Valores y vectores característicos

Cortés Rosas Jesús Javier, González Cárdenas Miguel Eduardo
Pinilla Morán Víctor Damián, Salazar Moreno Alfonso
Tovar Pérez Víctor Hugo *

2019

Resumen

Esta publicación pertenece al proyecto *Plataforma educativa para Análisis Numérico*, realizado con al apoyo del Programa UNAM-DGAPA-PAPIME PE105717.

El polinomio, los valores y vectores característicos son datos de una matriz que permiten resolver diversos problemas que provienen del modelado de sistemas físicos. Al respecto, el método de Krylov permite obtener los coeficientes del polinomio característico mientras que el método de las potencias proporciona el mayor y menor valores característicos. Son una alternativa iterativa para los procesos analíticos del Álgebra lineal.

El objetivo de este trabajo es presentar la aportación del análisis numérico a la obtención del polinomio (o ecuación), valores y vectores característicos; se presentará un método que pueda desarrollarse como algoritmo y, en consecuencia, desarrollado por medio de un programa de cómputo.

En este sentido, para reconocer los contenidos se debe disponer de los antecedentes necesarios del álgebra lineal en cuanto a los tópicos antes mencionados.

Las matrices de orden n, x, n no singulares poseen un *polinomio característico*; sus raíces son llamadas *valores característicos* y cada uno tiene asociado un *vector característico.*

El polinomio característico de la matriz A se obtiene por medio de la expresión:

$$|A - \lambda I| = 0 \tag{1}$$

El resultado de esta determinante es un polinomio en función de λ de grado igual al orden de la matriz A, en este caso, de orden n. Este *polinomio característico* posee n raíces o *valores característicos*; por lo cual, la matriz A de orden n posee n valores característicos.

El polinomio característico es de la forma:

$$a_0\lambda^n + a_1\lambda^{n-1} + a_2\lambda^{n-2} + ... + a_{n-1}\lambda + a_n = 0 \tag{2}$$

*Profesores de la División de Ciencias Básicas de la Facultad de Ingeniería de la UNAM.

1. Método de Krilov para obtener el polinomio característico

Este método se fundamenta en la aplicación del *Teorema de Cayley-Hamilton*, mismo que establece que toda matriz A verifica su ecuación característica:

$$F(A) = 0 \tag{3}$$

Es decir, si sustituimos a la matriz A en el polinomio (2), el resultado deberá ser cero. Sin embargo, operativamente es necesario hacer algunos comentarios. De inicio, la matriz A es de orden n, por lo cual la sustitución arrojará un sistema de n ecuaciones lineales; en consecuencia, el coeficiente a_0 deberá ser diferente de cero. Resulta conveniente hacer que este coeficiente sea la unidad, por lo cual se divide el polinomio entero por a_0, resultando:

$$\lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + ... + b_{n-1}\lambda + b_n = 0 \tag{4}$$

Donde los coeficientes b_i se obtienen como $b_i = \frac{a_i}{a_0}$. Aplicando el teorema de Cayley-Hamilton en el polinomio (4):

$$F(A) = A^n + b_1A^{n-1} + b_2A^{n-2} + ... + b_{n-1}A + b_nI = 0 \tag{5}$$

El polinomio (5) representa un sistema de ecuaciones lineales cuyas incógnitas son los coeficientes b_i. La solución de este sistema nos proporciona los coeficientes b_i que sustituidos en (4) nos da como resultado el polinomio característico de A.

Una forma sencilla de realizar este procedimiento es simplificar la elevación de la matriz A a las potencias necesarias. Esto se logra multiplicando la matriz A por un vector $\bar{y}$ compatible diferente de cero. Debe recordarse que la multiplicación de una matriz por un vector compatible arroja un vector.

