ELSEVIER Linear Algebra and its Applications 309 (2000) 45–56

LINEAR ALGEBRA AND ITS APPLICATIONS

www.elsevier.com/locate/laa

Absolute and relative perturbation bounds for invariant subspaces of matrices

Ilse C.F. Ipsen [*,1]

Department of Mathematics, Center for Research in Scientific Computation, North Carolina State University, P.O. Box 8205, Raleigh, NC 27695-8205, USA

Received 19 August 1998; accepted 14 May 1999

Submitted by J.L. Barlow

Abstract

Absolute and relative perturbation bounds are derived for angles between invariant subspaces of complex square matrices, in the two-norm and in the Frobenius norm. The absolute bounds can be considered extensions of Davis and Kahan's $\sin\theta$ theorem to general matrices and invariant subspaces of any dimension. The relative bounds are the most general relative bounds for invariant subspaces because they place no restrictions on the matrix or the perturbation. When the perturbed subspace has dimension one, the relative bound is implied by the absolute bound. © 2000 Elsevier Science Inc. All rights reserved.

AMS classification: 15A18; 15A42; 65F15

Keywords: Invariant subspace; Absolute error; Relative error; Eigenvalue separation; Angle between subspaces

1. Introduction

The goal is to derive bounds on the angle between an invariant subspace of a complex square matrix and a perturbed subspace. Two types of bounds will be derived, absolute and relative.

[*] E-mail: ipsen@math.ncsu.edu, http://www4.ncsu.edu/~ipsen/info.html

[1] This research was supported in part by NSF grants CCR-9400921 and DMS-9714811.

PII: S 0 0 2 4 - 3 7 9 5 (9 9) 0 0 1 0 4 - 4

Example 1. Let

$$A \equiv \begin{pmatrix} a & & \\ & b & \\ & & c \end{pmatrix}$$

be a diagonal matrix of order 3 with distinct eigenvalues a, b and c. An eigenvector associated with eigenvalue c is $(0\ 0\ 1)^{\mathrm{T}}$, where the superscript denotes the transpose.

Let

$$A + E_1 \equiv \begin{pmatrix} a & \epsilon & \epsilon \\ & b & \epsilon \\ & & c \end{pmatrix}$$

be a perturbed matrix with the same eigenvalues as A. An eigenvector of $A + E_1$ associated with eigenvalue c is

$$\begin{pmatrix} \frac{\epsilon}{c-a}\left(\frac{\epsilon}{c-b}+1\right) & \frac{\epsilon}{c-b} & 1 \end{pmatrix}^{\mathrm{T}}.$$

The difference between exact and perturbed eigenvector depends on $\epsilon/(c-a)$ and $\epsilon/(c-b)$. This suggests that the sine of their angle can be bounded in terms of

$$\|E_1\| / \min\{|c-a|, |c-b|\}, \tag{1.1}$$

as $\|E_1\| = \mathrm{O}(|\epsilon|)$.

Now consider the perturbed matrix

$$A + E_2 \equiv \begin{pmatrix} a & a\epsilon & a\epsilon \\ & b & b\epsilon \\ & & c \end{pmatrix},$$

again with the same eigenvalues as A. An eigenvector of $A + E_2$ associated with eigenvalue c is

$$\begin{pmatrix} \frac{\epsilon a}{c-a}\left(\frac{\epsilon b}{c-b}+1\right) & \frac{\epsilon b}{c-b} & 1 \end{pmatrix}.$$

The difference between the two eigenvectors depends on $\epsilon a/(c-a)$ and $\epsilon b/(c-b)$. This suggests that the sine of their angle can be bounded in terms of

$$\|A^{-1}E_2\| / \min\left\{\frac{|c-a|}{|a|}, \frac{|c-b|}{|b|}\right\}, \tag{1.2}$$

as $\|A^{-1}E_2\| = \mathrm{O}(|\epsilon|)$.

The bound (1.1) is an absolute bound and $\min\{|c-a|, |c-b|\}$ is an absolute eigenvalue separation, while (1.2) is a relative bound and

$$\min\left\{\frac{|c-a|}{|a|},\frac{|c-b|}{|b|}\right\}$$

is a relative eigenvalue separation.

