y se denota

$$\text{gen}(v_1, v_2, \ldots, v_n) = \{r_1 v_1 + r_2 v_2 + \ldots + r_n v_n | \; r_i \in R \; i=1,2,\ldots,n \}$$

5. Sea V un espacio vectorial, un conjunto de vectores $\{v_1, v_2,\ldots,v_k\}$ de V es linealmente dependiente si existe una relación de dependencia

$$r_1 v_1 + r_2 v_2 +\ldots+ r_n v_n = 0, \text{ para algún } r_i \neq 0$$

el conjunto es linealmente independiente si no existe dicha relación de dependencia, de modo que una combinación lineal de los v_i es el vector cero sólo si todos los coeficientes son cero.

6. Una lista finita de vectores distintos de cero en V forma un conjunto linealmente independiente si, y solo si, ningún vector de la lista se puede expresar como combinación lineal de sus predecesores.

7. Un conjunto de vectores $\{v_1, v_2,\ldots,v_k\}$ es una base para un subespacio S de un espacio vectorial V si es linealmente independiente y constituye un conjunto generador para S.

8. Cualquier subconjunto generador finito de un subespacio S, distinto de cero, de un espacio vectorial V se puede reducir, de ser necesario, a una base para S quitando los vectores cero y excluyendo después aquellos que son combinaciones lineales de predecesores.

9. Las siguientes afirmaciones son equivalentes para n vectores de R^n

a) Los vectores forman una base para R^n
b) Los vectores son linealmente independientes
c) Los vectores generan R^n
d) Una matriz que tenga a los vectores como columnas es invertible.

10. Sea S un subespacio de dimensión finita de un espacio vectorial V. Todas las bases de S tienen el mismo número de elementos. Este número, dim(S), se llama dimensión de S.

5.8 NOTAS HISTORICAS

LA IDEA DE UN ESPACIO N-DIMENSIONAL PARA N>3 se aceptó gradualmente durante el siglo XIX; por tanto, es difícil señalar una primera "invención" de este concepto. Entre los diversos usos tempranos de este concepto está su presencia en un trabajo del matemático ruso Mikhail Ostrogradskii (1801-1862) sobre el teorema de la divergencia, escrito en 1836; en los tratados geométricos de Hermann Grassmann (1809-1877) al principio de la década de 1840, y en un breve artículo de Arthur Cayley, en 1846. Lamentablemente, los dos primeros autores fueron virtualmente ignorados en vida. En particular,el trabajo de Grassman era bastante filosófico y muy difícil de leer. El artículo de Cayley simplemente decía que se podían generalizar ciertos resultados a dimensiones mayores que tres "sin recurrir a ninguna noción metafísica respecto a la posibilidad de un espacio de cuatro dimensiones". Sir William Rowan Hamilton (1805-1865), en una carta de 1841, también observó que "debe de ser posible, de una u otra manera introducir no sólo ternas sino poliadas (polyplets), de modo que en algún sentido se satisfaga la ecuación simbólica $a = (a_1, a_2, \ldots, a_n)$ donde a es un símbolo indicativo de un pensamiento (complejo) y $a_1, a_2, \ldots, a_n$ denotan n números reales, positivos o negativos".

AUNQUE LOS VECTORES DE R^n fueron tratados por los matemáticos y físicos durante la segunda mitad del siglo XIX (como sucedió con otros objetos que hoy consideramos vectores), hasta la aparición del tratado de Hermann Weyl (1885-1955) Space-Time Matter en 1918 no se imprimió una definición axiomática y abstracta de un espacio vectorial. Weyl escribió este libro como introducción a la teoría general de la relatividad de Einstein; en el capítulo 1 analizó la naturaleza del espacio euclideano y, como parte de ese análisis, formuló lo que ahora son los axiomas comunes de un espacio vectorial. Como trató sólo los espacios V de dimensión finita, incluyó el axioma de dimensionalidad: para algún número entero h hay h vectores linealmente independientes en V, pero todo conjunto de h+1 vectores es linealmente dependiente.
Los axiomas de espacio vectorial se conocían desde hacía años, pero en general se demostraban como consecuencia de otras definiciones de vectores. Por ejemplo, aparecen en el breve texto en italiano de Giuseppe Peano (1858-1932). Cálculo geométrico (1888), donde se explica el trabajo de Hermann Grassmann.

LA DESIGUALDAD DE SCHWARZ se debe independientemente a Augustin-Louis Cauchy (1789-1857), Hermann Amandus Schwarz (1843-1921) y Viktor Yakovlevich Bunyakovsky (1804-1889).

Primero se formuló como un teorema sobre coordenadas en un apéndice al texto de Cauchy de 1821 para su curso de análisis en la Ecole Polytechnique como sigue:

$$|a\alpha + a'\alpha' + a''\alpha'' + \ldots| \le \sqrt{a^2 + a'^2 + a''^2 + \ldots}\ \sqrt{\alpha^2 + \alpha'^2 + \alpha''^2 + \ldots}$$

Bunyakovsky demostró en 1859 la desigualdad para funciones, esto es, formuló el resultado

$$\left[\int_a^b f(x)\ g(x)\ dx\right]^2 \le \int_a^b f^2\ (x)\ dx \int_a^b g^2\ (x)\ dx$$

donde la integral del producto de las funciones f y g se puede considerar como el producto interior de las funciones continuas en [a,b]. Bunyakovsky fue vicepresidente de la Academia de Ciencias de San Petesburgo desde 1864 hasta su muerte. En 1875 la Academia estableció un premio matemático en su nombre, como reconocimiento a sus 50 años de enseñanza e investigación.

Schwarz formuló la desigualdad en 1884. En su caso, los vectores eran funciones Φ,Ω, de dos variables en una región T del plano, y el producto interior de éstas funciones estaba dado por la doble integral de su producto, donde se supone que ésta integral existe. Entonces la desigualdad establece que

$$\left|\iint_T \Phi\Omega \, dx \, dy\right| \leq \sqrt{\iint_T \Phi^2 \, dx \, dy} * \sqrt{\iint_T \Omega^2 \, dx \, dy}$$

Schwarz era el principal matemático de Berlín alrededor del cambio de siglo, el trabajo donde aparece la desigualdad está dedicado a una cuestión sobre superficies minimales.

CAPITULO 6

LOS CUATRO ESPACIOS FUNDAMENTALES

6.1 ESPACIOS ASOCIADOS A UNA MATRIZ.

Cuando tenemos un sistema de n ecuaciones con n incógnitas podemos resolverlo escribiendolo en forma matricial y la matriz de coeficientes correspondiente se lleva a una matriz triangular superior a través de eliminación gaussiana y se procede a resolver el sistema por sustitución regresiva.

Para el caso de sistemas con m ecuaciones y n incógnitas o viceversa se procede a efectuar eliminación gaussiana, pero la matriz resultante no es una matriz triangular superior estrictamente hablando, lo que se tiene es una matriz en forma escalonada

La forma final U es "triangular superior", donde los pivotes no están en la diagonal principal y las entradas distintas de cero quedan en forma escalonada.

$$U = \begin{bmatrix} * & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Antes de estudiar todas las posibilidades de encontrar soluciones a este tipo de sistemas conviene que estudiemos ciertos espacios asociados a las matrices, los resultados proporcionarán un procedimiento sencillo para elaborar bases reduciendo una matríz determinada a la forma escalonada, y en consecuencia tendremos la posibilidad de analizar soluciones a tales sistemas.

<u>DEFINICION 6.1</u> Considere la matríz

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

los vectores

$$\begin{array}{l} r_1 = (a_{11} \quad a_{12} \ \cdots \ a_{1n}) \\ r_2 = (a_{21} \quad a_{22} \ \cdots \ a_{2n}) \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ r_m = (a_{m1} \quad a_{m2} \ \cdots \ a_{mn}) \end{array}$$

que se forman con los renglones de A, se denominan los <u>vectores renglón de A</u>. El subespacio de R^n generado por los vectores renglón se llama <u>espacio de los renglones de A</u>.

<u>Ejemplo 6.1</u>

Los vectores renglón de la matríz

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

son $r_1 = (2,\ 1,\ 0)$ y $r_2 = (3,\ -1,\ 4)$

El siguiente teorema es de gran utilidad para los cálculos que se efectuarán posteriormente

<u>TEOREMA 6.1</u>

Las operaciones elementales en los renglones no alteran el espacio de los renglones de una matriz.

Supongamos que se quiere elaborar una base para el subespacio de R^n generado por determinado conjunto de vectores $\{v_1,\ v_2, \ldots, v_m\}$, entonces si se forma la matríz B cuyos vectores renglón son $v_1,\ v_2, \ldots, v_m$, el problema se reduce a encontrar una base para el espacio de los renglones de B, por ejemplo, se puede obtener una base para el espacio generado por los vectores:

$$\begin{array}{l} v_1 = (1,\ -2,\ 0,\ 0,\ 3) \\ v_2 = (2,\ -5,\ -3,\ -2,\ 6) \\ v_3 = (0,\ 5,\ 15,\ 10,\ 0) \\ v_4 = (2,\ 6,\ 18,\ 8,\ 6) \end{array}$$

encontrando una base para el espacio de los renglones de la matríz

$$B = \begin{bmatrix} 1 & -2 & 0 & 0 & 3 \\ 2 & -5 & -3 & -2 & 6 \\ 0 & 5 & 15 & 10 & 0 \\ 2 & 6 & 18 & 8 & 6 \end{bmatrix}$$

Por el teorema 6.1, se deduce que el espacio de los renglones de una matríz no se altera si la matríz se lleva a la forma escalonada

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

U es la forma escalonada de B. Dado que el espacio de los renglones de B coincide con los de U, una base para el espacio de los renglones de U será también una base para el espacio de los renglones de B. Es fácil verificar que los vectores renglón de U, diferentes de cero

$$r_1 = (1, -2, 0, 0, 3)$$
$$r_2 = (0, 1, 3, 2, 0)$$
$$r_3 = (0, 0, 1, 1, 0)$$

son linealmente independientes, por lo tanto éstos vectores forman una base para el espacio de los renglones de U y en consecuencia para el espacio de los renglones de B lo que nos conduce al siguiente teorema.

TEOREMA 6.2

Los vectores renglón diferentes de cero, en la forma escalonada de una matríz A, constituyen una base para el espacio de los renglones de A.

Nota 1: Las columnas que contienen los pivotes son linealmente independientes.

Nota 2: La dimensión del espacio de los renglones es igual al rango de la matríz.

DEFINICION 6.2 El espacio columna de una matríz de orden mxn es el espacio generado por las columnas. Es un subespacio del espacio m-dimensional R^m. Si m=n entonces tanto el espacio renglón como el espacio columna son subespacios de R^n, incluso pueden ser el mismo subespacio. Al espacio columna a menudo se le llama el recorrido de A : R(A), Como la idea de recorrido de una función (FIGURA 6.0)

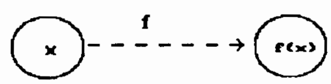


Figura 6.0

en éste caso la función es f(x) = Ax. Cuyo dominio consta de todas las x en R^n y el recorrido todos los posibles vectores Ax. Esto es, todas las b para las cuales puede resolverse $Ax = b$, esto es igual a todas las posibles combinaciones de las columnas: el recorrido es el espacio columna.

<u>Ejemplo 6.2</u>

Sea la matríz

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

las columnas de esta matríz son linealmente dependientes, ya que la segunda columna es tres veces la primera. Los renglones son también linealmente dependientes, el tercero es dos veces el segundo menos cinco veces el primero.

Si escalonamos la matríz A, los renglones distintos de cero deben ser linealmente independientes.

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

Además si elegimos las columnas que contienen los pivotes, también son linealmente independientes (en éste caso las columnas 1 y 3 son linealmente independientes).

Las cuatro columnas generan el espacio columna por definición sin embargo no son l.i., por lo tanto proponemos que las columnas que contengan los pivotes distintos de cero son una base del espacio columna.

Los dos primeros renglones (por el teorema 6.2) son una base para el espacio renglón. De todo ésto podemos concluir que los r renglones distintos de cero de una matríz escalonada U son l.i. y también lo son las r columnas que contienen los pivotes distintos de cero.

La forma escalonada entonces se puede describir de la siguiente manera:

1. Primero vienen los renglones distintos de cero (de haber alguno se habría intercambiado)

2. Debajo de cada pivote hay una columna de ceros, que se obtiene por eliminación.

3. Cada pivote está a la derecha del pivote del renglón anterior, esto produce la figura escalonada.

Nos interesa localizar las soluciones, si existen de $Ax = b$ donde A es una matriz de mxn, $x \epsilon R^n$ y $b \epsilon R^m$, las columnas de A las podemos ver como vectores de R^m, esto es

$$A = (a_1,\ a_2, \ldots, a_n) \text{ con } a_i \epsilon R^m$$

donde Ax podemos verla como una combinación lineal de los vectores columna de A, esto es

$$Ax = a_1\ x_1 + a_2\ x_2 + \ldots + a_n\ x_n$$

de donde $Ax = b$ tiene solución si el vector $b \epsilon R^m$ puede expresarse como combinación lineal de los vectores columna de A de donde

1. Si $m > n$ entonces los vectores columna son vectores de R^m y el número de vectores columna de A es menor que m de donde el espacio generado por las columnas de A es un subespacio de R^m (de a lo mas dimensión n)

por ejemplo

$$\begin{aligned} 3x_1 + 2x_2 &= 3 \\ 5x_1 + 6x_2 &= 1 \\ 3x_1 - 4x_2 &= 5 \end{aligned}$$

tenemos entonces que (3, 5, 3) y (2, 6, 4) son vectores en R^3 donde no pueden generar a R^3 pero si un subespacio de R^3, si son linealmente independientes un plano en R^3 y si no una recta.

Tenemos que considerar además que:

a) el sistema no tenga solución, lo que sucede si b no se encuentra en el subespacio generado por las columnas de A

b) que tenga solución única, si los vectores columna de A son l.i. y forman por lo tanto una base del espacio generado por las columnas de A.

c) que tenga una infinidad de soluciones, ésto ocurre si la dimensión del espacio generado por las columnas es menor que n, por lo cual los vectores columna son linealmente dependientes y conforman un conjunto generador (pero no base, ya que generan un subespacio de dimensión menor que n) de donde si b está en el espacio generado por las columnas de A puede representarse en una infinidad de formas de combinaciones lineales de las columnas de A.

2. Si m<n los vectores columna de A no pueden ser una base para R^m ya que son mas de los que se necesitan y para R^n solo pueden ser base para un subespacio.

Ejemplo 6.3

Comenzamos con el caso homogéneo Ax = 0 donde A es la matriz

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ 1 & -3 & 3 & 0 \end{bmatrix}$$

que se puede reducir a Ux = 0 de la siguiente manera:

$$Ux = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

las incógnitas x, y, z, w van en dos grupos. Un grupo consta de las variables básicas, aquellas que corresponden a las columnas con pivotes distintos de cero (en este caso la 1a. y la 3a.) asi que x, z son variables básicas. El otro grupo consta de las variables libres, que corresponden a las columnas sin pivotes, estas son la 2a y la 4a. columnas, entonces y, w son las variables libres.

Para encontrar la solución general a Ux = 0 (o Ax = 0) asignamos valores arbitrarios a las variables libres esto es, y, w entonces

$$\begin{aligned} 3z + w = 0 \quad &\text{-->} \quad z = -1/3\ w \\ x + 3y + 3z + 2w = 0 \quad &\text{-->} \quad x = -3y - w \end{aligned}$$

hay una infinidad de soluciones al sistema con dos parámetros libres w, y. La solución general es la combinación

$$x = \begin{bmatrix} -3y - w \\ y \\ -1/3\ w \\ w \end{bmatrix} = y \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

si damos valores a y , y a w, y = 1, w = 0 tenemos la solución (-3, 1, 0, 0), si y = 0, w = 1 tenemos (-1, 0, -1/3, 1). Todas las soluciones son combinación lineal de éstas dos.

Entonces las soluciones de $Ax = 0$ forman un subespacio bidimensional, el espacio nulo de A, del espacio de 4 dimensiones de todos los vectores x. Este espacio nulo lo podemos describir como un "plano" generado por (-3, 1, 0, 0) y (-1, 0, -1/3, 1). Las combinaciones de estos vectores forman un conjunto cerrado bajo la suma y el producto por un escalar y todas estas combinaciones constituyen el espacio nulo.

En el caso de $n > m$, no debe haber mas de m pivotes distintos de cero, entonces debe haber al menos n-m variables libres por lo cual: Cada sistema homogéneo $Ax = 0$ tiene una solución no trivial si tiene mas incógnitas que ecuaciones ($n > m$): existe alguna solución x diferente de la trivial $x = 0$.

Para el caso no homogéneo $b \neq 0$ tenemos $Ax = b$ aplicando operaciones elementales tenemos $Ux = c$:

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{bmatrix}$$

No está claro que el sistema tenga solución ya que el tercer renglón de la matriz es igual a cero entonces

$$b_3 - 2b_2 + 5b_1 = 0$$

el conjunto de vectores obtenibles b no es todo el espacio tridimensional. Es posible resolver $Ax = b$ si b está en el espacio columna de A generado por las columnas:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 6 \\ -3 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 2 \\ 5 \\ 0 \end{bmatrix}$$

aunque son 4 vectores, solo obtenemos un plano en el espacio tridimensional (ya que la 2a columna es 3 veces la primera y la cuarta es la primera mas una fraccion de la tercera). Son estas columnas las que no tienen pivotes.

Si suponemos que b está en el plano generado por el espacio columna por ejemplo $b = (1, 5, 5)$ efectuando eliminación y fijando las variables libres obtenemos

$$\begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

$$\begin{array}{lrcl} \to & 3z + w = 3 & \to & z = 1 - 1/3\, w \\ & x + 3y + 3z + 2w = 1 & \to & x = -2 - 3y - w \end{array}$$

de donde

$$x = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + y \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + w \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

si comparamos la solución con la del caso homogéneo Ax = 0 la única diferencia es la inclusión del vector (-2, 0, 1, 0) que es una solución de las ecuaciones dadas, es una solución particular de Ax = b, y la solución general x es una suma de esta solución particular y la solución general de Ax = 0. Geométricamente las soluciones se encuentran de nuevo en un plano en el espacio de 4 dimensiones, pero ahora no constituyen un subespacio, ya que el plano no pasa por el origen, el origen x = 0 no es una solución cuando $b \neq 0$. El plano es paralelo al espacio nulo que se tenía antes, pero desplazado por el vector que da la solución particular.

Resumiendo tenemos:

El própósito original de la eliminación era simplificar un sistema de ecuaciones lineales sin alterar ninguna de las soluciones. Se reduce el sistema Ax = 0 a Ux = 0 y éste proceso es reversible. Por lo tanto el espacio nulo de A es el mismo que el espacio nulo de U. De las m aparentes restricciones impuestas por la m ecuaciones Ax = 0, sólo r son independientes, éstan especificadas por cualesquiera r renglones de U distintos de cero.

