A GENERALIZED COURANT-FISCHER MINIMAX THEOREM

HAIM AVRON, ESMOND NG, AND SIVAN TOLEDO

1. Introduction

A useful tool for analyzing the spectrum of an Hermitian matrix is the *Courant-Fischer Minimax Theorem* [2].

Theorem 1.1. *(Courant-Fischer Minimax Theorem) Suppose that $S \in \mathbb{C}^{n\times n}$ is an Hermitian matrix, then*

$$\lambda_k(S) = \min_{\dim(U)=k} \max_{\substack{x \in U \\ x \neq 0}} \frac{x^*Sx}{x^*x}$$

and

$$\lambda_k(S) = \max_{\dim(V)=n-k+1} \min_{\substack{x \in V \\ x \neq 0}} \frac{x^*Sx}{x^*x}$$

where $\lambda_k(S)$ is the k'th largest eigenvalue of S.

The goal of this short communication it to present a generalization of Theorem 1.1, which we refer to as the *Generalized Courant-Fischer Minimax Theorem.* We now state the theorem, and we give a proof in the next section.

Theorem 1.2. *(Generalized Courant-Fischer Minimax Theorem) Suppose that $S \in \mathbb{C}^{n\times n}$ is an Hermitian matrix and that $T \in \mathbb{C}^{n\times n}$ is an Hermitian positive semidefinite matrix such that $\mathrm{null}(T) \subseteq \mathrm{null}(S)$. For $1 \leq k \leq \mathrm{rank}(T)$ we have*

$$\lambda_k(S,T) = \min_{\substack{\dim(U)=k \\ U \perp \mathrm{null}(T)}} \max_{x\in U} \frac{x^*Sx}{x^*Tx}$$

and

$$\lambda_k(S,T) = \max_{\substack{\dim(V)=rank(T)-k+1 \\ V \perp \mathrm{null}(T)}} \min_{x\in V} \frac{x^*Sx}{x^*Tx}\,.$$

2. Proof

We begin by stating and proving a generalization of the Courant-Fischer Theorem for pencils of Hermitian positive definite matrices.

Date: August 2008.

Theorem 2.1. *Let $S, T \in \mathbb{C}^{n\times n}$ be Hermitian matrices. If T is also positive definite then*

$$\lambda_k(S,T) = \min_{\dim(U)=k} \max_{x\in U} \frac{x^*Sx}{x^*Tx}$$

and

$$\lambda_k(S,T) = \max_{\dim(V)=n-k+1} \min_{x\in V} \frac{x^*Sx}{x^*Tx}\,.$$

Proof. Let $T = L^*L$ be the Cholesky factorization of B. Let U be some k-dimensional subspace of $\mathbb{C}^n$, let $x \in U$, and let $y = Lx$. Since T is Hermitian positive definite (hence nonsingular), the subspace $W = \{Lx : x \in U\}$ has dimension k. Similarly, for any k-dimensional subspace W, the subspace $U = \left\{L^{-1}x : x \in W\right\}$ has dimension k. We have

$$\frac{x^*Sx}{x^*Tx} = \frac{x^*L^*L^{-*}SL^{-1}Lx}{x^*L^*Lx} = \frac{y^*L^{-*}SL^{-1}y}{y^*y}\,.$$

By applying the Courant-Fischer to $L^{-*}SL^{-1}$, we obtain

$$\begin{aligned}\lambda_k(L^{-*}SL^{-1}) &= \min_{\dim(W)=k} \max_{y\in W} \frac{y^*L^{-*}SL^{-1}y}{y^*y}\\ &= \min_{\dim(U)=k} \max_{x\in S} \frac{x^*Sx}{x^*Tx}\,.\end{aligned}$$

The generalized eigenvalues of (S,T) are exactly the eigenvalues of $L^{-*}SL^{-1}$ so the first equality of the theorem follows. The second equality can be proved using a similar argument. ☐

Before proving the generalized version of the Courant-Fischer Minimax Theorem we show how to convert an Hermitian positive semidefinite problem to an Hermitian positive definite problem.

Lemma 2.2. *Let $S, T \in \mathbb{C}^{n\times n}$ be Hermitian matrices. Assume that T is also a positive semidefinite and that* $\mathrm{null}(T) \subseteq \mathrm{null}(S)$. *For any $Z \in \mathbb{C}^{n\times \mathrm{rank}(T)}$ with range(Z) = range(T), the determined generalized eigenvalues of (S,T) are exactly the generalized eigenvalues of (Z^*SZ, Z^*TZ).*

Proof. We first show that Z^*TZ has full rank. Suppose that $Z^*TZv = 0$. We have $TZv \in \mathrm{null}(Z^*)$. Therefore, $TZv \perp \mathrm{range}(Z) = \mathrm{range}(T)$. Obviously $TZv \in \mathrm{range}(T)$, so we must have $v = 0$. Since $\mathrm{null}(Z^*TZ) = \{0\}$, the matrix Z^*TZ has full rank.

Suppose that λ is a determined eigenvalue of (S,T). We will show that it is a determined eigenvalue of (Z^*SZ, Z^*TZ). The pencil (Z^*SZ, Z^*TZ) has exactly $\mathrm{rank}(Z^*TZ)$ determined eigenvalues. We will show that Z^*TZ is full rank, so the pencil (Z^*SZ, Z^*TZ) has exactly $\mathrm{rank}(T)$ eigenvalues. Since the pencil (S,T) has exactly $\mathrm{rank}(T)$ determined eigenvalues, each of them an eigenvalue of (Z^*SZ, Z^*TZ), this will conclude the proof.

