Last Time (10/23/12)

Example of Low Rank Perturbation

Relationship Between Eigenvalues and Principal Submatrices: We started with the bordering case, and then began to generalize.

Before we commence today’s warm-up, let us begin by restating and proving the last theorem from the previous class (because we will then use it in the warm-up). This theorem generalizes eigenvalue interlacing to principal submatrices of smaller sizes.

4.6 Eigenvalue interlacing for principal submatrices, continued

4.6.1 Theorem. (Eigenvalue Interlacing for Principal Submatrices) Let $A \in M_n$ be Hermitian, and let $A_r \in M_r$ be a principal submatrix of $A$, where $r \in \mathbb{N}, r \leq n$. Then for $1 \leq k \leq r$, one has that:

$$
\lambda_k(A) \leq \lambda_k(A_r) \leq \lambda_{k+n-r}(A)
$$

Proof. Suppose $A_r$ is a principal submatrix of $A$, as above. Then one can choose a sequence from $A$ of "nested" principal submatrices $A_{r+1}, A_{r+2}, ..., A_{n-1}, A_n = A$, where $A_k \in M_k$, which satisfies:

- $A_r$ is a principal submatrix of $A_{r+1}$
- $A_{r+1}$ is a principal submatrix of $A_{r+2}$
- $A_{r+2}$ is a principal submatrix of $A_{r+3}$
- ...
- $A_{n-1}$ is a principal submatrix of $A_n = A$

It is not hard to see that each $A_k$ above satisfies the conditions for the corollary of our "bordering" theorem with respect to $A_{k+1}$, for $k \in \{r, r+1, ..., n-1\}$. Thus, by applying the corollary in an iterative fashion, we obtain:

$$
\lambda_k(A) \leq \lambda_k(A_{n-1}) \leq \cdots \leq \lambda_k(A_r)
$$

Similarly, we can apply the corollary to get:

$$
\lambda_k(A_r) \leq \lambda_{k+1}(A_{r+1}) \leq \cdots \leq \lambda_{k+n-r}(A_{r+n-r}) = \lambda_{k+n-r}(A)
$$
This completes the proof.

\[\square\]

**Warm-up**

We will define and then develop a test for positive definiteness.

**4.6.2 Definition.** A Hermitian matrix \(A \in M_n\) is **positive-definite** if all of its eigenvalues are strictly positive.

**4.6.3 Theorem.** A Hermitian matrix \(A_n \in M_n\) is positive-definite iff each of its leading principal submatrices \(\{A_r\}_{r=1}^n\) satisfies that \(\det(A_r) > 0\).

**Note:** We say a principal submatrix \(A_r\) of \(A_n\) is "leading" if it shares its top-left corner with \(A_n:\)

\[
A_n = \begin{pmatrix}
A_r & * & \cdots & *\\
* & \ddots & \cdots & * \\
\vdots & \ddots & \ddots & \ddots \\
* & \cdots & * & \ddots \\
& \cdots & * & \ddots \\
& & \cdots & * \\
& & & \ddots
\end{pmatrix}
\]

**Proof.**

(\(\Rightarrow\)): Suppose \(A_n\) is positive-definite. Then, using to the usual convention where \(\lambda_1(A_n)\) is the smallest eigenvalue of \(A_n\), we have:

\[
\lambda_1(A_n) > 0.
\]

Hence, using the preceding theorem, we have for each leading principal submatrix \(A_r, r \in \{1, 2, \ldots, n\}\), that:

\[
0 < \lambda_1(A_n) \leq \lambda_1(A_r)
\]

Thus:

\[
0 \leq \lambda_k(A_r) \text{ for all } k \in \{1, 2, \ldots, r\}.
\]

So the implication is shown now, because:

\[
\det(A_r) = \prod_{k=1}^{r} \lambda_k(A_r) > 0
\]

(\(\Leftarrow\)): Now suppose \(\det(A_r) > 0\) for \(r \in \{1, 2, \ldots, n\}\). In order to show that all of \(A_r'\)s eigenvalues are strictly positive, we will proceed inductively on its dimension, \(n\).

