A NOTE ON UNIFYING ABSOLUTE AND RELATIVE PERTURBATION BOUNDS

ILSE C. F. IPSEN*

Abstract. Perturbation bounds for invariant subspaces and eigenvalues of complex matrices are presented that lead to absolute as well as a large class of relative bounds. In particular it is shown that absolute bounds (such as those by Davis-Kahan, Bauer-Fike and Hoffman-Wielandt) and some relative bounds are special cases of 'universal' bounds. As a consequence, we obtain a new relative bound for subspaces of normal matrices, which contains a deviation of the matrix from (positive-) definiteness. We also investigate how row scaling affects eigenvalues and their sensitivity to perturbations, and we illustrate how the departure from normality can affect the condition number (with respect to inversion) of the scaled eigenvectors.

Key words. invariant subspace, eigenvalue, absolute bound, relative bound, subspace angle, departure from normality

AMS subject classification. 15A18, 15A42, 65F15, 65F35

1. Introduction.

Traditionally perturbation bounds for eigenvalues bound the absolute error in the perturbed eigenvalue. In contrast, the newer relative perturbation bounds bound a measure of relative error [9]. Similarly, absolute bounds for invariant subspaces bound the angle between original and perturbed subspace in terms of an absolute eigenvalue difference, while relative bounds contain a relative eigenvalue difference [11].

Usually one is interested in the differences between absolute and relative bounds. For instance, under what circumstances is a relative bound tighter than an absolute bound? Here we focus instead on the similarities, and in particular on the 'heritage' of the bounds. For general purpose perturbation bounds, i.e. those that do not exploit structure such as symmetry or grading of the matrix, we exhibit 'universal' bounds that lead to absolute as well as a large class of relative bounds.

In §2 notation and facts for invariant subspaces are established. The universal subspace bound is proved in §3, and §4 presents existing bounds that are special cases of the universal bound. In §5 we derive a universal eigenvalue bound for diagonalizable matrices in the two-norm, and in §6 in the Frobenius norm. The effect of row scaling on eigenvalues and their perturbation bounds is investigated in §7.

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 the 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. Invariant Subspaces.

Let A be a complex square matrix. A subspace $\mathcal{S}$ is an *invariant subspace* of A if $Ax \in \mathcal{S}$ for every $x \in \mathcal{S}$ [6, §1.1], [18, §I.3.4]. Let the perturbed matrix $A+E$ have an invariant subspace $\hat{\mathcal{S}}$. As in [10, §2] set

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

where P is the orthogonal projector onto $\mathcal{S}^\perp$, and $\hat{P}$ is the orthogonal projector onto $\hat{\mathcal{S}}$. When $\dim(\mathcal{S}) = \dim(\hat{\mathcal{S}})$, the singular values of $P\hat{P}$ are the sines of the principal

*Center for Research In Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205, USA (`ipsen@math.ncsu.edu`, http://www4.ncsu.edu/~ipsen/). This research was supported in part by NSF grant DMS-9714811.

angles between $\mathcal{S}$ and $\hat{\mathcal{S}}$ [7, §12.4.3], [18, Theorem I.5.5]. The goal is to bound $\|\sin\Theta\|$, where $\|\cdot\|$ is the two-norm or the Frobenius norm.

We make frequent use of the following fact [6, (1.5.5)], [12, Theorem 5.8.4],

$$PA = PAP, \qquad (A+E)\hat{P} = \hat{P}(A+E)\hat{P}. \tag{2.1}$$

The first equality holds because $\mathcal{S}^\perp$ is an invariant subspace of A^* [18, Theorem V.1.1].

3. A Universal Subspace Bound. For general complex matrices, a basis-free bound for $\|\sin\Theta\|$ is derived.

Let A and $A+E$ be complex, non-singular matrices. Define a separation between A and $A+E$, with regard to the subspaces $\mathcal{S}$ and $\hat{\mathcal{S}}$ as

$$\mathrm{sep}_{k,l} \equiv \min_{\|Z\|=1, PZ\hat{P}=Z} \|P\left(A^{1-k}Z(A+E)^{-l} - A^{-k}Z(A+E)^{1-l}\right)\hat{P}\|,$$

where k and l are real numbers, and the powers are to be interpreted according to [8, Definition 6.2.4]. In all results to follow we assume $\mathrm{sep}_{k,l} > 0$. The likelyhood of this happening is discussed in Remark 3.1 below.

