CS724: Topics in Algorithms
Variational Results in Linear Algebra

Prof. Dan A. Simovici

Theorem

(Ky Fan's Theorem) *Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix such that spec$(A) = \{\lambda_1, \ldots, \lambda_n\}$, where $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ and let $V = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n)$ be a matrix whose columns consists of the corresponding unit eigenvectors of A,*

Let $\{\mathbf{x}_1, \ldots, \mathbf{x}_q\}$ be an orthonormal set of vectors in $\mathbb{R}^n$. For any positive integer $q \leqslant n$, the sums $\sum_{i=1}^{q} \lambda_i$ and $\sum_{i=1}^{q} \lambda_{n+1-i}$ equal, respectively, the maximum and minimum of $\sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j$.

The maximum of $\sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j$ is obtained by choosing the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_q$ as the first q columns of V; the minimum is obtained by assigning to $\mathbf{x}_1, \ldots, \mathbf{x}_q$ the last q columns of V.

Proof

Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal set of eigenvectors of A and let $X = (\mathbf{x}_1 \ \cdots \ \mathbf{x}_q) \in \mathbb{C}^{n \times q}$. Note that, by hypothesis, X is an orthogonal matrix.

The vectors $\mathbf{x}_i$ can be expressed as linear combinations of the vectors in V as

$$\mathbf{x}_i = \mathbf{v}_1 b_{1i} + \cdots + \mathbf{v}_n b_{ni} = V\mathbf{b}_i$$

for $1 \leqslant i \leqslant q$, or in matrix form as

$$(\mathbf{x}_1 \ \cdots \ \mathbf{x}_q) = (\mathbf{v}_1 \ \cdots \ \mathbf{v}_n) B$$

where $B = (\mathbf{b}_1 \ \cdots \ \mathbf{b}_q) \in \mathbb{C}^{q \times q}$. More succinctly, we have for $X = VB$. Note that $X'X = B'V'VB = B'B = I_q$, so B is also orthogonal.

Proof cont'd

We have

$$
\begin{aligned}
& \mathbf{x}_j' A \mathbf{x}_j \\
& = \mathbf{b}_j' V' A V \mathbf{b}_j \\
& = \mathbf{b}_j' \text{diag}(\lambda_1, \cdots, \lambda_n) \mathbf{b}_j = \sum_{i=1}^{n} \lambda_i b_{ij}^2 \\
& = \lambda_q \sum_{p=1}^{n} b_{pj}^2 + \sum_{p=1}^{q} (\lambda_p - \lambda_q) b_{pj}^2 + \sum_{p=q+1}^{n} (\lambda_p - \lambda_q) b_{pj}^2 .
\end{aligned}
$$

This implies $\mathbf{x}_j' A \mathbf{x}_j \leqslant \lambda_q + \sum_{p=1}^{q} (\lambda_p - \lambda_q) b_{pj}^2$. Therefore,

$$
\sum_{i=1}^{q} \lambda_i - \sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j \geqslant \sum_{i=1}^{q} (\lambda_i - \lambda_q) \left(1 - \sum_{j=1}^{q} b_{ij}^2 \right) .
$$

We have $\sum_{j=1}^{q} b_{ij}^2 \leqslant \| \mathbf{x}_i \|^2 = 1$, so $\sum_{i=1}^{q} (\lambda_i - \lambda_q) \left(1 - \sum_{j=1}^{q} b_{ij}^2\right) \geqslant 0$.
The left member becomes 0, when $\mathbf{x}_i = \mathbf{v}_i$, so $\sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j \leqslant \sum_{i=1}^{q} \lambda_i$.
The maximum of $\sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j$ is obtained when $\mathbf{x}_j = \mathbf{v}_j$ for $1 \leqslant j \leqslant q$, that is, when X consists of the first q columns of V that correspond to eigenvectors of the top k largest eigenvalues.
The argument for the minimum is similar.

Observe that the orthonormality condition of the set $\{\mathbf{x}_1, \ldots, \mathbf{x}_q\}$ can be expressed as $Y_q' Y_q = I_q$, where $Y_q \in \mathbb{C}^{n\times q}$ is the matrix $Y_q = (\mathbf{x}_1 \ \cdots \ \mathbf{x}_q)$. Also, the sum $\sum_{j=1}^{q} \mathbf{x}_j' A \mathbf{x}_j$ equals $trace(Y_q' A Y_q)$. Therefore, Ky Fan's Theorem implies that the sums $\sum_{i=1}^{q} \lambda_i$ and $\sum_{i=1}^{q} \lambda_{n+1-i}$ are, respectively, the maximum and minimum of $trace(Y_q' A Y_q)$, where $Y_q' Y_q = I_q$.

