Part 1 Diagonal Entries, Eigenvalues & Singular Values

Recall that the singular values $s_1(A) \geq \cdots \geq s_n(A)$ of $A \in M_n$ are the nonnegative square roots of the eigenvalues of A^*A. Finding bounds and estimates of eigenvalues, singular values, and diagonal entries of matrices are important in many applications. We will study some basic results and techniques. Here are some facts about eigenvalues and singular values.

- Let $A \in M_n$. There are unitary $U, V \in M_n$ such that $UAV = \mathrm{diag}\,(s_1(A), \ldots, s_n(A))$.
- Let $A \in M_n$. There is a unitary U such that U^*AU is in (upper or lower) triangular form with the eigenvalues of A arranged in any specific order on the diagonal.
- A matrix $A \in M_n$ is normal (Hermitian) if and only if A is unitarily similar to a (real) diagonal matrix.

1 Real symmetric and complex Hermitian Matrices

Theorem 1.1 (Courant-Fischer-Weyl) *Let $A \in H_n$ have eigenvalues $a_1 \geq \cdots \geq a_n$. Then for $1 \leq k \leq n$,*

$$a_k = \max_{\dim W = k} \; \min_{v \in W,\; v^*v=1} v^*Av = \min_{\dim W = n-k+1} \; \max_{v \in W, v^*v=1} v^*Av.$$

When $k = 1$, we have the following corollary, which is known as the Rayleigh principle.

Corollary 1.2 (Rayleigh Principle) *Let $A \in H_n$. For any unit vector $x \in \mathbf{F}^n$, we have*

$$\lambda_1(A) \geq x^*Ax \geq \lambda_n(A).$$

Corollary 1.3 (Cauchy's Interlacing Inequalities) *Let $A \in H_n$ have eigenvalues $a_1 \geq \cdots \geq a_n$. Suppose B is a $m \times m$ principal submatrix of A with eigenvalues $b_1 \geq \cdots \geq b_m$. Then*

$$a_j \geq b_j \geq a_{n-m+j} \qquad \text{for } j = 1, \ldots, m.$$

Theorem 1.4 (Fan-Pall) *Suppose $a_1 \geq \cdots \geq a_n$ and $b_1 \geq \cdots \geq b_m$ with $1 \leq m \leq n$ satisfy*

$$a_j \geq b_j \geq a_{n-m+j} \qquad \text{for } j = 1, \ldots, m.$$

Then there exists $A \in H_n$ with eigenvalues $a_1, \ldots, a_n$ such that the leading $m \times m$ principal submatrix of A has eigenvalues $b_1 \geq \cdots \geq b_m$.

Problem 1.5 Let $1 \leq m < n$. Determine the conditions on $a_1, \ldots, a_n, b_1, \ldots, b_m \in \mathbf{C}$ for the existence of a normal matrix with eigenvalues $a_1, \ldots, a_n$ and a leading principal normal submatrix with eigenvalues $b_1, \ldots, b_m$.

Definition 1.6 *Let* $a = (a_1, \ldots, a_n)$ *and* $b = (b_1, \ldots, b_n)$ *be real vectors. We say that* b *is* **weakly majorized** *by* a*, denoted by* $b \prec_w a$*, if the sum of the* k *largest entries of* b *is not larger than the sum of the* k *largest entries of* a *for* $k \in \{1, \ldots, n\}$*. In addition, if the sum of the entries of the two vectors are the same, we say that* b *is majorized by* a*, denoted by* $b \prec a$*.*

Theorem 1.7 (Schur-Horn) *There exists* $A \in H_n$ *with eigenvalues* $a_1 \geq \cdots \geq a_n$ *and diagonal entries* $d_1, \ldots, d_n$ *if and only if*

$$(d_1, \ldots, d_n) \prec (a_1, \ldots, a_n).$$

Problem 1.8 Determine the relation between the eigenvalues and diagonal entries of a normal matrix.

2 General matrices

Theorem 2.1 *Suppose* $x, y \in \mathbf{R}^n$ *and* $z \in \mathbf{C}^n$*.*

- *There exists* $T \in M_n$ *with entries of* z *as eigenvalues and entries of* x *as eigenvalues of* $(T + T^*)/2$ *if and only if* $\mathrm{Re}(z) \prec x$*.*
- *There exists* $T \in M_n$ *with entries of* z *as eigenvalues and entries of* y *as eigenvalues of* $(T - T^*)/(2i)$ *if and only if* $\mathrm{Im}(z) \prec y$*.*

Problem 2.2 Given $x, y \in \mathbf{R}^n$ and $z \in \mathbf{C}^n$, determine the necessary and sufficient conditions for $T \in M_n$ with entries of z as eigenvalues, entries of x as eigenvalues of $(T + T^*)/2$ and entries of y as eigenvalues of $(T - T^*)/(2i)$.

