Acta Numerica (1998), *pp.* 1–000

Relative Perturbation Results for Matrix Eigenvalues and Singular Values

Ilse C.F. Ipsen*
Department of Mathematics
North Carolina State University
Raleigh, NC 27695-8205, USA
ipsen@math.ncsu.edu
http://www4.ncsu.edu/~ipsen/info.html

It used to be good enough to bound absolute errors of matrix eigenvalues and singular values. Not anymore. Now it is fashionable to bound relative errors. We present a collection of relative perturbation results which have emerged during the past ten years.
No need to throw away all those absolute error bounds, though. Deep down the derivation of many relative bounds can be based on absolute bounds. This means, relative bounds are not always better. They may just be better sometimes – and exactly when depends on the perturbation.

CONTENTS

1. Introduction

Are you concerned about accuracy of matrix eigenvalues or singular values, especially the ones close to zero? If so, this paper is for you!

We present error bounds for eigenvalues and singular values that can be much tighter than the traditional bounds, especially when these values have

* Work supported in part by grant CCR-9400921 from the National Science Foundation.

small magnitude. Our goal is to give some intuition for what the bounds mean and why they hold.

Suppose you have to compute an eigenvalue of a complex square matrix A. Numerical software usually produces a number $\hat{\lambda}$ that is not the desired eigenvalue. So you ask yourself how far away is $\hat{\lambda}$ from an eigenvalue of A? If $\hat{\lambda}$ was produced by a reliable (i.e. backward stable) numerical method, there is a round off error analysis to assure you that $\hat{\lambda}$ is an eigenvalue of a nearby matrix $A + E$, where E is small in some sense. Then you can use perturbation theory to estimate the error in $\hat{\lambda}$. For instance, when A is diagonalisable the Bauer-Fike theorem bounds the absolute distance between $\hat{\lambda}$ and a closest eigenvalue λ of A by

$$|\lambda - \hat{\lambda}| \leq \kappa(X) \, \|E\|, \tag{1.1}$$

where $\kappa(X) \equiv \|X\| \, \|X^{-1}\|$ is the condition number of an eigenvector matrix X of A.

The quantity $|\lambda - \hat{\lambda}|$ represents an absolute error. Traditional perturbation theory assesses the quality of a perturbed eigenvalue by bounding absolute errors. However there are practical situations where small eigenvalues have physical meaning and should be determined to high relative accuracy. Such situations include computing modes of vibration in a finite element context, and computing energy levels in quantum mechanical systems (Demmel, Gu, Eisenstat, Slapničar, Veselič and Drmač 1997, §1). Absolute error bounds cannot cope with relative accuracy, especially when confronted with small eigenvalues or singular values. The following section explains why.

1.1. Why Absolute Bounds Don't Do the Job

If we want relative accuracy, we need relative error bounds. The simplest way to generate a relative error bound is to divide an absolute error bound by an eigenvalue. For instance, dividing the absolute error bound (1.1) by a non-zero eigenvalue λ produces the relative error bound

$$\frac{|\lambda - \hat{\lambda}|}{|\lambda|} \leq \frac{\|E\|}{|\lambda|}.$$

Unlike the absolute bound, though, the relative bound depends on λ. This has several disadvantages. First, each eigenvalue has a different relative bound. Second, the relative bound is smaller for eigenvalues λ that are large in magnitude than for those that are small in magnitude. Third, the relative bound can be pessimistic for eigenvalues of small magnitude, as the following example illustrates.

Example 1.1 The mere act of storing a diagonal matrix

$$A = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

in floating point arithmetic produces a perturbed matrix

$$A + E = \begin{pmatrix} \lambda_1(1+\epsilon_1) & & \\ & \ddots & \\ & & \lambda_n(1+\epsilon_n) \end{pmatrix},$$

where $|\epsilon_i| \leq \epsilon$ and $\epsilon > 0$ reflects the machine accuracy. According to the absolute perturbation bound (1.1), the error in an eigenvalue $\hat{\lambda}$ of $A+E$ is bounded by

$$\min_i |\lambda_i - \hat{\lambda}| \leq \|E\| = \max_k |\lambda_k \epsilon_k| \leq \epsilon \max_k |\lambda_k|.$$

This bound is realistic for eigenvalues of largest magnitude: If $\hat{\lambda}$ is closest to an eigenvalue λ_{max} of largest magnitude among all eigenvalues of A, then

$$\frac{|\lambda_{max} - \hat{\lambda}|}{|\lambda_{max}|} \leq \epsilon.$$

Since the relative error in all eigenvalues does not exceed ϵ, the bound is tight in this case.

However the bound is too pessimistic for eigenvalues of smallest magnitude: If $\hat{\lambda}$ is closest to an eigenvalue λ_{min} of smallest magnitude among all eigenvalues of A, then

$$\frac{|\lambda_{min} - \hat{\lambda}|}{|\lambda_{min}|} \leq \epsilon \frac{|\lambda_{max}|}{|\lambda_{min}|}.$$

The bound is much larger than ϵ when the magnitude of the eigenvalues varies widely. Since the relative error does not exceed ϵ, the bound is not tight. □

There are algorithms whose relative error bounds do not depend on the eigenvalues. These algorithms compute all eigenvalues or singular values to high relative accuracy, even those of small magnitude: the dqds algorithm for singular values of bidiagonal matrices (Fernando and Parlett 1994, Parlett 1995), for instance, as well as Jacobi methods for eigenvalues of symmetric positive-definite matrices and for singular values (Demmel 1997, §5.4.3), (Mathias 1995*a*). Absolute perturbation bounds cannot account for this phenomenon.

