Matrix Theory, Math6304
Lecture Notes from October 11, 2012

taken by Da Zheng

4 Variational characterization of eigenvalues, continued

We recall from last class that given a Hermitian matrix, we can obtain its largest (resp. smallest) eigenvalue by maximizing (resp. minimizing) the corresponding quadratic form over all the unit vectors. In fact, due to the following theorem by Courant and Fischer, we can obtain any eigenvalue of a Hermitian matrix through the "min-max" or "max-min" formula.

4.2 The Courant-Fischer Theorem

4.2.1 Theorem (Courant-Fischer)**.** *Suppose $A \in M_n$ is Hermitian, and for each $1 \leq k \leq n$, let $\{S_k^\alpha\}_{\alpha \in I_k}$ denote the set of all $k-$dimensional linear subspaces of $\mathbb{C}^n$. Also, enumerate the n eigenvalues $\lambda_1, \cdots, \lambda_n$ (counting multiplicity) in increasing order, i.e. $\lambda_1 \leq \cdots \leq \lambda_n$. Then, we have*

(i).

$$\min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \lambda_k,$$

(ii).

$$\max_{\alpha \in I_{n-k+1}} \min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \lambda_k.$$

Before starting the proof, we denote an or $u_1, \cdots, u_n$, which is orhonormal, as the eigenvectors of A corresponding to $\lambda_1, \cdots, \lambda_n$ respectively. .

Proof. (i). First, let $W = \text{span}\{u_1, \cdots, u_n\}$, then $\dim W = n-k+1$. So, for any $k-$dimensional subspace S_k^α, we should have $\dim(S_k^\alpha \cap W) \geq 1$. This is because of the equality: $\dim(S_k^\alpha + W) = \dim S_k^\alpha + \dim W - \dim(S_k^\alpha \cap W)$, and of course $\dim(S_k^\alpha + W) \leq n$.

Now, choose $x \in \left(S_k^\alpha \cap W\right) \setminus \{0\}$, note that $x = \sum_{j=k}^n \langle x, u_j\rangle u_j$ and $Au_j = u_j$, then it follows that

$$\frac{\langle Ax, x\rangle}{\|x\|^2} = \frac{\left\langle \sum_{j=k}^n \lambda_j \langle x, u_j\rangle u_j, x\right\rangle}{\|x\|^2}$$

$$
\begin{aligned}
&= \frac{\sum_{j=k}^{n} \lambda_j \langle x, u_j\rangle \langle u_j, x\rangle}{\|x\|^2} \\
&= \frac{\sum_{j=k}^{n} \lambda_j |\langle x, u_j\rangle|^2}{\|x\|^2} \\
&\geq \lambda_k \frac{\sum_{j=k}^{n} |\langle x, u_j\rangle|^2}{\|x\|^2} \\
&= \lambda_k.
\end{aligned}
$$

Here, we use the fact that $\lambda_k \leq \lambda_{k+1} \leq \cdots \leq \lambda_n$ and $\|x\|^2 = \sum_{j=k}^{n} |\langle x, u_j\rangle|^2$.

Thus, for any S_k^α,

$$\sup_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} \geq \lambda_k.$$

Indeed, for any S_k^α, it is easy to check that

$$\sup_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \sup_{x \in S_k^\alpha, \|x\|=1} \langle Ax, x\rangle,$$

and since $\{x \in S_k^\alpha : \|x\| = 1\}$ is compact, supremum is attained. So we have

$$\max_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \sup_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} \geq \lambda_k.$$

On the other hand, consider a particular $k-$dimensional subspace $S_k^\alpha = \text{span}\{u_1, \cdots, u_k\}$, then

$$\frac{\langle Ax, x\rangle}{\|x\|^2} = \frac{\sum_{j=1}^{k} \lambda_j |\langle x, u_j\rangle|^2}{\|x\|^2} \leq \lambda_k.$$

Hence, choosing $x = u_k$, we obtain $\max_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \lambda_k$.

This also implies that the minimum of $\max_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2}$ over all $\alpha \in I_k$ is also attained, and we conclude that

$$\min_{\alpha \in I_k} \max_{x \in S_k^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \lambda_k.$$

(ii). The proof of this formula follows from the same idea as (i), and we shall omit the similar details. We first choose $W = \text{span}\{u_1, \cdots, u_k\}$, so $\dim W = k$. Then, for any subspace S_{n-k+1}^α, $\dim(S_k^\alpha \cap W) \geq 1$.