If \(n = 1\), there is nothing to show, because then \(A_n = A_r\) has only one eigenvalue given by \(\lambda_1(A_n) = \det(A_r) > 0\).

Now suppose the implication holds for all matrices up to size \(n\). (That is to say, for \(l \in \{1, 2, \ldots, n\}\), if all of the leading principal submatrices of \(A_l\) have positive determinant, then \(A_l\) is positive definite.) Now let \(A_{n+1} \in M_{n+1}\) and suppose it satisfies that \(\det(A_r) > 0\) for
If we can show that $A_{n+1}$ is positive definite (i.e., all its eigenvalues are positive), then we will be done. Our strategy will be to show that every leading principal submatrix $A_r$ of $A_{n+1}$ must also be positive-definite, which will then imply that $A_{n+1}$ (as a leading principal submatrix of itself) must be positive-definite as well.

To this end, we will induce over $r$. (Note: This is a sort of “nested induction”, since we are now inducing over $r$ to show than $n$ implies $n+1$.)

If $r=1$, then once again there is nothing to show, since then its only eigenvalue is given by $\lambda_1(A_1) = \text{det}(A_1) > 0$.

Now suppose, for $1 \leq r \leq k$, that all leading principal submatrices $A_r$ of $A_{n+1}$ are positive-definite. That is to say, let’s assume that $A_r$ has strictly positive eigenvalues, for $1 \leq r \leq k$.

Now let us examine $A_{k+1}$. First, observe that we have a bordering situation:

$$A_{k+1} = \begin{pmatrix}
A_k & 
\ast \\
\vdots & 
\ast \\
\ast & 
\ast \\
\ast & 
\ast
\end{pmatrix}$$

Now, combine the bordering theorem from last class with our inductive assumption that $A_k$ is positive definite to get:

$$0 < \lambda_{j-1}(A_k) \leq \lambda_j(A_{k+1}), \text{ for } j \in \{2, \ldots, k+1\} \quad (1)$$

Furthermore, it is a standard fact that:

$$\text{det}(A_{k+1}) = \prod_{j=1}^{k+1} \lambda_j(A_{k+1}) = \lambda_1(A_{k+1}) \prod_{j=2}^{k+1} \lambda_j(A_{k+1}) \quad (2)$$

Finally, we combine (1) and (2) with our assumption that $\text{det}(A_{k+1}) > 0$ to get:

$$0 < \frac{\text{det}(A_{k+1})}{\prod_{j=2}^{k+1} \lambda_j(A_{k+1})} = \lambda_1(A_{k+1})$$

Thus, $A_{k+1}$ is positive definite, which shows inductively that all leading principal submatrices of $A_{n+1}$ must be positive-definite. In turn, this shows that $A_{n+1}$, as a principal submatrix of itself, must be positive-definite. Finally, this completes our “first” induction step, allowing us to conclude the implication.

This completes the warm-up. Now let’s resume the discussion of eigenvalue interlacing for principal submatrices.

We can also relate the eigenvalues of a Hermitian matrix $A$ with the eigenvalues of principal submatrices of matrices which are uniformly equivalent to $A$. This relationship is made clear in the following corollary.
4.6.4 Corollary. Suppose $A \in M_n$ is Hermitian and $\{u_1, u_2, \ldots, u_r\} \subseteq \mathbb{C}^n$ is an orthogonal system, with $r \leq n$. Now set $B_r = [(Au_j, u_i)]_{i,j=1}^r$ and order the eigenvalues of $A$ and $B_r$ as usual. Then:

$$\lambda_k(A) \leq \lambda_k(B_r) \leq \lambda_{k+n-r}(A).$$

Proof. Let us begin by using the extension theorem and then the Gramm-Schmidt algorithm to obtain an orthonormal basis, $\{u_1, u_2, \ldots, u_r, u_{r+1}, \ldots, u_n\}$, for $\mathbb{C}^n$. This gives us the following unitarily equivalent matrix:

$$A' = [(Au_j, u_i)]_{i,j=1}^n = U^*AU,$$

where $U = \left( \begin{array}{cccc} u_1 & u_2 & \cdots & u_n \end{array} \right)$ is unitary.

