Contents lists available at SciVerse ScienceDirect

Linear Algebra and its Applications

Perturbation of multiple eigenvalues of Hermitian matrices

Ren-Cang Li [a,1], Yuji Nakatsukasa [b,2], Ninoslav Truhar [c,*,3], Wei-guo Wang [d,4]

[a] *Department of Mathematics, University of Texas at Arlington, P.O. Box 19408, Arlington, TX 76019-0408, USA*
[b] *School of Mathematics, The University of Manchester, Manchester M13 9PL, UK*
[c] *Department of Mathematics, J.J. Strossmayer University of Osijek, Trg Lj. Gaja 6, 31000 Osijek, Croatia*
[d] *School of Mathematical Sciences, Ocean University of China, 238 Songling Road, Qingdao 266100, PR China*

ARTICLE INFO

ABSTRACT

This paper is concerned with the perturbation of a multiple eigenvalue μ of the Hermitian matrix $A = \text{diag}(\mu I, A_{22})$ when it undergoes an off-diagonal perturbation E whose columns have widely varying magnitudes. When some of E's columns are much smaller than the others, some copies of μ are much less sensitive than any existing bound suggests. We explain this phenomenon by establishing individual perturbation bounds for different copies of μ. They show that when $A_{22} - \mu I$ is definite the ith bound scales quadratically with the norm of the ith column, and in the indefinite case the bound is necessarily proportional to the product of E's ith column norm and E's norm. An extension to the generalized Hermitian eigenvalue problem is also presented.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Consider the eigenvalue problem for Hermitian matrix $\widetilde{A}$:

$$\widetilde{A} = \begin{matrix} m \\ n \end{matrix} \begin{pmatrix} \overset{m}{A_{11}} & \overset{n}{E^*} \\ E & A_{22} \end{pmatrix}, \quad A_{11} = \mu I_m, \tag{1.1}$$

where the superscript "·*" takes the complex conjugate transpose of a matrix or a vector, and I_m (or simply I later if its dimension is clear from the context) is the $m \times m$ identity matrix. If E is a zero block, then μ is a multiple eigenvalue with multiplicity m. In general, if E is small then $\widetilde{A}$ has m eigenvalues close to μ. In fact more can be said qualitatively. Let η be the eigenvalue gap between $A_{11} = \mu I$ and A_{22} defined as

$$\eta = \min_{\nu \in \mathrm{eig}(A_{22})} |\mu - \nu|, \tag{1.2}$$

where $\mathrm{eig}(A_{22})$ is the set of the eigenvalues of A_{22}, and let

$$\varepsilon = \|E\|_2, \tag{1.3}$$

where $\| \cdot \|_2$ is either the spectral norm of a matrix or the ℓ_2-norm of a vector. The main result in [1] says $\widetilde{A}$ has m eigenvalues $\theta_1, \ldots, \theta_m$ such that

$$|\mu - \theta_j| \leq \frac{2\varepsilon^2}{\eta + \sqrt{\eta^2 + 4\varepsilon^2}} \quad \text{for } 1 \leq j \leq m. \tag{1.4}$$

The right-hand side of (1.4) is of second order in ε if $\eta > 0$ and is never larger than ε. As confirmed by the 2-by-2 example in [1], in general these inequalities cannot be improved without knowing more information on E than just $\varepsilon = \|E\|_2$.

Suppose now that we do have additional information on E. For example, consider the case where one of the columns of E is zero for which $\theta_i = \mu$ for some i. Can we derive bounds that reflect this – a zero column leads to some θ_i being μ? A possible and quick answer can be given as follows: first zero out the jth column of E, and then use the well-known perturbation theorem (attributed to Lidskii, Weyl, Wiedlandt and Mirsky in various forms [2, pp. 196–205]) to conclude that $\widetilde{A}$ has an eigenvalue θ that differs from μ by no more than $\|E_{(:,j)}\|_2$, where $E_{(:,j)}$ denotes the jth column of E. It obviously implies that if E's jth column is a zero column, then μ must be an eigenvalue of $\widetilde{A}$. But there are two drawbacks to this quick answer:

1. $\|E_{(:,j)}\|_2$ can be potentially (much) larger than the right-hand side of (1.4), making the estimate less favorable to (1.4).
2. This does not imply that $\widetilde{A}$ has m eigenvalues θ_j such that $|\mu - \theta_j| \leq \|E_{(:,j)}\|_2$ because some of the θ by this argument could be the same eigenvalues of $\widetilde{A}$, as mentioned in [3, Section 11.5].

The purpose of this article is to develop a theory that will reflect the effect of disparity in the magnitudes of the columns of E on the eigenvalues of $\widetilde{A}$, unlike (1.4), through establishing different bounds for the m eigenvalues of $\widetilde{A}$ closest to μ.

