

MÉTODOS NUMÉRICOS II

SEGUNDO DE GRADO EN MATEMÁTICAS, CURSO 2019/20

FACULTAD DE CIENCIAS DE LA UNIVERSIDAD DE MÁLAGA

DPTO. DE ANÁLISIS MATEMÁTICO, ESTADÍSTICA E INVESTIGACIÓN OPERATIVA, Y MATEMÁTICA APLICADA
PROF. FRANCISCO JOSÉ PALMA MOLINA (ÁREA DE CONOCIMIENTO DE MATEMÁTICA APLICADA)

TEMA 3

Cálculo de Valores y Vectores Propios

Los objetivos de este tema son:

- motivar la necesidad de los métodos numéricos para el cálculo de los valores y los vectores propios de una matriz;
- condicionamiento de un problema de valores propios;
- métodos clásicos para el cálculo de los valores propios: Leverrier, Krilov, etc.;
- métodos iterativos para el cálculo de algunos o todos los valores y los vectores propios: potencia, Jacobi, etc.

1. Planteamiento del problema

Comenzamos resaltando que el problema del cálculo de los valores y los vectores propios de una matriz es un problema importante por sus muchas *aplicaciones prácticas*: teoría de ecuaciones en derivadas parciales, estudio de las vibraciones y oscilaciones, deformaciones y tensiones de una estructura, determinación de los ejes de las cónicas y cuádricas, optimización, etc.

Problema 1 Dada $A \in \mathcal{M}_n(\mathbb{K})$, calcular $\lambda \in \mathbb{C}$ (valor propio) y $X \in \mathbb{C}^n$ (vector propio asociado) tal que

$$AX = \lambda X.$$

Recordamos que una matriz $A \in \mathcal{M}_n(\mathbb{K})$ tiene exactamente n valores propios (quizás múltiples y probablemente complejos aunque la matriz sea de coeficientes reales), que son las n raíces del *polinomio característico* de la matriz (determinante formal)

$$p(\lambda) = \det(A - \lambda I) = (-1)^n (\lambda^n + a_1 \lambda^{n-1} + \cdots + a_{n-1} \lambda + a_n) \in \mathcal{P}_n[\lambda].$$

En consecuencia los *métodos de cálculo* de los valores propios de una matriz deben de ser necesariamente *iterativos*, pues en caso contrario se contradice la teoría de Galois en cuanto $n \geq 5$.

En ciertas ocasiones nos centramos sólo en el cálculo de los valores propios de una matriz, ya que, una vez obtenidos éstos, los vectores propios se pueden calcular resolviendo el sistema lineal homogéneo (con solución no trivial)

$$(A - \lambda I)X = 0.$$

En muchas aplicaciones prácticas, no interesa calcular *todos* los valores propios (y sus vectores propios asociados), sino solamente *unos cuantos*, normalmente el de módulo mayor o menor, o el más próximo a un valor dado.

Los *métodos de cálculo de los valores propios* se pueden dividir en dos tipos:

- *métodos clásicos*: son aquellos que calculan previamente el polinomio característico (mediante un método directo), para posteriormente calcular sus raíces (mediante un método iterativo, como el de Newton, de Bernoulli, etc.); actualmente estos métodos están en desuso;
- *métodos modernos*: son aquellos que calculan directamente algunos o todos los valores propios, mediante un método iterativo, sin pasar previamente por el cálculo del polinomio característico.

En general, no parece que sea una *buena táctica* la de los métodos clásicos debido a la *acumulación de errores*. El siguiente *ejemplo*, debido a *Wilkinson*, pone de manifiesto esta observación.