Since $A$ and $A'$ are uniformly equivalent, they have the same eigenvalues. Finally, since $B_r$ is clearly a principal submatrix of $A'$, the desired inequality follows by the preceding interlacing theorem for principal submatrices.

Now we can formulate a generalized, more geometric version of Raleigh-Ritz.

4.6.5 Corollary. For a subspace $S_r \subseteq \mathbb{C}^n$ with $\dim(S_r) = r$, let $P_{S_r}$ denote the orthogonal projection onto $S_r$, and let $A \in M_n$ be Hermitian with eigenvalues $\{\lambda_j\}_{j=1}^n$ in nondecreasing order. Then:

$$\lambda_1(A) + \lambda_2(A) + \ldots + \lambda_r(A) = \min_{S_r} \text{tr}[P_{S_r}A]$$

and

$$\lambda_{n-r+1}(A) + \lambda_{n-r+2}(A) + \ldots + \lambda_n(A) = \max_{S_r} \text{tr}[P_{S_r}A].$$

Note: As with the Courant-Fisher theorem, we are minimizing/maximizing over all possible subspaces of the given dimension, $r$.

Proof. Given a subspace $S_r$ as above, let $\beta = \{u_1, u_2, \ldots, u_r\}$ be a corresponding orthonormal basis. Then the action determined by orthogonally projecting a vector $x \in \mathbb{C}^n$ onto $S_r$ is given by:

$$P_{S_r}x = \sum_{j=1}^r \langle x, u_j \rangle u_j$$

It can then be shown that $P_{S_r} = UU^*$, where:

$$U = [u_1 | u_2 | \cdots | u_n] \in M_{n,r}.$$ 

Now, extend $\beta$ to an orthonormal basis $\beta' = \{u_1, \ldots, u_r, u_{r+1}, \ldots, u_n\}$ and define $\tilde{U}$ to be the unitary matrix given as:

$$\tilde{U} = [u_1 | u_2 | \cdots | u_r | u_{r+1} | \cdots | u_n] \in M_n.$$
Now let \( A \in M_n \) be Hermitian. Then, we have:

\[
\text{tr}[P_S A] = \text{tr}[\tilde{U} P_S A U] \quad \text{(because trace is unitarily invariant)}
\]

\[
= \sum_{j=1}^{n} \langle P_S u_j, u_j \rangle \quad \text{(this is how the sum of the diagonal entries are given)}
\]

\[
= \sum_{j=1}^{n} \langle Au_j, P_S u_j \rangle \quad \text{(because } P_S \text{ is self-adjoint)}
\]

\[
= \sum_{j=1}^{r} \langle Au_j, P_S u_j \rangle \quad \text{(because } P_S u_j = 0 \text{ unless } j \leq r)
\]

\[
= \text{tr}[B_r]
\]

Here, \( B_r = \langle Au_j, u_i \rangle_{i,j=1}^{r} \), and it is related to \( A \) as in the previous corollary, which results in the following \( r \) inequalities:

\[
\lambda_k(A) \leq \lambda_k(B_r), \text{ for } 1 \leq k \leq r
\]

Summing these \( r \) inequalities, we obtain a new inequality:

\[
\lambda_1(A) + ... + \lambda_r(A) \leq \lambda_1(B_r) + ... + \lambda_r(B_r)
\]

\[
= \text{tr}[B_r] \quad \text{(because the trace of a square matrix is the sum of its eigenvalues)}
\]

\[
= \text{tr}[P_S A] \quad \text{(by the above equality)}
\]

