BIT 38 (1998), 001–003.

RELATIVE PERTURBATION RESULTS FOR EIGENVALUES AND EIGENVECTORS OF DIAGONALISABLE MATRICES *

S.C. EISENSTAT[1] and I.C.F. IPSEN[2] †

[1] *Department of Computer Science, Yale University,*
P. O. Box 208285, New Haven, CT 06520-8285, USA.
email: eisenstat-stan@cs.yale.edu

[2] *Department of Mathematics, North Carolina State University,*
P. O. Box 8205, Raleigh, NC 27695-8205, USA.
email: ipsen@math.ncsu.edu
http://www.ncsu.edu/~ipsen/info.html

Abstract.

Let $\hat{\lambda}$ and $\hat{x}$ be a perturbed eigenpair of a diagonalisable matrix A. The problem is to bound the error in $\hat{\lambda}$ and $\hat{x}$. We present one absolute perturbation bound and two relative perturbation bounds.

The absolute perturbation bound is an extension of Davis and Kahan's $\sin\theta$ Theorem from Hermitian to diagonalisable matrices. The two relative perturbation bounds assume that $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a perturbed matrix D_1AD_2, where D_1 and D_2 are non-singular, but D_1AD_2 is not necessarily diagonalisable. We derive a bound on the relative error in $\hat{\lambda}$ and a $\sin\theta$ theorem based on a relative eigenvalue separation. The perturbation bounds contain both the deviation of D_1 and D_2 from similarity and the deviation of D_2 from identity.

AMS subject classification: 15A18, 15A42, 15A60, 65F15, 65F35, 65G99

Key words: eigenvalue, eigenvector, condition number, relative error, diagonalisable matrix, angle between subspaces.

1 Introduction.

The protagonist is a complex, diagonalisable matrix A with eigendecomposition

$$A = X\Lambda X^{-1},$$

where Λ is a diagonal matrix whose diagonal elements λ_i are the eigenvalues of A. The complex number $\hat{\lambda}$, the 'perturbed eigenvalue', approximates an eigenvalue of A; and the unit vector $\hat{x}$, the 'perturbed eigenvector', approximates an eigenvector. The problem is to bound the error in $\hat{\lambda}$ and $\hat{x}$. Our results extend the work in [8, 7] and are stronger and simpler than the results in [10, 11].

First we bound $\sin\theta$, where θ is the angle between $\hat{x}$ and the eigenspace corresponding to those eigenvalues closest to $\hat{\lambda}$ (Theorem 3.1). This is an extension

*Received April 1996. Revised January 1998.

†This work was partially supported by NSF grant CCR-9400921.

of Davis and Kahan's $\sin\theta$ Theorem [5, §6], [6, §2] from Hermitian to diagonalisable matrices. Compared to the Hermitian case, the new bound contains an additional factor κ_2, the condition number of a subset of the left eigenvectors. The bound is exact rather than first-order and holds without any assumptions on the perturbations.

Next we derive two relative perturbation bounds based on a multiplicative perturbation model. We assume that $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a perturbed matrix D_1AD_2, where D_1 and D_2 are non-singular matrices and D_1AD_2 need not be diagonalisable. Our bound on the relative error in $\hat{\lambda}$ contains the deviation of D_1 and D_2 from similarity (Theorem 5.1). It is tight when the perturbation is a similarity transformation $D_1 = D_2^{-1}$ or when $\hat{\lambda} = 0$.

From the absolute $\sin\theta$ theorem described above we derive a relative bound on $\sin\theta$ that consists of two summands: the deviation of D_2 from identity; and the deviation of D_1 and D_2 from similarity, amplified by κ_2 and a relative eigenvalue separation (Theorem 6.2). We conclude that the eigenvectors of diagonalisable matrices are well-conditioned when the perturbation is caused by a similarity transformation; and that the null vectors of diagonalisable matrices are well-conditioned when the perturbation is multiplicative (Corollary 6.3).

Our relative perturbation bounds are no stronger than the traditional absolute perturbation bounds because we use the traditional bounds to derive them.

