ABSOLUTE AND RELATIVE PERTURBATION BOUNDS FOR IN Variant SUBSPACES OF MATRICES

ILSE C. F. IPSEN*

10 August 98

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 are extensions of Davis and Kahan’s sinθ theorem to general matrices and invariant subspaces of any dimension. When the perturbed subspace has dimension one, the relative bound is implied by the absolute bound. The relative bounds presented here are the most general relative bounds for invariant subspaces because they place no restrictions on the matrix or the perturbation.

Key words. invariant subspace, condition number, separation between matrices, absolute error, relative error, eigenvalue separation, angle between subspaces

AMS subject classification. 15A12, 15A18, 15A42, 15A69, 65F15, 65F35

1. Introduction. Absolute and relative perturbation bounds are derived for angles between invariant subspaces of a complex square matrix $A$ and a perturbed matrix $A+E$, in the two-norm and in the Frobenius norm. The relative bounds presented here are the most general relative bounds because they place no restrictions on the original matrix $A$, the perturbation $E$, or the dimensions of the subspaces. The bounds are similar in spirit to Stewart’s bounds for invariant subspaces [15, 16] and, in the case of normal matrices, they reduce to the sinθ theorem of Davis and Kahan [4, 5].

The bounds presented in this paper demonstrate the following.

1. Relative bounds for invariant subspaces always exist, for any non-singular matrix $A$ and any perturbation $E$ (however whether the bounds are small depends on the condition numbers and on the size of $\|A^{-1}E\|$). In this sense relative bounds appear to be no more special than absolute bounds.

2. When the perturbed eigenspace has dimension one, the absolute bound implies the relative bound.

3. Absolute and relative bounds share the same subspace condition numbers, namely the conditioning of the perturbed subspace basis and the conditioning of an unwanted left subspace basis. This suggests that invariant subspaces exhibit the same sensitivity to basis conditioning, in the absolute as well as in the relative sense.

After defining the problem in Section 2, we show in Section 3 that the absolute bound implies a relative bound when the perturbed subspace has dimension one. For a perturbed subspace of arbitrary dimension absolute and relative bounds are derived in the Frobenius norm in Section 4, and in the two-norm in Section 5. The paper ends with a review of the literature in Section 6.

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

---

* Center for Research In Scientific Computation, Department of Mathematics, North Carolina State University, P. O. Box 8205, Raleigh, NC 27695-8205 (ipsen@math.ncsu.edu, http://www4.ncsu.edu/~ipsen/info.html). This research was supported in part by NSF grants CCR-9400921 and DMS-9714811.
2. The Problem. Let $A$ be a complex square matrix with invariant subspace \( \text{range}(X_1) \), that is,
\[
AX_1 = X_1B_1
\]
for some non-singular $B_1$. The perturbed subspace \( \text{range}(\hat{X}) \) approximates \( \text{range}(X_1) \),
\[
(A + E)\hat{X} = \hat{X}\hat{\lambda},
\]
for some $\hat{\lambda}$. Here $\hat{\lambda}$ is a diagonal matrix and $\hat{X}$ has full column rank. The number of columns in $\hat{X}$ can be different from that of $X_1$, hence $\text{range}(X_1)$ and $\text{range}(\hat{X})$ do not necessarily have the same dimensions. The goal is to bound the angles between the perturbed subspace $\text{range}(\hat{X})$ and the desired subspace $\text{range}(X_1)$.

The sines of the angles between $\text{range}(\hat{X})$ and $\text{range}(X_1)$ are determined from orthonormal bases as follows. Decompose $[9, \S4.8] A = XBX^{-1}$, where
\[
B = \begin{pmatrix} B_1 & B_{12} \\ B_{21} & B_2 \end{pmatrix}, \quad X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad X^{-1} = \begin{pmatrix} Y_1^* \\ Y_2^* \end{pmatrix}.
\]

Let the columns of $Q_2$ be an orthonormal basis for the unwanted left invariant subspace $\text{range}(Y_2)$ and the columns of $\hat{Q}$ an orthonormal basis for the perturbed subspace $\text{range}(\hat{X})$. The singular values of
\[
\sin \Theta \equiv Q_2^*\hat{Q}
\]
are the principal angles between $\text{range}(\hat{X})$ and $\text{range}(X_1)$ $[8, \S12.4.3], [13, \S3]$.

