

4. Autovalores y autovectores

4.1 Conceptos básicos

Definición 4.1 [VARIEDAD LINEAL INVARIANTE]

Una variedad lineal V de \mathbf{K}^n se dice que es *invariante* para una aplicación A si cualquier vector $x \in V$ se transforma, mediante A , en otro vector de V

$$Ax \in V \quad \forall x \in V$$

Las variedades invariantes más simples son aquellas que tienen dimensión 1, por lo que una base está constituida por un único vector $v \neq 0$.

Podemos observar que, en ese caso, cualquier vector $x \in V$ puede expresarse de la forma $x = \alpha v$ con $\alpha \neq 0$ si $x \neq 0$ y que, por tratarse de una variedad invariante, ha de ser $Ax = \beta v$, pero entonces:

$$Ax = A\alpha v = \alpha Av = \beta v \implies Av = \frac{\beta}{\alpha}v = \lambda v$$

Definición 4.2 [AUTOVALORES Y AUTOVECTORES DE UNA MATRIZ]

Las variedades invariantes de dimensión 1 vienen determinadas por un vector $v \neq 0$ tal que $Av = \lambda v$.

Estos vectores reciben en nombre de *autovectores* o *vectores propios* de la matriz A y los correspondientes valores de λ reciben el nombre de *autovalores* o *valores propios* de A .

Proposición 4.1 *Si $v \neq 0$ es un autovector de la matriz A asociado al autovalor λ , cualquier vector v proporcional a él, es también autovector de A asociado a λ .*

Demostración. Por ser v un autovector asociado a λ se tiene que $Av = \lambda v$.

Si $x = \alpha v$ se verifica que

$$Ax = A\alpha v = \alpha Av = \alpha \lambda v = \lambda(\alpha v) = \lambda x \quad \blacksquare$$

La proposición anterior establece que lo que caracteriza a un autovector no es su norma ni su sentido, sino su dirección.

Proposición 4.2 *Autovectores asociados a autovalores diferentes son linealmente independientes*

Demostración. Sean u y v dos autovectores de la matriz A asociados a los autovalores λ y μ respectivamente con $\lambda \neq \mu$.

Si u y v fuesen linealmente dependientes se verificaría que $u = \alpha v$ con $\alpha \neq 0$. Entonces:

$$\lambda u = Au = A(\alpha v) = \alpha Av = \alpha \mu v = \mu(\alpha v) = \mu u$$

y por tanto $(\lambda - \mu)u = 0$, pero dado que $u \neq 0$ se tiene que $\lambda = \mu$ en contra de la hipótesis de que $\lambda \neq \mu$, lo que prueba el resultado. \blacksquare

Los autovalores de una matriz son, en general, números complejos. Sin embargo, las matrices hermíticas (en el caso real, las simétricas) tienen todos sus autovalores reales como prueba el siguiente teorema.

Teorema 4.3 *Los autovalores de una matriz hermítica son todos reales y autovectores correspondientes a dos autovalores diferentes son ortogonales.*

Demostración. Sea A una matriz hermítica, es decir, una matriz tal que $A^* = A$ y sea λ un autovalor de A asociado al autovector x . Se verifica entonces que $Ax = \lambda x$.

Multiplicando la expresión anterior, por la izquierda por x^* obtenemos que

$$x^* Ax = x^* \lambda x = \lambda x^* x \quad (4.1)$$

Trasponiendo y conjugando la expresión (4.1) obtenemos

$$(x^* Ax)^* = (\lambda x^* x)^* \implies x^* A^* x = \bar{\lambda} x^* x \implies x^* Ax = \bar{\lambda} x^* x \quad (4.2)$$

Si comparamos las expresiones (4.1) y (4.2) obtenemos que

$$\lambda x^* x = \bar{\lambda} x^* x$$

y dado que x es un autovector (un vector no nulo) sabemos que $x^*x = \|x\|^2 \neq 0$, por lo que podemos dividir por x^*x para obtener que $\lambda = \bar{\lambda}$, es decir, $\lambda \in \mathbf{R}$.

Por otra parte, si x e y son autovectores asociados a dos autovalores $\lambda \neq \mu$ de una matriz hermítica A se verifica que

$$Ax = \lambda x \implies (Ax)^* = (\lambda x)^* \implies x^*A = \lambda x^*$$

$$x^*Ay = \lambda x^*y \implies x^*\mu y = \lambda x^*y \implies \mu x^*y = \lambda x^*y \implies (\lambda - \mu)x^*y = 0$$

y dado que $\lambda \neq \mu \implies \lambda - \mu \neq 0$ obtenemos que

$$x^*y = 0 \iff x \perp y \quad \blacksquare$$

Teorema 4.4 *Si A una matriz regular y λ un autovalor suyo asociado al autovector v , se verifica que $1/\lambda$ es un autovalor de A^{-1} asociado al mismo autovector v .*

Demostración. Si λ es un autovalor de A asociado a v sabemos que $Av = \lambda v$. Al tratarse de una matriz invertible ($\det(A) \neq 0$) sus autovalores son todos no nulos, por lo que podemos dividir por λ y multiplicar por A^{-1} la última igualdad para obtener que $A^{-1}v = \frac{1}{\lambda}v$ es decir, $1/\lambda$ es un autovalor de A^{-1} asociado a v . ■

Teorema 4.5 *Si λ es un autovalor de una matriz A asociado a un autovector v y α una constante cualquiera se verifica que $\lambda - \alpha$ es un autovalor de la matriz $A - \alpha I$ asociado al mismo autovector v .*

Demostración. Sabemos, por hipótesis, que $Av = \lambda v$, por lo que $Av - \alpha v = \lambda v - \alpha v$ y, por tanto,

$$(A - \alpha I)v = (\lambda - \alpha)v$$

es decir, $\lambda - \alpha$ es un autovalor de $A - \alpha I$ asociado a v . ■

Teorema 4.6 *Si λ es un autovalor de una matriz A asociado a un autovector v y α (una constante cualquiera) no es autovalor de A entonces $1/(\lambda - \alpha)$ es un autovalor de la matriz $(A - \alpha I)^{-1}$ asociado al autovector v .*

La demostración se basa en los Teoremas 4.4 y 4.5 y se deja al lector.

Sea A una matriz cuadrada de orden n y supongamos que existen V_1, V_2, \dots, V_k subespacios invariantes tales que $\mathbf{K}^n = V_1 \oplus V_2 \oplus \dots \oplus V_k$.

Si $\{x_{i1}, x_{i2}, \dots, x_{in_i}\}$ es una base de V_i se tiene que

$$\mathcal{B} = \{x_{11}, \dots, x_{1n_1}, x_{21}, \dots, x_{2n_2}, \dots, x_{k1}, \dots, x_{kn_k}\}$$

constituye una base de \mathbf{K}^n . Si P es la matriz del cambio de base de la base canónica a la base \mathcal{B} , se tiene que

$$P^{-1}AP = \begin{pmatrix} J_1 & \Theta & \cdots & \Theta \\ \Theta & J_2 & \cdots & \Theta \\ \vdots & \vdots & \ddots & \vdots \\ \Theta & \Theta & \cdots & J_k \end{pmatrix}$$

donde cada J_i es una caja cuadrada de dimensión $r_i = \dim(V_i)$.

La justificación es que $Ax_{ij} = (0, \dots, 0, \alpha_{i1}, \dots, \alpha_{in_i}, 0, \dots, 0)_B^T$ con $1 \leq i \leq k$, $1 \leq j \leq n_i$ y, por tanto, puede verse fácilmente que

$$AP = P \begin{pmatrix} J_1 & \Theta & \cdots & \Theta \\ \Theta & J_2 & \cdots & \Theta \\ \vdots & \vdots & \ddots & \vdots \\ \Theta & \Theta & \cdots & J_k \end{pmatrix}$$

Definición 4.3 [MATRIZ DIAGONALIZABLE]

Una matriz A se dice que es *diagonalizable*, si existe una matriz P , llamada *matriz de paso* tal que $P^{-1}AP = D$ siendo $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ la matriz diagonal cuyos elementos diagonales son los autovalores de la matriz A .

Es obvio que si A posee un autovector es porque existe un vector $v \neq 0$ tal que $Av = \lambda v$, o lo que es lo mismo, tal que $(\lambda I - A)v = 0$. Por tanto, el sistema $(\lambda I - A)x = 0$ es compatible y además indeterminado (infinitas soluciones), ya que si $v \neq 0$ es solución del sistema, cualquier vector proporcional a él también lo es.

