Boletín de Matemáticas
Volumen X (1976). págs. 61 - 77

APLICACIONES DEL ALGEBRA LINEAL (*)

JOSE A. BARRETO G.

PRIMERA PARTE.

Elipse, hipérbola, etc. : Una aplicación de la teoría de vectores y valores característicos. Diagonalización de formas cuadráticas.

La ecuación :

1.1 $$ax^2 + 2kxy + by^2 = 1$$

representa

a) Una elipse, ó

b) Una hipérbola, ó

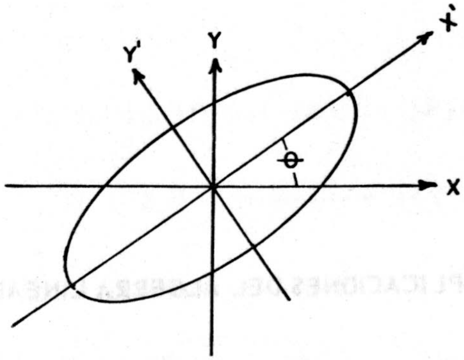
c) Un par de rectas paralelas (1), cuyos ejes de simetría están rotados con respecto a los ejes de coordenadas X-Y un ángulo θ.

Un método para determinar el ángulo θ se puede ver en [1, págs. 318-320] o en [2, págs. 378-379] .

---

(1) Se pueden presentar el caso especial $a=k=b=0$. En tal caso es obvio que no existe un punto $\begin{bmatrix} x \\ y \end{bmatrix}$ que satisfaga 1.1.

(*) Conferencia presentada por el autor en el V Coloquio Colombiano de Matemáticas; Medellín , 1975, *N.* del *E.*



Es conveniente expresar el lugar geométrico 1.1 con respecto a los ejes $X'-Y'$ en los cuales la curva se puede estudiar a la luz de ecuaciones como

a) $$\frac{x'^2}{c^2} + \frac{y'^2}{d^2} = 1 ,$$

1.2 b) $$\frac{x'^2}{c^2} - \frac{y'^2}{d^2} = \pm 1,$$

c) $$cx'^2 = 1 \quad \text{ó} \quad dx'^2 = 1$$

las cuales corresponden a los casos elipse, hipérbola y par de rectas paralelas respectivamente.

Si $\begin{bmatrix} x \\ y \end{bmatrix}$ y $\begin{bmatrix} x' \\ y' \end{bmatrix}$ representan las coordenadas de un punto P con respecto a los ejes $X-Y$ y $X'-Y'$ respectivamente, tenemos [1] que :

1.3 $$X = RX'$$

en donde

$$X = \begin{bmatrix} x \\ y \end{bmatrix} \quad , \quad X' = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad , \text{ y}$$

$$1.4 \qquad R = \begin{bmatrix} \cos\theta & -\operatorname{sen}\theta \\ \operatorname{sen}\theta & \cos\theta \end{bmatrix}$$

La matriz R de 1.4 es una matriz ortogonal [2].

La ecuación 1.1 se puede reescribir como

$$1.5 \qquad X^T A X = 1^{(3)} \quad ,$$

En donde A es la matriz simétrica [4].

$$1.6 \qquad A = \begin{bmatrix} a & k \\ k & b \end{bmatrix}$$

Sustituyendo 1.3 en 1.5 obtenemos [5]

$$1.7 \qquad (X')^T D\, X' = 1$$

en donde

$$1.8 \qquad D = R^T A\, R.$$

La ecuación 1.7 es la ecuación del lugar geométrico, dado por 1.1, con respecto a los ejes $X' - Y'$.

---

(2) Una matriz cuadrada R con elementos reales es una matriz ortogonal sii:
 a) R es una matriz no singular
 b) $R^{-1} = R^T$.

Las matrices ortogonales tienen gran importancia en análisis numérico y el hecho de que R sea ortogonal es sólo la manifestación explícita de un teorema sobre triangularización de matrices cuadradas [2, pág. 318] por medio de transformaciones semejantes.

(3) Representaremos a la matriz $[c]$ por el número c. Para cada matriz $X = [x_{i_j}]$ con elementos reales, se define la matriz traspuesta de X como la matriz $X^T = [y_{i_j}]$ donde $y_{i_j} = x_{j_i}$.

(4) Una matriz A es una matriz simétrica sii $A^T = A$.