3. A Single Perturbed Eigenvector. For a one-dimensional perturbed eigenspace it is shown that the absolute bound implies a relative bound. The bounds are expressed in the two norm.

Since the perturbed subspace $\hat{X}$ consists of only one column and $\hat{\lambda}$ is a scalar we write instead $\hat{x}$ and $\hat{\lambda}$, respectively. That is,
\[
(A + E)\hat{x} = \hat{\lambda}\hat{x}, \quad \|\hat{x}\| = 1.
\]
The principal angle between $\text{range}(\hat{x})$ and $\text{range}(X_1)$ is $0 \leq \theta \leq \pi$. The subsequent absolute bound on $\sin \theta$ is expressed in terms of the separation between $B_2$ and the perturbed eigenvalue $\hat{\lambda}$,
\[
\text{abssep}(B_2, \hat{\lambda}) \equiv 1/\| (B_2 - \hat{\lambda}I)^{-1} \|.
\]
Here, as in $[15, \S2]$ and in the alternative definition in $[16, \S4.3]$, the separation between two matrices is based on the two-norm, rather than the Frobenius norm as in $[8, \S 7.2.4], [20, \S1]$.

**Theorem 3.1.** If $\text{abssep}(B_2, \hat{\lambda}) > 0$ then
\[
\sin \theta \leq \kappa(Y_2) \frac{\|E\|}{\text{abssep}(B_2, \hat{\lambda})},
\]
where $\kappa(Y_2) \equiv \|Y_2\| \|Y_2^T\|$.  

**Proof.** Replacing $A$ by the partition (2.1) in $(A + E)\hat{x} = \hat{\lambda}\hat{x}$ and looking at the second block row gives
\[
(B_2 - \hat{\lambda}I) Y_2^*\hat{x} = -Y_2^*E\hat{x}.
\]
Take norms and use the QR decomposition $Y_2 = Q_2R_2$ from Section 2,

$$
\|Y_2\| \|E\| \geq \|Y_2^*E\| \geq \frac{\sin \theta}{\|R_2^{-1}\| \|((B_2 - \lambda I)^{-1})\|},
$$

where $\|R_2^{-1}\| = \|Y_2^*\|$. 

The above bound contains no explicit dependence on $B_1$ or $B_{12}$. It only depends on the conditioning $\kappa(Y_2)$ of the basis for the subspace range($X_2$) and on $B_2$. This suggests that a subspace in which we have no interest can affect the sensitivity of the desired subspace. The condition number $\kappa(Y_2)$ does not exceed the condition number of the similarity transformation, $\kappa(Y_2) \leq \kappa(X)$.

In the special case when $A$ is diagonalizable one can choose $X$ to be an eigenvector matrix, and $B_2 = \Lambda_2$ to be a diagonal matrix. The separation $\text{abssep}(B_{22}, \hat{\lambda})$ reduces to the eigenvalue separation

$$
\text{absgap}(\Lambda_2, \hat{\lambda}) \equiv 1/\|\Lambda_2 - \hat{\lambda}\| = \min_{\lambda \in \Lambda_2} |\lambda - \hat{\lambda}|.
$$

Thus, when $A$ is diagonalizable Theorem 3.1 becomes [7, Corollary 4.3],

$$
\sin \theta \leq \kappa(Y_2) \frac{\|E\|}{\text{absgap}(\Lambda_2, \hat{\lambda})}.
$$

When $A$ is normal, there is a unitary eigenvector matrix $X$, hence $\kappa(Y_2) = 1$ and Theorem 3.1 gives the same bound as Davis and Kahan’s $\sin \theta$ Theorem Theorem [4, §6], [5, §2].

