Relative perturbation bounds for the eigenvalues of diagonalizable and singular matrices – Application of perturbation theory for simple invariant subspaces

Yimin Wei \(^{a,b,*,1}\), Xiezhang Li \(^{c,2}\), Fanbin Bu \(^{d}\), Fuzhen Zhang \(^{e}\)

\(^{a}\) School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China
\(^{b}\) Key Laboratory of Mathematics for Nonlinear Sciences (Fudan University), Ministry of Education, Shanghai, 200433, PR China
\(^{c}\) Department of Mathematical Sciences, Georgia Southern University, Statesboro, GA 30460, USA
\(^{d}\) Institute of Mathematics, Fudan University, Shanghai 200433, PR China
\(^{e}\) Division of Math, Science and Technology, Nova Southeastern University, 3301 College Avenue, Fort Lauderdale, FL 33314, USA

Received 17 July 2005; accepted 16 June 2006
Available online 4 August 2006
Submitted by S. Kirkland

Abstract

Perturbation bounds for the relative error in the eigenvalues of diagonalizable and singular matrices are derived by using perturbation theory for simple invariant subspaces of a matrix and the group inverse of a matrix. These upper bounds are supplements to the related perturbation bounds for the eigenvalues of diagonalizable and nonsingular matrices.

© 2006 Elsevier Inc. All rights reserved.

AMS classification: 15A60; 65F35

Keywords: Diagonalizable and singular matrix; Simple invariant subspaces; Separation function; Relative perturbation bound

* Corresponding author. Address: School of Mathematical Sciences, Fudan University, Shanghai 200433, PR China. E-mail addresses: ymwei@fudan.edu.cn (Y. Wei), xli@georgiasouthern.edu (X. Li), fanbinbu@yahoo.com.cn (F. Bu), zhang@nova.edu (F. Zhang).

1 The work of this author was supported by the National Natural Science Foundation of China under grant 10471027 and Shanghai Education Committee.

2 Partial work of this author was finished when the author visited the Key Laboratory of Mathematics for Nonlinear Sciences of Fudan University.

0024-3795/ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2006.06.015
1. Introduction

Let $A$ be an $n \times n$ diagonalizable complex matrix with an eigendecomposition $A = X_A A X_A^{-1}$, where $X_A = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ and $X_A$ is an eigenmatrix of $A$. Assume that $E$ is a perturbation matrix of $A$ and that $\mu$ is an eigenvalue of $A + E$. The well-known Bauer–Fike theorem [1] reveals a bound for the absolute error between $\mu$ and the closest eigenvalue of $A$ by

$$\min_i |\lambda_i - \mu| \leq \kappa(X_A) \|E\|_2,$$

where $\kappa(X_A) = \|X_A\|_2 \|X_A^{-1}\|_2$ is the condition number of $X$ and $\|\cdot\|_2$ is the matrix 2-norm. Under an additional assumption that $A$ is nonsingular, Eisenstat and Ipsen [4] showed the following relative perturbation bound

$$\min_i \frac{|\lambda_i - \mu|}{|\lambda_i|} \leq \kappa(X_A) \|A^{-1}E\|_2. \quad (1.1)$$

They concluded that relative perturbation bounds are not necessarily stronger than absolute bounds, and the conditioning of an eigenvalue in the relative sense is the same as in the absolute sense. Recently, the conclusion has been extended to the case of singular and diagonalizable matrices by Eisenstat [3]. In that paper, the author developed an upper bound for $\mu > \kappa(X_A) \|E\|_2$ (i.e., $\mu$ is too large in magnitude for the zero eigenvalue of $A$ to satisfy the Bauer–Fike bound),

$$\min_{\lambda_i \neq 0} \frac{|\lambda_i - \mu|}{|\lambda_i|} \leq \sqrt{1 + \alpha^2 \kappa(X_A)} \|A^# E\|_2,$$  

where $\alpha = \kappa(X_A) \|E\|_2 / \sqrt{|\mu|^2 - (\kappa(X_A) \|E\|_2)^2}$ and $A^#$ denotes the group inverse of $A$. From (1.2) the author also derived a uniform upper bound

$$\min_{\lambda_i \neq 0} \frac{|\lambda_i - \mu|}{|\lambda_i|} \leq \sqrt{2 \kappa(X_A)} \|A^# E\|_2 \quad (1.3)$$

for $|\mu| \geq \sqrt{2 \kappa(X_A)} \|E\|_2$. In this paper, a completely new approach through perturbation theory of simple invariant subspaces is developed to derive a relative perturbation bound, which is an improvement on our earlier results [5,10,11].