Absolute error bounds are well suited for describing the accuracy of fixed point arithmetic. But fixed point arithmetic has been replaced by floating point arithmetic, especially on general purpose machines where many eigenvalue and singular value computations are carried out nowadays. The accuracy of floating point arithmetic is best described by relative errors. In the absence of underflow and overflow, a number α is represented as a floating point number

$$\hat{\alpha} \equiv \alpha(1+\epsilon_\alpha), \qquad \text{where} \quad |\epsilon_\alpha| \leq \epsilon,$$

and $\epsilon > 0$ reflects the machine accuracy. In IEEE arithmetic, for instance, $\epsilon \approx 10^{-7}$ in single precision and $\epsilon \approx 10^{-16}$ in double precision. Therefore the accuracy of floating point arithmetic can be described by relative error bounds of the form

$$|\hat{\alpha} - \alpha| \leq |\alpha|\,\epsilon \qquad \text{or} \qquad |\hat{\alpha} - \alpha| \leq |\hat{\alpha}|\,\epsilon.$$

Absolute error bounds cannot model this situation.

And even if you never require high relative accuracy from your small eigenvalues or singular values, you can still profit from it. It turns out that intermediate quantities computed to high relative accuracy can sometimes speed up subsequent computations. For instance, computing eigenvalues of a real, symmetric, tridiagonal matrix to high relative accuracy can accelerate eigenvector computations because the time consuming process of orthogonalising eigenvectors can be shortened or even avoided (Dhillon, Fann and Parlett 1997).

Now that we have established the need for 'genuine' relative error bounds beyond any shadow of a doubt, it's time to find out what kind of relative bounds are out there.

1.2. Overview

Relative error bounds have been derived in the context of two different perturbation models:

- *Additive* perturbations (§§2, 3, 4) represent the perturbed matrix as $A + E$.
- *Multiplicative* perturbations (§§5, 6, 7) represent the perturbed matrix as D_1AD_2, where D_1 and D_2 are non-singular matrices.

The traditional absolute error bounds are derived in the context of additive perturbations.

We group the bounds for eigenvalues (§2, §5) and for singular values (§3, §6) according to a loose order of increasing specialisation:

- Bauer-Fike-type:
  Two-norm bounds on the distance between a perturbed eigenvalue and a closest exact eigenvalue.
- Hoffman-Wielandt-type:
  Frobenius norm bounds on the sum (of squares) of all distances between

perturbed eigenvalues and corresponding exact eigenvalues, where perturbed and exact eigenvalues are paired up in a one-to-one fashion. Similar for singular values.

- Weyl-type:
  Two norm bounds on the largest distance between a perturbed eigenvalue and the corresponding exact eigenvalue, where the ith largest perturbed eigenvalue is paired up with the ith largest exact eigenvalue. Similar for singular values.

There are several different ways to normalise an absolute error $|\lambda-\hat{\lambda}|$ and turn it into a relative error. We present bounds for the following relative error measures

$$\frac{|\lambda-\hat{\lambda}|}{|\lambda|}, \qquad \frac{|\lambda-\hat{\lambda}|}{\sqrt{|\lambda|\,|\hat{\lambda}|}}, \qquad \frac{|\lambda-\hat{\lambda}|}{\sqrt[p]{|\lambda|^p+|\hat{\lambda}|^p}},$$

where $1 \leq p \leq \infty$ is an integer. For instance, the traditional relative error $|\lambda-\hat{\lambda}|/|\lambda|$ can be larger or smaller than the second error measure, while it is never smaller than the third. Detailed relationships among the different measures are discussed in (Li 1994a, §2). Since the measures are essentially proportional to each other we disregard any differences among them.

Sections §§4 and 7 discuss applications of additive and multiplicative perturbations.

1.3. Notation

We use two norms: the two-norm

$$\|A\| = \max_{x\neq 0} \frac{\|Ax\|}{\|x\|}, \qquad \text{where} \quad \|x\| = \sqrt{x^*x},$$

and the superscript $*$ denotes the conjugate transpose; and the Frobenius norm

$$\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2},$$

where a_{ij} are the elements of the matrix A. The identity matrix of order n is

$$I = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} = (\, e_1 \quad \ldots \quad e_n \,)$$

with columns e_i.

For a matrix A we denote by range(A) its column space, A^{-1} its inverse and $A^\dagger$ its Moore-Penrose inverse. The absolute value matrix $|A|$ has elements $|a_{ij}|$. A matrix inequality of the form $|A| \leq |B|$ is meant element-wise, i.e. $|a_{ij}| \leq |b_{ij}|$ for all i and j.

2. Additive Perturbations for Eigenvalues

Let A be a complex square matrix. We want to bound the absolute and relative errors in the eigenvalues of the perturbed matrix $A+E$. In the process we show that relative error bounds are as natural as absolute error bounds, and that many relative bounds are implied by absolute bounds.

