# Lecture 5: Eigenvalues of Hermitians Matrices

This lecture takes a closer look at Hermitian matrices and at their eigenvalues. After a few generalities about Hermitian matrices, we prove a minimax and maximin characterization of their eigenvalues, known as Courant–Fischer theorem. We then derive some consequences of this characterization, such as Weyl theorem for the sum of two Hermitian matrices, an interlacing theorem for the sum of two Hermitian matrices, and an interlacing theorem for principal submatrices of Hermitian matrices.

## 1 Basic properties of Hermitian matrices

We recall that a matrix $A \in \mathcal{M}_n$ is called Hermitian if $A^* = A$ and skew-Hermitian if $A^* = -A$, and we note that $A$ is Hermitian if and only if $iA$ is skew-Hermitian. We have observed earlier that the diagonal entries of a Hermitian matrix are real. This can also be viewed as a particular case of the following result.

**Proposition 1.** Given $A \in \mathcal{M}_n$,

$$[A \text{ is Hermitian}] \iff [\langle Ax, x\rangle = x^* A x \in \mathbb{R} \text{ for all } x \in \mathbb{C}^n].$$

*Proof.* $\Rightarrow$ If $A$ is Hermitian, then, for any $x \in \mathbb{C}^N$,

$$\langle Ax, x\rangle = \langle x, A^* x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle},$$

so that $\langle Ax, x\rangle \in \mathbb{R}$.
$\Leftarrow$ Suppose that $\langle Ax, x\rangle \in \mathbb{R}$ for all $x \in \mathbb{C}^n$. For any $u, v \in \mathbb{C}^N$, we have

$$\underbrace{\langle A(u+v), u+v\rangle}_{\in\mathbb{R}} = \underbrace{\langle Au, u\rangle}_{\in\mathbb{R}} + \underbrace{\langle Av, v\rangle}_{\in\mathbb{R}} + \langle Au, v\rangle + \langle Av, u\rangle, \qquad \text{so that } \langle Au, v\rangle + \langle Av, u\rangle \in \mathbb{R}.$$

Taking $u = e_j$ and $v = e_k$ yields

$$a_{k,j} \ + \ a_{j,k} \in \mathbb{R}, \qquad \text{thus } \mathrm{Im}(a_{k,j}) = -\mathrm{Im}(a_{j,k}),$$

then taking $u = ie_j$ and $v = e_k$ yields

$$ia_{k,j} - ia_{j,k} \in \mathbb{R}, \qquad \text{thus } \mathrm{Re}(a_{k,j}) = \mathrm{Re}(a_{j,k}).$$

Altogether, this gives $a_{k,j} = \overline{a_{j,k}}$ for all $j, k \in [1:n]$, i.e., $A = A^*$. □

**Proposition 2.** Any matrix $A \in \mathcal{M}_n$ can be uniquely written in the form

$$A = H + S, \qquad \text{where } H \in \mathcal{M}_n \text{ is Hermitian and } S \in \mathcal{M}_n \text{ is skew-Hermitian,}$$
$$A = H_1 + iH_2, \qquad \text{where } H_1, H_2 \in \mathcal{M}_n \text{ are both Hermitian.}$$

*Proof.* If $A = H + S$ with $H$ Hermitian and $S$ skew-Hermitian, then $A^* = H^* + S^* = H + S$. By adding and subtracting these two relations, we derive $H = (A + A^*)/2$ and $S = (A - A^*)/2$, hence $H$ and $S$ are uniquely determined. Moreover, with $H$ and $S$ given above, it is readily verified that $H$ is Hermitian, that $S$ is skew-Hermitian, and that $A = H + S$. For the second statement, we use the fact that $H_2$ is Hermitian if and only if $S := iH_2$ is skew-Hermitian. □

## 2 Variational characterizations of eigenvalues

We now recall that, according to the spectral theorem, if $A \in \mathcal{M}_n$ is Hermitian, there exists a unitary matrix $U \in \mathcal{M}_n$ and a real diagonal matrix $D$ such that $A = UDU^*$. The diagonal entries of $D$ are the eigenvalues of $A$, which we sort as

$$\lambda_1^\uparrow(A) \le \lambda_2^\uparrow(A) \le \cdots \le \lambda_n^\uparrow(A).$$