Consideremos una matriz $A \in \mathcal{M}_{20}(\mathbb{K})$ cuyo polinomio característico sea

$$p(\lambda) = \lambda^{20} - 210\lambda^{19} + \cdots + 20! = (\lambda - 1)(\lambda - 2) \cdots (\lambda - 20).$$

Evidentemente sus valores propios son $\lambda_i = i$, $i = 1, 2, \dots, 20$. Sin embargo, si en el cálculo del polinomio característico cometemos un pequeño error, de manera que obtenemos

$$\bar{p}(\lambda) = p(\lambda) - 2^{-23}\lambda^{19};$$

las raíces de este nuevo polinomio, redondeadas a 5 cifras decimales, son

$$\begin{aligned} \lambda_1 &= 1,00000, \\ \lambda_2 &= 2,00000, \\ \lambda_3 &= 3,00000, \\ \lambda_4 &= 4,00000, \\ \lambda_5 &= 5,00000, \\ \lambda_6 &= 6,00001, \\ \lambda_7 &= 6,99997, \\ \lambda_8 &= 8,00727, \\ \lambda_9 &= 8,91725, \\ \lambda_{10} &= 20,84691, \\ \lambda_{11,12} &= 10,09527 \pm 0,64350i, \\ \lambda_{13,14} &= 11,79363 \pm 1,65233i, \\ \lambda_{15,16} &= 13,99236 \pm 2,51883i, \\ \lambda_{17,18} &= 16,73074 \pm 2,81262i, \\ \lambda_{19,20} &= 19,50244 \pm 1,94033i. \end{aligned}$$

Como puede observarse, un pequeñísimo error en el cálculo del polinomio característico, ha provocado cambios sustanciales en el cálculo de los valores propios.

2. Localización de los valores propios

Proposición 2 (Teorema de Gerschgorin-Hadamard) *Los autovalores de una matriz $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ verifican que*

$$\text{sp}(A) \subset \bigcup_{i=1}^n \left\{ z \in \mathbb{C} : |z - a_{i,i}| \leq \sum_{j=1; j \neq i}^n |a_{i,j}| \right\}.$$

Una matriz $A = (a_{i,j}) \in \mathcal{M}_n(\mathbb{K})$ se dice de diagonal estrictamente dominante si

$$|a_{i,i}| > \sum_{j=1; j \neq i}^n |a_{i,j}|, \quad i = 1, 2, \dots, n;$$

una consecuencia del teorema anterior es que toda matriz de diagonal estrictamente dominante es necesariamente inversible pues el 0 no pertenece a la unión de las bolas dadas por el teorema de Gerschgorin-Hadamard.

3. Condicionamiento de un problema de valores propios

Analizamos como las *variaciones en los datos del problema* (en este caso los elementos de la matriz) pueden *afectar a los resultados* (los valores propios).

Consideremos el siguiente *ejemplo* que pone de manifiesto como matrices con elementos muy similares pueden tener valores propios diferentes. Sea

$$A(\varepsilon) = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & \varepsilon \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix} \in \mathcal{M}_{40}(\mathbb{K}).$$

Evidentemente, si $\varepsilon = 0$ el único valor propio es $\lambda = 0$ con multiplicidad 40.

Sin embargo, si $\varepsilon = 10^{-40}$ entonces el polinomio característico es

$$p(\lambda) = \lambda^{40} - 10^{-40},$$

por lo que los valores propios verifican

$$|\lambda| = 10^{-1};$$

como puede observarse, una mínima variación en los datos, se ha amplificado en los resultados con un factor de 10^{39} .

La situación es aún más grave si se piensa que algunos ordenadores tomarían 10^{-40} como 0.

Teorema 3 (Teorema de Bauer-Fike) Sea $A \in \mathcal{M}_n(\mathbb{K})$ diagonalizable, es decir, tal que existe $P \in \mathcal{M}_n(\mathbb{C})$ inversible con $P^{-1} A P = D = \text{diag}_{i=1,2,\dots,n} \{\lambda_i\}$. Sea $\|\cdot\|$ una norma matricial subordinada tal que $\|\text{diag}_{i=1,2,\dots,n} \{d_i\}\| = \max_{i=1,2,\dots,n} |d_i|$. Entonces, dada $\delta A \in \mathcal{M}_n(\mathbb{K})$, se verifica que

$$\text{sp}(A + \delta A) \subset \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - \lambda_i| \leq \text{cond}(P) \|\delta A\|\} .$$

■

La hipótesis hecha sobre la norma considerada, no es en la práctica restrictiva, pues las normas $\|\cdot\|_p$, $p = 1, 2, \infty$, la verifican.