Next, choose any $x \in \left(S_{n-k+1}^\alpha \cap W\right) \setminus \{0\}$, $\frac{\langle Ax, x\rangle}{\|x\|^2} \leq \lambda_k$, and therefore

$$\min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} \leq \lambda_k.$$

Now, we again choose a particular $S_{n-k+1} = \text{span}\{u_k, \cdots, u_n\}$, and this gives

$$\min_{x \in S_{n-k+1}^\alpha \setminus \{0\}} \frac{\langle Ax, x\rangle}{\|x\|^2} = \lambda_k.$$

So finally we conclude that

$$\max_{\alpha \in I_{n-k+1}} \min_{x \in S^{\alpha}_{n-k+1} \setminus \{0\}} \frac{\langle Ax, x \rangle}{\|x\|^2} = \lambda_k.$$

□

4.2.2 Remark. We can compare this result with theorem 4.2.11 in Horn and Johnson's "Matrix Analysis", which uses vectors to prove the "min-max" and "max-min" formulae, but the idea is essentially the same.

4.3 Eigenvalue estimates for sums of matrices

Next, we shall introduce several theorems and corollaries that can be considered as consequences of the Courant-Fischer's theorem. The first theorem, by Weyl, allows us to obtain a lower and upper bound for the kth eigenvalue of $A + B$.

4.3.3 Theorem (Weyl)**.** *Let $A, B \in M_n$ be both Hermitian, and $\{\lambda_j(A)\}_{j=1}^n$, $\{\lambda_j(B)\}_{j=1}^n$ and $\{\lambda_j(A+B)\}_{j=1}^n$ denote the sets of eigenvalues of A, B, and $A+B$ in increasing order, respectively. Then, for any $1 \le k \le n$,*

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

Notice that by symmetry, we naturally also have

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

Proof. By Rayleigh-Ritz's theorem, we know that

$$\lambda_1(B) \le \frac{\langle Bx, x \rangle}{\|x\|^2} \le \lambda_n(B), \quad \text{for } x \neq 0.$$

Then, by Courant-Fischer's theorem, for any $1 \le k \le n$,

$$\begin{aligned}
\lambda_k(A + B) &= \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_k \setminus \{0\}} \frac{\langle (A + B)x, x \rangle}{\|x\|^2} \\
&= \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_k \setminus \{0\}} \left(\frac{\langle Ax, x \rangle}{\|x\|^2} + \frac{\langle Bx, x \rangle}{\|x\|^2} \right) \\
&\ge \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_k \setminus \{0\}} \left(\frac{\langle Ax, x \rangle}{\|x\|^2} + \lambda_1(B) \right) \\
&= \lambda_1(B) + \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_k \setminus \{0\}} \frac{\langle Ax, x \rangle}{\|x\|^2} \\
&= \lambda_k(A) + \lambda_1(B).
\end{aligned}$$

For the other inequality, we apply similar argument and immediately obtain $\lambda_k(A + B) \le \lambda_k(A) + \lambda_n(B)$, and the proof is finished here. □

4.3.4 Remark. It is interesting to mention that for some special B, both lower and upper bounds can be attained in the above theorem we have just proved. For example, let $\{u_1, \cdots, u_n\}$ be the orthonormal set of eigenvectors of A with $Au_k = \lambda_k u_k$ for $1 \le k \le n$, and consider the rank one projection $B = au_k u_k^*$, where $a > 0$ or $a < 0$. Then, $A = UDU^*$, where $D = \text{diag}\{\lambda_1, \cdots, \lambda_n\}$, and we can assume that $\lambda_1, \cdots, \lambda_n$ is listed in increasing order. Also, it is easy to see that $B = UD'U^*$, where $D' = \text{diag}\{0, \cdots, 0, a, 0, \cdots, 0\}$ (a appears in the kth place). Then, we immediately have $\lambda_k(A+B) = \lambda_k(A) + \lambda_n(B) = \lambda_k(A) + a$ if $a > 0$; $\lambda_k(A+B) = \lambda_k(A) + \lambda_1(B) = \lambda_k(A) + a$ if $a < 0$.

The following corollary is called the *monotonicity theorem*, which refines the lower bound in Weyl's theorem by assuming B is positive semidefinite.

4.3.5 Definition. $B \in M_n$ is called *positive semidefinite*, if it is Hermitian and $\langle Bx, x\rangle \ge 0$ for all $x \in \mathbb{C}^n$.

Notice that $\langle Bx, x\rangle \ge 0$ for all $x \in \mathbb{C}^n$ indeed implies that B is Hermitian. However, for the real case, we have to impose that B is symmetric.

4.3.6 Corollary (Weyl)**.** *Adopt all the assumptions and notations in the above Weyl's theorem, if we further that suppose B is positive semideifnite, then*

$$\lambda_k(A) \le \lambda_k(A+B), \quad \textit{for all } 1 \le k \le n.$$

Proof. Since B is positive semidefinite, $\lambda_j(B) \ge 0$ for all $1 \le j \le n$, so the corollary follows from Weyl's theorem directly. □

The following theorem discusses the relationship between eigenvalues of a Hermitian matrix and those of the rank one perturbation of it, which is called the *interlacing theorem*. This is still an application of Courant-Fischer's theorem.

