Lecture 4: Outline

- Chapter 4: Hermitian and symmetric matrices, Congruence

Magnus Jansson/Mats Bengtsson

Lecture 4: Hermitian matrices

Def: A matrix $A = [a_{ij}] \in M_n$ is *Hermitian* if $A = A^*$.
A is *skew-Hermitian* if $A = -A^*$.

Simple observations:

1. If A is Hermitian, then A^k and A^{-1} are Hermitian.
2. $A + A^*$ and AA^* are Hermitian and $A - A^*$ is skew-Hermitian for all $A \in M_n$.
3. Any $A \in M_n$ can be decomposed uniquely as $A = B + iC = B + D$ where B, C are Hermitian and D skew-Hermitian. In fact
$$B = \frac{1}{2}(A + A^*) \qquad D = iC = \frac{1}{2}(A - A^*)$$
4. A Hermitian matrix in M_n is completely described by n^2 real valued parameters.

Hermitian matrices cont'd

A is Hermitian iff

- x^*Ax is real for all $x \in \mathbf{C}^n$
- A is normal with real eigenvalues
- S^*AS is Hermitian for all $S \in M_n$

All eigenvalues of a Hermitian matrix are real and it has a complete set of orthonormal eigenvectors (the last fact follows as a special case of the spectral theorem for normal matrices).

Thm (spectral): $A \in M_n$ is Hermitian iff it is unitarily diagonalizable to a real diagonal matrix. A matrix A is real symmetric iff it can be diagonalized by a real orthogonal matrix to a real diagonal matrix.

Commutation of Hermitian matrices

Let $\mathcal{F}$ be a family of Hermitian matrices. Then all $A \in \mathcal{F}$ are simultaneously unitarily diagonalizable iff $AB = BA$ for all $A, B \in \mathcal{F}$.

Positive definiteness

A Hermitian matrix $A \in M_n$ is

Positive definite if $x^*Ax > 0$ for all $x \in \mathbb{C}^n$, $x \neq 0$.

Positive semidefinite if $x^*Ax \geq 0$ for all $x \in \mathbb{C}^n$, $x \neq 0$.

Negative definite if $x^*Ax < 0$ for all $x \in \mathbb{C}^n$, $x \neq 0$.

Negative semidefinite if $x^*Ax \leq 0$ for all $x \in \mathbb{C}^n$, $x \neq 0$.

Indefinite if there are $y, z \in \mathbb{C}^n$ with $y^*Ay < 0 < z^*Az$.

Much more on positive (semi)definiteness in Chapter 7

Quadratic forms

Bilinear form in two variables $Q(x,y) = y^TAx$

Sesquilinear form in two variables $Q(x,y) = y^*Ax$

Quadratic form Both $Q(x) = x^TAx$ and $Q(x) = x^*Ax$ are commonly called quadratic forms. See homework on the need to require A to be symmetric/hermitian.

Non-homogeneous quadratic form $x^TAx + b^Tx + c$ or $x^*Ax + \mathrm{Re}\{b^*x\} + c$.

Homogenization Extend the vector with a scalar constant,

$$x^*Ax + \mathrm{Re}\{b^*x\} + c = \tilde{x}^* \underbrace{\begin{bmatrix} A & \frac{b}{2} \\ \frac{b^T}{2} & c \end{bmatrix}}_{\tilde{A}} \tilde{x}, \text{ where } \tilde{x} = \begin{bmatrix} x \\ 1 \end{bmatrix}$$

Variational characterization of eigenvalues

Let $A \in M_n$ be Hermitian with eigenvalues $\lambda_1 \leq \cdots \leq \lambda_n$.

Thm (Rayleigh-Ritz):

$$\lambda_1 = \min_{x \neq 0} \frac{x^*Ax}{x^*x} = \min_{x^*x=1} x^*Ax$$

$$\lambda_n = \max_{x \neq 0} \frac{x^*Ax}{x^*x} = \max_{x^*x=1} x^*Ax$$

Thm (Courant-Fischer): Let S denote a subspace of $\mathbf{C}^n$. Then,

$$\lambda_k = \min_{\{S:\dim[S]=k\}} \max_{\substack{x \in S \\ x \neq 0}} \frac{x^*Ax}{x^*x}$$

$$\lambda_k = \max_{\{S:\dim[S]=n-k+1\}} \min_{\substack{x \in S \\ x \neq 0}} \frac{x^*Ax}{x^*x}$$

