Capítulo 5

El problema de valores propios de una matriz

Discutiremos ahora el problema de la localización y la determinación numérica de los valores propios reales y complejos de una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ (o $\mathbf{A} \in \mathbb{R}^{n\times n}$) y los vectores propios asociados, o sea el problema de determinar los ceros del polinomio

$$p_n(\lambda; \mathbf{A}) := \det(\mathbf{A} - \lambda \mathbf{I})$$

y la solución de los sistemas homogéneos $(\mathbf{A} - \lambda_i \mathbf{I})\mathbf{x}_i = 0$. Formalmente, la solución del problema es dada por (a) la determinación de los coeficientes de $p_n(\lambda; \mathbf{A})$, (b) la computación (exacta o aproximada) de sus ceros y (c) la solución de los sistemas lineales homogéneos. Sin embargo, en la práctica, este camino es absolutamente inútil, bajo el aspecto del esfuerzo computacional tanto que bajo el de la estabilidad numérica. Para ilustrar el último punto, consideremos un pequeño ejemplo.

Ejemplo 5.1. *La matriz*

$$\mathbf{A} = \begin{bmatrix} 1000 & 1 \\ 1 & 1000 \end{bmatrix}$$

tiene los valores propios $\lambda_1 = 1001$ y $\lambda_2 = 999$. Ahora, si modificamos $\mathbf{A}$ a

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1000{,}001 & 1 \\ 1 & 1000 \end{bmatrix},$$

obtenemos $\lambda_1 = 1001{,}00050\ldots$ y $\lambda_2 = 999{,}00050\ldots$. Sabemos que

$$p_2(\lambda, \mathbf{A}) = \lambda^2 - 2000\lambda + 999999,$$
$$p_2(\lambda, \tilde{\mathbf{A}}) = \lambda^2 - 2000{,}001\lambda + 1000000.$$

Ahora, si el coeficiente 10^6 en $p_2(\lambda, \tilde{\mathbf{A}})$ se cambia a 1000002 (correspondiente a la magnitud de errores de redondeo en una aritmética con 6 dígitos), el polinomio modificado tiene los ceros

$$\lambda^2 - 2000{,}001\lambda + 1000002 = 0 \iff \lambda = 1000{,}0005 \pm 0{,}99999927\mathrm{i},$$

es decir que la influencia del error en los coeficientes de $p_2(\lambda : \mathbf{A})$ es casi 2000 veces mayor que la de los errores en la matriz original.

5.1. La localización de valores propios y la sensitividad del problema

Para la siguiente discusión es útil recordar el siguiente teorema.

Teorema 5.1. *Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$ y $\|\cdot\|$ una norma matricial inducida por una norma vectorial. Entonces cada valor propio $\lambda(\mathbf{A})$ satisface $|\lambda(\mathbf{A})| \leqslant r_\sigma(\mathbf{A}) \leqslant \|\mathbf{A}\|$.*

Teorema 5.2 (Los círculos de Gershgorin)**.** *Para una matriz $\mathbf{A} \in \mathbb{C}^{n\times n}$ definimos los círculos*

$$\mathcal{K}_i := \left\{ \lambda \in \mathbb{C} \;\middle|\; |\lambda - \alpha_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |\alpha_{ij}| \right\}, \quad \bar{\mathcal{K}}_i := \left\{ \lambda \in \mathbb{C} \;\middle|\; |\lambda - \alpha_{ii}| \leqslant \sum_{j=1, j\neq i}^{n} |\alpha_{ji}| \right\},$$

para $i = 1, \ldots, n$. Sea $\lambda(\mathbf{A})$ un valor propio de $\mathbf{A}$, entonces

$$\lambda(\mathbf{A}) \in \left(\bigcup_{i=1}^{n} \mathcal{K}_i \right) \cap \left(\bigcup_{i=1}^{n} \bar{\mathcal{K}}_i \right).$$