DEFINICION 6.3 El espacio nulo de A: N(A) tiene dimensión n-r y consta de todos aquellos vectores x para los cuales el sistema homogéneo Ax = 0 (o su equivalente Ux = 0) tiene solución. Se obtiene una base para el espacio nulo reduciendo el sistema original al sistema Ux = 0, que tiene n-r variables libres, correspondientes a las columnas de U que no contienen pivotes. Al espacio nulo también se le llama el núcleo de A y su dimensión n-r es la nulidad de A.

De todo lo expuesto hasta aqui podemos concluir que A no tiene el mismo espacio columna que U. La eliminación no altera el espacio de los renglones ni el espacio nulo, pero las columnas son completamente diferentes. Sin embargo si el conjunto de columnas de A es independiente, entonces lo mismo es cierto para las correspondientes columnas de U y viceversa, entonces U nos puede ser útil en términos de indicarnos cuales son las columnas de A que forman una base.

Para encontrar una base de R(A) consideremos el siguiente resultado:

TEOREMA 6.3

La dimensión del espacio columna R(A) es igual al rango r, que también es igual a la dimensión del espacio fila: el número de columnas independientes es igual al número de renglones independientes. Una base de R(A) está formada por aquellas r columnas de A correspondientes en U, a las columnas que contienen los pivotes distintos de cero.

Este teorema también afirma para el caso de matrices cuadradas que si los renglones de una matríz cuadrada son l.i. también lo son las columnas.

DEFINICION 6.4 El espacio nulo de A^t o espacio nulo izquierdo : $N(A^t)$ es un subespacio de R^m que consta de todos aquellos vectores y tales que $A^t y = 0$ ó $y^t A = 0$ y se le llama a y el vector nulo izquierdo de A.

Es fácil encontrar la dimensión de $N(A^t)$. Sabiendo que:

rango + nulidad = dimensión del espacio columna + dimensión del espacio nulo = número de columnas

Esta regla también se aplica para A^t la cual tiene m columnas. Pero el rango de los renglones = rango de las columnas = r, entonces

$$r + \dim N(A^t) = m$$
$$\dim N(A^t) = m - r$$

Esquemáticamente se puede ver en la figura 6.1

Sabemos ahora cual es la dimensión de los cuatro espacios asociados a una matríz y resumimos éste resultado en el siguiente teorema

TEOREMA 6.4: TEOREMA FUNDAMENTAL DEL ALGEBRA LINEAL (1a. Parte)

1. $R(A^t)$ = espacio renglón de A: dimensión = r
2. $N(A)$ = espacio nulo de A: dimensión = n-r
3. $R(A)$ = espacio columna de A: dimensión = r
4. $N(A^t)$ = espacio nulo izquierdo de A: dimensión = m-r

tal que:

$$n = \dim R(A) + \dim N(A)$$
$$m = \dim R(A^t) + \dim N(A^t)$$

Este teorema se ilustra con el siguiente ejemplo:

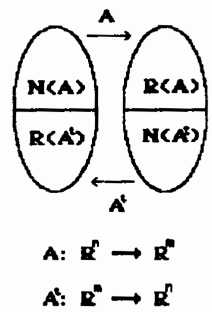


Figura 6.1

Ejemplo 6.3

Encuentre la dimensión y construya una base de los cuatros subespacios asociados a la matriz

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix}$$

Solución:

1.- El espacio de los renglones de A: $R(A^t)$ escalonando la matriz A tenemos:

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \approx \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

entonces el vector (0,1,4,0) forma una base de los renglones de A y Dim $[R(A^t)]=1$.

2.- El espacio de las columnas, buscamos las columnas cuyos pivotes sean distintos de cero y vemos que la unica es la segunda columna $(1,2)^t$ de donde :

Dim (R(A))= 1 y una de sus bases $(1,2)^t$

3.- Espacio nulo de A. N(A) con dimensión n-r => 4-1=3 y para determinar una base, procedemos a resolver el sistema Ax=0, (también puede hacerse con Ux=0)

$$Ax = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

de donde

$$x_2 + 4x_3 = 0 \quad => \quad x_2 = -4x_3$$

sea x_3 =t entonces x_2 =-4t y haciendo x_1 = u y x_4 = v se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} u \\ -4t \\ t \\ v \end{bmatrix} = t\begin{bmatrix} 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} + u\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

por lo cual

$$\begin{bmatrix} 0 \\ -4 \\ 1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

forman una base para N(A), y dim (N(A)) = n - r = 4 - 1 = 3 que son el número de vectores en la base.

4.- Espacio nulo izquierdo de A :N(A^t) con dimensión m-r => 2-1= 1. Para encontrar una base procedemos a resolver A^t y = 0

$$A^t\, y = \begin{bmatrix} 0 & 0 \\ 1 & 2 \\ 4 & 8 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde

$y_1 = -2y_2$ haciendo $y_2 = t$
$y_1 = -2t$

por lo cual una base de N(A^t) está dada por

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

y dim $(N(A^t)) = m - r = 2 - 1 = 1$

Ejemplo 6.4

Encuentre la dimensión y construya una base de los cuatro subespacios asociados a la matiz

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

Solución:

1.- El espacio de los renglones de A: R (A^t)

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix} \approx \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = U$$

entonces los vectores {(1, 3, 3, 2), (0, 0, 3, 1)} forman una base de los renglones de A y Dim $R(A^t) = 2$

2.- El espacio de las columnas, buscamos las columnas cuyos pivotes sean distintos de cero:

$$\begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \quad \begin{bmatrix} 3 \\ 9 \\ 3 \end{bmatrix}$$

Dim R(A)= 2

3.- Espacio nulo de A. N(A) con dimensión n-r => 4-2=2 y para determinar una base, procedemos a resolver el sistema Ax=0, (también puede hacerse con Ux=0)

$$Ax = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

de donde

$3x_3 + x_4 = 0 \quad => \quad x_3 = -1/3\ x_4$

sea $x_4 = t$ entonces $x_3 = -1/3t$ y haciendo $x_1 = -3x_2 - 3x_3 - 2x_4$ y $x_2 = u$ se obtiene:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -3u-t \\ u \\ -t/3 \\ t \end{bmatrix} = u \begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{bmatrix}$$

por lo cual

$$\begin{bmatrix} -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} -1 \\ 0 \\ -1/3 \\ 0 \end{bmatrix}$$

forman una base para N(A).

4.- Espacio nulo izquierdo de A :$N(A^t)$ con dimensión m-r => 3 - 2 = 1. Para encontrar una base procedemos a resolver $A^t\ y = 0$

$$A^t\ y = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 0 & 0 \\ 3 & 3 & 0 \\ 2 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_3 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

por lo cual una base de $N(A^t)$ está dada por

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = t \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Hemos visto hasta aqui que para encontrar una base del espacio renglón y una base para el espacio columna se procede a escalonar la matriz A en una matriz U, es entonces factible preguntar si éstas matrices están relacionadas como antes, mediante una matriz triangular inferior L, tal que A = LU, la respuesta es si existe una matriz L tal que A = LU, por ejemplo:

Si se tienen las siguientes matrices A y su forma escalonada U

$$A = \begin{bmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{bmatrix}$$

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

entonces la matriz L esta dada por:

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

y se verifica que A = LU, note además que L no es rectangular sino cuadrada. Es una matriz del mismo orden m = 3 que el número de renglones en A y en U. En este caso como se requirió de un intercambio de renglones entonces se introduce una matriz de permutación P con lo que enunciamos el siguiente teorema:

TEOREMA 6.5

A cada matriz A de orden mxn corresponde una matriz de permutación P, una matriz triangular inferior L con diagonal unitaria y una matriz escalonada U de orden mxn tales que PA = LU

MATRICES DE RANGO UNO

Un ejemplo de una matriz de rango uno, r = 1 está dada por la siguiente matriz:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix}$$

Cada renglón es un múltiplo del primero de modo que el espacio renglón es unidimensional. Se puede escribir toda la matriz de la siguiente manera:

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 2 \\ 8 & 4 & 4 \\ -2 & -1 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 4 \\ -1 \end{bmatrix} [2, 1, 1]$$

como el producto de un vector columna con un vector fila, además todas las columnas son múltiplos del mismo vector, el espacio columna tiene dimensión $r = 1$.

Cualquier matriz de rango uno se puede factorizar en la forma simple $A = uv^t$. Todos los renglones son múltiplos del mismo vector v^t y todas las columnas son múltiplos de un mismo vector u.

6.2 SOLUCION DE M ECUACIONES EN N INCOGNITAS

Ya conocemos el proceso de eliminación para matrices cuadradas. La eliminación misma se realiza directamente sin mayor cambio, pero hay algunas diferencias al obtener la solución mediante la sustitución regresiva.

Consideremos la ecuación escalar $ax = b$ (un sistema de una ecuación con una incógnita), podemos considerar las siguientes tres alternativas:

1. Si $a \neq 0$ entonces existe la solución $x = b/a$ para cualquier b y ésta solución es única. Este es el caso <u>no singular</u> , de una matríz de 1x1 invertible.

2. Si $a=0$ y $b=0$ hay una infinidad de soluciones, cualquier x satisface $0x = 0$. Este es el caso <u>indeterminado</u>, existe solución pero no es única.

3. Si $a=0$ y $b\neq0$ no hay solución a $0x = b$ éste es el caso <u>inconsistente</u>.

Estas alternativas pueden ocurrir para matrices cuadradas. Con una matríz rectangular no podemos tener existencia y también unicidad para cada b. Esto se puede ver a través de la existencia de la inversa.

Sabemos que una matriz A tiene inversa si existen $BA = I$ y $AC = I$ donde $B = C = A^{-1}$. Ahora, del rango de una matriz es fácil decir que matrices tienen realmente estas inversas, se puede decir en general que existe una inversa sólo cuando el rango es lo mas grande posible.

El rango siempre satisface $r \leq m$ y $r \leq n$ ya que una matriz de mxn no puede tener mas de m filas independientes o de n columnas independientes. Entonces si $r = m$ existe una inversa derecha y si $r = n$ existe una inversa izquierda

<u>DEFINICION 6.5</u> Sea A matriz de orden mxn. Una matriz G de nxm (si existe) tal que $GA = I_n$ se dice inversa izquierda de A. Asimismo una matriz H de orden nxm (si existe) tal que $AH = I_m$ se dice inversa derecha.

Ejemplo 6.5

Sea la matriz de 2x1 siguiente

$$A = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

calcule (si existe) la inversa izquierda y derecha de A

Solución:

Sea $G = [g_1, g_2]$ si $GA = 1$

$$[g_1, g_2] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = g_1 + g_2 = 1$$

si $g_1 = \alpha$ entonces $g_2 = 1-\alpha$ de donde $G = [\alpha, 1-\alpha]$ para $\alpha \epsilon R$

Sea $H = [h_1, h_2]$ tal que $AH = I$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} [h_1, h_2] = \begin{bmatrix} h_1 & h_2 \\ h_1 & h_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

pero $h_1 = 1$ y $h_1 = 0$ por lo que se afirma que no existe matriz derecha además el rango de $A = 1 < 2 = m$

TEOREMA 6.6

EXISTENCIA: El sistema $Ax = b$ tiene al menos una solución x para toda b si y sólo si las columnas de A generan R^m, entonces $r = m$. En éste caso existe una inversa derecha H de nxm tal que $AH = I_m$, la matriz identidad de orden m. Esto es posible sólo si $m \leq n$.

UNICIDAD: El sistema $Ax = b$ tiene a lo sumo una solución x para cada b si y sólo si las columnas son linealmente independientes, entonces $r=n$. En éste caso existe una inversa izquierda G de nxm tal que $GA = I_n$, la matriz identidad de orden n. Esto es posible sólo si $m \geq n$.

En el primer caso, una posible solución es $x = Hb$, ya que entonces $Ax = AHb = b$ pero habrá otras soluciones si hay otras inversas derechas.

En el segundo caso, si existe una solución a $Ax = b$, tiene que ser $x = GAx = Gb$, pero no puede haber otra solución.

Ejemplo 6.6

Considere la siguiente matriz A de 2x3 y de rango 2

$$A = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix}$$

verifique si existen inversa izquierda o derecha de A.

Solución:

Como r = m = 2 el teorema garantiza una inversa derecha H:

$$AH = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1/4 & 0 \\ 0 & 1/5 \\ c_{31} & c_{32} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

de hecho hay muchas inversas derechas, ya que el último renglón de H es totalmente arbitrario, entonces si escribimos un sistema Ax = b con esta matriz A se puede afirmar que existe solución.

Para ver si existe inversa izquierda solo hay que verificar que r = n, en este caso r = 2 < 3 por lo que no existe inversa izquierda.

TEOREMA 6.7

Una matriz A de nxn es invertible si y solo si rango (A) = n

Ejemplo 6.7

Determine si la matriz siguiente es invertible hallando su rango.

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

Solución:

Escalonando A se tiene:

$$\begin{bmatrix} 4 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 3 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 1 & -12 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & -14 \end{bmatrix}$$

por lo que rango A = 3 y A es invertible.

6.3 SUBESPACIOS ORTOGONALES.

Recordemos que la longitud de un vector $\|x\|$ en R^n está dada por:

$$\|x\| = \sqrt{x_1^2 + x_2^2 + x_3^2 + \ldots + x_n^2}$$

geométricamente ésto se obtiene aplicando la fórmula de Pitágoras n-1 veces, añadiendo en cada paso una dimensión mas.

Si tenemos dos vectores x, y ¿Cómo podemos saber si son perpendiculares?

En el caso de R^2 la respuesta se puede obtener usando trigonometría en este caso x es ortogonal a y si forman un triángulo rectángulo y usando Pitágoras se tiene:

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

para R^n se tiene

$$(x_1^2 + x_2^2 + \ldots + x_n^2) + (y_1^2 + y_2^2 + \ldots + y_n^2) = (x_1 - y_1)^2 + \ldots + (x_n - y_n)^2$$

desarrollando el lado derecho se tiene:

$$(x_1^2 + x_2^2 + \ldots + x_n^2) - 2\,(x_1y_1 + \ldots + x_ny_n) + (y_1^2 + y_2^2 + \ldots + y_n^2)$$

asi la igualdad se cumple si los "términos cruzados" son iguales a cero:

$$x_1\, y_1 + \ldots + x_n\, y_n = 0$$

esta cantidad no es otra que el producto interior de $x^t y$ en R^n por lo que se puede afirmar que es igual a cero sólo si x y y son ortogonales.

Existe una relación sencilla entre independencia y ortogonalidad: si los vectores son mutuamente ortogonales entonces son linealmente independientes.

En el espacio tridimensional ordinario se representan los subespacios mediante rectas o planos que pasan por el origen y, en casos extremos por el origen mismo o por todo el espacio.

DEFINICION 6.6 Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector $v \epsilon V$ es ortogonal a cada vector $w \epsilon W$: $v^T w = 0$ para toda v y w.

Ejemplo 6.8.

Suponga que V es el plano generado por $v_1 = (1,0,0,0)$ y $v_2 = (1,1,0,0)$ y W es una recta generada por $w_1 = (0,0,4,5)$, como $w_1{}^T v_1 = 0$ y $w_2{}^T v_2 = 0$ entonces la recta W será ortogonal a todo el plano V.

En el caso de los cuatro subespacios fundamentales asociados a una matríz dos son subespacios de R^n : el espacio nulo y el espacio renglón o fila y los otros dos están en R^m, como se muestra en la figura 6.2

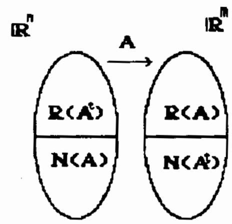


Figura 6.2

asi tenemos el siguiente resultado:

TEOREMA 6.8

Para cualquier matríz A de orden mxn, el espacio nulo y el espacio fila son subespacios ortogonales de R^n. Análogamente, el espacio nulo izquierdo y el espacio columna son subespacios ortogonales de R^m.

demostración:

Supongamos que $w \epsilon N(A)$ y $v \epsilon R(A^t)$, entonces $Aw = 0$ y v es de la forma $v = A^t x$ para algún vector x (esto es, v es una combinación de las columnas de A^t). Por lo tanto

$$w^t v = w^t (A^t x) = (w^t A^t) x = (Aw)^t x = 0^t x = 0$$

Ejemplo 6.9

Sean la matríz A y su correspondiente U como sigue:

$$A = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 2 & 8 & 0 \end{bmatrix} \qquad U = \begin{bmatrix} 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

la segunda columna es básica y las otras tres son variables libres, por lo tanto resolviendo para $Ux = 0$, se obtiene que una base para $N(A)$ está dada por :

$$\{[1,0,0,0]^t, [0,-4,1,0]^t, [0,0,0,1]^t\}$$

la cual debe ser ortogonal a los renglones de A.

El espacio columna de A es unidimensional y está generado por la única columna básica $[1,2]^t$. Por otro lado hallamos una base del espacio nulo izquierdo, dada por el vector $y^t = [-2,1]$, el cual es ortogonal al espacio columna:

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

Sin embargo cabe aclarar que el espacio nulo $N(A)$ no solo contiene algunos de los vectores ortogonales al espacio fila, contiene todos esos vectores. El espacio nulo se formó con todas las soluciones de $Ax = 0$.

DEFINICION 6.7 Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por $V^{\perp}$.

Asi se tiene entonces que el espacio nulo $N(A)$ es el complemento ortogonal de $R(A^T)$: $N(A) = (R(A^T))^{\perp}$. Al mismo tiempo es válida la relación opuesta: el espacio fila $R(A^T)$ contiene todos los vectores que son ortogonales al espacio nulo.

Al aplicar el mismo razonamiento a A^T, se produce el resultado dual: el espacio nulo izquierdo $N(A^T)$ y el espacio columna $R(A)$ son complementos ortogonales entre si en R^m. Esto completa la segunda parte del teorema fundamental del álgebra lineal.

TEOREMA 6.9:TEOREMA FUNDAMENTAL DEL ALGEBRA LINEAL. segunda parte.

$$N(A) = (R(A^T))^{\perp}, \qquad R(A^T) = N(A)^{\perp}$$
$$N(A^T) = (R(A))^{\perp}, \qquad R(A) = (N(A^T))^{\perp}.$$

La última igualdad significa que $Ax = b$ tiene una solución si y sólo si b es ortogonal a $N(A^T)$; b está en el espacio columna si y sólo si es ortogonal a cada solución y de la ecuación homogénea transpuesta $A^T y = 0$.

Sin embargo dos subespacios V y W pueden ser ortogonales sin ser complementos ortogonales entre si. En R^3, la recta V generada por (1,0,0) es ortogonal a la recta W generada por (0,0,1), pero V no es igual $W^{\perp}$. El complemento ortogonal de W es un subespacio bidimensional (un plano) que contiene todos los vectores de la forma $(x_1, x_2, 0)$, esto es todos los generados por los vectores (1,0,0) y (0,1,0). La recta V sólo puede ser parte de $W^{\perp}$ porque su dimensión es muy pequeña. Sin embargo, si las dimensiones son correctas, los dos subespacios ortogonales son necesariamente complementos ortogonales, asi fue en el caso del espacio fila y el espacio nulo. Además si $W = V^{\perp}$, esto asegura que las dimensiones son las correctas y automáticamente $V = W^{\perp}$. Simplemente se descompone el espacio en dos partes perpendiculares V y W

<u>TEOREMA 6.10</u>

Si V y W son subespacios de R^n, entonces cualquiera de las siguientes condiciones los obliga a ser complementos ortogonales entre si:

1. $W = V^{\perp}$ (W consta de todos los vectores ortogonales a V)
2. $V = W^{\perp}$ (V consta de todos los vectores ortogonales a W)
3. V y W son ortogonales y $\dim V + \dim W = n$.

