Variational characterizations of eigenvalues

MATH 6610 Lecture 04

September 9, 2020

Hermitian matrices

If $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian, recall that:

- $\boldsymbol{A}$ is unitarily diagonalizable, $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^*$
- The spectrum of $\boldsymbol{A}$ is real-valued
- (Hermitian) Positive-definite matrices have matrix square roots

(and unique spd square roots)

Hermitian matrices

If $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ is Hermitian, recall that:

- $\boldsymbol{A}$ is unitarily diagonalizable, $\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^*$
- The spectrum of $\boldsymbol{A}$ is real-valued
- (Hermitian) Positive-definite matrices have matrix square roots

Today: Variational characterizations of eigenvalues for Hermitian matrices.

Rayleigh quotients, I

Let $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ be a matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x}\rangle}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} = \frac{x^* A x}{x^* x}$$

$$R_A(x) = R_A\left(\frac{x}{\|x\|_2}\right)$$

(i.e. $R_A(x)$ is invariant under scaling of x)

Rayleigh quotients, I

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

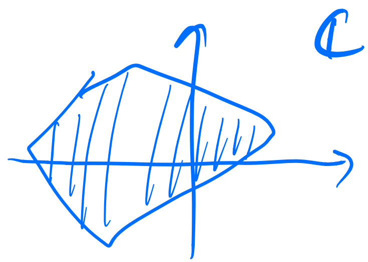
The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

- The numerical range of $\boldsymbol{A}$ is the set of all possible values of $R_{\boldsymbol{A}}$:

$$W_{\boldsymbol{A}}(\mathbb{C}^n) := R_{\boldsymbol{A}}\left(\mathbb{C}^n \backslash \{\boldsymbol{0}\}\right).$$

$W_A(\mathbb{C}^n) \subset \mathbb{C}$



Rayleigh quotients, I

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

- The numerical range of $\boldsymbol{A}$ is the set of all possible values of $R_{\boldsymbol{A}}$:

$$W_{\boldsymbol{A}}(\mathbb{C}^n) := R_{\boldsymbol{A}}\left(\mathbb{C}^n \backslash \{\boldsymbol{0}\}\right).$$

- The numerical range contains the spectrum of $\boldsymbol{A}$.

Let v be an(y) eigenvector of A associated to eigenvalue λ. Then $R_A(v) = \dfrac{v^* A v}{\|v\|_2^2} = \lambda \underline{\dfrac{v^* v}{\|v\|_2^2}} = \lambda$

Rayleigh quotients, I

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\boldsymbol{0}\}$ be a vector.

The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

- The numerical range of $\boldsymbol{A}$ is the set of all possible values of $R_{\boldsymbol{A}}$:

$$W_{\boldsymbol{A}}(\mathbb{C}^n) := R_{\boldsymbol{A}}\left(\mathbb{C}^n \backslash \{\boldsymbol{0}\}\right).$$

- The numerical range contains the spectrum of $\boldsymbol{A}$.
- If $\boldsymbol{A}$ is Hermitian, then $\lambda_{\min}(\boldsymbol{A}) \leqslant R_{\boldsymbol{A}}(\boldsymbol{x}) \leqslant \lambda_{\max}(\boldsymbol{A})$.

Upper inequality: $A = U \Lambda U^*$, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$

$\lambda_j \leq \lambda_{j+1}$. $R_A(x) = \dfrac{\langle x, Ax \rangle}{\langle x, x \rangle} = \dfrac{\langle U^* x, \Lambda U^* x \rangle}{\langle U^* x, U^* x \rangle}$

Let $y = U^* x$

Then: $$R_A(x) = \frac{\langle y, \Lambda y \rangle}{\|y\|_2^2} = \frac{\sum_{j=1}^{n} \lambda_j |y_j|^2}{\|y\|_2^2} \leq \frac{\sum_{j=1}^{n} \lambda_n |y_j|^2}{\|y\|_2^2}$$

$$= \lambda_n \; (= \lambda_{max}(A))$$

Rayleigh quotients, I

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be a matrix, and let $\boldsymbol{x} \in \mathbb{C}^n \backslash \{\mathbf{0}\}$ be a vector.