We utilize this notation for the rest of the lecture, although we may sometimes just write $\lambda_j^\uparrow$ instead of $\lambda_j^\uparrow(A)$ when the context is clear. Note also that the columns $u_1, \ldots, u_n$ of $U$ form an orthonormal system and that $Au_j = \lambda_j^\uparrow(A) u_j$ for all $j \in [1:n]$. We start by observing that

$$\lambda_1^\uparrow(A)\|x\|_2^2 \le \langle Ax, x\rangle \le \lambda_n^\uparrow(A)\|x\|_2^2, \tag{1}$$

with the leftmost inequality becoming an equality if $x = u_1$ and the rightmost inequality becoming an equality if $x = u_n$. The argument underlying the observation (1) will reappear several times (sometimes without explanation), so we spell it out here. It is based on the expansion of $x \in \mathbb{C}^n$ on the orthonormal basis $(u_1, \ldots, u_n)$, i.e.,

$$x = \sum_{j=1}^n c_j u_j \qquad \text{with} \qquad \sum_{j=1}^n c_j^2 = \|x\|_2^2.$$

We now simply write

$$\langle Ax, x\rangle = \langle \sum_{j=1}^n c_j A u_j, \sum_{j=1}^n c_j u_j \rangle = \langle \sum_{j=1}^n c_j \lambda_j^\uparrow u_j, \sum_{j=1}^n c_j u_j \rangle = \sum_{j=1}^n \lambda_j^\uparrow c_j^2 = \begin{cases} \ge \lambda_1^\uparrow \sum_{j=1}^n c_j^2 = \lambda_1^\uparrow \|x\|_2^2, \\ \le \lambda_n^\uparrow \sum_{j=1}^n c_j^2 = \lambda_n^\uparrow \|x\|_2^2. \end{cases} \tag{2}$$

The inequalities (1) (with the cases of equality) can also be expressed as $\lambda_1^\uparrow = \min_{\|x\|_2=1} \langle Ax, x\rangle$ and $\lambda_n^\uparrow = \max_{\|x\|_2=1} \langle Ax, x\rangle$, which is known as Rayleigh–Ritz theorem. It is a particular case of Courant–Fischer theorem stated below.

**Theorem 3.** For $A \in \mathcal{M}_n$ and $k \in [1:n]$,

$$\lambda_k^\uparrow(A) = \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle = \max_{\dim(V)=n-k+1} \min_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle. \tag{3}$$

**Remark.** This can also be stated with $\dim(V) \ge k$ and $\dim(V) \ge n-k+1$, respectively, or (following the textbook) as

$$\lambda_k^\uparrow(A) = \min_{w_1,\ldots,w_{n-k} \in \mathbb{C}^n} \max_{\substack{x \perp w_1,\ldots,w_{n-k} \\ \|x\|_2=1}} \langle Ax, x\rangle = \max_{w_1,\ldots,w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_k \\ \|x\|_2=1}} \langle Ax, x\rangle.$$

*Proof of Theorem 3.* We only prove the first equality — the second is left as an exercise. To begin with, we notice that, with $U := \mathrm{span}[u_1, \dots, u_k]$, we have

$$\min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle \le \max_{\substack{x \in U \\ \|x\|_2=1}} \langle Ax, x\rangle \le \lambda_k^{\uparrow}(A),$$

where the last inequality follows from an argument similar to (2). For the inequality in the other direction, we remark that our objective is to show that for any $k$-dimensional linear subspace $V$ of $\mathbb{C}^n$, there is $x \in V$ with $\|x\|_2 = 1$ and $\langle Ax, x\rangle \ge \lambda_k^{\uparrow}(A)$. Considering the subspace $W := \mathrm{span}[u_k, \dots, u_n]$ of dimension $n - k + 1$, we have

$$\dim(V \cap W) = \dim(V) + \dim(W) - \dim(V \cup W) \ge k + n - k + 1 - n = 1.$$

Hence we may pick $x \in V \cup W$ with $\|x\|_2 = 1$. The inequality $\langle Ax, x\rangle \ge \lambda_k^{\uparrow}(A)$ follows from an argument similar to (2). $\square$

We continue with some applications of Courant–Fischer theorem, starting with Weyl theorem.