Theorem

(Rayleigh-Ritz Theorem) *Let* $A \in \mathbb{C}^{n \times n}$ *be a Hermitian matrix and let* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *be its eigenvalues, where* $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$*. Define the* Rayleigh-Ritz function $ral_A : \mathbb{R}^n - \{\mathbf{0}\} \longrightarrow \mathbb{R}$ *as*

$$ral_A(\mathbf{x}) = \frac{\mathbf{x}^H A \mathbf{x}}{\mathbf{x}^H \mathbf{x}}$$

Then,

$$\lambda_1 \geqslant ral_A(\mathbf{x}) \geqslant \lambda_n$$

for $\mathbf{x} \in \mathbb{C}^n - \{\mathbf{0}_n\}$.

Proof

Since A is Hermitian, there exists a unitary matrix P and a diagonal matrix T such that $A = P^{\mathsf{H}} T P$ and the diagonal elements of T are the eigenvalues of A, that is, $T = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$. This allows us to write

$$\mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{x}^{\mathsf{H}} P^{\mathsf{H}} T P \mathbf{x} = (P\mathbf{x})^{\mathsf{H}} T P \mathbf{x} = \sum_{j=1}^{n} \lambda_i |(P\mathbf{x})_i|^2,$$

which implies

$$\lambda_1 \parallel P\mathbf{x} \parallel^2 \geqslant \mathbf{x}^{\mathsf{H}} A \mathbf{x} \geqslant \lambda_n \parallel P\mathbf{x} \parallel^2 .$$

Proof cont'd

Since P is unitary we also have

$$\parallel P\mathbf{x} \parallel^2 = \mathbf{x}^{\mathsf{H}} P^{\mathsf{H}} P\mathbf{x} = \mathbf{x}^{\mathsf{H}}\mathbf{x},$$

which implies

$$\lambda_1 \mathbf{x}^{\mathsf{H}}\mathbf{x} \geqslant \mathbf{x}^{\mathsf{H}} A\mathbf{x} \geqslant \lambda_n \mathbf{x}^{\mathsf{H}}\mathbf{x},$$

for $\mathbf{x} \in \mathbb{C}^n$.

Corollary

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix and let $\lambda_1, \lambda_2, \ldots, \lambda_n$ be its eigenvalues, where $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. We have

$$\begin{aligned} \lambda_1 &= \max\{\mathbf{x}^H A\mathbf{x} \mid \mathbf{x}^H\mathbf{x} = 1\}, \\ \lambda_n &= \min\{\mathbf{x}^H A\mathbf{x} \mid \mathbf{x}^H\mathbf{x} = 1\}. \end{aligned}$$

Proof.

Note that if $\mathbf{x}$ is an eigenvector that corresponds to λ_1, then $A\mathbf{x} = \lambda_1\mathbf{x}$, so $\mathbf{x}^{\mathrm{H}}A\mathbf{x} = \lambda_1\mathbf{x}^{\mathrm{H}}\mathbf{x}$; in particular, if $\mathbf{x}^{\mathrm{H}}\mathbf{x} = 1$ we have $\lambda_1 = \mathbf{x}^{\mathrm{H}}A\mathbf{x}$, so

$$\lambda_1 = \max\{\mathbf{x}^{\mathrm{H}}A\mathbf{x} \mid \mathbf{x}^{\mathrm{H}}\mathbf{x} = 1\}.$$

The equality for λ_n can be shown in a similar manner. □

We discuss next an important result that is a generalization of Rayleigh-Ritz Theorem.

Theorem

(Courant-Fisher Theorem) *Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix having the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. We have*

$$\lambda_k = \min_{dim(S)=n-k+1} \max_{\mathbf{x}} \{\mathbf{x}^H A \mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\},$$

and

$$\lambda_k = \max_{dim(S)=k} \min_{\mathbf{x}} \{\mathbf{x}^H A \mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\},$$

where S ranges over the subspaces of $\mathbb{C}^n$.