Notation.

$\|\cdot\|$ denotes the two-norm. If V is a matrix of full column rank and the superscript * denotes the conjugate transpose then $V^\dagger \equiv (V^*V)^{-1}V^*$ is the Moore Penrose-inverse of V and $\kappa(V) \equiv \|V\|\,\|V^\dagger\|$ is the two-norm condition number.

2 Setting the Stage.

To establish a correspondence between perturbed and exact quantities, we partition the eigenvalues of A so that Λ_1 contains all eigenvalues closest to $\hat{\lambda}$ and Λ_2 contains the remaining eigenvalues; i.e.,

$$\Lambda = \begin{pmatrix} \Lambda_1 & \\ & \Lambda_2 \end{pmatrix}$$

where

$$\|\Lambda_1 - \hat{\lambda} I\| = \min_i |\lambda_i - \hat{\lambda}|, \qquad \text{and} \qquad \|\Lambda_1 - \hat{\lambda} I\| < 1/\|(\Lambda_2 - \hat{\lambda} I)^{-1}\|.$$

Thus $\|\Lambda_1 - \hat{\lambda} I\|$ represents the absolute error in $\hat{\lambda}$. The 'absolute gap'

$$\text{absgap} \equiv 1/\|(\Lambda_2 - \hat{\lambda} I)^{-1}\| = \min_i |(\Lambda_2)_{ii} - \hat{\lambda}|$$

is the absolute separation of $\hat{\lambda}$ from the eigenvalues of Λ_2. It approximates the absolute separation of the desired eigenvalues Λ_1 from the remaining eigenvalues Λ_2.

We partition the eigenvectors conformally with Λ,

$$X = \begin{pmatrix} X_1 & X_2 \end{pmatrix} \qquad \text{and} \qquad X^{-1} = \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix}.$$

Then $\hat{x}$ can be regarded as approximating a unit vector in range(X_1). Since Λ_1 consists of all eigenvalues on a circle of minimal radius around $\hat{\lambda}$, range(X_1) is in general associated with several distinct, possibly multiple eigenvalues.

To determine how well the unit vector $\hat{x}$ approximates a vector in the right eigenspace range(X_1), we measure the largest principal angle $0 \leq \theta \leq \pi/2$ between range$(\hat{x})$ and range(X_1) [9, §12.4.3]. Define the orthogonal projector onto range$(X_1)^\perp$, the orthogonal complement of range(X_1),

$$P \equiv I - (X_1^\dagger)^* X_1^*.$$

Then $\sin\theta$ is the orthogonal projection of $\hat{x}$ onto range$(X_1)^\perp$ [5, §6], [6, p. 10], [16, (2.3)],

$$\sin\theta = \|P\hat{x}\|. \tag{2.1}$$

To derive our $\sin\theta$ theorem, we express P in terms of a basis Y_2 for the left eigenspace range(Y_2).

Lemma 2.1. *The orthogonal projector P satisfies*

$$P = (Y_2^\dagger)^* Y_2^*.$$

Proof. $X^{-1}X = I$ implies range$(Y_2) = $ range$(X_1)^\perp$. Thus

$$P = I - (X_1^\dagger)^* X_1^* = (Y_2^\dagger)^* Y_2^*.$$

□

3 A $\sin\theta$ Theorem for Diagonalisable Matrices.

The traditional perturbation bound for the eigenvectors of a Hermitian matrix is Davis and Kahan's $\sin\theta$ Theorem [5, §6], [6, §2],

$$\sin\theta \leq \|r\|/\text{absgap},$$

where $r \equiv (A - \hat{\lambda}I)\hat{x}$ is the residual. That is, for a Hermitian matrix the condition number of the perturbed eigenvector is inversely proportional to the separation between $\hat{\lambda}$ and the remaining eigenvalues Λ_2. We extend this bound to diagonalisable matrices.

Theorem 3.1.

$$\sin\theta \leq \kappa_2\, \|r\|/\text{absgap},$$

where $\kappa_2 \equiv \kappa(Y_2) \geq 1$.