Applications of C-F thm

Thm: If $A, B \in M_n$ are Hermitian, then if $j + k \geq n + 1$

$$\lambda_{j+k-n}(A+B) \leq \lambda_j(A) + \lambda_k(B)$$

and if $j + k \leq n + 1$

$$\lambda_j(A) + \lambda_k(B) \leq \lambda_{j+k-1}(A+B)$$

Applications cont'd

Thm: If $A, B \in M_n$ are Hermitian, then

$$\lambda_k(A) + \lambda_1(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_n(B)$$

Interlacing theorem: Let $z \in \mathbf{C}^n$ and $A \in M_n$ be Hermitian. Then, for $k = 1, 2, \ldots, n-1$:

$$\lambda_k(A + zz^*) \le \lambda_{k+1}(A) \le \lambda_{k+1}(A + zz^*)$$
$$\lambda_k(A) \le \lambda_k(A + zz^*) \le \lambda_{k+1}(A)$$

$$\lambda_k(A - zz^*) \le \lambda_k(A) \le \lambda_{k+1}(A - zz^*)$$
$$\lambda_k(A) \le \lambda_{k+1}(A - zz^*) \le \lambda_{k+1}(A)$$

Applications cont'd

Interlacing theorem for bordered matrices:
Let $A \in M_n$ be Hermitian, $y \in \mathbf{C}^n$, $a \in \mathbf{R}$ and define

$$\hat{A} = \begin{bmatrix} A & y \\ y^* & a \end{bmatrix}$$

Then with $\lambda_i \in \sigma(A)$ and $\hat{\lambda}_i \in \sigma(\hat{A})$

$$\hat{\lambda}_1 \le \lambda_1 \le \hat{\lambda}_2 \le \cdots \le \hat{\lambda}_n \le \lambda_n \le \hat{\lambda}_{n+1}$$

The Poincaré separation theorem

Let $A \in M_n$ be Hermitian, let $U \in M_{n,r}$ be a matrix with $r \le n$ orthonormal columns and define $B_r = U^*AU$. Then

$$\lambda_k(A) \le \lambda_k(B_r) \le \lambda_{k+n-r}(A); \qquad k = 1, 2, \ldots, r$$

Application:

$$\min_{U,\ U^*U=I_r} \operatorname{Tr}(U^*AU) = \sum_{k=1}^{r} \lambda_k(A)$$

$$\max_{U,\ U^*U=I_r} \operatorname{Tr}(U^*AU) = \sum_{k=1}^{r} \lambda_{k+n-r}(A)$$

Note that equality is obtained by choosing the columns of U as suitable eigenvectors of A.

Generalized Rayleigh Quotients

Let $A \in M_n$ be Hermitian and $B \in M_n$ be Hermitian positive definite. Consider the following **generalized eigenvalue problem**

$$Ax = \lambda Bx$$

with eigenvalues $\lambda_1 \le \cdots \le \lambda_n$. Then,

$$\lambda_1 = \min_{x \ne 0} \frac{x^*Ax}{x^*Bx} = \min_{x^*Bx \ge 1} x^*Ax$$

$$\lambda_n = \max_{x \ne 0} \frac{x^*Ax}{x^*Bx} = \max_{x^*Bx \le 1} x^*Ax$$

Solve the generalized eigenvalue problem in Matlab using `[E,Lambda]=eig(A,B);`
Note: Elements of `Lambda` not sorted.

Majorization

Def: Let $\alpha = [\alpha_i] \in \mathbf{R}^n$ and $\beta = [\beta_i] \in \mathbf{R}^n$ with sorted versions, $\alpha_{j_1} \leq \alpha_{j_2} \leq \cdots \leq \alpha_{j_n}$ and $\beta_{m_1} \leq \beta_{m_2} \leq \cdots \leq \beta_{m_n}$.
If

$$\sum_1^n \alpha_i = \sum_1^n \beta_i$$

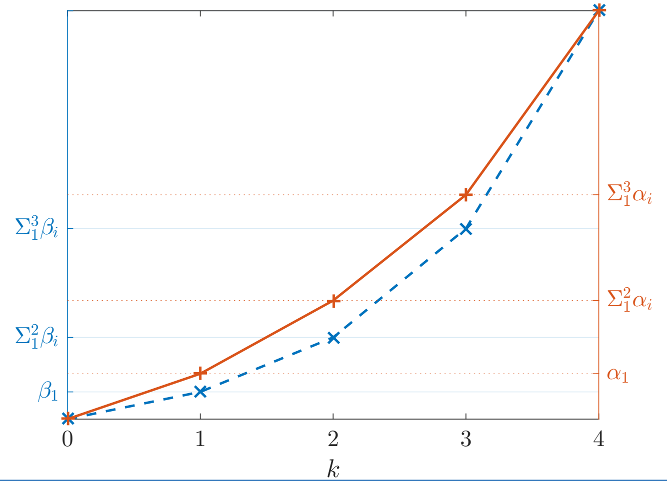
and

$$\sum_{i=1}^k \beta_{m_i} \leq \sum_{i=1}^k \alpha_{j_i} \quad \text{for all } k = 1, 2, \ldots, n,$$

then the vector β *majorizes* the vector α.