For a perturbed matrix $A+E$ we derive absolute and relative bounds on the angle between a true and a perturbed subspace, in the two-norm and in the Frobenius norm. The absolute bounds contain $\|E\|$ and an absolute separation, while the relative bounds contain $\|A^{-1}E\|$ and a relative separation. This means, absolute bounds measure sensitivity with regard to perturbations E, while the relative bounds measure sensitivity with regard to perturbations $A^{-1}E$. The estimates provided by the two types of bounds can be very different. Which bound to use depends on the matrix and on the perturbation.

Among relative bounds for subspaces, the bounds presented here are the most general because they place no restriction on the original matrix A (other than non-singularity), the perturbation E, or the dimensions of the subspaces. They demonstrate the following:

1. Relative bounds for invariant subspaces always exist, for any non-singular matrix A and any perturbation E. In this sense relative bounds appear to be no more special than absolute bounds.
2. When the perturbed eigenspace has dimension one, the relative bound can be derived from the absolute bound. Hence relative bounds are not necessarily stronger than absolute bounds.

After stating the problem in Section 2, we derive in Section 3 absolute and relative bounds that make no reference to any basis, and we show that the absolute bound implies the relative bound when the perturbed subspace has dimension one. In Section 4 the bounds are expressed in terms of subspace bases. In Section 5 diagonalizable matrices are considered and the separations are expressed in terms of eigenvalues.

Notation. I is the identity matrix; $\|\cdot\|_2$ is the two-norm; $\|\cdot\|_F$ the Frobenius norm; and $\|\cdot\|$ stands for both norms. The conjugate transpose of a matrix A is A^*; and $A^\dagger$ is its Moore–Penrose inverse. The condition number with respect to inversion of a full-rank matrix Y is $\kappa(Y)\equiv\|Y\|_2\|Y^\dagger\|_2$.

2. Statement of the problem

Let A be a complex square matrix. A subspace $\mathscr{S}$ is an *invariant subspace* of A if $Ax\in\mathscr{S}$ for every $x\in\mathscr{S}$ [7, Section 1.1; 14, Section I.3.4]. Applications involving invariant subspaces are given in [7].

Let the perturbed matrix $A+E$ have an invariant subspace $\hat{\mathscr{S}}$, whose dimension may be different from that of $\mathscr{S}$. The *gap* or *distance* between $\mathscr{S}$ and $\hat{\mathscr{S}}$ is defined as [2, p. 202; 7, Section 13.1] $\|P_{\mathscr{S}}-\hat{P}\|$, where $P_{\mathscr{S}}$ is the orthogonal projector onto $\mathscr{S}$, and $\hat{P}$ is the orthogonal projector onto $\hat{\mathscr{S}}$. In the two-norm one has [2, Exercise VII.1.11; 19, Section 1]

$$\|P_{\mathscr{S}}-\hat{P}\|_2 = \|P\hat{P}\|_2,$$

where P is the orthogonal projector onto $\mathscr{S}^{\perp}$. In the Frobenius norm [2, p. 202] implies

$$\|P_{\mathscr{S}}-\hat{P}\|_{\mathrm{F}}^2 = 2\|P\hat{P}\|_{\mathrm{F}}^2 + \dim(S) - \dim(\hat{\mathscr{S}}).$$

This means, to bound $\|P_{\mathscr{S}}-\hat{P}\|$ it suffices to bound $\|P\hat{P}\|$. When $\dim(S)=\dim(\hat{\mathscr{S}})$, the singular values of $P\hat{P}$ are the sines of the principal angles between $\mathscr{S}$ and $\hat{\mathscr{S}}$ [8, Section 12.4.3; 14, Theorem I.5.5]. Therefore we set (see also [19, Section 1])

$$\sin\Theta \equiv P\hat{P}.$$

The goal in this paper is to bound $\|\sin\Theta\|$, where $\|\cdot\|$ is the two-norm or the Frobenius norm.

3. Bounds without subspace bases

Absolute and relative perturbation bounds are derived for invariant subspaces of complex square matrices. The purpose is to show that there are bounds that make no reference to subspace bases and to provide a unifying framework for all subsequent bounds.