Suponiendo cualquiera de éstas tres condiciones equivalentes, cada vector x puede descomponerse de una sola manera en una suma $x = v + w$ con $v \epsilon V$ y $w \epsilon W$. Estas componentes, las proyecciones de x en V y W son ortogonales $v^T w = 0$.

Una x arbitraria se descompone en $x_r + x_n$, y A transforma la componente x_r del espacio fila en un vector $Ax_r = Ax$ en el espacio columna, mientras que transforma la componente x_n del espacio nulo en cero.

<u>Ejemplo 6.10</u>

Sea la matriz A como sigue:

$$A = \begin{bmatrix} 0 & 1 & 4 \\ 0 & 2 & 8 \end{bmatrix}$$

su correspondiente forma escalonada queda como sigue:

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

de aqui que dim $R(A^t) = 1$, dim $N(A) = 2$

una base de $R(A^t) = \{ (0, 1, 4)\}$

una base de N(A) está dada por {(1, 0, 0),(0, 4, 1)} entonces cualquier $x \in R^3$ se puede escribir como:

$$x = \underbrace{\alpha_1 \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix}}_{\text{base de } R(At)} + \underbrace{\alpha_2 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + \alpha_3 \begin{bmatrix} 0 \\ 4 \\ 1 \end{bmatrix}}_{\text{base de } N(A)}$$

6.4 ESPACIOS CON PRODUCTO INTERIOR Y EL PROCESO DE GRAM-SCHMIDT.

<u>DEFINICION 6.8</u> Un producto interior en un espacio vectorial V es una función que a cada par de vectores x y y en V asocia un número real $\langle x, y\rangle$ de tal manera que se satisfacen los siguientes axiomas, para todos los vectores x, y, z en V y todos los escalares k.

1. $\langle x, y\rangle = \langle y, x\rangle$ axioma de la simetría
2. $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$ axioma de la aditividad
3. $\langle kx, y\rangle = k \langle x, y\rangle$ axioma de la homogeneidad
4. $\langle x, x\rangle \geq 0$ y $\langle x, x\rangle = 0 \iff x = 0$ axioma de positividad.

Un espacio vectorial que tiene definido un producto interior se denomina espacio vectorial con producto interior.

Las siguientes propiedades adicionales son una consecuencia inmediata de los cuatro axiomas del producto interior:

a. $\langle 0, x\rangle = \langle x, 0\rangle = 0$
b. $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$
c. $\langle x, ky\rangle = k\langle x, y\rangle$

<u>Ejemplo 6.11</u>

Considere el espacio R^n con producto interior dado como

$$\langle x, y\rangle = x^t\, y = \sum_{i=1}^{n} x_i\, y_i$$

donde x, y son vectores columna en R^n.

Note que $<., \ ,>$ o $<.|.>$ es un producto interior ya que:

1. $<x, y> = x^t y = y^t x = <y, x>$
2. $<x+y, z> = (x+y)^t z = (x^t + y^t) z = x^t z + y^t z = <x, z> + <y, z>$
3. $<kx, y> = (kx)^t y = kx^t y = k <x, y>$
4. $<x, x> = x^t x = \Sigma x_i^2 \geq 0$ si $\Sigma x_i^2 = 0$ entonces $x = 0 \in R^n$

Ejemplo 6.12

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces

$$<x, y> = 3x_1 y_1 + 2x_2 y_2$$

define un producto interior.

Solución:

Para verificar ésta afirmación observe que:

1. $<x, y> = 3x_1 y_1 + 2x_2 y_2 = 3y_1 x_1 + 2y_2 x_2 = <y, x>$
2. $<x+y, z> = 3(x_1 + y_1)z_1 + 2(x_2 + y_2)z_2 = (3x_1 z_1 + 2x_2 z_2) + (3x_1 z_1 + 2x_2 z_2) = <x, z> + <y, z>$
3. $<kx, y> = 3(kx_1)y_1 + 2(kx_2)y_2 = k(3x_1 y_1 + 2x_2 y_2) = k<x, y>$
4. $<x, x> = 3x_1 x_1 + 2x_2 x_2 = 3x_1^2 + 2x_2^2 \geq 0$ y $3x_1^2 + 2x_2^2 = 0$ si y solo si $x_1 = x_2 = 0$.

El producto interior de éste ejemplo es distinto al producto interior de R^2 lo que muestra que un espacio vectorial puede tener mas de un producto interior definido en el.

Ejemplo 6.13

Sean $x = <x_1, x_2>$, $y = <y_1, y_2>$ vectores en R^2, entonces

$$<x, y> = x_1 + y_1$$

verifique si $<x, y>$ define un producto interior.

Solución:

Para verificar ésta afirmación observe que:

1. $<x, y> = x_1 + y_1 = y_1 + x_1 = <y, x>$
2. $<x+y, z> = (x_1 + y_1) + z_1 = (x_1 + z_1) + y_1 = <x+z, y> \neq <x, z> + <y, z>$

por lo que se puede concluir que $<x, y>$ asi definido no es un producto interior.

Ejemplo 6.14

(Caso de polinomios) Sea P_n el espacio vectorial de los polinomios de grado $\leq$ n. Un producto interior en éste espacio está dado por:

$$<x, y> = \int_a^b x(t)\ y(t)\ dt$$

1.
$$<x, y> = \int_a^b x(t)\ y(t)\ dt = \int_a^b y(t)\ x(t)\ dt = <y, x>$$

2.
$$<x, y+z> = \int_a^b x(t)\ [y(t)+ z(t)]\ dt = \int_a^b [x(t)\ y(t)+ x(t)\ z(t)]\ dt$$

$$= \int_a^b x(t)\ y(t)\ dt + \int_a^b x(t)\ z(t)\ dt = <x, y> + <y, z>$$

3.
$$<kx, y> = \int_a^b kx(t)\ y(t)\ dt = k \int_a^b x(t)\ y(t)\ dt = k<x, y>$$

4.
$$<x, x> = \int_a^b x^2(t)\ dt = 0 \quad \text{si y sólo si } x = 0$$

DEFINICION 6.9 Si V es un espacio con producto interior, entonces se llama un espacio normado si existe una función $\| \ \|: V \dashrightarrow R$ tal que:

a) $\|x + y\| \leq \|x\| + \|y\|$
b) $\|\alpha x\| = |\alpha|\ \|x\|$
c) $\|x\| \geq 0$ y $\|x\| = 0 \iff x = 0$

Ejemplo 6.15

Muestre que el espacio $V = R^n$ es normado con cualquiera de las siguientes normas:

a) $\|x\|_1 = \sum_{i=1}^{n} |x_i|$

b) $\|x\|_2 = (\sum_{i=1}^{n} x_i^2)^{½}$

c) $\|x\|_\infty = \max \{|x_i| \ \ i = 1, 2, \ldots, n\}$

Solución:

Vamos a demostrar que efectivamente a) b) y c) son normas, esto es:

a) $\|x + y\|_1 = \sum_{i=1}^{n} |x_i + y_i| \leq \sum_{i=1}^{n} |x_i| + \sum_{i=1}^{n} |y_i| = \|x\|_1 + \|y\|_1$

$\|\alpha x\| = \sum_{i=1}^{n} |\alpha x_i| = |\alpha| \sum_{i=1}^{n} x_i = |\alpha| \ \|x\|_1$

$\|x\| = \sum_{i=1}^{n} |x_i| \geq 0$ y $\|x\| = 0 \iff x = 0.$

b) $\|x + y\|_2 = \sum_{i=1}^{n} (x_i + y_i)^2 = \sum_{i=1}^{n} (x_i^2 + 2x_i\, y_i + y_i^2) =$

$= \sum_{i=1}^{n} x_i^2 + 2 \sum_{i=1}^{n} x_i\, y_i + \sum_{i=1}^{n} y_i^2 = \|x\|_1 + \|y\|_1$

$\|\alpha x\| = \sum_{i=1}^{n} |\alpha x_i| = |\alpha| \sum_{i=1}^{n} x_i = |\alpha| \ \|x\|_1$

$\|x\| = \sum_{i=1}^{n} |x_i| \geq 0$ y $\|x\| = 0 \iff x = 0.$

c) $\|x + y\|_\infty = \max \{ |x_i + y_i|,\ i= 1,2,\ldots,n\} \leq \max \{ |x_i| + |y_i|\ i = 1,2,\ldots,n\} \leq \max \{|x_i|\ i=1,2,\ldots,n\} + \max \{|y_i|\ i = 1,2,\ldots,n\} = \|x\|_\infty + \|y\|_\infty$

$\|\alpha x\|_\infty = \max \{|\alpha x_i|\ i = 1,2,\ldots,n\} = \max \{|\alpha| |x_i|\ i = 1,2,\ldots,n\} = |\alpha| \max\{|x_i|\ i = 1,\ldots,n\} = \alpha \ \|x\|_\infty$

6.5 ANGULO EN LOS ESPACIOS VECTORIALES, PROYECCIONES ORTOGONALES Y PROCESO DE GRAM-SCHMIDT.

Ya vimos en la sección anterior que el hecho de que dos vectores x y y sean ortogonales se puede expresar como $x^t y = 0$ donde x y son vectores columna.

Considere el espacio vectorial R^2 y los vectores a y b donde sus longitudes están dadas por $\|a\|$ y $\|b\|$ respectivamente. Sea α el ángulo que forma el vector a con el eje de las x y β el ángulo del vector b con el mismo eje como se ilustra en la figura 6.3

se definen entonces el seno y coseno de α como

$$\operatorname{sen} \alpha = \frac{a_2}{\|a\|} \qquad \cos \alpha = \frac{a_1}{\|a\|}$$

y para β se tiene

$$\operatorname{sen} \beta = \frac{b_2}{\|a\|} \qquad \cos \beta = \frac{b_1}{\|a\|}$$

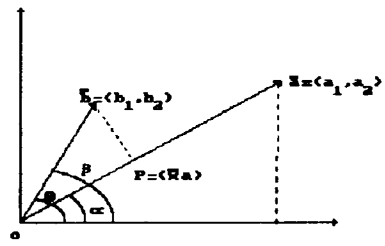


Figura 6.3

si $\theta = \beta - \alpha$ entonces usando una identidad trigonométrica se tiene:

$$\cos \theta = \cos \beta \cos \alpha + \operatorname{sen} \beta \operatorname{sen} \alpha = a_1 b_1$$

por lo tanto el coseno del ángulo entre dos vectores cualesquiera a y b está dado por

$$\cos \theta = \frac{\langle a, b\rangle}{\|a\| \, \|b\|}$$

Si queremos encontrar la proyección de un punto del vector b, sobre el vector a, este punto debe ser algún múltiplo p = xa del vector a y el problema se reduce a calcular el coeficiente x. Todo lo que necesitamos es el hecho de que la recta desde b hasta el punto mas cercano p = xa es ortogonal al vector a, como se puede ver en la figura 6.3, entonces:

$$(b - xa) \perp a \quad \text{o} \quad \langle a, b-xa\rangle = 0 \quad \text{de donde}$$

$$x = \frac{\langle a, b\rangle}{\langle a, a\rangle} = \frac{\langle a, b\rangle}{\|a\|^2}$$

entonces la proyección p del punto b sobre la recta generada por el vector a está dada por

$$p = \frac{\langle a, b\rangle}{\|a\|^2} a$$

de esta relación se puede encontrar nuevamente la desigualdad de Cauchy- Schwarz ya que la distancia al cuadrado del punto a la recta es:

$$\left\| b - \frac{<a, b>}{\|a\|^2} a \right\|^2 = \frac{\|b\|^2 \; \|a\|^2 - <a, b>}{\|a\|^2}$$

como la distancia de un punto a una recta es mayor o igual a cero lo mismo sucede con el cuadrado de la distancia entonces

$$\|b\|^2 \; \|a\|^2 - <a, b> \; \geq 0$$

$\|b\|^2 \; \|a\|^2 \; \geq \; <a, b>$ que es la desigualdad de Schwarz.

Al seleccionar una base se busca que ésta simplifique la solución del problema que se trate, en la mayoría de los casos, la mejor selección será la base en la cual todos los vectores son ortogonales entre si, es asi como definimos:

DEFINICION 6.10 Se dice que un conjunto de vectores en un espacio con producto interior es ortonormal si cada vector del conjunto tiene norma 1 y si además cualesquiera dos vectores distintos en el conjunto son ortogonales.

Ejemplo 6.16

Sean

$v_1 = (0, 1, 0)$, $v_2 = (1/\sqrt{2}, 0, 1/\sqrt{2})$, $v_3 = (1/\sqrt{2}, 0, -1/\sqrt{2})$, el conjunto $S = \{v_1, v_2, v_3\}$ es ortonormal si en R^3 se tiene que con el producto interior euclídeano:

$$<v_1, v_2> = <v_2, v_3> = <v_1, v_3> = 0$$

y además

$$\|v_1\| = \|v_2\| = \|v_3\| = 1$$

Ejemplo 6.17

Si v es un vector en un espacio con producto interior, y v es diferente de cero, entonces el vector

$$\frac{1}{\|v\|} v \qquad \text{tiene norma 1}$$

ya que aplicando $\|ku\| = |k| \; \|u\|$ se tiene

$$\left\| \frac{1 \; v}{\|v\|} \right\| = \frac{1}{\|v\|} \|v\| = 1$$

al proceso de multiplicar un vector v diferente de cero por el recíproco de su longitud para obtener un vector de norma 1 se denomina <u>normalización de v.</u>

<u>TEOREMA 6.11: TEOREMA DE LA BASE ORTONORMAL (GRAM-SCHMIDT)</u>

Todo subespacio distinto de cero en R^n tiene una base ortonormal.

<u>demostración:</u>

Queremos producir de dos vectores independientes ay b dos vectores perpendiculares v_1 y v_2. Claramente, el primero puede ir en la dirección de a $v_1 = a$. El problema es entonces encontrar un segundo vector que sea perpendicular.

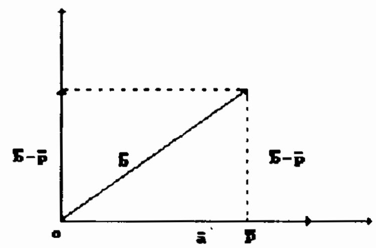


Figura 6.4

En la figura 6.4 se puede ver que ese vector es b-p ya que se le sustrajo la proyección de b en la dirección de a, entonces el segundo vector queda:

$$v_2 = b - p = b - \frac{\langle a, b\rangle}{\|a\|^2} a = b - \frac{\langle v_1, b\rangle}{\|v_1\|^2} v_1$$

si hay un tercer vector independiente entonces la idea es la misma, sustraemos las proyecciones de c en la dos direcciones v_1 y v_2

$$v_3 = c - \frac{\langle v_1, c\rangle}{\|v_1\|^2} v_1 - \frac{\langle v_2, c\rangle}{\|v_2\|^2} v_2$$

es inmediato que v_3 es perpendicular a v_1 y v_2.

para que estos vectores sean ortonormales deben transformarse en vectores unitarios entonces

$$q_1 = \frac{v_1}{\|v_1\|} \qquad q_2 = \frac{v_2}{\|v_2\|} \qquad q_3 = \frac{v_3}{\|v_3\|}$$

este proceso se puede resumir de la siguiente manera:

Es posible convertir cualquier conjunto de vectores independientes $a_1, a_2, \ldots, a_n$ en un conjunto de vectores ortogonales mediante el proceso de Gram-Schmidt: primero se fija $a_1 = v_1$ después cada v_i es ortogonal a las $v_1, \ldots, v_{i-1}$ precedentes:

$$v_i = a_i - \frac{<v_1, a_i>}{\|v_1\|^2} v_1 - \ldots - \frac{<v_{i-1}, a_i>}{\|v_{i-1}\|^2} v_2$$

<u>Ejemplo 6.18</u>

Aplique el proceso de Gram-Schmidt a los siguientes vectores independientes: $a_1 = [1, 1, 0]$, $a_2 = [1, 0, 1]$, $a_3 = [0, 1, 1]$

<u>Solución:</u>

Sea $a_1 = v_1$, v_2 y v_3 se calculan:

$$v_2 = a_2 - \tfrac{1}{2} v_1 = [\tfrac{1}{2}, -\tfrac{1}{2}, 1]$$

$$v_3 = a_3 - \tfrac{1}{2} v_1 - 1/3\, v_2 = [-2/3, 2/3, 2/3]$$

los vectores ortonormales finales son:

$$q_1 = \sqrt{1/2}\, [1, 1, 0]$$
$$q_2 = \sqrt{2/3}\, [\tfrac{1}{2}, -\tfrac{1}{2}, 1]$$
$$q_3 = \sqrt{3/4}\, [-2/3, 2/3, 2/3]$$

6.6 MATRICES DE PROYECCION Y MINIMOS CUADRADOS

Considere ahora la proyección p de un vector sobre un subespacio general $S = \{a_1, a_2, \ldots, a_k\}$ de R^n donde las a_i son vectores independientes de R^n. Los detalles se elaboran para el caso $k = 2$, pero los cálculos son igualmente válidos para el caso general. Recordemos que $\{a_1, a_2\}$ corresponde a un plano en R^n que contiene al origen, y sus miembros son todas las combinaciones lineales de a_1 y a_2.

La figura 6.5 muestra al vector proyección p de b en $\{a_1, a_2\}$

Observe que p satisface:

1. p debe estar en el subespacio $S = \{a_1, a_2\}$ y
2. b-p debe ser perpendicular a cada vector de S.