Notamos que el condicionamiento de un problema de valores propios depende del condicionamiento de las matrices de paso a la forma diagonal, pero no del condicionamiento de la propia matriz A .

Definición 4 Se llama *condicionamiento de la matriz diagonalizable* $A \in \mathcal{M}_n(\mathbb{K})$, respecto del cálculo de sus valores propios, a

$$\Gamma(A) = \inf \{ \text{cond}(P) : P \in \mathcal{M}_n(\mathbb{C}) \text{ matriz de paso a la forma diagonal} \} .$$

Resulta entonces evidente que

$$\text{sp}(A + \delta A) \subset \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - \lambda_i| \leq \Gamma(A) \|\delta A\|\} .$$

Si la matriz A es normal, entonces es diagonalizable con matriz de paso unitaria, por lo que $\Gamma_2(A) = 1$, ya que $\text{cond}_2(P) = 1$ para P unitaria, y ésta es la mejor situación posible; en consecuencia para matrices normales podemos escribir que

$$\text{sp}(A + \delta A) \subset \bigcup_{i=1}^n \{z \in \mathbb{C} : |z - \lambda_i| \leq \|\delta A\|_2\} .$$

Observación 5 Si las matrices A y $A + \delta A$ son simétricas-hermíticas (por lo que sus valores propios son reales), se puede probar que en cada una de las bolas cerradas del Teorema 3 existe un valor propio de la matriz perturbada. ■

4. Método de Leverrier

El *método de Leverrier* es un método directo para calcular los coeficientes del polinomio característico de una matriz.

Se *basa* en las conocidas *fórmulas de Newton* (que relacionan los coeficientes de un polinomio con las sumas y productos de sus raíces) y en las propiedades de los valores propios.

Descripción del método.

Sea $A \in \mathcal{M}_n(\mathbb{K})$ y sean $\lambda_i \in \text{sp}(A)$, $i = 1, 2, \dots, n$, sus valores propios. Su polinomio característico

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

lo escribimos como

$$p(\lambda) = (-1)^n (\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + a_3 \lambda^{n-3} + \dots + a_n);$$

como además

$$\begin{aligned} p(\lambda) &= (-1)^n \prod_{i=1}^n (\lambda - \lambda_i) \\ &= (-1)^n \left[\lambda^n - \left(\sum_{i=1}^n \lambda_i \right) \lambda^{n-1} + \left(\sum_{1 \leq i < j \leq n} \lambda_i \lambda_j \right) \lambda^{n-2} \right. \\ &\quad \left. - \left(\sum_{1 \leq i < j < k \leq n} \lambda_i \lambda_j \lambda_k \right) \lambda^{n-3} + \dots + (-1)^n \prod_{i=1}^n \lambda_i \right], \end{aligned}$$

identificando términos y utilizando las propiedades de los autovalores, se obtiene que

$$\begin{aligned} a_1 &= - \sum_{i=1}^n \lambda_i = -\text{tr}(A), \\ a_2 &= \sum_{1 \leq i < j \leq n} \lambda_i \lambda_j = \frac{1}{2} (\text{tr}(A)^2 - \text{tr}(A^2)), \\ a_3 &= - \sum_{1 \leq i < j < k \leq n} \lambda_i \lambda_j \lambda_k = \frac{1}{6} (3 \text{tr}(A) \text{tr}(A^2) - 2 \text{tr}(A^3) - \text{tr}(A)^3), \end{aligned}$$

y así sucesivamente.

5. Método de Krilov

El *método de Krilov* es un método directo para calcular los coeficientes del polinomio característico de una matriz.

Se *basa* en el *Teorema de Cayley-Hamilton* que dice que si

$$p(\lambda) = (-1)^n (\lambda^n + a_1 \lambda^{n-1} + a_2 \lambda^{n-2} + a_3 \lambda^{n-3} + \cdots + a_n);$$

es el polinomio característico de una matriz $A \in \mathcal{M}_n(\mathbb{K})$, entonces

$$p(A) = (-1)^n (A^n + a_1 A^{n-1} + a_2 A^{n-2} + a_3 A^{n-3} + \cdots + a_n I) = 0.$$

Descripción del método.