Se verifica entonces que $\text{rg}(\lambda I - A) < n$ (donde n representa el orden de la matriz) y, por tanto $\det(\lambda I - A) = 0$.

Definición 4.4 [POLINOMIO CARACTERÍSTICO]

Al polinomio $p(\lambda) = \det(\lambda I - A)$ se le denomina *polinomio característico* de la matriz A y a la ecuación $p(\lambda) = 0$ *ecuación característica*.

Al ser $P(\lambda) = \det(\lambda I - A)$ su obtención conlleva el cálculo de un determinante de orden n que no es numérico, pues contiene al parámetro λ , lo cual no es fácil de realizar.

MÉTODO INTERPOLATORIO PARA LA OBTENCIÓN DEL POLINOMIO CARAC-
TERÍSTICO
$$\begin{array}{ll} \lambda = \lambda_1 & \implies \lambda_1^n + a_1 \lambda_1^{n-1} + \cdots + a_{n-2} \lambda_1^2 + a_{n-1} \lambda_1 + a_n = \det(\lambda_1 I - A) \\ \lambda = \lambda_2 & \implies \lambda_2^n + a_1 \lambda_2^{n-1} + \cdots + a_{n-2} \lambda_2^2 + a_{n-1} \lambda_2 + a_n = \det(\lambda_2 I - A) \\ & \dots\dots\dots \\ & \dots\dots\dots \\ \lambda = \lambda_n & \implies \lambda_n^n + a_1 \lambda_n^{n-1} + \cdots + a_{n-2} \lambda_n^2 + a_{n-1} \lambda_n + a_n = \det(\lambda_n I - A) \end{array}$$

La solución del sistema resultante nos proporciona los coeficientes del polinomio característico.

Ejemplo 4.1 Dada la matriz $A = \begin{pmatrix} 1 & 3 & 2 & 5 \\ 4 & 2 & -1 & 0 \\ 0 & 1 & 0 & 1 \\ 2 & -2 & -1 & 1 \end{pmatrix}$, su polinomio ca-

racterístico es de grado cuatro $P(\lambda) = \lambda^4 + a_1\lambda^3 + a_2\lambda^2 + a_3\lambda + a_4$, por lo que vamos a dar a λ cuatro valores:

$$\lambda = 0 \quad \Rightarrow \quad a_4 = \det(-A) = 41$$

$$\lambda = 1 \quad \Rightarrow \quad 1 + a_1 + a_2 + a_3 + a_4 = \det(I - A) = 74$$

$$\lambda = -1 \quad \Rightarrow \quad 1 - a_1 + a_2 - a_3 + a_4 = \det(-I - A) = -20$$

$$\lambda = 2 \quad \Rightarrow \quad 16 + 8a_1 + 4a_2 + 2a_3 + a_4 = \det(2I - A) = 67$$

dado que $a_4 = 41$ el sistema se reduce a

$$\left. \begin{array}{l} a_1 + a_2 + a_3 = 32 \\ -a_1 + a_2 - a_3 = -62 \\ 8a_1 + 4a_2 + 2a_3 = 10 \end{array} \right\} \Rightarrow a_1 = -4, a_2 = -15 \text{ y } a_3 = 51$$

por lo que su polinomio característico es

$$P(\lambda) = \lambda^4 - 4\lambda^3 - 15\lambda^2 + 51\lambda + 41$$

□

4.2 Matrices normales

Un ejemplo de matriz, de orden n , *no diagonalizable* lo constituyen las matrices de la forma

$$A = \left(\begin{array}{c|ccc} 0 & & & \\ 0 & & I_{n-1} & \\ 0 & & & \\ \hline 0 & 0 & 0 & 0 \end{array} \right)$$

cuyo polinomio característico es $P_A(\lambda) = \lambda^n$ y por tanto, con sus n autovalores nulos.

Si introducimos en la matriz una perturbación y en vez de la matriz A tomamos la matriz

$$B = A + E = \left(\begin{array}{c|ccc} 0 & & & \\ 0 & & I_{n-1} & \\ 0 & & & \\ \hline \varepsilon & 0 & 0 & 0 \end{array} \right)$$

siendo $\varepsilon > 0$ muy pequeño (incluso más pequeño que la resolución del ordenador) se obtiene como polinomio característico $P_B(\lambda) = \lambda^n + (-1)^n \varepsilon$ que posee n raíces distintas (las raíces n -ésimas de $(-1)^{n-1} \varepsilon$), por lo resultan muy sensibles a cualquier perturbación que se produzca en la matriz, por pequeña que ésta sea.

Nos preguntamos entonces,

¿Cómo afecta una pequeña perturbación a los autovalores de una matriz diagonalizable?

Teorema 4.7 Sean A una matriz diagonalizable, $B = A + E$ una perturbación de dicha matriz, λ_i los autovalores de la matriz A y μ uno cualquiera de los autovalores de la matriz B con $\lambda_i \neq \mu$.

En las condiciones anteriores se tiene que

$$\min_i |\mu - \lambda_i| \leq \|P\| \|P^{-1}\| \|E\| = \|E\| \kappa(P)$$

donde P representa a la matriz de paso que diagonaliza a la matriz A .

Demostración. Por ser A diagonalizable existe P tal que $P^{-1}AP = D$.

Sea x un autovector de B asociado a μ . Entonces $Bx = \mu x$ o lo que es lo mismo:

$$\begin{aligned} (A + E)x = \mu x &\implies (\mu I - A)x = Ex \implies (\mu I - PDP^{-1})x = Ex \implies \\ P(\mu I - D)P^{-1}x &= Ex \implies (\mu I - D)(P^{-1}x) = P^{-1}Ex = P^{-1}EPP^{-1}x \end{aligned}$$

Supuesto que $\mu \neq \lambda_i$ cualquiera que sea $i = 1, 2, \dots, n$, la matriz $\mu I - D$ es regular, por lo que

$$P^{-1}x = (\mu I - D)^{-1}(P^{-1}EP)P^{-1}x$$

y para cualquier norma multiplicativa se tiene

$$\|P^{-1}x\| \leq \|(\mu I - D)^{-1}\| \|P^{-1}EP\| \|P^{-1}x\|$$

es decir

$$1 \leq \|(\mu I - D)^{-1}\| \|P^{-1}EP\|$$

Dado que $(\mu I - D)^{-1} = \text{diag} \left(\frac{1}{\mu - \lambda_1}, \dots, \frac{1}{\mu - \lambda_n} \right)$ tenemos que

$$1 \leq \max_i \left\{ \left| \frac{1}{\mu - \lambda_i} \right| \right\} \|P^{-1}\| \|E\| \|P\|$$

por lo que

$$\min_i |\mu - \lambda_i| \leq \|P^{-1}\| \|P\| \|E\| = \|E\| \kappa(P) \quad \blacksquare$$

Obsérvese que la perturbación cometida en los autovalores depende del número de condición de la matriz de paso P y dado que esta no es única.

Trataremos de elegir, entre todas las posibles matrices de paso, aquella cuyo número de condición sea mínimo.

Corolario 4.8 *Si A es unitariamente diagonalizable (diagonalizable mediante una matriz de paso unitaria), la perturbación producida en los autovalores es menor o igual que la perturbación producida en la matriz.*

$$\min_i |\mu - \lambda_i| \leq \|E\|$$

Por tanto, las mejores matrices para el cálculo efectivo de sus autovalores y autovectores son las diagonalizables unitariamente y éstas reciben el nombre de *matrices normales*.

Es evidente que no podemos estudiar si una matriz es normal calculando sus autovalores y autovectores para comprobar que se puede diagonalizar por semejanza mediante una matriz de paso unitaria (matriz constituida por una base ortonormal de autovectores).

Es necesario, por tanto, encontrar alguna forma de detectar si una matriz es, o no es, normal *sin necesidad de calcular sus autovalores y autovectores* para ver si es posible diagonalizarla mediante una matriz de paso unitaria.

CARACTERIZACIÓN DE LAS MATRICES NORMALES

Proposición 4.9 *Sea T una matriz triangular. Si $T^*T = TT^*$ entonces T es diagonal.*

Demostración. Probaremos que se trata de una matriz diagonal por inducción en el orden de la matriz.

Para $n = 1$ es obvio. Para $n = 2$ es $T = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \implies T^* = \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix}$.

$$T^*T = \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = \begin{pmatrix} |a|^2 & \bar{a}b \\ a\bar{b} & |b|^2 + |c|^2 \end{pmatrix}$$

$$TT^* = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} \bar{a} & 0 \\ \bar{b} & \bar{c} \end{pmatrix} = \begin{pmatrix} |a|^2 + |b|^2 & b\bar{c} \\ \bar{b}c & |c|^2 \end{pmatrix}$$

Dado que $T^*T = TT^*$, igualando ambas matrices se obtiene que $|b|^2 = 0$ es decir $b = 0$ y, por tanto, T es diagonal.