We have shown "generally" that this inequality must always hold. It remains to show that the inequality actually becomes an equality when we minimize over all possible \( r \)-dimensional subspaces, \( S_r \). To this end, suppose our subspace \( S_r \) was chosen specially so that its orthonormal basis, \( \beta = \{u_1, ..., u_r\} \), satisfies that each \( u_i \) is an eigenvector corresponding to \( \lambda_i(A) \). In this special case, \( B_r \) looks like:

\[
B_r = [\langle Au_j, u_i \rangle_{i,j=1}^{r}]
\]

\[
= [\langle \lambda_j(A) u_j, u_i \rangle_{i,j=1}^{r}]
\]

\[
= [\lambda_j(A) \langle u_j, u_i \rangle]_{i,j=1}^{r}
\]

\[
= [\lambda_j(A) \delta_{i,j}]_{i,j=1}^{r}
\]

\[
= \begin{pmatrix}
\lambda_1(A) & 0 & \cdots & 0 \\
0 & \lambda_2(A) & \cdots & 0 \\
0 & 0 & \ddots & 0 \\
0 & 0 & \cdots & \lambda_r(A)
\end{pmatrix}
\]

Hence, for this specially chosen subspace, we obtain:

\[
\lambda_1(A) + ... + \lambda_r(A) = \text{tr}[B] = \text{tr}[P_S A]
\]
It is in this sense that we mean:

$$\lambda_1(A) + \lambda_2(A) + ... + \lambda_r(A) = \min_{S_r} tr[P_{S_r}A]$$

Now we’ll show the other claim, which is:

$$\lambda_{n-r+1}(A) + \lambda_{n-r+2}(A) + ... + \lambda_n(A) = \max_{S_r} tr[P_{S_r}A]$$

First, let us make two observations:

$$\lambda_{n+1-j}(A) = -\lambda_j(-A)$$

And:

$$\max_{S_R} tr[P_{S_r}A] = -\min_{S_R} tr[P_{S_r}(-A)]$$

Combining these two observations with the first result, we see that:

$$\lambda_{n-r+1}(A) + \lambda_{n-r+2}(A) + ... + \lambda_n(A) = \lambda_{(n+1)-r}(A) + \lambda_{(n+1)-(r-1)}(A) + ... + \lambda_{(n+1)-1}(A)$$

$$= -\lambda_r(-A) - \lambda_{(r-1)}(-A) - ... - \lambda_1(-A)$$

$$= -(\min_{S_r} tr[P_{S_r}(-A)])$$

$$= \max_{S_R} tr[P_{S_r}A]$$

This concludes the proof. \qed

4.7 Majorization

Because our results for spectrum analysis up to this point have been heavily dependent on minimizing/maximizing various mathematical objects (ie, Raleigh-Ritz, Courant-Fisher, the preceding corollary, etc.), the following question emerges naturally: what kind of information can we extrapolate about eigenvalues without solving optimization problems? It turns out that, in some cases, one can determine more information about a matrix by studying sums of subsets of its eigenvalues than by studying the individual eigenvalues themselves. This idea motivates the following definition.

4.7.6 Definition. Let \( \{\alpha_j\}_{j=1}^n \) and \( \{\beta_j\}_{j=1}^n \) be real sequences. Then we say \( \beta \) majorizes \( \alpha \), or \( \alpha \preceq \beta \), if \( \sum_{i=1}^k \beta_{m_i} \geq \sum_{i=1}^k \alpha_{j_i} \) for all \( 1 \leq k \leq n \), where, for such a given \( k \), \( m_i \) and \( j_i \) are permutations such that \( \alpha_{j_1} \leq \alpha_{j_2} \leq ... \leq \alpha_{j_n} \) and \( \beta_{m_1} \leq \beta_{m_2} \leq ... \leq \beta_{m_n} \)