Theorem 2.3 (Weyl-Horn) *There exists* $A \in M_n$ *with singular values* $s_1 \geq \cdots \geq s_n$ *and eigenvalues* $\lambda_1, \ldots, \lambda_n$ *with* $|\lambda_1| \geq \cdots \geq |\lambda_n|$ *if and only if* $\prod_{j=1}^n |\lambda_j| = \prod_{j=1}^n s_j$ *and*

$$\prod_{j=1}^k |\lambda_j| \leq \prod_{j=1}^k s_j, \qquad k = 1, \ldots, n-1.$$

Theorem 2.4 (Thompson-Sing) *There exists* $A \in M_n$ *with singular values* $s_1 \geq \cdots \geq s_n$ *and diagonal entries* $d_1, \ldots, d_n$ *with* $|d_1| \geq \cdots \geq |d_n|$ *if and only if*

$$\sum_{j=1}^{n-1} |d_j| - |d_n| \leq \sum_{j=1}^{n-1} s_j - s_n$$

and

$$\sum_{j=1}^k |d_j| \leq \sum_{j=1}^k s_j, \qquad k = 1, \ldots, n,$$

Theorem 2.5 (Thompson) *There exists a complex symmetric matrix $A \in M_n$ with singular values $s_1 \geq \cdots \geq s_n$ and diagonal entries $d_1 \geq \cdots \geq d_n$ if and only if*

(a) $\sum_{j=1}^{k} d_j \leq \sum_{j=1}^{k} s_j$ *for* $j \in \{1, \ldots, n\}$,

(b) $\sum_{j=1}^{k-1} d_j - \sum_{j=k}^{n} d_j \leq \sum_{j=1}^{n} s_j - 2s_k$ *for* $k \in \{1, \ldots, n\}$, *and*

(c) $\sum_{j=1}^{n-3} d_j - d_{n-2} - d_{n-1} - d_n \leq \sum_{j=1}^{n-2} s_j - s_{n-1} - s_n$ *in case* $n \geq 3$.

Problem 2.6 *The proof of Thompson is very long. It would be nice to find a short proof.*

3 Compound matrices

Definition 3.1 *If A is $m \times n$ and $k \leq \min\{m, n\}$, then the* k**th compound matrix** *of A is the $\binom{m}{k} \times \binom{n}{k}$ matrix $C_k(A)$ with entries equal to* $\det A[i_1, \ldots, i_k; j_1, \ldots, j_k]$ *with $1 \leq i_1 < \cdots < i_k \leq m$ and $1 \leq j_1 < \cdots < j_k \leq n$, arranged in lexicographic order; the* k**th additive compound matrix** *of A is the $\binom{m}{k} \times \binom{n}{k}$ matrix $\Delta_k(A)$ appeared as the coefficient matrix of t in the expansion*

$$C_k(I + tA) = I + t\Delta_k(A) + \cdots + t^k C_k(A).$$

Theorem 3.2 *Let $A \in M_n$ have eigenvalues $a_1, \ldots, a_n$, and singular values $s_1, \ldots, s_n$. Then $C_k(A)$ has eigenvalues $a_{j_1} \cdots a_{j_k}$ and singular values $s_{j_1} \cdots s_{j_k}$ with $1 \leq j_1 < \cdots < j_k$. Moreover, $\Delta_k(A)$ has eigenvalues $a_{j_1} + \cdots + a_{j_k}$ with $1 \leq j_1 < \cdots < j_k$.*

4 Bounds for zeros of polynomials

Definition 4.1 *Suppose A is the* **companion matrix** *of the monic polynomial $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$, i.e., the $(1, j)$ entry of A is $-a_j$ for $j = 1, \ldots, n$; the $(j, j-1)$ entry of A is 1 for $j = 2, \ldots, n$; all other entries of A are zero.*