Some preliminary results about the perturbation theory of simple invariant spaces of a matrix and the separation function of two square matrices are introduced in Section 2. A relative perturbation bound with respect to norms for nonzero eigenvalues of diagonalizable and singular matrices is derived with an application of perturbation theory of simple invariant spaces in Section 3.

2. Preliminaries

Let $A \in \mathbb{C}^{n \times n}$ be a diagonalizable and singular matrix with rank($A$) = $r < n$. There are, from [6, Theorem 5.1.5], a unitary matrix $[X_1 \quad Y_2]$ with $[X_1 \quad X_2]^{-1} = [Y_1 \quad Y_2]^H$ such that the spectral resolution of $A$ is given by

$$[Y_1^H \quad Y_2^H] A[X_1 \quad X_2] = \begin{bmatrix} C & O \\ O & O_{n-r} \end{bmatrix},$$  

where $C \in \mathbb{C}^{r \times r}$ is nonsingular and $O_{n-r}$ is the zero matrix of order $n - r$. The group inverse $A^#$ (see [2,8]) is given by

$$A^# = [X_1 \quad X_2] \begin{bmatrix} C^{-1} & O \\ O & O_{n-r} \end{bmatrix} \begin{bmatrix} Y_1^H \\ Y_2^H \end{bmatrix}. \quad (2.2)$$
Let $E$ be a perturbation matrix of $A$. As $E$ approaches zero, each eigenvalue of $A + E$ approaches the corresponding eigenvalue of $A$. Then a relative perturbation bound is meaningful only for those nonzero eigenvalues of $A$. On the other hand, we would estimate an upper bound of those eigenvalues of $A + E$ which approach zero. For this purpose, we will apply perturbation theory of simple invariant subspaces of a square matrix due to Stewart and Sun [6, Chapter V].

Writing
\[
\begin{bmatrix}
Y^H_1 \\
Y^H_2
\end{bmatrix} E [X_1 \ X_2] = \begin{bmatrix}
F_{11} & F_{12} \\
F_{21} & F_{22}
\end{bmatrix},
\]
we have
\[
\begin{bmatrix}
Y^H_1 \\
Y^H_2
\end{bmatrix} (A + E) [X_1 \ X_2] = \begin{bmatrix}
C + F_{11} & F_{12} \\
F_{21} & F_{22}
\end{bmatrix}.
\]

It follows that
\[
X_1 Y^H_1 = AA^#, \quad X_2 Y^H_2 = I - AA^#,
\]
\[
C^{-1} = Y^H_1 A^# X_1,
\]
\[
F_{i,j} = Y^H_i E X_j, \quad i, j = 1, 2.
\]

Recall $[X_1 \ X_2]$ is a unitary matrix. With the notations above, the following relations are easily derived from (2.1) to (2.4):
\[
\|AA^#\|_2 = \|I - AA^#\|_2 \quad \text{if $A$ is a nonzero singular matrix (see [9])},
\]
\[
\|C^{-1}\|_2 = \|Y^H_1 A^# X_1\|_2 = \|X_1 Y^H_1 A^# X_1\|_2 = \|A^# X_1\|_2 \leq \|A^#\|_2,
\]
\[
\|F_{11}\|_2 \leq \|AA^#\|_2, \quad \|F_{12}\|_2 = \|AA^# E (I - AA^#)\|_2,
\]
\[
\|F_{21}\|_2 \leq \|E\|_2, \quad \|F_{22}\|_2 \leq \|E (I - AA^#)\|_2.
\]