2.1. Bauer-Fike-Type Bounds for Diagonalisable Matrices

The Bauer-Fike theorem bounds the distance between an eigenvalue of $A+E$ and a closest eigenvalue of A. The matrix A must be diagonalisable, while $A+E$ does not have to be.

Let $A = X\Lambda X^{-1}$ be an eigendecomposition of A, where

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

and λ_i are the eigenvalues of A. Let $\hat{\lambda}$ be an eigenvalue of $A+E$.

The Bauer-Fike Theorem for the two-norm (Bauer and Fike 1960, Theorem IIIa) bounds the absolute error,

$$\min_i |\lambda_i - \hat{\lambda}| \leq \kappa(X)\, \|E\|. \tag{2.1}$$

The relative version of the Bauer-Fike Theorem below requires in addition that A be non-singular.

Theorem 2.1 If A is diagonalisable and non-singular then

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

where $\kappa(X) \equiv \|X\|\, \|X^{-1}\|$.

Proof. (Eisenstat and Ipsen 1997, Corollary 3.2)

The idea is to 'divide' by the eigenvalues of A and apply the absolute error bound (2.1).

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

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

This means, 1 is an eigenvalue of $\hat{\lambda}A^{-1} - A^{-1}E$. The matrix $\hat{\lambda}A^{-1}$ has the same eigenvector matrix as A and its eigenvalues are $\hat{\lambda}/\lambda_i$. Apply the Bauer-Fike Theorem (2.1) to $\hat{\lambda}A^{-1}$ and to the perturbed matrix $\hat{\lambda}A^{-1} - A^{-1}E$. $\square$

If we interpret the amplifier $\kappa(X)$ as a condition number for the eigenvalues of A then absolute and relative error bounds have the same condition number. This means in the eyes of the Bauer-Fike Theorem an eigenvalue is as sensitive in the absolute sense as it is in the relative sense. A comparison of the absolute bound (2.1) and the relative bound in Theorem 2.1 shows that the absolute error is bounded in terms of the absolute perturbation E, while the relative error is bounded in terms of the relative perturbation $A^{-1}E$.

But $A^{-1}E$ is not the only way to express a relative perturbation. Why keep A^{-1} on the left of E? Why not move it to the right, or distribute it on both sides of E? Splitting $A = A_1 A_2$ and sandwiching E between the two factors, like $A_1^{-1} E A_2^{-1}$, results in undreamt-of possibilities for relative perturbations.

Theorem 2.2 Let A be diagonalisable and non-singular. If $A = A_1 A_2$ where A_1 and A_2 commute then

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \kappa(X)\ \|A_1^{-1} E A_2^{-1}\|.$$

Proof. (Eisenstat and Ipsen 1997, Corollary 3.4)

The idea is to apply Theorem 2.1 to the similarity transformations

$$A_2 A A_2^{-1} \qquad \text{and} \qquad A_2 (A + E) A_2^{-1}.$$

Fortunately, similarity transformations preserve eigenvalues. And the commutativity of A_1 and A_2 prevents the similarity from changing A,

$$A_2\ A\ A_2^{-1} = A_2\ (A_1 A_2)\ A_2^{-1} = A_2 A_1 = A_1 A_2 = A.$$

Therefore we retain the condition number of the original eigenvector matrix X. $\square$

When $A_1 = A$ and $A_2 = I$, Theorem 2.2 reduces to the relative bound in Theorem 2.1. Setting $A_1 = I$ and $A_2 = A$ gives (Eisenstat and Ipsen 1997, Corollary 3.5)

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

This bound includes as a special case (Veselić and Slapničar 1993, Theorem 3.17). Another popular choice for A_1 and A_2 is a square root $A^{1/2}$ of A. In this case Theorem 2.2 gives (Eisenstat and Ipsen 1997, Corollary 3.6)

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \kappa(X)\ \|A^{-1/2} E A^{-1/2}\|.$$

2.2. Bauer-Fike-Type Bounds for Normal Matrices

Normal matrices have unitary eigenvector matrices but, in contrast to Hermitian or real symmetric matrices, their eigenvalues are not necessarily real.

Normal matrices include Hermitian, skew-Hermitian, real symmetric, real skew-symmetric, diagonal, unitary and real orthogonal matrices.

Since the condition number of a unitary eigenvector matrix equals one, the Bauer-Fike theorem applied to a normal matrix A simplifies. The absolute error bound (2.1) becomes

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

while the corresponding relative bound requires again that A be non-singular.

Theorem 2.3 Let A be normal and non-singular. If $A = A_1 A_2$, where A_1 and A_2 commute, then

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \|A_1^{-1} E A_2^{-1}\|.$$

Proof. Follows immediately from Theorem 2.2. □

Therefore eigenvalues of normal matrices are well-conditioned, in the absolute as well as in many relative senses. The relative bound in Theorem 2.3 is tight for diagonal matrices A and component-wise perturbations E, like those in Example 1.1.