The absolute bound in Theorem 3.1 implies the following relative bound, which is expressed in terms of a relative two-norm separation between $B_2$ and $\hat{\lambda}$,

$$
\text{relsep}(B_{22}, \hat{\lambda}) \equiv 1/\|B_2(\hat{\lambda} - \hat{\lambda})\|.
$$

**Corollary 3.2.** If $\text{relsep}(B_{22}, \hat{\lambda}) > 0$ and $A$ is non-singular then

$$
\sin \theta \leq \kappa(Y_2) \frac{\|A^{-1}E\|}{\text{relsep}(B_{22}, \hat{\lambda})}.
$$

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

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

Note that $\hat{x}$ is an eigenvector of $\hat{A} + \hat{E}$ associated with eigenvalue 1, and that $\hat{A}$ can be decomposed with the same similarity transformation as $A$, $\hat{A} = X (\hat{\lambda}B^{-1}) X^{-1}$. Applying the absolute bound Theorem 3.1 to $(\hat{A} + \hat{E})\hat{x} = \hat{x}$ yields

$$
\sin \theta \leq \kappa(Y_2) \frac{\|\hat{E}\|}{\text{abssep}(\hat{\lambda}B_2^{-1}, 1)},
$$

where

$$
\text{abssep}(\hat{\lambda}B_2^{-1}, 1) = \frac{1}{\|((\hat{\lambda}B_2^{-1} - I)^{-1})\|} = \frac{1}{\|B_2(\hat{\lambda} - \hat{\lambda})\|} = \text{relsep}(B_2, \hat{\lambda}).
$$


The absolute and the relative bound have the same condition number \( \kappa(Y_2) \).

In the special case when \( A \) is diagonalizable, one can choose \( B_2 = \Lambda_2 \) to be diagonal and the relative separation \( \text{relsep}(B_{22}, \hat\lambda) \) reduces to the relative eigenvalue separation

\[
\text{relgap}(\Lambda_2, \hat\lambda) \equiv 1/\|\Lambda_2(\Lambda_2 - \hat\lambda I)^{-1}\| = \min_{\lambda \in \Lambda_2} \frac{|\lambda - \hat\lambda|}{|\lambda|}.
\]

Thus when \( A \) is diagonalizable Corollary 3.2 becomes

\[
\sin \theta \leq \kappa(Y_2) \frac{\|A^{-1}E\|}{\text{relgap}(\Lambda_2, \hat\lambda)}.
\]

4. Frobenius Norm Bounds for a Perturbed Subspace. Absolute and relative Frobenius-norm bounds are derived for a perturbed subspace range(\( \hat{X} \)) of arbitrary dimension.

The absolute Frobenius norm separation between \( B_2 \) and \( \hat{\Lambda} \) is

\[
\text{abssep}(B_2, \hat{\Lambda}) \equiv \min_{\|Z\|_F = 1} \|B_2 Z - \hat{\Lambda} Z\|_F.
\]

The bound below can be considered an extension of Davis and Kahan’s \( \sin \theta \) theorem [4, §6], [5, §2] to general matrices and invariant subspaces of any dimension.

**Theorem 4.1.** If \( \text{abssep}(B_2, \hat{\Lambda}) > 0 \) then

\[
\|\sin \Theta\|_F \leq \kappa(Y_2) \frac{\|E\|_F}{\text{abssep}(B_2, \hat{\Lambda})},
\]

where \( \kappa(\hat{X}) \equiv \|\hat{X}\| \|\hat{X}^*\| \).