THEOREM 3.1. *If A and $A+E$ are non-singular and if $\mathrm{sep}_{k,l} > 0$ then*

$$\|\sin\Theta\| \leq \|A^{-k}E(A+E)^{-l}\|/\mathrm{sep}_{k,l}.$$

Proof. From $E = (A+E) - A$ follows

$$A^{-k}E(A+E)^{-l} = A^{-k}(A+E)^{1-l} - A^{1-k}(A+E)^{-l},$$

and (2.1) implies

$$PA^{-k}E(A+E)^{-l}\hat{P} = PA^{-k}\ \sin\Theta\ (A+E)^{1-l}\hat{P} - PA^{1-k}\ \sin\Theta\ (A+E)^{-l}\hat{P}.$$

Hence

$$\|A^{-k}E(A+E)^{-l}\| \geq \|PA^{-k}E(A+E)^{-l}\hat{P}\| \geq \mathrm{sep}_{k,l}\|\sin\Theta\|,$$

since $\sin\Theta = P\sin\Theta\hat{P}$. ◻

The next lemma expresses the separation in terms of eigenvalues when the matrices are diagonalizable. Let A and $A+E$ be diagonalizable. Then there are matrices Y and $\hat{X}$ with linearly independent columns so that $\mathcal{S}^\perp = \mathrm{range}(Y)$, $\hat{\mathcal{S}} = \mathrm{range}(\hat{X})$, and

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

where Λ and $\hat{\Lambda}$ are diagonal. Denote the two-norm condition numbers of these bases by, respectively,

$$\kappa \equiv \|Y\|_2\,\|Y^\dagger\|_2, \qquad \hat{\kappa} \equiv \|\hat{X}\|_2\,\|\hat{X}^\dagger\|_2.$$

In the case of diagonalizable matrices the Frobenius-norm separation can be bounded in terms of an eigenvalue separation.

LEMMA 3.2. *If A and $A+E$ are diagonalizable and non-singular then in the Frobenius norm*

$$\mathrm{sep}_{k,l} \geq \frac{1}{\kappa\hat{\kappa}} \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|^k\,|\hat{\lambda}|^l},$$

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

Proof. The proof is similar to those for the Sylvester equations in [14, §2.4].

Let $Y = QR$ and $\hat{X} = \hat{Q}\hat{R}$ be QR decompositions where Q and $\hat{Q}$ have orthonormal columns, and R and $\hat{R}$ are non-singular. Let Z_0 be a matrix that attains the minimum in $\mathrm{sep}_{k,l}$. From $P = QQ^*$ and $\hat{P} = \hat{Q}\hat{Q}^*$ follows in the Frobenius norm

$$\begin{aligned}
\mathrm{sep}_{k,l} &= \|R^{-*}\left(\Lambda^{1-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{-l} - \Lambda^{-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{1-l}\right)\hat{R}^{-1}\|_F \\
&\geq \frac{\|\Lambda^{1-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{-l} - \Lambda^{-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{1-l}\|_F}{\|R\|_2\,\|\hat{R}\|_2} \\
&\geq \frac{1}{\kappa\hat{\kappa}} \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|^k\,|\hat{\lambda}|^l},
\end{aligned}$$

where the last inequality is obtained by considering individual elements of the matrix inside the norm, summing them up according to $\|M\|_F^2 = \sum_{i,j}|M_{ij}|^2$ and using the fact

$$\|Y^*Z_0\hat{X}\|_F \geq \frac{\|PZ_0\hat{P}\|_F}{\|R^{-1}\|_2\,\|\hat{R}^{-1}\|_2} = \frac{1}{\|R^{-1}\|_2\,\|\hat{R}^{-1}\|_2}.$$

□

Consequently, the bound in Theorem 3.1 can be expressed in terms of an eigenvalue separation when the matrices are diagonalizable:

COROLLARY 3.3. *If A and $A+E$ are diagonalizable then*

$$\|\sin\Theta\|_F \leq \kappa\hat{\kappa}\,\|A^{-k}E(A+E)^{-l}\|_F/\min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\lambda|^k\,|\hat{\lambda}|^l},$$

provided the denominator is positive. Here the minimum ranges over all diagonal elements λ of Λ and all diagonal elements $\hat{\lambda}$ of $\hat{\Lambda}$.

REMARK 3.1. *When A and $A+E$ are normal it is unlikely that $\mathrm{sep}_{k,l} = 0$, unless $\sin\Theta = 0$ or the eigenvalue separation is zero.*

The reason is as follows. Consider the proof of Lemma 3.2, when Y and $\hat{X}$ have orthogonal columns, and set $W \equiv Y^*\hat{X}$. Then for $\mathrm{sep}_{k,l}$ in the Frobenius norm,

$$\begin{aligned}
\mathrm{sep}_{k,l}^2 &= \|\Lambda^{1-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{-l} - \Lambda^{-k}\,Y^*Z_0\hat{X}\,\hat{\Lambda}^{1-l}\|_F^2 \\
&\leq \|\Lambda^{1-k}\,W\,\hat{\Lambda}^{-l} - \Lambda^{-k}\,W\,\hat{\Lambda}^{1-l}\|_F^2 = \sum_{i,j}\left|\frac{(\Lambda_{ii}-\hat{\Lambda}_{jj})w_{ij}}{\Lambda_{ii}^k\hat{\Lambda}_{jj}^l}\right|^2,
\end{aligned}$$

where Z_0 is replaced by $\sin\Theta = YY^*\hat{X}\hat{X}^* = YW\hat{X}^*$ in the upper bound. The sum is zero if $\Lambda_{ii} = \hat{\Lambda}_{jj}$ for all i and j, or if $W = 0$.