Este vector $\bar{y}$ puede ser libremente elegido, proponiéndose que su conformación permita realizar de mejor forma las operaciones. Una buena elección es escoger al vector con la forma:

$$\bar{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Ubicando al elemento 1 en una posición estratégica de acuerdo con los coeficientes de A de tal forma que, aprovechando las multiplicaciones por cero, se minimicen las operaciones.

Atendiendo a la anterior recomendación, el sistema queda la forma:

$$A^n\bar{y} + b_1A^{n-1}\bar{y} + b_2A^{n-2}\bar{y} + ... + b_{n-1}A\bar{y} + b_n\bar{y}I = 0 \tag{6}$$

Finalmente, el sistema de ecuaciones puede ser resuelto por el método de preferencia.

1.1. Ejemplo de aplicación

Sea la matriz:

$$A = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ -2 & 3 & -1 \end{bmatrix} \tag{7}$$

De acuerdo con la ecuación (4), el polinomio característico tendrá la forma:

$$\lambda^3 + b_1\lambda^2 + b_2\lambda + b_3 = 0 \tag{8}$$

De acuerdo con la ecuación (6), el sistema de ecuaciones tendrá la forma:

$$A^3\bar{y} + b_1 A^2\bar{y} + b_2 A\bar{y} + b_3\bar{y}I = 0 \tag{9}$$

Se propone al vector $\bar{y}$ de acuerdo con la forma recomendada:

$$\bar{y} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \tag{10}$$

Multiplicando:

$$A\bar{y} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ -2 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} \tag{11}$$

multiplicando:

$$A^2\bar{y} = A \cdot A\bar{y} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ -2 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} \tag{12}$$

$$A^3\bar{y} = A \cdot A^2\bar{y} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ -2 & 3 & -1 \end{bmatrix} \cdot \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} = \begin{bmatrix} -1 \\ 4 \\ -4 \end{bmatrix} \tag{13}$$

Sustituyendo los resultados (11), (12) y (13) en el esquema (9) se conforma el sistema de ecuaciones lineales:

$$\begin{bmatrix} -1 \\ 4 \\ -4 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \\ 6 \end{bmatrix} \cdot b_1 + \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \cdot b_2 + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot b_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{14}$$

Que se expresa en forma más coloquial como:

$$\begin{array}{rcrcrcr} -b_1 & + & b_2 & + & b_3 & = & 1 \\ & & 2b_2 & & & = & -4 \\ 6b_1 & + & 2b_2 & & & = & 4 \end{array} \tag{15}$$

Cuya solución es:

$$\begin{array}{rcr} b_1 & = & 0 \\ b_2 & = & -2 \\ b_3 & = & 3 \end{array} \tag{16}$$

Finalmente, sustituyendo en (8) se obtiene la ecuación característica de la matriz A:

$$\lambda^3 - 2\lambda + 3 = 0 \tag{17}$$

1.2. Conclusiones

Krilov debe utilizarse en conjunto con un método de solución de sistemas de ecuaciones lineales. Esta unión arroja una alternativa muy apropiada cuando se trata de evitar resolver determinantes de orden mayor considerando que se debe hacer en forma analítica.

2. Método de las potencias

La manera ortodoxa de obtener los valores característicos de la una matriz A es resolver las raíces de su polinomio característico. El método de las potencias ofrece una opción para obtener el mayor y el menor valor característico de la matriz A de orden nxn sin la necesidad de disponer de la ecuación característica.

El método de las potencias es un proceso iterativo de aproximaciones sucesivas, por lo cual, además de la matriz A deberá conocerse una tolerancia preestablecida.