Demostración. Sean $\lambda \in \sigma(\mathbf{A})$ y $\mathbf{x}$ el vector propio asociado, es decir $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, e $i \in \{1, \ldots, n\}$ tal que $|x_i| = \|\mathbf{x}\|_\infty$. Entonces, la componente i de la ecuación vectorial $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ es

$$\sum_{j=1}^{n} \alpha_{ij} x_j = \lambda x_i \iff \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_{ij} x_j + \alpha_{ii} x_i = \lambda x_i,$$

lo cual entrega que

$$\lambda - \alpha_{ii} = \sum_{\substack{j=1 \\ j\neq i}}^{n} \alpha_{ij} \frac{x_j}{x_i}.$$

Dado que $x_i \neq 0$, podemos concluir que

$$|\lambda - \alpha_{ii}| \leqslant \sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_{ij}| \frac{|x_j|}{|x_i|} \leqslant \sum_{\substack{j=1 \\ j\neq i}}^{n} |\alpha_{ij}|,$$

y luego se toma en cuenta que $\mathbf{A}$ y $\mathbf{A}^*$ tienen los mismos valores propios. ■

Dado que las matrices $\mathbf{A}$ y $\mathbf{D}^{-1}\mathbf{A}\mathbf{D}$ poseen los mismos valores propios, a veces se puede precisar el resultado del Teorema 5.2 significativamente.

Ejemplo 5.2. *La matriz*

$$\mathbf{A} = \begin{bmatrix} 1 & 10^{-3} & 10^{-4} \\ 10^{-3} & 2 & 10^{-3} \\ 10^{-4} & 10^{-3} & 3 \end{bmatrix}$$

es simétrica, entonces sus valores propios son reales, y según el Teorema 5.2, cada valor propio λ de $\mathbf{A}$ satisface

$$\lambda \in [1 - 0{,}0011, 1 + 0{,}0011] \cup [2 - 0{,}002, 2 + 0{,}002] \cup [3 - 0{,}0011, 3 + 0{,}0011].$$

Ahora, usando las matrices

$$\mathbf{D}_1 := \operatorname{diag}(1, 100, 10), \quad \mathbf{D}_2 := \operatorname{diag}(100, 1, 100), \quad \mathbf{D}_3 := \operatorname{diag}(10, 100, 1),$$

obtenemos las siguientes inclusiones respectivas:

$$\lambda \in U_1 := [1 - 2\times 10^{-5}, 1 + -2\times 10^{-5}] \cup [2 - 0{,}11, 2 + 0{,}11] \cup [3 - 0{,}0011, 3 + 0{,}0011],$$
$$\lambda \in U_2 := [1 - 0{,}1001, 1 + 0{,}1001] \cup [2 - 2\times 10^{-5}, 2 + 2\times 10^{-5}] \cup [3 - 0{,}1001, 3 + 0{,}1001],$$

$$\lambda \in U_3 := [1 - 0{,}0011, 1 + 0{,}0011] \cup [2 - 0{,}02, 2 + 0{,}02] \cup [3 - 2 \times 10^{-5}, 3 + 2 \times 10^{-5}],$$

lo que implica

$$\lambda \in U_1 \cap U_2 \cap U_3 = \bigcup_{i=1}^{3} [i - 2 \times 10^{-5}, i + 2 \times 10^{-5}].$$

Teorema 5.3. *Consideremos las hipótesis del Teorema 5.2. Sea $\{i_1, \ldots, i_n\}$ una permutación de $\{1, \ldots, n\}$ y*

$$(\mathcal{K}_{i_1} \cup \ldots \cup \mathcal{K}_{i_m}) \cap \mathcal{K}_{i_s} = \varnothing \quad \textit{para } s = m+1, \ldots, n.$$

Entonces $\mathcal{K}_{i_1} \cup \ldots \cup \mathcal{K}_{i_m}$ contiene exactamente m valores propios de $\mathbf{A}$, contados con su multiplicidad, es decir, cada componente de conectividad por camino de $\mathcal{K}_{i_1} \cup \ldots \cup \mathcal{K}_{i_m}$ contiene tantos valores propios de $\mathbf{A}$ que círculos.