For the sake of convenience, throughout this paper η and ε are always defined by (1.2) and (1.3), respectively, and set

$$\epsilon_j = \|E_{(:,i_j)}\|_2 \quad \text{for } 1 \leq j \leq m, \tag{1.5}$$

where $\{i_1, i_2, \ldots, i_m\}$ is the permutation of $\{1, 2, \ldots, m\}$ such that

$$\epsilon_1 \leq \epsilon_2 \leq \cdots \leq \epsilon_m. \tag{1.6}$$

It is well-known that $\epsilon_m \leq \varepsilon \leq \sqrt{m}\,\epsilon_m$. The eigenvalues of E^*E are $\tau_1, \tau_2, \ldots, \tau_m$, arranged in ascending order:

$$0 \leq \tau_1 \leq \tau_2 \leq \cdots \leq \tau_m. \tag{1.7}$$

We will also use $X \prec Y$ ($X \preceq Y$) for two Hermitian matrices of the same size to mean $Y - X$ is positive definite (semi-definite), and $X \succ Y$ ($X \succeq Y$) to mean $Y \prec X$ ($Y \preceq X$). In particular, $X \succ 0$ ($X \succeq 0$) means that X is positive definite (semi-definite).

Our perturbation bounds are actually presented in terms of τ_j, the eigenvalues of E^*E. They can be easily turned into bounds in terms of ϵ_j, because of Lemma 3.1 below, in order to serve our purpose of developing a theory that reflects the effect of disparity in the magnitudes of the columns of E.

The rest of this paper is organized as follows. We first investigate specific examples in Section 2, which provide insights into possible bounds that could be expected. In Section 3 we give our main results, in which we separately deal with the cases where $A_{22} - \mu I$ is definite or indefinite. For the indefinite case, we give asymptotic estimates that are correct up to fourth-order terms, as well as strict bounds that are slightly larger than the asymptotic estimates. In Section 4 we describe how our bounds can be extended to the generalized eigenvalue problem. Finally we summarize our conclusions in Section 5.

2. Motivational examples

The examples below will shape our expectation on possible effects of different magnitudes of the columns of E on the eigenvalues of $\widetilde{A}$ nearest 0.

Example 2.1. Consider the 4-by-4 matrix $\widetilde{A}$ given by

$$E = \begin{pmatrix} 3 \cdot 10^{-4} & -2 \cdot 10^{-2} \\ 2 \cdot 10^{-4} & 10^{-2} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & E^* \\ E & A_{22} \end{pmatrix}.$$

In this case $A_{11} = 0$, i.e., $\mu = 0$ in (1.1), and $\eta = 1$. The two eigenvalues of $\widetilde{A}$ closest to 0 are approximately

$$1.632172864323117 \cdot 10^{-7} \quad \text{and} \quad -3.000632552828267 \cdot 10^{-4}, \tag{2.1}$$

which are about $\epsilon_1^2 = \|E_{(:,1)}\|_2^2 = 1.3 \cdot 10^{-7}$ and $\epsilon_2^2 = \|E_{(:,2)}\|_2^2 = 5.0 \cdot 10^{-4}$, respectively.

The inequality (1.4) says $\widetilde{A}$ has two eigenvalues that differ from 0 by no more than $4.9978 \cdot 10^{-4}$. This estimate is very good for the second eigenvalue in (2.1) but not so for the first one which is about less than the square of the estimate. The quick answer, on the other hand, says $\widetilde{A}$ has an eigenvalue that differs from 0 by no more than $\epsilon_1 = 3.6056 \cdot 10^{-4}$ and an eigenvalue from 0 by no more than $\epsilon_2 = 2.2361 \cdot 10^{-2}$, providing even worse estimates than by (1.4). □

Example 2.1 may lead us to believe that there are m properly ordered eigenvalues $\theta_1, \ldots, \theta_m$ of $\widetilde{A}$ with each difference $|\mu - \theta_j|$ being of second order in $\epsilon_j = \|E_{(:,j)}\|_2$ if $\eta > 0$. Later we will show this is indeed true if $A_{22} - \mu I$ is definite, but not so in the general case as we can see by the next example.

Example 2.2. Consider the 4-by-4 matrix $\widetilde{A}$ given by

$$E = \begin{pmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & E^* \\ E & A_{22} \end{pmatrix},$$

where both δ_i are real numbers and $|\delta_i| \le 1$. The characteristic equation of $\widetilde{A}$ is

$$\lambda^4 - (\delta_1^2 + \delta_2^2 + 1)\lambda^2 + \delta_1^2\delta_2^2 = 0,$$

whose two smallest eigenvalues in magnitude satisfy

$$|\lambda| = \sqrt{\frac{2}{\delta_1^2 + \delta_2^2 + 1 + \sqrt{[1 + (\delta_1 + \delta_2)^2][1 + (\delta_1 - \delta_2)^2]}}} \, |\delta_1\delta_2|.$$

Thus $|\lambda|/|\delta_1\delta_2| = 1 + O(\delta_1^2 + \delta_2^2)$. It follows that the smaller $|\lambda|$ can be made arbitrarily larger than $O(\min\{\delta_1^2, \delta_2^2\})$. □

3. Main results

Throughout this section, $\widetilde{A}$ is Hermitian and given by (1.1). Without loss of generality, we assume

$$\mu = 0.$$

Since by assumption μ is not an eigenvalue of A_{22}, A_{22} is non-singular as a result of assuming $\mu = 0$, and the gap η as defined by (1.2) now is

$$\eta = 1/\|A_{22}^{-1}\|_2.$$

For any $\lambda \notin \text{eig}(A_{22})$, set

$$X = \begin{pmatrix} I & -E^*(A_{22} - \lambda I)^{-1} \\ 0 & I \end{pmatrix}.$$