(5) Aquí utilizamos la propiedad

$$(RX')^T = (X')^T R^T$$

La matriz D de 1.7 es una matriz diagonal lo cual se puede concluir ya que D es una matriz simétrica (véase 1.8) y el coeficiente de $X'Y'$ en 1.7 debe ser igual a cero [6].

Como R es una matriz ortogonal, la ecuación 1.8 se puede expresar como

1.9 $$D = R^{-1} A R$$

A partir de 1.9 y utilizando el hecho de que D es una matriz diagonal se concluye que los elementos de la diagonal de D son precisamente los valores característicos de A. Es decir : si

1.10 $$D = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}$$

entonces λ_1 y λ_2 son los valores característicos de A.

Además, los vectores columna R_1 y R_2 de R [7] son nada menos que vectores característicos de A correspondientes a los valores característicos λ_1 y λ_2 (en su orden) [8].

A partir de 1.10 y 1.7 concluimos que la ecuación del lugar geométrico dado por 1.1 con respecto a los ejes $X-Y$ es, con respecto a los ejes $X'-Y'$

1.11 $$\lambda_1 x'^2 + \lambda_2 y'^2 = 1 .$$

Por lo tanto para conocer la ecuación del lugar geométrico con respecto a los ejes $X'-Y'$ basta calcular los valores característicos de la matriz A.

Los vectores columna R_1 y R_2 los cuales son a su vez vectores característicos

(6) El hecho de que R sea una matriz ortogonal no es por lo tanto fortuito ni irrelevante.

(7) R_i es el i-ésimo vector columna de R para i = 1,2.

(8) Un método para hallar valores y vectores característicos de la matriz A consiste en determinar las matrices D y R sin utilizar el polinomio característico de A.

asociados a los valores característicos λ_1 y λ_2 (en su orden), señalan la dirección de los ejes $X'-Y'$ ya que los vectores cuyas componentes son

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

con respecto a tales ejes, son vectores en la dirección de los mismos y de acuerdo con 1.3 las expresiones de tales vectores con respecto a los ejes $X-Y$ son

$$R\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{y} \quad R\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{es decir, } R_1 \text{ y } R_2$$

Ejemplo [2] .

Para el lugar geométrico dado por

$$x^2 + 4xy - 2y^2 = 1 ,$$

tenemos que

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

Los valores característicos de A son $\lambda_1 = 2$ y $\lambda_2 = -3$. Un par de vectores propios V_1 y V_2 asociados con λ_1 y λ_2 respectivamente son

$$V_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \text{y} \quad V_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

Luego la ecuación de la cónica con respecto a dos ejes de coordenadas $X'-Y'$ en la dirección de V_1 y V_2 es

$$2x'^2 - 3y'^2 = 1,$$

por lo tanto el lugar geométrico es una hipérbola.

El mismo método se puede extender para el estudio de las superficies cuádricas [1] del tipo

$$X^T A X = c$$

en donde X es un vector con 3 componentes.

Ejemplo :

Para la superficie

$$3x^2 + 2y^2 + 3z^2 - 2xy - 2yz = 1,$$

la matriz A es

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

Los valores propios de A son

$$\lambda_1 = 1, \ \lambda_2 = 3, \ \lambda_3 = 4$$

Los vectores

$$V_1 = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \quad V_2 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad V_3 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

son vectores propios correspondientes a los valores propios λ_1, λ_2 y λ_3 respectivamente.

Por lo tanto la superficie es un elipsoide cuya ecuación es

$$x^2 + 3y^2 + 4z^2 = 1$$

con respecto a los ejes determinados por V_1, V_2 y V_3.

Diagonalización de formas cuadráticas.

En general una forma *cuadrática* en n variables reales $x_1, x_2, \ldots, x_n$, asociada con una matriz *real* A, es una cantidad escalar consistente de sumas de elementos del tipo

$$a_{ij}\, x_i\, x_j \; .$$

La forma cuadrática es por lo tanto

$$X^T A X$$

en donde

$$A = [\, a_{ij} \,] \qquad \text{y} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

Como siempre se puede hallar una matriz *simétrica* B tal que [11]

$$X^T A X = X^T B X \, ,$$

se puede suponer sin pérdida de generalidad que la matriz A es simétrica.