For the relative bound to remain in effect, A_1 and A_2 have to commute. Our choices for 'commuting factorisations' have been so far:

$$(A_1, A_2) = (A, I), \qquad (A_1, A_2) = (I, A), \qquad (A_1, A_2) = (A^{1/2}, A^{1/2}).$$

But since A is normal there is another commuting factorisation: the polar factorisation. Every square matrix A has a polar factorisation $A = HU$, where $H \equiv (AA^*)^{1/2}$ is Hermitian positive-semidefinite and U is unitary (Horn and Johnson 1985, Theorem 7.3.2). The matrix H is always unique, while U is only unique when A is non-singular. In particular, when A is Hermitian positive-definite $H = A$ and U is the identity. We use the fact that polar factors of normal non-singular matrices commute in the following sense (Eisenstat and Ipsen 1997, Lemma 4.2)

$$HU = UH = H^{1/2} U H^{1/2}.$$

Theorem 2.4 If A is normal and non-singular, with Hermitian positive-definite polar factor H then

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \|H^{-1/2} E H^{-1/2}\|.$$

Proof. (Eisenstat and Ipsen 1997, Theorem 4.3)

Since $A = H^{1/2} U H^{1/2}$, we can set $A_1 = H^{1/2} U$ and $A_2 = H^{1/2}$ in Theorem 2.3 to get

$$\|A_1^{-1} E A_2^{-1}\| = \|U^* H^{-1/2} E H^{-1/2}\| = \|H^{-1/2} E H^{-1/2}\|.$$

□

Therefore the eigenvalues of a normal matrix have the same relative error bound as the eigenvalues of its positive-definite polar factor. This suggests that the eigenvalues of a normal matrix are as well conditioned as the eigenvalues of its positive-definite polar factor. More generally we conclude that eigenvalues of normal matrices are no more sensitive than eigenvalues of Hermitian positive-definite matrices.

2.3. Hoffman-Wielandt-Type Bounds for Diagonalisable Matrices

The Hoffman-Wielandt theorem establishes a one-to-one pairing between all eigenvalues of A and $A+E$ and bounds the sum of all pairwise distances in the Frobenius norm. This requires not only A but also $A+E$ to be diagonalisable.

Let A and $A+E$ be diagonalisable matrices with eigendecompositions $A = X\Lambda X^{-1}$ and $A+E = \hat{X}\hat{\Lambda}\hat{X}^{-1}$, respectively. The eigenvalues are

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \qquad \hat{\Lambda} = \begin{pmatrix} \hat{\lambda}_1 & & \\ & \ddots & \\ & & \hat{\lambda}_n \end{pmatrix}.$$

The extension of the Hoffman-Wielandt theorem from normal to diagonalisable matrices (Elsner and Friedland 1995, Theorem 3.1) bounds the absolute error,

$$\sqrt{\sum_{i=1}^{n} |\lambda_i - \hat{\lambda}_{\tau(i)}|^2} \leq \kappa(\hat{X})\,\kappa(X)\,\|E\|_F \tag{2.2}$$

for some permutation τ. The condition numbers $\kappa(X)$ and $\kappa(\hat{X})$ are expressed in the two-norm to make the bound tighter, since the two-norm never exceeds the Frobenius norm.

We can obtain a relative Hoffman-Wielandt-type bound from a stronger version of (2.2) that deals with eigenvalues of matrix products. To this end write the perturbed matrix as $AC+E$, where C must have the same eigenvector matrix as $AC+E$. The bound (2.2) is the special case where $C = I$. The eigendecomposition of C is

$$C = \hat{X}\Gamma\hat{X}^{-1}, \qquad \text{where} \qquad \Gamma = \begin{pmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{pmatrix}.$$

The eigendecompositions of A and the perturbed matrix remain the same,

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

The stronger Hoffman-Wielandt-type bound below bounds the sum of squares of absolute errors in the products of the eigenvalues of A and C.

Lemma 2.1 If A, C and $AC + E$ are diagonalisable then there exists a permutation τ such that

$$\sqrt{\sum_{i=1}^{n} |\lambda_i \gamma_{\tau(i)} - \hat{\lambda}_{\tau(i)}|^2} \leq \kappa(\hat{X})\,\kappa(X)\,\|E\|_F.$$

Proof. (Eisenstat and Ipsen 1997, Theorem 6.1) □

Now we are ready for the relative bound. The stronger absolute bound in Lemma 2.1 implies a relative version of the original Hoffman-Wielandt-type bound (2.2), provided A is non-singular.

Theorem 2.5 Let A and $A + E$ be diagonalisable. If A is non-singular then there exists a permutation τ so that

$$\sqrt{\sum_{i=1}^{n} \left(\frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|}{|\lambda_i|}\right)^2} \leq \kappa(\hat{X})\,\kappa(X)\,\|A^{-1}E\|_F.$$

Proof. (Eisenstat and Ipsen 1997, Corollary 6.2)
Since $A^{-1}(A + E) - A^{-1}E = I$ we can set

$$\bar{A} \equiv A^{-1}, \qquad C \equiv A + E, \qquad \bar{E} \equiv -A^{-1}E.$$

Then $\bar{A}$ is diagonalisable with eigenvector matrix X and eigenvalues λ_i^{-1}; C is diagonalisable with eigenvector matrix $\hat{X}$ and eigenvalues $\hat{\lambda}_i$; and $\bar{A}C + \bar{E} = \hat{X}I\hat{X}^{-1}$ is diagonalisable, where the eigenvalues are 1 and one can choose $\hat{X}$ as an eigenvector matrix. Applying Lemma 2.1 to $\bar{A}$, C and $\bar{E}$ gives

$$\sum_{i=1}^{n} |\lambda_i^{-1}\hat{\lambda}_{\tau(i)} - 1|^2 \leq \kappa(\hat{X})^2 \kappa(X)^2 \|A^{-1}E\|_F^2.$$