**Theorem 4.** Let $A, B \in \mathcal{M}_n$ be Hermitian matrices. For $k \in [1:n]$,

$$\lambda_k^{\uparrow}(A) + \lambda_1^{\uparrow}(B) \le \lambda_k^{\uparrow}(A+B) \le \lambda_k^{\uparrow}(A) + \lambda_n^{\uparrow}(B).$$

*Proof.* We use Courant–Fischer theorem and inequality (1) to write

$$\begin{aligned}\lambda_k^{\uparrow}(A+B) &= \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \Big(\langle (A+B)x, x\rangle\Big) = \min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \Big(\langle Ax, x\rangle + \underbrace{\langle Bx, x\rangle}_{\le \lambda_n^{\uparrow}(B)}\Big) \\ &\le \Big(\min_{\dim(V)=k} \max_{\substack{x \in V \\ \|x\|_2=1}} \langle Ax, x\rangle\Big) + \lambda_n^{\uparrow}(B) = \lambda_k^{\uparrow}(A) + \lambda_n^{\uparrow}(B).\end{aligned}$$

This establishes the rightmost inequality. We actually use this result to prove the leftmost inequality by replacing $A$ with $A + B$ and $B$ with $-B$, namely

$$\lambda_k^{\uparrow}(A) = \lambda_k^{\uparrow}(A + B + (-B)) \le \lambda_k^{\uparrow}(A+B) + \lambda_n^{\uparrow}(-B) = \lambda_k^{\uparrow}(A+B) - \lambda_1^{\uparrow}(B).$$

A rearrangement gives the desired result. $\square$

**Corollary 5.** For Hermitian matrices $A, B \in \mathcal{M}_n$, if all the eigenvalues of $B$ are nonnegative (i.e., $\langle Bx, x\rangle \ge 0$ for all $x \in \mathbb{C}^n$, or in other words $B$ is positive semidefinite), then,

$$\lambda_k^{\uparrow}(A) \le \lambda_k^{\uparrow}(A+B) \qquad \text{for all } k \in [1:n].$$

Weyl theorem turns out to be the particular case $k = 1$ of Lidskii's theorem stated below with the sequences of eigenvalues arranged in nonincreasing order rather than nondecreasing order.

**Theorem 6.** For Hermitian matrices $A, B \in \mathcal{M}_n$, if $1 \le j_1 < \ldots < j_k \le n$, then

$$\sum_{\ell=1}^{k} \lambda_{j_\ell}^{\downarrow}(A+B) \le \sum_{\ell=1}^{k} \lambda_{j_\ell}^{\downarrow}(A) + \sum_{\ell=1}^{k} \lambda_{\ell}^{\downarrow}(B).$$

*Proof.* Replacing $B$ by $B - \lambda_{k+1}^{\downarrow}(B)I$, we may assume that $\lambda_{k+1}^{\downarrow}(B) = 0$. By the spectral theorem, there exists a unitary matrix $U \in \mathcal{M}_n$ such that

$$B = U\mathrm{diag}\big[\underbrace{\lambda_1^{\downarrow}(B), \ldots, \lambda_k^{\downarrow}(B)}_{\ge 0}, \underbrace{\lambda_{k+1}^{\downarrow}(B), \ldots, \lambda_n^{\downarrow}(B)}_{\le 0}\big]U^*.$$

Let us introduce the Hermitian matrices

$$B^+ := U\mathrm{diag}\big[\lambda_1^{\downarrow}(B), \ldots, \lambda_k^{\downarrow}(B), 0, \ldots, 0\big]U^*, \quad B^- := U\mathrm{diag}\big[0, \ldots, 0, -\lambda_{k+1}^{\downarrow}(B), \ldots, -\lambda_n^{\downarrow}(B)\big]U^*.$$