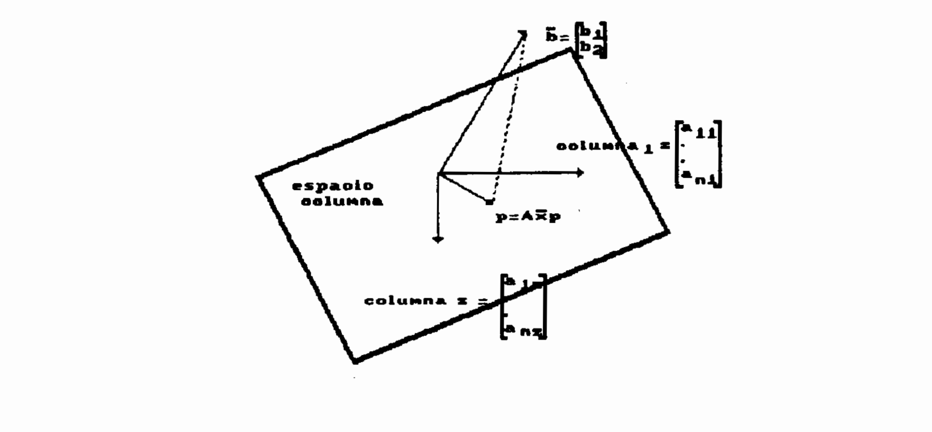


Figura 6.5

Si escribimos los vectores de R^n como vectores columna, el subespacio S es el espacio columna de la matriz A de nx2 cuyas columnas son a_1 y a_2 entonces como $Ax = b$ tiene solución si y sólo si b está en el espacio columna de A, todos los vectores de S tienen la forma Ax donde $x = [x_1, x_2]^t$. Como p está en el espacio S entonces $p = Ax_p$ donde $x_p = [r_1, r_2]^t$, r_1, r_2 escalares, como $b - Ax_p$ debe ser perpendicular a cada vector de S entonces el producto punto de $b- Ax_p$ y Ax debe ser igual a cero, para toda x.

$$<b-Ax_p, Ax> = 0$$

o bien

$$(Ax)^t \ (b - Ax_p) = x^t(A^tb - A^tAx_p) = 0$$

es decir, el producto punto de los vectores x y $A^tb - A^tAx_p$ debe ser cero para todos los vectores x, pero esto solo sucede si

$$A^t b - A^tA \ x_p = 0$$

La matriz A^tA de 2x2 es invertible* pues tiene el mismo rango que A, entonces despejando x_p se tiene

$$x_p = (A^tA)^{-1} A^t b$$

Denotando el vector proyección $p = Ax_p$ de b sobre S por medio de $p = b_s$ y escribiendo b como un vector columna obtenemos la forma general:

Sea $S = \{a_1, \ldots, a_k\}$ un subespacio de R^n y sea la matriz A cuyas columnas son los vectores $a_1, a_2, \ldots, a_k$. El vector proyección de b en R^n sobre S está dado por :

$$b_s = A(A^tA)^{-1} A^tb$$

este resultado se puede usar para una proyección de un vector en otro, por ejemplo:

Ejemplo 6.19

Sea el vector a = (2, 4, 3), encuentre el vector proyección sobre a del vector b = (1, 2, 3)

Solución:

$$p = \frac{\langle a, b\rangle}{\|a\|^2}\, a = \frac{19}{29}\,(2, 4, 3)$$

Considerando la fórmula $b_s = A(A_tA)^{-1} A^t b$ se tiene que la matriz A consiste de una sola columna que es el vector a, haciendo cálculos se tiene:

$$A^tA = 29$$

$$b_s = A(A^tA)^{-1} A^tb = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \frac{1}{29} (2, 4, 3) \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \frac{1}{29} \begin{bmatrix} 4 & 8 & 6 \\ 8 & 16 & 12 \\ 6 & 12 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} =$$

$$= \frac{1}{29} \begin{bmatrix} 38 \\ 76 \\ 57 \end{bmatrix}$$

A la matriz $P_s = A(A^tA)^{-1}A^t$ se le conoce como la matriz de proyección para el subespacio S.

Ejemplo 6.20

Encuentre la matriz de proyección para el plano x_2x_3 de R^3.

Solución:

El plano x_2x_3 es el subespacio S generado por los vectores e_2e_3 que son las columnas de la siguiente matriz A.

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

donde $A^tA = I$ y por lo tanto

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

asi P_s proyecta cada vector b en R^3 en el plano x_2x_3 de la siguiente manera:

$$P_s \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ b_2 \\ b_3 \end{bmatrix}$$

Ejemplo 6.21

Sean la matriz A y el vector b como sigue:

$$A = \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 4 \\ 3 \\ 9 \end{bmatrix}$$

encuentre la proyección del vector b en el sobre el plano generado por los vectores columna de la matriz A.

Solución:

El espacio columna de A es el plano x-y en el espacio tridimensional ya que ambas columnas terminan con un cero. La proyección de b sobre este plano no altera las componentes x, y que son 4 y 3 pero desaparecerá la componente z y se puede confirmar.

$$A^tA = \begin{bmatrix} 1 & 1 & 0 \\ 2 & 5 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 5 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 7 \\ 7 & 29 \end{bmatrix}$$

$$(A^tA)^{-1} = (1/9) \begin{bmatrix} 29 & -7 \\ -7 & 2 \end{bmatrix}$$

$$b_s = \begin{bmatrix} 4 \\ 3 \\ 0 \end{bmatrix}$$

* TEOREMA 6.12

Para cualquier matriz A de mxn, de rango r, la matriz simétrica A^tA de nxn también tiene rango r.

demostración:

Trabajando con los espacios nulos se tiene que si v es cualquier vector solución del sistema Ax = 0, de modo que Av = 0 entonces, al multiplicar a la izquierda ambos miembros de esta última ecuación por A^t, vemos que v también es una solución al sistema $(A^t)Ax = 0$. Recíprocamente, supongamos que $(A^t)\ Aw = 0$ para un vector w de nx1. Entonces,

$$0 = w^t\ [(A^t)Aw] = (Aw)^t(Aw)$$

que se puede escribir como $\|Aw\|^2$. Esto es, Aw es un vector de magnitud cero y, en consecuencia, $Aw = 0$. Resulta así que A y $(A^t)A$ tienen el mismo espacio nulo. Como tienen el mismo número de columnas, se concluye que el rango de A y el rango de $(A^t)A$ son iguales.

De este teorema se desprende el siguiente resultado: Si A tiene columnas linealmente independientes de modo que $r = n$, entonces A^tA es una matriz cuadrada, simétrica e invertible.

MATRICES ORTOGONALES

Una matriz ortogonal es simplemente una matriz cuadrada con columnas ortonormales. Se usará la letra Q para denotar una matriz ortogonal y a $q_1,\ldots,q_n$ para denotar sus columnas.

Como Q es cuadrada de nxn entonces si sus columnas son independientes, esto es su rango r es tal que $r = n$ y es invertible, entonces: si Q^t es inversa izquierda entonces es la inversa y $QQ^t = I$

Una matriz ortogonal tiene las siguientes propiedades:

$$Q^tQ = I = QQ^t$$

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

Q no solo tiene las columnas ortonormales, sino que también sus fils son ortonormales, en otras palabras si Q es ortogonal también lo es Q^t.

Además cumple: La multiplicación por una matriz ortogonal Q preserva la longitud

$$\|Qx\| = \|x\| \qquad \text{para todo vector } x$$

y también preserva los productos internos:

$$(Qx)^t\ (Qy) = x^t\ y \text{ para todos los vectores } x,y.$$

<u>Ejemplo 6.22</u>

Sea la matriz Q como sigue:

$$Q = \begin{bmatrix} \cos\theta & -\text{sen}\ \theta \\ \text{sen}\ \theta & \cos\theta \end{bmatrix}$$

su transpuesta es igual a su inversa

$$Q^t = Q^{-1} = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

Q rota cada vector en un ángulo θ y Q^t lo rota de regreso en $-\theta$

Ejemplo 6.23

Verifique que la matriz siguiente es una matriz ortogonal y encuentre su inversa.

$$A = (1/7) \begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

Solución:

$$AA^t = I$$

$$A^{-1} = A^t = (1/7) \begin{bmatrix} 2 & 3 & 6 \\ 3 & -6 & 2 \\ 6 & 2 & -3 \end{bmatrix}$$

METODO DE LOS MINIMOS CUADRADOS

Supongamos que se obtienen datos de mediciones de la forma (a_i, b_i) mediante observación o experimentación y se ubican como puntos de datos en el plano x-y. Es conveniente hallar una relación matemática $Y = f(x)$ que represente razonablemente bien los datos, de manera que podamos efectuar predicciones de valores no medidos. Dependiendo de la naturaleza del experimento y de la configuración de los puntos de datos localizados, podemos decidir acerca de un tipo apropiado de función $y = f(x)$ como una función lineal, una función cuadrática o exponencial, entre otras. Se pueden presentar problemas como el siguiente:

De acuerdo con la ley de Hooke, la distancia que se estira un resorte es proporcional a la fuerza aplicada. Supongamos que colocamos 4 pesos distintos, a_1, a_2, a_3 y a_4 sucesivamente en la parte inferior de un resorte. Medimos las cuatro longitudes b_1, b_2, b_3 y b_4 del resorte estirado y suponga que se obtienen los datos de la siguiente tabla:

a_i = peso en gramos	2.0	4.0	5.0	6.0
b_i = longitud en cms.	2.5	4.5	7.0	8.5

Debido a la ley de Hooke, esperamos que los puntos de datos (a_i, b_i) estén cerca de alguna recta con ecuación

$$y = f(x) = r_0 + r_1 x$$

donde r_0 es la longitud del resorte y r_1 es la constante del resorte. estoes, si nuestras mediciones fueran exactas y el resorte ideal, tendríamos $b_i = r_0 + r_1 a_i$ para valores espacíficos r_0 y r_1.

Como solamente tenemos las dfos incógnitas r_0 y r_1, bastarán dos mediciones para hallarlas, sin embargo en la práctica esperamos tener algún error en las mediciones físicas, por lo tanto hacemos mas mediciones de las técnicamente necesarias con la esperanza de que en general, los errores se cancelen unos con otros. La sustitución de cada punto de datos (a_i, b_i) en la ecuación da una ecuación lineal con dos incógnitas r_0 y r_1, asi los 4 puntos de datos del problema, dan lugar a un sistema lineal de 4 ecuaciones con 2 incógnitas. Dicho sistema lineal con mas ecuaciones que incógnitas se llama <u>sobredeterminado</u>, y se espera que dicho sistema sea inconsistente, entonces debemos encontrar valores para las incógnitas r_0 y r_1 que estén los mas cerca posible, de satisfacer las 4 ecuaciones. Geométricamente esto equivale a encontrar la recta en el plano que esté mas cerca de pasar por los 4 puntos de datos.

En general, consideramos el problema de hallar una recta o función lineal $f(x) = r_0 + r_1 x$ que ajuste "mejor" los daots (a_i, b_i) para $i = 1, 2, \ldots, m$ donde $m>2$. Si no hubiera error en las mediciones y los datos fueran realmente lineales, entonces para algún r_0 y r_1 se tendría:

$$b_i = r_0 + r_1 a_i \qquad \text{para } i = 1,2,\ldots,m$$

Estas m ecuaciones en las 2 incógnitas r_0 y r_1 forman un sistema sobredeterminado de ecuaciones que probablemente no tenga solución. Pero nuestros puntos satisfacen realmente un sistema de aproximaciones lineales que se pueden expresar como:

$$\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ b_n \end{bmatrix} \approx \begin{bmatrix} 1 & a_1 \\ 1 & a_2 \\ \cdots & \cdots \\ \cdots & \cdots \\ 1 & a_n \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix} \tag{6.1}$$

o simplemente $b \approx Ar$.

Tratamos de encontrar un vector solución óptimo r para el sistema (6.1). Para cada vector r el vector error $Ar - b$ mide a que distancua esta el sistema (6.1) de un sistema con solución r. Las normas de las componentes del vector $Ar - b$ representan las distancias $d_i = r_0 + r_1 a_i - b_i$ como se muestra en la figura 6.6

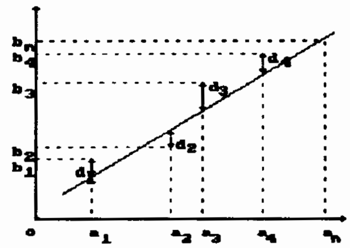


Figura 6.6

queremos entonces minimizar nuestro vector de error Ar - b. Minimizaremos la longitud $\|Ar - b\|$ del vector error, pero esto equivale a minimizar $\|Ar - b\|^2$ o lo que es lo mismo, minimizar

$$d_1^2 + d_2^2 + \ldots + d_m^2$$

de aqui el nombre de mínimos cuadrados.

Si a_1 y a_2 denotan las columnas de A en el sistema (6.1) entonces el vector $Ar = r_0\, a_1 + r_1\, a_2$ está en el espacio columna de A.

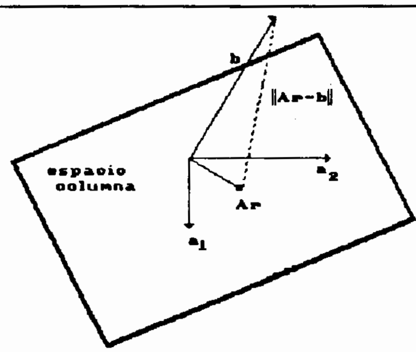


Figura 6.7

Entonces de la figura 6.7, resulta claro que de todos los vectores Ar en el espacio columna el que minimiza $\|Ar - b\|$ es la proyección b_s = Ar de b en el espacio columna S. Entonces $Ar = A(A^tA)^{-1}A^t\, b$; y de aqui se tiene que el vector solución r que es óptimo está dado por:

$$r = (A^tA)^{-1} A^t b.$$

Ahora si, volviendo a los datos del problema, hacemos el ajuste de mínimos cuadrados por medio de una recta, esto es:

$$y = r_0 + r_1 x$$

formamos el sistema $y \approx Ar$

$$\begin{bmatrix} 2.5 \\ 4.5 \\ 7 \\ 8.5 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 4 & 5 & 6 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 1 & 4 \\ 1 & 5 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}$$

$$(A^tA)^{-1} = (1/35) \begin{bmatrix} 81 & -17 \\ -17 & 4 \end{bmatrix}$$

entonces $r = (A^tA)^{-1} A^t b = [-0.9,\ 1.5]^t$

de donde la recta que mejor ajusta los datos esta dada por la ecuación $y = -0.9 + 1.5\ x$

Ejemplo 6.24

Si se varía la carga que se aplica a una estructura y se mide la deformación que produce donde x es la carga, y es la lectura del medidor de la deformación. A menos que la carga sea tan grande que el material se haga plástico lo normal en la teoría de la elasticidad es la relación lineal $y = r_0 + r_1 x$. Dadas las siguientes mediciones encuentre la recta que mejor se ajusta.

x	0	1	3	4
y	0	1	2	5

Solución:

formamos el sistema $y \approx Ar$

$$\begin{bmatrix} 0 \\ 1 \\ 2 \\ 5 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} r_0 \\ r_1 \end{bmatrix}$$

$$A^tA = \begin{bmatrix} 4 & 17 \\ 17 & 81 \end{bmatrix}$$

$$(A^tA)^{-1} = (1/20) \begin{bmatrix} 13 & -4 \\ -4 & 2 \end{bmatrix}$$

entonces $r = (A^tA)^{-1} A^t b = [-0.2, 1.1]^t$

de donde la recta que mejor ajusta los datos esta dada por la ecuación $y = -0.2 + 1.1 x$

Ejemplo 6.25

Una población de conejos de una gran isla se estimó todos los años desde 1981 hasta 1984 y se obtuvieron los datos que se listan en la siguiente tabla:

a_i = año observado	1	2	3	4
b_i = # de conejos en unidades de 1000	3	4.5	8	17
z_i = Ln b_i	1.1	1.5	2.08	2.83

Sabiendo que el crecimiento de la población es exponencial en ausencia de enfermedades, predadores, hambre, etc. se espera que una función exponencial

$$b_i \approx re^{sa_i} \qquad \text{o} \qquad b = re^{sa}$$

sea la mejor representación de los datos. Encuentre lafunción exponencial que mejor ajusta los datos y con ella haga una proyección de la población de la isla para 1991.

Solución:

Observe que usando logaritmos es posible convertir esta función exponencial a la forma lineal

$$\text{Ln } b = \text{Ln } r + s(a)$$

de donde en la tabla se agraga el renglón de $z = \ln (b_i)$

Entonces como es lineal la función asi escrita se usa

$$x = (A^tA)^{-1}A^t b$$

donde

$$\begin{bmatrix}1.1\\1.5\\2.08\\2.83\end{bmatrix} \approx \begin{bmatrix}1 & 1\\1 & 2\\1 & 3\\1 & 4\end{bmatrix} \begin{bmatrix}r_0\\r_1\end{bmatrix}$$

$$A^tA = \begin{bmatrix}4 & 10\\10 & 30\end{bmatrix}$$

$$(A^tA)^{-1} = \begin{bmatrix}3/2 & -1/2\\-1/2 & 1/5\end{bmatrix}$$

entonces $(A^tA)^{-1}\ A^t\ b = [0.435,\ 0.577]^t = [\ln r,\ s]^t$

por lo tanto $\ln b = \ln r + s(a_i) = 0.435 + 0.577\ a$ de donde

$$b = re^{sa} = e^{0.435}\ e^{0.577a} = 1.54\ e^{0.577a}$$

en términos de funciones se tiene $f(x) = 1.54\ e^{0.577x}$

a partir de esta función se puede proyectar la población de conejos para 1991 haciendo $f(11)\ 1000 = 570\ 778$ conejos.

APROXIMACION CUADRATICA

Ahora, pretendemos ajustar una curva cuadrática a los n puntos de datos. Recuerde que una cuadrática en x es cualquier expresion de la forma:

$$y = a + bx + cx^2$$

que representa a una parábola en el plano. Si los n puntos dados estuvieran en la parábola se tendría:

$$\begin{aligned} y_1 &= a + bx_1 + cx_1^2 \\ y_2 &= a + bx_2 + cx_2^2 \\ &\cdots\cdots\cdots\cdots \\ &\cdots\cdots\cdots\cdots \\ y_n &= a + bx_n + cx_n^2 \end{aligned} \tag{6.2}$$

para

$$y = \begin{bmatrix}y_1\\y_2\\\cdot\cdot\\y_n\end{bmatrix} \qquad A = \begin{bmatrix}1 & x_1 & x_1^2\\ \cdots & \cdots & \cdots\\ 1 & x_n & x_n^2\end{bmatrix} \qquad r = \begin{bmatrix}a\\b\\c\end{bmatrix}$$

entonces (6.2) se puede escribir como:

$$y = Ar$$

igual que antes, si los puntos dados no están todos en la parábola entonces $y - Ar \neq 0$ y se tiene que para cualquier r nuestro problema vuelve a ser:

Encontrar un vector r en R^3 tal que $\|y - Ar\|$ sea mínimo. y usando un razonamiento similar al anterior se tiene:

$$r = (A^tA)^{-1} A^t y$$

Ejemplo 6.26

En una reciente exhibición de yates se hicieron las observaciones listadas en la tabla siguiente, donde se relacionan los precios b_i de las embarcaciones y sus pesos a_i

a_i = peso en tons.	2	4	5	8
b_i = precio en unidades de 10 000	1	3	5	12

Al localizar los puntos de datos (a_i, b_i) como se muestra en la figura 6.8 podemos esperar que una función cuadrática de la forma

$$y = f(x) = r_0 + r_1 x + r_2 x^2$$

ajuste bien los datos

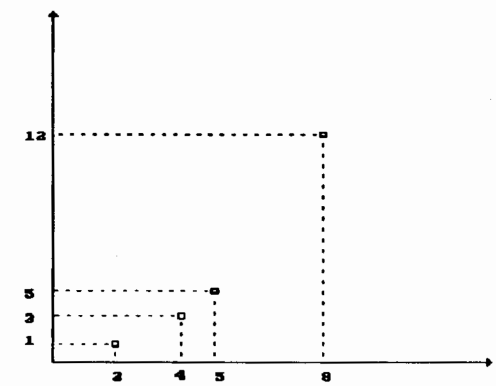


Figura 6.8

Solución:

formamos el sistema $y \approx Ar$

$$\begin{bmatrix} 1 \\ 3 \\ 5 \\ 12 \end{bmatrix} \approx \begin{bmatrix} 1 & 2 & 4 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \\ 1 & 8 & 64 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

$$A^tA = \begin{vmatrix} 4 & 19 & 109 \\ 19 & 109 & 709 \\ 109 & 709 & 4993 \end{vmatrix}$$

$$(A^tA)^{-1} = (1/5400) \begin{bmatrix} 12744 & -3348 & -6624 \\ -4464 & 2538 & 3744 \\ 360 & -270 & -360 \end{bmatrix} \begin{bmatrix} 2628 \\ -1818 \\ 270 \end{bmatrix}$$

entonces $y = (A^tA)^{-1} A^t b = [0.207,\ 0.01,\ 0.183]^t$

asi la función cuadrática que mejor aproxima los datos en el sentido de los mínimos cuadrados es:

$$y = 0.207 + 0.01x + 0.183x^2$$

para graficar esta función se hace:

a_i	b_i	$f(a_i)$
2	1	0.959
4	3	3.175
5	5	4.832
8	12	11.999

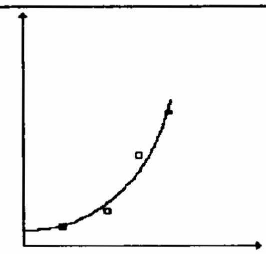


Figura 6.9

6.7 LA PSEUDOINVERSA Y LA DESCOMPOSICION EN VALOR SINGULAR.

En èste punto cabe hacernos la siguiente pregunta: ¿Cuál es la solución óptima x para el sistema inconsistente $Ax = b$? deberíamos tener una regla que especifíque x dada cualquier matriz A y cualquier lado derecho b.