□

2.4. Hoffman-Wielandt-Type Bounds for Hermitian Matrices

When A and $A+E$ are Hermitian the permutation in the Hoffman-Wielandt Theorem is the identity, provided exact and perturbed eigenvalues are numbered as

$$\lambda_n \leq \ldots \leq \lambda_1, \qquad \hat{\lambda}_n \leq \ldots \leq \hat{\lambda}_1.$$

The Hoffman-Wielandt Theorem for Hermitian matrices (Bhatia 1997, Exercise III.6.15) (Löwner 1934) bounds the absolute error by

$$\sqrt{\sum_{i=1}^{n} |\lambda_i - \hat{\lambda}_i|^2} \leq \|E\|_F. \tag{2.3}$$

The relative bound below requires in addition that A and $A+E$ be positive-definite.

Theorem 2.6 If A and $A+E$ be Hermitian positive-definite then

$$\sqrt{\sum_{i=1}^{n} \frac{|\lambda_i - \hat{\lambda}_i|^2}{\lambda_i \hat{\lambda}_i}} \leq \|A^{-1/2} E A^{-1/2} \,(I + A^{-1/2} E A^{-1/2})^{-1/2}\|_F.$$

Proof. (Li 1994*a*, Theorem 3.2), (Li and Mathias 1997, Proposition 3.4') □

As consequence a small $\|A^{-1/2} E A^{-1/2}\|_F$ guarantees a small eigenvalue error. If one does not mind dealing with majorisation theory one can derive bounds that are stronger than Theorem 2.6 and hold for any unitarily invariant norm (Li and Mathias 1997, Proposition 3.4, (3.19)).

2.5. Weyl-Type Bounds

Weyl's perturbation theorem (Bhatia 1997, Corollary III.2.6) bounds the worst distance between the ith eigenvalues of Hermitian matrices A and $A+E$ in the two-norm,

$$\max_{1 \leq i \leq n} |\lambda_i - \hat{\lambda}_i| \leq \|E\|. \tag{2.4}$$

The absolute bound (2.4) implies a relative bound, provided that A is positive-definite. There is no restriction on E other than being Hermitian.

Theorem 2.7 Let A and $A+E$ be Hermitian. If A is also positive-definite then

$$\max_{1 \leq i \leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \|A^{-1/2} E A^{-1/2}\|.$$

Proof. (Eisenstat and Ipsen 1997, Corollary 5.2), (Mathias 1994, Theorem 2.3)

We reproduce the proof from (Eisenstat and Ipsen 1997) because it explains how the absolute bound (2.4) implies the relative bound. Fix an index i. Let $\hat{x}$ be an eigenvector of $A+E$ associated with $\hat{\lambda}_i$, i.e.

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

Multiplying $(\hat{\lambda}_i I - E)\hat{x} = A\hat{x}$ by $A^{-1/2}$ on both sides gives

$$(\bar{A} + \bar{E})\, z = z,$$

where

$$\bar{A} \equiv \hat{\lambda}_i A^{-1}, \qquad \bar{E} \equiv -A^{-1/2} E A^{-1/2}, \qquad z \equiv A^{1/2} \hat{x}.$$

Hence 1 is an eigenvalue of $\bar{A} + \bar{E}$.

We show that it is actually the $(n-i+1)$st eigenvalue and argue as in the proof of (Eisenstat and Ipsen 1995, Theorem 2.1). Since $\hat{\lambda}_i$ is the ith eigenvalue of $A + E$, 0 must be the ith eigenvalue of

$$(A+E) - \hat{\lambda}_i I = A^{1/2} \left(I - \bar{A} - \bar{E}\right) A^{1/2}.$$

But this is a congruence transformation because square-roots of positive-definite matrices are Hermitian. Congruence transformations preserve the inertia. Hence 0 is the ith eigenvalue of $I - \bar{A} - \bar{E}$, and 1 is the $(n-i+1)$st eigenvalue of $\bar{A} + \bar{E}$.

Applying Weyl's Theorem (2.4) to $\bar{A}$ and $\bar{A} + \bar{E}$ gives

$$\max_{1 \le j \le n} \left| \frac{\hat{\lambda}_i}{\lambda_{n-j+1}} - \mu_j \right| \le \|\bar{E}\| = \|A^{-1/2} E A^{-1/2}\|,$$

where μ_j are the eigenvalues of $\bar{A} + \bar{E}$. When $j = n - i + 1$, then $\mu_j = 1$ and we get the desired bound. $\square$

The following example illustrates what the relative bound in Theorem 2.7 looks like when E is a component-wise relative perturbation.

