FORMULACION MATRICIAL DEL PROCESO DE GRAM-SCHMIDT O COMO CALCULAR VALORES PROPIOS DE UNA MATRIZ

Por Humberto Madrid de la Vega*

Los valores propios de una matriz A de orden n son los ceros del polinomio característico de orden n

$$p(\lambda) = \det(A - \lambda I)$$

En los cursos de Algebra Lineal esta caracterización teórica es generalmente la única que se utiliza, pero determinar el polinomio característico y luego encontrar sus raíces, es demasiado complicado cuando el orden de la matriz es grande. Ya que a menudo se presenta en la práctica el problema de determinar los valores propios de una matriz, este esquema teórico no es aplicable.

El presente trabajo pretende mostrar un método para calcular valores propios de una matriz: el algoritmo *QR*. Nos limitaremos únicamente a la idea del método para no complicar el trabajo. Además se menciona literatura adecuada para el que se interese en profundizar en el algoritmo.

Por motivos de simplicidad trabajaremos con vectores linealmente independientes y matrices reales y no singulares.

El Proceso de Gram-Schmidt.

Sean $\mathbf{x}_1, \ldots, \mathbf{x}_n$ vectores linealmente independientes en R^n.

Recordemos que el proceso de Gram-Schmidt consiste en construir un conjunto $\mathbf{y}_1, \ldots, \mathbf{y}_n$ de vectores ortonormales a partir de los vectores $\mathbf{x}_i$ $(i = 1, \ldots, n)$. Primero definimos

(1) $$\mathbf{y}_1 = \frac{x_i}{x_i}$$

Luego construimos

(2) $$u_2 = \mathbf{x}_2 - a_{12}\mathbf{y}_1$$

* Profesor de la Facultad de Ciencias de la UNAM. Ayudante de Investigador CIMASS. UNAM.

aquí se escoge a_{12}, de suerte que u_2 sea ortogonal a y_1, entonces definimos

(3) $$y_2 = \frac{u_2}{|u_2|}$$

En general, en el paso k-ésimo construimos

(4) $$u_k = x_k - a_{k-1,k}\, y_{k-1} - \ldots - a_{1k}\, y_1.$$

donde las a_{ik} se toman de manera tal que u_k sea ortogonal a las y_i ($i = 1,\ldots, n$). Definimos entonces

(5) $$y_k = \frac{u_k}{|u_k|} \qquad\qquad k = 1, \ldots, n$$

Por construcción las y_k constituyen un conjunto ortonormal.

La Formulación Matricial.

De (1) obtenemos

(6) $x_1 = a_{11} y_1$, en donde $a_{11} = |x_1|$.

De (3) obtenemos

$u_2 = a_{22} y_2$, en donde $a_{22} = |u_2|$. Esto junto con (2) nos da

$a_{22} y_2 = x_2 - a_{12} y_1$, de lo que resulta:

(7) $$x_2 = a_{12} y_1 + a_{22} y_2.$$

En general, de (5) y (4) obtenemos

(8) $$x_k = a_{1k} y_1 + a_{2k} y_2 + \ldots + a_{kk} y_k \qquad (k = 1,\ldots, n),$$

en donde $a_{kk} = |u_k|$.

Ahora veremos que las ecuaciones dadas por (8) podemos formularlas como una ecuación matricial. Para esto, sea A la matriz cuyas columnas son las x_i; Q la matriz cuyas columnas son las y_i; y, finalmente, sea R la matriz siguiente:

$$R = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{nn} \end{pmatrix}$$

Proposición: Si A, Q y R son las matrices descritas arriba, entonces

$$A = QR.$$

En efecto, la primera columna de A se obtiene multiplicando los renglones de Q por la primera columna de R, y esto nos da:

$$x_1 = a_{11} y_1.$$

En general, la columna k-ésima de A se obtiene multiplicando los renglones de Q por la k-ésima columna de R, y esto nos da:

$x_k = a_{1k} y_1 + a_{2k} y_2 + \dots + a_{kk} y_k$, lo que coincide con (8).

Así queda demostrada nuestra Proposición.