Proof. Since our goal is a residual bound for the angle, we write

$$\begin{aligned} r &= (A - \hat{\lambda}I)\hat{x} = X \begin{pmatrix} \Lambda_1 - \hat{\lambda}I & \\ & \Lambda_2 - \hat{\lambda}I \end{pmatrix} X^{-1}\hat{x} \\ &= X \begin{pmatrix} \Lambda_1 - \hat{\lambda}I & \\ & \Lambda_2 - \hat{\lambda}I \end{pmatrix} \begin{pmatrix} \hat{c} \\ \hat{s} \end{pmatrix}, \end{aligned}$$

where $\hat{c} \equiv Y_1^* \hat{x}$ and $\hat{s} \equiv Y_2^* \hat{x}$. Multiplying both sides by $(\Lambda_2 - \hat{\lambda} I)^{-1} Y_2^*$ gives

$$\hat{s} = (\Lambda_2 - \hat{\lambda} I)^{-1}\, Y_2^* r.$$

From Lemma 2.1 follows

$$P\hat{x} = (Y_2^\dagger)^* Y_2^* \hat{x} = (Y_2^\dagger)^* \hat{s} = (Y_2^\dagger)^* (\Lambda_2 - \hat{\lambda} I)^{-1}\, Y_2^* r.$$

At last (2.1) implies

$$\sin\theta = \|P\hat{x}\| \le \|(Y_2^\dagger)^*\|\, \|(\Lambda_2 - \hat{\lambda} I)^{-1}\|\, \|Y_2^*\|\, \|r\|.$$

□

The worse conditioned the left eigenvectors Y_2 and hence X^{-1}, the larger κ_2 is likely to be. If Y_2 has orthonormal columns, which occurs when A is normal, then $\kappa_2 = 1$ and Theorem 3.1 reduces to Davis and Kahan's $\sin\theta$ Theorem. Although Davis and Kahan considered only Hermitian matrices, the ideas for this extension to diagonalisable matrices are already present in [6, Theorem 6.1] (essentially their factor $1/\epsilon$ corresponds to our κ_2).

Related bounds on the angle between a vector and a subspace appear in [2], [3, (11)], [4, (2.11)], [14, Theorem 4.1], [15, Theorem 4.11], [16, Theorem 2.2], [13, Corollary 1], and [18, (3.54.14)].

4 Multiplicative Perturbations.

Assume that the perturbed quantities $\hat{\lambda}$ and $\hat{x}$ are an exact eigenvalue and eigenvector of a matrix $D_1 A D_2$, where D_1 and D_2 are non-singular matrices; i.e.,

$$(D_1 A D_2)\, \hat{x} = \hat{\lambda} \hat{x}.$$

The matrix $D_1 A D_2$ is not required to be diagonalisable. When $D_1 = D_2^{-1}$ then $D_1 A D_2$ is a similarity transformation of A, which means that A and $D_1 A D_2$ have the same eigenvalues. When $D_1 = D_2^*$ then $D_1 A D_2$ is a congruence transformation of A, which means A and $D_1 A D_2$ have the same inertia when A is Hermitian.

The derivation of relative perturbation bounds for such *multiplicative perturbations* is based on the following idea. From $(D_1 A D_2)\hat{x} = \hat{\lambda}\hat{x}$ follows

$$A(D_2\hat{x}) = \hat{\lambda} D_1^{-1} D_2^{-1} (D_2 \hat{x}).$$

Setting $z \equiv D_2\hat{x}/\|D_2\hat{x}\|$ gives

$$Az = \hat{\lambda} D_1^{-1} D_2^{-1} z.$$

Thus the residual for $\hat{\lambda}$ and z is

$$f \equiv Az - \hat{\lambda} z = \hat{\lambda}\,(D_1^{-1} D_2^{-1} - I) z = \hat{\lambda}\,(D_1^{-1} - D_2)\hat{x}/\|D_2\hat{x}\|.$$