El "truco" para simplificar la forma cuadrática es hallar una matriz ortogonal $P(P^{-1} = P^T)$ que diagonalice a A [12], es decir, tal que

$$P^{-1} A P = P^T A P = D$$

---

(11) Tómese $B = \frac{A + A^T}{2}$ y téngase en cuenta que $X^T A X = (X^T A X)^T = X^T A^T X$.

(12) Tal matriz P siempre existe [2] ya que A es una matriz simétrica.

donde

$$D = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}$$

es una matriz diagonal y $\lambda_1, \lambda_2, \ldots, \lambda_n$ son los valores característicos de la matriz A.

Si se toma Y tal que

$$X = PY$$

(El vector Y señala las coordenadas del vector X en un nuevo eje), se tiene :

$$X^T AX = (PY)^T A(PY) = Y^T P^T APY = Y^T DY = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

y la forma cuadrática se ha *diagonalizado*. Recuérdese que para *diagonalizar* una forma cuadrática no es necesario hallar P explícitamente sino hallar los valores propios $\lambda_1, \lambda_2, \ldots, \lambda_n$ de la matriz A quedando dicha forma reducida a

$$\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2$$

Bibliografía

[1] Apostol, Tom M. [1962]. *"Calculus"*. Vol. 1. Blaisdell Publishing Company, New York.

[2] Noble, Ben [1969]. *"Applied Linear Algebra"*. Prentice Hall, Inc., Englewood Cliffs, New Jersey.

SEGUNDA PARTE.

Aproximación Lineal por Mínimos Cuadrados.

A) Introducción.

Problema 1. Si para determinar la temperatura t en un punto dado de la ciudad, a determinada hora se efectúan mediciones $t_1, t_2, \ldots, t_m$ un buen criterio para determinar t es lograr que minimice

$$r_1^2 + r_2^2 + \cdots + r_m^2 \tag{1}$$

en donde

$$\begin{aligned} t - t_1 &= r_1 \\ t - t_2 &= r_2 \\ &\vdots \\ t - t_m &= r_m \end{aligned} \tag{2}$$

Definiendo

$$A = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad b = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ \cdot \\ t_m \end{bmatrix}, \qquad r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix} \tag{3}$$

Definiendo como es usual

$$\| r \| = \sqrt{r_1^2 + r_2^2 + \cdots + r_m^2} \tag{4}$$

Concluimos que nuestro problema se puede expresar como :

Hallar $t \in \mathbb{R}$, tal que minimice

(5) $$\| b - tA \|$$

Definiendo

(6) $$x = [\, t \,]$$

el problema se puede reformular como :

Hallar $x \in R^1$, tal que minimice

(7) $$\| b - Ax \|$$

Problema 2. Para un cuerpo que parte de una posición inicial e_o (tiempo $t_o = 0$) en línea recta, con movimiento uniforme y velocidad (constante) v_o se tiene que e (espacio), e_o, v_o y t (tiempo) están relacionados por

(8) $$e = e_o + v_o t$$

Uno de los problemas que se propone a los estudiantes de física es determinar e_o y v_o a partir de mediciones sucesivas $e_1, e_2, \ldots, e_m$ del espacio recorrido por un móvil en tiempos $t_1, t_2, \ldots, t_m$.

Como generalmente (y necesariamente ya que se trata de "corregir" errores inherentes a los sistemas de medición usados) $m > 2$, el sistema de m ecuaciones

(9) $$\begin{aligned} e_1 &= e_o + v_o t_1 \\ e_2 &= e_o + v_o t_2 \\ &\vdots \\ e_m &= e_o + v_o t_m \end{aligned}$$

no tiene solución.

Indudablemente, la inconsistencia de (9) es producida por errores propios del experimento o por otros factores. Un criterio muy usado es determinar e_o y v_o en base a que minimicen

$$(10) \qquad r_1^{\,2} + r_2^{\,2} + \cdots + r_m^{\,2}$$

en

$$(11) \qquad \begin{aligned} e_1 - (e_o + v_o\, t_1) &= r_1 \\ e_2 - (e_o + v_o\, t_2) &= r_2 \\ &\vdots \\ e_m - (e_o + v_o\, t_m) &= r_m \end{aligned}$$

Esto equivale a hallar una recta $y = e_o + v_o\, t$ que aproxime los datos obtenidos experimentalmente tomando como punto esencial el criterio (10). (Véase el gráfico adjunto).