Then

$$X(\widetilde{A} - \lambda I)X^* = \begin{pmatrix} (-\lambda)I - E^*(A_{22} - \lambda I)^{-1}E & \\ & A_{22} - \lambda I \end{pmatrix}, \tag{3.1}$$

and thus

$$\det(\widetilde{A} - \lambda I) = \det\left(-E^*(A_{22} - \lambda I)^{-1}E - \lambda I\right) \times \det(A_{22} - \lambda I). \tag{3.2}$$

From this we see that any eigenvalue $\widetilde{\lambda}$ of $\widetilde{A}$ not in $\text{eig}(A_{22})$ is a root of

$$\det(-E^*(A_{22} - \lambda I)^{-1}E + (-\lambda)I). \tag{3.3}$$

Recall from (1.1) that for E sufficiently small in magnitude, the eigenvalues of $\widetilde{A}$ consist of two subsets: one spawned from m copies of μ and another from the eigenvalues in $\text{eig}(A_{22})$ upon being moved by E. Hence $\widetilde{A}$ has m eigenvalues close to 0 and these m eigenvalues are zeros of (3.3) near 0. Note that for $|\lambda|\|A_{22}^{-1}\|_2 = |\lambda|/\eta < 1$ we can write $(A_{22} - \lambda I)^{-1} = \sum_{j=0}^{\infty} \lambda^j A_{22}^{-j-1}$, so for such λ we have

$$-E^*(A_{22} - \lambda I)^{-1}E + (-\lambda)I = -\sum_{j=0}^{\infty} \lambda^j E^* A_{22}^{-j-1} E + (-\lambda)I. \tag{3.4}$$

Theorem 3.1. *Let $\widetilde{A}$ be a Hermitian matrix of form (1.1) with $\mu = 0$.*

1. *Assume $\varepsilon < \sqrt{3/4}\,\eta$. Then*
   (a) *$\widetilde{A}$ has exactly m eigenvalues θ_j in the open interval $(-\eta/2, \eta/2)$, and moreover*

$$|\theta_j| \leq \frac{2\varepsilon^2}{\eta + \sqrt{\eta^2 + 4\varepsilon^2}}, \tag{3.5}$$

   for $1 \leq j \leq m$;
   (b) *The function (3.3) has exactly m zeros in $(-\eta/2, \eta/2)$ and these zeros are precisely the eigenvalues θ_j of $\widetilde{A}$.*
2. *$\widetilde{A}$ has m eigenvalues $\theta_j = \vartheta_j + O(\varepsilon^4/\eta^2)$, where ϑ_j for $1 \leqslant j \leqslant m$ are the eigenvalues of $-E^* A_{22}^{-1} E$. In particular, if $\eta = O(1)$, then $\theta_j = \vartheta_j + O(\varepsilon^4)$.*

Proof. Since $4t^2/(1 + \sqrt{1 + 4t^2}) < 1$ if $t^2 < 3/4$, we have

$$\frac{2\varepsilon^2}{\eta + \sqrt{\eta^2 + 4\varepsilon^2}} < \frac{\eta}{2} \quad \text{if} \quad \frac{\varepsilon}{\eta} < \sqrt{\frac{3}{4}}.$$

By the main result of [1], we conclude that $\widetilde{A}$ has exactly m eigenvalues θ_j in the open interval $(-\eta/2, \eta/2)$ and (3.5) holds.

Item 1(b) is a consequence of Item 1(a), (3.2) and $\det(A_{22} - \lambda I) \neq 0$ for $\lambda \in (-\eta/2, \eta/2)$.

The expression in (3.4) is equal to $-E^* A_{22}^{-1} E + (-\lambda)I$, up to $O(\varepsilon^4/\eta^2)$, for $|\lambda| = O(\varepsilon^2/\eta)$. Since by (1.4) $\widetilde{A}$ has exactly m eigenvalues no larger than $O(\varepsilon^2/\eta)$ in magnitude, we conclude that $\theta_j = \vartheta_j + O(\varepsilon^4/\eta^2)$ for $1 \leqslant j \leqslant m$. □

Example 2.1 (revisit). The eigenvalues of $-E^* A_{22}^{-1} E$ are

$$1.632173307879875 \cdot 10^{-7}, \quad -3.002132173307880 \cdot 10^{-4}$$

which are extremely close to the exact values given in (2.1). □

Theorem 3.1 gives asymptotic estimates for θ_j in terms of ϑ_j. In the subsections that follow, we will establish bounds that reflect the effect of disparity in the magnitudes of the columns of E. To this end, we normalize the columns of E by their ℓ_2-norms to get

$$E = E_0 D, \tag{3.6}$$

where

$$D = \mathrm{diag}(\|E_{(:,1)}\|_2, \|E_{(:,2)}\|_2, \ldots, \|E_{(:,m)}\|_2), \tag{3.7a}$$

$$(E_0)_{(:,j)} = \begin{cases} E_{(:,j)}/\|E_{(:,j)}\|_2, & \text{if } E_{(:,j)} \neq 0, \\ 0, & \text{if } E_{(:,j)} = 0. \end{cases} \tag{3.7b}$$

Lemma 3.1. Let $\tau_1, \tau_2, \ldots, \tau_m$ be the eigenvalues of $E^* E$, arranged in ascending order as in (1.7), and let ϵ_j be defined as in (1.5) and (1.6). Then

$$\tau_j \leq \|E_0\|_2^2 \, \epsilon_j^2 \leq m \, \epsilon_j^2. \tag{3.8}$$

Proof. Use $0 \preceq E^* E = D E_0^* E_0 D \preceq \|E_0\|_2^2 D^2$ to get

$$\tau_j \leq \|E_0\|_2^2 \, D_{(i_j, i_j)}^2 = \|E_0\|_2^2 \, \epsilon_j^2.$$

The second inequality is due to $\|E_0\|_2 \leq \sqrt{m}$. □

Next, we separately consider the cases according to whether A_{22} is definite or not. All bounds will be given in terms of τ_j. Corresponding bounds in terms of ϵ_j can then be easily derived by using (3.8).