2.1. Definición del método de las potencias

Sea el sistema homogéneo de ecuaciones lineales:

$$\begin{array}{ccccccccccc} (a_{11}-\lambda)x_1 & + & a_{12}x_2 & + & a_{13}x_3 & + & ... & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & (a_{22}-\lambda)x_2 & + & a_{23}x_3 & + & ... & + & a_{2n}x_n & = & b_2 \\ a_{31}x_1 & + & a_{32}x_2 & + & (a_{33}-\lambda)x_3 & + & ... & + & a_{3n}x_n & = & b_3 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{n1}x_1 & + & a_{n2}x_2 & + & a_{n3}x_3 & + & ... & + & (a_{nn}-\lambda)x_n & = & b_n \end{array} \tag{18}$$

Que puede representarse como:

$$|A - \lambda I|\bar{x} = 0 \tag{19}$$

A partir de la expresión (19) pueden obtenerse los valores característicos, ya que representa a la ecuación característica de la matriz A; el tratamiento para obtener el valor mayor o el menor es ligeramente diferente, pero en ambos casos proporciona el vector característico respectivo.

2.2. Mayor valor característico

La expresión (19) puede acomodarse de la siguiente forma:

$$A\bar{x} - \lambda\bar{x} = 0$$

$$A\bar{x} = \lambda\bar{x} \tag{20}$$

Para la expresión (20) se percibe la forma característica de un proceso de aproximaciones sucesivas. El método propone utilizar un vector inicial $\bar{x}_{(0)} \neq 0$ compatible y multiplicarlo por la matriz A. El resultado será un nuevo vector $\bar{x}_{(1)}$ el cual será normalizado utilizando su elemento mayor; este es utilizado para la normalización será una aproximación al mayor valor característico λ_i y el vector normalizado será su vector asociado. El proceso se repetirá hasta que la diferencia entre dos aproximaciones cumpla con una tolerancia preestablecida.

Para completar el método, la ecuación (20) debe expresarse en forma iterativa:

$$A\bar{x}_{(k)} = \lambda_{(k+1)}\bar{x}_{(k+1)} \tag{21}$$

donde k es el número de la iteración y $k = 1, 2, 3, ...$

2.3. Menor valor característico

Retomamos la ecuación (19), misma que se premultiplica por la matriz inversa de A, que es A^{-1}.

$$A^{-1}A\bar{x} = A^{-1}\lambda\bar{x}$$

La que resulta en:

$$\bar{x} = A^{-1}\lambda\bar{x}$$

Dividiendo entre λ:

$$\frac{1}{\lambda}\bar{x} = A^{-1}\bar{x} \tag{22}$$

Para completar el proceso, haciendo a (22) una ecuación iterativa:

$$A^{-1}\bar{x}_{(k)} = \frac{1}{\lambda_{(k+1)}}\bar{x}_{(k+1)} \tag{23}$$

donde k es el número de la iteración y $k = 1, 2, 3, ...$

El tratamiento para obtener el menor valor característico es similar al utilizado para el mayor valor, a diferencia del uso de la matriz inversa A^{-1} y que el factor de normalización representa al recíproco del menor valor característico de la matriz A. El proceso se repetirá hasta que la diferencia entre dos aproximaciones cumpla con una tolerancia preestablecida.

2.4. Ejemplo de aplicación

Sea la matriz A:

$$A = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} \tag{24}$$

Se calculará inicialmente el mayor valor característico y posteriormente el menor. La primera acción será definir un vector inicial $\bar{x}_{(0)} \neq 0$ compatible que debe diseñarse de forma práctica y que permita el menor número de operaciones. Se propone el vector:

$$\bar{x}_{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Realizando la primera iteración:

$$A \cdot \bar{x}_{(0)} = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

El elemento de mayor valor en este vector resultante representa la primera aproximación al mayor valor característico, es decir, $\lambda_{(0)} = 2$; normalizando el vector resultado de la multiplicación se obtiene la siguiente aproximación:

$$\lambda_{(0)} = 2 \qquad \bar{x}_{(1)} = \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix}$$

Repitiendo el proceso:

$$A \cdot \bar{x}_{(1)} = \begin{bmatrix} 2 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0{,}5 \\ 0{,}5 \end{bmatrix} = \begin{bmatrix} 3{,}5 \\ 3 \\ 3 \end{bmatrix}; \quad \lambda_{(1)} = 3{,}5 \qquad \bar{x}_{(2)} = \begin{bmatrix} 1 \\ 0{,}8571 \\ 0{,}8571 \end{bmatrix}$$