Since $\mathscr{S}^{\perp}$ is an invariant subspace of A^* [14, Theorem V.1.1], the associated projector P satisfies [7, (1.5.5); 9, Theorem 5.8.4]

$$PA = PAP, \quad (A+E)\hat{P} = \hat{P}(A+E)\hat{P}. \tag{3.1}$$

The absolute bound is expressed in terms of an absolute separation between A and $A+E$,

$$\mathrm{abssep} \equiv \mathrm{abssep}_{\{A,A+E\}} \equiv \min_{\|Z\|=1,\, PZ\hat{P}=Z} \|PAZ - Z(A+E)\hat{P}\|.$$

Theorem 3.1. *If* abssep >0 *then*

$$\|\sin\Theta\| \leqslant \|E\|/\mathrm{abssep}.$$

Proof. Let $\hat{X}$ be a matrix whose columns span $\hat{\mathscr{S}}$, so $\hat{\mathscr{S}}=\mathrm{range}(\hat{X})$. Then there exists a unique matrix $\hat{B}$ such that [14, Theorem I.3.9]

$$(A+E)\hat{X} = \hat{X}\hat{B}.$$

Multiply on the left by P and on the right by $\hat{X}^\dagger$ and use the fact that $\hat{P} = \hat{X}\hat{X}^\dagger$ [14, Theorem III.1.3], then

$$-PE\hat{P} = PA\hat{P} - P\hat{X}\hat{B}\hat{X}^\dagger = PA\hat{P} - P(A+E)\hat{P}.$$

From (3.1) follows

$$-PE\hat{P} = PA\sin\Theta - \sin\Theta(A+E)\hat{P},$$

and $\sin\Theta = P\sin\Theta\hat{P}$ implies

$$\|E\| \geqslant \|PE\hat{P}\| \geqslant \mathrm{abssep}\|\sin\Theta\|. \qquad \square$$

The relative bound is expressed in terms of a relative separation between A and $A+E$,

$$\mathrm{relsep} \equiv \mathrm{relsep}_{\{A,A+E\}} \equiv \min_{\|Z\|=1, PZ\hat{P}=Z} \|PA^{-1}(PAZ - Z(A+E)\hat{P})\|,$$

provided A is non-singular.

Theorem 3.2. *If A is non-singular and* relsep > 0 *then*

$$\|\sin\Theta\| \leqslant \|A^{-1}E\|/\mathrm{relsep}.$$

Proof. As in the proof of Theorem 3.1, $(A+E)\hat{X} = \hat{X}\hat{B}$ for some $\hat{B}$, and $\hat{P} = \hat{X}\hat{X}^\dagger$. Multiply $(A+E)\hat{X} = \hat{X}\hat{B}$ on the left by PA^{-1} and on the right by $\hat{X}^\dagger$,

$$-PA^{-1}E\hat{P} = P\hat{P} - PA^{-1}\hat{X}\hat{B}\hat{X}^\dagger = \sin\Theta - PA^{-1}(A+E)\hat{P}.$$

Then (3.1) implies

$$\begin{aligned} -PA^{-1}E\hat{P} &= \sin\Theta - PA^{-1}\sin\Theta(A+E)\hat{P} \\ &= PA^{-1}PA\sin\Theta - PA^{-1}\sin\Theta(A+E)\hat{P} \\ &= PA^{-1}(PA\sin\Theta - \sin\Theta(A+E)\hat{P}) \end{aligned}$$

and $\sin\Theta = P\sin\Theta\hat{P}$ implies

$$\|A^{-1}E\| \geqslant \|PA^{-1}E\hat{P}\| \geqslant \mathrm{relsep}\|\sin\Theta\|. \qquad \square$$

Theorem 3.2 shows that relative subspace perturbation bounds exist for any non-singular matrix.

