Matrix Computations
Chapter 4: Eigenvalues, Eigenvectors, and Eigendecomposition

Section 4.6 More on Variational Characterizations of Eigenvalues

Jie Lu
ShanghaiTech University

Courant-Fischer Min-Max Theorem (Revisit)

For $\mathbf{A} \in \mathbb{H}^{n\times n}$, let $\lambda_k(\mathbf{A})$ denote the kth largest eigenvalue of $\mathbf{A}$, i.e.,

$$\lambda_{\min}(\mathbf{A}) := \lambda_n(\mathbf{A}) \le \cdots \le \lambda_1(\mathbf{A}) =: \lambda_{\max}(\mathbf{A})$$

For simplicity, we may also write $\lambda_{\min} := \lambda_n \le \cdots \le \lambda_1 =: \lambda_{\max}$

Theorem

For any $\mathbf{A} \in \mathbb{H}^{n\times n}$ and $k = 1, \ldots, n$,

$$\begin{aligned}\lambda_k(\mathbf{A}) &= \max_{\substack{\mathcal{S}\subseteq\mathbb{C}^n:\\ \dim(\mathcal{S})=k}} \min_{\substack{\mathbf{y}\in\mathcal{S},\\ \mathbf{y}\neq\mathbf{0}}} \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}} \\ &= \min_{\substack{\mathcal{S}\subseteq\mathbb{C}^n:\\ \dim(\mathcal{S})=n-k+1}} \max_{\substack{\mathbf{y}\in\mathcal{S},\\ \mathbf{y}\neq\mathbf{0}}} \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}}\end{aligned}$$

$R_{\mathbf{A}}(\mathbf{y}) = \frac{\mathbf{y}^H\mathbf{A}\mathbf{y}}{\mathbf{y}^H\mathbf{y}}$, $\mathbf{y} \neq \mathbf{0}$ is the Rayleigh–Ritz quotient, or Rayleigh quotient

This section focuses on variational characterizations of eigenvalues of real symmetric matrices ($\mathbb{S}^n$)

Rayleigh-Ritz Theorem

A special case of Courant-Fischer Min-Max Theorem

Theorem (Rayleigh-Ritz)

For any $\mathbf{A} \in \mathbb{S}^n$,

$$\lambda_{\min} \|\mathbf{x}\|_2^2 \leq \mathbf{x}^T \mathbf{A} \mathbf{x} \leq \lambda_{\max} \|\mathbf{x}\|_2^2$$

where the equalities can be attained when $\mathbf{x}$ *is an eigenvector associated with* $\lambda_{\min}$ *and* $\lambda_{\max}$, *respectively*

- Even without Courant-Fischer Min-Max Theorem, we may prove this using eigendecomposition $\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^T$, $\mathbf{V}$ real orthogonal

More Results from Courant-Fischer

Let $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{z} \in \mathbb{R}^n$

- (Weyl) $\lambda_k(\mathbf{A}) + \lambda_n(\mathbf{B}) \leq \lambda_k(\mathbf{A} + \mathbf{B}) \leq \lambda_k(\mathbf{A}) + \lambda_1(\mathbf{B}), \; k = 1, \ldots, n$

More Results from Courant-Fischer (cont'd)

Let $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{z} \in \mathbb{R}^n$

- (Interlacing) $\lambda_{k+1}(\mathbf{A}) \leq \lambda_k(\mathbf{A} \pm \mathbf{z}\mathbf{z}^T) \leq \lambda_{k-1}(\mathbf{A})$ for proper k

More Results from Courant-Fischer (cont'd)

Let $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$, $\mathbf{z} \in \mathbb{R}^n$

- If $\mathrm{rank}(\mathbf{B}) \leq r$, then $\lambda_{k+r}(\mathbf{A}) \leq \lambda_k(\mathbf{A}+\mathbf{B}) \leq \lambda_{k-r}(\mathbf{A})$ for proper k

- (Weyl) $\lambda_{j+k-1}(\mathbf{A}+\mathbf{B}) \leq \lambda_j(\mathbf{A}) + \lambda_k(\mathbf{B})$ for proper j, k

- For any semi-orthogonal $\mathbf{U} \in \mathbb{R}^{n \times r}$,
  $\lambda_{k+n-r}(\mathbf{A}) \leq \lambda_k(\mathbf{U}^T\mathbf{A}\mathbf{U}) \leq \lambda_k(\mathbf{A})$ for proper k

Extend Variational Characterization to Sum of Eigenvalues

Theorem

For any $\mathbf{A} \in \mathbb{S}^n$,

$$\sum_{i=1}^{r} \lambda_i = \max_{\substack{\mathbf{U} \in \mathbb{R}^{n \times r} \\ \|\mathbf{u}_i\|_2 = 1 \ \forall i, \ \mathbf{u}_i^T \mathbf{u}_j = 0 \ \forall i \neq j}} \sum_{i=1}^{r} \mathbf{u}_i^T \mathbf{A} \mathbf{u}_i = \max_{\substack{\mathbf{U} \in \mathbb{R}^{n \times r} \\ \mathbf{U}^T \mathbf{U} = \mathbf{I}}} \operatorname{tr}(\mathbf{U}^T \mathbf{A} \mathbf{U})$$

- This can be proved using $\lambda_k(\mathbf{U}^T \mathbf{A} \mathbf{U}) \leq \lambda_k(\mathbf{A})$, but we may try another way of proof to get better understanding of trace, which uses the fact that

$$\max_{\substack{\mathbf{U} \in \mathbb{R}^{n \times r} \\ \mathbf{U}^T \mathbf{U} = \mathbf{I}}} \operatorname{tr}(\mathbf{U}^T \mathbf{A} \mathbf{U}) = \max_{\substack{\mathbf{U} \in \mathbb{R}^{n \times r} \\ \mathbf{U}^T \mathbf{U} = \mathbf{I}}} \operatorname{tr}(\mathbf{U}^T \Lambda \mathbf{U})$$

Other Extensions

(Von Neumann) For any $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$,

$$\mathrm{tr}(\mathbf{AB}) \leq \sum_{i=1}^{n} \lambda_i(\mathbf{A})\lambda_i(\mathbf{B})$$

(Lidskii) For any $\mathbf{A}, \mathbf{B} \in \mathbb{S}^n$ and any $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq n$,

$$\sum_{j=1}^{k} \lambda_{i_j}(\mathbf{A} + \mathbf{B}) \leq \sum_{j=1}^{k} \lambda_{i_j}(\mathbf{A}) + \sum_{j=1}^{k} \lambda_{i_j}(\mathbf{B})$$