El cálculo del error absoluto entre las dos aproximaciones a λ es:

$$|\lambda_1 - \lambda_0| = |3{,}5 - 2| = 1{,}5$$

En los siguientes cuadros se presenta el proceso iterativo completo:

Cuadro 1 Iteraciones 4 a 8

	x_0		x_1		x_2		x_3		x_4
	1,00000	2,00000	1,00000	3,50000	1,00000	4,57143	1,00000	4,90625	1,00000
	0,00000	1,00000	0,50000	3,00000	0,85714	4,42857	0,96875	4,87500	0,99363
	0,00000	1,00000	0,50000	3,00000	0,85714	4,42857	0,96875	4,87500	0,99363
$\lambda_{(k)}$		2,00000		3,50000		4,57143		4,90625	
Tol				1,50000		1,07143		0,33482	

Cuadro 2 Iteraciones 4 a 8

	x_4		x_5		x_6		x_7		x_8
	1,00000	4,98089	1,00000	4,99616	1,00000	4,99923	1,00000	4,99985	1,00000
	0,99363	4,97452	0,99872	4,99488	0,99974	4,99898	0,99995	4,99980	0,99999
	0,99363	4,97452	0,99872	4,99488	0,99974	4,99898	0,99995	4,99980	0,99999
$\lambda_{(k)}$		4,98089		4,99616		4,99923		4,99985	
Tol		0,07464		0,01527		0,00307		0,00061	

Cuadro 3 Iteraciones 4 a 8

	x_8		x_9
	1,00000	4,99997	1,00000
	0,99999	4,99996	1,00000
	0,99999	4,99996	1,00000
$\lambda_{(k)}$		4,99997	
Tol		0,00012	

En el cuadro 3 se presenta en forma aislada la última iteración, de la cual se concluye que con una tolerancia de 0,00012, el mayor valor característico de la matriz A en $\lambda = 4{,}99997 \approx 5$ y su vector asociado es:

$$\bar{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \tag{25}$$

Calculando ahora el menor valor característico. Para la matriz A, su inversa es:

$$A^{-1} = \begin{bmatrix} 0{,}8 & -0{,}4 & -0{,}2 \\ -0{,}2 & 0{,}6 & -0{,}2 \\ -0{,}2 & -0{,}4 & 0{,}8 \end{bmatrix} \tag{26}$$

Se propone el vector inicial $\bar{x}_{(0)} \neq 0$ compatible:

$$\bar{x}_{(0)} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

Realizando la primera iteración:

$$A^{-1} \cdot \bar{x}_{(0)} = \begin{bmatrix} 0{,}8 & -0{,}4 & -0{,}2 \\ -0{,}2 & 0{,}6 & -0{,}2 \\ -0{,}2 & -0{,}4 & 0{,}8 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0{,}8 \\ -0{,}2 \\ -0{,}2 \end{bmatrix}$$

El elemento de mayor valor en este vector resultante representa al recíproco de la primera aproximación al menor valor característico, es decir, $\frac{1}{\lambda_{(0)}} = 0{,}8$; normalizando el vector resultado de la multiplicación se obtiene la siguiente aproximación:

$$\frac{1}{\lambda_{(0)}} = 0{,}8 \qquad \bar{x}_{(1)} = \begin{bmatrix} 1 \\ -0{,}25 \\ -0{,}25 \end{bmatrix}$$