It is now shown that the absolute bound implies the relative bound when the perturbed subspace has dimension one. Since $\hat{X}$ consists of only one column, and $\hat{B}$ is a scalar one can write instead

$$(A+E)\hat{x} = \hat{\lambda}\hat{x}.$$

Using $\hat{P} = \hat{x}\hat{x}^*/\hat{x}^*\hat{x}$, (3.1) and $PZ\hat{P} = Z$ one can express Theorem 3.1 as

$$\sin\Theta \leqslant \|E\|/\text{abssep}, \quad \text{where} \quad \text{abssep} = \min_{\|Z\|=1} \|P(A-\hat{\lambda}I)Z\|,$$

and Theorem 3.2 as

$$\sin\Theta \leqslant \|A^{-1}E\|/\text{relsep}, \quad \text{where} \quad \text{relsep} = \min_{\|Z\|=1} \|PA^{-1}(A-\hat{\lambda}I)Z\|.$$

Theorem 3.3. *If* $\hat{\mathscr{S}}$ *has dimension one then Theorem* 3.1 *implies Theorem* 3.2.

Proof. $(A+E)\hat{x} = \hat{\lambda}\hat{x}$ implies $(\tilde{A}+\tilde{E})\hat{x} = \hat{x}$, where

$$\tilde{A} \equiv \hat{\lambda}A^{-1}, \quad \tilde{E} \equiv -A^{-1}E.$$

Note that $\tilde{A}$ and $\tilde{A}+\tilde{E}$ are associated with the same projectors P and $\hat{P}$, respectively, as A and $A+E$.

Theorem 3.1 implies Theorem 3.2 because applying the absolute bound to $(\tilde{A}+\tilde{E})\hat{x} = 1\cdot\hat{x}$ yields the relative bound. In particular, the norm in abssep is

$$\|P(\tilde{A}-1\cdot I)Z\| = \|P(\hat{\lambda}A^{-1}-I)Z\| = \|PA^{-1}(A-\hat{\lambda}I)Z\|,$$

which is equal to the norm in relsep. □

Since the relative bound is derived by means of the absolute bound one cannot necessarily conclude that relative perturbation bounds are stronger than absolute bounds. However there are classes of matrices and perturbations where relative bounds can be much sharper than absolute bounds.

Example 2. Consider a special case of Example 1, where

$$A = \begin{pmatrix} 10^{-k} & & \\ & 2\cdot 10^{-k} & \\ & & 10^{k}, \end{pmatrix}, \quad k > 0.$$

Suppose $\mathscr{S} = \text{range}(1\ 0\ 0)^{\mathrm{T}}$ is approximated by the subspace associated with the smallest eigenvalue $\hat{\lambda} = 10^{-k}$ of

$$A+E = \begin{pmatrix} 10^{-k} & & \\ \epsilon 10^{-k} & 2\cdot 10^{-k} & \\ \epsilon 10^{k} & \epsilon 10^{k} & 10^{k}, \end{pmatrix}$$

where $\epsilon > 0$. In this case

$$\mathscr{S}^{\perp} = \text{range}\begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathscr{P} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The absolute bound contains

$$P(A - \hat{\lambda} I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 10^{-k} & 0 \\ 0 & 0 & 10^{k} - 10^{-k} \end{pmatrix}.$$

Hence in the two-norm abssep $\approx 10^{-k}$. Since $\|E\|_2 \approx \epsilon 10^k$, the absolute bound is

$$\|\sin \Theta\|_2 \leqslant \|E\|_2/\text{abssep} \approx \epsilon 10^{2k}.$$

In contrast, the relative bound contains

$$PA^{-1}(A - \hat{\lambda} I) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 - 10^{-2k} \end{pmatrix}.$$

Hence in the two-norm relsep ≈ 1. Since $\|A^{-1}E\|_2 \approx \epsilon$, the relative bound is

$$\|\sin \Theta\|_2 \leqslant \|A^{-1}E\|_2/\text{relsep} \approx \epsilon.$$

In this case the relative bound is sharper by a factor of 10^{2k} than the absolute bound.

4. Bounds with subspace bases

The absolute and relative bounds from the previous section are expressed in terms of subspace bases.