3.1. Positive (negative) definite A_{22}

Theorem 3.2. For Hermitian matrix $\widetilde{A}$ as in (1.1) with $\mu = 0$, suppose $\varepsilon < \sqrt{3/4}\,\eta$. If A_{22} is positive (negative) definite, then $\widetilde{A}$ has m nonpositive (nonnegative) eigenvalues $\theta_1, \ldots, \theta_m$ arranged in ascending order satisfying

$$0 \leq -\theta_{m-j+1} \leq \frac{2\tau_j}{\eta + \sqrt{\eta^2 + 4\tau_j}}, \quad \text{if } A_{22} \succ 0, \tag{3.9a}$$

$$0 \leq \theta_j \leq \frac{2\tau_j}{\eta + \sqrt{\eta^2 + 4\tau_j}}, \quad \text{if } A_{22} \prec 0, \tag{3.9b}$$

for $1 \leq j \leq m$.

Proof. The case in which $A_{22} \prec 0$ can be turned into the case in which $A_{22} \succ 0$ by considering $-\widetilde{A}$ instead. Suppose that $A_{22} \succ 0$, i.e., A_{22} is positive definite. Set

$$B(t) = -E^*(A_{22} - tI)^{-1}E \tag{3.10}$$

for $t \in \mathbb{R}$. By Theorem 3.1 and the assumption $\varepsilon < \sqrt{3/4}\,\eta$, we know $\widetilde{A}$ has exactly m eigenvalues in $(-\eta/2, \eta/2)$ and these m eigenvalues are the zeros of $\det(B(t) - tI)$ in $(-\eta/2, \eta/2)$. Since for any $t \in (-\eta/2, \eta/2)$, $0 \prec A_{22} - tI$ and thus $B(t) \preceq 0$; so

$$B(t) - tI \prec 0 \quad \text{for } t \in (0, \eta/2).$$

Therefore the m eigenvalues of $\widetilde{A}$ are in $(-\eta/2, 0]$. Denote them by

$$-\eta/2 < \theta_1 \le \theta_2 \le \cdots \le \theta_m \le 0.$$

Also denote by

$$\lambda_1(t) \le \lambda_2(t) \le \cdots \le \lambda_m(t) \le 0 \tag{3.11}$$

the m eigenvalues of $B(t)$ for $t \in (-\eta/2, 0]$. They are continuous. The fixed points of $\lambda_i(t)$ within $t \in (-\eta/2, 0]$ give all the θ_j. In fact, we have $\lambda_j(\theta_j) = \theta_j$. This is because $\lambda_j(t)$ is a decreasing function for $t \in (-\eta/2, 0]$ and thus $\lambda_j(t) = t$ has a unique solution on $(-\eta/2, 0]$. Hence θ_j is the jth smallest eigenvalue of $B(\theta_j)$. This implies that $|\theta_j| = -\theta_j$ is the jth largest eigenvalue of $-B(\theta_j)$. Since

$$-B(\theta_j) = E^*(A_{22} - \theta_j I)^{-1}E \preceq \frac{E^*E}{\eta + |\theta_j|},$$

we have

$$|\theta_j| \le \frac{\tau_{m-j+1}}{\eta + |\theta_j|} \quad \text{implying} \quad |\theta_j| \le \frac{2\tau_{m-j+1}}{\eta + \sqrt{\eta^2 + 4\tau_{m-j+1}}}$$

which gives (3.9a). $\quad\square$

Remark 3.1. Since the right-hand sides in (3.9) are increasing as τ_j does, replacing τ_j by its upper bound in (3.8) yields bounds on $|\theta_j|$ in terms of ϵ_j, the norms of E's columns.

Example 3.1. Consider the 4-by-4 matrix $\widetilde{A}$ given by

$$E = \begin{pmatrix} 3 \cdot 10^{-4} & -2 \cdot 10^{-2} \\ 2 \cdot 10^{-4} & 10^{-2} \end{pmatrix}, \quad A_{22} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \widetilde{A} = \begin{pmatrix} 0 & E^* \\ E & A_{22} \end{pmatrix}. \tag{3.12}$$
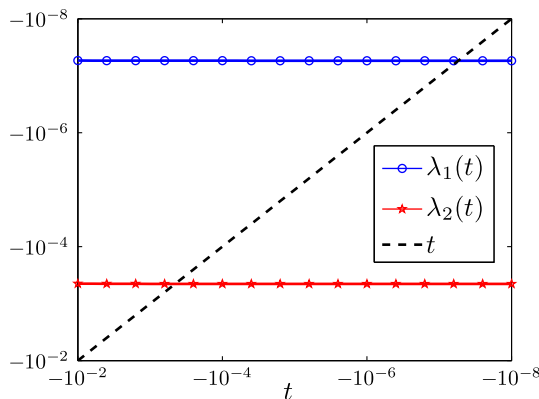

Fig. 1. The log–log scale plot of $\lambda_i(t)$ for $\widetilde{A}$ in (3.12).