Now let μ be an eigenvalue of (Z^*SZ, Z^*TZ). It must be determined, since Z^*TZ has full rank. Let y be the corresponding eigenvector, $Z^*SZy = \mu Z^*TZy$, and let $x = Zy$. Now there are two cases. If $\mu = 0$, then $SZy = Sx = 0$ (since Z^* has full rank and at least as many columns as rows). The vector x is in $\mathrm{range}(Z) = \mathrm{range}(T)$, $Tx \neq 0$. This implies that $\mu = 0$ is also a determined eigenvalue of (S,T).

If $\mu \neq 0$, the analysis is a bit more difficult. Clearly, $TZy \in \text{range}(T) = \text{range}(Z)$. But $\text{range}(Z) = \text{range}(Z^{*+})$ [1, Proposition 6.1.6.vii], so $Z^{*+}Z^*TZy = TZy$ [1, Proposition 6.1.7]. We claim that $SZy \in \text{range}(Z)$. If it is not, it Zy must be in $\text{null}(T) \subseteq \text{null}(S)$, but μ would have to be zero. Therefore, we also have $Z^{*+}Z^*SZy = SZy$, so by multiplying $Z^*SZy = \mu Z^*TZy$ by Z^{*+} we see that μ is an eigenvalue of (S, T). □

We are now ready to prove Theorem 1.2, the generalization of the Courant-Fischer Minimax Theorem. The technique is simple: we use Lemma 2.2 to reduce the problem to a smaller-sized full-rank problem, apply Theorem 2.1 to characterize the determined eigenvalues in terms of subspaces, and finally show a complete correspondence between the subspaces used in the reduced pencil and subspaces used in the original pencil.

Proof. (Theroem 1.2) Let $Z \in \mathbb{C}^{n \times \text{rank}(T)}$ have $\text{range}(Z) = \text{range}(T)$. We have

$$\lambda_k(S,T) = \lambda_k(Z^*SZ, Z^*TZ) = \min_{\dim(W)=k} \max_{x \in W} \frac{x^*Z^*SZx}{x^*Z^*TZx}$$

and

$$\lambda_k(S,T) = \lambda_k(Z^*SZ, Z^*TZ) = \max_{\dim(W)=\text{rank}(T)-k+1} \min_{x \in W} \frac{x^*Z^*TZx}{x^*Z^*TZx} .$$

The leftmost equality in each of these equations follows from Lemma 2.2 and the rightmost one follows from Theorem 2.1.

We now show that for every k-dimensional subspace $U \subseteq \mathbb{C}^n$ with $U \perp \text{null}(T)$, there exists a k-dimensional subspace $W \subseteq \mathbb{C}^{\text{rank}(T)}$ such that

$$\left\{ \frac{x^*Sx}{x^*Tx} : x \in U \right\} = \left\{ \frac{y^*Z^*SZy}{y^*Z^*TZy} : y \in W \right\} ,$$

and vice versa. The validity of this claim establishes the min-max side of the theorem.

We first need to show that $k \leq \text{rank}(T)$. This is true because every vector in U is in $\text{range}(T)$, so its dimension must be at most $\text{rank}(T)$.

Define $W = \left\{ y \in \mathbb{C}^{\text{rank}(T)} : Zy \in U \right\}$. Let $b_1, \ldots, b_k$ be a basis for U. Because $U \perp \text{null}(T)$, $b_j \in \text{range}(T)$, so there is a y_j such that $Zy_j = b_j$. Therefore, dimension of W is at most k. Now let the vectors y_i's be a basis of W and define $b_i = Zy_i$. The b_i's span U, so there are at most k of them, so the dimension of W is at least k. Therefore, it is exactly k.

Every $x \in U$ is orthogonal to $\text{null}(T)$, so it must be in $\text{range}(T)$. There exist a $y \in \mathbb{C}^{\text{rank}(T)}$ such that $Zy = x$. So we have $x^*Sx/x^*Tx = y^*Z^*SZy/y^*Z^*TZy$. Combining with the fact that $y \in W$, we have shown inclusion of one side. Now suppose $y \in W$. Define $x = Zy \in U$. Again we have $x^*Sx/x^*Tx = y^*Z^*SZy/y^*Z^*TZy$, which shows the other inclusion.

Now we will show that for every k-dimensional subspace W there is a subspace U that satisfies the claim. Define $U = \{Zy : y \in W\}$. Because Z has full rank, $\dim(U) = k$. Also, $U \subseteq \text{range}(Z) = \text{range}(T)$ so $U \perp \text{null}(T)$. The equality of the Raleigh-quotient sets follows from taking $y \in W$ and $x = Zy \in U$ or vice versa.

References

[1] Denis S. Bernstein. *Matrix Mathematics: Theory, Facts, and Formulas with Applications to Linear Systems Theory*. Princeton University Press, 2005.

[2] Gene H. Golub and Charles F. Van Loan. *Matrix Computations (3rd ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996.

$\square$

Current address: Haim Avron and Sivan Toledo: School of Computer Science, Tel-Aviv University, Esmond Ng: Lawrence Berkely National Laboratory