The definite generalized eigenvalue problem:

A reworked perturbation theory

Chi-Kwong Li
Department of Mathematics
College of William and Mary
Williamsburg, Virginia 23187-8795

Joint work with: Roy Mathias

Notation and definitions

Let A and B be $n \times n$ Hermitian matrices.

They form a definite pair if the Crawford number

$$c(A, B) = \min\{|x^*(A + iB)x| : x \in \mathbb{C}^n,\ x^*x = 1\} > 0.$$

If a real ordered pair (a, b) satisfies $a^2 + b^2 = 1$ and

$$bAx = aBx \text{ for some nonzero } x \in \mathbb{C}^n,$$

then (a, b) is a *normalized generalized eigenvalue* of (A, B).

Remark

If $B = I$, we have the usual eigenvalue problem.
If B is positive definite, we have $bB^{-1}Ax = ax$,
which is essentially the usual eigenvalue problem.

Proposition For a definite pair (A, B) there exists an invertible matrix X such that

$$X^*(A + iB)X = \operatorname{diag}(a_1 + ib_1, \ldots, a_n + ib_n),$$

where (a_j, b_j) are normalized generalized eigenvalues of (A, B).

Some nice properties [Li and Mathias, 1998]

Theorem [Uniqueness]

The normalized eigenvalues are unique up to ordering.

Theorem [max-min characterization]

If $C = A + iB$ and $X^*CX = \operatorname{diag}(e^{i\theta_1}, \dots, e^{i\theta_n})$ with $\pi > \theta_1 \geq \cdots \geq \theta_n > 0$ then for any $1 \leq j_1 < \cdots < j_k \leq n$,

$$\theta_{j_1} + \cdots + \theta_{j_k}$$

$$= \sup_{\substack{V_1 \subseteq \cdots \subseteq V_k \\ \dim V_t = j_t}} \inf_{\substack{y_t \in W_t \\ \det(y_r^* y_s) > 0}} \sum_{t=1}^k \arg(y_t^* C y_t).$$

Theorem [Interlacing inequalities]
Suppose $C = A + iB$ and $X^*CX = \operatorname{diag}(e^{i\theta_1}, \ldots, e^{i\theta_n})$ with $\pi > \theta_1 \geq \cdots \geq \theta_n > 0$. If Z is $n \times n - 1$ such that Z^*Z is invertible, and $Z^*CZ = \operatorname{diag}(e^{i\phi_1}, \ldots, e^{i\phi_{n-1}})$ with $\pi > \phi_1 \geq \cdots \geq \phi_{n-1} > 0$, then

$$\theta_1 \geq \phi_1 \geq \theta_2 \geq \cdots \theta_{n-1} \geq \phi_{n-1} \geq \theta_n.$$

Perturbation bounds

Let $r(E+iF) = \max\{|x^*(E+iF)x| : x \in \mathbb{C}^n,\ x^*x = 1\}$

be the *numerical radius* of $E+iF$.

Proposition Let (A,B) be a definite pair, and (E,F) be a Hermitian pair so that

$$\varepsilon = r(E+iF)/c(A,B) < 1.$$

Then $(\tilde{A},\tilde{B}) = (A,B) + (E,F)$ is a definite pair.

Theorem [Li and Mathias, 2004] Let (A,B) be a definite pair. Then

$$c(A,B) = \inf\{r(E+iF) : (A+E, B+F) \text{ is indefinite}\}.$$

Suppose $s_1(X) \geq \cdots \geq s_n(X)$ are the singular values of X. Denote by $s(X) = (s_1(X), \ldots, s_n(X))$.

For $x, y \in \mathbb{R}^{1\times n}$, $x \prec_w y$ if the sum of the k largest entries of x is not larger than that of y for $k = 1, \ldots, n$.

Theorem Let (A, B) be a definite pair, and (E, F) be a Hermitian pair so that $\varepsilon = r(E + iF)/c(A, B) < 1$.

The matrices $A + iB$ and $\tilde{A} + i\tilde{B}$ are $*$-congruent to

$$e^{\phi}\text{diag}\,(e^{i\theta_1}, \ldots, e^{i\theta_n}) \quad \text{and} \quad e^{\phi}\text{diag}\,(e^{i\tilde{\theta}_1}, \ldots, e^{i\tilde{\theta}_n})$$

such that $\pi \geq \theta_1 \geq \cdots \geq \theta_n > 0$, $\pi > \tilde{\theta}_1 \geq \cdots \geq \tilde{\theta}_n > 0$,

$(|\tilde{\theta}_1 - \theta_1|, \ldots, |\tilde{\theta}_n - \theta_n|) \prec_w \left(1 + \frac{1}{2\sqrt{1-\varepsilon}}\right) \frac{1}{c(A,B)} s([E|F])$.

A symmetric norm (symmetric gauge function) on $\mathbb{R}^{1\times n}$ is an absolute norm satisfying $\|x\| = \|xP\|$ for any permutation matrix P. Examples include ℓ_p norms with $p \in [1, \infty]$.

Let $x, y \in \mathbb{R}^{1\times n}$. Then $x \prec_w y$ if and only if $\|x\| \le \|y\|$ for all symmetric norms $\|\cdot\|$.