We next introduce simple invariant subspaces of $A$ and the separation function (see [6, Chapter V] and [7]). It follows from (2.1) that
\[
AX_1 = X_1 C \quad \text{and} \quad AX_2 = X_2 O_{n-r}.
\]

Let $\sigma(M)$ denote the spectrum of a square matrix $M$. Since $\sigma(C)$ and $\sigma(O_{n-r})$ are disjoint, the ranges of $X_1$ and $X_2$ are called simple invariant subspaces of $A$. The separation function of two arbitrary square matrices $S$ and $H$ is defined as
\[
\text{sep}_2(S, H) := \begin{cases} 
\min_{\|W\|_2 = 1} \|WS - HW\|_2 & \text{if } \sigma(S) \cap \sigma(H) = \emptyset, \\
0 & \text{otherwise}.
\end{cases}
\]

With notations in (2.1) and (2.3), it follows from [6, Theorem 5.2.8] that if
\[
\text{sep}_2(C, O_{n-r}) - \|F_{11}\|_2 - \|F_{22}\|_2 > 2\sqrt{\|F_{12}\|_2 \|F_{21}\|_2},
\]
there is a unique matrix $P$ satisfying
\[
\|P\|_2 < \frac{2\|F_{21}\|_2}{\text{sep}_2(C, O_{n-r}) - \|F_{11}\|_2 - \|F_{22}\|_2},
\]
such that
\[
\begin{bmatrix}
Y^H_1 \\
Y^H_2 - PY^H_1
\end{bmatrix} (A + E) [X_1 + X_2 P \ X_2] = \begin{bmatrix}
C + F_{11} + F_{12} P & F_{12} \\
O & F_{22} - PF_{12}
\end{bmatrix}.
\]
3. Relative bounds with respect to norms

In this section, we derive relative perturbation bounds for the eigenvalues of diagonalizable and singular matrix with an application of perturbation theory of simple invariant subspaces.

Lemma 3.1. Let $C$ be nonsingular, and let $O$ be a square null matrix. Then
\begin{equation}
\text{sep}_2(C, O) = \frac{1}{\|C^{-1}\|_2}.
\end{equation}

Proof. The result follows directly from the definition of $\text{sep}_2(C, O)$ and the fact that $\|WC\|_2 \geq \|C^{-1}\|_2 \|W\|_2$, with equality for a $W$ whose rows are multiples of left singular vectors corresponding to the smallest singular value of $C$. □

Eq. (3.1) will be used in the proof of the following main result of this paper.

Theorem 3.2. Let $A$ be a diagonalizable and singular matrix, and let $E$ be a perturbation matrix. If
\begin{equation}
\|E\|_2 < \frac{1}{4\|A\|^2_2 \|AA^\#\|^2_2},
\end{equation}
then for $\mu \in \sigma(A + E)$, either
\begin{equation}
|\mu| \leq \|E(I - AA^\#)\|_2 + p \|A^\#E(I - AA^\#)\|_2
\end{equation}
or
\begin{equation}
\min_{0 \neq \lambda \in \sigma(A)} \frac{|\lambda - \mu|}{|\lambda|} \leq \kappa_2(X_A) \left\{ \|A^\#E\|_2 + p \|A^\#E(I - AA^\#)\|_2 \right\},
\end{equation}
where $X_A$ is an eigenmatrix of $A$ and
\begin{equation}
p = \frac{2\|E\|_2}{\|A\|^2_2 - \|AA^\#E\|_2 - \|E(I - AA^\#)\|_2}.
\end{equation}

Proof. Suppose that $A \in \mathbb{C}^{n \times n}$ has an eigendecomposition,
\begin{equation}
X_A^{-1}AX_A = \begin{bmatrix} D & O \\ O & O_{n-r} \end{bmatrix},
\end{equation}
where $D$ is nonsingular diagonal matrix of order $r, 1 \leq r \leq n-1$, and $X_A$ is an eigenmatrix of $A$. Let $X_A = [S_1 \ S_2]$ and $X_A^{-1} = [T_1 \ T_2]^H$, where $S_1, T_1 \in \mathbb{C}^{n \times r}$. Suppose that $S_1 = X_1R_1$ and $T_2 = Y_2R_2$ are $QR$ decompositions of $S_1$ and $T_2$, respectively. We denote $X_2 := S_2R_2^H$ and $Y_1 := T_1R_1^H$.