Proof

There exists a unitary matrix U and a diagonal matrix D such that $A = U^{\mathsf{H}} D U$ and the diagonal elements of D are the eigenvalues of A, that is, $D = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$.
We prove initially that

$$\lambda_k = \min_{dim(S)=n-k+1} \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} D \mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\}.$$

For $\dim(S) = n - k + 1$ define $\tilde{S}$ as the set of unit vectors in the subspace S, that is,

$$\tilde{S} = \{\mathbf{y} \in \mathbb{C}^n \mid y \in S \text{ and } \parallel \mathbf{y} \parallel = 1\}$$

and $\hat{S} = S \cap \langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle$. We have $\hat{S} = S \cap \langle \mathbf{e}_1, \ldots, \mathbf{e}_k \rangle \neq \{\mathbf{0}_n\}$ because otherwise the dimension of the subspace generated by $S \cup \{\mathbf{e}_1, \ldots, \mathbf{e}_k\}$ would exceed $n + 1$.

Proof cont'd

Therefore, $\hat{S}$ consists of vectors of $\tilde{S}$ having the form

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_k \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

such that $\sum_{i=1}^{k} y_i^2 = 1$. So, if $\dim(S) = n - k + 1$ we have

$$\mathbf{y}^{\mathsf{H}} D \mathbf{y} = \sum_{i=1}^{k} \lambda_i |y_i|^2 \geqslant \lambda_k \sum_{i=1}^{k} |y_i|^2 = \lambda_k$$

for all $\mathbf{y} \in \hat{S}$.

Proof cont'd

Since $\hat{S} \subseteq \tilde{S}$ it follows that $\max_{\mathbf{y} \in \tilde{S}} \mathbf{y}^{\mathsf{H}} D\mathbf{y} \geqslant \max_{\mathbf{y} \in \hat{S}} \mathbf{y}^{\mathsf{H}} D\mathbf{y} \geqslant \lambda_k$, so

$$\min_{dim(S)=n-k+1} \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} D\mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\} \geqslant \lambda_k.$$

Let now S be $S = \langle \mathbf{e}_1, \ldots, \mathbf{e}_{k-1} \rangle^{\perp}$. Clearly, $\dim(S) = n - k + 1$. A vector $\mathbf{y} \in S$ has the form

$$\mathbf{y} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ y_k \\ \vdots \\ y_n \end{pmatrix}$$

Therefore,

$$\mathbf{y}^{\mathsf{H}} D \mathbf{y} = \sum_{i=k}^{n} \lambda_i |y_i|^2 \leq \lambda_i \sum_{i=k}^{n} |y_i|^2 = \lambda_i$$

for all $\mathbf{y} \in \{\mathbf{y} \in S \mid \parallel \mathbf{y} \parallel_2 = 1\}$. This implies

$$\min_{dim(S)=n-k+1} \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} D \mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\} \leqslant \lambda_k,$$

which yields the desired equality.

The matrices A and D have the same eigenvalues. Also $\mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{x}^{\mathsf{H}} A \mathbf{x} = \mathbf{x}^{\mathsf{H}} U^{\mathsf{H}} D U \mathbf{x} = (U\mathbf{x})^{\mathsf{H}} D (U\mathbf{x})$ and $\parallel U\mathbf{x} \parallel_2 = \parallel \mathbf{x} \parallel_2$, because U is a unitary matrix. This yields the first equality of the theorem. The proof of the second part of the theorem is entirely similar.

Another form of Courant-Fishers Theorem can be obtained by observing that every p-dimesional subspace S of $\mathbb{C}^n$ is the orthogonal space of an $(n-p)$-dimensional subspace. Therefore, for each p-dimensional subspace S there is a sequence of $n-p$ vectors $\mathbf{w}_1, \ldots, \mathbf{w}_{n-p}$ (which is a basis of $S^\perp$) such that $S = \{\mathbf{x} \in \mathbb{C}^n \mid \mathbf{x} \perp \mathbf{w}_1, \ldots, \mathbf{x} \perp \mathbf{w}_{n-p}\}$.