Let Y and $\hat{X}$ be respective bases for $\mathscr{S}^{\perp}$ and $\hat{\mathscr{S}}$, that is,

$$\mathscr{S}^{\perp} = \text{range}(Y), \qquad Y^{\dagger} Y = I, \qquad Y^* A = BY^*$$

for some B, and

$$\hat{\mathscr{S}} = \text{range}(\hat{X}), \qquad \hat{X}^{\dagger}\hat{X} = I, \qquad (A + E)\hat{X} = \hat{X}\hat{B}$$

for some $\hat{B}$. Let $Y = QR$ and $\hat{X} = \hat{Q}\hat{R}$ be QR decompositions where the columns of Q and $\hat{Q}$ are orthonormal bases for $\mathscr{S}^{\perp}$ and $\hat{\mathscr{S}}$, respectively. Then

$$Q^* A = (R^{-*} B R^*) Q^*, \qquad (A + E)\hat{Q} = \hat{Q}(\hat{R}\hat{B}\hat{R}^{-1}). \tag{4.1}$$

The absolute separation is expressed in terms of

$$\text{abssep}(F, G) \equiv \min_{\|Z\|=1} \|FZ - ZG\|,$$

which is the same as the separation defined in [14, Theorem V.2.1]. The following bound shows that this separation is weaker than the corresponding separation abssep in Section 3.

Theorem 4.1. *In Theorem* 3.1

$$\text{abssep} \geqslant \text{abssep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}),$$

hence

$$\|\sin\Theta\| \leqslant \|E\|/\text{abssep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}).$$

Proof. Let Z_0 attain the minimum in abssep, i.e.,

$$\text{abssep} = \|PAZ_0 - Z_0(A+E)\hat{P}\|, \quad \|Z_0\| = 1, \quad PZ_0\hat{P} = Z_0.$$

Then $P = QQ^{*}$, $\hat{P} = \hat{Q}\hat{Q}^{*}$ and (4.1) imply

$$\begin{aligned}\text{abssep} &= \|Q(R^{-*}BR^{*})Q^{*}Z_0 - Z_0\hat{Q}(\hat{R}\hat{B}\hat{R}^{-1})\hat{Q}^{*}\| \\ &= \|(R^{-*}BR^{*})(Q^{*}Z_0\hat{Q}) - (Q^{*}Z_0\hat{Q})(\hat{R}\hat{B}\hat{R}^{-1})\|.\end{aligned}$$

From $\|Q^{*}Z_0\hat{Q}\| = \|PZ_0\hat{Q}\| = \|Z_0\| = 1$ follows

$$\text{abssep} \geqslant \text{abssep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}). \qquad \square$$

The relative separation is expressed in terms of

$$\text{relsep}(F, G) \equiv \min_{\|Z\|=1} \|F^{-1}(FZ - ZG)\|.$$

Again, this separation is weaker than the corresponding separation relsep in Section 3, as the result below shows.

Theorem 4.2. *In Theorem* 3.2

$$\text{relsep} \geqslant \text{relsep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}),$$

hence

$$\|\sin\Theta\| \leqslant \|A^{-1}E\|/\text{relsep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}).$$

Proof. The proof is similar to that of Theorem 4.1. $\square$

For instance, in the special case when the columns of Y and $\hat{X}$ are already orthonormal one can choose $R = I$ and $\hat{R} = I$, and the bounds in Theorems 4.1 and 4.2 simplify,

$$\|\sin \Theta\| \leqslant \|E\|/\mathrm{abssep}(B, \hat{B}), \quad \|\sin \Theta\| \leqslant \|A^{-1}E\|/\mathrm{relsep}(B, \hat{B}).$$

A result in the same spirit as the absolute bound is [14, Theorem V.2.1].

To prepare for the next section, we remove information about the basis from the separation. Let $\kappa(F) \equiv \|F\|_2\|F^{-1}\|_2$ be a two-norm condition number with respect to inversion. Since

$$\mathrm{abssep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}) \geqslant \frac{\mathrm{abssep}(B, \hat{B})}{\kappa(Y)\kappa(\hat{X})}$$

[14, p. 245, Problem 1], one gets

$$\|\sin \Theta\| \leqslant \kappa(Y)\kappa(\hat{X})\|E\|/\mathrm{abssep}(B, \hat{B}). \tag{4.2}$$

Similarly,

$$\mathrm{relsep}(R^{-*}BR^{*}, \hat{R}\hat{B}\hat{R}^{-1}) \geqslant \frac{\mathrm{relsep}(B, \hat{B})}{\kappa(Y)\kappa(\hat{X})}$$

leads to

$$\|\sin \Theta\| \leqslant \kappa(Y)\kappa(\hat{X})\|A^{-1}E\|/\mathrm{relsep}(B, \hat{B}). \tag{4.3}$$

5. (Partially) diagonalizable matrices

The bounds (4.2) and (4.3) are expressed in terms of eigenvalues.