4. Existing Subspace Bounds. We show that specific values for k and l in Theorem 3.1 and Corollary 3.3 lead to existing bounds. We also derive a new relative bound for normal matrices that reduces to an existing bound in the special case of Hermitian positive-definite matrices.

Let A and $A+E$ be complex square matrices.

Case $k = l = 0$. Theorem 3.1 is identical to the absolute bound [10, Theorem 3.1],

$$\|\sin\Theta\| \leq \|E\|/\mathrm{sep}_{0,0},$$

where

$$\mathrm{sep}_{0,0} = \min_{\|Z\|=1, PZ\hat{P}=Z} \|PAZ - Z(A+E)\hat{P}\|,$$

since (2.1) and $Z = PZ\hat{P}$ imply $PZ(A+E)\hat{P} = Z(A+E)\hat{P}$.

When A and $A+E$ are diagonalizable, Theorem 3.1 implies the Frobenius norm bound [10, Theorem 5.1], [11, Theorem 3.4]

$$\|\sin\Theta\|_F \leq \kappa\hat{\kappa}\, \|E\|_F / \min_{\lambda\in\Lambda, \hat{\lambda}\in\hat{\Lambda}} |\lambda - \hat{\lambda}|.$$

When A and $A+E$ are normal, one obtains one of Davis and Kahan's $\sin\Theta$ Theorems [2, §6], [3, §2],

$$\|\sin\Theta\|_F \leq \|E\|_F / \min_{\lambda\in\Lambda, \hat{\lambda}\in\hat{\Lambda}} |\lambda - \hat{\lambda}|.$$

Case $k = 1$, $l = 0$. Theorem 3.1 is identical to the relative bound [10, Theorem 3.2],

$$\|\sin\Theta\| \leq \|A^{-1}E\|/\mathrm{sep}_{1,0},$$

where

$$\mathrm{sep}_{1,0} = \min_{\|Z\|=1, PZ\hat{P}=Z} \|PA^{-1}\left(PAZ - Z(A+E)\hat{P}\right)\|,$$

because (2.1) and $PZ\hat{P} = Z$ imply $PZ\hat{P} = PA^{-1}PAZ$.

When A and $A+E$ are diagonalizable Theorem 3.1 implies the Frobenius norm bound [10, Theorem 5.1], [11, Theorem 3.4]

$$\|\sin\Theta\|_F \leq \kappa\hat{\kappa}\, \|A^{-1}E\|_F / \min_{\lambda\in\Lambda, \hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{|\lambda|}.$$

Case $k = l = 1/2$. Theorem 3.1 reduces to the relative bound

$$\|\sin\Theta\| \leq \|A^{-1/2}E(A+E)^{-1/2}\|/\mathrm{sep}_{\frac{1}{2},\frac{1}{2}},$$

where

$$\mathrm{sep}_{\frac{1}{2},\frac{1}{2}} = \min_{\|Z\|=1, PZ\hat{P}=Z} \|PA^{-1/2}\left(PAZ - Z(A+E)\hat{P}\right)(A+E)^{-1/2}\hat{P}\|.$$

When A and $A+E$ are diagonalizable Theorem 3.1 implies

$$\|\sin\Theta\|_F \leq \kappa\hat{\kappa}\, \|A^{-1/2}E(A+E)^{-1/2}\|_F / \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{|\lambda\,\hat{\lambda}|}}.$$

When A and $A+E$ are also normal then Theorem 3.1 implies the following relative Frobenius norm bound, which contains a quantity δ that can be interpreted as a deviation of $A+E$ from definiteness.

THEOREM 4.1. *If A and $A+E$ are normal and non-singular,* $\mathrm{sep}_{\frac{1}{2},\frac{1}{2}} > 0$*, and*

$$\eta_2 \equiv \|A^{-1/2}EA^{-1/2}\|_2 < 1$$

then

$$\|\sin\Theta\|_F \leq \delta\, \frac{\eta_F}{\sqrt{1-\eta_2}} \,/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{|\lambda\hat{\lambda}|}},$$

where

$$\eta_F \equiv \|A^{-1/2}EA^{-1/2}\|_F, \qquad \delta \equiv \|A^{1/2}\hat{U}A^{-1/2}\|_2^{1/2},$$

and $\hat{U}$ is a unitary polar factor of $A+E$.