It follows that $[X_1 \ X_2]^{-1} = [Y_1 \ Y_2]^H$, $[X_1 \ Y_2]^H = [X_1 \ Y_2]^{-1}$, and the spectral resolution of $A$ is given by
\begin{equation}
\begin{bmatrix} Y_1^H \\ Y_2^H \end{bmatrix} A[X_1 \ X_2] = \begin{bmatrix} C & O \\ O & O \end{bmatrix},
\end{equation}
where $C = R_1DR_1^{-1}$ is nonsingular. So $X_C := R_1$ is an eigenmatrix of $C$. Moreover, we have
\begin{equation}
\kappa(X_C) \leq \kappa(X_A).
\end{equation}
In fact, \( X_A = [X_1 X_C \quad X_2 R_2^{-H}] \) and \( X_A^{-1} = [Y_1 X_C^{-H} \quad Y_2 R_2]^{H} \), which implies that

\[
\begin{align*}
\|X_C\|_2 &= \|X_1 X_C\|_2 \leq \|X_A\|_2, \\
\|X_C^{-1}\|_2 &= \|X_C^{-1} Y_1^H X_1\|_2 \leq \|X_C^{-1} Y_1^H\|_2 \|X_1\|_2 = \|Y_1 X_C^{-H}\|_2 \leq \|X_A^{-1}\|_2.
\end{align*}
\]

With the assumption (3.2) and relations in (2.5), it is obvious that

\[
\|E\|_2 < \frac{1}{4\|A\|_2 \|AA^#\|_2} \leq \frac{1}{4\|C^{-1}\|_2 \|AA^#\|_2}.
\]

It follows from relations in (2.5) and (3.1) that

\[
\|F_1\|_2 + \|F_2\|_2 + 2\sqrt{\|F_1\|_2 \|F_2\|_2} \leq 4\|E\|_2 \|AA^#\|_2 < \|C^{-1}\|_2^{-1} = \text{sep}_2(C, O).
\]

Then according to (2.7), there exists a unique matrix \( P \) satisfying

\[
\|P\|_2 < \frac{2\|F_2\|_2}{\|C^{-1}\|_2^{-1} \|F_1\|_2 - \|F_2\|_2} \leq \frac{2\|E\|_2}{\|A\|_2^{-1} \|AA^#\|_2 - \|E(I - AA^#)\|_2} = p
\]

such that

\[
(A + E) \quad \text{similar to} \quad \begin{bmatrix} C + F_{11} + F_{12} P & F_{12} \\ O & F_{22} - P F_{12} \end{bmatrix}.
\]

So,

\[
\sigma(A + E) = \sigma(F_2 - P F_{12}) \cup \sigma(C + F_{11} + F_{12} P).
\]

We first consider the case when \( \mu \in \sigma(F_2 - P F_{12}) \). It follows from (2.5) and (3.7) that

\[
|\mu| \leq \|F_2 - P F_{12}\|_2 \leq \|F_2\|_2 + \|P\|_2 \|F_{12}\|_2 \leq \|E(I - AA^#)\|_2 + p \|AA^# E(I - AA^#)\|_2,
\]

which proves the content of (3.3).

For the second case when \( \mu \in \sigma(C + F_{11} + F_{12} P) \). According to (3.5), the set of all nonzero eigenvalues of \( A \) and the spectrum of \( C \) are identical. With Eisenstat and Ipsen’s result in (1.1), and (3.6) above, we have

\[
\min_{0 \neq \lambda \in \sigma(A)} \frac{|\lambda - \mu|}{|\lambda|} = \min_{\lambda \in \sigma(C)} \frac{|\lambda - \mu|}{|\lambda|} \leq \kappa(X_C) \|C^{-1}(F_{11} + F_{12} P)\|_2 \leq \kappa(X_A) \|C^{-1}(F_{11} + F_{12} P)\|_2.
\]

