9 Matrix spectra

Thursday, October 1, 2020 4:15 AM

Last time we saw results about the matrices corresponding to normal operators under appropriate orthonormal bases. These theorems have matrix analogues.

For real matrices $A \in \mathbb{R}^{n \times n}$

Normal: $AA^T = A^TA$, $A = PDP^T$, P orthogonal,

Thm 16.8/16.18

$$D = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{pmatrix}, \quad D_j \text{ is } 1 \times 1 \text{ or } D_j = \begin{pmatrix} \lambda_j & -\mu_j \\ \mu_j & \lambda_j \end{pmatrix}, \quad \lambda_j, \mu_j \in \mathbb{R}, \ \mu_j > 0.$$

Symmetric: $A^T = A$, $A = PDP^T$, P orthogonal,

Thm 16.9/16.19

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad \lambda_i \in \mathbb{R}.$$

Skew-symmetric $A^T = -A$, $A = PDP^T$, P orthogonal

Thm 16.10/16.20

$$D = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{pmatrix}, \quad D_j = [0] \text{ or } D_j = \begin{pmatrix} 0 & -\mu_j \\ \mu_j & 0 \end{pmatrix}, \quad \mu_j \in \mathbb{R}, \ \mu_j > 0.$$

eigenvalues are $\pm i\mu_j$ or 0.

Orthogonal $AA^T = A^TA = I_n$, $A = PDP^T$, P orthogonal

Thm 16.11/16.21

$$D = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & \cdots \end{pmatrix} \quad D_j = [1] \text{ or } D_j = [-1]$$

or $\begin{bmatrix} \cos\theta_j & -\sin\theta_j \\ \cdots \end{bmatrix}$

$$D = \begin{pmatrix} D_1 & & 0 \\ & \ddots & \\ 0 & & D_p \end{pmatrix} \qquad D_j = [1] \text{ or } D_j = [-1]$$

$$\text{or } D_j = \begin{bmatrix} \cos\theta_j & -\sin\theta_j \\ \sin\theta_j & \cos\theta_j \end{bmatrix},$$

$$0 < \theta_j < \pi.$$

eigenvalues are $\cos\theta_j \pm i \sin\theta_j$, 1, or -1.

For complex matrices $\mathbb{C}^{n \times n}$, (Thm 16.12 / 16.22)

normal: $AA^* = A^*A$, $A = UDU^*$, U unitary, D diagonal, $d_{ii} \in \mathbb{C}$

Hermitian: $A^* = A$, $A = UDU^*$, U unitary, D diagonal, $d_{ii} \in \mathbb{R}$

skew-Hermitian: $A^* = -A$, $A = UDU^*$, U unitary, D diagonal, $d_{ii} = ir$, $r \in \mathbb{R}$.

unitary: $AA^* = A^*A = I_n$, $A = UDU^*$, U unitary, D diagonal, $|d_{ii}| = 1$.

Rayleigh-Ritz and Eigenvalue interlacing

Define the Rayleigh ratio of a symmetric matrix A and nonzero vector x is

$$R(A)(x) = \frac{x^T A x}{x^T x}$$

(or Hermitian matrix $R(A)(x) = \frac{x^* A x}{x^* x}$)

This ratio is used in characterization of eigenvalues of A.

(16.23)

Prop. 16.11 (Rayleigh-Ritz) If $A \in \mathbb{R}^{n\times n}$ with eigenvalues $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_n$ and if $(u_1, \dots, u_n)$ is any orthonormal basis of eigenvectors of A, where u_i is a unit eigenvector associated with λ_i, then

$$\max_{x \neq 0} \frac{x^T A x}{x^T x} = \lambda_n$$

with maximum attained for $x = u_n$, and

$$\max_{\substack{x \neq 0 \\ x \in \{u_{n-k+1}, \dots, u_n\}^\perp}} \frac{x^T A x}{x^T x} = \lambda_{n-k} \qquad (\text{with max at } x = u_{n-k})$$

Equivalently, if $V_k = \operatorname{span}\{u_1, \dots, u_k\}$, then

$$\lambda_k = \max_{\substack{x \neq 0 \\ x \in V_k}} \frac{x^T A x}{x^T x}.$$

proof. Note

$$\max_{x \neq 0} \frac{x^T A x}{x^T x} = \max_x \left\{ x^T A x \mid x^T x = 1 \right\}.$$

and

$$\max_{\substack{x \neq 0 \\ x \in \{u_{n-k+1}, \dots, u_n\}^\perp}} \frac{x^T A x}{x^T x} = \max_x \left\{ x^T A x \,\middle|\, x \in \{u_{n-k+1}, \dots, u_n\}^\perp \wedge (x^T x = 1) \right\}.$$

Letting $x = \sum_{i=1}^n x_i u_i$, $\quad x^T A x = \sum_{i=1}^n \lambda_i x_i^2$.