Corollary Continue to use the hypotheses and notation in the last theorem. For any symmetric norm $\|\cdot\|$ on $\mathbb{R}^n$,

$$\|(|\tilde{\theta}_1 - \theta_1|, \ldots, |\tilde{\theta}_n - \theta_n|)\| \le \left(1 + \tfrac{1}{2\sqrt{1-\varepsilon}}\right) \tfrac{1}{c(A,B)} \|s([E|F])\|.$$

A new approach

Proposition Suppose (A, B) is a definite pair. There is an invertible X with unit columns such that

$$X^*(A + iB)X = \text{diag}\,(a_1 + ib_1, \ldots, a_n + ib_n). \qquad (1)$$

Remarks

(1) Note that (a_j, b_j) are generalized eigenvalue pairs, not necessarily normalized.

(2) The pairs (a_j, b_j) are not determined uniquely. They will be if additional assumption is imposed on X.

For example, we may use any one of the following:

(i) X is chosen so that $(a_j, b_j) = (a_k, b_k)$ whenever

$$\arg(a_j + ib_j) = \arg(a_k + ib_k).$$

(ii) $\det(X^*X)$ is max (min) among all X satisfying (1).

A new approach

Consider $(\tilde{A}, \tilde{B}) = (A, B) + (E, F)$.

Suppose X has linearly independent unit columns and $X^*(A + iB)X = \operatorname{diag}(z_1, \dots, z_n)$ is diagonal, where $z_j = d_j e^{i\theta_j}$ for $j = 1, \dots, n$, such that $\pi > \theta_1 > \cdots > \theta_n > 0$.

We continue to assume $r = r(E + iF) < c(A, B)$.

Applying the mapping $T \mapsto X^*TX$ to each matrix, we may and we will assume that (A, B) are diagonal matrices.

Note that $\|E\|$ and $\|F\|$ will be increased by a factor of n. It still worths the price.

Theorem [Perturbation of normalized generalized eigenvalues] Set $u_j = \theta_j + \sin^{-1}(r/d_j)$ and $l_j = \theta_j - \sin^{-1}(r/d_j)$.

Rearrange the entries of $(u_1, \ldots, u_n)$ and $(l_1, \ldots, l_n)$ in descending order to get $(\tilde{u}_1, \ldots, \tilde{u}_n)$ and $(\tilde{l}_1, \ldots, \tilde{l}_n)$. Assume that $(\tilde{A}, \tilde{B})$ has normalized generalized eigenvalues such that $\pi > \tilde{\theta}_1 \geq \cdots \geq \tilde{\theta}_n > 0$. Then

$$\tilde{l}_j \leq \theta_j \leq \tilde{u}_j \quad \text{for } j = 1, \ldots, n.$$

Consequently,

$$|\theta_j - \tilde{\theta}_j| \leq \sin^{-1}(r/d_{\min}) \quad \text{for } j = 1, \ldots, n.$$

Other authors use $c(A,B)$ instead of $d_{\min}$

which may give much worse bounds.

For example, if $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $B = \begin{pmatrix} \varepsilon & 0 \\ 0 & \varepsilon \end{pmatrix}$,

then $d_{\min} = \sqrt{1+\varepsilon^2}$ can be much larger than $c(A,B) = \varepsilon$.

So, $\sin^{-1}(r/d_{\min})$ could be much smaller than $\sin^{-1}(r/c(A,B))$

Theorem [A quadratic bound]

Suppose $E = \begin{pmatrix} 0_m & R \\ R^* & 0_{n-m} \end{pmatrix}$ and $F = \begin{pmatrix} 0_m & S \\ S^* & 0_{n-m} \end{pmatrix}$.

For any $k \in \{1, \ldots, n\}$, let

$$d = \min\{d_j : j = m+1, \ldots, n\} \qquad \text{and}$$

$$\Delta_k = \min\{|d_j \sin(\theta_j - \tilde{\theta}_k)| : j = 1, \ldots, m\}.$$

Assume that $\Delta_k > 0$ and $r^2/\Delta_k < c(A, B)$. Then

$$|\theta_k - \tilde{\theta}_k| \leq \sin^{-1}\left(r^2/d\Delta_k\right).$$

Eigenvector perturbation

Theorem Suppose v is a generalized eigenvector corresponding to a normalized generalized eigenvalue $(\cos\tilde{\theta}, \sin\tilde{\theta})$ of $(\tilde{A}, \tilde{B})$. Let

$$\tilde{\Delta} = \min\{|d_j \sin(\tilde{\theta} - \theta_k)| : j \neq k\}.$$

If $r < \tilde{\Delta}$, then there is j such that

$$\|v - e_j\| \leq r/[\tilde{\Delta} + r].$$

We have the result for the non-diagonal case.

Theorem Suppose $A+iB$ is not in diagonal form, and X has linearly independent unit columns $X^*(A+iB)X$ is in diagonal form. Suppose v is a generalized eigenvector corresponding to a normalized generalized eigenvalue $(\cos\tilde{\theta}, \sin\tilde{\theta})$ of $(\tilde{A}, \tilde{B})$. Let

$$\tilde{\Delta} = \min\{|d_j \sin(\tilde{\theta} - \theta_k)| : j \neq k\}.$$

If $\|X\|^2 r < \tilde{\Delta}$, then there is j such that

$$\|v - X_j\| \leq \|X\|^2 r / [\tilde{\Delta} + \|X\|^2 r].$$

There are perturbation bounds on eigensubspace.

We omit their discussion because of its technicality.

Key message

Our approach can give much stronger bounds.

Question

Can one refine the results and get prettier theorems?

Thank you for your attention!