The upper bounds

$$\|f\| \le |\hat{\lambda}|\, \|I - D_1^{-1} D_2^{-1}\| \qquad \text{and} \qquad \|f\| \le |\hat{\lambda}|\, \|D_1^{-1} - D_2\|/\|D_2\hat{x}\|$$

consist of two parts: the factor $|\hat{\lambda}|$, which is responsible for the perturbation bound being relative, and the deviation of D_1 and D_2 from similarity. The factor $\|D_1^{-1} - D_2\|$ is an absolute deviation from similarity, while the factor $\|I - D_1^{-1}D_2^{-1}\|$ constitutes a relative deviation because

$$I - D_1^{-1}D_2^{-1} = (D_2 - D_1^{-1})D_2^{-1}$$

is a difference relative to D_2.

5 Eigenvalues of Multiplicative Perturbations.

We bound the relative error in $\hat{\lambda}$ in terms of the relative deviation from similarity of the matrices D_1 and D_2.

THEOREM 5.1. *The perturbed eigenvalue $\hat{\lambda}$ satisfies*

$$\|\Lambda_1 - \hat{\lambda}I\| \leq |\hat{\lambda}|\, \kappa(X)\, \|I - D_1^{-1}D_2^{-1}\|.$$

PROOF. Applying the Bauer-Fike Theorem with residual bound [1, Theorem IIIa], [12, §3.2.7] to the residual f for $\hat{\lambda}$ and z gives

$$\|\Lambda_1 - \hat{\lambda}I\| = \min_i |\lambda_i - \hat{\lambda}| \leq \kappa(X)\, \|f\|,$$

and bound $\|f\|$ as above. □

In the case of a similarity transformation $D_1 = D_2^{-1}$, the eigenvalues are preserved and the bound above is zero, which is tight. When $\hat{\lambda} = 0$, A itself must be singular and the bound is zero, which is again tight.

Multiplicative bounds for Hermitian matrices appear in [8, Theorem 2.1] and [10, Theorem 3.1], and for diagonalisable matrices in [17, Theorem 3.17].

6 Eigenvectors of Multiplicative Perturbations.

We use the absolute perturbation bound in Theorem 3.1 to derive bounds on $\sin\theta$ in terms of a 'relative gap', i.e., a relative separation of $\hat{\lambda}$ from the remaining eigenvalues Λ_2. We define the relative gap via its inverse,

$$1/\text{relgap} \equiv |\hat{\lambda}|/\text{absgap} \geq 0,$$

to ensure that the definition is valid even when $\hat{\lambda} = 0$.

Since multiplicative perturbations lead to a residual for $z = D_2\hat{x}/\|D_2\hat{x}\|$ rather than for $\hat{x}$, a straightforward application of Theorem 3.1 yields a bound on the angle ϕ between z and $\text{range}(X_1)$ rather than the desired bound on θ. The following lemma shows how to adjust this bound.

LEMMA 6.1. *Let $0 \leq \phi \leq \pi/2$ be the largest principal angle between z and* $\text{range}(X_1)$. *Then*

$$\sin\theta \leq \|D_2\hat{x}\|\, \sin\phi + \|D_2 - I\|,$$

and

$$\sin\theta \leq \sin\phi + \|D_2 - I\|.$$

PROOF. To derive the first bound write

$$P\hat{x} = PD_2\hat{x} \;+\; P(I-D_2)\hat{x} = \|D_2\hat{x}\|\, Pz \;+\; P(I-D_2)\hat{x},$$

and note that

$$\sin\theta = \|P\hat{x}\| \qquad \text{and} \qquad \sin\phi = \|Pz\|.$$

The second bound follows from the first if $\|D_2\hat{x}\| \leq 1$, so assume that $\|D_2\hat{x}\| > 1$. Then

$$\sin\theta = \|P\hat{x}\| \leq \|Pz\| \;+\; \|P(z-\hat{x})\| \leq \sin\phi \;+\; \|z-\hat{x}\|.$$