They have only nonnegative eigenvalues and satisfy $B = B^+ - B^-$. According to Corollary 5,

$$\lambda_j^{\downarrow}(A+B^+) \ge \lambda_j^{\downarrow}(A) \qquad \text{and} \qquad \lambda_j^{\downarrow}(A+B^+) = \lambda_j^{\downarrow}(A+B+B^-) \ge \lambda_j^{\downarrow}(A+B).$$

It follows that

$$\sum_{\ell=1}^{k} \big(\lambda_{j_\ell}^{\downarrow}(A+B) - \lambda_{j_\ell}^{\downarrow}(A)\big) \le \sum_{\ell=1}^{k} \big(\lambda_{j_\ell}^{\downarrow}(A+B^+) - \lambda_{j_\ell}^{\downarrow}(A)\big) \le \sum_{j=1}^{n} \big(\lambda_{j}^{\downarrow}(A+B^+) - \lambda_{j}^{\downarrow}(A)\big)$$
$$= \mathrm{tr}\,(A+B^+) - \mathrm{tr}\,(A) = \mathrm{tr}\,(B^+) = \sum_{\ell=1}^{k} \lambda_\ell^{\downarrow}(B).$$

This is just a rearrangement of the desired result. □

# 3 Interlacing theorems

The two results of this section locate the eigenvalues of a matrix derived from a matrix $A$ relatively to the eigenvalues of $A$. They are both consequences of Courant–Fischer theorem.

**Theorem 7.** Let $A \in \mathcal{M}_n$ be a Hermitian matrix and $A_s$ be an $s \times s$ principal submatrix of $A$, $s \in [1:n]$. Then, for $k \in [1:s]$,

$$\lambda_k^{\uparrow}(A) \le \lambda_k^{\uparrow}(A_s) \le \lambda_{k+n-s}^{\uparrow}(A).$$

**Remark.** The terminology of interlacing property is particularly suitable in the case of an $(n-1) \times (n-1)$ principal submatrix $\widetilde{A}$, since we then have

$$\lambda_1^{\uparrow}(A) \le \lambda_1^{\uparrow}(\widetilde{A}) \le \lambda_2^{\uparrow}(A) \le \lambda_2^{\uparrow}(\widetilde{A}) \le \lambda_3^{\uparrow}(A) \le \cdots \le \lambda_{n-1}^{\uparrow}(A) \le \lambda_{n-1}^{\uparrow}(\widetilde{A}) \le \lambda_n^{\uparrow}(A).$$

As for the particular case $s = 1$, it gives

$$\lambda_1^{\uparrow}(A) \le a_{j,j} \le \lambda_n^{\uparrow}(A) \qquad \text{for all } j \in [1:n],$$

which is also a consequence of (1) with $x = e_j$.

*Proof of Theorem 7.* Suppose that the rows and columns of $A$ kept in $A_s$ are indexed by a set $S$ of size $s$. For a vector $x \in \mathbb{C}^s$, we denote by $\widetilde{x} \in \mathbb{C}^n$ the vector whose entries on $S$ equal those of $x$ and whose entries outside $S$ equal zero. For a linear subspace $V$ of $\mathbb{C}^s$, we define $\widetilde{V} := \{\widetilde{x}, x \in V\}$, which is a subspace of $\mathbb{C}^n$ with dimension equal the dimension of $V$. Given $k \in [1:s]$, Courant–Fischer theorem implies that, for all linear subspace $V$ of $\mathbb{C}^s$ with $\dim(V) = k$,

$$\max_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A_s x, x\rangle = \max_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x}\rangle = \max_{\substack{\widetilde{x} \in \widetilde{V} \\ \|\widetilde{x}\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x}\rangle \geq \lambda_k^{\uparrow}(A).$$