Proof. Lemma 3.2 implies

$$\mathrm{sep}_{\frac{1}{2},\frac{1}{2}} \geq \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{|\lambda\hat{\lambda}|}}.$$

For the remaining factor in the bound we show

$$\|A^{-1/2}E(A+E)^{-1/2}\|_F \leq \delta\, \eta_F/\sqrt{1-\eta_2},$$

similar to [11, Theorem 3.6]. Start with

$$\|A^{-1/2}E(A+E)^{-1/2}\|_F \leq \eta_F\, \|A^{1/2}(A+E)^{-1/2}\|_2.$$

Let $A = UH$ and $A+E = \hat{U}\hat{H}$ be polar factorizations, where U and $\hat{U}$ are unitary, and H and $\hat{H}$ Hermitian positive-definite. We use the fact that polar factors of normal matrices commute [4, Lemma 3.2] and

$$A^{1/2}(A^{1/2})^* = H = (A^{1/2})^*A^{1/2}$$

for any normal, nonsingular matrix A. If $\lambda_i(A)$ denotes an eigenvalue of A then

$$\begin{aligned}
\|A^{1/2}(A+E)^{-1/2}\|_2^2 &= \|A^{1/2}\hat{H}^{-1}(A^{1/2})^*\|_2 = \max_i |\lambda_i(\hat{H}^{-1}H)| \\
&= \max_i |\lambda_i(\hat{U}(A+E)^{-1}AU^*)| \\
&= \max_i |\lambda_i(\hat{U}A^{-1/2}(I+A^{-1/2}EA^{-1/2})^{-1}A^{1/2}U^*)| \\
&= \max_i |\lambda_i((I+A^{-1/2}EA^{-1/2})^{-1}A^{1/2}U^*\hat{U}A^{-1/2})| \\
&\leq \|(I-A^{-1/2}EA^{-1/2})^{-1}\|_2\, \|A^{1/2}U^*\hat{U}A^{-1/2}\|_2 \\
&\leq \frac{1}{1-\eta_2}\, \|A^{1/2}\hat{U}A^{-1/2}\|_2,
\end{aligned}$$

where the last inequality follows from [7, Lemma 2.3.3] and the fact that $A^{1/2}$ and U^* commute. □

When A is Hermitian positive-definite, $\eta_2 < 1$, and when $A+E$ is positive-definite $\delta = 1$ (because $\hat{U} = I$). Thus, when A and $A+E$ are Hermitian positive-definite, Theorem 4.1 implies the relative Frobenius norm bound [15, Theorem 1], [14, Theorem 3.3] (see also [16, Theorem 1]),

$$\|\sin\Theta\|_F \leq \frac{\eta_F}{\sqrt{1-\eta_2}} \Big/ \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{\sqrt{\lambda\hat{\lambda}}}.$$

For general, Hermitian indefinite matrices A and $A+E$ it is not easy to compare the bound in Theorem 4.1 with existing bounds, such as the ones in [17, §3.2], because they are expressed in terms of hyperbolic eigenvector matrices and different relative gaps. However, the bound in [17, Corollary 2], for instance, contains an amplification factor $\kappa(A)^{1/2}$, while our amplification factor is only bounded by

$$\delta \leq \kappa(A)^{1/4}.$$

In this sense, we expect Theorem 4.1 to be no worse than the bounds in [17]. In particular, $\delta \approx 1$ for well-conditioned matrices.

Case $k=0,\ l=1$. Now the perturbed, instead of the true eigenvalue is in the denominator of the separation

$$\|\sin\Theta\| \leq \|E(A+E)^{-1}\|/\mathrm{sep}_{0,1},$$

where

$$\mathrm{sep}_{0,1} = \min_{\|Z\|=1, PZ\hat{P}=Z} \|\left(PAZ - Z(A+E)\hat{P}\right)(A+E)^{-1}\hat{P}\|.$$

When A and $A+E$ are diagonalizable Theorem 3.1 implies

$$\|\sin\Theta\|_F \leq \kappa\hat{\kappa}\,\|E(A+E)^{-1}\|_F / \min_{\lambda\in\Lambda,\hat{\lambda}\in\hat{\Lambda}} \frac{|\lambda-\hat{\lambda}|}{|\hat{\lambda}|}.$$

5. A Universal Eigenvalue Bound in the Two-Norm. We bound, in the two-norm, the distance of a single perturbed eigenvalue $\hat{\lambda}$ to the eigenvalues of a diagonalizable matrix A.

Let A be a complex, non-singular, diagonalizable matrix, and $A+E$ a complex non-singular matrix with eigenvalue $\hat{\lambda}$. Let $A = X\Lambda X^{-1}$ be an eigenvalue decomposition of A, where

$$\Lambda = \begin{pmatrix} \ddots & & \\ & \lambda_i & \\ & & \ddots \end{pmatrix},$$

is a diagonal matrix whose diagonal elements are the eigenvalues λ_i of A, and

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

where $\hat{x}$ is a non-zero vector. Let

$$\kappa \equiv \|X\|_2\,\|X^{-1}\|_2$$

be the two-norm condition number with respect to inversion of the eigenvector matrix X.