Furthermore, it follows from (2.4) and (2.5) that

\[
\|C^{-1}(F_{11} + F_{12} P)\|_2 = \left\| Y_1^H A^# X_1 \left( Y_1^H E X_1 + Y_1^H E X_2 P \right) \right\|_2 \leq \|A^# E\|_2 + \|P\|_2 \|A^# E(I - AA^#)\|_2.
\]

Thus, (3.4) follows by replacing \( \|P\|_2 \) with its upper bound in (3.7). We complete the proof of the theorem. □
Remark. There is a tighter bound than (3.4) for small \( \|P\|_2 \). Writing \( C = CX_C \cdot X_C^{-1} \equiv C_1C_2 \) and using [4, Theorem 2.3] and the fact that \( C_2C_1 = X_C^{-1} \cdot CX_C = D \) is diagonal, we have

\[
\min_{\lambda \in \sigma(C)} \frac{|\lambda - \mu|}{|\lambda|} \leq \|C_1^{-1}(F_{11} + F_{12}P)C_2^{-1}\|_2 = \|X_C^{-1}C_1^{-1}(F_{11} + F_{12}P)X_C\|_2.
\]

But with the notation above,

\[
X_C^{-1}(F_{11} + F_{12}P)X_C = R_1^{-1} \cdot Y_1^H A^# X_1 \left( Y_1^H EX_1 + Y_1^H EX_2 \cdot P \right) R_1
\]

\[
= R_1^{-1}Y_1^H \cdot A^# AA^# E \left( X_1 R_1 \cdot I + X_2 R_2^{-H} \cdot R_2^H PR_1 \right)
\]

\[
= T_1^H \cdot A^# E \cdot [S_1 \ S_2] \left[ \begin{array}{c} I \\ R_1^H PR_1 \end{array} \right].
\]

Thus,

\[
\min_{0 \neq \lambda \in \sigma(A)} \frac{|\lambda - \mu|}{|\lambda|} = \min_{\lambda \in \sigma(C)} \frac{|\lambda - \mu|}{|\lambda|} \leq \|X_A^{-1}\|_2 \|A^# E\|_2 \|X_A\|_2 \left( 1 + \|R_2^H PR_1\|_2^2 \right) \leq \kappa(X_A) \|A^# E\|_2 \left( 1 + \frac{1}{2} \kappa(X_A)^2 \|P\|_2^2 \right),
\]

(3.10)
since \( \|R_2^H PR_1\|_2 \leq \|T_2^H\|_2 \|P\|_2 \|S_1\|_2 \leq \kappa(X_A) \|P\|_2 \).

We remark that for a sufficiently small \( \|E\|_2 \), the upper bounds in (3.4) and (3.10) are dominated by \( \kappa(X_A) \|A^# E\|_2 \). The term \( \kappa(X_A) \) could be interpreted as an approximate condition number of the relative perturbation error for eigenvalues of \( A \) up to the first order of \( \|E\|_2 \). Inequality (1.2) gives an upper bound for a specific \( \mu \) which is known. (1.3) gives a simple uniform upper bound.

However, it follows from (3.7) that if

\[
\|A^#\|_2 \|AA^#\|_2 \|E\|_2 < \frac{2 - \sqrt{2}}{4},
\]

then \( p \|AA^#\|_2 < \sqrt{2} - 1 \), which implies that \( \|A^# E\|_2 + p \|A^# E(I - AA^#)\|_2 < \sqrt{2} \|A^# E\|_2 \). Thus, in this case, (3.4) gives a better bound than (1.3). In an analogous way, if \( \|P\|_2 \) is small such that \( \kappa(X_A) \|P\|_2 < \sqrt{2} \left( \sqrt{2} - 1 \right) \), then \( 1 + \frac{\kappa(X_A)^2 \|P\|_2^2}{2} < \sqrt{2} \). Thus, in this case, (3.10) gives a better bound than (1.3).

Acknowledgements

The authors thank Professor S. Kirkland and referees for their valuable suggestions and very detailed comments on earlier drafts of the manuscript. The first author would like to thank Professors Ren-cang Li and Qiang Ye for their useful discussions and hospitality during his visit at University of Kentucky in 2000.

References