Example 2.1 (Mathias 1994, pages 6, 7)

Let's first subject a single diagonal element of a Hermitian positive-definite matrix A to a component-wise relative perturbation. Say, a_{jj} is perturbed to $a_{jj}(1+\epsilon)$. The perturbed matrix is $A + E$, where $E = \epsilon e_j e_j^T$ and e_j is the jth column of the identity matrix. Then

$$\begin{aligned} \|A^{-1/2} E A^{-1/2}\| &= |\epsilon| \, \|A^{-1/2} e_j e_j^T A^{-1/2}\| = |\epsilon| \, |e_j^T A^{-1/2} A^{-1/2} e_j| \\ &= |\epsilon| \, (A^{-1})_{jj}, \end{aligned}$$

where $(A^{-1})_{jj}$ is the jth diagonal element of A^{-1} (which is positive since A is positive-definite). The relative error bound in Theorem 2.7 is

$$\max_{1 \le i \le n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \le |\epsilon| \, (A^{-1})_{jj}.$$

This means, a small relative error in a diagonal element of a Hermitian positive-definite matrix causes only a small relative error in the eigenvalues if the corresponding diagonal element of the inverse is not much larger than one.

Next we'll subject a pair of off-diagonal elements to a component-wise relative perturbation. Say, a_{jk} and a_{kj} are perturbed to $a_{jk}(1+\epsilon)$ and $a_{kj}(1+\epsilon)$, respectively. The perturbed matrix is $A+E$, where $E = \epsilon(e_j e_k^T +$

$e_j e_k^T$). In this case we get

$$\|A^{-1/2} E A^{-1/2}\| \leq 2|\epsilon| \sqrt{(A^{-1})_{jj}(A^{-1})_{kk}}.$$

The relative error bound in Theorem 2.7 becomes

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq 2|\epsilon| \sqrt{(A^{-1})_{jj}(A^{-1})_{kk}}.$$

This means, a small relative error in a pair of off-diagonal elements of a Hermitian positive-definite matrix causes only a small relative error in the eigenvalues if the product of the corresponding diagonal elements in the inverse is not much larger than one. □

The bound below, for a different error measure, is similar to the Frobenius norm bound in Theorem 2.6.

Theorem 2.8 If A and $A+E$ are Hermitian positive-definite then

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{\sqrt{\lambda_i \hat{\lambda}_i}} \leq \|A^{-1/2} E A^{-1/2}\ (I + A^{-1/2} E A^{-1/2})^{-1/2}\|.$$

Proof. (Li 1994*a*, Theorem 3.2), (Li and Mathias 1997, Proposition 3.4') □

2.6. Weyl-Type Bounds for More Restrictive Perturbations

It is possible to get a Weyl-type bound for eigenvalues of Hermitian matrices without officially asking for positive-definiteness. The price to be paid, however, is a severe restriction on E to prevent perturbed eigenvalues from switching sign.

Theorem 2.9 Let A and $A+E$ be Hermitian. If $0 < \epsilon_l \leq \epsilon_u$ and

$$\epsilon_l\, x^* A x \leq x^* E x \leq \epsilon_u\, x^* A x \quad \text{for all } x$$

then

$$\epsilon_l\, \lambda_i \leq \hat{\lambda}_i - \lambda_i \leq \epsilon_u\, \lambda_i, \qquad 1 \leq i \leq n.$$

Proof. This is a consequence of (Barlow and Demmel 1990, Lemma 1).

The Minimax Principle for eigenvalues of Hermitian matrices (Bhatia 1997, Corollary III.1.2) implies

$$\lambda_i = \max_{\dim(S)=i} \min_{x\in S} \frac{x^* A x}{x^* x} = \min_{x\in S_0} \frac{x^* A x}{x^* x},$$

for some subspace S_0 of dimension i. Then

$$\hat{\lambda}_i = \max_{\dim(S)=i} \min_{x\in S} \frac{x^*(A+E)x}{x^* x} \geq \min_{x\in S_0} \frac{x^*(A+E)x}{x^* x} = \frac{x_0^*(A+E)x_0}{x_0^* x_0}$$

for some $x_0 \in S_0$. The above expression for λ_i and the assumption imply

$$\frac{x_0^*(A+E)x_0}{x_0^* x_0} \geq \min_{x \in S_0} \frac{x^* A x}{x^* x} + \epsilon_l \frac{x_0^* A x_0}{x_0^* x_0} \geq \lambda_i + \epsilon_l \lambda_i,$$

where we have also used the fact $\epsilon_l > 0$. Hence $\epsilon_l \lambda_i \leq \hat{\lambda}_i - \lambda_i$. The upper bound is proved similarly using the characterisation

$$\lambda_i = \min_{\dim(S)=n-i+1} \max_{x \in S} \frac{x^* A x}{x^* x}.$$

□

Therefore the relative error in the eigenvalues lies in the same interval as the relative perturbation. The relative error bound in Theorem 2.9 implies

$$(1+\epsilon_l)\,\lambda_i \leq \hat{\lambda}_i \leq (1+\epsilon_u)\,\lambda_i,$$

where ϵ_l and ϵ_u are positive. Hence $\hat{\lambda}_i$ has the same sign as λ_i, and $|\hat{\lambda}_i| > |\lambda_i|$. Thus the restriction on the perturbation is strong enough that it not only forces A and $A+E$ to have the same inertia, but it also pushes the perturbed eigenvalues farther from zero than the exact eigenvalues.

The restriction on the perturbation in the following bound is slightly weaker. It uses the polar factor technology from Theorem 2.4.

Theorem 2.10 Let A and $A+E$ be Hermitian. If H is the positive-semidefinite polar factor of A and if for some $0 < \epsilon < 1$

$$|x^* E x| \leq \epsilon\, x^* H x \qquad \text{for all } x$$

then

$$|\lambda_i - \hat{\lambda}_i| \leq \epsilon\, |\lambda_i|.$$

Proof. This is a consequence of (Veselić and Slapničar 1993, Theorem 2.1).