Theorem 4.2 (Gershgorin) *Suppose $A = (a_{ij})) \in M_n$. Let $R_i = \sum_{j \neq i} |a_{ij}|$ be the deleted row sum, and let $G_i = \{\mu \in \mathbf{C} : |\mu - a_{ii}| \leq R_i\}$ be a Gershgorin disk. Then $\sigma(A) \subseteq \cup_{j=1}^{n} G_j$.*

Theorem 4.3 *Suppose $p(z) = z^n + a_1 z^{n-1} + \cdots + a_n$, and μ is a zero of μ. Then*

$$|\mu| \leq \max\{1 + |a_j| : 1 \leq j \leq n\} \quad \text{and} \quad |\mu| \leq \max\left\{1, \sum_{j=1}^{n} |a_j|\right\}.$$

If $a_n \neq 0$, one can estimate the lower bound for $|\mu|$ by studying the companion matrix of $q(z) = z^n p(1/z)/a_n = z^n + (a_{n-1}/a_n) z^{n-1} + \cdots + 1/a_n$.

Moreover, if A is the companion matrix of $p(z)$, then there are unitary matrices U and V such that $UAV = \begin{pmatrix} a_n & \gamma \\ 0 & 1 \end{pmatrix} \oplus I_{n-2}$ *with* $\gamma = \sqrt{\sum_{j=1}^{n-1} |a_j|^2}$, *and hence*

$$|\mu| \leq s_1(A) = \frac{1}{2}\left\{\sqrt{\gamma^2 + (1+|a_n|)^2} + \sqrt{\gamma^2 + (1-|a_n|)^2}\right\}.$$

If A is singular, $|\mu| \leq \sqrt{\gamma^2+1}$. If A is invertible, one can get the lower bound $|\mu| \geq s_n(A)$.

Problem 4.4 (Sendov Conjecture) Suppose A is a complex circulant matrix with all eigenvalues in the closed unit disk $\{\mu \in \mathbf{C} : |\mu| \leq 1\}$. If λ is an eigenvalue of A and B is obtained from $A - \lambda I$ by removing its last row and last column. Show that B has an eigenvalue in the closed unit disk.

Remark 4.5 *Note that if A' is obtained from A by removing its first row and first column, then the eigenvalues of A' are the zeros of $p'(z)$ for $p(z) = \det(zI - A)$.*

5 Functions preserving majorization

To obtain more matrix inequalities from the basic ones one can use Schur convex functions.

Definition 5.1 *A nonnegative matrix $A \in M_n$ is* **doubly stochastic** *if all row sums and column sums equal 1. If A is doubly stochastic and has two nonzero off-diagonal entries, then A is a* **pinching** *matrix.*

Proposition 5.2 *Let $x, y \in \mathbf{R}^{1\times n}$. The following are equivalent.*

(a) $x \prec y$.

(b) *There exist pinching matrices $T_1, \dots, T_m$ with $m < n$ such that $x = yT_1 \cdots T_m$.*

(c) *There exists a doubly stochastic matrix A such that $x = yA$.*

Proposition 5.3 *Let $f : \mathbf{R} \to \mathbf{R}$, and $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbf{R}^{1\times n}$.*

(a) *If f is convex and $x \prec y$, then*

$$(f(x_1), \dots, f(x_n)) \prec_w (f(y_1), \dots, f(y_n)).$$

(b) *If f is convex increasing and $x \prec_w y$, then*

$$(f(x_1), \dots, f(x_n)) \prec_w (f(y_1), \dots, f(y_n)).$$

Definition 5.4 *Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ have nonnegative entries. If the product of the k largest entries of x is not larger than that of y for $k = 1, \ldots, n$, we say that x is* **weakly log majorized** *by y, denoted by $x \prec_{wlog} y$. In addition, if the product of the entries of the two vectors are the same, then we say that x is* **log majorized** *by y, denoted by $x \prec_{log} y$.*

Proposition 5.5 *If $x \prec_{wlog} y$ then $x \prec_w y$.*

Definition 5.6 *A function $f : \mathbf{R}^n \to \mathbf{R}$ is Schur-convex if $f(u) \leq f(v)$ whenever $u \prec v$.*

Example 5.7 (a) *If $f : \mathbf{R} \to \mathbf{R}$ is convex, then for any $k \in \{1, \ldots, n\}$,*

$$\phi_k(x_1, \ldots, x_n) = \max\{f(x_{j_1}) + \cdots + f(x_{j_k}) : 1 \leq j_1 < \cdots < j_k \leq n\}$$

is Schur convex. In addition, if g is increasing, then $\phi(x) \leq \phi(y)$ whenever $x \prec_w y$.