Supongamos ahora que el teorema es cierto para cualquier matriz triangular con $T^*T = TT^*$ de orden n y vamos a probarlo para otra de orden $n + 1$.

$$\text{Sean } T = \left(\begin{array}{c|ccc} a_1 & a_2 & \cdots & a_{n+1} \\ \hline 0 & & & \\ \vdots & & T_n & \\ 0 & & & \end{array} \right) \text{ y } T^* = \left(\begin{array}{c|ccc} \bar{a}_1 & 0 & \cdots & 0 \\ \hline \bar{a}_2 & & & \\ \vdots & & T_n^* & \\ \bar{a}_{n+1} & & & \end{array} \right)$$

$$TT^* = \left(\begin{array}{c|c} \sum_{i=1}^{n+1} |a_i|^2 & \\ \hline & T_n T_n^* \end{array} \right) \quad T^*T = \left(\begin{array}{c|c} |a_1|^2 & \\ \hline & T_n^* T_n \end{array} \right)$$

De la igualación de ambas obtenemos que

$$TT^* = T^*T \implies |a_2|^2 + \cdots + |a_{n+1}|^2 = 0 \implies a_2 = a_3 = \cdots = a_{n+1} = 0$$

Como, además, es $T_n T_n^* = T_n^* T_n$, por hipótesis de inducción sabemos que T_n es diagonal y, por tanto, T es diagonal. ■

Teorema 4.10 [TEOREMA DE SCHUR]

Cualquier matriz cuadrada A es unitariamente semejante a una triangular superior T . Es decir, existe una unitaria U tal que

$$U^* A U = T.$$

Demostración. Aplicamos inducción en el orden de la matriz A .

Si $n = 1$ es obvio. Supuesto cierto para n vamos a probarlo para una matriz de orden $n + 1$.

Sean $A \in \mathbf{K}^{(n+1) \times (n+1)}$, λ un autovalor de A y x un autovector asociado a λ con $\|x\| = 1$.

Consideremos la matriz $P = (x \ e_2 \cdots e_n)$ en la que sus columnas x, e_2, \dots, e_n constituyen una base ortonormal de \mathbf{K}^{n+1}

$$P^* A P = \left(\begin{array}{c|c} \lambda & \alpha_{ij} \\ \hline \Theta & A_n \end{array} \right)$$

Sea $Q = \left(\begin{array}{c|c} 1 & \Theta \\ \hline \Theta & U_n \end{array} \right)$ en donde U_n es la matriz unitaria que, por hipótesis de inducción, verifica que $U_n^* A_n U_n = \text{Triangular superior}$.

Si consideremos la matriz $U = PQ$ es fácil comprobar que $U^* A U = Q^* P^* A P Q$ es una triangular superior. ■

Teorema 4.11 [CARACTERIZACIÓN DE LAS MATRICES NORMALES]

Una matriz A es normal si, y sólo si, $AA^ = A^*A$.*

Demostración. Supongamos que $AA^* = A^*A$. Por el Teorema 4.10 sabemos que existe una matriz unitaria U tal que $U^* A U = T$ con T triangular superior. Entonces $A = UTU^*$, por lo que

$$\left. \begin{aligned} AA^* &= (UTU^*)(UTU^*)^* = UTU^*UT^*U^* = UTT^*U^* \\ A^*A &= (UTU^*)^*(UTU^*) = UT^*U^*UTU^* = UT^*TU^* \end{aligned} \right\} \implies TT^* = T^*T$$

por lo que al ser T una matriz normal y triangular, la Proposición 4.9 nos asegura que es diagonal y, por tanto, A es diagonalizable unitariamente, es decir, es normal.

Recíprocamente, si A es unitariamente diagonalizable (normal) existe una matriz unitaria U tal que $U^* A U = D$ o lo que es lo mismo, $A = UDU^*$.

$$\left. \begin{aligned} AA^* &= UDU^*(UDU^*)^* = UDD^*U^* = U \text{diag}(|d_1|^2, \dots, |d_n|^2)U^* \\ A^*A &= (UDU^*)^*UDU^* = UD^*DU^* = U \text{diag}(|d_1|^2, \dots, |d_n|^2)U^* \end{aligned} \right\} \implies AA^* = A^*A \quad \blacksquare$$

¿Cómo podemos prever el comportamiento de los autovalores ante una perturbación cuando la matriz no es normal?

En otras palabras, ¿cómo estudiar el *condicionamiento* de una matriz para el cálculo de sus autovalores?

• DESVIACIÓN DE LA NORMALIDAD

Por el Teorema 4.10 cualquier matriz cuadrada A es unitariamente semejante a una triangular superior T donde

$$T = \begin{pmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix} = \begin{pmatrix} t_{11} & 0 & \cdots & 0 \\ 0 & t_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & t_{nn} \end{pmatrix} + \begin{pmatrix} 0 & t_{12} & \cdots & t_{1n} \\ 0 & 0 & \cdots & t_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}$$

Es decir, $T = D + M$ y, obviamente, A es normal si, y sólo si, $M = \Theta$.

La norma de la matriz M recibe el nombre de *desviación de la normalidad* y nos va a medir el condicionamiento del problema.

Además, $\|M\|_2$ no depende de U ni de T sino sólo de A ya que

$$\|M\|_2^2 = \|T\|_2^2 - \|D\|_2^2 = \|U^*AU\|_2^2 - \|D\|_2^2 = \|A\|_2^2 - \sum_{i=1}^n |\lambda_i|^2 \implies$$

$$\|M\|_2 = \sqrt{\|A\|_2^2 - \sum_{i=1}^n |\lambda_i|^2}$$

- **MATRIZ CONMUTATRIZ**

Hemos visto que A es normal si, y sólo si,

$$A^*A = AA^* \iff A^*A - AA^* = \Theta$$

Se denomina *matriz conmutatriz* de una matriz cuadrada A a la matriz

$$C(A) = A^*A - AA^*$$

Evidentemente, A no es normal si $A^*A - AA^* \neq \Theta$.

La matriz $C(A) = A^*A - AA^*$ se le llama *matriz conmutatriz* y su norma euclídea también nos permite medir el condicionamiento del problema.

Ambas formas de medir el condicionamiento están relacionadas como muestra el siguiente resultado:

Teorema 4.12 *Dada una matriz $A \in \mathbf{K}^{n \times n}$ se verifica:*

- $\frac{\|C(A)\|_2^2}{6\|A\|_2^2} \leq \|M\|_2^2$ *Desigualdad de Eberlein.*
- $\|M\|_2^2 \leq \sqrt{\frac{n^3 - 3}{12}} \|C(A)\|_2^2$ *Desigualdad de Heurici.*

Teorema 4.13 [TEOREMA ESPECTRAL PARA MATRICES HERMÍTICAS]

Sea A una matriz hermítica de orden n . Existe una base de \mathbf{C}^n constituida por autovectores de A . Dicho de otra forma, toda matriz hermítica es diagonalizable por semejanza (es normal).

4.3 Métodos iterados para la obtención de autovalores y autovectores

Los métodos que veremos a continuación tratan de encontrar una sucesión convergente cuyo límite nos permitirá conocer los autovalores y autovectores de una matriz dada. El primero que estudiaremos consiste en buscar una sucesión de vectores que convergen a la dirección del autovector asociado al autovalor de mayor módulo de la matriz dada.

Vamos a comenzar, por ello, a estudiar cómo podemos aproximar el autovalor correspondiente a un determinado autovector cuando lo que conocemos es una aproximación de éste.

COCIENTE DE RAYLEIGH

Si nos limitamos a calcular la sucesión (x_n) hasta obtener una aproximación x adecuada del autovector podemos obtener el autovalor resolviendo el sistema vectorial $\lambda x = Ax$. Este sistema resulta, en general, incompatible por no ser x exactamente un autovector sino sólo una aproximación, por lo que la solución que mejor se ajusta es la pseudosolución del sistema, que nos dará una aproximación del autovalor λ .

$$\lambda x^*x = x^*Ax \implies \lambda = \frac{x^*Ax}{x^*x}$$

Definición 4.5 [COCIENTE DE RAYLEIGH]

Dada una matriz cuadrada A y un vector x , se denomina *cociente de Rayleigh* del vector x respecto de la matriz A al cociente $\frac{x^*Ax}{x^*x}$.