Sea $X \in \mathbb{K}^n - \{0\}$ tal que los vectores $\{A^i X\}_{i=0}^{n-1}$ sean linealmente independientes, es decir formen una base de \mathbb{K}^n (esto no tiene por qué ser cierto).

Entonces del Teorema de Cayley-Hamilton se deduce que

$$A^n X + a_1 A^{n-1} X + a_2 A^{n-2} X + a_3 A^{n-3} X + \cdots + a_n X = 0,$$

de donde

$$a_1 A^{n-1} X + a_2 A^{n-2} X + a_3 A^{n-3} X + \cdots + a_n X = -A^n X;$$

es decir, podemos calcular los coeficientes del polinomio característico $a_i \in \mathbb{K}$, $i = 1, 2, \dots, n$, que son las coordenadas del vector $-A^n X$ respecto de la base anterior, resolviendo un sistema lineal.

6. Método de la potencia

El método de la potencia es un método iterativo muy simple que permite calcular una aproximación del valor propio de módulo mayor de una matriz, así como del vector propio asociado. Determinadas variantes de este método, como el de la potencia inversa con o sin desplazamiento, permiten también obtener aproximaciones de otros valores propios (el de módulo menor, el más próximo a un valor dado, etc.).

Un ejemplo.

Consideremos la matriz

$$A = \begin{pmatrix} 10 & 0 \\ -9 & 1 \end{pmatrix}.$$

Sus valores y vectores propios son:

Valores propios	$\lambda_1 = 1$	$\lambda_2 = 10$
Ecuaciones cartesianas de los vectores propios	$\{x_1 = 0\}$	$\{x_1 + x_2 = 0\}$
Vector propio	$X_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$X_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

En consecuencia esta matriz es diagonalizable, siendo

$$B^{-1} A B = D \iff \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 10 & 0 \\ -9 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 10 \end{pmatrix}.$$

La “idea” del método de la potencia es tomar un vector $X_0 \in \mathbb{R}^2$ “cualquiera” y construir la sucesión $X_1 = A X_0$, $X_2 = A X_1 = A^2 X_0$ y, en general, $X_k = A X_{k-1} = A^k X_0$. Así, por ejemplo, obtenemos

$$X_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad X_1 = \begin{pmatrix} 20 \\ -17 \end{pmatrix}, \quad X_2 = \begin{pmatrix} 200 \\ -197 \end{pmatrix}, \quad X_3 = \begin{pmatrix} 2000 \\ -1997 \end{pmatrix},$$

$$X_4 = \begin{pmatrix} 20000 \\ -19997 \end{pmatrix}, \quad X_5 = \begin{pmatrix} 200000 \\ -199997 \end{pmatrix}, \quad \dots, \quad X_k = \begin{pmatrix} 2 \times 10^k \\ 3 - 2 \times 10^k \end{pmatrix}, \quad \dots$$

Puede observarse entonces que:

- la dirección de X_k “tiende” hacia la de los vectores propios asociados a λ_2 ;
- $X_{k+1} \approx 10 X_k \iff X_{k+1} \approx \lambda_2 X_k$;
- $\|X_{k+1}\| \approx |\lambda_2| \|X_k\|$;
- se puede fácilmente alcanzar un overflow.

Algoritmo del método de la potencia.

Sea $A \in \mathcal{M}_n(\mathbb{K})$.

Se toma $X_0 \in \mathbb{C}^n - \{0\}$ unitario “cualquiera.”

Para $k = 0, 1, 2, \dots$, se hace

$$\begin{aligned} Y_{k+1} &= A X_k, \\ \lambda_{k+1}^j &= \frac{(Y_{k+1})_j}{(X_k)_j} \text{ para } j = 1, 2, \dots, n, \text{ supuesto que } (X_k)_j \neq 0, \\ X_{k+1} &= \frac{1}{\|Y_{k+1}\|} Y_{k+1}. \end{aligned}$$

Como puede observarse, con el objeto de evitar el overflow, en cada iteración se realiza un proceso de normalización de manera que todos los vectores de la sucesión $\{X_k\}_{k \in \mathbb{N}}$ son unitarios.