(b) *If $g : \mathbf{R}^n \to \mathbf{R}$ is convex, then $\phi(x) = \max\{g(Px) : P$ is a permutation matrix$\}$, is Schur-convex. In addition, if g is increasing, then $\phi(x) \leq \phi(y)$ whenever $x \prec_w y$.*

(c) *The* **variance** *function on real vectors $x = (x_1, \ldots, x_n)$ defined by*

$$V(x) = \frac{1}{n}\sum_{j=1}^{n}(x_j - \bar{x})^2 \quad \text{with} \quad \bar{x} = \left(\sum_{j=1}^{n} x_j\right)/n$$

is Schur-convex.

(d) *The* **entropy** *function on nonnegative vectors $x = (x_1, \ldots, x_n)$ is defined by*

$$H(x) = -\sum_{j=1}^{j} x_j \log x_j,$$

where by convention $t \log t = 0$ if $t = 0$. Then $-H(x)$ is Schur -convex, i.e., $H(x)$ is Schur-concave,

Definition 5.8 *For $k \in \{1, \ldots, n\}$, define the kth* **elementary symmetric function** *by*

$$E_k(x_1, \ldots, x_n) = \sum_{1 \leq j_1 < \cdots < j_k \leq n} x_{j_1} \cdots x_{j_k}.$$

Proposition 5.9 *Let $k \in \{2, \ldots, n\}$, the functions $-E_k(x_1, \ldots, x_n)$ and*

$$E_{k-1}(x_1, \ldots, x_n)/E_k(x_1, \ldots, x_n)$$

are Schur convex on positive vectors $x = (x_1, \ldots, x_n)$.

6 Inequalities on determinant and related functions

Theorem 6.1 *Let f be a Schur-convex function on real vectors $x = (x_1, \dots, x_n)$. If $A \in H_n$ has diagonal entries $d_1, \dots, d_n$ and eigenvalues $a_1, \dots, a_n$, then*

$$f(d_1, \dots, d_n) \leq f(a_1, \dots, a_n).$$

Theorem 6.2 *Let $A \in H_n$ have diagonal entries $d_1, \dots, d_n$ and eigenvalues $a_1, \dots, a_n > 0$. Then*

$$\frac{E_n(d_1, \dots, d_n)}{E_n(a_1, \dots, a_n)} \geq \frac{E_{n-1}(d_1, \dots, d_n)}{E_{n-1}(a_1, \dots, a_n)} \geq \cdots \geq \frac{E_1(d_1, \dots, d_n)}{E_1(a_1, \dots, a_n)} = 1.$$

Corollary 6.3 *Suppose $A \in M_n$. Then $|\det(A)| \leq \prod_{j=1} \|A_i\|$, where A_i is the ith column of A.*

Theorem 6.4 *If $A \in H_n$ is positive semi-definite, then*

$$(\det A)^{1/n} = \inf\{\mathrm{tr}\,(AB)/n : B \text{ is positive definite with } \det(B) = 1\}.$$

If A is invertible, then the infimum is attainable.

Corollary 6.5 (Minkowski) *Let $A, B \in H_n$ be positive. Then*

$$(\det(A+B))^{1/n} \geq (\det A)^{1/n} + (\det B)^{1/n}.$$

7 Exercises

1. Please make suggestions and comments to the lecture.

2. Provide the missing details of the proofs of various results such as Example 5.7, Proposition 5.9, and results in Section 6, etc.

3. Prove that if $A \in M_n$ is complex symmetric, there is unitary U such that $U^t AU = \mathrm{diag}\,(s_1(A), \dots, s_n(A))$. Hint: Choose unit vector u so that $\mathrm{Re}(u^t Au)$ is largest possible. Show that $U^t AU = [s_1] \oplus A_2$ if U has u as the first column. Induct on A_2.

4. Let $A \in H_n$. Show that a unit vector $v \in \mathbf{C}^n$ satisfies $v^* Av = \lambda_j(A)$ for $j = 1$ or n if and only if v is an eigenvector of $\lambda_j(A)$.

5. Let $A, B \in H_n$. Use Theorem 1.1 to show that

   $$\lambda_{i+j-1}(A+B) \leq \lambda_i(A) + \lambda_j(B), \qquad i + j - 1 \leq n.$$

   Formulate and prove a lower bound on $\lambda_{i+j-1}(A+B)$.