A partir de (8) hemos llegado a descomponer la matriz A (cuyas columnas son las x_i) en la forma

(9) $$A = QR,$$

donde Q es una matriz ortogonal, es decir, sus columnas constituyen un conjunto ortonormal de vectores (equivalentemente: $Q^t = Q^{-1}$) y R es una matriz triangular superior.

A la descomposición (9) se le llama descomposición QR.

Por otro lado, si una matriz A (no singular) se puede descomponer en $A = QR$, siendo Q y R matrices ortogonal y triangular superior respectivamente; entonces, por los argumentos usados en la demostración de la proposición, se puede llegar a la ecuación (5), esto es, las columnas de Q serán las mismas y_i que se obtienen por el proceso de Gram-Schmidt.

Resumiendo:

Sean $x_1, \dots, x_n$ vectores linealmente independientes en R^n. Construir a partir de estos vectores, por el proceso de Gram-Schmidt, un conjunto ortonormal de vectores $y_1, \dots, y_n$, es equivalente a descomponer la matriz A, cuyas columnas son las x_i, en el producto

$$A = QR,$$

siendo Q una matriz ortogonal, cuyas columnas son las y_i, y R una matriz triangular superior.

Aplicación al Cálculo de Valores Propios.

Supongamos que $A = A_1$ es no singular, entonces sus columnas son linealmente independientes y podemos descomponer A_1 en

$$A_1 = Q_1 R_1.$$

Considérese ahora la matriz

$$A_2 = R_1 Q_1.$$

Como Q_1 y R_1 son no singulares (¿por qué?), entonces

$$A_2 = R_1 A_1 R_1^{-1} = Q_1^{-1} A_1 Q_1,$$

esto es, A_2 es similar a A_1. Por lo tanto, tiene los mismos valores propios que A_1.

Por otro lado, por ser A_2 no singular, se puede escribir como

$$A_2 = Q_2 R_2.$$

Si formamos la matriz

$$A_3 = R_2 Q_2,$$

se tendrá que A_3 es similar a A_2. Por lo tanto a A_1, y tendrá entonces los mismos valores propios que A_1.

Podemos luego descomponer a A_3 como

$$A_3 = Q_3 R_3$$

y continuar el proceso. Se tendrá, pues, una sucesión de matrices A_n, donde si

$$A_n = Q_n R_n,$$

el elemento A_{n+1} se obtiene en la forma:

$$A_{n+1} = R_n Q_n,$$

y cada A_n tiene los mismos valores propios que A_1.

Se puede demostrar que bajo ciertas condiciones la sucesión A_n converge a una matriz triangular superior A, similar a A_1, y entonces los elementos de la diagonal de A serán los valores propios de A_1.

Este método para encontrar los valores propios de una matriz es conocido como el algoritmo QR. Este algoritmo converge cuando A es simétrica o cuando tiene valores propios de distinto módulo, y aún en casos más generales. El lector interesado puede consultar la bibliografía marcada con los números I, III, V y VI.

El algoritmo QR es usado frecuentemente en la práctica, aunque con ciertas modificaciones en muchos casos. Por ejemplo, para descomponer A en el producto QR, es demasiado laborioso usar el proceso de Gram-Schmidt y a menudo se usan otros métodos mucho más eficientes como los debidos a Givens (V) y Householder (II).

Por último es importante mencionar que también hay modificaciones al algoritmo para hacer que la convergencia sea más rápida (III y IV).

BIBLIOGRAFIA:

I FRANKLIN, J. N., *Matrix Theory*, Prentice-Hall, (1968).

II PARLETT, B., *The LU and QR Algorithms*, en A. Ralston and H. S Wilf, Eds. "Mathematical Methods for Digital Computers", Vol. II Wiley (1967).

III NOBLE, B., *Applied Linear Algebra*, Prentice-Hall, (1969).

IV PARLETT, B., *The development and use of methods of LR type*, SIAM Review, Vol. 6, (1964), 275-295.

V RALSTON, A. *A First Course in Numerical Analysis* McGraw Hil (1965).

VI WILKINSON, J. H., *Convergence of the LR, QR and related algo rithms*, The Computer Journal, Vol. 8, (1965), 77-84.