THEOREM 5.1. *If A is diagonalizable then*

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|^k \, |\hat{\lambda}|^l} \leq \kappa \, \|A^{-k} E (A+E)^{-l}\|_2.$$

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

$$\begin{aligned}\hat{x} &= -(A - \hat{\lambda} I)^{-1} E \hat{x} = -(A - \hat{\lambda} I)^{-1} A^k \, A^{-k} E (A+E)^{-l} \, (A+E)^l \hat{x} \\ &= -(\hat{\lambda}^{-l} A^{1-k} - \hat{\lambda}^{1-l} A^{-k})^{-1} \, A^{-k} E (A+E)^{-l} \hat{x}.\end{aligned}$$

Now apply the eigenvalue decomposition of A, take norms on both sides and use the fact that

$$\frac{1}{\|(\hat{\lambda}^{-l}\Lambda^{1-k} - \hat{\lambda}^{1-l}\Lambda^{-k})^{-1}\|_2} = \min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|^k \, |\hat{\lambda}|^l}.$$

□

Several existing bounds follow as special cases from Theorem 5.1.

Case $k = l = 0$. Theorem 5.1 is identical to one of the absolute bounds by Bauer and Fike [1, Theorem IIIa]

$$\min_i |\lambda_i - \hat{\lambda}| \leq \kappa \, \|E\|_2.$$

Case $k = 1$, $l = 0$. Theorem 5.1 is identical to the relative bound [4, Corollary 2.2]

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \kappa \, \|A^{-1} E\|_2.$$

Case $k = 0$, $l = 1$. Theorem 5.1 is identical to a relative bound, where the perturbed eigenvalue is in the denominator,

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\hat{\lambda}|} \leq \kappa \, \|E (A+E)^{-1}\|_2.$$

6. A Universal Eigenvalue Bound in the Frobenius Norm. We bound, in the Frobenius norm, the distances of all eigenvalues of $A+E$ to those of A.

Let A and $A+E$ be a complex, non-singular, diagonalizable matrices. Denote by

$$A = X \Lambda X^{-1}, \qquad A + E = \hat{X} \hat{\Lambda} \hat{X}^{-1}$$

eigenvalue decompositions, where

$$\Lambda = \begin{pmatrix} \ddots & & \\ & \lambda_i & \\ & & \ddots \end{pmatrix}, \qquad \hat{\Lambda} = \begin{pmatrix} \ddots & & \\ & \hat{\lambda}_i & \\ & & \ddots \end{pmatrix}$$

are diagonal matrices whose diagonal elements are the eigenvalues of A and $A+E$, respectively. Also let

$$\kappa \equiv \|X\|_2 \, \|X^{-1}\|_2, \qquad \hat{\kappa} \equiv \|\hat{X}\|_2 \, \|\hat{X}^{-1}\|_2$$

be the two-norm condition numbers with respect to inversion of the eigenvector matrices X and $\hat{X}$, respectively.

THEOREM 6.1. *If A and $A+E$ are diagonalizable then there is permutation τ such that*

$$\sqrt{\sum_i \left(\frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|}{|\lambda_i|^k\,|\hat{\lambda}_{\tau(i)}|^l}\right)^2} \leq \kappa\hat{\kappa}\;\|A^{-k}E(A+E)^{-l}\|_F.$$

Proof. The proof proceeds analogously to the one for [4, Theorem 5.1]. □

Several existing bounds are special cases of Theorem 6.1.

Case $k = l = 0$. Theorem 6.1 is identical to the absolute bound of the extended Hoffman-Wielandt theorem by Elsner and Friedland [5, Theorem 3.1],

$$\sqrt{\sum_i |\lambda_i - \hat{\lambda}_{\tau(i)}|^2} \leq \kappa\hat{\kappa}\;\|E\|_F.$$

Case $k = 1$, $l = 0$. Theorem 6.1 is identical to the relative bound [4, Corollary 5.2], as well as the multiplicative bound [13, Theorem 2.1′] with $D_1 = I$ and $D_2 = I + A^{-1}E$,

$$\sqrt{\sum_i \left(\frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|}{|\lambda_i|}\right)^2} \leq \kappa\hat{\kappa}\;\|A^{-1}E\|_F.$$

7. Effect of Scaling. We examine how row scaling affects eigenvalues and their perturbation bounds.

The motivation is the following. In the two-norm bound for $k = 1$, $l = 0$ in §5,

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \kappa\;\|A^{-1}E\|_2,$$

the term $A^{-1}E$ is invariant under row scaling, because if we row-scale A and $A+E$ to DA and $D(A+E)$ for some non-singular D, then $(DA)^{-1}(DE) = A^{-1}E$. Hence the row-scaled matrices have the same relative backward error as the original matrices. This is also true for the corresponding Frobenius norm bound in §6. Two questions arise: First, how do the eigenvalues change under row scaling? Second, how does the condition number κ of the eigenvectors change under row scaling?