The Rayleigh Quotient (of $\boldsymbol{A}$ at $\boldsymbol{x}$) is the scalar,

$$R_{\boldsymbol{A}}(\boldsymbol{x}) := \frac{\langle \boldsymbol{A}\boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

- The numerical range of $\boldsymbol{A}$ is the set of all possible values of $R_{\boldsymbol{A}}$:

$$W_{\boldsymbol{A}}(\mathbb{C}^n) := R_{\boldsymbol{A}}\left(\mathbb{C}^n \backslash \{\mathbf{0}\}\right).$$

- The numerical range contains the spectrum of $\boldsymbol{A}$.
- If $\boldsymbol{A}$ is Hermitian, then $\lambda_{\min}(\boldsymbol{A}) \leqslant R_{\boldsymbol{A}}(\boldsymbol{x}) \leqslant \lambda_{\max}(\boldsymbol{A})$.

We can also consider the image of the Rayleigh quotient but only on a subspace V:

$$W_{\boldsymbol{A}}(V) := R_{\boldsymbol{A}}\left(V \backslash \{\mathbf{0}\}\right).$$

Rayleigh quotients, II

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $W_{\boldsymbol{A}}(V)$, is some subset of $\mathbb{R}$.

Rayleigh quotients, II

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $W_{\boldsymbol{A}}(V)$, is some subset of $\mathbb{R}$.

- The minimum of $W_{\boldsymbol{A}}(V)$ can be $\lambda_{\min}(\boldsymbol{A})$. (if V contains minimal eigenvector of A)
  What is the largest possible minimum value?

Rayleigh quotients, II

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian. Consider a subspace $V \subset \mathbb{C}^n$.

The image of the V under the Rayleigh quotient, $W_{\boldsymbol{A}}(V)$, is some subset of $\mathbb{R}$.

- The minimum of $W_{\boldsymbol{A}}(V)$ can be $\lambda_{\min}(\boldsymbol{A})$.
  What is the largest possible minimum value?
- The maximum of $W_{\boldsymbol{A}}(V)$ can be $\lambda_{\max}(\boldsymbol{A})$.
  What is the smallest possible maximum value?

The "min-max" theorem

Theorem (Courant-Fischer-Weyl "min-max")

Let $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ be Hermitian, with eigenvalues $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$. Then for each $1 \leqslant k \leqslant n$,

$$(\star) \qquad \lambda_k = \min_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_{\boldsymbol{A}}(V)$$

$k = n \Rightarrow V = \mathbb{C}^n \rightarrow \lambda_n = \max W_A(\mathbb{C}^n).$

$$\lambda_k = \max_{\substack{V \subset \mathbb{C}^n \\ \dim V = n-k+1}} \min W_{\boldsymbol{A}}(V)$$

In addition, if $(\boldsymbol{u}_j)_{j=1}^n$ are the eigenvectors associated with $(\lambda_j)_{j=1}^n$, then:

- $V = \text{span}\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\}$ *achieves the outer minimum*
- $V = \text{span}\{\boldsymbol{u}_k, \ldots, \boldsymbol{u}_n\}$ *achieves the outer maximum*

we'll prove first min-max equation. $(\star)$.

$$(\circledast) \quad \lambda_k = \min_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_A(V).$$

We'll accomplish this by showing the $\leq$ and $\geq$ inequalities.

Let $V = \operatorname{span}\{u_1, \dots, u_k\}$ $\left((\lambda_j, u_j)_{j=1}^n \text{ are eigenpairs of } A \right)$.

$\boxed{\max W_A(V)}$

$x \in V$: $R_A(x) = \dfrac{\langle x, Ax \rangle}{\langle x, x \rangle}$ $\quad x \in V \Rightarrow \boxed{x = \sum_{j=1}^k c_j u_j}$

$$= \frac{\langle x, \sum_{j=1}^k c_j \lambda_j u_j \rangle}{\langle x, x \rangle} = \frac{\sum_{j=1}^k |c_j|^2 \lambda_j}{\langle x, x \rangle}$$

$$= \frac{\sum_{j=1}^k |c_j|^2 \lambda_j}{\sum_{j=1}^k |c_j|^2} \leq \lambda_k \frac{\sum_{j=1}^k |c_j|^2}{\sum_{j=1}^k |c_j|^2} = \lambda_k.$$

$\max W_A(V) = \lambda_k$

$$\min_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_A(V) \leq \lambda_k. \qquad (\leq)$$

(already showed $\lambda_k \geq \min\limits_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_A(V)$)

Let V be any k-dim subspace of $\mathbb{C}^n$.

span$\{u_k, u_{k+1} \dots u_n\}$ is a dimension $(n-k+1)$ space.