The assumption implies

$$x^*(A - \epsilon H)x \leq x^*(A+E)x \leq x^*(A+\epsilon H)x.$$

If $A = X \Lambda X^*$ is an eigendecomposition of A then, because A is Hermitian, the polar factor is $H = X|\Lambda|X^*$, where $|\Lambda|$ is the matrix whose elements are the absolute values of Λ. Hence A and H have the same eigenvectors. A min-max argument as in the proof of Theorem 2.9 establishes

$$\lambda_i - \epsilon|\lambda_i| \leq \hat{\lambda}_i \leq \lambda_i + \epsilon|\lambda_i|.$$

□

Therefore the relative error in the eigenvalues is small if the relative perturbation with regard to the polar factor is small. Since $0 < \epsilon < 1$ the relative error bound implies that $\hat{\lambda}_i$ has the same sign as λ_i. Hence the assumptions in Theorem 2.10 ensure that $A+E$ has the same inertia as A.

Theorem 2.10 applies to component-wise relative perturbations. Matrix inequalities below of the form $|E| \leq \epsilon |A|$ are to be interpreted element-wise.

Corollary 2.1 Let A and $A+E$ be Hermitian, and $|E| \leq \epsilon\, |A|$ for some $\epsilon > 0$. If H is the positive-definite polar factor of A and for some $\eta > 0$

$$\epsilon\, |x|^*|A||x| \leq \eta\, x^*Hx \qquad \text{for all } x$$

then

$$|\lambda_i - \hat{\lambda}_i| \leq \eta\, |\lambda_i|.$$

Proof. This is a consequence of (Veselić and Slapničar 1993, Theorem 2.11).

We merely need to verify that the assumptions of Theorem 2.10 hold,

$$|x^*Ex| \leq |x|^*|E||x| \leq \epsilon\, |x|^*|A||x| \leq \eta\, x^*Hx.$$

□

2.7. Congruence Transformations for Positive-Definite Matrices

All the bounds we have presented so far for positive-definite matrices contain the term $A^{-1/2}EA^{-1/2}$. Since A is Hermitian positive-definite, $A^{1/2}$ is Hermitian, which makes $A^{-1/2}EA^{-1/2}$ Hermitian. This in turn implies that the two-norm and Frobenius norm of $A^{-1/2}EA^{-1/2}$ are invariant under congruence transformations. We say that two square matrices A and M are *congruent* if $A = D^*MD$ for some non-singular matrix D. If A is Hermitian positive-definite, so is M because congruence transformations preserve the inertia.

We start out by showing that the bound in Theorem 2.7 is invariant under congruence transformations.

Corollary 2.2 Let A and be Hermitian positive-definite and $A+E$ Hermitian. If $A = D^*MD$ and $A+E = D^*(M+F)D$, where D is non-singular, then

$$\max_{1 \leq i \leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \|M^{-1/2}FM^{-1/2}\|.$$

Proof. (Mathias 1994, Theorem 2.4)

Start with the bound in Theorem 2.7,

$$\max_{1 \leq i \leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \|A^{-1/2}EA^{-1/2}\|.$$

Positive-definiteness is essential here. Since A is Hermitian positive-definite, it has a Hermitian square-root $A^{1/2}$. Hence $A^{-1/2}EA^{-1/2}$ is Hermitian. This implies that the norm is an eigenvalue,

$$\|A^{-1/2}EA^{-1/2}\| = \max_{1 \leq j \leq n} |\lambda_j(A^{-1/2}EA^{-1/2})|.$$

Now comes the trick. Since eigenvalues are preserved under similarity transformations, we can reorder the matrices in a circular fashion until all grading matrices have cancelled each other out,

$$\begin{aligned}\lambda_j(A^{-1/2}EA^{-1/2}) &= \lambda_j(A^{-1}E) = \lambda_j(D^{-1}\, M^{-1}F\, D) = \lambda_j(M^{-1}F)\\ &= \lambda_j(M^{-1/2}FM^{-1/2}).\end{aligned}$$

At last recover the norm,

$$\max_{1\leq j\leq n} |\lambda_j(M^{-1/2}FM^{-1/2})| = \|M^{-1/2}FM^{-1/2}\|.$$

□

Corollary 2.2 extends (Barlow and Demmel 1990, Theorem 2) and (Demmel and Veselič 1992, Theorem 2.3) to a larger class of matrices. It suggests that the eigenvalues of $A+E$ have the same error bound as the eigenvalues of $M+F$. We can interpret this to mean that the eigenvalues of a Hermitian positive-definite matrix behave as well as the eigenvalues of any matrix congruent to A. The example below illustrates this.