Podemos, por tanto, obtener una aproximación del autovector por un método iterado para más tarde aproximar el autovalor mediante el cociente de Rayleigh.

4.3.1 Método de la potencia simple y variantes

MÉTODO DE LA POTENCIA SIMPLE

Sea $A \in \mathbf{R}^{n \times n}$ una matriz diagonalizable y supongamos que sus autovalores verifican que:

$$|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$$

en cuyo caso diremos que λ_1 es el *autovalor dominante* de la matriz.

Sea $\mathcal{B} = \{x_1, x_2, \dots, x_n\}$ una base de \mathbf{R}^n formada por autovectores asociados a $\lambda_1, \lambda_2, \dots, \lambda_n$ respectivamente. Se verifica entonces que

$$A^2 x_i = A(Ax_i) = A(\lambda_i x_i) = \lambda_i Ax_i = \lambda_i(\lambda_i x_i) = \lambda_i^2 x_i$$

por lo que es fácil probar, por inducción, que

$$A^k x_i = \lambda_i^k x_i \quad \text{para cualquier } i = 1, 2, \dots, n$$

Dado un vector $z_0 \in \mathbf{R}^n$ se define, a partir de él, la sucesión (z_n) con

$$z_n = Az_{n-1} = A^2 z_{n-2} = \dots = A^n z_0.$$

Si las coordenadas del vector z_0 respecto de la base \mathcal{B} son $(\alpha_1, \alpha_2, \dots, \alpha_n)$ se tiene que $z_0 = \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n$, por lo que

$$\begin{aligned} z_k &= A^k z_0 = A^k(\alpha_1 x_1 + \dots + \alpha_n x_n) = \alpha_1 A^k x_1 + \dots + \alpha_n A^k x_n = \\ &= \lambda_1^k \alpha_1 x_1 + \dots + \lambda_n^k \alpha_n x_n = \lambda_1^k \left[\alpha_1 x_1 + \sum_{i=2}^n \left(\frac{\lambda_i}{\lambda_1} \right)^k \alpha_i x_i \right] \end{aligned}$$

Dado que $|\lambda_1| > |\lambda_i|$ se tiene que $\lim_{k \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_1} \right)^k = 0 \quad \forall i = 2, 3, \dots, n$.

Se verifica entonces que $\lim_{k \rightarrow \infty} \frac{z_k}{\lambda_1^k} = \alpha_1 x_1$ que es un autovector de A asociado al autovalor λ_1 .

Si k es suficientemente grande, se tiene

$$Az_k = z_{k+1} \approx \lambda_1^{k+1} \alpha_1 x_1 = \lambda_1(\lambda_1^k \alpha_1 x_1) = \lambda_1 z_k$$

por lo que la sucesión (z_n) nos proporciona un método para aproximar el autovalor λ_1 .

Teniendo en cuenta que, para valores de k suficientemente grandes se verifica que

$$Az_k = z_{k+1} \approx \lambda_1 z_k \tag{4.3}$$

el método no convergerá a un vector en concreto ya que

$$z_{k+1} \approx \lambda_1 z_k \neq z_k.$$

Ahora bien, dado que z_{k+1} y z_k tienen la misma dirección, ambos representan al autovector buscado.

Geométricamente la matriz A transforma al vector z_k (para valores de k suficientemente grandes, en otro proporcional a él y dado que se verifica la ecuación (4.3) ese factor de proporcionalidad es precisamente el autovalor buscado λ_1 .

ESCALADO

Si $|\lambda_1| > 1$ la sucesión (z_n) converge a una dirección determinada (la del autovector), pero las normas de los vectores van creciendo llegando a diverger a un vector de coordenadas infinitas. Si, por el contrario, $|\lambda_1| < 1$, la sucesión (z_n) converge al vector nulo. Es decir: en ninguno de los dos casos podremos calcular ni el autovalor ni el autovector.

Como sólo nos interesa la dirección y los vectores de la sucesión (z_n) convergen a una determinada dirección, podemos dividir o multiplicar en cada paso el vector z_k por una constante sin que ello modifique su dirección. Con ello conseguiremos que la sucesión (z_n) converja a un autovector asociado a λ_1 . Este proceso de modificar las normas de los vectores recibe el nombre de *escalado*.

La cantidad por la que se escala puede ser arbitraria, pero lo más útil es escalar por la coordenada de mayor módulo.

De esa forma los vectores escalados w_k y $w_{k+1} = \frac{z_k}{\|z_k\|_\infty}$ al ser muy próximos tendrán la coordenada de mayor módulo en la misma posición y además valdrán 1 por lo que se evita el problema de la divergencia a infinito o la convergencia al vector nulo. Además, dado que llamando α al factor de proporcionalidad se verifica que

$$\alpha w_k = z_k = A w_{k-1}$$

el factor de proporcionalidad α existente entre los vectores z_k y w_k es precisamente el autovalor λ_1 buscado. En otras palabras, escalando los vectores por su coordenada de mayor módulo, el proceso converge a un autovector cuya coordenada de mayor módulo (antes de escalar) es el autovalor buscado.

Ejemplo 4.2 Para calcular, por el método de la potencia simple, el autovalor de mayor valor absoluto de la matriz

$$A = \begin{pmatrix} 6 & 2 & 5 \\ 2 & 2 & 3 \\ 5 & 3 & 6 \end{pmatrix}$$

partiendo del vector $z_0 = (1 \ 1 \ 1)^T$ obtenemos:

$$\begin{aligned}
 z_1 &= \begin{pmatrix} 13.0000 \\ 7.0000 \\ 14.0000 \end{pmatrix} & z_2 &= \begin{pmatrix} 11.5714 \\ 5.8571 \\ 12.1429 \end{pmatrix} & z_3 &= \begin{pmatrix} 11.6824 \\ 5.8706 \\ 12.2118 \end{pmatrix} & z_4 &= \begin{pmatrix} 11.7013 \\ 5.8748 \\ 12.2254 \end{pmatrix} \\
 z_5 &= \begin{pmatrix} 11.7039 \\ 5.8753 \\ 12.2273 \end{pmatrix} & z_6 &= \begin{pmatrix} 11.7042 \\ 5.8754 \\ 12.2275 \end{pmatrix} & z_7 &= \begin{pmatrix} 11.7042 \\ 5.8754 \\ 12.2275 \end{pmatrix}
 \end{aligned}$$

El vector z_7 es una aproximación del autovector asociado al autovalor dominante. Su coordenada de mayor módulo es 12.2275, por lo que $\lambda_1 \approx 12.2275$.

No obstante, dado que no hemos obtenido el límite sino una aproximación, no todas sus coordenadas tendrán el mismo factor de proporcionalidad (variará en poco, pero variará), por lo que es más correcto decir: si z_7 es un autovector asociado a λ es porque $Az_7 = \lambda z_7$, sistema que, en general es incompatible, por lo que buscamos λ resolviendo el sistema superdeterminado en término de mínimos cuadrados, es decir, mediante el cociente de Rayleigh.

$$\lambda_1 = \frac{z_7^T A z_7}{z_7^T z_7} = 12.22753579693696.$$

□

ALGORITMO DE LA POTENCIA SIMPLE

```

z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z;
    z = A*w;
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z = z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)

```

Ejecutando este algoritmo con MatLab para la matriz del Ejemplo 4.2 obtenemos la aproximación 12.22753579693706 del autovalor con un error del orden de $1.986027322597819 \cdot 10^{-15}$ en 17 iteraciones.

Una forma de acelerar el proceso es la siguiente: una vez calculado el valor de w_n (vector z_n escalado) aproximar el autovector mediante el sistema homogéneo $(A - \lambda I)w_n = 0$ donde λ es la aproximación obtenida para el autovalor mediante el cociente de Rayleigh aplicado a la matriz A y el vector w_n . Esta aproximación del autovector es la que utilizaremos en el siguiente paso del método de la potencia simple.

Hay que tener cuidado ya que puede darse el caso de que al aproximar el autovalor mediante el cociente de Rayleigh (cuando el vector w_n está aún distante del autovector) nos acerquemos a otro autovalor de la matriz y, al final, el proceso nos conduzca a un autovalor diferente del dominante.

ALGORITMO DE LA POTENCIA SIMPLE MEJORADO

```

z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z;
    z = A*w;
    l=(z'*A*z)/(z'*z);
    if rcond(A-l*eye(length(A)))>10^(-15)
        z = (A-l*eye(length(A)))\z;
    end
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z= z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)

```


Ejecutando este algoritmo con MatLab para la matriz del Ejemplo 4.2 obtenemos la aproximación 12.22753579693706 del autovalor con un error del orden de $2.664535259100376 \cdot 10^{-15}$ en tan sólo 4 iteraciones, frente a las 17 que eran necesarias con algoritmo anterior.

