

# Principio minimax para los valores propios de matrices hermitianas

El propósito de estos apuntes (escritos por Egor Maximenko, Rogelio Rocha Hernández, Eliseo Sarmiento Rosales y Zacarías Sánchez Francisco Javier) es explicar una demostración del Teorema 6 que se conoce como el principio minimax para los valores propios y se asocia con los nombres Rayleigh, Ritz, Courant, Fischer y Weyl. Este teorema da una caracterización variacional de los valores propios de matrices hermitianas.

## Descomposición espectral de matrices hermitianas (repasso)

Primero recordamos el teorema sobre la descomposición espectral de matrices hermitianas. Dada una matriz compleja  $A$ , denotamos por  $A^*$  su adjunta (es decir, la transpuesta conjugada). Decimos que una matriz  $A$  es *hermitiana* o *autoadjunta* si  $A^* = A$ . Denotamos por  $\mathcal{HM}_n(\mathbb{C})$  al conjunto de todas las matrices hermitianas de orden  $n$ :

$$\mathcal{HM}_n(\mathbb{C}) := \{A \in \mathcal{M}_n(\mathbb{C}) : A^* = A\}.$$

**1 Teorema.** *Sea  $A \in \mathcal{HM}_n(\mathbb{C})$ . Entonces existen una base ortonormal  $(u_1, \dots, u_n)$  del espacio  $\mathbb{C}^n$  y una lista  $(\lambda_1, \dots, \lambda_n)$  de números reales tales que*

$$\lambda_1 \leq \dots \leq \lambda_n$$

y

$$Au_j = \lambda_j u_j \quad (j \in \{1, \dots, n\}).$$

En otras palabras, dada una matriz autoadjunta  $A$ , existe una base ortonormal del espacio  $\mathbb{C}^n$  formada de vectores propios de la matriz  $A$ . Los números  $\lambda_1, \dots, \lambda_n$  en el Teorema 1 son los valores propios de  $A$  escritos en el orden ascendente y contados con sus multiplicidades. Estos números están determinados de manera única por la matriz  $A$ . Los denotamos por  $\lambda_1(A), \dots, \lambda_n(A)$ .

El Teorema 1 se puede escribir también en términos de *matrices unitarias*. Es fácil ver que el Teorema 1 da un *criterio* de matrices hermitianas. No explicamos estos hechos de manera más detallada porque no los utilizamos en este tema.

## Cociente de Rayleigh

Sea  $A \in \mathcal{HM}_n(\mathbb{C})$ . Entonces para cada  $x \in \mathbb{C}^n$  el número  $x^*Ax$  es real. Esto se verifica fácilmente usando propiedades de la operación  $*$  y tratando este número como una matriz de tamaño  $1 \times 1$ :

$$\overline{x^*Ax} = (x^*Ax)^* = x^*A^*(x^*)^* = x^*Ax.$$

La función  $R_A : \mathbb{C}^n \setminus \{\mathbf{0}_n\} \rightarrow \mathbb{R}$  definida mediante la regla

$$R_A(x) = \frac{x^*Ax}{x^*x}$$

se conoce como el *cociente de Rayleigh*.

Notamos que si  $x \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}$  y  $\lambda \in \mathbb{C} \setminus \{0\}$ , entonces

$$R_A(\lambda x) = \frac{|\lambda|^2 x^* A x}{|\lambda|^2 x^* x} = R_A(x).$$

Esto implica que el cociente de Rayleigh se puede definir y estudiar en el espacio proyectivo  $\mathbb{P}(\mathbb{C}^n)$ , pero no vamos a desarrollar esta idea.

Escribimos  $S < \mathbb{C}^n$  cuando  $S$  es un subespacio de  $\mathbb{C}^n$ . Denotamos por  $C$  a la esfera unitaria del espacio  $\mathbb{C}^n$ :

$$C = \{x \in \mathbb{C}^n : \|x\| = 1\}.$$

**2 Proposición.** *Sea  $S < \mathbb{C}^n$ . Entonces el conjunto  $\{R_A(x) : x \in S \setminus \{\mathbf{0}_n\}\}$  coincide con el conjunto  $\{R_A(x) : x \in S \cap C\}$  y tiene un valor mínimo y un valor máximo.*

*Demostración.* Para cada  $x \in S \setminus \{\mathbf{0}_n\}$ , el vector  $z$  definido mediante la fórmula  $z = \frac{x}{\|x\|}$  tiene propiedades  $z \in C \cap S$  y  $R_A(x) = R_A(z)$ , por eso

$$\{R_A(x) : x \in S \setminus \{\mathbf{0}_n\}\} \subseteq \{R_A(z) : z \in S \cap C\}.$$