Theorem

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix having the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. We have

$$\begin{aligned}\lambda_k &= \min_{\mathbf{w}_1,\ldots,\mathbf{w}_{k-1}} \max_{\mathbf{x}} \{\mathbf{x}^H A\mathbf{x} \mid \mathbf{x} \perp \mathbf{w}_1, \ldots, \mathbf{x} \perp \mathbf{w}_{k-1} \text{ and } \parallel \mathbf{x} \parallel_2 = 1\},\\ &= \max_{\mathbf{w}_1,\ldots,\mathbf{w}_{n-k}} \min_{\mathbf{x}} \{\mathbf{x}^H A\mathbf{x} \mid \mathbf{x} \perp \mathbf{w}_1, \ldots, \mathbf{x} \perp \mathbf{w}_{n-k}, \text{ and } \parallel \mathbf{x} \parallel_2 = 1\}.\end{aligned}$$

An interesting special case occurs when $A \in \mathbb{C}^{n\times n}$ is a Hermitian matrix which has the least eigenvalue equal to 0 and the corresponding eigenvector $\mathbf{1}_n$. In this case, the second smallest eigenvalue λ_2 is given by

$$\lambda_2 = \min_{dim(S)=n-1} \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A\mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\}, \tag{1}$$

and

$$\lambda_2 = \max_{dim(S)=2} \min_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A\mathbf{x} \mid \mathbf{x} \in S \text{ and } \parallel \mathbf{x} \parallel_2 = 1\}, \tag{2}$$

where S is a subspace of $\mathbb{C}^n$.

Lemma

Let $\{\{i_1, \ldots, i_k\}$ *be a subset of the set* $\{1, \ldots, n\}$, *where* $i_1 < \cdots < i_k$.
Let $A \in \mathbb{C}^{n\times n}$ *be a matrix and let* $B = A\begin{bmatrix} i_1 \cdots i_k \\ i_1 \cdots i_k \end{bmatrix} \in \mathbb{C}^{k\times k}$ *be a principal submatrix of* A. *Let* $\mathbf{y} \in \mathbb{C}^k$ *and let* $f : \{1, \ldots, k\} \longrightarrow \{1, \ldots, n\}$ *be an injective function.*
Define $\mathbf{x} \in \mathbb{C}^n$ *such that*

$$x_r = \begin{cases} y_i & \text{if } f(i) = r, \\ 0 & \text{otherwise} \end{cases}$$

for $1 \leqslant r \leqslant n$. *We have* $\mathbf{y}^H B\mathbf{y} = \mathbf{x}^H A\mathbf{x}$.

Proof

Observe that if r does not belong to the range of f then $x_r = 0$. The definition of $\mathbf{x}$ implies

$$\begin{aligned}
\mathbf{x}^{\mathsf{H}} A \mathbf{x} &= \sum_{r=1}^{n} \sum_{s=1}^{n} \overline{x}_r a_{rs} x_s \\
&= \sum_{i=1}^{k} \sum_{j=1}^{k} \overline{y}_i b_{ij} y_j \\
&\quad (\text{if } f(i) = r \text{ and } f(j) = s) \\
&= \mathbf{y}^{\mathsf{H}} B \mathbf{y}.
\end{aligned}$$

Theorem

(Interlacing Theorem) *Let* $A \in \mathbb{C}^{n\times n}$ *be a Hermitian matrix and let* $B = A\begin{bmatrix} i_1 \cdots i_k \\ i_1 \cdots i_k \end{bmatrix}$ *be a principal submatrix of* A*,* $B \in \mathbb{C}^{k\times k}$*. If* $spec(A) = \{\lambda_1, \ldots, \lambda_n\}$ *and* $spec(B) = \{\mu_1, \ldots, \mu_k\}$*, where* $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ *and* $\mu_1 \geqslant \cdots \geqslant \mu_k$*, then* $\lambda_j \geqslant \mu_j \geq \lambda_{n-k+j}$ *for* $1 \leqslant j \leqslant k$*.*