7.1. Effect of Scaling on Eigenvalues. We determine relations between the eigenvalues of A and DA.

Let A and D be complex matrices of order n, and let λ_i be the eigenvalues of A and μ_i the eigenvalues of DA, ordered in decreasing magnitude,

$$|\lambda_n| \leq \ldots \leq |\lambda_1|, \qquad |\mu_n| \leq \ldots \leq |\mu_1|.$$

First, the eigenvalue products of A and DA differ by the determinant of D,

$$\mu_1 \cdots \mu_n = \det(D)\;\lambda_1 \cdots \lambda_n,$$

a consequence of $\det(DA) = \det(D)\det(A)$. This equality suggests that the change in eigenvalues is determined mostly by D alone, without the influence of other factors, such as the eigenvector conditioning κ.

Second, if A is normal, then the ratio of corresponding eigenvalues is bounded by $\|D\|$,

$$|\mu_i| \leq \|D\| \, |\lambda_i|, \qquad 1 \leq i \leq n,$$

which follows from the singular value product inequalities [8, Theorem 3.3.16(d)].

Third, when A is only diagonalizable, the corresponding bound turns into a relation between partial eigenvalue products,

$$|\mu_1 \cdots \mu_i| \leq (\kappa \, \|D\|)^i \, |\lambda_1 \cdots \lambda_i|, \qquad 1 \leq i \leq n,$$

which follows from [8, Theorem 3.3.2]. However this bound is not likely to be tight due to the presence of the eigenvector condition number κ.

7.2. Effect of Scaling on Eigenvector Condition Number: Matrices of Order 2. We examine the effect of row scaling on the condition number with respect to inversion of the eigenvectors.

In particular, we want to know how the condition number for the eigenvectors of DA compares to κ, the condition number of the eigenvectors of A. If the two eigenvector condition numbers have the same order of magnitude then the perturbation bounds for the eigenvalues of A and DA provide similar estimates. In this case, the eigenvalues of the scaled matrix DA are about as sensitive to perturbations as the eigenvalues of A, and the scaling has not done any harm.

In general, by how much can the condition numbers for eigenvectors of DA and A differ? To get a feeling for the condition number of the eigenvectors of a scaled matrix, we first consider matrices of order 2.

The Original Problem. Consider a non-singular diagonalizable triangular matrix

$$A \equiv \begin{pmatrix} \lambda_1 & \eta \\ & \lambda_2 \end{pmatrix}, \qquad \text{where} \quad \lambda_1 \neq \lambda_2, \quad \lambda_1\lambda_2 \neq 0.$$

An eigendecomposition is $A = X\Lambda X^{-1}$, where

$$\Lambda \equiv \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix}, \qquad X \equiv \begin{pmatrix} 1 & \xi \\ & 1 \end{pmatrix}, \qquad \xi \equiv \frac{\eta}{\lambda_2 - \lambda_1}.$$

Since $\|X\|_F = \|X^{-1}\|_F = \sqrt{2 + |\xi|^2}$, the Frobenius norm condition number of the eigenvectors is

$$\kappa_F(X) \equiv \|X\|_F \, \|X^{-1}\|_F = 2 + \left| \frac{\eta}{\lambda_2 - \lambda_1} \right|^2.$$

The condition number is small if $|\eta| \lesssim |\lambda_1 - \lambda_2|$. This means the eigenvalues of A are well-conditioned if the non-normality η is not much larger than the absolute eigenvalue separation.

The Scaled Problem. The row scaling is given by a non-singular diagonal matrix

$$D \equiv \begin{pmatrix} d_1 & \\ & d_2 \end{pmatrix}, \qquad \text{where} \quad d_1 d_2 \neq 0.$$

We also assume that $d_1\lambda_1 \neq d_2\lambda_2$, so

$$DA = \begin{pmatrix} d_1\lambda_1 & d_1\eta \\ & d_2\lambda_2 \end{pmatrix}$$

is diagonalizable with distinct eigenvalues. An eigendecomposition is $DA = \tilde{X}\tilde{\Lambda}\tilde{X}^{-1}$ with

$$\tilde{X} \equiv \begin{pmatrix} 1 & \tilde{\xi} \\ & 1 \end{pmatrix},$$

and

$$\tilde{\xi} \equiv \frac{d_1 \eta}{d_2 \lambda_2 - d_1 \lambda_1} = -\frac{\eta}{\lambda_1} \frac{1}{1-\omega}, \qquad \text{where} \quad \omega \equiv \frac{d_2}{d_1} \frac{\lambda_2}{\lambda_1}.$$

The factor $|\eta|/|\lambda_1|$ can be interpreted as a relative departure of A from normality, while ω is a measure for the eigenvalue separation of DA.