Según lo intuitivo por el ejemplo anterior, las sucesiones $\{\lambda_k^j\}_{k \in \mathbb{N}}$, $j = 1, 2, \dots, n$, formadas por los cocientes de las componentes j -ésimas de los vectores $Y_{k+1} = A X_k$ y X_k , deben converger hacia el valor propio de módulo mayor.

Convergencia del método de la potencia.

Teorema 6 Sea $A \in \mathcal{M}_n(\mathbb{K})$ diagonalizable con $\text{sp}(A) = \{\lambda_i : i = 1, 2, \dots, n\}$ y sea $\{Z_i\}_{i=1}^n$ una base unitaria de vectores propios, donde Z_1 está asociado a λ_1 , siendo $|\lambda_1| > |\lambda_i|$, $i = 2, 3, \dots, n$. Sea $X_0 \in \mathbb{C}^n - \{0\}$ unitario tal que $X_0 = \sum_{i=1}^n \alpha_i Z_i$ con $\alpha_1 \neq 0$, y a partir de él construimos la sucesión $\{X_k\}_{k \in \mathbb{N}}$ mediante el algoritmo del método de la potencia. Entonces:

1. $\lim_{k \rightarrow \infty} \left(\frac{\bar{\lambda}_1}{|\lambda_1|} \right)^k X_k$ existe y es un vector propio asociado a λ_1 ;
2. $\lim_{k \rightarrow \infty} \|A X_k\| = |\lambda_1|$;
3. $\lim_{k \rightarrow \infty} \frac{(Y_{k+1})_j}{(X_k)_j} = \lambda_1$, $j = 1, 2, \dots, n$, si $(X_k)_j \neq 0$.

Corolario 7 Bajo las mismas hipótesis del teorema anterior, suponiendo ahora que el autovalor de módulo mayor tiene multiplicidad $p > 1$, es decir, $\lambda_1 = \dots = \lambda_p$, con $|\lambda_1| > |\lambda_i|$, $i = p+1, \dots, n$, y con $X_0 = \sum_{i=1}^n \alpha_i Z_i$ con $|\alpha_1| + |\alpha_2| + \dots + |\alpha_p| > 0$, se obtienen los mismos resultados.

En el caso de matrices no diagonalizables, pero con valor propio dominante único (excepto multiplicidad) se obtienen resultados análogos pero en la demostración hay que recurrir a la forma de Jordan de la matriz.

En el teorema se demuestra que $\left(\frac{\bar{\lambda}_1}{|\lambda_1|} \right)^k X_k$ converge (hacia un vector propio asociado a λ_1), pero excepto si $\lambda_1 > 0$, en cuyo caso $\left(\frac{\bar{\lambda}_1}{|\lambda_1|} \right)^k = 1$, la sucesión $\{X_k\}_{k \in \mathbb{N}}$ no converge.

Se puede probar que las sucesiones anteriores convergen con orden 1 y que la constante de reducción asintótica del error es $\left| \frac{\lambda_2}{\lambda_1} \right|$ ($\left| \frac{\lambda_{p+1}}{\lambda_1} \right|$ en el caso de multiplicidad p), de donde el interés de que el valor propio dominante esté “aislado.”

Método de la potencia inversa.

Supuesta $A \in \mathcal{M}_n(\mathbb{K})$ inversible (además de otras hipótesis necesarias para la convergencia), el *método de la potencia inversa* consiste en aplicar el método de la potencia a A^{-1} , con lo que calculamos el valor propio de módulo mayor de A^{-1} que es el inverso del valor propio de módulo menor de A .

El algoritmo correspondiente es:

Se toma $X_0 \in \mathbb{C}^n - \{0\}$ unitario “cualquiera.”