Repitiendo el proceso:

$$A^{-1} \cdot \bar{x}_{(1)} = \begin{bmatrix} 0{,}8 & -0{,}4 & -0{,}2 \\ -0{,}2 & 0{,}6 & -0{,}2 \\ -0{,}2 & -0{,}4 & 0{,}8 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -0{,}25 \\ -0{,}25 \end{bmatrix} = \begin{bmatrix} 0{,}95 \\ -0{,}3 \\ -0{,}3 \end{bmatrix}; \quad \frac{1}{\lambda_{(1)}} = 0{,}95 \qquad \bar{x}_{(2)} = \begin{bmatrix} 1 \\ -0{,}3298 \\ -0{,}3158 \end{bmatrix}$$

El cálculo del error absoluto entre las dos aproximaciones a $\frac{1}{\lambda}$ es:

$$\left| \frac{1}{\lambda_1} - \frac{1}{\lambda_0} \right| = |0{,}95 - 0{,}8| = 0{,}15$$

En los siguientes cuadros se presenta el proceso iterativo completo:

Cuadro 4 Iteraciones 0 a 4

	x_0		x_1		x_2		x_3		x_4
	1,00000	0,80000	1,00000	0,95000	1,00000	0,98947	1,00000	0,99787	1,00000
	0,00000	-0,20000	-0,25000	-0,30000	-0,31579	-0,32632	-0,32979	-0,33191	-0,33262
	0,00000	-0,20000	-0,25000	-0,30000	-0,31579	-0,32632	-0,32979	-0,33191	0,33262
$\frac{1}{\lambda_{(k)}}$		0,80000		0,95000		0,98947		0,99787	
Tol				0,15000		0,03947		0,00840	

Se concluye que con una tolerancia de 0,00007, el menor valor característico de la matriz A es:

$$\frac{1}{\lambda} = 0{,}99998 \Rightarrow \lambda = \frac{1}{0{,}99998} = 1{,}00002 \approx 1 \tag{27}$$

Cuadro 5 Iteraciones 4 a 7

	x_4		x_5		x_6		x_7
	1,00000	0,99957	1,00000	0,99991	1,00000	0,99998	1,00000
	-0,33262	-0,33305	-0,33319	-0,33328	-0,33330	-0,33332	-0,33333
	-0,33262	-0,33305	-0,33319	-0,33328	-0,33330	-0,33332	-0,33333
$\frac{1}{\lambda_{(k)}}$		0,99957		0,99991		0,99998	
Tol		0,00170		0,00034		0,00007	

Y su vector asociado es:

$$\bar{x} = \begin{bmatrix} 1 \\ -0{,}33333 \\ -0{,}33333 \end{bmatrix} \approx \begin{bmatrix} 1 \\ -\frac{1}{3} \\ -\frac{1}{3} \end{bmatrix} \tag{28}$$

Como información adicional, la ecuación característica de la matriz A detallada en (24) es:

$$-\lambda^3 + 7\lambda^2 - 11\lambda + 5 = 0 \tag{29}$$

Al resolver este polinomio se verificara la coincidencia con los resultado expresados en (25) y (27).

2.5. Conclusiones

El método de las potencias, junto con Krilov, permite obtener, a través de recursos de cómputo, los resultados característicos de matrices. Si bien no se especifican criterios de convergencia dado que se trata de métodos de aproximaciones sucesivas, es necesario aclarar que la característica fundamental para obtener resultados se limita a que la matriz cumpla con los requisitos específicos marcados por el álgebra matricial.

Referencias

Borras, H., Duran, R., y Iriarte, R. (1984). *Apuntes de métodos numéricos* (F. de Ingeniería UNAM, Ed.).

Burden, R., y Faires, D. (2011). *Análisis numérico* (C. Learning, Ed.).

García B., S. (2017). *Métodos numéricos.*

Gerald, C. (1991). *Análisis numérico* (Alfaomega, Ed.).

Gerald, C., y Wheatley, P. (2000). *Análisis numérico con aplicaciones* (P. Hall, Ed.).

Olivera Salazar, A. (s.f.). *Métodos numéricos* (Limusa, Ed.).

Sandoval, H. (2017). *Métodos numéricos.*