Then $V \cap \text{span}\{u_k, \dots u_n\} \neq \phi$.

$\Rightarrow \exists\ v \neq 0$ s.t. $v = \sum\limits_{j=k}^{n} c_j u_j$, $v \in V$.

$$R_A(v) = \frac{\langle v, Av \rangle}{\langle v, v \rangle} = \frac{\sum\limits_{j=k}^{n} |c_j|^2 \lambda_j}{\sum\limits_{j=k}^{n} |c_j|^2} \geq \lambda_k$$

$\Rightarrow \max\limits_{x \in V} R_A(x) = \max W_A(V) \geq \lambda_k \quad (\forall\ k\text{-dim}V)$.

$\Rightarrow \min\limits_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_A(V) \geq \lambda_k$.

$\Rightarrow \lambda_k = \min\limits_{\substack{V \subset \mathbb{C}^n \\ \dim V = k}} \max W_A(V)$ ✓ (✯)

Cauchy interlacing theorem

A matrix $\boldsymbol{B}$ is a compression of $\boldsymbol{A}$ if $\boldsymbol{B} = \boldsymbol{Q}^* \boldsymbol{A} \boldsymbol{Q}$ for some $\boldsymbol{Q} \in \mathbb{C}^{n \times r}$ with othonormal columns.

orth

$A \in \mathbb{C}^{n \times n}$ $\qquad$ $B = Q^* A Q \in \mathbb{C}^{r \times r}$

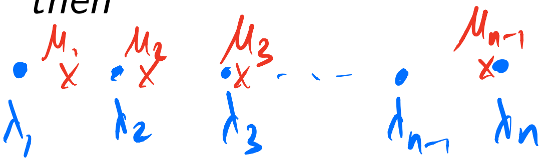
Cauchy interlacing theorem

A matrix $\boldsymbol{B}$ is a compression of $\boldsymbol{A}$ if $\boldsymbol{B} = \boldsymbol{Q}^*\boldsymbol{A}\boldsymbol{Q}$ for some $\boldsymbol{Q} \in \mathbb{C}^{n\times r}$ with othonormal columns.

Just one consequence of the min-max theorem:

Theorem (Cauchy interlacing)

Let $\boldsymbol{B} \in \mathbb{C}^{(n-1)\times(n-1)}$ *be a compression of a Hermitian matrix* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$*. If* $\boldsymbol{A}$ *has eigenvalues* $\lambda_1 \leqslant \lambda_2 \leqslant \ldots \leqslant \lambda_n$*, and* $\boldsymbol{B}$ *has eigenvalues* $\mu_1, \ldots, \mu_{n-1}$*, then*



$$\lambda_j \leqslant \mu_j \leqslant \lambda_{j+1},$$

for all $j = 1, \ldots, n-1$.

B is Hermitian $\Rightarrow$ μ_j are $\mathbb{R}$-valued

Assume $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_{n-1}$.

Proof:

Recall: if $U = \text{span}\{u_1, \dots u_j\}$, where

$(\mu_{\cancel{j}\,k}, u_{\cancel{j}\,k})_{\cancel{j}\,k=1}^{n-1}$ are eigen pairs of B.

$$\mu_j = \max_{x \in U \setminus \{0\}} R_B(x) = \max_{x \in U \setminus \{0\}} \frac{\langle Bx, x \rangle}{\langle x, x \rangle}$$

defn of B, and $Q^*Q = I$.

Then: $$\mu_j = \max_{x \in U \setminus \{0\}} \frac{\langle Bx, x \rangle}{\langle x, x \rangle} = \max_{x \in U \setminus \{0\}} \frac{\langle AQx, Qx \rangle}{\langle Qx, Qx \rangle}$$

$$= \max_{\underbrace{y \in QU \setminus \{0\}}_{\text{is } j\text{-dim subspace}}} \frac{\langle Ay, y \rangle}{\langle y, y \rangle},$$ where $QU := \{Qy \mid y \in U\}$

$\dim QU = j$

$$\geq \min_{\dim V = j} \max_{y \in V} \frac{\langle Ay, y \rangle}{\langle y, y \rangle} = \lambda_j$$

Courant-Fischer-Weyl.

$\Rightarrow \mu_j \geq \lambda_j$.

Other inequality: $U := \text{span}\{u_j, u_{j+1} \dots u_{n-1}\}$

Consider $\mu_j = \min_{x \in U \setminus \{0\}} R_B(x)$.

and perform similar computations.

...... $\square$