Para $k = 0, 1, 2, \dots$, se hace

$$\begin{aligned} Y_{k+1} &= A^{-1} X_k \iff AY_{k+1} = X_k, \\ \lambda_{k+1}^j &= \frac{(Y_{k+1})_j}{(X_k)_j} \text{ para } j = 1, 2, \dots, n, \text{ supuesto que } (X_k)_j \neq 0, \\ X_{k+1} &= \frac{1}{\|Y_{k+1}\|} Y_{k+1}. \end{aligned}$$

Como puede comprobarse, en cada iteración debemos resolver un sistema lineal (si bien todos ellos tienen la misma matriz, por lo que se puede realizar sólo una factorización LU , Cholesky, etc.).

Método de la potencia desplazada.

Dada $A \in \mathcal{M}_n(\mathbb{K})$, el *método de la potencia desplazada* consiste en aplicar el método de la potencia a $A - \mu I$, con lo que calculamos el valor propio de módulo mayor de $A - \mu I$; recordamos que los autovalores de $A - \mu I$ son los autovalores de A menos μ .

Por último, se pueden combinar el método de la potencia con el de la potencia desplazada, para calcular otros valores (y vectores) propios de la matriz, en particular el valor propio más próximo a un valor dado.

7. Método de Jacobi

Se trata de un método iterativo para calcular todos los autovalores y autovectores de una matriz simétrica.

Idea de base

Sea $A \in \mathcal{M}_n(\mathbb{R})$ simétrica; por tanto es diagonalizable con matriz de paso ortogonal, es decir, existe $O \in \mathcal{M}_n(\mathbb{R})$ ortogonal tal que

$$O^t A O = D, \quad \text{siendo} \quad D = \text{diag}\{\lambda_i: i = 1, 2, \dots, n\} \quad \text{con} \quad \lambda_i \in \text{sp}(A)$$

además las columnas de O son los autovectores de A , que por tanto forman una base ortonormal. En resumen, A es semejante (por transformación unitaria) a D , por lo que

$$\|A\|_E = \|D\|_E.$$

La idea es ir haciendo cero en posiciones no diagonales de A mediante sucesivas transformaciones elementales unitarias, es decir,

$$\begin{aligned} A_1 &= A, \\ A_2 &= \Omega_1^t A_1 \Omega_1 \quad \text{con} \quad \Omega_1 \in \mathcal{M}_n(\mathbb{R}) \quad \text{ortogonal}, \\ &\dots \\ A_{k+1} &= \Omega_k^t A_k \Omega_k \quad \text{con} \quad \Omega_k \in \mathcal{M}_n(\mathbb{R}) \quad \text{ortogonal}, \\ &\dots \end{aligned}$$

de manera que la sucesión $\{A_k\}_{k \in \mathbb{N}}$ converja hacia D ; como además

$$A_{k+1} = \Omega_k^t A_k \Omega_k = \dots = \Omega_k^t \dots \Omega_1^t A_1 \Omega_1 \dots \Omega_k = O_k^t A O_k,$$

siendo $O_k = \Omega_1 \dots \Omega_k$, se debería verificar también que la sucesión $\{O_k\}_{k \in \mathbb{N}}$ convergiera hacia O .

El paso de A_k a A_{k+1} , es decir la construcción de Ω_k , se hace imponiendo que A_{k+1} anule dos elementos simétricos no diagonales de A_k , es decir, siendo (p, q) (con $p \neq q$) tales que $a_{p,q}^k = a_{q,p}^k \neq 0$, entonces $a_{p,q}^{k+1} = a_{q,p}^{k+1} = 0$; evidentemente este proceso de “anular” elementos de A_k al pasar a A_{k+1} es reversible, en el sentido de que hay elementos nulos de A_k que pasan a ser no nulos en A_{k+1} (de no ocurrir esto se trataría de un método directo), pero la idea es que los “pasos hacia delante” sean más grandes que los “pasos hacia atrás” de manera que se consiga la convergencia buscada.