**Proof.** Replacing \( A \) by the partition (2.1) in \( (A + E)\hat{X} = \hat{X}\hat{\Lambda} \) and looking at the second block row yields

\[
B_2 Z - Z \hat{\Lambda} = -Y_2^* E\hat{X}, \quad \text{where} \quad Z \equiv Y_2^* \hat{X}.
\]

Taking norms gives

\[
\|Y_2\| \|E\|_F \|\hat{X}\| \geq \|Y_2^* E \hat{X}\|_F \geq \text{abssep}(B_2, \hat{\Lambda}) \|Z\|,
\]

where we have used \( \|AB\|_F \leq \|A\| \|B\|_F \). The QR decompositions \( Y_2 = Q_2 \hat{R}_2 \) and \( \hat{X} = Q \hat{R} \) lead to

\[
\|Z\|_F = \|R_2^* Q_2^* Q \hat{R}\|_F \geq \frac{\|\sin \Theta\|_F}{\|R_2\| \|\hat{R}^{-1}\|},
\]

where \( \|R_2^{-1}\| = \|Y_2^*\| \) and \( \|\hat{R}^{-1}\| = \|\hat{X}^*\| \).

Compared to the single vector case in Theorem 3.1, Theorem 4.1 also contains the conditioning \( \kappa(\hat{X}) \) of the basis for the perturbed subspace.

When the decomposition of \( A \) in (2.1) is a Schur decomposition then Theorem 4.1 has the same spirit as [8, Theorem 7.2.4], [16, Theorem 4.11], and [17, Theorem V.2.1], where the separation is between \( B_1 \) and \( B_2 \), rather than between \( B_1 \) and \( \hat{\Lambda} \).

In the special case when \( A \) is diagonalizable one can choose \( B_2 = \Lambda_2 \) to be diagonal, and the separation \( \text{abssep}(B_2, \Lambda) \) amounts to an eigenvalue separation [15, Theorem 2.4], [16, Theorem 4.7]

\[
\text{absgap}(\Lambda_2, \hat{\Lambda}) \equiv \min_{\lambda \in \Lambda_2, \hat{\lambda} \in \hat{\Lambda}} |\lambda - \hat{\lambda}|.
\]
Thus when \( A \) is diagonalizable Theorem 4.1 becomes
\[
\| \sin \Theta \|_F \leq \kappa(Y_2) \frac{\| E \|_F}{\text{absgap}(\Lambda_2, \hat{\Lambda})}.
\]

For the relative bound we define the relative separation between \( B_2 \) and \( \hat{\Lambda} \) as
\[
\text{relsep}(B_2, \hat{\Lambda}) \equiv \min_{\lambda, \hat{\lambda}} \frac{1}{\| B_2(B_2 - \hat{\lambda}I) \|}.
\]

**Theorem 4.2.** If \( \text{relsep}(B_2, \hat{\Lambda}) > 0 \) and \( A \) is non-singular then
\[
\| \sin \Theta \|_F \leq \kappa(Y_2) \frac{\| A^{-1}E \|_F}{\text{relsep}(B_2, \hat{\Lambda})}.
\]

**Proof.** \((A + E)\hat{\Lambda} = \hat{\Lambda} \hat{\Lambda} \text{ implies } \hat{\Lambda} - A^{-1} \hat{\Lambda} \hat{\Lambda} = -A^{-1}E \hat{\Lambda} \). Replacing \( A \) by the partition (2.1), multiplying by \( X^{-1} \) on the left, and looking at the second block row yields
\[
Z - B_2^{-1} Z \hat{\Lambda} = -Y_2^* A^{-1} E X, \quad \text{where } Z \equiv Y_2^* \hat{\Lambda}.
\]

The \( j \)th column is
\[
(I - \hat{\lambda}_j B_2^{-1}) Z_j = -Y_2^* A^{-1} E \hat{\Lambda} X_j.
\]

Taking two-norms,
\[
\| Y_2^* A^{-1} E \hat{\Lambda} X_j \| \geq \frac{\| Z_j \|}{\| B_2(B_2 - \hat{\lambda}I) \|}
\]

and summing up the squares of the two-norms for all columns gives
\[
\| Y_2^* A^{-1} E \hat{\Lambda} \|_F^2 \geq \| Z \|_F^2 \text{ relsep}(B_2, \hat{\Lambda})^2.
\]