Demostración. Sean $\mathbf{D} = \operatorname{diag}(\alpha_{11}, \ldots, \alpha_{nn})$ y $\mathbf{B}(\tau) := \mathbf{D} + \tau(\mathbf{A} - \mathbf{D})$, $0 \leqslant \tau \leqslant 1$, es decir, $\mathbf{B}(0) = \mathbf{D}$ y $\mathbf{B}(1) = \mathbf{A}$. Todos los valores propios de $\mathbf{B}(\tau)$ estan contenidos en $\mathcal{K}_1(\tau) \cup \ldots \cup \mathcal{K}_n(\tau)$, donde definimos

$$\mathcal{K}_i(\tau) = \left\{ z \in \mathbb{C} \;\middle|\; |z - \alpha_{ii}| \leqslant \tau \sum_{j=1, j\neq i}^{n} |\alpha_{ij}| \right\}, \quad i = 1, \ldots, n.$$

Obviamente, el Teorema 5.3 es válido para $\mathbf{B}(0)$, además, los valores propios dependen de forma continua de τ (ver Lema 5.1 abajo). Pero como $\mathcal{K}_{i_1}(0) \cup \cdots \cup \mathcal{K}_{i_m}(0)$ contiene exactamente m valores propios de $\mathbf{B}(0)$, y

$$\forall \tau \in [0,1] : \left(\bigcup_{j=1}^{m} \mathcal{K}_{i_j}(\tau) \right) \cap \mathcal{K}_{i_s}(1) \subset \left(\bigcup_{j=1}^{m} \mathcal{K}_{i_j}(1) \right) \cap \mathcal{K}_{i_s}(1) = \varnothing,$$

entonces $\mathcal{K}_{i,1}(\tau) \cup \cdots \cup \mathcal{K}_{i,m}(\tau)$ contiene exactamente m valores propios para $0 \leqslant \tau \leqslant 1$. ■

Lema 5.1. *Sean $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$, $\lambda_1, \ldots, \lambda_n$ los valores propios de $\mathbf{A}$, contados con su multiplicidad, y $\lambda'_1, \ldots, \lambda'_n$ los valores propios de $\mathbf{B}$ contados con su multiplicidad. Sean*

$$\varrho := \text{máx}\{|\alpha_{ij}|, |\beta_{ij}| : 1 \leqslant i, j \leqslant n\}, \quad \delta := \frac{1}{n\varrho} \sum_{i=1}^{n} \sum_{j=1}^{n} |\alpha_{ij} - \beta_{ij}|.$$

Entonces existe una enumeración de los valores propios λ_i y λ'_i tal que a cada λ_i corresponde un valor λ'_i con

$$|\lambda_i - \lambda'_i| \leqslant 2(n+1)^2 \varrho \sqrt[n]{\delta}.$$

Demostración. Ver A.M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces, Academic Press, 3rd ed., 1973, pp. 334–335, 276–279. ■

Ejemplo 5.3. *En virtud del Teorema 5.3, podemos mejorar ahora el resultado del Ejemplo 5.2: los valores propios de la matriz $\mathbf{A}$ en este ejemplo pueden ser enumerados de tal forma que*

$$\lambda_i \in [i - 2 \times 10^{-5}, i + 2 \times 10^{-5}], \quad i = 1, 2, 3.$$

Cuando conocemos un vector propio $\mathbf{x}$ aproximado (de hecho, como tal vector podemos usar cualquier vector $\mathbf{x}$ con $\mathbf{Ax} \neq 0$), podemos usar el cuociente de Rayleigh

$$R(\mathbf{x}; \mathbf{A}) := \frac{\mathbf{x}^*\mathbf{Ax}}{\mathbf{x}^*\mathbf{x}}$$

para definir una aproximación del valor propio correspondiente, para la cual también podemos definir una inclusión.