Para cada b, Ax debe estar en el espacio de las columnas de A ya que es uan combinación de las columnas ponderadas por las componentes de x. Por lo tanto, la elección óptima Ax es el punto p en el espacio columna que está mas cerca a la b dada. En otras palabras, tenemos que proyectar a b sobre el espacio columna:

$$Ax = b_s = p$$

Claramente x está determinada cunado hay sólo una combinación de las columnas de A que producen p, los pesos de esta combinación serán las componentes de x. Si se tiene solución única existen varias condiciones equivalentes:

a) las columnas de A son l.i.
b) el espacio nulo de A solo contiene al cero.
c) el rango de A es n
d) la matriz cuadrada $A^T A$ es invertible.

Si estas condiciones no son válidas, entonces x no está determinada en forma única y

$$x = (A^TA)^{-1} A^T b$$

es la expresión de dicha solución. Si A es una matriz invertible entonces x coincide con la única solución al sistema original

$$x = A^{-1} (A^T)^{-1} A^T b = A^{-1} b \qquad (6.3)$$

pero si A no es invertible definimos la pseudoinversa A+ (o inversa de Moore- Penrose), por lo tanto, si A es invertible $A+ = A^{-1}$.

Cuando la matriz A cumple con las cuatro condiciones anteriores, la pseudoinversa es la inversa izquierda que aparece en (6.3):

$$A+ = (A^TA)^{-1} A^T$$

pero la pseudoinversa queda por definirse cuando no son válidas las condiciones a)-d) y x no está determinada en forma única por $Ax = p$. Tenemos que escoger uno de los muchos vectores que satisfacen $Ax = p$ y la elección será la solución óptima $x = a + b$ al sistema inconsistente $Ax = b$.

Esta elección se efectúa de acuerdo a la regla siguiente: la solución óptima de entre todas las soluciones de Ax = p es aquella con longitud mínima. La clave para encontrarla es recordar que el espacio fila y el espacio nulo de A son complementos ortogonales en R^n. Esto signifca que cualquier vector puede descomponerse en dos piezas perpendiculares, su proyección sobre el espacio fila y su proyección sobre el espacio nulo. Suponiendo que aplicamos esta descomposición a una de las soluciones x_0 de la ecuación Ax = p. Entonces $x_0 = x_r + x_n$ donde x_r está en el espacio fila y x_n está en el espacio nulo. Hay ahora tres puntos importantes:

1.- La componente x_r es una solución de Ax = p ya que $Ax_n = 0$,

$$Ax_0 = A(x_r + x_n) = Ax_r = p$$

2.- Todas las soluciones de Ax = p comparten ésta misma componente x_r en el espacio fila y difieren solamente en la componente x_n en el espacio nulo. La solución general es la suma de una solución particular x_r y una solución arbitraria x_n de la ecuación homogénea.

3.- La longitud de dicha solución $x_r + x_n$ obedece la ley de Pitágoras, ya que las dos componentes son ortogonales:

$$\| x_r + x_n \|^2 = \|x_r\|^2 + \|x_n\|^2$$

Las conclusiones que de aqui se desprenden son:

La solución que tiene longitud mínima es x_r. Deberíamos elegir como cero a la componente en el espacio nulo, dejando una solución situada completamente en el espacio fila.

La solución óptima en mínimos cuadrados para cualquier sistema Ax = b es el vector x_r determinado por las dos condiciones:

(1) Ax es igual a la proyección de b sobre el espacio columna de A.
(2) x está en el espacio fila de A.

La matriz que resuelve Ax = b es la pseudoinversa A+ definida por $x = A^+b$.

Una mejor manera de entender la pseudoinversa es viendola geométricamente, como se muestra en la figura 6.10.

La matriz A+ combina el efecto de dos pasos separados: proyecta a b sobre el punto p y despuès encuentra el único vector x en el espacio fila que resuelve Ax = p. Un caso extremo es cuando b es perpendicular al espacio columna, en otras palabras, cuando b está en el espacio nulo izquierdo. Entonces p = 0 y x = 0, A+ envía todo b-p a cero. En el otro extremo tenemos a b dentro del espacio columna entonces p = b y encontramos x al "invertir" A (ya se

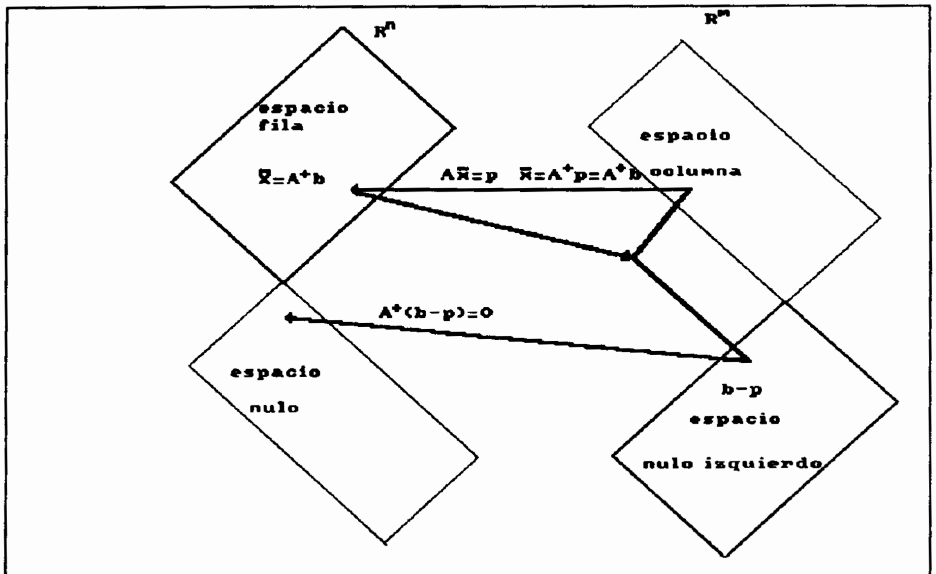


Figura 6.10

había dicho que A es invertible si la consideramos como una aplicación de su espacio fila en su espacio columna con A+ como la inversa).

Una b arbitraria está entre estos dos extremos: la componente p se invierte para dar x y la otra componente b-p se aniquila. De esta descripción y de la figura 6.10 podemos listar algunas propiedades básicas de la pseudoinversa:

1. A+ es una matriz de nxm comienza con el vector $b \in R^m$ y produce el vector $x \in R^n$.
2. El espacio columna de A+ es el espacio fila de A y el espacio fila de A+ es el espacio columna de A, de aqui que rango A = rango A+.
3. La pseudoinversa de A+ es A.
4. En general AA+ $\neq$ I, ya que es posible que A no tenga inversa derecha, pero AA+ siempre es igual a la proyección P sobre el espacio columna:

$$AA^+b = Ax = p = Pb \quad \text{o}$$
$$AA^+ = P$$

Ejemplo 6.27

Sea la matriz A una matriz no invertible:

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

Observe que esta no es otra que la matriz de proyección que manda cualquier punto de R^3 al plano x-y. De esta forma el espacio columna y el espacio fila de A coinciden con el plano x-y en R^3 que contiene todos los vectores (x, y, 0). El espacio nulo es el eje z, que es ortogonal al espacio fila. Para encontrar x se proyecta b sobre el espacio columna:

Si

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad \text{entonces} \quad p = Pb = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

y resolviendo Ax = p

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$$

se tiene:

$$x_1 = b_1 \qquad x_2 = b_2 \qquad \text{y } x_3 \text{ es arbitraria}$$

se elige entre esta familia infinita de soluciones aquella que tenga longitud mínima, donde la tercer componente debe ser cero, esto deja a $x = (b_1, b_2, 0)^T$ que está en el espacio fila (el plano x-y) y la pseudoinversa A+ está dada por A misma. Esto sucede porque A actúa como la matriz identidad en la manera como aplica su espacio fila en su espacio columna y la pseudoinversa ignora todo lo demás. Resumiendo, se tien que la solución óptima de Ax = b que es un conjunto inconsistente

$$\begin{aligned} x_1 + 0x_2 + 0x_3 &= b_1 \\ 0x_1 + x_2 + 0x_3 &= b_2 \\ 0x_1 + 0x_2 + 0x_3 &= b_3 \end{aligned}$$

es satisfacer las dos primeras ecuaciones y hacer $x_3 = 0$.

6.8 RESUMEN.

1. La dimensión de los cuatro espacios asociados a una matríz es

1. $R(A^t)$ = espacio renglón de A: dimensión = r
2. $N(A)$ = espacio nulo de A: dimensión = n-r
3. $R(A)$ = espacio columna de A: dimensión = r
4. $N(A^t)$ = espacio nulo izquierdo de A: dimensión = m-r

tal que:

$$n = \dim R(A) + \dim N(A)$$
$$m = \dim R(A^t) + \dim N(A^t)$$

2. El sistema $Ax = b$ tiene al menos una solución x para toda b si y sólo si las columnas de A generan R^m, entonces $r = m$. En éste caso existe una inversa derecha H de nxm tal que $AH = I_m$, la matriz identidad de orden m. Esto es posible sólo si $m \le n$.

El sistema $Ax = b$ tiene a lo sumo una solución x para cada b si y sólo si las columnas son linealmente independientes, entonces $r=n$. En éste caso existe una inversa izquierda G de nxm tal que $GA = I_n$, la matriz identidad de orden n. Esto es posible sólo si $m \ge n$.

3. Dos subespacios V y W del mismo espacio R^n son ortogonales si cada vector $v \in V$ es ortogonal a cada vector $w \in W$: $v^T w = 0$ para toda v y w.

4. Dado un subespacio V de R^n, el espacio de todos los vectores ortogonales a V es el complemento ortogonal de V y se denota por $V^{\perp}$. Asi se tiene entonces

$$N(A) = (R(A^T))^{\perp}, \qquad R(A^T) = N(A)^{\perp}$$
$$N(A^T) = (R(A))^{\perp}, \qquad R(A) = (N(A^T))^{\perp}.$$

5. Un producto interior en un espacio vectorial V es una función que a cada par de vectores x y y en V asocia un número real $<x, y>$ de tal manera que se satisfacen los siguientes axiomas, para todos los vectores x, y, z en V y todos los escalares k.

1. $<x, y> = <y, x>$ axioma de la simetría
2. $<x+y, z> = <x, z> + <y, z>$ axioma de la aditividad
3. $<kx, y> = k <x, y>$ axioma de la homogeneidad
4. $<x, x> \ge 0$ y $<x, x> = 0 \iff x = 0$ axioma de positividad.

6. Es posible convertir cualquier conjunto de vectores independientes $a_1, a_2, \ldots, a_n$ en un conjunto de vectores ortogonales mediante el proceso de Gram-Schmidt: primero se fija $a_1 = v_1$ después cada v_i es ortogonal a las $v_1, \ldots, v_{i-1}$ precedentes:

$$v_i = a_i - \frac{<v_1, a_i>}{\|v_1\|^2} v_1 - \cdots - \frac{<v_{i-1}, a_i>}{\|v_{i-1}\|^2} v_2$$

7. Sea $Ax = b$ un sistema lineal de m ecuaciones con n incógnitas, donde m>n (un sistema sobredeterminado) y el rango de A es n. La solución de mínimos cuadrados del sistema correspondiente $Ax \approx b$ de aproximaciones lineales es el vector $x = r$ que minimiza la magnitud del vector error.

Es decir, de todos los vectores Ar en el espacio columna el que minimiza $\|Ar - b\|$ es la proyección $b_s = Ar$ de b en el espacio columna S. Entonces $Ar = A(A^tA)^{-1}A^t\, b$; y de aqui se tiene que el vector solución r que es óptimo está dado por:

$$r = (A^tA)^{-1}\, A^t\, b.$$

6.9 NOTAS HISTORICAS.

EL RANGO DE UNA MATRIZ lo definió en 1879 Georg Frobenius (1849-1917) como sigue: si se anulan todos los determinantes de grado (r+1), pero no todos los de grado r, entonces r es el rango de la matríz. Frobenius usó este concepto para tratar las cuestiones de formas canónicas para ciertas matrices de enteros y las soluciones de ciertos sistemas de congruencias lineales.

Por otro lado James Sylvester definió en 1884 la nulidad para matrices cuadradas, como sigue: la nulidad de una matríz de nxn es i si todo menor (determinante) de orden n-i+1 (y, por tanto, de todo orden superior) es igual a 0 e i es el mayor de los números para los cuales esto es cierto. Aqui, Sylvester estaba interesado, como en buena parte de su carrera matemática, en descubrir invariantes, esto es, propiedades de objetos matemáticos particulares que no cambian bajo tipos específicos de transformación. El procedió a probar lo que llamó una de las leyes cardinales en la teoría de matrices: que la nulidad del producto de dos matrices no es menor que la nulidad de cualquier factor ni mayor que la suma de las nulidades de los factores.

EL PROCESO DE GRAM-SCHMIDT debe su nombre al matemático danés Jorge P. Gram (1850-1916), y al alemán Erhard Schmidt (1876-1959). Lo publicó primero Gram en 1883 en un artícluo titulado "Desarrollo de series usando el método de los mínimos cuadrados", y Schmidt lo publicó de nuevo en 1907, con una cuidadosa demostración, en un trabajo sobre ecuaciones integrales. De hecho Schmidt incluso se refirió al trabajo de Gram. Para Schmidt igual que para Gram, los vectores eran funciones continuas defindias en un intervalo [a,b] con el producto interno de dos funciones σ y Φ dado por la integral de su producto en ese intervalo. Sin embargo Schmidt fue más explícito que Gram al escribir el proceso con gran detalle y probar que el conjunto de funciones Φ_i derivado de su conjunto original σ_i, era de hecho, un conjunto ortonormal.

Schmidt, que estuvo en la Universidad de Berlín desde 1917 hasta su muerte, es mas conocido por su trabajo decisivo sobre los espacios de Hilbert, espacios de sucesiones de números complejos de cuadrado sumable. De hecho aplicó el proceso de Gram-Schmidt a conjuntos de vectores en esos espacios para ayudar a desarrollar condiciones necesarias y suficientes para que dichos conjuntos sean linealmente independientes.

UNA TECNICA MUY CERCANA A LA DE LOS MINIMOS CUADRADOS es la que desarrolló Roger Cotes (1682-1716), el genial matemático que editó la segunda edición de los **Principia de Isaac Newton**, en una obra que trataba errores en las observaciones astronómicas, escrita alrededor de 1715.

Sin embargo, fue Carl Gauss, a los 16 años el primero en formular el principio completo, mientras ajustaba aproximaciones relacionadas con la distribución de los números primos. Más tarde Gauss, afirmo que durante años había usado a menudo el método, por ejemplo en sus cálculos sobre las órbitas de asteroides. Gauss publicó el método en 1809 e hizo una exposición definitiva 14 años después.

Por otro lado, debemos a Adrien-Marie Legendre (1752-1833), fundador de la teoría de las funciones elípticas, la primer publicación del método de los mínimos cuadrados en un trabajo de 1806, sobre la determinación de las órbitas de cometas. Después de la publicación de Gauss en 1809, Legendre le escribió censurándolo por reclamar el método como propio. Todavía en 1827, Legendre seguía acusando a Gauss de apropiarse de los descubrimientos de otros. De hecho el problema radicaba en el fallo de Gauss de no publicar a tiempo sus descubrimientos; solamente los mencionaba cuando ya habían sido publicados por otros.

CAPITULO 7

VALORES Y VECTORES PROPIOS.

7.1 VALORES Y VECTORES PROPIOS.

Sea $T: V \dashrightarrow V$ una transformación lineal. En una gran variedad de aplicaciones resulta útil encontrar un vector $v \in V$ tal que Tv y v sean paralelos. Esto es, buscamos un vector v y un escalar α tal que:

$$Tv = \alpha v \tag{7.1}$$

si $v \neq 0$ y α satisface (7.1) entonces α se conoce como un valor propio o característico de T y v es un vector propio o característico de T correspondiente al valor propio α.

<u>DEFINICION 7.1</u> Si A es una matriz de nxn, entonces un escalar α es un valor propio de A si hay un vector v distinto de cero tal que

$$Av = \alpha v$$

el vector $v \neq 0$ se llama un vector propio de A correspondiente a un valor propio α

<u>Ejemplo 7.1</u>

Sea A tal que

$$A = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix}$$

entonces

$$A \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

de ésta froma $\alpha_1 = 1$ es un valor propio (característico) de A correspondiente al vector propio (característico) $v_1 = [2, 1]^t$, análogamente

$$A \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 10 & -18 \\ 6 & -11 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -6 \\ -4 \end{bmatrix} = -2 \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

por lo que $\alpha_2 = -2$ es un valor propio de A con su correspondiente vector propio $v_2 = [3, 2]^t$

Sea α un valor propio de A, entonces existe un vector no nulo $v = (x_1, \ldots, x_n)^t \neq 0$ tal que $Av = \alpha v = \alpha Iv$ o

$$(A - \alpha I)v = 0 \qquad (7.2)$$

si A es de orden nxn entonces la ecuación (7.2) es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero) el sistema tiene soluciones no triviales entonces $\det(A - \alpha I) = 0$

A esta ecuación se le denomina la <u>ecuación característica de A.</u>

Si $A = (a_{ij})$ entonces la ecuación anterior se puede escribir

$$\begin{bmatrix} a_{11} - \alpha & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \alpha & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & & & a_{nn} - \alpha \end{bmatrix} = 0$$

si desarrollamos el determinante obtenemos una expresión polinomial $p(\alpha)$ de grado n con coeficientes que incluyen a_{ij}, es decir

$$\det(A - \alpha I) = p(\alpha)$$

el polinomio $p(\alpha)$ es el polinomio característico de la matriz A. Los valores propios de A son precisamente las soluciones de la ecuación característica $p(\alpha) = 0$.