Este método sólo nos permite calcular, en caso de existir, el autovalor dominante de una matriz. Existen sin embargo otras variantes del método de la potencia simple que nos permiten calcular cualquiera de sus autovalores a partir de una aproximación de ellos. Debido a la existencia de dichas variantes es por lo que el método estudiado anteriormente es conocido como *método de la potencia simple* para distinguirlo de los que estudiaremos a continuación.

MÉTODO DE LA POTENCIA INVERSA

Sea $A \in \mathbf{R}^{n \times n}$ una matriz diagonalizable regular para la que sus autovalores verifican que:

$$0 < |\lambda_1| < |\lambda_2| \leq \dots \leq |\lambda_n| \quad (4.4)$$

Los autovalores $\mu_1, \mu_2, \dots, \mu_n$ de la matriz A^{-1} son los inversos de los autovalores de A ,

$$\mu_i = \frac{1}{\lambda_i} \quad \text{para } i = 1, 2, \dots, n.$$

Invirtiendo la expresión (4.4) nos queda

$$\frac{1}{|\lambda_1|} > \frac{1}{|\lambda_2|} \geq \dots \geq \frac{1}{|\lambda_n|} \iff |\mu_1| > |\mu_2| \geq \dots \geq |\mu_n|$$

es decir, el autovalor de menor valor absoluto de la matriz A se corresponde con el de mayor valor absoluto de A^{-1} , por lo que aplicando el método de la potencia simple a la matriz A^{-1} obtenemos el autovalor de menor valor absoluto de la matriz A .

Obsérvese que debemos ir calculando, en cada paso, los vectores $z_n = A^{-1}\omega_{n-1}$ y $\omega_n = \frac{z_n}{\|z_n\|_\infty}$. Pues bien, el cálculo de z_n se realiza resolviendo, por alguno de los métodos estudiados, el sistema $Az_n = \omega_{n-1}$ lo que nos evita calcular A^{-1} y arrastrar los errores que se cometan a lo largo de todo el proceso.

Este método es conocido como *método de la potencia inversa* y nos permite calcular, en caso de existir, el autovalor de menor valor absoluto de una matriz invertible A .

Ejemplo 4.3 Para calcular, por el método de la potencia inversa, el autovalor de menor valor absoluto de la matriz del Ejemplo 4.2 aplicamos el método de la potencia inversa y partiendo del vector $z_0 = (1 \ 1 \ 1)^T$ obtenemos:

$$\begin{aligned} z_1 &= \begin{pmatrix} 0.5000 \\ 1.5000 \\ -1.0000 \end{pmatrix} & z_2 &= \begin{pmatrix} 1.6667 \\ 4.3333 \\ -3.6667 \end{pmatrix} & z_3 &= \begin{pmatrix} 1.8846 \\ 4.7308 \\ -4.0769 \end{pmatrix} & z_4 &= \begin{pmatrix} 1.9106 \\ 4.7724 \\ -4.1220 \end{pmatrix} \\ z_5 &= \begin{pmatrix} 1.9140 \\ 4.7777 \\ -4.1278 \end{pmatrix} & z_6 &= \begin{pmatrix} 1.9144 \\ 4.7784 \\ -4.1285 \end{pmatrix} & z_7 &= \begin{pmatrix} 1.9145 \\ 4.7785 \\ -4.1286 \end{pmatrix} & z_8 &= \begin{pmatrix} 1.9145 \\ 4.7785 \\ -4.1287 \end{pmatrix} \\ z_9 &= \begin{pmatrix} 1.9145 \\ 4.7785 \\ -4.1287 \end{pmatrix} \end{aligned}$$

Tenemos, por tanto, que z_9 es una aproximación al autovector asociado al autovalor de menor valor absoluto de la matriz A , por lo que el cociente de Rayleigh nos proporcionará una aproximación de éste.

$$\lambda_3 = \frac{z_9^T A z_9}{z_9^T z_9} = 0.20927063325837.$$

□

ALGORITMO DE LA POTENCIA INVERSA

```

z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z;
    z = A\w;
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z = z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)

```

Al igual que hicimos con el método de la potencia simple, podemos acelerar la convergencia del método de la potencia inversa.

ALGORITMO DE LA POTENCIA INVERSA MEJORADO

```
z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z;
    z = A\w;
    l=(z'*A*z)/(z'*z);
    if rcond(A-l*eye(length(A)))>10^(-15)
        z = (A-l*eye(length(A)))\z;
    end
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z = z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)
```

Para la matriz de nuestro Ejemplo 4.2 obtenemos, con MatLab, mediante el algoritmo del método de la potencia inversa, la aproximación 0.20927063325837 del autovalor de menor módulo con un error $5.119756492560000 \cdot 10^{-16}$ en 18 iteraciones, mientras que con el mejorado, obtenemos, en tan solo 8 iteraciones, la misma aproximación con un error del orden de $9.723368807669973 \cdot 10^{-16}$.

Veamos, por último, una nueva variante de este método nos permite calcular un autovalor cualquiera de una matriz regular A a partir de una aproximación suya. Sin embargo, hay que tener en cuenta que si el autovalor que vamos a aproximar es múltiple o existen otros con el mismo módulo aparecen dificultades que sólo podremos solventar utilizando otros métodos.

MÉTODO DE LA POTENCIA INVERSA CON DESPLAZAMIENTO

Consideremos una matriz A regular tal que todos sus autovalores tengan diferente módulo y supongamos conocida una aproximación α de un determinado autovalor λ_k . Si la aproximación es buena se verificará que $|\lambda_k - \alpha| < |\lambda_i - \alpha|$ para cualquier $i = 1, 2, \dots, n$ con $i \neq k$.

Dado que $1/(\lambda_i - \alpha)$ son los autovalores de $A - \alpha I$ y ésta posee un autovalor $(\lambda_k - \alpha)$ de menor valor absoluto que los demás, el método de la potencia inversa nos proporcionará el valor $\mu_k = \frac{1}{\lambda_k - \alpha}$ pudiéndose, a partir de éste último, hallar el valor de λ_k .

Esta variante es conocida como *método de la potencia inversa con desplazamiento*.

Ejemplo 4.4 Supongamos ahora que sabemos que el otro autovalor de la matriz A del Ejemplo 4.2 es aproximadamente 1.5.

Para calcular dicho autovalor, por el método de la potencia inversa con desplazamiento, aplicamos el método de la potencia inversa a la matriz

$$(A - 1.5I) = \begin{pmatrix} 4.5 & 2 & 5 \\ 2 & 0.5 & 3 \\ 5 & 3 & 4.5 \end{pmatrix}$$

Partiendo del vector $z_0 = (1 \ 1 \ 1)^T$ obtenemos:

$$\begin{aligned} z_1 &= \begin{pmatrix} -3.1429 \\ 2.5714 \\ 2.0000 \end{pmatrix} & z_2 &= \begin{pmatrix} -15.8701 \\ 13.8442 \\ 8.5455 \end{pmatrix} & z_3 &= \begin{pmatrix} -15.8499 \\ 13.7466 \\ 8.5663 \end{pmatrix} & z_4 &= \begin{pmatrix} -15.8233 \\ 13.7272 \\ 8.5501 \end{pmatrix} \\ z_5 &= \begin{pmatrix} -15.8244 \\ 13.7280 \\ 8.5508 \end{pmatrix} & z_6 &= \begin{pmatrix} -15.8244 \\ 13.7279 \\ 8.5508 \end{pmatrix} & z_7 &= \begin{pmatrix} -15.8244 \\ 13.7279 \\ 8.5508 \end{pmatrix} \end{aligned}$$

Tenemos, por tanto, que z_7 es una aproximación al autovector asociado al autovalor buscado, por lo que

$$\lambda_2 = \frac{z_7^T A z_7}{z_7^T z_7} = 1.56319356980456.$$

□

ALGORITMO DE LA POTENCIA INVERSA CON DESPLAZAMIENTO

```
z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z
    z = (A- $\alpha$ I)\w;
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z = z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)
```

Para la matriz del Ejemplo 4.2 y conociendo que tiene un autovalor próximo a 1.5, el algoritmo de la potencia inversa con desplazamiento, ejecutado con MatLab, nos proporciona la aproximación 1.56319356980457 del autovalor con un error de $3.845925372767128 \cdot 10^{-16}$ en 12 iteraciones.

También, en este caso, podemos hacer la misma mejora que en los anteriores algoritmos.