Now extract \( \sin \Theta \) from \( Z \) as in the proof of Theorem 4.1. \( \Box \)

In the special case when \( A \) is diagonalizable, one can choose \( B_2 = \Lambda_2 \) to be diagonal and the relative separation \( \text{relsep}(B_2, \hat{\Lambda}) \) reduces to the relative eigenvalue separation
\[
\text{relgap}(\Lambda_2, \hat{\Lambda}) \equiv \min_{\lambda, \hat{\lambda} \in \Lambda_2, \hat{\Lambda}} \frac{|\lambda - \hat{\lambda}|}{|\lambda|}.
\]

Thus when \( A \) is diagonalizable Theorem 4.2 becomes
\[
\sin \theta \leq \kappa(Y_2) \frac{\| A^{-1}E \|_F}{\text{relgap}(\Lambda_2, \hat{\Lambda})}.
\]

The absolute bound in Theorem 4.1 and the relative bound in Theorem 4.2 have the same eigenvector condition numbers \( \kappa(Y_2) \) and \( \kappa(\hat{\Lambda}) \).
5. Two-Norm Bounds for a Perturbed Subspace. Absolute and relative two-norm bounds are derived for a perturbed subspace range($\hat{X}$) of arbitrary dimension. The proofs in this section resemble the proofs of [10, Lemmas 2.3, 2.5].

The two-norm bounds require a stronger eigenvalue separation than the Frobenius norm bounds: The perturbed eigenvalues must be strictly separated from the singular values of $B_2$. The absolute separation is

$$\text{absgap}(B_2, \hat{\Lambda}) \equiv \max \left\{ \frac{1}{\|B_2^{-1}\|} - \|\hat{\Lambda}\|, \frac{1}{\|\hat{\Lambda}^{-1}\|} - \|B_2\| \right\}.$$ 

**Theorem 5.1.** If $\text{absgap}(B_2, \hat{\Lambda}) > 0$ then

$$\|\sin \Theta\| \leq \kappa(Y_2) \frac{\kappa(\hat{X})}{\text{absgap}(B_2, \hat{\Lambda})} \|E\|.$$

**Proof.** As in the proof of Theorem 4.1 derive

$$B_2 Z - Z \hat{\Lambda} = -Y_2^* E \hat{X}, \quad \text{where} \quad Z \equiv Y_2^* \hat{X}.$$

Suppose $\text{absgap}(B_2, \hat{\Lambda}) = \frac{1}{\|B_2^{-1}\|} - \|\hat{\Lambda}\| > 0$ (the other case is similar). Then

$$\|Y_2^* E \hat{X}\| = \|B_2 Z - Z \hat{\Lambda}\| \geq \|B_2 Z\| - \|Z \hat{\Lambda}\| \geq \|Z\| \text{absgap}(B_2, \hat{\Lambda}).$$

Bound $\|Z\|$ in terms of $\|\sin \Theta\|$ as in the proof of Theorem 4.1. □

The relative two-norm bound also requires a stronger relative gap,

$$\text{relgap}(B_2, \hat{\Lambda}) \equiv \max \left\{ 1 - \|B_2^{-1}\| \|\hat{\Lambda}\|, 1 - \|\hat{\Lambda}^{-1}\| \|B_2\| \right\}.$$ 

**Theorem 5.2.** If $\text{relgap}(B_2, \hat{\Lambda}) > 0$ and $A$ is non-singular then

$$\|\sin \Theta\| \leq \kappa(Y_2) \frac{\kappa(\hat{X})}{\text{relgap}(B_2, \hat{\Lambda})} \|A^{-1} E\|.$$

**Proof.** As in the proof of Theorem 4.2 derive

$$Z - B_2^{-1} Z \hat{\Lambda} = -Y_2^* A^{-1} E \hat{X}, \quad \text{where} \quad Z \equiv Y_2^* \hat{X}.$$