<u>Ejemplo 7.2</u>

Sea $A = I$, entonces para todo vector $v \epsilon R^n$, $Av = Iv = v$, de ésta manera I es el único valor característico de A y cada vector v es un vector característico de I.

<u>Ejemplo 7.3</u>

Encuentre los vectores propios de la matriz A:

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

<u>Solución:</u>

El polinomio característico de A es

$$\det(A - \alpha I) = \begin{vmatrix} 3 - \alpha & 2 \\ 2 & -\alpha \end{vmatrix} = \alpha^2 - 3\alpha - 4 = 0$$

$\alpha^2 - 3\alpha - 4 = (\alpha - 4)(\alpha + 1) = 0$ de donde $\alpha_1 = -1$ y $\alpha_2 = 4$ son valores propios de A

Ejemplo 7.4

Encuentre los valores propios de la matriz A

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

Solución:

$$\det(A - \alpha I) = \begin{vmatrix} 2-\alpha & 1 & 0 \\ -1 & -\alpha & 1 \\ 1 & 3 & 1-\alpha \end{vmatrix} \begin{array}{l} = -(\alpha - 2)(\alpha^2 - \alpha - 2) \\ = -(\alpha - 2)(\alpha - 2)(\alpha + 1) \end{array}$$

$-(\alpha-2)(\alpha-2)(\alpha+1) = 0$ de donde $\alpha_1 = -1$ y $\alpha_2 = \alpha_3 = 2$ son valores propios de A

En general si se tiene una matriz A de orden nxn

$$A = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & & & a_{nn} \end{vmatrix} = 0$$

entonces

$$p(\alpha) = \det(A - \alpha I) = \begin{vmatrix} a_{11} - \alpha & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} - \alpha & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & & & a_{nn} - \alpha \end{vmatrix} = 0$$

y $p(\alpha)$ se puede escribir como:

$$p(\alpha) = \alpha^n + b_{n-1}\,\alpha^{n-1} + \ldots + b_1\,\alpha + b_0 = 0 \qquad (7.3)$$

la ecuación (7.3) tiene n raíces, varias de las cuales pueden repetirse. Si $\alpha_1, \ldots, \alpha_m$ son las diferentes raíces de (73.) con multiplicidades $r_1, r_2, \ldots, r_m$ entonces (7.3) se puede factorizar para obtener:

$$p(\alpha) = (\alpha-\alpha_1)^{r1}\,(\alpha-\alpha_2)r^2 \ldots (\alpha-\,\alpha_m)^{rm} = 0$$

los números $r_1, r_2, \ldots, r_m$ se llaman <u>multiplicidades algebráicas</u> de los valores propios $\alpha_1, \ldots, \alpha_m$.

Ejemplo 7.5

Todo polinomio de grado n con coeficientes reales o complejos tiene n raices exactamente (contando multiplicidades), por ejemplo el polinomio $(\alpha \;-\; I)^5$ tiene cinco raices todas iguales al número uno. Puesto que todo valor característico de A es una raiz de la ecuación característica de A se tiene:

DEFINICION 7.2 Si contamos multiplicidades, cada matriz de nxn tiene exactamente n valores caracteristicos.

TEOREMA 7.1 Si A es una matriz de orden nxn entonces las siguientes afirmaciones son equivalentes:

a) α es un valor característico de A
b) El sistema (A - αI)v = 0 tiene soluciones no triviales.
c) Existe un vector v e n R^n v $\neq$ 0 tal que Av = αv

Si α es un valor característico de A entonces el espacio solución del sistema de ecuaciones (A - αI)v = 0 se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico se denominan los vectores característicos de A correspondientes a α.

INTERPRETACION GEOMETRICA

Si v es un vector característico de A correspondiente a α entonces Av = αv. Por lo tanto, la multiplicación por A mapea a v en un múltiplo escalar de si mismo por consiguiente dependiendo del valor de α, la multiplicación por A dilata a v, lo contrae o invierte su dirección, ver figura 7.1

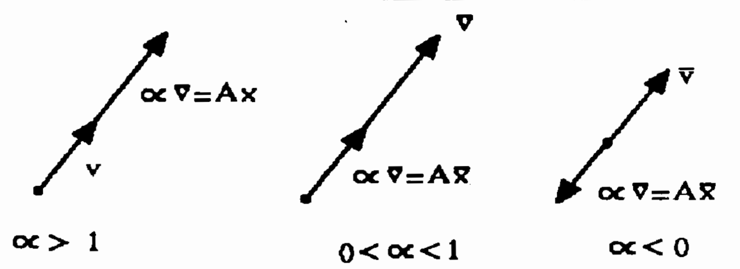


Figura 7.1

Ejemplo 7.6

Sea la matriz A como sigue:

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

encuentre los valores y los vectores propios

Solución

Sea

$$\det (A - \alpha I) = \begin{vmatrix} 4 - \alpha & 1 \\ 0 & 4 - \alpha \end{vmatrix} = (\alpha - 4)^2 = 0$$

$\alpha = 4$ es el valor característico de multiplicidad 2 de donde

$$(A - 4I)v = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} v_2 \\ 0 \end{bmatrix}$$

lo que implica que $v_1 = [1, 0]^t$ es un vector propio.

Ejemplo 7.7

Sea la matriz A como sigue:

$$A = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

encuentre los valores y vectores característicos de A.

Solución:

Sea

$$\det (A - \alpha I) = \begin{vmatrix} 4 - \alpha & 0 \\ 0 & 4 - \alpha \end{vmatrix} = (\alpha - 4)^2 = 0$$

$\alpha = 4$ es el valor característico de multiplicidad 2 de donde $Av = 4v$ lo que implica que $v_1 = [1, 0]$ y $v_2 = [0, 1]$ generan el espacio característico de A.

DEFINICION 7.3 Sea α un valor característico de A. El subespacio E_α se denomina el espacio característico de A correspondiente al valor característico α

TEOREMA 7.2 Sea α un valor característico de la matriz A de orden nxn y sea $E_\alpha = \{v : Av = \alpha v\}$ entonces E_α es un subespacio de R^n (en general de E^n)

demostración

Si $Av = \alpha v$ entonces $(A - \alpha I)\, v = 0$ de donde E_α es el espacio nulo de la matriz $A - \alpha I$, la cual es un subespacio de R^n

TEOREMA 7.3 Sea A una matriz de nxn y sea $\alpha_1, \alpha_2, \ldots, \alpha_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, \ldots, v_m$. Entonces v_i son linealmente independientes o bien, los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. (la demostración se hace por inducción).

Ejemplo 7.8

Sea la matriz A como sigue

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

encuentre los valores y vectores propios asi como el espacio característico correspondiente.

Solución:

Sea $\det(A - \alpha I) = -\alpha^3 + 6\alpha^2 + 15\alpha + 8 = -(\alpha + 1)^2 (\alpha - 8) = 0$

Lo que implica $\alpha_1 = 8$ y $\alpha_2 = -1$ con multiplicidad algebráica 2, para $\alpha_1 = 8$ se obtiene $v_1 = [2, 1, 2]^t$ y $E_8 = \{[2, 1, 2]\}$ para $\alpha_2 = -1$ se obtiene $v_2 = [1, -2, 0]^t$ y $v_3 = [0, -2, 1]^t$ con

$$E_{-1} = \left\{ \begin{bmatrix} 1 \\ -2 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ -2 \\ 1 \end{bmatrix} \right\}$$

En los ejemplos que hemos visto encontramos un valor característico con una multiplicidad algebráica de dos o mas. Sin embargo, el número de vectores característicos linealmente independientes no necesariamente es igual a la multiplicidad algebráica del valor característico (como en el ejemplo 7.6) de aqui que:

DEFINICION 7.4 Sea α un valor cracterístico de A entonces la multiplicidad geométrica de α es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz $A - \alpha I$) esto es:

$$\text{muliplicidad geométrica de } \alpha = \dim E_\alpha$$

Si A es una matriz de 2x2 y α un valor característico con multiplicidad algebráica 2 entonces la multiplicidad geométrica de α es menor o igual a 2 puesto que puede haber al menos dos vectores l.i. en un espacio de dos dimensiones.

Si A es de 3x3 con dos valores característicos α_1 y α_2 de multiplicidad algebráica 1 y 2 respectivamente entonces la multiplicidad geométrica de α_2 es menor o igual a 2.

TEOREMA 7.4 Sea α un valor característico de A entonces:

$$\text{multiplicidad geometrica de } \alpha \leq \text{multiplicidad algebraica de } \alpha$$

7.2 DIAGONALIZACION

TEOREMA 7.5 Sea A una matriz de orden nxn, entonces A tiene n vectores caracteristicos linealmente independientes si y solo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos l.i. si todos sus valores caracteristicos son distintos (pues si no, podrían ser menores que n)

Como ya se ha visto en varios ejemplos, algebráicamente y desde el punto de vista de los cálculos, el problema de los valores propios es mucho más difícil que Ax = b. Sin embargo se pueden usar algunos resultados para que sea mas fácil trabajar con ellos.

DEFINICION 7.5 La suma de los n valores propios es igual a la suma de las n entradas de la diagonal de A:

$$\alpha_1 +...+ \alpha_n = a_{11} +...+a_{nn}$$

Esta suma es la traza de A. Además, el producto de los n valores propios es igual al determinante de A.

Ejemplo 7.9

Sea la matriz simétrica

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

con valores propios $\alpha_1 = 0$, $\alpha_2 = 1$ y $\alpha_3 = 3$ entonces

$$0 + 1 + 3 = 1 + 2 + 1 = 4 \quad \text{y} \quad \det A = 0$$

NOTA: No deben confundirse los valores propios de una matriz y sus entradas diagonales. Normalmente son completamente diferentes. Sin embargo se tiene la siguiente definición:

DEFINICION 7.6 Si la matriz A es triangular (puede ser superior y en particular diagonal) entonces los valores propios $\alpha_1,..,\alpha_n$ son exactamente los mismos que las entradas de la diagonal $a_{11},...,a_{nn}$.

Ejemplo 7.10

Sea la matriz A como sigue:

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

su polinomio característico es:

$$\det \begin{bmatrix} 1-\alpha & 1/4 & 0 \\ 0 & 3/4-\alpha & 0 \\ 0 & 0 & 1/2-\alpha \end{bmatrix} = (1-\alpha)(3/4-\alpha)(1/2-\alpha)$$

el determinante es el producto de las entradas diagonales. Obviamente las raices son $\alpha_1 = 1$, $\alpha_2 = 3/4$ y $\alpha_3 = 1/2$, los valores propios ya estaban colocados a lo largo de la diagonal principal.

Este ejemplo,en el cual se pueden encontrar los valores propios por inspección señala lo mas importante de éste tema:**transformar a una matriz A en una matriz diagonal o triangular sin cambiar sus valores propios.**(La factorización LU no no sirve para este caso ya que los valores propios de U no son los valores propios de A)

Otra situación en que los cálculos son fáciles es: si ya tenemos los valores y vectores propios de una matriz A entonces los valores propios de A^2 son exactamente $\alpha_1^2, \ldots, \alpha_n^2$ y cada vector propio de A es también un vector propio de A^2

<u>TEOREMA 7.6</u> Supongamos que la matriz A de nxn tiene n vectores propios linealmente independientes. Entonces si se eligen éstos vectores como las columnas de una matriz S, se sigue que $S^{-1}AS$ es una matriz diagonal Λ con los valores propios de A en su diagonal.

$$S^{-1}AS = \Lambda = \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix}$$

<u>demostración</u>

Colocamos los vectores propios x_i en las columnas de S y se calcula el producto AS, una columna a la vez:

$$AS = A \begin{bmatrix} \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ \alpha_1 x_1 & \alpha_2 x_2 & \cdots & \alpha_n x_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

entonces

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ \alpha_1 x_1 & \alpha_2 x_2 & \cdots & \alpha_n x_n \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_n \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \alpha_1 & & & \\ & \alpha_2 & & \\ & & \ddots & \\ & & & \alpha_n \end{bmatrix} = S\Lambda$$

$$AS = S\Lambda \quad \text{o} \quad S^{-1}AS = \Lambda \quad \text{o} \quad A = S\Lambda S^{-1}$$

la matriz S es invertible ya que supusimos que sus columnas (los vectores propios son l.i.)

Observación 1

Si la matriz A no tiene valores propios repetidos (los números $\alpha_1, \ldots, \alpha_n$ son distintos) entonces los n vectores propios son independientes y por lo tanto cualquier matriz con valores propios distintos puede diagonalizarse.

Observación 2

La matriz diagonalizada S no es única, ya que un vector propios x puede multiplicarse por una constante y seguir siendo vector propio, por lo tanto se puedne multiplicar las columnas de S por cualquier constante distinta de cero y producir una nueva diagonalización S.

Observación 3

La ecuación AS = SΛ es válida si las columnas de S son los vectores propios de A y no de otra manera. Supongamos que la primer columna de S es algún vector y, entonces la primer columna de SA es α_1y. Si esto va a corresponder con la primer columna de AS que debido a la multiplicación de matrices es Ay, entonces y debe ser un vector propio Ay = α_1 y

Observación 4

No todas las matrices poseen n vectores propios l.i. y por lo tanto no todas las matrices son diagonalizables.

Ejemplo 7.11

Sea la matriz A como sigue:

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

sus valores propios son $\alpha_1 = \alpha_2 = 0$ ya que es triangular

$$\det (A - \alpha I) = \begin{vmatrix} -\alpha & 1 \\ 0 & -\alpha \end{vmatrix} = \alpha^2$$

si x es un vector propio entonces

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

o

$$x = \begin{bmatrix} x_1 \\ 0 \end{bmatrix}$$

aunque $\alpha = 0$ es un valor propio de multiplicidad algebráica 2 sólo tiene un espacio unidimensional de vectores propios, la multiplicidad geométrica de éste valor propio es 1 y no podemos construir S.

Como $\alpha_1 = \alpha_2 = 0$ A tendría que ser la matriz de ceros. Pero si $S^{-1}AS = 0$ entonces $SS^{-1}AS = 0$ lo que implica $ASS^{-1} = 0$ y de aqui se deduce que $A = 0$ lo que resulta imposible. Por lo tanto no existe S tal que $S^{-1}AS = A$

Algunas matrices con valores propios repetidos pueden diagonalizarse. Si se repite un valore propio α, m veces, entonces para que se pueda diagonalizar A debe haber m vectores propios correspondientes.

<u>Ejemplo 7.12</u>

Diagonalice la siguiente matriz A

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

<u>Solución:</u>

Los valores propios de A son $\alpha_1 = 1$ y $\alpha_2 = 5$.

Para $\alpha_2 = 5$ se tienen los vectores propios

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

y para $\alpha_1 = 1$ se tiene

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

de donde $\{v_1, v_2, v_3\}$ es un conjunto de vectores l.i. entonces

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

diagonaliza a A

$$S^{-1}AS = \begin{bmatrix} -1/2 & 1/2 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 3 & -2 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix} \begin{bmatrix} -1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

el orden de las columnas de S no altera los valores propios sólo su orden.

TEOREMA 7.7 Una matriz A de nxn es diagonalizable si y sólo si la multiplicidad algebráica de cada valor propio es igual a su multiplicidad geométrica.

7.3 PROPIEDADES DE VALORES Y VECTORES PROPIOS.

Sea la matriz A de orden nxn

1. Si α es un valor propio de A con v como vector propio correspondiente, entonces Λ^k es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.

2. Si α es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/\alpha$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.

3. Si α es un valor propio de A, entonces el conjunto E_α es un subespacio de R^n.

COROLARIO 7.1 Sean A y S matrices de orden nxn, entonces $A^k = A \Lambda^k S^{-1}$ para cada entero positivo k.

demostracion

De $S^{-1}AS = \Lambda$ obtenemos $A = S \Lambda S^{-1}$ asi

$$A^k = \underbrace{(S\Lambda S^{-1})(S\Lambda S^{-1}) \ \ldots \ (S\Lambda S^{-1})}_{k \text{ veces}}$$

como los términos adyacentes $S^{-1}S$ se cancelan queda

$$A^k = S\Lambda^k S^{-1}$$

Matrices Simetricas y Diagonalizacion Ortogonal

TEOREMA 7.8 Sea A una matriz simétrica real de nxn. Entonces los valores propios de A son reales. Para cualquier matriz simétrica la multiplicidad algebraica de cada valor propio es igual a su multiplicidad geometrica, de modo que toda matriz simétrica es diagonalizable.

Ya vimos que vectores caracteristicos correspondientes a valores caracteristicos diferentes son l.i. Para matrices simétricas el resultado es mas fuerte los vectores caracteristicos de una matriz simétrica correspondientes a valores diferentes son ortogonales.

TEOREMA 7.9 Sea A una matriz simetrica real de nxn. Si α_1 y α_2 son valores caracteristicos distintos con correspondientes vectores caracteristicos reales v_1 y v_2 entonces v_1 y v_2 son ortogonales.

demostración Calculamos

$$Av_1\ v_2 = \alpha_1 v_1\ v_2 = \alpha_1(v_1\ v_2)$$

y

$$Av_1\ v_2 = v_1\ A^t v_2 = v_1\ Av_2 = v_1\ (\alpha_2 v_2) = \alpha_2(v_1 v_2)$$

entonces $\alpha_1(v_1\ v_2) = \alpha_2(v_1\ v_2)$ como $\alpha_1 \neq \alpha_2$

y de aqui $(v_1\ v_2) = 0$ l.q.q.d.

TEOREMA 7.10 Sea A una matriz simétrica real de nxn. Entonces A tiene n vectores caracteristicos reales ortonormales.

Este teorema nos dice que si A es simétrica entonces R^n tiene una base $B = \{u_1, u_2, \ldots, u_n\}$ de vectores caracteristicos ortonormales de A. Sea Q la matriz cuyas columnas son $u_1, u_2, \ldots, u_n$ entonces Q es una matriz ortogonal.

DEFINICION 7.7 Una matriz A de nxn se dice diagonalizable ortogonalmente si existe una matriz ortogonal Q tal que

$$Q^t A Q = \Lambda$$

donde $\Lambda = \text{diag.}\ (\alpha_1, \ldots, \alpha_n)$ los valores caracteristicos de A. como Q es ortogonal $Q^t = Q^{-1}$ lo que implica $Q^{-1}AQ = \Lambda$

TEOREMA 7.11 Sea A una matriz real de nxn. Entonces A es diagonalizable ortogonalmente si y solo si A es simétrica.

Antes de ver ejemplos, enunciaremos los 3 pasos que sirven para encontrar la matriz ortogonal Q que diagonaliza a la matriz simétrica A.

