University of Maryland College Park

Institute for Advanced Computer Studies TR–2003–116
Department of Computer Science TR–4548

An Elsner-Like Perturbation Theorem for Generalized Eigenvalues*

G. W. Stewart†

December 2003

ABSTRACT

In 1985 Elsner established a general bound on the distance between an eigenvalue of a matrix and the closest eigenvalue of a perturbation of that matrix. In this note, we show that a similar result holds for the generalized eigenvalue problem.

*This report is available by anonymous ftp from `thales.cs.umd.edu` in the directory `pub/reports` or on the web at `http://www.cs.umd.edu/~stewart/`.

†Department of Computer Science and Institute for Advanced Computer Studies, University of Maryland, College Park, MD 20742 (`stewart@cs.umd.edu`). This work was supported in part by the National Science Foundation under grant CCR0204084.

An Elsner-Like Perturbation Theorem for Generalized Eigenvalues

G. W. Stewart

ABSTRACT

In 1985 Elsner established a general bound on the distance between an eigenvalue of a matrix and the closest eigenvalue of a perturbation of that matrix. In this note, we show that a similar result holds for the generalized eigenvalue problem.

Let A be a matrix of order n and let $\tilde{A} = A + E$ be a perturbation of A. Elsner's theorem [1] essentially states that if λ is an eigenvalue of A, then there is an eigenvalue $\tilde{\lambda}$ of $\tilde{A}$ satisfying

$$|\tilde{\lambda} - \lambda| \leq (\|A\| + \|\tilde{A}\|)^{1-\frac{1}{n}} \|E\|^{\frac{1}{n}},$$

where $\|\cdot\|$ denotes the spectral norm. The theorem is remarkable in several ways. It is general, requiring no hypotheses about A or λ. It puts no restrictions on the size of E. The bound is completely symmetric in A and $\tilde{A}$. The ingredients in the bound can be computed or bounded knowing only $\|A\|$ and $\|E\|$. Finally, it is sharp in the sense that the exponent $\frac{1}{n}$ of $\|E\|$ is the best possible. The price to be paid for this generality is that for small E and for most matrices the bound in unrealistically large — though by no means unuseful (see [2, 3]).

The purpose of this note is to prove an analogue of Elsner's theorem for the generalized eigenvalue problem $Ax = \lambda Bx$. Some of the nice features of Elsner's theorem will have to go by the board. We will lose some symmetry and we will have to assume that the pairs (A, B) and $(\tilde{A}, \tilde{B})$ are regular in the sense defined below.

We will begin by stating the generalized eigenvalue problem in projective form. We will then discuss the metrics we will use to measure distance between matrix pencils and their eigenvalues. Next we will introduce a condition that insures that a perturbation does not destroy regularity. The final preliminary is the introduction of the generalized Schur decomposition, after which we will state and prove our version of Elsner's theorem. More detail on this background material can be found in [4, 5].

We will call a pair (A, B) of $n \times n$ matrices a matrix pencil of order n. It is regular if

$$\det(\nu A - \mu B) \text{ is not identically zero.}$$

If $(\mu, \nu) \neq 0$ and $\det(\nu A - \mu B) = 0$ we will call the set $\langle \mu, \nu \rangle = \{(\tau\mu, \tau\nu) \colon \tau \in \mathbb{C}\}$ an eigenvalue of the pencil. The advantage of this projective representation is that $\langle 1, 0 \rangle$, which represents an infinite eigenvalue of the pencil is placed on an even footing with the other eigenvalues. Note for later reference that if $\langle\mu,\nu\rangle$ is an eigenvalue of (A,B), then there is a nonzero eigenvector x satisfying $(\nu A-\mu B)x=0$.