Note that

$$\begin{aligned} \|(D_2-I)\hat{x}\|^2 &= \|(D_2\hat{x}-z)+(z-\hat{x})\|^2 \\ &= \|z-\hat{x}\|^2 + \|D_2\hat{x}-z\|^2 + 2\,\mathrm{Re}\,(D_2\hat{x}-z)^*(z-\hat{x}). \end{aligned}$$

Since $\hat{x}$ and z are unit vectors,

$$2\,\mathrm{Re}\,(D_2\hat{x}-z)^*(z-\hat{x}) = 2\,(\|D_2\hat{x}\|-1)\,\mathrm{Re}\,(1-z^*\hat{x}) \geq 0.$$

Thus

$$\|(D_2-I)\hat{x}\|^2 \geq \|z-\hat{x}\|^2,$$

and the result follows from the inequality $\|(D_2-I)\hat{x}\| \leq \|D_2-I\|$. $\square$

In general any bound on $\sin\theta$ derived from Lemma 6.1 must consist of two summands. The theorem below bounds the first summand by the (absolute or relative) deviation of D_1 and D_2 from a similarity transformation, amplified by κ_2 and by the relative eigenvalue separation; and bounds the second summand by the (absolute and relative) deviation of the similarity transformation from the identity.

THEOREM 6.2. *Let*

$$\alpha_1 \equiv \|D_1^{-1} - D_2\|, \qquad \alpha_2 \equiv \|I - D_1^{-1}D_2^{-1}\|,$$

and $\kappa_2 \equiv \kappa(Y_2)$ *be defined as in Theorem 3.1. Then*

$$\sin\theta \leq \kappa_2\,\frac{\min\{\alpha_1,\,\alpha_2\}}{\mathrm{relgap}} \;+\; \|I-D_2\|,$$

where $\alpha_1 \leq \alpha_2$ *if* $\|D_2^{-1}\| \leq 1$ *and* $\alpha_2 \leq \alpha_1$ *if* $\|D_2\| \leq 1$.

PROOF. Applying Theorem 3.1 to the residual f for $\hat{\lambda}$ and z gives

$$\sin\phi \leq \kappa_2\,\|f\|/\mathrm{absgap}.$$

To derive the bound with $\alpha_1 \leq \alpha_2$, bound $\|f\|$ by

$$\|f\| \leq |\hat{\lambda}|\,\|D_1^{-1} - D_2\|/\|D_2\hat{x}\| = |\hat{\lambda}|\,\alpha_1/\|D_2\hat{x}\|$$

and use the first bound in Lemma 6.1. To derive the bound with $\alpha_2 \leq \alpha_1$, bound $\|f\|$ by

$$\|f\| \leq |\hat{\lambda}|\,\|I - D_1^{-1}D_2^{-1}\| = |\hat{\lambda}|\,\alpha_2$$

and use the second bound in Lemma 6.1. Finally, $\|D_2\| \leq 1$ implies

$$\alpha_1 = \|D_2 - D_1^{-1}\| \leq \|I - D_1^{-1}D_2^{-1}\| \, \|D_2\| \leq \alpha_2,$$

while $\|D_2^{-1}\| \leq 1$ implies

$$\alpha_2 = \|I - D_1^{-1}D_2^{-1}\| \leq \|D_2 - D_1^{-1}\| \, \|D_2^{-1}\| \leq \alpha_1.$$

□

In some cases the first summand in the bound of Theorem 6.2 can be omitted.

COROLLARY 6.3. *If $D_1 = D_2^{-1}$ or $\hat{\lambda} = 0$, then*

$$\sin\theta = \|P(I - D_2)\hat{x}\| \leq \|I - D_2\|.$$

PROOF. Consider the case $D_1 = D_2^{-1}$. Then $D_1AD_2\hat{x} = \hat{\lambda}\hat{x}$ implies $AD_2\hat{x} = \hat{\lambda}D_2\hat{x}$, i.e., $\hat{\lambda}$ and $D_2\hat{x}$ are an exact eigenpair of A. According to the partitioning defined in §2, we must have $\Lambda_1 = \hat{\lambda}I$, so $D_2\hat{x}$ is in range(X_1) = range$(Y_2)^{\perp}$. Thus $PD_2\hat{x} = 0$ and