ALGORITMO DE LA POTENCIA INVERSA CON DESPLAZAMIENTO MEJORADO

```

z = ones(length(A),1);
w = zeros(length(A),1);
n = 0;
while norm(z-w)>10^(-14)
    n = n + 1;
    w = z;
    z = (A-alpha*I)\w;
    l=(z'*A*z)/(z'*z);
    if rcond(A-l*eye(length(A)))>10^(-15)
        z = (A-l*eye(length(A)))\z;
    end
    for i = 1:length(A)
        if abs(z(i)) == max(abs(z))
            break
        end
    end
    z = z/z(i);
end
n
autovalor = (z'*A*z)/(z'*z)
Error = norm(A*z-autovalor*z)

```

Este algoritmo mejorado nos proporciona la aproximación 1.56319356980456 del autovalor con un error de $5.978733960281817 \cdot 10^{-16}$ en tan solo 3 iteraciones.

4.3.2 Algoritmo *QR* de Francis

Dada una matriz $A \in \mathbf{K}^{n \times n}$, se pretende encontrar una sucesión de matrices $(A_n)_{n \in \mathbf{N}}$ convergente a la matriz triangular de Schur T (en cuya diagonal aparecen los autovalores de la matriz A).

La construcción de dicha sucesión se hace de la siguiente manera: $A_0 = A$.

Supuesta encontrada A_n , se realiza la factorización $A_n = Q_n R_n$ mediante matrices de Householder y $A_{n+1} = R_n Q_n$. En otras palabras:

$$\left. \begin{aligned}
A_0 &= A = Q_0 R_0 \\
A_1 &= R_0 Q_0 = Q_0^* A Q_0 = Q_1 R_1 \\
A_2 &= R_1 Q_1 = Q_1^* Q_0^* A Q_0 Q_1 = Q_2 R_2 \\
&\vdots \\
A_{n+1} &= R_n Q_n = Q_n^* \cdots Q_1^* Q_0^* A Q_0 Q_1 \cdots Q_n = Q_{n+1} R_{n+1} \\
&\vdots
\end{aligned} \right\} \Rightarrow$$

$$A_{n+1} = (Q_0 Q_1 \cdots Q_n)^* A (Q_0 Q_1 \cdots Q_n)$$

Al ser A_n semejante a A cualquiera que sea $n \in \mathbf{N}$, todas las matrices de las sucesión tienen los mismos autovalores.

Si $\lambda_1, \lambda_2, \dots, \lambda_n$ son los autovalores de una matriz A tales que

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0$$

el algoritmo converge.

En general, si existen autovalores de igual módulo, *la sucesión converge a una matriz triangular por cajas* en la que cada caja contiene a los autovalores del mismo módulo.

En otras palabras,

- Si todos los autovalores tienen distinto módulo el límite de la sucesión es una matriz *triangular superior* en la que los elementos de su diagonal son los autovalores de la matriz.
- Si existen autovalores de igual módulo la matriz límite es una matriz *triangular superior por cajas* en la que cada caja de orden k de su diagonal es una matriz cuyos autovalores son todos los k autovalores de igual módulo de la matriz A .

Así, por ejemplo, si una matriz A de orden 3 tiene un autovalor real y dos autovalores complejos conjugados (por tanto de igual módulo) de distinto módulo que el autovalor real se llegaría a una matriz del tipo

$$\left(\begin{array}{c|cc}
a_{11} & a_{12} & a_{13} \\
\hline
0 & a_{22} & a_{23} \\
0 & a_{32} & a_{33}
\end{array} \right)$$

donde a_{11} es el autovalor real y los autovalores de la matriz $\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$ son los dos autovalores complejos de la matriz A .

De forma más general, si A es una matriz cualquiera (normal o no), al aplicarle el algoritmo, la sucesión converge (*en la mayoría de los casos*) a su forma de Schur.

Ejemplo 4.5 Para el cálculo de los autovalores de la matriz $A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix}$

se tiene:

$$\begin{aligned} A_0 = A &= \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} & A_1 &= \begin{pmatrix} 7 & -1.9415 & 0.6671 \\ -1.9415 & -0.5385 & 0.1850 \\ 0.6671 & 0.1850 & -0.4615 \end{pmatrix} \\ A_2 &= \begin{pmatrix} 7.5034 & 0.3365 & 0.0344 \\ 0.3365 & -1.1474 & -0.1174 \\ 0.0344 & -0.1174 & -0.3554 \end{pmatrix} & A_3 &= \begin{pmatrix} 7.5162 & -0.0530 & 0.0016 \\ -0.0530 & -1.1758 & 0.0348 \\ 0.0016 & 0.0348 & -0.3404 \end{pmatrix} \\ A_4 &= \begin{pmatrix} 7.5165 & 0.0083 & 0.0001 \\ 0.0083 & -1.1775 & -0.0100 \\ 0.0001 & -0.0100 & -0.3390 \end{pmatrix} & A_5 &= \begin{pmatrix} 7.5165 & -0.0013 & 0.0000 \\ -0.0013 & -1.1776 & 0.0029 \\ 0.0000 & 0.0029 & -0.3389 \end{pmatrix} \\ A_6 &= \begin{pmatrix} 7.5165 & 0.0002 & 0 \\ 0.0002 & -1.1776 & -0.0008 \\ 0 & -0.0008 & -0.3389 \end{pmatrix} & A_7 &= \begin{pmatrix} 7.5165 & 0 & 0 \\ 0 & -1.1776 & 0 \\ 0 & 0 & -0.3389 \end{pmatrix} \end{aligned}$$

Por lo que los autovalores de la matriz A son:

$$-1.1776 \quad -0.3389 \quad \text{y} \quad 7.5165 \quad \square$$

Ejemplo 4.6 La matriz $A = \begin{pmatrix} 2 & 3 & 5 \\ 1 & 4 & 6 \\ -1 & 3 & 1 \end{pmatrix}$ tiene por autovalores 7 y $\pm\sqrt{2}$.

Al tener dos autovalores de igual módulo no converge a una matriz triangular, sino a la matriz triangular por cajas.

Realizando el proceso con MatLab hemos obtenido en la iteración 31 que

$$A_{31} = \left(\begin{array}{c|cc} 7.0000 & 6.7412 & -0.6901 \\ -0.0000 & 0.5797 & 2.4374 \\ 0.0000 & 0.6826 & -0.5798 \end{array} \right)$$

por lo que sabemos que un autovalor es 7 y los otros dos son los autovalores de la matriz

$$B = \begin{pmatrix} 0.5797 & 2.4374 \\ 0.6826 & -0.5798 \end{pmatrix}$$

cuyos autovalores son $\pm\sqrt{2}$. □

4.3.3 Método de Jacobi para matrices simétricas reales

Comenzaremos viendo el método para dimensión dos y, más tarde, generalizaremos para una dimensión cualquiera.

Dada una matriz simétrica y real $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ el método trata de buscar un valor para α de tal forma que la rotación $P = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}$ haga que la matriz $P^{-1}AP$ sea diagonal.

$$P^{-1}AP = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}$$

Se debe verifica, por tanto, que

$$(a - c) \cos \alpha \sin \alpha + b(\cos^2 \alpha - \sin^2 \alpha) = 0$$

es decir:

$$b(\cos^2 \alpha - \sin^2 \alpha) = (c - a) \cos \alpha \sin \alpha$$

Si $c = a$ basta tomar $\alpha = \frac{\pi}{4}$.

Si $c \neq a$ podemos dividir por $c - a$ y obtenemos que

$$\frac{2b}{c - a} = \frac{2 \cos \alpha \sin \alpha}{\cos^2 \alpha - \sin^2 \alpha} = \frac{\sin 2\alpha}{\cos 2\alpha} = \operatorname{tg} 2\alpha$$

Llamando $m = \frac{2b}{c - a}$ se tiene que $\operatorname{tg} 2\alpha = m$, por lo que

$$t = \operatorname{tg} \alpha = \frac{-1 \pm \sqrt{1 + m^2}}{m}$$

y, a partir del valor de t obtenido, se pueden calcular

$$\cos \alpha = \frac{1}{\sqrt{1 + t^2}} \quad \text{y} \quad \sin \alpha = \frac{t}{\sqrt{1 + t^2}}$$

valores, estos últimos, que nos permiten determinar la matriz P .