4.3.7 Theorem. *Let $A \in M_n$ be Hermitian, $z \in \mathbb{C}^n$, $\{\lambda_j(A)\}$ and $\{\lambda_j(A \pm zz^*)\}$ be both in increasing order, then*

(i).

$$\lambda_k(A \pm zz^*) \le \lambda_{k+1}(A) \le \lambda_{k+2}(A \pm zz^*), \quad \textit{for } 1 \le k \le n-2,$$

(ii).

$$\lambda_k(A) \le \lambda_{k+1}(A \pm zz^*) \le \lambda_{k+2}(A), \quad \textit{for } 1 \le k \le n-2.$$

endsection

Proof. By Courant-Fischer's theorem,

$$\begin{aligned}
\lambda_{k+2}(A \pm zz^*) &= \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_{k+2} \setminus \{0\}} \frac{\langle (A \pm zz^*)x, x\rangle}{\|x\|^2} \\
&= \min_{\alpha \in I_k} \max_{x \in S^{\alpha}_{k+2} \setminus \{0\}} \left(\frac{\langle Ax, x\rangle}{\|x\|^2} \pm \frac{|\langle x, z\rangle|^2}{\|x\|^2} \right) \\
&\ge \min_{\alpha \in I_k} \max_{\substack{x \in S^{\alpha}_{k+2} \setminus \{0\} \\ x \perp z}} \frac{\langle Ax, x\rangle}{\|x\|^2} \quad (*)
\end{aligned}$$

$$\geq \min_{\substack{\alpha\in I_k \\ z\perp S_{k+1}^\alpha}} \max_{x\in S_{k+1}^\alpha\setminus\{0\}} \frac{\langle Ax,x\rangle}{\|x\|^2} \quad (**)$$

$$\geq \min_{\alpha\in I_k} \max_{x\in S_{k+1}^\alpha\setminus\{0\}} \frac{\langle Ax,x\rangle}{\|x\|^2} \quad (***)$$

$$= \lambda_{k+1}(A).$$

Here, the inequality (*) and (***) are trivial. For inequality (**), note that $x \in S_{k+2}^\alpha \setminus \{0\}$ and $x \perp z$ is equivalent to $x \in S_{k+2}^\alpha \cap (\text{span}\{z\})^\perp$, and again by the equality $\dim(S_k^\alpha + W) = \dim S_k^\alpha + \dim W - \dim(S_k^\alpha \cap W)$, we know that $\dim(S_{k+2}^\alpha \cap (\text{span}\{z\})^\perp) = k+2$ or $k+1$.

Then, we see that for each $S_{k+2}^\alpha \cap (\text{span}\{z\})^\perp$, we can extract a $k+1-$dimensional subspace S_{k+1}^α such that $z \perp S_{k+1}^\alpha$, and therefore

$$\max_{\substack{x\in S_{k+2}^\alpha\setminus\{0\} \\ x\perp z}} \frac{\langle Ax,x\rangle}{\|x\|^2} \geq \max_{\substack{x\in S_{k+1}^\alpha\setminus\{0\} \\ z\perp S_{k+1}^\alpha}} \frac{\langle Ax,x\rangle}{\|x\|^2},$$

because on the right side we maximize over a subspace of that on the left.

Now, (**) becomes clear. As we finally want the minimum, and for each of the maximums in (*) over which we try to take the minimum, there exists a maximum in (**) which is smaller, we know that $\geq$ in (**) should hold.

For the other inequality of (i), we apply the analogous argument, again by the "max-min" formula in the Courant-Fischer's theorem,

$$\lambda_k(A \pm zz^*) = \max_{\alpha\in I_{n-k+1}} \min_{x\in S_{n-k+1}^\alpha\setminus\{0\}} \frac{\langle (A\pm zz^*)x,x\rangle}{\|x\|^2}$$

$$\leq \max_{\alpha\in I_{n-k+1}} \min_{\substack{x\in S_{n-k+1}^\alpha\setminus\{0\} \\ x\perp z}} \frac{\langle (A\pm zz^*)x,x\rangle}{\|x\|^2}$$

$$= \max_{\alpha\in I_{n-k+1}} \min_{\substack{x\in S_{n-k+1}^\alpha\setminus\{0\} \\ x\perp z}} \frac{\langle Ax,x\rangle}{\|x\|^2}$$

$$\leq \max_{\substack{\alpha\in I_{n-k} \\ z\perp S_{n-k}^\alpha}} \min_{x\in S_{n-k}^\alpha\setminus\{0\}} \frac{\langle (A\pm zz^*)x,x\rangle}{\|x\|^2} \quad (\Delta)$$

$$\leq \max_{\alpha\in I_{n-k}} \min_{x\in S_{n-k}^\alpha\setminus\{0\}} \frac{\langle (A\pm zz^*)x,x\rangle}{\|x\|^2}$$

$$= \lambda_{k+1}(A).$$

Here, inequality (Δ) follows from the similar argument as we did for (**). Thus, we proved (i).

(ii). This is indeed a direct corollary of (i), by modifying the indices. So the proof of the *interlacing theorem* is done.

□