$$\sin\theta = \|P\hat{x}\| = \|P(D_2\hat{x} - \hat{x})\| \leq \|I - D_2\|.$$

Next consider the case $\hat{\lambda} = 0$. Then $D_1AD_2\hat{x} = 0 \cdot \hat{x}$ implies $D_2^{-1}AD_2\hat{x} = 0 \cdot \hat{x}$, since D_1 and D_2 are non-singular. Hence $\hat{\lambda}$ and $\hat{x}$ are an exact eigenpair of a similarity transformation of A, and we are back to the first case. □

For similarity transformations $D_1 = D_2^{-1}$, the bounds in Corollary 6.3 are essentially tight. The error angle is bounded by the relative deviation of D_2 from identity, without any amplification by κ_2 or by a relative gap. Therefore we can say that the eigenvectors of diagonalisable matrices are well-conditioned when the perturbation is caused by a similarity transformation.

Moreover, when $\hat{\lambda} = 0$, the bounds in Corollary 6.3 are again essentially tight. This means the null vectors of diagonalisable matrices are also well-conditioned when the perturbation is multiplicative.

Multiplicative bounds for Hermitian matrices are derived in [7, Theorems 2.1, 3.1] and [11, Theorem 3.1 and 3.2].

Acknowledgement.

We thank Axel Ruhe and Krešimir Veselić for helpful discussions, and Ren-Cang Li and the referees for suggestions that improved the paper.

REFERENCES

1. F. BAUER AND C. FIKE, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–41.
2. R. BHATIA, C. DAVIS, AND A. MCINTOSH, *Perturbation of spectral subspaces and solution of linear operator equations*, Linear Algebra Appl., 52/53 (1983), pp. 45–57.
3. Z. BOHTE, *A posteriori error bounds for eigensystems of matrices*, in Numerical Methods and Approximation Theory III, University of Niš, Yugoslavia, 1988, pp. 45–67.

4. ———, *Computable error bounds for approximate eigensystems of matrices*, in VII Conference on Applied Mathematics, R. Scitovski, ed., University of Osijek, Croatia, 1990, pp. 29–37.
5. C. Davis and W. Kahan, *Some new bounds on perturbation of subspaces*, Bull. Amer. Math. Soc., 75 (1969), pp. 863–8.
6. C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation, III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.
7. S. Eisenstat and I. Ipsen, *Relative perturbation bounds for eigenspaces and singular vector subspaces*, in Applied Linear Algebra, SIAM, Philadelphia, 1994, pp. 62–5.
8. ———, *Relative perturbation techniques for singular value problems*, SIAM J. Numer. Anal., 32 (1995), pp. 1972–88.
9. G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins Press, Baltimore, second ed., 1989.
10. R. Li, *Relative perturbation theory: (I) eigenvalue variations*, LAPACK working note 84, Computer Science Department, University of Tennessee, Knoxville, 1994. Revised May 1997.
11. ———, *Relative perturbation theory: (II) eigenspace variations*, LAPACK working note 85, Computer Science Department, University of Tennessee, Knoxville, 1994. Revised May 1997.
12. J. Ortega, *Numerical Analysis, A Second Course*, Academic Press, New York, 1972.
13. A. Ruhe, *Perturbation bounds for means of eigenvalues and invariant subspaces*, BIT, 10 (1970), pp. 343–54.
14. G. Stewart, *Error bounds for approximate invariant subspaces, of closed linear operators*, SIAM J. Numer. Anal., 8 (1971), pp. 796–808.
15. ———, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review, 15 (1973), pp. 727–64.
16. J. Varah, *Computing invariant subspaces of a general matrix when the eigensystem is poorly conditioned*, Math. Comp., 24 (1970), pp. 137–49.
17. K. Veselić and I. Slapničar, *Floating-point perturbations of Hermitian matrices*, Linear Algebra Appl., 195 (1993), pp. 81–116.
18. J. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford University Press, 1965.