ALGORITMO DE CÁLCULO

a) Dada $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$

a.1) Si $b = 0$ FIN.

a.2) Si $b \neq 0$ y $a = c$

$$P = \begin{pmatrix} \sqrt{2}/2 & \sqrt{2}/2 \\ -\sqrt{2}/2 & \sqrt{2}/2 \end{pmatrix} \quad \text{y} \quad P^{-1}AP = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}. \quad \text{FIN.}$$

b) $m = \frac{2b}{c-a} \quad t = \frac{-1 \pm \sqrt{1+m^2}}{m}.$

c) $\cos \alpha = \frac{1}{\sqrt{1+t^2}} \quad \text{sen } \alpha = \frac{t}{\sqrt{1+t^2}}.$

$$P = \begin{pmatrix} \cos \alpha & \text{sen } \alpha \\ -\text{sen } \alpha & \cos \alpha \end{pmatrix} \quad \text{y} \quad P^{-1}AP = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix}. \quad \text{FIN.}$$

Para dimensiones superiores sea P la matriz diagonal por bloques

$$P = \begin{pmatrix} I_{p-1} & & \\ & T_{q-p+1} & \\ & & I_{n-q} \end{pmatrix}$$

en la que

$$T_{q-p+1} = \begin{pmatrix} \cos \alpha & & \text{sen } \alpha \\ & I_{q-p-1} & \\ -\text{sen } \alpha & & \cos \alpha \end{pmatrix}$$

y todos los demás elementos nulos.

- Si $a_{pq} = 0$, hacemos $\cos \alpha = 1$ y $\text{sen } \alpha = 0$.
- Si $a_{pp} = a_{qq}$, hacemos $\cos \alpha = \text{sen } \alpha = \frac{\sqrt{2}}{2}$.
- Si $a_{pp} \neq a_{qq}$ llamando $m = \frac{2a_{pq}}{a_{qq} - a_{pp}}$ y $t = \frac{-1 \pm \sqrt{1+m^2}}{m}$ hacemos

$$\cos \alpha = \frac{1}{\sqrt{1+t^2}} \quad \text{y} \quad \text{sen } \alpha = \frac{t}{\sqrt{1+t^2}}$$

Entonces, los elementos que ocupan los lugares pq y qp de la matriz $P^{-1}AP$ son nulos.

El método de Jacobi consiste en buscar el elemento a_{pq} con $p \neq q$ de mayor módulo, anularlo mediante una matriz P_1 , buscar el siguiente elemento de mayor módulo y anularlo mediante otra matriz P_2 y así sucesivamente hasta diagonalizar la matriz A .

Dado que las matrices P_k , o matrices de Jacobi, son ortogonales, se verifica que $P_k^{-1} = P_k^T$ por lo que se trata, en definitiva, de buscar P_1, P_2, \dots, P_k de la forma descrita anteriormente, para obtener

$$P_k^T \cdots P_1^T AP_1 \cdots P_K = D$$

Lo más importante de este método es que el cálculo de las matrices P_i puede hacerse simultáneamente ya que no se requiere el conocimiento de ninguna de ellas en el cálculo de cualquier otra. Es decir, *el método es paralelizable*.

Ejemplo 4.7 Consideremos la matriz $A = \begin{pmatrix} 1 & -2 & 3 & 1 \\ -2 & 0 & 4 & 2 \\ 3 & 4 & 1 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix}$ que es

simétrica y real.

El elemento extradiagonal de mayor valor absoluto es $a_{23} = 4$, por lo que comenzamos haciendo

$$m = \frac{2a_{23}}{a_{33} - a_{22}} \quad t = \frac{-1 + \sqrt{1 + m^2}}{m} \quad \cos \alpha = \frac{1}{\sqrt{1 + t^2}} \quad \sin \alpha = \frac{t}{\sqrt{1 + t^2}}$$

$$P_{23} = I_4 \quad p_{22} = p_{33} = \cos \alpha \quad p_{23} = -p_{32} = \sin \alpha$$

obteniéndose

$$P_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0.7497 & 0.6618 & 0 \\ 0 & -0.6618 & 0.7497 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

por lo que

$$P_{23}^T AP_{23} = \begin{pmatrix} 1 & -3.4848 & 0.9254 & 1 \\ -3.4848 & -3.5311 & 0 & 0.8376 \\ 0.9254 & 0 & 4.5311 & 2.0733 \\ 1 & 0.8376 & 2.0733 & 0 \end{pmatrix}$$

Obsérvese que la transformación de semejanza sólo ha afectado a las filas y columnas 2 y 3 dejando invariantes a los demás elementos.

Si queremos aplicar el método a la matriz $P_{23}^T A P_{23}$ obtenida, utilizando como referencia el elemento $a_{14} = 1$, podemos observar que todos los cálculos necesarios para obtener la nueva matriz P_{14} podíamos haberlos realizados al mismo tiempo que los realizados para P_{23} ya que sólo necesitamos los valores de los elementos $(1, 1)$, $(1, 4)$, $(4, 1)$ y $(4, 4)$ de la matriz $P_{23}^T A P_{23}$, que por haber quedado invariantes son los mismos que los de la matriz A .

Realizando los cálculos necesarios de forma análoga al caso anterior obtenemos

$$P_{14} = \begin{pmatrix} 0.8507 & 0 & 0 & -0.5257 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0.5257 & 0 & 0 & 0.8507 \end{pmatrix}$$

que aplicaríamos a $P_{23}^T A P_{23}$ para obtener $P_{14}^T P_{23}^T A P_{23} P_{14} = P^T A P$ donde la matriz P viene dada por $P = P_{23} P_{14}$.

Ahora bien, la matriz

$$P = P_{23} P_{14} = \begin{pmatrix} 0.8507 & 0 & 0 & -0.5257 \\ 0 & 0.7497 & 0.6618 & 0 \\ 0 & -0.6618 & 0.7497 & 0 \\ 0.5272 & 0 & 0 & 0.8507 \end{pmatrix}$$

puede ser escrita directamente sin necesidad de construir previamente P_{23} y P_{14} y multiplicarlas.

En resumen, se pueden enviar a un ordenador los valores de a_{22} , $a_{23} = a_{32}$ y a_{33} para que calcule el ángulo de giro correspondiente y nos devuelva los valores de $p_{22} = p_{33}$ y $p_{23} = -p_{32}$, mientras que de forma simultánea enviamos a otro ordenador los valores de a_{11} , a_{14} y a_{44} y nos devolverá los de $p_{11} = p_{44}$ y $p_{14} = -p_{41}$.

En el ordenador central construimos la matriz P con los datos recibidos y calculamos

$$P^T A P = \begin{pmatrix} 1.6180 & -2.5240 & 1.8772 & 0 \\ -2.5240 & -3.5311 & 0 & 2.5445 \\ 1.8772 & 0 & 4.5311 & 1.2771 \\ 0 & 2.5445 & 1.2771 & -0.6180 \end{pmatrix}$$

que volveremos a renombrar con A .

A la vista de la matriz enviaríamos al **Ordenador 1** los datos de los elementos a_{22} , $a_{24} = a_{42}$ y a_{44} mientras que enviaríamos al **Ordenador 2** los de a_{11} , $a_{13} = a_{31}$ y a_{33} que nos devolverían los elementos necesarios para construir una nueva matriz

$$P = \begin{pmatrix} 0.8981 & 0 & 0.4399 & 0 \\ 0 & 0.8661 & 0 & 0.5016 \\ -0.4399 & 0 & 0.8981 & 0 \\ 0 & -0.5016 & 0 & 0.8651 \end{pmatrix}$$

y a partir de ella

$$P^T A P = \begin{pmatrix} 0.6986 & -1.6791 & 0 & -1.6230 \\ -1.6791 & -5.0065 & -1.5358 & 0 \\ 0 & -1.5358 & 5.4506 & 0.4353 \\ -1.6230 & 0 & 0.4353 & 0.8573 \end{pmatrix}$$

Reiterando el proceso llegamos a la matriz

$$\begin{pmatrix} 2.5892 & 0 & 0 & 0 \\ 0 & -5.6823 & 0 & 0 \\ 0 & 0 & 5.7118 & 0 \\ 0 & 0 & 0 & -0.6188 \end{pmatrix}$$

cuyos elementos diagonales son los autovalores de la matriz original. \square

4.4 Reducción del problema a matrices hermiticas

Los autovalores de una matriz normal cualquiera pueden ser reales o complejos mientras que los de una matriz hermitica son todos reales, por lo que si pudiéramos transformar el cálculo de los autovalores de una matriz normal en los de otra hermitica habríamos simplificado el problema. Es más, si pudiéramos transformarlo en el cálculo de los autovalores de una matriz simétrica real, podríamos trabajar en el campo real en vez de hacerlo en el campo complejo y, entre otras cosas, podríamos utilizar el método de Jacobi a partir de una matriz normal cualquiera, previa transformación en un problema de autovalores de una matriz simétrica real.