Note: The notation is not standardized, some texts (including 1st edition of Horn&Johnson) use the opposite definition.

Illustration of the definition, β majorizes α



Majorization cont'd

Thm: Let $A \in M_n$ be Hermitian. The vector of eigenvalues majorizes the vector of diagonal elements.

Converse thm: If the vector $\lambda \in \mathbf{R}^n$ majorizes the vector $a \in \mathbf{R}^n$ then there exists a real symmetric matrix $A \in M_n(\mathbf{R})$ with a_i as diagonal elements and λ_i as eigenvalues.

Thm: Let $A, B \in M_n$ be Hermitian and let $\lambda(A)$ be the vector of eigenvalues of A etc. The vector $\lambda(A) + \lambda(B)$ majorizes the vector $\lambda(A + B)$.

More to read on majorization

- Albert W. Marshall, Ingram Olkin, and Barry C Arnold. *Inequalities: Theory of Majorization and Its Applications.* Springer, New York, 2nd edition, 2011.
- Eduard Jorswieck and Holger Boche. Majorization and matrix-monotone functions in wireless communications. *Foundations and Trends® in Communications and Information Theory*, 3(6):553–701, 2007.
- Daniel P. Palomar and Yi Jiang. MIMO transceiver design via majorization theory. *Foundations and Trends® in Communications and Information Theory*, 3(4-5):331–551, 2007.

Complex symmetric matrices

Autonne-Takagi factorization: If $A \in M_n$ is symmetric, then $A = U\Sigma U^T$. Here, $U \in M_n$ and unitary, $\Sigma = diag\{\sigma_1, \ldots, \sigma_n\}$ is real and nonnegative. The columns of U can be taken as an orthonormal set of eigenvectors to $A\bar{A}$ and σ_i is the square root of an eigenvalue of $A\bar{A}$.

Thm: Every matrix $A \in M_n$ is similar to a symmetric matrix.

Thm: Let $A \in M_n$. There exist a nonsingular matrix S and a unitary matrix U such that $(US)A(\bar{U}S)^{-1}$ is a diagonal matrix with nonnegative elements.

Congruence

Def: Let $A, B \in M_n$ and S a nonsingular matrix.
If $B = SAS^*$, then B is *-congruent to A.
If $B = SAS^T$, then B is ^{T}-congruent to A.

Both congruence relations induce equivalence classes:

1. A is congruent to A
2. If A is congruent to B, then B is congruent to A.
3. If A is congruent to B and B is congruent to C, then A is congruent to C.

Inertia

Def: Let $A \in M_n$ be Hermitian. The *inertia* of A is the ordered triple

$$i(A) = (i_+(A), i_-(A), i_0(A))$$

where the entries correspond to the number of positive, negative and zero eigenvalues of A, respectively.
Note that the rank of A equals $i_+(A) + i_-(A)$.
The *signature* of A is $i_+(A) - i_-(A)$.

Canonical form/Sylvester's law of inertia

If $A \in M_n$ is Hermitian, then we can decompose it as

$$A = SI(A)S^*$$

where S is nonsingular and $I(A)$ is the *inertia matrix*

$$I(A) = \text{diag}(1 \ \ldots \ 1 \ -1 \ \ldots \ -1 \ 0 \ldots 0)$$

Thm (Syl): Let $A, B \in M_n$ be Hermitian. Then $A = SBS^*$ for a nonsingular matrix $S \in M_n$ iff A and B have the same inertia.

Quantitative Inertia Result / ^{T}-congruence

Thm: (Ostrowski) Let $A, S \in M_n$ where A is Hermitian. Let the eigenvalues be arranged in nondecreasing order. For each $k = 1, \ldots, n$ there exists a real number θ_k such that $\lambda_1(SS^*) \leq \theta_k \leq \lambda_n(SS^*)$ and

$$\lambda_k(SAS^*) = \theta_k \lambda_k(A)$$

Thm: Let $A, B \in M_n$ be symmetric matrices (real or complex). There is a nonsingular matrix $S \in M_n$ such that $A = SBS^T$ iff A and B have the same rank.

More about diagonalization by congruence: Thm 4.5.17 (4.5.15 in old ed.)