The eigenvector condition number $\kappa_F(\tilde{X}) = 2 + |\tilde{\xi}|^2$ indicates how sensitive the eigenvalues of DA are to perturbations in the matrix. Since, by assumption, $\omega \neq 1$, we distinguish two cases.

$|\omega| < 1$:

$$\tilde{\xi} = -\frac{\eta}{\lambda_1} \left(1 + \mathcal{O}(\omega)\right),$$

and the condition number for the eigenvectors of DA is bounded by

$$\kappa_F(\tilde{X}) \leq 2 + \left|\frac{\eta}{\lambda_1}\right|^2 \left(1 + \mathcal{O}(|\omega|)\right).$$

$|\omega| > 1$:

$$\tilde{\xi} = -\frac{\eta}{\lambda_1} \frac{1}{\omega} \left(1 + \mathcal{O}\left(\frac{1}{\omega}\right)\right),$$

and the condition number for the eigenvectors of DA is bounded by

$$\kappa_F(\tilde{X}) \leq 2 + \left|\frac{\eta}{\lambda_1}\right|^2 \frac{1}{|\omega|^2} \left(1 + \mathcal{O}\left(\frac{1}{|\omega|}\right)\right).$$

We conclude that for diagonalizable triangular matrices of order 2, the condition number of the eigenvectors of DA is governed by the relative departure from normality $|\eta|/|\lambda_1|$ of A. If the relative departure from normality of A is moderate or low then eigenvector matrices of any row scaling DA are well-conditioned with respect to inversion and the eigenvalues of the row scaled matrix DA are well-conditioned. When $|d_2\lambda_2| > |d_1\lambda_1|$ (i.e. $|\omega| > 1$) the scaling can even improve the condition number of the eigenvectors.

Therefore the conditioning of the eigenvalues of a scaled 2×2 triangular matrix is governed by the relative departure from normality of the original matrix.

7.3. Effect of Scaling on Eigenvector Condition Number: Matrices of Order n. We extend the above observations for matrices of order 2 to matrices of order n.

The Original Problem. Consider the diagonalizable triangular matrix

$$A \equiv \begin{array}{cc} & \begin{array}{cc} n-k & k \end{array} \\ \begin{array}{c} n-k \\ k \end{array} & \begin{pmatrix} T_1 & N \\ & T_2 \end{pmatrix} \end{array}$$

of order n, where T_1 and T_2 are triangular, and the eigenvalues of T_1 are different from those of T_2. A similarity transformation to block diagonal form is $A = X\Lambda X^{-1}$, where

$$\Lambda \equiv \begin{pmatrix} T_1 & \\ & T_2 \end{pmatrix}, \qquad X \equiv \begin{pmatrix} I & X_1 \\ & I \end{pmatrix},$$

and X_1 satisfies $X_1 T_2 - T_1 X_1 = N$. The condition number of the similarity transformation is

$$\kappa_F(X) \equiv \|X\|_F \, \|X^{-1}\|_F = n + \|X_1\|_F^2.$$

To extract X_1, consider one column of $X_1 T_2 - T_1 X_1 = N$ at a time and stack up the columns. The result is a non-singular, block-lower triangular system of order $k(n-k)$, which in the case $k = 3$ looks like

$$\left[\begin{pmatrix} (T_2)_{11} I & & \\ (T_2)_{12} I & (T_2)_{22} I & \\ (T_2)_{13} I & (T_2)_{23} I & (T_2)_{33} I \end{pmatrix} - \begin{pmatrix} T_1 & & \\ & T_1 & \\ & & T_1 \end{pmatrix}\right] \begin{pmatrix} X_1 e_1 \\ X_1 e_2 \\ X_1 e_3 \end{pmatrix} = \begin{pmatrix} N e_1 \\ N e_2 \\ N e_3 \end{pmatrix},$$

where e_i is the *ith* column of I. With $\otimes$ the Kronecker product and $\mathrm{vec}(A)$ the vector of columns of A [8, §4.2], one can write $X_1 T_2 - T_1 X_1 = N$ as [8, §4.3]

$$\left[(T_2^T \otimes I) - (I \otimes T_1)\right] \mathrm{vec}(X_1) = \mathrm{vec}(N).$$