If $x^T x = 1$, then $\sum_{i=1}^n x_i^2 = 1$

$$\Rightarrow x^T A x = \sum_{i=1}^n \lambda_i x_i^2 \le \lambda_n \left(\sum_{i=1}^n x_i^2 \right) = \lambda_n.$$

So $\max\limits_{x^Tx=1} x^TAx \leq \lambda_n$.

Further, clearly $u_n^TAu_n = \lambda_n$, so $\max\limits_{x^Tx=1} x^TAx = \lambda_n$.

Similarly, $x \in \{u_{n-k+1}, \ldots, u_n\}^\perp$ and $x^Tx = 1$ iff

$x_{n-k+1} = \cdots = x_n = 0$ and $\sum\limits_{i=1}^{n-k} x_i^2 = 1$.

Thus, $x^TAx = \sum\limits_{i=1}^{n-k} \lambda_i x_i^2 \leq \lambda_{n-k}\left(\sum\limits_{i=1}^{n-k} x_i^2\right) = \lambda_{n-k}$.

Further, $u_{n-k}^T A u_{n-k} = \lambda_{n-k}$.

Thus, $$\max_{\substack{x^Tx=1 \\ x\in\{u_{n-k+1},\ldots,u_n\}^\perp}} x^TAx = \lambda_{n-k}.$$

Prop. 16.12/16.24 (Rayleigh-Ritz)

If A is a symmetric $n\times n$ matrix with eigenvalues with eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and if $(u_1, \ldots, u_n)$ is any orthonormal basis of eigenvectors of A where u_i is a unit eigenvector associated with λ_i, then

$$\min_{x\neq 0} \frac{x^TAx}{x^Tx} = \lambda_1. \qquad (\text{with minimum attained for } x = u_1)$$

and

$$\min_{\substack{x\neq 0, \\ x\in\{u_1,\ldots,u_{i-1}\}^\perp}} \frac{x^TAx}{x^Tx} = \lambda_i \qquad (\text{with minimum attained for } x = u_i)$$

Equivalently, if $W_k = V_{k-1}^\perp$ denotes the subspace spanned by $(u_k, \ldots, u_n)$ (with $V_0 = (0)$), then

$$\lambda_k = \min_{\substack{x\neq 0 \\ x\in W_k}} \frac{x^TAx}{x^Tx} = \min_{\substack{x\neq 0 \\ x\in V_{k-1}^\perp}} \frac{x^TAx}{x^Tx}.$$

$$\lambda_k = \min_{\substack{x \neq 0 \\ x \in W_k}} \frac{\qquad}{x^T x} = \min_{\substack{x \neq 0 \\ x \in V_{k-1}^{\perp}}} \frac{\qquad}{x^T x}.$$

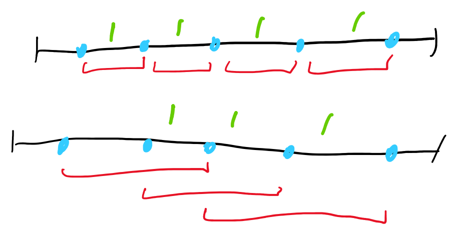
We can use these Rayleigh-Ritz theorems to compare eigenvalues of symmetric matrices A and $B = R^T A R$, where $R^T R = I$ and R is rectangular.

Define 16.5 Given symmetric $A \in \mathbb{R}^{n \times n}$ and symmetric $B \in \mathbb{R}^{m \times m}$, with $m \leq n$, if $\lambda_1 \leq \cdots \leq \lambda_n$ are eigenvalues of A and $\mu_1 \leq \cdots \leq \mu_m$ are eigenvalues of B, then the eigenvalues of B *interlace* the eigenvalues of A if

$$\lambda_i \leq \mu_i \leq \lambda_{n-m+i}, \quad i = 1, \ldots, n.$$

e.g. if $n = 5$, $m = 3$, then *interlacing* means

$$\lambda_1 \leq \mu_1 \leq \lambda_3$$
$$\lambda_2 \leq \mu_2 \leq \lambda_4$$
$$\lambda_3 \leq \mu_3 \leq \lambda_5$$



Prop. 16.13/16.25 Let $A \in \mathbb{R}^{n \times n}$ symmetric, $B \in \mathbb{R}^{m \times m}$ symmetric, $m \leq n$,
$R \in \mathbb{R}^{n \times m}$ s.t. $R^T R = I$ and $B = R^T A R$.