(12)

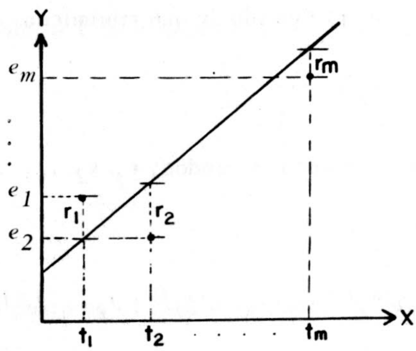


Figura 1

Denotando

$$(13) \qquad A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \quad b = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix}, \quad x = \begin{bmatrix} e_o \\ v_o \end{bmatrix}, \quad r = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{bmatrix}$$

el problema propuesto se puede replantear como :

Hallar $x \in \mathbb{R}^2$ y tal que minimice

$$(14) \qquad \| b - Ax \|$$

Problema 3. Un problema típico de aproximación en "análisis numérico" es el siguiente.

Dada una función

$$\varphi : \mathbb{R}^k \to \mathbb{R}$$

y funciones

$$\psi_i : \mathbb{R}^k \to \mathbb{R} \quad , \qquad i = 1, 2, \ldots, n ,$$

se desea determinar $\lambda_1, \lambda_2, \ldots, \lambda_n$ en $\mathbb{R}$ tales que

$$\lambda_1 \psi_1(x) + \lambda_2 \psi_2(x) + \cdots + \lambda_n \psi_n(x)$$

aproxime a $\varphi(x)$ (para todo x), en el sentido de que minimicen

$$(15) \qquad r_1^2 + r_2^2 + \cdots + r_m^2$$

Los r_i de (15) han sido definidos tomando m nodos $x_1, x_2, \ldots, x_m$ como punto de referencia así :

$$(16) \quad \varphi(x_j) - (\lambda_1 \psi_1(x_j) + \lambda_2 \psi_2(x_j) - \ldots + \lambda_n \psi_n(x_j)) = r_j , \qquad j = 1, 2, \ldots, m$$

Con la notación :

$$(17) \qquad A = \begin{bmatrix} \psi_1(x_1) & \psi_2(x_1) & \ldots & \psi_n(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \ldots & \psi_n(x_2) \\ \vdots & & & \vdots \\ \psi_1(x_m) & \psi_2(x_m) & \ldots & \psi_n(x_m) \end{bmatrix}, \quad x = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad b = \begin{bmatrix} (x_1) \\ (x_2) \\ \vdots \\ \varphi(x_m) \end{bmatrix}$$

este problema se formula así :

Hallar $x \in R^n$ y tal que minimice

$$(18) \qquad || b - Ax || .$$

B) Formulación general.

El fin de esta charla es presentar un método para resolver el siguiente problema de *aproximación lineal por mínimos cuadrados.*

A partir de una matriz A de orden mxn y un vector b con m componentes, determine $x \in \mathbf{R}^n$ tal que minimice

$$(19) \qquad || Ax - b ||$$

C) Solución.

El método de solución será presentado a partir de nuestros problemas 1,2.

Para el problema 1 de acuerdo con (5), hay que explorar la recta

$$L \quad : \quad t \begin{bmatrix} 1 \\ 1 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

hasta hallar un punto en el cual la distancia a b sea mínima.

Este punto está señalado precisamente por la proyección b_1 de b sobre la recta como lo muestra la figura.

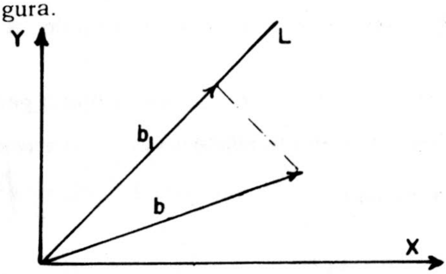


Figura 2

Es decir que resolver el problema 1, se reduce a hallar las soluciones de

$$(20) \qquad tA = b_1$$

o en la nomenclatura $x = [\,t\,]$, las soluciones de

$$(21) \qquad Ax = b_1$$

La ecuación (21) sería ideal para "despejar" a x si no fuera por el grave inconveniente de que nosotros no conocemos a b_1 .