Proposición 4.14 [COMPONENTES HERMÍTICAS DE UNA MATRIZ]

Dada cualquier matriz cuadrada A , existen dos matrices hermíticas H_1 y H_2 tales que $A = H_1 + iH_2$ las cuales reciben el nombre de *componentes hermíticas* de la matriz A .

Demostración. Si se quiere que $A = H_1 + iH_2$ con H_1 y H_2 hermíticas, se tendrá que $A^* = H_1^* - iH_2^* = H_1 - iH_2$, por lo que resolviendo el sistema resultante se tiene que

$$H_1 = \frac{A + A^*}{2} \quad H_2 = \frac{A - A^*}{2i}$$

además, la solución es única. ■

Teorema 4.15 Sea $A = H_1 + iH_2$ una matriz normal con H_1 y H_2 hermíticas. Si $x \neq 0$ es un autovector de A asociado al autovalor $\alpha + i\beta$ ($\alpha, \beta \in \mathbf{R}$) se verifica que

$$H_1x = \alpha x \quad y \quad H_2x = \beta x$$

Demostración. Por ser A normal, existe una matriz unitaria U tal que $U^*AU = D$

$$D = U^*AU = \begin{pmatrix} \alpha + i\beta & 0 & \cdots & 0 \\ 0 & \alpha_2 + i\beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n + i\beta_n \end{pmatrix} \Rightarrow$$

$$A = U \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix} U^* + iU \begin{pmatrix} \beta & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_n \end{pmatrix} U^* \Rightarrow$$

$$H_1 = U \begin{pmatrix} \alpha & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{pmatrix} U^* \quad H_2 = U \begin{pmatrix} \beta & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \beta_n \end{pmatrix} U^*$$

donde H_1 y H_2 son hermíticas. Se tiene entonces que α es un autovalor de H_1 asociado a x (primera columna de U) y β un autovalor de H_2 asociado también a x . ■

Las componentes hermíticas H_1 y H_2 de una matriz cuadrada A nos proporcionan otro método para estudiar si la matriz A es, o no es, normal.

Teorema 4.16 *Una matriz cuadrada A es normal si, y sólo si, sus componentes hermíticas conmutan.*

$$A \text{ es normal} \iff H_1 H_2 = H_2 H_1$$

Demostración.

$$H_1 H_2 = \frac{A + A^*}{2} \cdot \frac{A - A^*}{2i} = \frac{A^2 - AA^* + A^*A + (A^*)^2}{4i}$$

$$H_2 H_1 = \frac{A - A^*}{2i} \cdot \frac{A + A^*}{2} = \frac{A^2 + AA^* - A^*A + (A^*)^2}{4i}$$

- Si A es normal se verifica que $AA^* = A^*A$ por lo que

$$\left. \begin{aligned} H_1 H_2 &= \frac{A + A^*}{2} \cdot \frac{A - A^*}{2i} = \frac{A^2 + (A^*)^2}{4i} \\ H_2 H_1 &= \frac{A - A^*}{2i} \cdot \frac{A + A^*}{2} = \frac{A^2 + (A^*)^2}{4i} \end{aligned} \right\} \implies H_1 H_2 = H_2 H_1$$

- Si $H_1 H_2 = H_2 H_1$ se verifica que

$$A^2 - AA^* + A^*A + (A^*)^2 = A^2 + AA^* - A^*A + (A^*)^2$$

por lo que

$$-AA^* + A^*A = A^2 + AA^* - A^*A \iff 2A^*A = 2AA^* \iff A^*A = AA^*$$

y, por tanto, A es normal.

4.5 Reducción del problema a matrices simétricas reales

Sea H una matriz hermítica y sean $h_{ij} = a_{ij} + ib_{ij}$ con $1 \leq i, j \leq n$ sus elementos.

Podemos descomponer la matriz H de la forma $H = A + iB$ donde A y B son matrices reales con $A = (a_{ij})$ y $B = (b_{ij})$.

Por ser H hermítica se verifica que

$$A + iB = H = H^* = A^* - iB^* = A^T - iB^T \text{ (por ser } A \text{ y } B \text{ reales)} \implies \begin{cases} A = A^T \\ B = -B^T \end{cases}$$

Sea $x = a + ib$ un autovector de H asociado al autovalor α (recuérdese que $\alpha \in \mathbf{R}$ por ser H una matriz hermítica). Se verifica entonces que

$$\left. \begin{aligned} Hx &= (A + iB)(a + ib) = (Aa - Bb) + i(Ab + Ba) \\ \alpha x &= \alpha(a + ib) = \alpha a + i\alpha b \end{aligned} \right\}$$

$$Hx = \alpha x \implies \begin{cases} Aa - Bb = \alpha a \\ Ab + Ba = \alpha b \end{cases}$$

por lo que

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \alpha \begin{pmatrix} a \\ b \end{pmatrix}$$

Obsérvese además que la matriz real $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ es simétrica, ya que

$$\begin{pmatrix} A & -B \\ B & A \end{pmatrix}^T = \begin{pmatrix} A^T & B^T \\ -B^T & A^T \end{pmatrix} = \begin{pmatrix} A & -B \\ B & A \end{pmatrix}$$

Los autovalores de la matriz $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ son, por tanto, los mismos que los de H y si un autovector de $\begin{pmatrix} A & -B \\ B & A \end{pmatrix}$ es $(x_1, \dots, x_n, x_{n+1}, \dots, x_{2n})$, el correspondiente autovector de la matriz H es $((x_1 + ix_{n+1}), \dots, (x_n + ix_{2n}))$.

Dada una matriz normal calcularemos sus autovalores de la siguiente forma:

ALGORITMO DE CÁLCULO DE LOS AUTOVALORES DE UNA MATRIZ NORMAL

Paso 1 Calcular las componentes hermíticas

$$H_1 = \frac{A + A^*}{2} \quad \text{y} \quad H_2 = \frac{A - A^*}{2i}$$

de la matriz A .

Paso 2 Calcular los autovalores λ_i ($1 \leq i \leq n$) de la primera componente hermítica H_1 y sus autovectores asociados v_i reduciendo previamente el problema al caso de matrices reales y simétricas.

Paso 3 Calcular los cocientes de Rayleigh $\mu_i = \frac{v_i^* H_2 v_i}{v_i^* v_i}$.

Como los vectores v_i ($1 \leq i \leq n$) son también autovectores de H_2 , los μ_i obtenidos son sus autovalores asociados.

Paso 4 Los autovectores de A son los mismos vectores v_i y sus autovalores asociados vienen dados por $\lambda_i + i\mu_i$.

4.6 Aplicación al cálculo de las raíces de un polinomio

Consideremos el polinomio

$$P(x) = a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n$$

Dado que el polinomio característico de la matriz

$$A = \left(\begin{array}{cccc|c} -a_1/a_0 & -a_2/a_0 & \cdots & -a_{n-1}/a_0 & -a_n/a_0 \\ & & & & 0 \\ & & I_{n-1} & & \vdots \\ & & & & 0 \end{array} \right)$$

(donde I_{n-1} representa la matriz unidad de orden $n-1$) es precisamente $P(x)$, la raíces de dicho polinomio coinciden con los autovalores de la matriz A , por lo que dichas raíces pueden ser obtenidas, en bloque, mediante un método iterado de cálculo de autovalores.

Así, por ejemplo, MATLAB en su comando **roots(P)** para calcular las raíces de un polinomio P construye, en primer lugar, la matriz A definida anteriormente y se limita luego a calcular sus autovalores aplicando internamente el comando **eig(A)** de cálculo de los autovalores de una matriz. Este comando lo que hace, en primer lugar, es aplicar un algoritmo para llevar la matriz A a una matriz de Hessenberg y posteriormente aplicarle el algoritmo QR mediante transformaciones de Householder.

Bibliografía

- [1] Burden, R.L. y Faires, J.D. *Análisis Numérico (Sexta edición)*. Internacional Thomson Ed. 1998.
- [2] Golub, G.H. y Van Loan, C.F. *Matrix Computations (Third edition)*. Johns Hopkins University Press
- [3] Hager, W. *Applied Numerical Linear Algebra*. Ed. Prentice-Hall International. 1988.
- [4] Kincaid, D. y Cheney, W. *Análisis Numérico*. Ed. Addison-Wesley Iberoamericana. 1994.
- [5] Noble, D. y Daniel, J.W. *Álgebra Lineal Aplicada*. Ed. Prentice-Hall. 1989.
- [6] Watkins, D.S. *Fundamentals of MATRIX Computations*. John Wiley & Sons. 1991.