Let A and $A + E$ be partially diagonalizable, in the sense that B and $\hat{B}$ are diagonal. To make this clear, write

$$Y^{*}A = \Lambda Y^{*}, \qquad (A+E)\hat{X} = \hat{X}\hat{\Lambda},$$

with Λ and $\hat{\Lambda}$ diagonal. We use the notation

$$\min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda - \hat{\lambda}| \qquad \text{and} \qquad \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{|\lambda|},$$

where the minima range over all diagonal elements λ of Λ and all diagonal elements $\hat{\lambda}$ of $\hat{\Lambda}$.

Inequalities (4.2) and (4.3) readily lead to Frobenius norm bounds.

Theorem 5.1. *If Λ and $\hat{\Lambda}$ are diagonal then*

$$\|\sin \Theta\|_{\mathrm{F}} \leqslant \kappa(Y)\kappa(\hat{X})\|E\|_{\mathrm{F}}/\min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda - \hat{\lambda}|.$$

If, in addition, A is non-singular, then

$$\|\sin\Theta\|_{\mathrm{F}} \leqslant \kappa(Y)\kappa(\hat{X})\|A^{-1}E\|_{\mathrm{F}} / \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}.$$

Proof. The absolute bound follows from the fact that [14, p. 245, Problem 3] .

$$\operatorname{abssep}_{\mathrm{F}}(\Lambda,\hat{\Lambda}) = \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}|,$$

see also [14, p. 245, Problem 4]. With regard to the relative bound,

$$\|Z-\Lambda^{-1}Z\hat{\Lambda}\|_{\mathrm{F}}^2 = \sum_{i,j}\left|1-\frac{\hat{\lambda}_j}{\lambda_i}\right|^2 |z_{ij}|^2 \geqslant \min_{i,j}\left|1-\frac{\hat{\lambda}_j}{\lambda_i}\right|^2 \|Z\|_{\mathrm{F}}^2$$

implies

$$\operatorname{relsep}_{\mathrm{F}}(\Lambda,\hat{\Lambda}) \geqslant \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}. \qquad \square$$

In the particular case when $\dim(\hat{\mathscr{S}})=1$, the absolute bound in Theorem 5.1 reduces to [6, Theorem 3.1].

One can convert the Frobenius norm bounds in Theorem 5.1 to the two-norm. Let n be the order of A and use the fact that $\|E\|_{\mathrm{F}} \leqslant \sqrt{n}\|E\|_2$ [8, (2.3.7)].

Corollary 5.2. *If Λ and $\hat{\Lambda}$ are diagonal then*

$$\|\sin\Theta\|_2 \leqslant \sqrt{n}\kappa(Y)\kappa(\hat{X})\|E\|_2 / \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} |\lambda-\hat{\lambda}|.$$

If, in addition, A is non-singular, then

$$\|\sin\Theta\|_2 \leqslant \sqrt{n}\kappa(Y)\kappa(\hat{X})\|A^{-1}E\|_2 / \min_{\lambda\in\Lambda,\ \hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|}.$$

The absolute bound above is essentially [17, Theorem 2.2].

The factor $\sqrt{n}$ can be removed if the eigenvalues are more strongly separated. Set

$$\delta_a \equiv \max\{\min_{\lambda\in\Lambda}|\lambda| - \max_{\hat{\lambda}\in\hat{\Lambda}}|\hat{\lambda}|, \min_{\hat{\lambda}\in\hat{\Lambda}}|\hat{\lambda}| - \max_{\lambda\in\Lambda}|\lambda|\}$$

and

$$\delta_r \equiv \max\left\{\frac{\min_{\lambda\in\Lambda}|\lambda| - \max_{\hat{\lambda}\in\hat{\Lambda}}|\hat{\lambda}|}{\min_{\lambda\in\Lambda}|\lambda|}, \frac{\min_{\hat{\lambda}\in\hat{\Lambda}}|\hat{\lambda}| - \max_{\lambda\in\Lambda}|\lambda|}{\min_{\hat{\lambda}\in\hat{\Lambda}}|\hat{\lambda}|}\right\}.$$

In this case (4.2) and (4.3) imply the following two-norm bounds.