To measure the size of matrix (and scalar) pairs we will use the norm

$$\|(E,F)\|=\sqrt{\|E\|^2+\|F\|^2},$$

where on the right $\|\cdot\|$ denotes a the spectral norm or Frobenius norm. One reason for using this norm is the following inequality:

$$\max_{\|(\alpha,\beta)\|=1}\|\beta E+\alpha F\|\le\|(E,F)\|. \tag{1}$$

In fact,

$$\begin{aligned}\|\beta E+\alpha F\| &\le |\beta|\|E\|+|\alpha|\|F\|\\ &\le \sqrt{|\alpha|^2+|\beta|^2}\sqrt{\|E\|^2+\|F\|^2}\quad\text{(by the Cauchy inequality)}\\ &= \sqrt{\|E\|^2+\|F\|^2}.\end{aligned}$$

Turning now to eigenvalues, we will measure the distance between the eigenvalues $\langle\mu,\nu\rangle$ and $\langle\tilde\mu,\tilde\nu\rangle$ by the chordal metric

$$\chi(\langle\mu,\nu\rangle,\langle\tilde\mu,\tilde\nu\rangle)=\frac{|\mu\tilde\nu-\nu\tilde\mu|}{\|(\mu,\nu)\|\|(\tilde\mu,\tilde\nu)\|}.$$

The utility of this metric is that it, like the projective representation, makes no distinction between finite and infinite eigenvalue.

We now turn to the preservation of regularity under perturbations. Ideally we would like to determine the smallest perturbation that makes the pencil in question irregular. Unfortunately, this is an unsolved problem, and we must be content with a bound on perturbations that do not destroy regularity. One such bound is the number

$$\gamma(A,B)=\max_{\|(\alpha,\beta)\|=1}\sigma_{\min}(\beta A-\alpha B), \tag{2}$$

where $\sigma_{\min}(X)$ denotes the smallest singular value of X. To see this, note that if (A,B) is regular $\gamma(A,B)>0$. Now suppose $\|(E,F)\|<\gamma$ and let α and β maximize the right-hand side of (2). Then by (1), $\|\beta E-\alpha F\|<\sigma_{\min}(\beta A-\alpha B)$, and hence by the Schmidt–Eckart–Young–Mirsky theorem $\det[\beta(A+E)-\alpha(B+F)]\ne 0$, so that the pencil $(A+E,B+F)$ is regular.

An important advantage of this measure is that it is insensitive to perturbations in its arguments. Specifically, it is easy to show that

$$\gamma(A+E,B+F)\ge\gamma(A,B)-\|(E,F)\|. \tag{3}$$

We now introduce the generalized Schur decomposition. Specifically, if (A, B) is regular, there are unitary matrices U and V such that

$$U^{\mathrm{H}}AV = S \quad \text{and} \quad U^{\mathrm{H}}BV = T$$

where S and T are upper triangular. The quatities $\langle \sigma_{ii}, \tau_{ii} \rangle$ are the eigenvalue of (A, B), which can be made to appear anywhere on the diagonals of S and T. An important consequence of this form is that

$$\gamma_i \equiv \|(\sigma_{ii}, \tau_{ii})\| \geq \gamma(S, T) = \gamma(A, B). \tag{4}$$

For if not, we could set $\sigma_{ii} = \tau_{ii} = 0$ and render the pencil (A, B) irregular by a perturbation whose norm is less than $\gamma(A, B)$ — a contradiction.

We are now in a position to state and prove our variant of Elsner's theorem.

Theorem. Let (A, B) and $(\tilde{A}, \tilde{B}) = (A + E, B + F)$ be regular matrix pairs, and let $\langle \mu, \gamma \rangle$ be an eigenvalue of (A, B), then there is an eigenvalue $\langle \tilde{\mu}, \tilde{\nu} \rangle$ of $(\tilde{A}, \tilde{B})$ satisfying

$$\chi(\langle \mu, \nu \rangle, \langle \tilde{\mu}, \tilde{\nu} \rangle) \leq \frac{\|A\|^{1-\frac{1}{n}} \|(E, F)\|^{\frac{1}{n}}}{\gamma(A, B)}. \tag{5}$$