Suppose $\text{relgap}(B_2, \hat{\Lambda}) = 1 - \|B_2^{-1}\| \|\hat{\Lambda}\|$ (the other case is similar), then

$$\|Y_2^* A^{-1} E \hat{X}\| = \|Z - B_2^{-1} Z \hat{\Lambda}\| \geq \|Z\| - \|B_2^{-1} Z \hat{\Lambda}\| \geq \|Z\| \text{relgap}(B_2, \hat{\Lambda}).$$

Bound $\|Z\|$ in terms of $\|\sin \Theta\|$ as in the proof of Theorem 4.1. □

The absolute and relative two-norm bounds have the same eigenvector condition numbers $\kappa(Y_2)$ and $\kappa(\hat{X})$ as the Frobenius norm bounds. Since $\text{resep}(B_2, \hat{\Lambda}) > 0$ if and only if $\text{abssep}(B_2, \hat{\Lambda}) > 0$, the relative two-norm bound in Theorem 5.2 holds if and only if the absolute bound in Theorem 5.1 holds. As in the previous sections, the matrix separations reduce to eigenvalue separations when $A$ is diagonalizable.
6. Existing Literature. In the context of absolute bounds, Bhatia, Davis and McIntosh [3] prove \( \sin \theta \) theorems for normal operators in Hilbert spaces. Their bounds are also of the form
\[
\sin \theta \leq c \| AX - XB \| / \text{absgap},
\]
where \( A \) and \( B \) are normal operators and \( X \) represents a perturbed subspace of any dimension. Determining the value of the positive constant \( c \) amounts to solving a minimization problem for functions in \( L_1 \) [2]. These bounds are also discussed in [1, VII.3].

For diagonalizable matrices Varah [19, Theorem 2.2] shows that if \( \| E \| \) is sufficiently small then
\[
\sin \theta \leq \kappa(X) \kappa(\| E \|) / \text{absgap}.
\]

For general, possibly defective matrices, Stewart [15, Theorem 4.1], [16, Theorem 4.11] derives a tan \( \Theta \) bound for the case when \( X \) consists of Ritz values and \( \| E \| \) is sufficiently small. And Ruhe [13, Corollary 1] bounds the sines of the angles between \( \text{range}(X) \) and singular vectors associated with the smallest singular values of \( A \). Here \( \text{absgap} \) is replaced by the singular value separation \( \sqrt{\sigma_n^2 + \sigma_{n+1}^2} \), where \( s \) is the dimension of \( \text{range}(X) \).

In the context of multiplicative perturbations, where the perturbed matrix is expressed as \( D_1 AD_2 \), relative perturbation bounds for invariant subspaces have been derived for Hermitian matrices [9, 10] and for diagonalizable matrices [7].

In the context of additive perturbations, where the perturbed matrix is represented as \( A + E \), relative bounds have been derived for Hermitian matrices. In [11, Theorem 2.7] an asymptotic bound for a single eigenvector of a Hermitian positive-definite matrix is derived in terms of the relative gap
\[
\min_{\lambda \in \mathbb{C}} \frac{|\lambda - \bar{\lambda}|}{\sqrt{\lambda \bar{\lambda}}}
\]
and the relative perturbation \( \| A^{1/2} E A^{1/2} \| \). A similar non-asymptotic bound is derived in [12, Theorem 1]. It is extended to subspaces of arbitrary dimension and unitarily invariant norms in [10, Theorems 3.3, 3.4].

The bounds in [14, Theorem 1], [18], [21, §2.1] hold for equally dimensioned invariant subspaces of Hermitian matrices in the context of component-wise relative perturbations and are expressed in terms of yet a different set of relative gaps.

In contrast to the bounds presented here, the existing relative bounds for additive perturbations have the advantage of being invariant under congruence transformations and grading. However, the bounds here are more general because they place no restrictions on the original matrix \( A \) or the perturbation \( E \). Also they are simpler and easier to interpret than the bounds in [14, 18, 21].

REFERENCES