Proof

Let $\{j_1, \ldots, j_q\} = \{1, \ldots, n\} - \{i_1, \ldots, i_k\}$, where $j_1 < \cdots < j_q$ and $k + q = n$. By Courant-Fisher Theorem we have

$$\lambda_j = \min_W \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A\mathbf{x} \mid \parallel \mathbf{x} \parallel_2 = 1 \text{ and } \mathbf{x} \in \langle W \rangle^{\perp}\},$$

where W ranges over sets of non-zero vectors in $\mathbb{C}^n$ containing $j-1$ vectors. Therefore,

$$\begin{aligned}\lambda_j &\geqslant \min_W \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A\mathbf{x} \mid \parallel \mathbf{x} \parallel_2 = 1 \text{ and } \mathbf{x} \in \langle W \rangle^{\perp} \\ &\quad \text{and } x \in \langle \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q} \rangle^{\perp}\} \\ &= \min_U \max_{\mathbf{y}} \{\mathbf{y}^{\mathsf{H}} B\mathbf{y} \mid \parallel \mathbf{y} \parallel_2 = 1 \text{ and } \mathbf{y} \in \langle U \rangle^{\perp} = \mu_j,\end{aligned}$$

where U ranges over sets of non-zero vectors in $\mathbb{C}^k$ containing $j-1$ vectors.

Proof cont'd

Again, by Courant-Fisher Theorem,

$$\lambda_{n-k+j} = \max_{Z} \min_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A \mathbf{x} \mid \parallel \mathbf{x} \parallel_2 = 1 \text{ and } \mathbf{x} \in \langle Z \rangle^{\perp}\},$$

where Z ranges over sets containing $k-j$ non-zero vectors in $\mathbb{C}^n$. Consequently,

$$\begin{aligned} \lambda_{n-k+j} &\leqslant \max_{Z} \min_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} A \mathbf{x} \mid \parallel \mathbf{x} \parallel_2 = 1 \text{ and } \mathbf{x} \in \langle Z \rangle^{\perp} \\ &\qquad \text{and } x \in \langle \mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_q} \rangle^{\perp}\} \\ &= \max_{S} \min_{\mathbf{y}} \{\mathbf{y}^{\mathsf{H}} B \mathbf{y} \mid \parallel \mathbf{y} \parallel_2 = 1 \text{ and } \mathbf{y} \in \langle S \rangle^{\perp}\} = \mu_j, \end{aligned}$$

where S ranges over the sets of non-zero vectors in $\mathbb{C}^k$ containing $n-j$ vectors.

Corollary

Let $A \in \mathbb{C}^{n\times n}$ be a Hermitian matrix and let $B = A\begin{bmatrix} i_1 & \cdots & i_k \\ i_1 & \cdots & i_k \end{bmatrix}$ be a principal submatrix of A, $B \in \mathbb{C}^{k\times k}$. The set spec$(B)$ contains no more positive eigenvalues than the number of positive eigenvalues of A and no more negative eigenvalues than the number of negative eigenvalues of A.

Theorem

Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix having the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. If $\mathbf{u}_1, \ldots, \mathbf{u}_n$ are eigenvectors that correspond to $\lambda_1, \ldots, \lambda_n$, respectively, $W = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ and $Z = \{\mathbf{u}_{k+2}, \ldots, \mathbf{u}_n\}$, then we have:

$$\begin{aligned} \lambda_{k+1} &= \max_{\mathbf{x}} \{\mathbf{x}^H A\mathbf{x} \mid \| \mathbf{x} \|_2 = 1 \text{ and } \mathbf{x} \in \langle W \rangle^{\perp}\} \\ &= \min_{\mathbf{x}} \{\mathbf{x}^H A\mathbf{x} \mid \| \mathbf{x} \|_2 = 1 \text{ and } \mathbf{x} \in \langle Z \rangle^{\perp}\}. \end{aligned}$$

UMASS BOSTON

Proof

If $A = U^{\mathsf{H}} D U$, where U is a unitary matrix and D is a diagonal matrix, then $\mathbf{u}_i$, the i^{th} column of U^{H} can be written as $\mathbf{u}_i = U^{\mathsf{H}} \mathbf{e}_i$. Therefore, by the second part of the proof of Courant-Fisher's theorem, we have $\mathbf{x} A \mathbf{x} \leqslant \lambda_{k+1}$ if $\mathbf{x}$ belongs to the subspace orthogonal to the subspace generated by the first k eigenvectors of A. Consequently, the Courant-Fisher Theorem implies the first equality of this theorem. The second equality can be obtained in a similar manner.