La contención  $\supseteq$  es obvia porque  $S \cap C \subseteq S \setminus \{\mathbf{0}_n\}$ . Siendo un subespacio de  $\mathbb{C}^n$  el conjunto  $S$  es un subconjunto cerrado de  $\mathbb{C}^n$ , por eso  $C \cap S$  es cerrado y acotado en  $\mathbb{C}^n$ . La función  $R_A$  es continua y alcanza sus valores mínimo y máximo en  $C \cap S$ .  $\square$

## Cociente de Rayleigh sobre combinaciones lineales de vectores propios

Si  $(u_1, \dots, u_n)$  es una base ortonormal de  $\mathbb{C}^n$  y  $x \in \mathbb{C}^n$ , entonces existe un único vector  $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{C}^n$  tal que

$$x = \sum_{j=1}^n \alpha_j u_j.$$

De hecho, las coordenadas  $\alpha_j$  se calculan mediante la fórmula  $\alpha_j = u_j^* x$ . Usando la propiedad  $u_j^* u_k = \delta_{j,k}$  es fácil ver que

$$\|x\|^2 = x^* x = \sum_{j=1}^n |\alpha_j|^2 = \|\alpha\|^2,$$

así que  $\|x\| = \|\alpha\|$ . En particular, esto implica que  $x \neq \mathbf{0}_n$  si y sólo si  $\alpha \neq \mathbf{0}_n$ .

**3 Lema** (expresión del cociente de Rayleigh en términos de la descomposición espectral). *Sea  $A \in \mathcal{HM}_n(\mathbb{C})$  y sea  $(u_1, \dots, u_n)$  una base ortonormal de valores propios de  $A$ , como en el Teorema 1. Más aún, sea*

$$x = \sum_{j=1}^n \alpha_j u_j, \tag{1}$$

donde  $\alpha_1, \dots, \alpha_n$  son algunos escalares complejos no todos cero. Entonces

$$R_A(x) = \frac{\sum_{j=1}^n \lambda_j(A) |\alpha_j|^2}{\sum_{j=1}^n |\alpha_j|^2}. \quad (2)$$

En particular, para cada  $j$  en  $\{1, \dots, n\}$

$$R_A(u_j) = \lambda_j(A). \quad (3)$$

*Demostración.* Usando las igualdades  $Au_j = \lambda_j(A)u_j$  y la ortonormalidad de los vectores  $u_1, \dots, u_n$ , obtenemos

$$x^*x = \sum_{j=1}^n |\alpha_j|^2, \quad x^*Ax = \sum_{j=1}^n \lambda_j(A) |\alpha_j|^2. \quad \square$$

El siguiente resultado se asocia habitualmente con los nombres de Rayleigh y Ritz.

**4 Teorema.** Sea  $A \in \mathcal{HM}_n(\mathbb{C})$ . Entonces

$$\lambda_1(A) = \min_{x \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}} R_A(x), \quad \lambda_n(A) = \max_{x \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}} R_A(x).$$

*Demostración.* Usamos la notación y el resultado del Lema 3. Dado un vector  $x \in \mathbb{C}^n \setminus \{\mathbf{0}_n\}$ , lo escribimos como (1). Para cada  $j \in \{1, \dots, n\}$

$$\lambda_1(A) \leq \lambda_j(A) \leq \lambda_n(A).$$

Multiplicamos por  $|\alpha_j|^2$ , sumamos por  $j$  y dividimos entre  $\|x\|^2$ . De esta manera obtenemos

$$\lambda_1(A) \leq R_A(x) \leq \lambda_n(A).$$

Por otro lado, aplicamos la igualdad (3) con  $j = 1$  y luego con  $j = n$ :

$$R_A(u_1) = \lambda_1(A), \quad R_A(u_n) = \lambda_n(A).$$

Hemos demostrado que la función  $R_A$  toma valores entre  $\lambda_1(A)$  y  $\lambda_n(A)$  y alcanza ambos estos valores.  $\square$

Notamos que la función  $R_A$  alcanza el valor  $\lambda_1(A)$  no sólo en el vector  $u_1$ , sino también en todos los múltiplos complejos no nulos del vector  $u_1$ . Más aún, si  $\lambda_1(A) = \lambda_2(A) = \lambda_3(A)$ , entonces, por ejemplo,  $R_A(7u_1 + 8u_2 + 9u_3) = \lambda_1(A)$ .

El siguiente resultado se puede ver como una generalización del Teorema 4 y un lema para el Teorema principal 6.