1. Encuentre una base para cada espacio caracteristico de A.
2. Encuentre una base ortonormal para cada espacio caracteristico de A usando el proceso de Gram-Schmidt.
3. Escriba Q como la matriz cuyas columnas son los vectores caracteristicos ortonormales obtenidos en el paso 2

Ejemplo 7.13

Encuentre una matriz ortogonal Q que diagonalice a la matriz A dada como sigue:

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

Solución:

1. det (A- I) = $\begin{vmatrix} 4- & 2 & 2 \\ 2 & 4- & 2 \\ 2 & 2 & 4- \end{vmatrix} = (\alpha-2)^2(\alpha-8) = 0$

por lo que $\alpha_1 = 2$ y $\alpha_2 = 8$

Para $\alpha_1 = 2$ se tiene

$$v_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \qquad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

forman una base para E_2. Aplicando Gram-Schmidt se tiene.

$$v'_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \qquad v'_2 = \begin{bmatrix} -1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

el espacio caracteristico correspondiente a $\alpha_2=8$ tiene como base

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

aplicando Gram-Schimdt a v_3 se tiene

$$v'_3 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

$$Q = \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{3} \end{bmatrix}$$

diagonaliza ortogonalmente a A.

$$A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \qquad S = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

7.4 ECUACIONES EN DIFERENCIAS.

Las ecuaciones en diferencias se dirigen a un numero finito de pasos finitos, mientras que una ecuacion diferencial lleva un numero infinito de pasos infinitesimales (sin embargo ambas teorias son paralelas).

Ejemplo 7.14

Supongamos que se invierten \$100 por 5 años a un interés de 6% Si se compone una vez al año, entonces el capital se multiplica por 1.06 y $P_{k+1} = 1.06P_k$ donde ésta es una ecuación en diferencias con un lapso de tiempo de un año. Relacione el capital despues de k+1 años con el capital del año anterior y es fácil de resolver, después de 5 años el capital original B=1000 se ha multiplicado 5 veces y

$$P_5=(1.06)^5 \quad P_0 = (1.06)_5 \; 1000 = \$ \; 1338.$$

Si el lapso de tiempo se reduce a un mes

$$P_{k+1} = (1 + 0.6/12)P_k.$$

y despues de 5 años o 60 meses.

$$P_{60} = (1 + 0.06/12)^{60} \; p_0 = (1.005)^{60} \; 1000 = 1349$$

si se compone diariamente el interés se tiene:

$$(1+0.06/365)^{5 \cdot 365} \; 1000 = \$ \; 1349.83.$$

Si se pretende componer continuamente el interés, se añade a cada instante, entonces se tiene un proceso de limite:

$$(1+0.06/n)^{5n} \; 1000 \; \text{--->} \quad e^{0.30} \; 1000 = \$ \; 1349.87$$

o se puede cambiar a una ecuacion diferencial que será el límite de la ecuación en diferencias $P_{k+1} = (1+0.06 \; \Delta t)P_k$, lo que implica

$$P_{k+1} = P_k + 0.06 \; \Delta t \; P_k$$
$$P_{k+1} - P_k = 0.06 \; \Delta t \; P_k$$

de donde

$$\frac{P_{k+1} - P_k}{\Delta t} = 0.06 \; P_k \quad \text{--->} \quad \frac{dp}{dt} = 0.06p.$$

cuya solucion es $p(t) = e^{0.06t} \; p_0$ para 5 anos se tiene \$ 1349.07.

Este ejemplo incluyó tanto ecuaciones en diferencias como ecuaciones diferenciales, con una tendiendo a la otra a medida que desapareció el lapso de tiempo. Pero hay muchas ecuaciones en diferencias por derecho propio y aqui nuestro siguiente ejemplo:

Ejemplo 7.15 (Los conejos de Fibonacci)

Suponga que las parejas de conejos recien nacidos no tienen descendencia durante su primer mes de vida, pero que a partir de ahi cada pareja produce otra nueva pareja cada mes. Comenzando con $F_1 = 1$ que es la pareja recien nacida en el primer mes, hallar el número F_k de parejas en el k-ésimo mes, suponiendo que no muere ningun conejo.

En el k-ésimo mes el número de parejas de conejos es

F_k = (numero de parejas mas el mes anterior) +
+ (numero de parejas recien nacidas en el k-ésimo mes)

Como nuestros conejos no tienen descendencia durante el primer mes de vida, vemos que el número de parejas recien nacidos en el k-ésimo mes es el número F_{k-2} de parejas vivas dos meces antes. Asi, podemos escribir la ecuacion anterior como

$$F_k = F_{k-1} + F_{k-2}$$

esta es una ecuacion en diferencias, conocida como Relacion de Fibonacci.

Es conveniente hacer $F_0 = 0$ para denotar 0 parejas en el mes 0 antes de la llegada de la primera pareja recien nacida. Asi la sucesión

$$F_0, F_1, F_2, \ldots, F_k.$$

para el número de parejas de conejos se convierte en la sucesión de Fibonacci

0,1,1,2,3,5,8,13,21,34,...

donde cada término, comenzando con $F_2 = 0+1=1$ es la suma de los dos términos anteriores. Para cualquier k particular, podemos calcular F_k prolongando lo suficiente la sucesión. Sin embargo esto puede ser una tarea tediosa aunque solo queramos calcular F_{30}.

Una forma de hacerlo, consiste en resolver la ecuación en diferencias $F_{k+2} = F_{k+1} + F_k$ y como primer paso podemos reducirla a una ecuacion $u_{k+1}=Au_k$, precisamente como el interés compuesto $P_{k+1} = 1.06\ P_k$ excepto que ahora la incógnita es un vector y el multiplicador A tiene que ser una matriz

si $u_k = \begin{vmatrix} F_{k+1} \\ F_k \end{vmatrix}$

entonces

$$F_{k+2} = F_{k+1} + F_k$$
$$F_{k+1} = F_{k+1}$$

se transforma en

$$u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$$

Formalmente la ecuación e diferencias $u_{k+1} = Au_k$ es fácil de resolver, como cada paso conlleva una multiplicación por A, la solucion u_k esta relacionada con el valor inicial u_0 mediante $u_k = A^k u_0$. Entonces el problema se reduce a encontrar A^k de un modo rápido y esto se puede resolver a través de los valores y vectores característicos de A.

<u>TEOREMA 7.12</u> Si puede diagonalizarse A, $A=S\Lambda S^{-1}$ entonces

$$u_k = A^k u_0 = S\Lambda^k S^{-1} u_0$$

o lo que es lo mismo

$$u_k = \begin{bmatrix} \vdots & & \vdots \\ x_1 & \cdots & x_n \\ \vdots & & \vdots \end{bmatrix} \begin{bmatrix} \Lambda_1^k & & \\ & \ddots & \\ & & \Lambda_n^k \end{bmatrix} S^{-1} u_0$$

donde $S^{-1}u_0$ es un vector columna de R^n, ya que S^{-1} tiene rango n de modo que sus columnas forman una base de R^n de aqui que

$$S^{-1}u_0 = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}$$

entonces

$$u_k = C_1 \Lambda_1^k x_1 + \dots + C_n \Lambda_1^k x_n$$

la solución general es una combinación de las soluciones especiales $\alpha^k_i x_i$ y los coeficientes c_i correspondientes con la condición inicial u_0 son :

$$C_1 \alpha^0_1 x_1 + \dots + C_n \alpha^0_n x_n = u_0 \quad \text{o} \quad Sc = u_0$$

en el caso específico de la ecuación de Fibonacci, el primer paso es diagonalizar la matriz A:

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

$$\det (A - \alpha I) = \alpha^2 - \alpha - 1$$

lo que implica $\alpha_1 = \dfrac{1+\sqrt{5}}{2}$ $\alpha_2 = \dfrac{1-\sqrt{5}}{2}$

$$A = S\Lambda S^{-1} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 & \\ & \alpha_2 \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

Con $F_0 = 0$ y $F_1 = 1$ se tiene $u_0 =$ [1, 0] entonces

$$\begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix} = u_k = A^k u_0 = S\ {}^k S^{-1} u_0 = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 & \\ & \alpha^k{}_2 \end{bmatrix} \begin{bmatrix} 1 & -\alpha_2 \\ -1 & \alpha_1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 & \\ & \alpha^k{}_2 \end{bmatrix} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2} = \begin{bmatrix} \alpha_1 & \alpha_2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \alpha^k{}_1 \\ -\alpha^k{}_2 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

$$= \begin{bmatrix} \alpha^{k+1}{}_1 & -\alpha^{k+1}{}_2 \\ \alpha^k{}_1 & \alpha^k{}_2 \end{bmatrix} \frac{1}{\alpha_1 - \alpha_2}$$

entonces el número de Fibonacci F_k es la segunda componente de este producto:

$$F_k = \frac{\alpha^k{}_1}{\alpha_1 - \alpha_2} - \frac{\alpha^k{}_2}{\alpha_1 - \alpha_2} = 1/\sqrt{5} \quad [(1+\sqrt{5}/2)^k - (1-\sqrt{5}/2)^k]$$

como el segundo termino $(1-\sqrt{5}/2]^k/\sqrt{5}$ siempre es menor que 1/2 solo debe moverse el primer termino al entero mas cercano, de modo que

$$F_k \approx 1/\sqrt{5}\ (1+\sqrt{5}/2)^k$$

o

F_k = entero mas cercano a $1/\sqrt{5}\ (1+\sqrt{5}/2)^k$ para toda k

por ejemplo si k=1000 entonces

F_{1000} = entero mas cercano a $1/\sqrt{5}\ (1+\sqrt{5}/2)^{1000}$

pero este número como se puede ver es muy grande y aún mas grande es F_{1001}, la razon F_{1000}/F_{1001} debe estar muy cerca de la cantidad $1+\sqrt{5}/2 \approx 1.1618$ que los griegos llamaron la "sección dorada"

en otras palabras $\alpha^k{}_2$ se vuelve insignificante comparado con $\alpha^k{}_1$ y la razón F_{k+1}/F_k se acerca a $\alpha_1^{k+1}/\alpha_1^k = \alpha_1$.

7.5 CADENAS DE MARKOV

Las Cadenas de Markov se utilizaron inicialmente para analizar procesos en física y meteorología, sin embargo las aplicaciones mas recientes incluyen el análisis de los movimientos de precios de bienes, el mantenimiento de maquinaria, el comportamiento de los animales en el laboratorio, la seleccion de productos por el consumidor, la longitud de las colas en un supermercado o un aeropuerto, en el manejo de inventarios en cuanto a nivel y variedad, y en la administración de plantas.

Para definir una cadena de Markov, considere un experimento con un espacio muestra finito $S=\{E_1,E_2,...,E_n\}$, y una secuencia (o cadena) de experimentos realizados. Se dice que el experimento esta en el estado E_i en el intento m-ésimo si E_i es el resultado del m-ésimo ensayo del experimento.

<u>DEFINICION 7.8</u> Una secuencia de intentos de m experimentos es una cadena de Markov si

a) El resultado del intento m-ésimo depende solo del resultado del intento (m-1)-ésimo y no de los resultados en los intentos anteriores, y

b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

Por ejemplo si el clima de hoy depende solamente del clima de ayer, entonces la observacion y predicción del clima es un problema de cadenas de Markov. Si la probabilidad de escoger una marca particular de coche la proxima vez que piense comprar uno depende unicamente del auto que ya se tiene, entonces el problema del patrón de ventas de automóviles es un problema que se puede resolver por cadenas de Markov.

Por otra parte, si el clima de hoy es determinado por el clima de varios dias anteriores, entonces no se trata de una cadena de Markov.

Una cadena de Markov se caracteriza por las probabilidades de que el sistema pase de un estado a otro en intentos sucesivos.

<u>Matriz de Transición de una Cadena de Markov.</u>

La matriz de transición de una cadena de Markov es la matriz de nxn de probabilidades $T=(p_{ij})$ cuya componente ij-esima p_{ij} es la probabilidad de que el sistema pase del estado E_i al estado E_j en intentos sucesivos del experimento por lo cual todos los componentes son no negativos y la suma de ellos por columna es 1.

<u>Ejemplo 7.16</u>

Cada año 1/10 de la gente que vive fuera de California se muda dentro y 2/10 de la gente que vive dentro se muda fuera.

Esto sugiere una ecuacion en diferencias. Comenzamos con y_0 gente que vive fuera y z_0 que vive dentro y al final del primer año hay

$$y_1 = 0.9\ y_0 + 0.2\ z_0$$
$$z_1 = 0.1\ y_0 + 0.8\ z_0$$

o bien

$$\begin{bmatrix} y_1 \\ z_1 \end{bmatrix} = \underbrace{\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}}_{T} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$

La matriz T es una matriz de transición ya que

$$0.9+0.1=1 \quad 0.2+0.8=1$$

y son todos no negativos. Por lo tanto las potencias T^k son no negativas.

Resolvemos esta ecuación en diferencias usando S $^kS^{-1}u_0$, posteriormente vemos si la población tiende a un estado estacionario. Esto es, que después de un tiempo las probabilidades de que el sistema se encuentre en cada uno de sus posibles estados son invariables con el tiempo. Para comenzar los cálculos, T tiene que diagonalizarse.

$$T = \begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix}$$

$$\det(T-\alpha I) = \alpha_2 - 1.7\ \alpha + 0.7$$

$$\alpha_1 = 1 \quad y \quad \alpha_2 = 0.7$$

$$A = S\Lambda S^{-1} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix} \begin{bmatrix} 1 & \\ & 0.7 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix}$$

podemos encontrar A^k y la distribución después de k años:

$$\begin{bmatrix} y_k \\ z_k \end{bmatrix} = A^k \begin{bmatrix} y_0 \\ z_0 \end{bmatrix} \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & -1/3 \end{bmatrix} \begin{bmatrix} 1^k & \\ & 0.7^k \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} y_0 \\ z_0 \end{bmatrix}$$
$$= y_0 z_0 \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} + (y_0 - 2z_0)\ (0.7)^k \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix}$$

Esta es la solución que queríamos, a largo plazo el factor $(0.7)^k$ se vuelve muy pequeño y la solución tiende a un estado limite.

$$\begin{bmatrix} y_\infty \\ z_\infty \end{bmatrix} = (y_0 + z_0) \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

La población total es todavía y_0+z_0, la misma que al principio pero en el límite 2/3 de ésta población está fuera de California y 1/3 está dentro. Esto es cierto independientemente de las distribuciones inicial. Si se comienza el ano con 2/3 fuera y 1/3 dentro, entonces termina de la misma manera:

$$\begin{bmatrix} 0.9 & 0.2 \\ 0.1 & 0.8 \end{bmatrix} \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix}$$

o $$Au_\infty = u_\infty$$

El estado estacionario es el vector propio de A correspondiente a $\alpha= 1$. La multiplicacion por A que nos lleva de un lapso de tiempo a otro, no altera u_∞.

En el ejemplo de California, si el individuo está fuera se mudará adentro con una probabilidad de 1/10, si está adentro entonces se mudará afuera con probabilidad 2/10. Su movimiento se vuelve un proceso aleatorio y la matriz A que lo gobierna es una matriz de transición. No sabemos donde está, pero cada año las componentes de $u_k = A^k u_0$ especifican la probabilidad de que esté fuera del estado y la probabilidad de que esté dentro. Estas probabilidades suman 1 el individuo tiene que estar en algún lado y nunca son negativas.

Por que $\alpha= 1$ es siempre un valor propio y porque su vector propio es el estado estacionario.

Como se vió en el caso de los conejos de Fibonacci se tiene que el vector propio o de información después del k-ésimo estado es igual a $A^k x_1$ para un vector inicial x_1 y una matriz diagonalizable A. Como es evidente un aspecto importante es si alguno de los valores propios de A tiene magnitud mayor que 1. Cuando éste es el caso, la magnitud de las componentes del vector de información pueden crecer exporencialmente como es el caso de la sucesión de Fibonacci donde $|\alpha_1|>1$ y se dice que es inestable.

Por otro lado, si todos los valores propios tienen magnitud menor que 1, las componentes del vector de información deben tender a cero a medida que k crece. Si las probabilidades de Markov decrecieran a cero, entonces la ecuacion sería estable, cosa que no sucede porque en cada estado deben sumar 1.

Dada cualquier ecuacion en diferencias $U_{k+1} = Au_k$ donde queremos estudiar su comportamiento cuando $k \dashrightarrow \infty$. Suponiendo que A puede diagonalizarse, la solución U_k será una combinación de soluciones para

$$U_k = S\Lambda^k S^{-1} U_0 = C_1 \alpha^k{}_1 x_1 + \ldots + C_n \alpha^k{}_n x_n$$

DEFINICION 7.9 La ecuacion en diferencias $U_{k+1} = Au_k$ es estable y $U_k \rightarrow 0$. Si todos los valores propios satisfacen $|\alpha_i| < 1$, es neutralmente estable y u_k está acotado si todo $|\alpha_i| \leq 1$ y es inestable (u_k no está acotado) cuando al menos un valor propio de A tiene $|\alpha_i| > 1$.

7.6 FORMAS CUADRATICAS

Las formas cuadráticas son una aplicación inmediata de los valores propios, el teorema del eje principal asegura que toda forma cuadratica se puede diagonalizar, este resultado tiene importantes aplicaciones a vibración de cuerpos elásticos, mecánica cuántica y circuitos eléctricos.

DEFINICION 7.10 Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0 \qquad (7.4)$$

donde a,b,c,d,e,f son números reales y al menos uno de ellos es diferente de cero, es una ecuación cuadrática en x,y. La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

EJEMPLO 7.17

En la ecuación cuadrática

$$3x^2 + 5xy - 7y^2 + 2x + 7 = 0$$

las constantes son:

$$a = 3 \quad b=5/2 \quad c=-7 \quad d=2 \quad e=0 \quad f=7.$$

EJEMPLO 7.18

ecuación cuadrática	forma cuadrática asociada.
$3x^2 + 5xy - 7y^2 + 2x + 7 = d$	$3x^2 + 5xy - 7y^2$
$4x^2 - 5y^2 + 8y + 9 = d$	$4x^2 - 5y^2$
$xy + y = d$	xy

las gráficas de las ecuaciones cuadráticas en x,y se llaman cónicas o secciones cónicas. Las cónicas mas importantes son la elipse, la circunferencia, la hipérbola y la parábola.

Una forma cuadrática con una variable x es un polinomio $f(x)= ax^2$ con $a \neq 0$. Una forma cuadrática en 2 variables x,y es un polinomio $f(x) = ax^2 + bxy + y^2$ donde a,b o c es $\neq 0$. El término cuadrático significa de grado 2, el término forma significa homogénea, es decir, cada sumando tiene un producto del mismo número de variables, a saber 2.

Se dice que una cónica tiene una posición normal relativa a los ejes de coordenadas si no se encuentra trasladada ni rotada.

EJEMPLO 7.19

La ecuación

$$\frac{x^2}{4} + \frac{y^2}{9} = 1$$

es de la forma

$$\frac{x^2}{k^2} + \frac{y^2}{l^2} = 1 \text{ con } k=2, \quad l=3$$

por lo tanto, su gráfica es una elipse en posición normal que intersecta al eje x en (-2,0), (2,0) y al eje y en (0,-3) y (0,3)

La ecuación $x^2 - 8y^2 = -16$ se puede expresar como $y^2/2 - x^2/16 = 1$ que es de la forma $y^2/k^2 - x^2/l^2 = 1$ con $k = \sqrt{2}$ $l=4$ por lo tanto su gráfica es una hipérbola en posición normal que corta al eje y en $(0, -\sqrt{2})$ y $(0, \sqrt{2})$.

La ecuación $5x^2 + 2y = 0$ se puede reescribir como $x^2 = (-2/5)$ y que es de la forma $x^2 = ky$ con $k=-2/5$ como $k < 0$ su gráfica es una parábola en posición normal que se abre hacia abajo.