The Scaled Problem. Now consider the row scaled matrix DA, where

$$D \equiv \begin{pmatrix} D_1 & \\ & D_2 \end{pmatrix},$$

and D_1 and D_2 are non-singular diagonal. We also assume that the eigenvalues of $D_1 T_1$ are different from those of $D_2 T_2$. Then

$$DA = \begin{pmatrix} D_1 T_1 & D_1 N \\ & D_2 T_2 \end{pmatrix}$$

can be reduced to block-diagonal form via a similarity transformation. That is, $DA = \tilde{X}\tilde{\Lambda}\tilde{X}^{-1}$ with

$$\tilde{X} \equiv \begin{pmatrix} I & \tilde{X}_1 \\ & I \end{pmatrix},$$

and $\tilde{X}_1$ satisfies

$$\tilde{X}_1 \, D_2 T_2 - D_1 T_1 \, \tilde{X}_1 = D_1 N,$$

or

$$D_1^{-1} \tilde{X}_1 D_2 T_2 - T_1 \tilde{X}_1 = N.$$

In Kronecker product form this is

$$\left[((D_2 T_2)^T \otimes D_1^{-1}) - (I \otimes T_1)\right] \mathrm{vec}(\tilde{X}_1) = \mathrm{vec}(N).$$

Solving for $\tilde{X}_1$ gives

$$\operatorname{vec}(\tilde{X}_1) = -(I-W)^{-1}\operatorname{vec}(T_1^{-1}N), \qquad \text{where} \quad W \equiv (D_2T_2)^T \otimes (D_1T_1)^{-1}.$$

As before, we interpret $\|\operatorname{vec}(T_1^{-1}N)\|_2 = \|T_1^{-1}N\|_F$ as a relative departure of A from (block) normality, while $\|W\|_2$ indicates how far the two sets of eigenvalues of DA are apart.

By assumption, the diagonal elements of D_1T_1 and D_2T_2 are different, so $\|W\|_2 \neq 1$. If $\|W\|_2 < 1$ then

$$\|\tilde{X}_1\|_F = \|\operatorname{vec}(\tilde{X}_1)\|_2 \leq \frac{\|\operatorname{vec}(T_1^{-1}N)\|_2}{1-\|W\|_2} = \frac{\|T_1^{-1}N\|_F}{1-\|W\|_2},$$

and the condition number for the similarity transformation is bounded by

$$\kappa_F(\tilde{X}) \leq n + \|T_1^{-1}N\|_F^2\,(1+\mathcal{O}(\|W\|))\,.$$

Hence, if the relative departure from (block) normality of A is moderate or low then eigenvector matrices of any row scaling DA are well-conditioned with respect to inversion and the eigenvalues of the row scaled matrix DA are well-conditioned.

REFERENCES

[1] F. Bauer and C. Fike, *Norms and exclusion theorems*, Numer. Math., 2 (1960), pp. 137–41.

[2] C. Davis and W. Kahan, *Some new bounds on perturbation of subspaces*, Bull. Amer. Math. Soc., 75 (1969), pp. 863–8.

[3] C. Davis and W. Kahan, *The rotation of eigenvectors by a perturbation, III*, SIAM J. Numer. Anal., 7 (1970), pp. 1–46.

[4] S. Eisenstat and I. Ipsen, *Three absolute perturbation bounds for matrix eigenvalues imply relative bounds*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 149–58.

[5] L. Elsner and S. Friedland, *Singular values, doubly stochastic matrices, and applications*, Linear Algebra Appl., 220 (1995), pp. 161–9.

[6] I. Gohberg, R. Lancaster, and L. Rodman, *Invariant Subspaces of Matrices with Applications*, John Wiley & Sons, New York, 1986.

[7] G. Golub and C. van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, third ed., 1996.

[8] R. Horn and C. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, 1991.

[9] I. Ipsen, *Relative perturbation results for matrix eigenvalues and singular values*, in Acta Numerica 1998, vol. 7, Cambridge University Press, Cambridge, 1998, pp. 151–201.

[10] ———, *Absolute and relative perturbation bounds for invariant subspaces of matrices*, Linear Algebra Appl., 309 (2000), pp. 45–56.

[11] ———, *An overview of relative* $\sin\Theta$ *theorems for invariant subspaces of complex matrices*, J. Comput. Appl. Math., 123 (2000), pp. 131–53. Invited Paper for the special issue *Numerical Analysis 2000: Vol. III – Linear Algebra*.

[12] P. Lancaster and M. Tismenetsky, *The Theory of Matrices, Second Edition*, Academic Press, Orlando, 1985.

[13] R. Li, *Relative perturbation theory: III. more bounds on eigenvalue variation*, Linear Algebra Appl., 266 (1997), pp. 337–45.

[14] ———, *Relative perturbation theory: II. eigenspace and singular subspace variations*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 471–92.

[15] T. Londrè and N. Rhee, *A note on relative perturbation bounds*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 357–61.

[16] R. Mathias and K. Veselić, *A relative perturbation bound for positive definite matrices*, Linear Algebra Appl., 270 (1998), pp. 315–21.

[17] I. Slapničar and N. Truhar, *Relative perturbation theory for hyperbolic eigenvalue problem*, Linear Algebra Appl., 309 (2000), pp. 57–72.

[18] G. Stewart and J. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, 1990.