Proof. We may assume without loss of generality that (A, B) is in generalized Schur form with $\langle \alpha_{11}, \beta_{11} \rangle = \langle \mu, \nu \rangle$. Let $\langle \tilde{\mu}, \tilde{\nu} \rangle$ be the eigenvalue of $(\tilde{A}, \tilde{B})$ that is closest to $\langle \mu, \nu \rangle$ in the chordal metric, and assume that $\|(\tilde{\mu}, \tilde{\nu})\| = 1$. Then

$$\begin{aligned}
\frac{|\det(\tilde{\nu}A - \tilde{\mu}B)|}{\gamma(A, B)^n} &= \frac{|\tilde{\nu}\alpha_{11} - \tilde{\mu}\beta_{11}|}{\gamma(A, B)} \times \cdots \times \frac{|\tilde{\nu}\alpha_{nn} - \tilde{\mu}\beta_{nn}|}{\gamma(A, B)} \\
&\geq \frac{|\tilde{\nu}\alpha_{11} - \tilde{\mu}\beta_{11}|}{\gamma_1} \times \cdots \times \frac{|\tilde{\nu}\alpha_{nn} - \tilde{\mu}\beta_{nn}|}{\gamma_n} \\
&= \chi(\langle \alpha_{11}, \beta_{11} \rangle, \langle \tilde{\mu}, \tilde{\nu} \rangle) \times \cdots \times \chi(\langle \alpha_{nn}, \beta_{nn} \rangle, \langle \tilde{\mu}, \tilde{\nu} \rangle) \\
&\geq \chi(\langle \alpha_{11}, \beta_{11} \rangle, \langle \tilde{\mu}, \tilde{\nu} \rangle)^n.
\end{aligned}$$

Now let $X = (x_1 \ \cdots \ x_n)$ be a unitary matrix with $(\tilde{\nu}\tilde{A} - \tilde{\mu}\tilde{B})x_1 = 0$. Then by Hadamard's inequality

$$|\det(\tilde{\nu}A - \tilde{\mu}B)| \leq \|(\tilde{\nu}A - \tilde{\mu}B)x_1\| \times \cdots \times \|(\tilde{\nu}A - \tilde{\mu}B)x_n\|$$

But

$$\|(\tilde{\nu}A - \tilde{\mu}B)x_1\| = \|[\tilde{\nu}(A - \tilde{A}) - \tilde{\mu}(B - \tilde{B})x_1]\| = \|(\tilde{\nu}E - \tilde{\mu}F)x_1\| \leq \|(E, F)\|.$$

On the other hand for $i \neq 1$

$$\|(\tilde{\nu}A - \tilde{\mu}B)x_i\| \leq \|(A, B)\|.$$

Consequently we have

$$\chi(\langle\mu,\nu\rangle,\langle\tilde{\mu},\tilde{\nu}\rangle)^n \leq \frac{\|(A,B)\|^{n-1}\|(E,F)\|}{\gamma(A,B)^n},$$

and (5) follows on taking nth roots. ■

The proof is along the lines of Elsner's. As mentioned above we have to restrict (A,B) and $(\tilde{A},\tilde{B})$ to be regular, but there are no restrictions on (E,F). Of course, if $\|(E,F)\| < \gamma(A,B)$, then (3) implies that $(\tilde{A},\tilde{B})$ is regular.

The chief difference between the two theorems is the appearance of $\gamma(A,B)$ in (5). Divisors of this kind are common in generalized eigenvalue bounds, and they reflect the fact that eigenvalues with small γ_i [see (4)] are extremely sensitive to small perturbations in the pencil. In fact, in the theorem we could replace $\gamma(A,B)$ with $\min_i \gamma_i$ — giving a potentially sharper bound. Unfortunately, the γ_i associated with a particular eigenvalue can change when the eigenvalue is moved to another place in the generalized Schur decomposition, so that the resulting bound would depend on the vagaries of how the decomposition was computed.

References

[1] L. Elsner. An optimal bound for the spectral variation of two matrices. *Linear Algebra and Its Applications*, 71:77–80, 1985.

[2] Z. Jia and G. W. Stewart. On the convergence of Ritz values, Ritz vectors, and refined Ritz vectors. Technical Report TR–3986, Department of Computer Science, University of Maryland, College Park, 1999.

[3] Z. Jia and G. W. Stewart. An analysis of the Rayleigh–Ritz method for approximating eigenspaces. *Mathematics of Computation*, 70:637–647, 2001.

[4] G. W. Stewart. *Matrix Algorithms II: Eigensystems.* SIAM, Philadelphia, 2001.

[5] G. W. Stewart and J.-G. Sun. *Matrix Perturbation Theory.* Academic Press, New York, 1990.