Taking the minimum over all $k$-dimensional subspaces $V$ gives $\lambda_k^{\uparrow}(A_s) \geq \lambda_k^{\uparrow}(A)$. Similarly, for all linear subspace $V$ of $\mathbb{C}^s$ with $\dim(V) = s - k + 1 = n - (k + n - s) + 1$,

$$\min_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A_s x, x\rangle = \min_{\substack{x \in V \\ \|x\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x}\rangle = \min_{\substack{\widetilde{x} \in \widetilde{V} \\ \|\widetilde{x}\|_2 = 1}} \langle A\widetilde{x}, \widetilde{x}\rangle \leq \lambda_{k+n-s}^{\uparrow}(A).$$

Taking the maximum over all $(s-k+1)$-dimensional subspaces $V$ gives $\lambda_k^{\uparrow}(A_s) \leq \lambda_{k+n-s}^{\uparrow}(A)$. □

**Theorem 8.** Let $A, B \in \mathcal{M}_n$ be Hermitian matrices with $\operatorname{rank}(B) = r$. Then, for $k \in [1 : n-2r]$,

$$\lambda_k^{\uparrow}(A) \leq \lambda_{k+r}^{\uparrow}(A + B) \leq \lambda_{k+2r}^{\uparrow}(A).$$

Before turning to the proof, we observe that the $n \times n$ Hermitian matrices of rank $r$ are exactly the matrices of the form

(4) $$B = \sum_{j=1}^{r} \mu_j v_j v_j^*, \qquad \mu_1, \ldots, \mu_r \in \mathbb{R} \setminus \{0\}, \quad (v_1, \ldots, v_r) \text{ orthonormal system.}$$

Indeed, by the spectral theorem, the $n \times n$ Hermitian matrices of rank $r$ are exactly the matrices of the form

(5) $$B = V \operatorname{diag}\big[\mu_1, \ldots, \mu_r, 0, \ldots, 0\big] V^*, \qquad \mu_1, \ldots, \mu_r \in \mathbb{R} \setminus \{0\}, \quad V^*V = I,$$

which is just another way of writing (4).

*Proof of Theorem 8.* Let $k \in [1 : n - r]$ and let $B \in \mathcal{M}_n$ be as in (4). We use Courant–Fischer theorem to derive

$$\begin{aligned}
\lambda_k^{\uparrow}(A) &= \max_{w_1,\ldots,w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_k \\ \|x\|_2 = 1}} \langle Ax, x\rangle \leq \max_{w_1,\ldots,w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_k,v_1,\ldots,v_r \\ \|x\|_2 = 1}} \langle Ax, x\rangle \\
&\leq \max_{w_1,\ldots,w_k \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_k,v_1,\ldots,v_r \\ \|x\|_2 = 1}} \langle (A+B)x, x\rangle \leq \max_{w_1,\ldots,w_{k+r} \in \mathbb{C}^n} \min_{\substack{x \perp w_1,\ldots,w_{k+r} \\ \|x\|_2 = 1}} \langle (A+B)x, x\rangle \\
&= \lambda_{k+r}^{\uparrow}(A+B).
\end{aligned}$$

This establishes the rightmost inequality. This inequality can also be used to establish the leftmost inequality. Indeed, for $k \in [1 : k - 2r]$, we have $k + r \in [1 : k - r]$, and it follows that $\lambda_{k+r}^{\uparrow}(A + B) \leq \lambda_{k+r+r}^{\uparrow}(A + B + (-B)) = \lambda_{k+2r}^{\uparrow}(A)$. □

## 4 Exercises

Ex.1: Prove the second inequality in (3).

Ex.2: Verify in details the assertions made in (4) and (5).

Ex.3: Exercise 1 p. 174

Ex.4: Exercise 3 p. 174

Ex.5: Exercise 6 p. 174

Ex.6: Exercise 11 p. 175

Ex.7: Exercise 13 p. 175

Ex.8: Exercise 4 p. 181

Ex.9: Exercise 2 p. 198

Ex.10: Exercise 5 p. 199

Ex.11: Exercise 6 p. 199

Ex.12: Exercise 7 p. 199

Ex.13: Exercise 14 p. 200

Ex.14: Exercise 17 p. 200