In this case $A_{11} = 0$, i.e., $\mu = 0$ in (1.1), and $\eta = 1$. The following table displays the eigenvalues θ_j of $\widetilde{A}$ nearest to 0, the eigenvalues ϑ_j of $-E^* A_{22}^{-1} E$, and the upper bounds in (3.9) and the ones after τ_j replaced by $m\epsilon_j^2$.

θ_j	ϑ_j	$\dfrac{2\tau_j}{\eta+\sqrt{\eta^2+4\tau_j}}$	$\dfrac{2m\epsilon_j^2}{\eta+\sqrt{\eta^2+4m\epsilon_j^2}}$
$-4.4986 \cdot 10^{-4}$	$-4.5006 \cdot 10^{-4}$	$4.9978 \cdot 10^{-4}$	$9.9900 \cdot 10^{-4}$
$-5.4438 \cdot 10^{-8}$	$-5.4438 \cdot 10^{-8}$	$9.7994 \cdot 10^{-8}$	$2.6000 \cdot 10^{-7}$

Thus our bounds are remarkably sharp. Let $\lambda_i(t)$ be as in the proof of Theorem 3.2 for this example. Fig. 1 plots $\lambda_1(t)$ and $\lambda_2(t)$ as functions of t. The intersections with the curve for t are the eigenvalues θ_1 and θ_2. Note that in Fig. 1 $\lambda_1(t)$ and $\lambda_2(t)$ appear to be nearly constants. That is because they decrease very slowly, which is a typical behavior of $\lambda_i(t)$ when $\varepsilon \ll \eta/2$. In fact it can be shown that $-\frac{\varepsilon^2}{\eta^2} \leqslant \frac{d\lambda_i(t)}{dt} \leqslant 0$ for $t \in (-\eta/2, 0]$ and $1 \leqslant i \leqslant m$. $\square$

3.2. Indefinite A_{22}

Consider now that $A_{22} - \mu I$ is indefinite. We will use the following result, which is a direct consequence of the proof of [4, Theorem 1].

Lemma 3.2. Let W be an ℓ-by-ℓ Hermitian matrix, and let $D = \mathrm{diag}(\delta_1, \delta_2, \ldots, \delta_\ell)$ with $|\delta_1| \leq |\delta_2| \leq \cdots \leq |\delta_\ell|$. Denote the eigenvalues of $D^* W D$ by $\omega_1, \ldots, \omega_\ell$ arranged such that $|\omega_1| \leq |\omega_2| \leq \cdots \leq |\omega_\ell|$. Then for $1 \leqslant i \leqslant \ell$

$$|\omega_i| \leqslant \min_{1 \leqslant j \leqslant \ell - i + 1} |\delta_{\ell-j+1} \delta_{i+j-1}| \, \|W\|_2$$
$$\leqslant |\delta_\ell \delta_i| \, \|W\|_2.$$

Two types of bounds will be proven for the eigenvalues θ_j of interest of $\widetilde{A}$: asymptotical bounds up to $O(\varepsilon^4)$, and strict bounds at a tradeoff of being slightly larger than the asymptotic bounds if higher order terms $O(\varepsilon^4)$ are ignored.

Lemma 3.3. Let ϑ_j for $1 \leqslant j \leqslant m$ be the eigenvalues of $-E^* A_{22}^{-1} E$ arranged such that

$$|\vartheta_1| \leq |\vartheta_2| \leq \cdots \leq |\vartheta_m|.$$

Then

$$|\vartheta_j| \leq \zeta_j \overset{\text{def}}{=} \frac{1}{\eta} \min_{1 \leqslant k \leqslant m-j+1} \sqrt{\tau_{m+1-k}\tau_{j+k-1}} \tag{3.13a}$$

$$\leq \frac{1}{\eta} \sqrt{\tau_m \tau_j}, \tag{3.13b}$$

where τ_i $(1 \leqslant i \leqslant m)$ are the eigenvalues of $E^* E$ as in Lemma 3.1.

Proof. Inequality (3.13b) follows from (3.13a) by simply picking $k = 1$ without the minimization. We now prove (3.13a). Let $E = U \Sigma V^*$ be the SVD of E, where U and V are unitary and

$$\Sigma = \begin{cases} \begin{pmatrix} \mathrm{diag}(\sqrt{\tau_1}, \sqrt{\tau_2}, \ldots, \sqrt{\tau_m}) \\ 0_{(n-m) \times m} \end{pmatrix}, & \text{if } n \geq m, \\ \left(\mathrm{diag}(\sqrt{\tau_{m-n+1}}, \sqrt{\tau_{m+n+2}}, \ldots, \sqrt{\tau_m}) \;\; 0_{n \times (m-n)} \right), & \text{if } n < m. \end{cases}$$

Note that in the case when $n < m$, $\tau_1 = \cdots = \tau_{m-n} = 0$. We have $E^* A_{22}^{-1} E = V \Sigma^* U^* A_{22}^{-1} U \Sigma V^*$ which has the same eigenvalues as $\Sigma^* U^* A_{22}^{-1} U \Sigma$. It can be proven that for either $n \geq m$ or $n < m$,

$$\Sigma^* U^* A_{22}^{-1} U \Sigma = DWD$$

for some matrix W satisfying $\|W\|_2 \leq 1/\eta$ and $D = \mathrm{diag}(\sqrt{\tau_1}, \ldots, \sqrt{\tau_m})$. Now apply Lemma 3.2 to complete the proof. $\square$

Remark 3.2. *When A_{22} is definite, we can get $|\vartheta_j| \leq \tau_j/\eta$ which is stronger than* (3.13a) *and thus* (3.13b).