**5 Lema.** Sea  $A \in \mathcal{HM}_n(\mathbb{C})$  y sea  $(u_1, \dots, u_n)$  una base ortonormal de  $\mathbb{C}^n$  tal que  $Au_j = \lambda_j(A)u_j$  para cada  $j \in \{1, \dots, n\}$ . Sean  $p, q \in \{1, \dots, n\}$  tales que  $p < q$ . Denotemos por  $E_{p,q}$  al subespacio generado por los vectores  $u_p, \dots, u_q$ :

$$E_{p,q} = \ell(u_p, \dots, u_q) = \left\{ x \in \mathbb{C}^n : \exists \alpha_p, \dots, \alpha_q \in \mathbb{C} \quad x = \sum_{j=p}^q \alpha_j u_j \right\}.$$

Entonces

$$\lambda_p(A) = \min_{x \in E_{p,q} \setminus \{\mathbf{0}_n\}} R_A(x), \quad \lambda_q(A) = \max_{x \in E_{p,q} \setminus \{\mathbf{0}_n\}} R_A(x).$$

*Demostración.* Para cada  $j \in \{p, \dots, q\}$  se cumplen las desigualdades

$$\lambda_p(A) \leq \lambda_j(A) \leq \lambda_q(A).$$

Si  $x \in E_{p,q} \setminus \{\mathbf{0}_n\}$ , entonces  $x$  se puede escribir como

$$x = \sum_{j=p}^q \alpha_j u_j,$$

donde entre los números  $\alpha_p, \dots, \alpha_q$  hay al menos uno distinto de cero. Luego

$$\lambda_p(A) \|x\|^2 = \lambda_p(A) \sum_{j=p}^q |\alpha_j|^2 \leq \sum_{j=p}^q \lambda_j(A) |\alpha_j|^2 \leq \lambda_q(A) \sum_{j=p}^q |\alpha_j|^2 = \lambda_q(A) \|x\|^2.$$

Dividiendo entre  $\|x\|^2$  obtenemos

$$\lambda_p(A) \leq R_A(x) \leq \lambda_q(A).$$

Por otro lado, la fórmula (3) con  $j = p$  y luego con  $j = q$  nos da

$$R_A(u_p) = \lambda_p(A), \quad R_A(u_q) = \lambda_q(A).$$

Hemos demostrado que la función  $R_A$  restringida a  $E_{p,q} \setminus \{\mathbf{0}_n\}$  toma valores entre  $\lambda_p(A)$  y  $\lambda_q(A)$  y alcanza ambos estos valores.  $\square$

## Resultado principal

**6 Teorema** (principio minimax para los valores propios de matrices hermitianas). *Sea  $A \in \mathcal{HM}_n(\mathbb{C})$  y sea  $k \in \{1, \dots, n\}$ . Entonces*

$$\lambda_k(A) = \min_{\substack{S < \mathbb{C}^n \\ \dim(S)=k}} \max_{x \in S \setminus \{\mathbf{0}_n\}} R_A(x). \quad (4)$$

*Demostración.* Sea  $S < \mathbb{C}^n$  tal que  $\dim(S) = k$ . Por la fórmula de Grassmann,

$$\dim(S \cap E_{k,n}) = \dim(S) + \dim(E_{k,n}) - \dim(S + E_{k,n}).$$

Como  $\dim(S + E_{k,n}) \leq n$ , obtenemos

$$\dim(S \cap E_{k,n}) \geq \dim(S) + \dim(E_{k,n}) - n = k + (n - k + 1) - n = 1,$$

así que  $S \cap E_{k,n}$  es un subespacio no nulo del espacio  $\mathbb{C}^n$ . Si  $x \in S \cap E_{k,n}$  y  $x \neq \mathbf{0}_n$ , entonces por el Lema 5 (aplicado con  $p = k$  y  $q = n$ ) tenemos la desigualdad  $R_A(x) \geq \lambda_k(A)$ . Por eso

$$\max_{x \in S \setminus \{\mathbf{0}_n\}} R_A(x) \geq \max_{x \in (S \cap E_{k,n}) \setminus \{\mathbf{0}_n\}} R_A(x) \geq \lambda_k(A).$$

Hemos demostrado que para cada  $S < \mathbb{C}^n$  con  $\dim(S) = k$  se tiene la desigualdad

$$\max_{x \in S \setminus \{\mathbf{0}_n\}} R_A(x) \geq \lambda_k(A). \quad (5)$$

Falta encontrar un  $S < \mathbb{C}^n$  con  $\dim(S) = k$  tal que la desigualdad (5) se convierta en una igualdad. Aplicamos el Lema 5 con  $p = 1$  y  $q = k$  y obtenemos

$$\max_{x \in E_{1,k} \setminus \{\mathbf{0}_n\}} R_A(x) = \lambda_k(A). \quad \square$$

**7 Ejercicio.** Revisar la demostración del Teorema 6 y demostrar que

$$\lambda_k(A) = \min_{\substack{S < \mathbb{C}^n \\ \dim(S) \geq k}} \max_{x \in S \setminus \{\mathbf{0}_n\}} R_A(x). \quad (6)$$