Teorema 5.4. *Sea $\mathbf{A} \in \mathbb{C}^{n\times n}$ similar a una matriz diagonal con los valores propios $\lambda_1, \ldots, \lambda_n$. Sean $\mathbf{x} \in \mathbb{C}^n$, $\mathbf{x} \neq 0$ y $\mathbf{Ax} \neq 0$. Sea $\lambda := R(\mathbf{x}; \mathbf{A})$. Entonces*

(i) $\forall c \in \mathbb{C} : \|\mathbf{Ax} - \lambda\mathbf{x}\|_2^2 \leqslant \|\mathbf{Ax} - c\mathbf{x}\|_2^2$.

(ii) *Existe un valor $\lambda_j \neq 0$, $\lambda_j \in \sigma(\mathbf{A})$, tal que*

$$\left| \frac{\lambda_j - \lambda}{\lambda_j} \right| \leqslant \frac{\|\mathbf{Ax} - \lambda\mathbf{x}\|_2}{\|\mathbf{Ax}\|_2} \operatorname{cond}_{\|\cdot\|_2}(\mathbf{U}),$$

donde $\mathbf{U} = \begin{bmatrix} \mathbf{u}_1 & \ldots & \mathbf{u}_n \end{bmatrix}$ es un sistema completo de vectores propios de $\mathbf{A}$.

(iii) *Si $\mathbf{A}$ es normal (es decir, $\mathbf{AA}^* = \mathbf{A}^*\mathbf{A}$), entonces existe un valor $0 \leqslant \lambda_j \in \sigma(\mathbf{A})$ tal que*

$$\left| \frac{\lambda_j - \lambda}{\lambda_j} \right| \leqslant \frac{\|\mathbf{Ax} - \lambda\mathbf{x}\|_2}{\|\mathbf{Ax}\|_2}.$$

Demostración.

(i) Supongamos que $\|\mathbf{x}\|_2 = 1$. Entonces

$$\begin{aligned} \|\mathbf{Ax} - c\mathbf{x}\|_2^2 &= (\mathbf{x}^*\mathbf{A}^* - \bar{c}\mathbf{x}^*)(\mathbf{Ax} - c\mathbf{x}) \\ &= \mathbf{x}^*\mathbf{A}^*\mathbf{Ax} - \bar{c}\mathbf{x}^*\mathbf{Ax} - c\mathbf{x}^*\mathbf{A}^*\mathbf{x} + |c|^2 \\ &= \|\mathbf{Ax}\|_2^2 + |c - \mathbf{x}^*\mathbf{Ax}|^2 - |\mathbf{x}^*\mathbf{Ax}|^2 \\ &\geqslant \|\mathbf{Ax}\|_2^2 - |\mathbf{x}^*\mathbf{Ax}|^2 \end{aligned}$$

con igualdad para $c = \lambda$.

(ii) Sean $\mathbf{U}^{-1}\mathbf{AU} = \operatorname{diag}(\lambda_1, \ldots, \lambda_n) =: \mathbf{\Lambda}$, $\mathbf{y} := \mathbf{U}^{-1}\mathbf{x}$ y

$$\operatorname{glb}(\mathbf{U}) := \sup\{\alpha \mid \alpha\|\mathbf{x}\| \leqslant \|\mathbf{Ux}\|\} = \frac{1}{\|\mathbf{U}^{-1}\|}.$$

Entonces

$$\begin{aligned} \frac{\|\mathbf{Ax} - \lambda\mathbf{x}\|_2}{\|\mathbf{Ax}\|_2} &= \frac{\|\mathbf{U}(\mathbf{\Lambda} - \lambda\mathbf{I})\mathbf{U}^{-1}\mathbf{x}\|_2}{\|\mathbf{U\Lambda U}^{-1}\mathbf{x}\|_2} \\ &\geqslant \frac{\operatorname{glb}(\mathbf{U})\|(\mathbf{\Lambda} - \lambda\mathbf{I})\mathbf{y}\|_2}{\|\mathbf{U}\|_2\|\mathbf{\Lambda y}\|_2} \end{aligned}$$