Theorem 3.3. *For Hermitian matrix $\widetilde{A}$ as in* (1.1) *with $\mu = 0$, suppose $\varepsilon < \sqrt{3/4}\,\eta$. Then $\widetilde{A}$ has m eigenvalues $\theta_1, \ldots, \theta_m$ arranged such that*

$$|\theta_1| \leq |\theta_2| \leq \cdots \leq |\theta_m| \tag{3.14}$$

satisfying

$$|\theta_j| \leq \zeta_j + O(\varepsilon^4), \tag{3.15}$$

where ζ_j is defined by (3.13a).

Proof. It is a consequence of Theorem 3.1 and Lemma 3.3. $\square$

Next we derive strict bounds, i.e., without the term $O(\varepsilon^4)$ in (3.15). One difficulty here is that $\lambda_i(t)$ is no longer monotonic. However, the fact remains true that if $\theta_i \in (-\eta/2, \eta/2)$ is an eigenvalue of

$$B(\theta_i) = -E^* (A_{22} - \theta_i I)^{-1} E,$$

then θ_i is also an eigenvalue of $\widetilde{A}$.

Lemma 3.4. *Let $B(t)$ be defined as in* (3.10) *with eigenvalues*

$$\lambda_1(t) \leq \lambda_2(t) \leq \cdots \leq \lambda_m(t) \tag{3.16}$$

each of which are piecewise differentiable[5]. *If $\varepsilon < \eta/2$, then*

$$\left\| \frac{dB(t)}{dt} \right\|_2 \leqslant \frac{4\varepsilon^2}{\eta^2} < 1 \quad and \quad \left| \frac{d\lambda_j(t)}{dt} \right| \leqslant \frac{4\varepsilon^2}{\eta^2} < 1 \quad for \quad t \in (-\eta/2, \eta/2). \tag{3.17}$$

Proof. We have

$$\begin{aligned} B(t) - B(t + \Delta t) &= E^* (A_{22} - tI)^{-1} E - E^* [A_{22} - (t + \Delta t)I]^{-1} E \\ &= E^* \left\{ (A_{22} - tI)^{-1} - [A_{22} - (t + \Delta t)I]^{-1} \right\} E \\ &= E^* (A_{22} - tI)^{-1} \left\{ I - \left[I - \Delta t (A_{22} - tI)^{-1} \right]^{-1} \right\} E. \end{aligned}$$

Therefore

$$\begin{aligned} \left\| \frac{B(t) - B(t + \Delta t)}{\Delta t} \right\|_2 &= \frac{\left\| E^* (A_{22} - tI)^{-1} \left\{ I - \left[I - \Delta t (A_{22} - tI)^{-1} \right]^{-1} \right\} E \right\|_2}{|\Delta t|} \\ &\leqslant \frac{\varepsilon \| (A_{22} - tI)^{-1} \|_2 \left\| I - \left[I - \Delta t (A_{22} - tI)^{-1} \right]^{-1} \right\|_2 \varepsilon}{|\Delta t|}. \end{aligned}$$

---

[5] By [5, Theorem 4.8], there are countable points in $(-\eta/2, \eta/2)$ such that between any two nearby points, each $\lambda_i(t)$ is differentiable.

Noting that for $t \in (-\eta/2, \eta/2)$, we have

$$\|(A_{22} - tI)^{-1}\|_2 < \frac{2}{\eta},$$

$$\left\| I - \left[I - \Delta t (A_{22} - tI)^{-1} \right]^{-1} \right\|_2 < \frac{1}{1 - |\Delta t| \cdot 2/\eta} - 1$$
$$= \frac{|\Delta t| \cdot 2/\eta}{1 - |\Delta t| \cdot 2/\eta},$$

and thus

$$\left\| \frac{B(t) - B(t + \Delta t)}{\Delta t} \right\|_2 \leqslant \frac{\varepsilon^2 \cdot 2/\eta \cdot \frac{|\Delta t| \cdot 2/\eta}{1 - |\Delta t| \cdot 2/\eta}}{|\Delta t|} = \frac{4\varepsilon^2}{\eta^2 (1 - |\Delta t| \cdot 2/\eta)}.$$

Let $\Delta t \to 0$ to get

$$\left\| \frac{dB(t)}{dt} \right\|_2 \leqslant \frac{4\varepsilon^2}{\eta^2} < 1,$$

since $\varepsilon < \eta/2$. Finally, we use the well-known perturbation theorem (attributed to Lidskii, Weyl, Wiedlandt and Mirsky in various forms [2, pp. 196–205]) to conclude that

$$\left| \frac{d\lambda_j(t)}{dt} \right| \leq \left\| \frac{dB(t)}{dt} \right\|_2 \leqslant \frac{4\varepsilon^2}{\eta^2} < 1,$$

as expected. $\square$

Theorem 3.4. *For Hermitian $\widetilde{A}$ as in (1.1), if $\varepsilon < \eta/2$, then $\widetilde{A}$ has m eigenvalues $\theta_1, \ldots, \theta_m$ (arranged as in (3.14)) satisfying*

$$|\theta_j| \leqslant \frac{\zeta_j}{1 - 4\rho^2}, \tag{3.18}$$

for $1 \leqslant j \leqslant m$, where $\rho = \varepsilon/\eta < 1/2$ and ζ_j is defined by (3.13a).