Then the following properties hold:

(a) The eigenvalues of B *interlace* the eigenvalues of A.

(b) If $\lambda_1 \leq \cdots \leq \lambda_n$ and $\mu_1 \leq \cdots \leq \mu_m$ are eigenvalues of A and B respectively, and if $\lambda_i = \mu_i$, then $\exists$ an eigenvector v of B with eigenvalue μ_i s.t. Rv is an eigenvector of A with eigenvalue λ_i.

proof. (a) Let $(u_1, \dots, u_n)$ be an orthonormal basis of eigenvectors of A
$(v_1, \dots, v_m)$ " " " " " " " B.

corresponding to resp. eigenvalues

Let $U_j = \text{span}\{u_1, \dots, u_j\} \subseteq \mathbb{R}^n$

$V_j = \text{span}\{v_1, \dots, v_j\} \subseteq \mathbb{R}^m$

Note $\dim(V_i) = i$ and $\dim(R^T U_{i-1}) \le i-1$.

Thus, $\exists v \neq 0$ s.t. $v \in V_i \cap (R^T U_{i-1})^\perp$ (since $\dim(R^T U_{i-1})^\perp \ge m - i + 1$)

Then $v^T \underbrace{R^T u_j}_{\in R^T U_{i-1}} = (Rv)^T u_j = 0$ for $j = 1, \dots, i-1$

$\Rightarrow Rv \in (U_{i-1})^\perp$ since Rv is orthogonal to all of u_j, $j = 1, \dots, i-1$.

By the Rayleigh-Ritz thm,

$$\lambda_i \le \frac{(Rv)^T A Rv}{(Rv)^T Rv} = \frac{v^T B v}{v^T v}.$$

And

$$\mu_i = \max_{\substack{x \neq 0 \\ x \in \{v_{i+1}, \dots, v_n\}^\perp}} \frac{x^T B x}{x^T x} = \max_{\substack{x \neq 0 \\ x \in \text{span}\{v_1, \dots, v_i\}}} \frac{x^T B x}{x^T x},$$

So $\frac{w^T B w}{w^T w} \le \mu_i$ for all $w \in V_i$.

$$\Rightarrow \lambda_i \le \frac{v^T B v}{v^T v} \le \mu_i.$$

But we can apply the same argument to $-A$ and $-B$, so

$$-\lambda_{n-m+i} \le -\mu_i$$

$$\Rightarrow \mu_i \le \lambda_{n-m+i}, \quad i = 1, \dots, m$$

$\Longrightarrow \quad \lambda_i \le \mu_i \le \lambda_{n-m+i}, \quad i = 1, \dots, m.$

(b) If $\lambda_i = \mu_i$, then

$$\lambda_i = \frac{(Rv)^T A Rv}{(Rv)^T Rv} = \frac{v^T B v}{v^T v} = \mu_i$$

Therefore, v must be an eigenvector for B and Rv an eigenvector for A, both for the eigenvalue $\lambda_i = \mu_i$. $\square$

Prop 16.14/16.26 (Poincare separation thm) (applications to quantum mechanics)

Let A be a $n \times n$ symmetric (or Hermitian) matrix, let $r \in \mathbb{Z}$ with $1 \le r \le n$, and let $(u_1, \dots, u_r)$ be r orthonormal vectors.

Let $B = (u_i^T A u_j)$ (an $r \times r$ matrix),

let $\lambda_1(A) \le \dots \le \lambda_n(A)$ be eigenvalues of A

$\lambda_1(B) \le \dots \le \lambda_r(B)$ be eigenvalues of B

Then $\lambda_k(A) \le \lambda_k(B) \le \lambda_{k+n-r}(A), \quad k = 1, \dots, r.$

proof. Immediate corollary of prev Prop. letting $R = (u_1 \cdots u_r)$.

We also have an immediate corollary of interlacing eigenvalues of matrix minors.

Let $P_1 \in \mathbb{R}^{n \times n-1}$ be defined by $I_{n\times n}[1:n; 1:n-1]$, the identity minus the last col.

Then $P_1^T P_1 = I$ and $B = P_1^T A P$ is $A[1:n-1; 1:n-1]$.

Then we get a genuine interlacing

$$\lambda_1 \le \mu_1 \le \lambda_2 \le \mu_2 \le \dots \le \mu_{n-2} \le \lambda_{n-1} \le \mu_{n-1} \le \lambda_n.$$

We get of course similar results for $B = A[1:r; 1:r]$, We get of course similar results for $B = A[1:r; 1:r]$,

$$\lambda_k \leq u_k \leq \lambda_{k+n-r} \quad , \quad k=1, \dots, r.$$