Proposición 8 Sean $1 \leq p < q \leq n$ y $\theta \in \mathbb{R}$; construimos

$$\Omega = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \cos \theta & \cdots & \sin \theta \\ & & \vdots & \ddots & \vdots \\ & & -\sin \theta & \cdots & \cos \theta \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}).$$

Entonces Ω es ortogonal. Si $A \in \mathcal{M}_n(\mathbb{R})$ es simétrica y definimos $B = \Omega^t A \Omega$, entonces B es también simétrica, con $\|B\|_E = \|A\|_E$. Además, si (p, q) son tales que $a_{p,q} = a_{q,p} \neq 0$, entonces existe un único $\theta \in (-\frac{\pi}{4}, \frac{\pi}{4}] - \{0\}$ tal que $b_{p,q} = b_{q,p} = 0$; concretamente θ es tal que

$$\cotg 2\theta = \frac{a_{q,q} - a_{p,p}}{2 a_{p,q}}$$

y se verifica también que

$$\sum_{i=1}^n b_{i,i}^2 = \sum_{i=1}^n a_{i,i}^2 + 2 a_{p,q}^2.$$

Observaciones:

- la matriz Ω es una matriz de rotación de ángulo θ en el plano de las p -ésimas y q -ésimas coordenadas;
- sólo las filas y columnas p y q cambian al pasar de A a B ;
- si bien $b_{p,q} = b_{q,p} = 0$, puede haber elementos nulos en A que son no nulos en B .

Algoritmo del método de Jacobi

Sea $A_1 = A$.

Construida A_k , $k = 1, 2, 3, \dots$, elegimos (estrategia para la elección) (p, q) (con $p < q$) tales que $a_{p,q}^k = a_{q,p}^k \neq 0$; se construye entonces

$$\Omega_k = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & \cos \theta_k & \cdots & \sin \theta_k \\ & & \vdots & \ddots & \vdots \\ & & -\sin \theta_k & \cdots & \cos \theta_k \\ & & & & \ddots & \\ & & & & & 1 \end{pmatrix} \in \mathcal{M}_n(\mathbb{R}),$$

donde $\theta_k \in (-\frac{\pi}{4}, \frac{\pi}{4}] - \{0\}$ es tal que

$$\cotg 2\theta_k = \frac{a_{q,q}^k - a_{p,p}^k}{2a_{p,q}^k},$$

y se define entonces $A_{k+1} = \Omega_k^t A_k \Omega_k$, de manera que $a_{p,q}^{k+1} = a_{q,p}^{k+1} = 0$.

Estrategias para la elección del par (p, q)

- Método de Jacobi clásico: consiste en elegir

$$|a_{p,q}^k| = \max_{i \neq j} |a_{i,j}^k|;$$

- método de Jacobi cíclico: consiste en elegir el par (p, q) de forma cíclica

$$(1, 2) \rightarrow (1, 3) \rightarrow \cdots \rightarrow (1, n) \rightarrow (2, 3) \rightarrow \cdots \rightarrow (n-1, n) \rightarrow (1, 2) \rightarrow \cdots$$

(si $a_{p,q}^k = 0$, se pasa al siguiente elemento; equivald'ía a tomar $\theta_k = 0$ o, lo que es lo mismo, $\Omega_k = I$) esta estrategia evita el tiempo de la búsqueda del par óptimo;

- método de Jacobi cíclico con umbral de precisión: es análogo al método anterior, pero se salta el par (p, q) si $|a_{p,q}^k| < \epsilon$, donde ϵ es el umbral de precisión que puede variar de un barrido a otro.

Convergencia del método de Jacobi

Teorema 9 *La sucesión $\{A_k\}_{k \in \mathbb{N}}$ de matrices obtenidas por el método de Jacobi clásico converge, y*

$$\lim_{k \rightarrow \infty} A_k = \text{diag}(\lambda_{\sigma(i)}),$$

donde σ es una permutación de $\{1, 2, \dots, n\}$.

Observación: para otras variantes del método de Jacobi (distintas estrategias) existen otros resultados análogos que demuestran la convergencia del método; igualmente para la matrices $O_k = \Omega_1 \cdots \Omega_k$, $k = 1, 2, 3, \dots$, se pueden probar resultados de convergencia hacia una matriz ortogonal cuyas columnas constituyen una base ortonormal de autovectores de la matriz A .