Proof. Instead of proving (3.18) directly, we shall prove that for any given $j \in \{1, \ldots, m\}$ there are j of θ_i's satisfying $|\theta_i| \leq \zeta_j/(1 - 4\rho^2)$. Thus (3.18) must hold.

Adopt the notations in Lemmas 3.3 and 3.4. By (3.17), for any $t \in (-\eta/2, \eta/2)$, we have

$$|\lambda_i(t) - \lambda_i(0)| \leqslant \int_0^t \left| \frac{d\lambda_i(\tau)}{d\tau} \right| d\tau \leqslant \frac{4\varepsilon^2 |t|}{\eta^2} = 4\rho^2 |t| \tag{3.19}$$

for $1 \leqslant i \leqslant m$. Let $\delta_j = \frac{\zeta_j}{1 - 4\rho^2}$. We claim that there are at least j of $\lambda_i(t)$ such that

$$\lambda_i(t) \in [-\delta_j, \delta_j] \quad \text{for all } t \in [-\delta_j, \delta_j]. \tag{3.20}$$

This means that each function $\lambda_i(t)$ maps the interval $t \in [-\delta_j, \delta_j]$ into itself. By Brouwer's fixed point theorem, each of such $\lambda_i(t)$ has a fixed point $t_i \in [-\delta_j, \delta_j]$ such that $\lambda_i(t_i) = t_i$. Hence, recalling (3.2) we see that t_i is an eigenvalue of $\widetilde{A}$. Note that the second inequality in (3.17) implies that t_i is a *unique* fixed point of $\lambda_i(t)$ in $(-\eta/2, \eta/2)$. Therefore all counted, $\widetilde{A}$ has at least j eigenvalues in $[-\delta_j, \delta_j]$.

It remains to show that there are at least j of $\lambda_i(t)$ satisfying (3.20). To see this, we notice

$$\vartheta_k \in [-\zeta_k, \zeta_k] \subseteq [-\zeta_j, \zeta_j] \subset [-\delta_j, \delta_j] \quad \text{for } 1 \leq k \leq j.$$

These ϑ_k for $1 \leqslant k \leqslant j$ are taken by j different $\lambda_i(t)$ at $t = 0$, i.e., $\vartheta_k = \lambda_{\ell_k}(0)$, where $\ell_k \in \{1, \ldots, m\}$ are distinct for $k \in \{1, \ldots, j\}$. We now prove that $\lambda_{\ell_k}(t)$ for $k \in \{1, \ldots, j\}$ are the j of $\lambda_i(t)$ satisfying (3.20). In fact, for $t \in [-\delta_j, \delta_j]$ and $k \in \{1, \ldots, j\}$

$$|\lambda_{\ell_k}(t)| \leq |\lambda_{\ell_k}(0)| + |\lambda_{\ell_k}(t) - \lambda_{\ell_k}(0)|$$
$$= |\vartheta_k| + |\lambda_{\ell_k}(t) - \lambda_{\ell_k}(0)|$$
$$\leq \zeta_j + 4\rho^2 \delta_j$$
$$= \delta_j,$$

as expected. □

Remark 3.3. Compared with (3.15) in Theorem 3.3, the bound in (3.18) removes the term $O(\varepsilon^4)$ at the expense of the factor $(1 - 4\rho^2)^{-1}$.

Example 2.1 (revisit). The following table displays the eigenvalues θ_j of $\widetilde{A}$ nearest to 0, the eigenvalues ϑ_j of $-E^* A_{22}^{-1} E$, and the upper bounds in (3.18) and the ones after τ_j replaced by $m\epsilon_j^2$.

θ_j	ϑ_j	$\dfrac{\zeta_j}{1-4\rho^2}$	$\dfrac{m\epsilon_j\epsilon_m}{1-4\rho^2}$
$-3.0006 \cdot 10^{-4}$	$-3.002 \cdot 10^{-4}$	$5.0103 \cdot 10^{-4}$	$1.0020 \cdot 10^{-3}$
$1.6322 \cdot 10^{-7}$	$1.6322 \cdot 10^{-7}$	$7.0140 \cdot 10^{-6}$	$1.6157 \cdot 10^{-5}$

The bounds are rather sharp. □

4. Possible extensions to the generalized eigenvalue problem

So far we have focused on the Hermitian eigenvalue problem (1.1). We now consider the following Hermitian definite generalized eigenvalue problem

$$\widetilde{A} = \begin{pmatrix} \mu B_{11} & E^* \\ E & A_{22} \end{pmatrix}, \quad \widetilde{B} = \begin{pmatrix} B_{11} & F^* \\ F & B_{22} \end{pmatrix}, \tag{4.1}$$

where $B_{ii} \succ 0$, and $\|F\|_2$ is sufficiently small [6] so that $\widetilde{B} \succ 0$ also.

If $E = F = 0$, then μ is an eigenvalue of the pencil $\widetilde{A} - \lambda \widetilde{B}$ of multiplicity m. In this section we outline how to develop perturbation bounds using what we have gotten in Section 3.