Theorem 5.3. *If* $\delta_a > 0$ *then*

$$\|\sin\Theta\|_2 \leqslant \kappa(Y)\kappa(\hat{X})\|E\|_2/\delta_a.$$

If, in addition, A is non-singular, then

$$\|\sin\Theta\|_2 \leqslant \kappa(Y)\kappa(\hat{X})\|A^{-1}E\|_2/\delta_r.$$

When A and $A+E$ are normal, the above absolute bound represents one of Davis and Kahan's $\sin\theta$ Theorems [3, Section 6; 4, Section 2].

Quite a few relative bounds have been derived for invariant eigenspaces of Hermitian matrices and Hermitian perturbations, for instance [1,5,10–13, 15,16,18]. For particular perturbations these bounds can be tighter than the ones presented here because they exploit the Hermitian structure, and many are invariant under congruence transformations and grading. However, the bounds here are the most general because they place no restrictions on the original matrix A or the perturbation E, and they are simple and easy to interpret.

In the context of multiplicative perturbations, where the perturbed matrix is expressed as D_1AD_2, relative two-norm perturbation bounds for invariant subspaces of diagonalizable matrices are derived in [6].

Acknowledgement

I thank the referees for their careful reading of the manuscript and many helpful suggestions.

References

[1] J. Barlow, I. Slapničar, Optimal perturbation bounds for the Hermitian eigenvalue problem, Technical Report 99-001, Department of Computer Science and Engineering, The Pennsylvania State University Park, PA, February 1999.
[2] R. Bhatia, Matrix analysis, Springer, New York, 1997.
[3] C. Davis, W. Kahan, Some new bounds on perturbation of subspaces, Bull. Amer. Math. Soc. 75 (1969) 863–868.
[4] C. Davis, W. Kahan, The rotation of eigenvectors by a perturbation III, SIAM J. Numer. Anal. 7 (1970) 1–46.
[5] S. Eisentat, I. Ipsen, Relative perturbation bounds for eigenspaces and singular vector subspaces, in: Applied Linear Algebra, SIAM, Philadelphia, 1994, pp. 62–65.
[6] S. Eisenstat, I. Ipsen, Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices, BIT 38 (1998) 502–509.
[7] I. Gohberg, R. Lancaster, L. Rodman, Invariant Subspaces of Matrices with Applications, Wiley, New York, 1986.
[8] G. Golub, C. van Loan, Matrix Computations, 3rd ed., The Johns Hopkins University Press, Baltimore, 1996.

[9] P. Lancaster, M. Tismenetsky, The Theory of Matrices, 2nd ed., Academic Press, Orlando, 1985.
[10] R. Li, Relative perturbation theory: (II) eigenspace and singular subspace variations, SIAM J. Matrix Anal. Appl. 20 (1999) 471–492.
[11] R. Li, G. Stewart, A New Relative Perturbation Theorem for Singular Subspaces, Submitted for publication.
[12] R. Mathias, Spectral perturbation bounds for positive definite matrices, SIAM J. Matrix Anal. Appl. 18 (1997) 959–980.
[13] I. Slapničar, K. Veselic, Perturbation of the eigenprojections of a factorised Hermitian matrix, Linear Algebra Appl. 218 (1995) 273–280.
[14] G. Stewart, J. Sun, Matrix Perturbation Theory, Academic Press, San Diego, 1990.
[15] N. Truhar, I. Slapničar, Relative perturbation of invariant subspaces, Math. Commun. 1 (1996) 169–174.
[16] N. Truhar, I. Slapničar, Relative perturbation bound for invariant subspaces of graded indefinite Hermitian matrices, Linear Algebra Appl., submitted.
[17] J. Varah, Computing invariant sunspaces of a general matrix when the eigensystem is poorly conditioned, Math. Comp. 24 (1970) 137–149.
[18] K. Veselič, I. Slapničar, Floating-point perturbations of Hermitian matrices, Linear Algebra Appl. 195 (1993) 81–116.
[19] H. Wimmer, Canonical angles of unitary spaces and perturbations of direct complements, Linear Algebra Appl. 287 (1999) 373–379.