Ninguna cónica en posición normal tiene términos cruzados (xy) en su ecuación. Si estos aparecen significa que la cónica se giró con respecto a su posición normal. Tampoco se tienen cónicas en posición normal que tengan simultáneamente un término en x^2 y en x o en y^2 y en y. Si tales términos aparecen significa que la cónica se trasladó de su posición normal, ver figura 7.2.

EJEMPLO 7.20

Dada la ecuación cuadrática.

$$2x^2 + y^2 -12x -4y + 18 = 0$$

tiene términos en x^2, x, y^2, y pero no en xy entonces su gráfica es una cónica que se trasladó pero no se giró: Agrupamos términos para obtener:

$$(2x^2 - 12x) + (y^2 - 4y) + 18 = 0$$

$$2(x^2 - 6x) + (y^2 - 4y) = -18$$

completando cuadrados (donde para completar el cuadrado de una expresión de la forma x^2+px se suma y se resta $(P/2)^2$

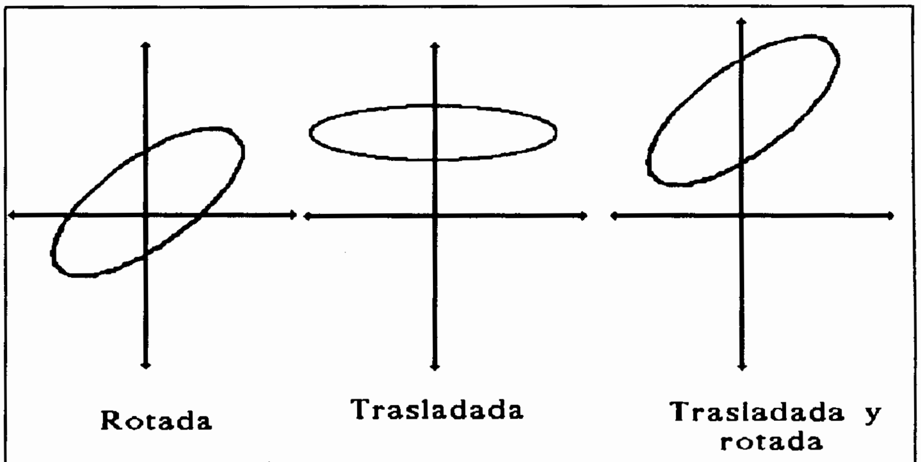


Figura 7.2

$$x^2 + px = x^2 + px + (p/2)^2 - (p/2)^2 = (x+p/2)^2 - (p/2)^2$$

se tiene entonces

$$2(x^2 - 6x + 9) + (y^2 - 4y + 4) = -18 + 18 + 4$$

$$2(x - 3)^2 + (y - 2)^2 = 4$$

Si se trasladan los ejes de coordenadas mediante

$$x' = x-3 \qquad y' = y-2$$

se tiene $\quad 2x'^2 + y'^2 = 4$

entonces

$$\frac{x'^2}{2} + \frac{y'^2}{4} = 1$$

que es una elipse en posición normal con respecto a x'y' (figura 7.3)

La ecuacion (7.4)

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

se puede expresar en forma matricial como

$$[x \quad y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + [d \quad e] \begin{bmatrix} x \\ y \end{bmatrix} + f = 0$$

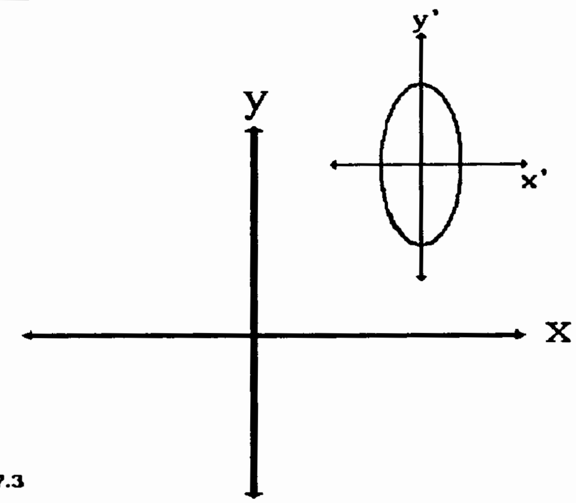


Figura 7.3

o sea

$$x^tAx + kx + f = 0 \qquad (7.5)$$

donde

$$x = \begin{bmatrix} x \\ y \end{bmatrix} \qquad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix} \qquad k = [d, \ \ e]$$

con esta notación, la forma cuadrática asociada a (7.5) es

$$x^tAx$$

La matriz simétrica A se llama la matriz de la forma cuadrática x^tAx

Ejemplo 7.21

La matriz de la forma cuadrática

$$3x^2 + 5xy - 7y^2$$

es

$$\begin{bmatrix} 3 & 5/2 \\ 5/2 & -7 \end{bmatrix}$$

Considere que la ecuación de una cónica C es

$$x^tAx + kx + f = 0 \qquad (7.6)$$

a continuación se muestra que siempre es posible rotar los ejes de coordenadas x-y de tal manera que la ecuacion de la cónica con respecto al sistema de coordenadas x'-y' no contenga un término en xy

<u>Etapa 1</u>

Encuentre la matriz

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

que diagonalice ortogonalmente a A.

<u>Etapa 2</u> Si es necesario, intercambiar las columnas de P de tal suerte que det(P) = 1. Esto garantiza que la transformación ortogonal de coordenadas (7.7) es una rotación

$$x = Px' \qquad (7.7)$$

<u>Etapa 3</u> Para obtener la ecuación de C con respecto al sistema x'- y' se sustituye (7.6) en (7.7)

$$(Px')^t \;\; A(Px') + k(Px') + f = 0$$

o

$$x^t \; (P^tAP) \; x' + (kP)x' + f = 0 \qquad (7.8)$$

como P diagonaliza ortogonalmente a A.

$$P^tAP = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix}$$

donde α_1 y α_2 son valores propios de A, por lo cual (7.8) se puede expresar como

$$[x' \;\; y'] \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + [d, \; c] \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} + f = 0$$

entonces

$$\alpha_1 x'^2 + \alpha_2 y'^2 + d'x' + e'y' + f = 0.$$

donde

$$d' = dP_{11} + eP_{21} \qquad e' = dP_{12} + eP_{22}.$$

El siguiente teorema resume todo el proceso.

7.13 TEOREMA DE LOS EJES PRINCIPALES PARA R^2

Sea

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

la ecuación de una cónica C y sea

$$x^tAx = ax^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas x' - y' tiene la forma

$$\alpha_1 x^{12} + \alpha_2 y^{12} + d'x' + e'y' + f = 0$$

donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución x = Px' donde P diagonaliza ortogonalmente a A.

Ejemplo 7.22

Describa la cónica C cuya ecuación está dada por

$$5x^2 - 4xy + 8y^2 + (20/\sqrt{5})x - (80/\sqrt{5})y + 4 = 0$$

la forma matricial de ésta ecuación es

$$x^tAx + kx + 4 = 0$$

donde

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix} \qquad k = [20/\sqrt{5},\ -80/\sqrt{5}]$$

$$\det(A - \alpha I) = (\alpha - 9)(\alpha - 4)$$

y v_1 v_2 son:

$$P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

diagonaliza ortogonalmente a A. Sustituyendo x = Px'

$$(Px')^t A(Px') + k(Px') + y = 0$$

es decir

$$(x'^t)(P^tAP)x' + (kP)x' + y = 0$$

dado que

$$P^tAP = \begin{bmatrix} 4 & \\ & 9 \end{bmatrix} \quad kP = [20/\sqrt{5}, -80/\sqrt{5}] \begin{bmatrix} 2/5 & -1/5 \\ 1/5 & 2/5 \end{bmatrix} = [-8, -36]$$

entonces

$$4x'^2 + 9y'^2 - 8x' - 36y' + 4 = 0$$

trasladando

$$4(x'^2 - 2x') + 9(y'^2 - 4y') = -4$$

completando cuadrados

$$4(x'^2 - 2x' + 1) + 9(y'^2 - 4y' + 4) = -4+4+36$$

o bién

$$4(x'-1)^2 + 9(y' -2)^2 = 36$$

sea $x'' = x' -1$ $\quad$ $y'' = y' -2$ las ecuaciones de traslación entonces

$$4x''^2 + 9y''^{12} = 36$$

o

$$\frac{x''^2}{9} + \frac{y''^2}{4} = 1$$

es la ecuación de una elipse

<u>Ejemplo 7.23</u>

Describa la cónica C cuya ecuación está dada por

$$5x^2 - 4xy + 8y^2 - 36 = 0$$

o bién

$$x^tAx - 36 = 0$$

donde

$$A = \begin{bmatrix} 5 & -2 \\ -2 & 8 \end{bmatrix}$$

det (A $-\alpha$I) = $(\alpha - 9)\ (\alpha - 4) = 0$

para $\alpha = 4$ se tiene

$$v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

para $\alpha = 9$

$$v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

por lo tanto

$$P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

digonaliza ortogonalmente a A además det P = 1, por lo cual se puede concluir que x = Px' es una rotación o bién:

$$(x')_t\ (P_tAP)x'-36 = 0$$

donde

$$P^tAP = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$$

$$[x'\ y'] \begin{bmatrix} 4 & \\ & 9 \end{bmatrix} \begin{bmatrix} x' \\ y' \end{bmatrix} -36 = 0$$

$$4x'^2 + 9y'^2 - 36 = 0$$

o

$$\frac{x'^2}{9} + \frac{y'^2}{4} = 1$$

es una elipse.

7.7 RESUMEN

1. Si A es una matriz de nxn, entonces un escalar α es un valor propio de A si hay un vector v distinto de cero tal que

$$Av = \alpha v$$

el vector $v \neq 0$ se llama un vector propio de A correspondiente a un valor propio α

2. Si A es de orden nxn entonces la ecuación

$$(A - \alpha I)v = 0$$

es un sistema homogéneo de n ecuaciones con n incógnitas, como por hipótesis (ya que v es un vector propio distinto de cero) el sistema tiene soluciones no triviales entonces det $(A - \alpha I) = 0$ A esta ecuación se le denomina la ecuación característica de A.

3. Si contamos multiplicidades, cada matriz de nxn tiene exactamente n valores caracteristicos.

4. Si A es una matriz de orden nxn entonces las siguientes afirmaciones son equivalentes:

a) α es un valor característico de A
b) El sistema ($A - \alpha I)v = 0$ tiene soluciones no triviales.
c) Existe un vector v e n R^n $v \neq 0$ tal que $Av = \alpha v$

5. Si α es un valor característico de A entonces el espacio solución del sistema de ecuaciones $(A - \alpha I)v = 0$ se denomina el espacio característico de A correspondiente a α, y los vectores diferentes de cero en el espacio característico se denominan los vectores característicos de A correspondientes a α.

6. Sea A una matriz de nxn y sea $\alpha_1, \alpha_2, \ldots, \alpha_m$ valores característicos diferentes de A con sus correspondientes vectores característicos $v_1, v_2, \ldots, v_m$. Entonces v_i son linealmente independientes o bien, los vectores característicos correspondientes a valores característicos diferentes son linealmente independientes. (la demostración se hace por inducción).

7. Sea α un valor cracterístico de A entonces la multiplicidad geométrica de α es la dimensión del espacio característico correspondiente a α (que es la nulidad de la matriz $A - \alpha I$) esto es:

$$\text{multiplicidad geométrica de } \alpha = \dim E_\alpha$$

8. Sea A una matriz de orden nxn, entonces A tiene n vectores caracteristicos linealmente independientes si y solo si la multiplicidad geométrica de todo valor característico es igual a la multiplicidad algebraica. En particular, A tiene n vectores característicos l.i. si todos sus valores caracteristicos son distintos (pues si no, podrían ser menores que n)

9. Sea la matriz A de orden nxn

9.1 Si α es un valor propio de A con v como vector propio correspondiente, entonces ^{k} es un valor propio de A^k, de nuevo con v como vector propio correspondiente, para cualquier entero positivo k.

9.2 Si α es un valor propio de una matriz invertible A con v como vector propio correspondiente, entonces $1/\alpha$ es un valor propio de A^{-1}, de nuevo con v como vector propio correspondiente.

9.3 Si α es un valor propio de A, entonces el conjunto E es un subespacio de R^n.

10. Una secuencia de intentos de m experimentos es una cadena de Markov si

a) El resultado del intento m-ésimo depende solo del resultado del intento (m-1)-ésimo y no de los resultados en los intentos anteriores, y

b) La probabilidad de pasar del estado E_i al estado E_j en dos intentos sucesivos del experimento permanece constante.

11. Se dice que una ecuación de la forma

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

donde a,b,c,d,e,f son números reales y al menos uno de ellos es diferente de cero, es una ecuación cuadrática en x,y. La expresión $ax^2 + 2by + cy^2$, se denomina la forma cuadrática asociada.

12. Sea

$$ax^2 + 2bxy + cy^2 + dx + cy + f = 0$$

la ecuación de una cónica C y sea

$$x^t Ax = ax^2 + 2bxy + cy^2$$

la forma cuadrática asociada. Entonces los ejes de coordenadas se pueden rotar de tal manera que la ecuación C con respecto al nuevo sistema de coordenadas $x' - y'$ tiene la forma

$$\alpha_1 x^{12} + \alpha_2 y^{12} + d'x' + e'y' + f = 0$$

donde α_1 y α_2 son valores propios de A. La rotación se puede efectuar mediante la sustitución $x = Px'$ donde P diagonaliza ortogonalmente a A.

7.8 NOTAS HISTORICAS

LA PRIMER APARICION DE LOS VALORES PROPIOS fue en relación con su uso para resolver ecuaciones diferenciales. En 1743 Leonhard Euler introdujo por primera vez el método estándar de resolución de una ecuación diferencial de n-ésimo orden con coeficientes constantes

$$y^{(n)} + a_{n-1}\, y^{(n-1)} + \ldots + a_1\, y^1 + a_0 y = 0$$

usando funciones de la forma $y = e^{at}$ donde a es una raíz de la ecuación característica:

$$z^n + a_{n-1}\, z^{n-1} + \ldots + a_1\, z + a_0 = 0$$

Esta es la misma ecuación que se obtiene al hacer las sustituciones $y_1 = y$, $y_2 = y'$, $y_3 = y''$,.., $y_n = y^{(n-1)}$,reemplazando la única ecuación de orden n por un sistema de n ecuaciones de primer orden

$$\begin{cases} y'_1 = & y_2 \\ y'_2 = & \quad y_3 \\ y'_n = & -a_0y_1 - a_1y_2 - \ldots -a_{n-1}y_n \end{cases}$$

y calculando la ecuación característica de la matríz de coeficientes de este sistema.

Unos 20 años más tarde, Lagrange dio una versión más explícita de esta misma idea, al hallar la solución de un sistema de ecuaciones diferenciales hallando las raices de lo equivalente a la ecuación característica de la matriz de coeficientes. El sistema particular de ecuaciones diferenciales surgió del examen de los "movimientos infinitesimales" de un sistema mecánico en la vecindad de su posición de equilibrio. En 1774, Lagrange resolvió un problema similar de mecánica celeste usando la misma técnica.

El nombre de **CADENAS DE MARKOV** se debe al matemático ruso Andrei Andreevich Markov (1856-1922) quien las definió por primera vez en un artículo de 1906 que trataba de la ley de los grandes números y posteriormente demostró muchos resultados estándar sobre ellas. Su interés en estas sucesiones se originó en las necesidades de la teoría de probabilidad. Markov nunca trató sus aplicaciones a las ciencias. Los únicos ejemplos reales que utilizó eran de textos literarios, donde los dos estados posibles eran vocales y consonantes. Para ilustrar sus resultados hizo un estudio estadístico de la alternancia de vocales y consonantes en el libro de Pushkin: Eugen Onegin.

Andrei Markov dió clase en la Universidad de San Petersburgo de 1880 a 1905, y se retiró para dar paso a matemáticos más jóvenes. Además de su trabajo en probabilidad, hizo contribuciones a campos como teoría de números, fracciones continuas y teoría de la aproximación. Fue un participante activo en el movimiento liberal ruso en la época anterior a la Primer Guerra Mundial, en varias ocasiones criticó públicamente la actuación de las autoridades estatales. En 1913, cuando como miembro de la Academia de Ciencias se le pidió participar en las pomposas ceremonias de celebración del 300 aniversario de la disnastía Romanov, prefirió organizar una conmemoración del 200 aniversario de la publicación de la ley de los grandes números de Jacobo Bernoulli.

FIBONACCI (1175 - 1230) El notable matemático italiano Leonardo de Pisa mejor conocido por su sobrenombre de Fibonacci, una abreviación de "filus Bonacci" que significa el hijo de Bonaccio de Pisa. Su padre, notario público de Bugía, en Argel, confió su hijo a las enseñanzas de un maestro de cálculo musulmán: dedicó "algunos días" al estudio del ábaco, aprendió los signos numéricos de los indios, y penetró en seguida en los secretos de la ciencia arábiga. Ulteriores viajes lo llevaron posteriormente a través de Oriente y de Europa, hacia Egipto, Siria; Grecia, Sicilia y Provenza. Por doquier planteó nuevas cuestiones a los más afamados maestros con los que medía sus fuerzas para perfeccionar su sabiduría. "Pero todo esto, y el algoritmo y los arcos de Pitágoras, sólo me parecieron otras tantas equivocaciones, comparadas con el método de los indios.

En el año 1202, a los 27 años de edad, fue cuando escribió Leonardo, resumiendo su sabiduría singular, un gran libro sobre el arte de calcular, el Liber abaci, en el que por primera vez un matemático crisitano ofrecía a sus contemporáneos una imágen completa, acabada y ordenada con la mayor claridad, del arte del cálculo tal y como éste se había desarrollado más allá del mundo europeo. Nos informa en ella que, habiendo aprendido en Berbería, el sistema númerico de los árabes lo transmitía "a la raza latina para que ésta no permaneciese más tiempo privada de él".

En el Liber abaci comezaba Leonardo explicando los signos numéricos, enseñando a continuación a calcular con los dedos, asi como las operaciones mas sencillas con los números enteros y las fracciones. Una gran parte de sus problemas se referían a casos prácticos. Asi, por ejemplo, en la sección 8 trataba " de los precios de las mercancías a largo plazo", en la 9, "del intercambio de mercancías y cosas semejantes"; en la 10, "de la aligación y compañía"; en la 11, "de la aleación y mezcla de las monedas"; en la 15, "de las reglas pertenecientes a la Geometría y de los problemas del Algebra y Almucabala". Es en éste mismo libro donde expone su ejemplo de los conejos y como se construye la sucesión que posteriormente llevará su nombre.

EN 1826, CAUCHY ANALIZO LAS FORMAS CUADRATICAS CON TRES VARIABLES, es decir, formas del tipo $Ax^2 + By_2 + Cz^2 + 2Dxy + 2Exz + 2Fyz$. Demostró que la ecuación característica formada a partir del determinante

$$\begin{vmatrix} A & D & E \\ D & B & F \\ E & F & C \end{vmatrix}$$

permanece igual bajo cualquier cambio de ejes rectangulares, lo que nosotros llamariamos un cambio coordenado ortogonal.