4.1. Special case: $B_{ii} = I$ and $\mu = 0$

In this case,

$$\widetilde{A} = \begin{pmatrix} 0 & E^* \\ E & A_{22} \end{pmatrix}, \quad \widetilde{B} = \begin{pmatrix} I_m & F^* \\ F & I_n \end{pmatrix}. \tag{4.2}$$

Assume that $\|F\|_2 < 1$. A similar approach to the one in [6, Section 2.1] can be applied as follows. We first let

$$X = \begin{pmatrix} I_m & -F^* \\ 0 & I_n \end{pmatrix}, \quad W = \begin{pmatrix} I_m & 0 \\ 0 & [I - FF^*]^{1/2} \end{pmatrix}, \tag{4.3}$$

and then let

$$\widehat{B} \stackrel{\text{def}}{=} X^* \widetilde{B} X = \begin{pmatrix} I_m & 0 \\ 0 & I - FF^* \end{pmatrix} = W^2, \tag{4.4a}$$

---

[6] For example, $\|F\|_2 < \min_i\{\sigma_{\min}(B_{ii})\}$ guarantees $\widetilde{B} \succ 0$, where $\sigma_{\min}(B_{ii})$ is the smallest singular value of B_{ii}.

$$\widehat{A} \stackrel{\text{def}}{=} X^* \widetilde{A} X = \begin{pmatrix} 0 & E^* \\ E & \widehat{A}_{22} \end{pmatrix}, \tag{4.4b}$$

where W is the unique Hermitian definite square root [7, Chapter 6] of $\widehat{B}$, and

$$\widehat{A}_{22} = A_{22} - EF^* - FE^*.$$

$\widetilde{A} - \lambda \widetilde{B}$ has the same eigenvalues as $W^{-1}\widehat{A}W^{-1} - \lambda I_N$. Since $W^{-1}\widehat{A}W^{-1}$ takes the form of (1.1), our theory in Section 3 applies to $W^{-1}\widehat{A}W^{-1}$, leading to various bounds.

4.2. General case

Now we consider the general case (4.1). Assume $\mu = 0$; otherwise we shall consider

$$(\widetilde{A} - \mu \widetilde{B}) - \lambda \widetilde{B}$$

instead. Suppose

$$\widetilde{A} = \begin{pmatrix} 0 & E^* \\ E & A_{22} \end{pmatrix}, \quad \widetilde{B} = \begin{pmatrix} B_{11} & F^* \\ F & B_{22} \end{pmatrix}. \tag{4.5}$$

Set $Y = \text{diag}(B_{11}^{-1/2}, B_{22}^{-1/2})$ to get

$$Y^* \widetilde{A} Y = \begin{pmatrix} 0 & \widehat{E}^* \\ \widehat{E} & \widehat{A}_{22} \end{pmatrix}, \quad Y^* \widetilde{B} Y = \begin{pmatrix} I_m & \widehat{F}^* \\ \widehat{F} & I_n \end{pmatrix}, \tag{4.6}$$

which reduces to the case in Section 4.1, where

$$\widehat{A}_{22} = B_{22}^{-1/2} A_{22} B_{22}^{-1/2}, \quad \widehat{F} = B_{22}^{-1/2} F B_{11}^{-1/2}, \quad \widehat{E} = B_{22}^{-1/2} E B_{11}^{-1/2}. \tag{4.7}$$

5. Conclusion

We established perturbation bounds for the multiple eigenvalue μ of Hermitian matrix A under a perturbation in the off-diagonal block:

$$A = \begin{pmatrix} \mu I_m & 0 \\ 0 & A_{22} \end{pmatrix} \quad \text{perturbed to} \quad \widetilde{A} = \begin{pmatrix} \mu I_m & E^* \\ E & A_{22} \end{pmatrix},$$

with an emphasis on the case where the magnitudes of the columns of E vary widely. We show that whether $A_{22} - \mu I_m$ is definite or not plays a major role: if it is (positive or negative) definite, then $\widetilde{A}$ has m eigenvalues θ_i $(1 \leqslant i \leqslant m)$ such that the ith difference $|\theta_i - \mu|$ is bounded by a quantity that is proportional to the square of the norm of E's ith column, but when $A_{22} - \mu I_m$ is indefinite the quantity is proportional to the product of the ith column norm and the norm of E. We also outline a possible extension to the Hermitian definite generalized eigenvalue problem.

Acknowledgment

References

[1] C.-K. Li, R.-C. Li, A note on eigenvalues of perturbed Hermitian matrices, Linear Algebra Appl. 395 (2005) 183–190.
[2] G.W. Stewart, J.-G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.

[3] B.N. Parlett, The Symmetric Eigenvalue Problem, SIAM, Philadelphia, 1998.
[4] Y. Nakatsukasa, On the condition numbers of a multiple eigenvalue of a generalized eigenvalue problem, Numer. Math., in press. arXiv:1008.1201.
[5] K.R. Parthasarathy, Eigenvalues of matrix-valued analytic maps, J. Austral. Math. Soc. Ser. A 26 (1978) 179–197.
[6] R.-C. Li, Y. Nakatsukasa, N. Truhar, S.-F. Xu, Perturbation of partitioned hermitian generalized eigenvalue problem, SIAM J. Matrix Anal. Appl. 32 (2) (2011) 642–663.
[7] N.J. Higham, Functions of Matrices: Theory and Computation, SIAM, Philadelphia, PA, USA, 2008.