Corollary

Let $A \in \mathbb{C}^{n \times n}$ be a Hermitian matrix having the eigenvalues $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. If $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are eigenvectors that correspond to $\lambda_1, \ldots, \lambda_k$, respectively, then a unit vector $\mathbf{x}$ that maximizes $\mathbf{x}^H A\mathbf{x}$ and belongs to the subspace orthogonal to the subspace generated by the first k eigenvectors of A is an eigenvector that corresponds to λ_{k+1}.

Proof.

Let $\mathbf{u}_1, \ldots, \mathbf{u}_n$ be the eigenvectors of A and let $\mathbf{x} \in \langle \mathbf{u}_1, \ldots, \mathbf{u}_k \rangle^{\perp}$ be a unit vector. We have $\mathbf{x} = \sum_{j=k+1}^{n} a_j \mathbf{u}_j$, and $\sum_{j=k+1}^{n} a_j^2 = 1$ which implies

$$\mathbf{x}^{\mathrm{H}} A\mathbf{x} = \sum_{j=k+1}^{n} \lambda_j a_j^2 = \lambda_{k+1}.$$

This, in turn, implies $a_{k+1} = 1$ and $a_{k+2} = \cdots = a_n = 0$, so $\mathbf{x} = \mathbf{u}_{k+1}$. □

Theorem

Let $A, B \in \mathbb{C}^{n\times}$ be two Hermitian matrices and let $E = B - A$. Suppose that the eigenvalues of A, B, E these are $\alpha_1 \geqslant \cdots \geqslant \alpha_n$, $\beta_1 \geqslant \cdots \geqslant \beta_n$, and $\epsilon_1 \geqslant \cdots \geqslant \epsilon_n$, respectively. Then, we have $\epsilon_n \leqslant \beta_i - \alpha_i \leqslant \epsilon_1$.

Proof

Note that E is Hermitian, so all matrices involved have real eigenvalues. By Courant-Fisher Theorem,

$$\beta_k = \min_W \max_{\mathbf{x}} \{\mathbf{x}^{\mathsf{H}} B\mathbf{x} \mid \| \mathbf{x} \|_2 = 1 \text{ and } \mathbf{w}_i^{\mathsf{H}}\mathbf{x} = 0 \text{ for } 1 \leqslant i \leqslant k-1\},$$

where $W = \{\mathbf{w}_1, \ldots, \mathbf{w}_{k-1}\}$. Thus,

$$\beta_k \leqslant \max_{\mathbf{x}} \mathbf{x}^{\mathsf{H}} B\mathbf{x} = \max_{\mathbf{x}} (\mathbf{x}^{\mathsf{H}} A\mathbf{x} + \mathbf{x}^{\mathsf{H}} E\mathbf{x}). \tag{3}$$

Let U be a unitary matrix such that $U^{\mathsf{H}} A U = \text{diag}(\alpha_1, \ldots, \alpha_n)$. Choose $\mathbf{w}_i = U\mathbf{e}_i$ for $1 \leqslant i \leqslant k-1$. We have $\mathbf{w}_i^{\mathsf{H}}\mathbf{x} = \mathbf{e}_i^{\mathsf{H}} U^{\mathsf{H}}\mathbf{x} = 0$ for $1 \leqslant i \leqslant k-1$.

Proof cont'd

Define $\mathbf{y} = U^{\mathsf{H}}\mathbf{x}$. Since U is an unitary matrix, $\parallel \mathbf{y} \parallel_2 = \parallel \mathbf{x} \parallel_2 = 1$. Observe that $\mathbf{e}_i^{\mathsf{H}}\mathbf{y} = y_i = 0$ for $1 \leqslant i \leqslant k$. Therefore, $\sum_{i=k}^{n} y_i^2 = 1$. This, in turn implies $\mathbf{x}^{\mathsf{H}}A\mathbf{x} = \mathbf{y}^{\mathsf{H}}U^{\mathsf{H}}AU\mathbf{y} = \sum_{i=k}^{n} \alpha_i y_i^2 \leqslant \alpha_k$.
From the Inequality (3) it follows that

$$\beta_k \leqslant \alpha_k + \max_{\mathbf{x}} \mathbf{x}^{\mathsf{H}}E\mathbf{x} \leqslant \alpha_k + \epsilon_n.$$

Since $A = B - E$, by inverting the roles of A and B we have $\alpha_k \leqslant \beta_k - \epsilon_1$, or $\epsilon_1 \leqslant \beta_k - \alpha_k$, which completes the argument.