Pasemos ahora al problema 2. Dada una matriz A de orden mxn, el conjunto

$$(22) \qquad \{ Ax \}$$

es precisamente el subespacio generado por las columnas de A es decir

$$(23) \qquad \{ Ax : x \in \mathbf{R}^n \} = \{ x_1 A_1 + x_2 A_2 + \cdots + x_n A_n \}$$

en donde

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}$$

y $A_1, A_2, \ldots, A_n$ son los n vectores columna de A.

Si A es de orden $3x2$, $\{ Ax : x \in \mathbf{R}^2 \}$ es el plano generado por las combinaciones lineales de los vectores columna A_1, A_2 como se aprecia enseguida

Luego hallar x minimizando

$$(24) \qquad \| b - Ax \|$$

es equivalente a resolver

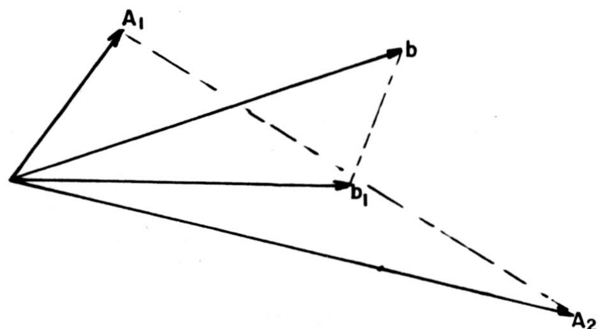


Figura 3

$$(25) \qquad Ax = b_1$$

en donde b_1 es la "proyección" de b sobre $\{Ax\}$.

Para hallar x utilizando (25) es necesario determinar primero a b_1.

Como se aprecia en las figuras (2) y (3), el vector $b - b_1$ es ortogonal al subespacio generado por los vectores columna de A $(\{Ax\})$ y en consecuencia

$A^T(b - b_1) = 0$. Es decir, que

$$(26) \qquad A^T b = A^T b_1$$

Toda solución de la ecuación

$$(27) \qquad Ax = b_1$$

presentada en (21) y (25) como la ecuación que resuelve el problema de aproximación lineal por mínimos cuadrados es una solución de

$$(28) \qquad A^T Ax = A^T b_1$$

La igualdad (26) nos permite asegurar que la solución del problema de los míni -

mos cuadrados es una solución de

$$(29) \qquad A^T Ax = A^T b$$

La solución al problema 1 es por lo tanto una de las soluciones de

$$(30) \qquad [1 \ 1 \ldots 1] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} x = [1 \ 1 \ldots 1] \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_m \end{bmatrix}$$

es decir, una solución de

$$(31) \qquad [m] \ x = [t_1 + t_2 + \cdots + t_m]$$

Como $x = [t]$, concluimos que la solución es

$$(32) \qquad t = \frac{t_1 + t_2 + \cdots + t_m}{m}$$

Al resolver el problema 2 a partir de los datos

$$t_1 = 1 \qquad e_1 = 1$$

$$t_2 = 2 \qquad e_2 = 3$$

$$t_3 = 3 \qquad e_3 = 6$$

de acuerdo con 29, la solución $x = \begin{bmatrix} e_o \\ v_o \end{bmatrix}$, es cuna solución de

$$(33) \qquad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} e_o \\ v_o \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 6 \end{bmatrix}$$

o lo que es lo mismo, una solución de

$$\begin{bmatrix} 2 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} e_o \\ v_o \end{bmatrix} = \begin{bmatrix} 10 \\ 25 \end{bmatrix}$$

lo que es $e_o = -5/3, \ v_o = 5/2$.

La solución del problema 2 con m puntos es una solución

$$\begin{bmatrix} m & \sum_{i=1}^{m} t_i \\ \sum_{i=1}^{m} t_i & \sum_{i=1}^{m} t_i^2 \end{bmatrix} \begin{bmatrix} e_o \\ v_o \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{m} e_i \\ \sum_{i=1}^{m} t_i e_i \end{bmatrix}$$

Bibliografía

[1] Kreider, Donald L. [1966] . "*An Introduction to Linear Analysis*". Addison Wesley Publishing Company, Inc. Reading, Massachusetts.

[2] Shilov, Georgi E. [1961] . "*An Introduction to the Theory of Linears Spaces*". Prentice Hall, Inc., Englewood Cliffs, New Jersey.

[3] Stewart, G. W. [1973] . "*Introduction to Matrix Computations*". Academic Press, New York.