Example 2.2 (Demmel and Veselič 1992, page 1211)
The matrix

$$A = \begin{pmatrix} 10^{40} & 10^{29} & 10^{19} \\ 10^{29} & 10^{20} & 10^{9} \\ 10^{19} & 10^{9} & 1 \end{pmatrix}$$

is symmetric positive-definite with eigenvalues (to six decimal places)

$$1.00000\cdot 10^{40}, \qquad 9.90000\cdot 10^{19}, \qquad 9.81818\cdot 10^{-1}.$$

If we write $A = DMD$, where

$$M = \begin{pmatrix} 1 & .1 & .1 \\ .1 & 1 & .1 \\ .1 & .1 & 1 \end{pmatrix}, \qquad D = \begin{pmatrix} 10^{20} & & \\ & 10^{10} & \\ & & 1 \end{pmatrix},$$

then the eigenvalues of M are (to six decimal places)

$$9.00000\cdot 10^{-1}, \qquad 9.00000\cdot 10^{-1}, \qquad 1.20000.$$

Corollary 2.2 implies that the widely varying eigenvalues of A, and in particular the very small ones, are as impervious to changes in M as the uniformly sized eigenvalues of M.

Already thirty years ago structural engineers considered congruence transformations like the one above where D is diagonal and all diagonal elements of M are equal to one (Rosanoff, Glouderman and Levy 1968, pages 1041, 1050). They observed that such an equilibration 'reduce[s] the ratio of extreme eigenvalues' (Rosanoff et al. 1968, page 1045), and that 'equilibration is of major importance in measurement of matrix conditioning' (Rosanoff et al. 1968, page 1059). □

From the circular reordering argument in the proof of Corollary 2.2 it also follows that the other bounds for positive-definite matrices are invariant under congruences. One bound is Theorem 2.8.

Corollary 2.3 Let A and $A+E$ be Hermitian positive-definite. If $A = D^*MD$ and $A+E = D^*(M+F)D$, where D is non-singular then

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{\sqrt{\lambda_i \hat{\lambda}_i}} \leq \|M^{-1/2}FM^{-1/2}\,(I + M^{-1/2}FM^{-1/2})^{-1/2}\|.$$

Proof. (Li 1994a, Theorem 3.2), (Li and Mathias 1997, Proposition 3.4') □

The other bound that is also invariant under congruences is the Frobenius norm bound Theorem 2.6.

Corollary 2.4 Let A and $A+E$ be Hermitian positive-definite. If $A = D^*MD$ and $A+E = D^*(M+F)D$, where D is non-singular then

$$\sqrt{\sum_{i=1}^{n} \frac{|\lambda_i - \hat{\lambda}_i|^2}{\lambda_i \hat{\lambda}_i}} \leq \|M^{-1/2}FM^{-1/2}\,(I + M^{-1/2}FM^{-1/2})^{-1/2}\|_F.$$

Proof. (Li 1994a, Theorem 3.2), (Li and Mathias 1997, Proposition 3.4')
Since the Frobenius norm sums up squares of eigenvalues, the bound from Theorem 2.8 can be written as

$$\|A^{-1/2}EA^{-1/2}\,(I + A^{-1/2}EA^{-1/2})^{-1/2}\|_F^2 = \sum_{i=1}^{n} \frac{\mu_i^2}{1+\mu_i}$$

where μ_i are the eigenvalues of the Hermitian matrix $A^{-1/2}EA^{-1/2}$. The circular reordering argument from the proof of Corollary 2.2 implies that μ_i are also the eigenvalues of $M^{-1/2}FM^{-1/2}$. □

One may wonder what's so interesting about congruence transformations. One can use congruence transformations to pull the grading out of a matrix (Barlow and Demmel 1990, §2), (Demmel and Veselič 1992, §1, §2.1), (Mathias 1995a). Consider the matrix A in Example 2.2. It has elements of widely varying magnitude that decrease from top to bottom. The diagonal matrix D removes the grading and produces a matrix M, where $M \equiv D^{-1}AD^{-1}$, all of whose elements have about the same order of magnitude and all of whose eigenvalues are of about the same size.

More generally we say that a Hermitian positive-definite matrix A is *graded*, or *scaled*, if $A = DMD^*$ and the eigenvalues of M vary much less in magnitude than the eigenvalues of A (Mathias 1995a, §1).

2.8. Congruence Transformations for Indefinite Matrices

Since the application of congruence transformations is not restricted to Hermitian positive-definite matrices, we may as well try to find out whether indefinite matrices are invariant under congruences. It turns out that the resulting error bounds are weaker than the ones for positive-definite matrices because they require stronger assumptions.

If we are a little sneaky (by extracting the congruence from the polar factor rather than the matrix proper) then the bound for normal matrices in Theorem 2.4 becomes invariant under congruences.

Corollary 2.5 Let A be normal and non-singular, with Hermitian positive-definite polar factor H. If D is non-singular and

$$E = DE_1D^*, \qquad H = DM_1D^*$$

then

$$\min_i \frac{|\lambda_i - \hat{\lambda}|}{|\lambda_i|} \leq \|M_1^{-1/2} E_1 M_1^{-1/2}\|.$$

Proof. (Eisenstat and Ipsen 1997, Corollary 4.4) □

This means the error bound for eigenvalues of a normal matrix is the same as the error bound for eigenvalues of the best scaled version of its positive-definite polar factor.

Let's return to Weyl-type bounds, but now under the condition that the congruence transformation is real diagonal. Theorem 2.10 leads to a bound that is essentially scaling invariant. It is similar to the one above, in the sense that the scaling matrix is extracted from its positive-definite polar factor. However now the perturbations are restricted to be component-wise relative.

Corollary 2.6 Let $A = DMD$ be non-singular Hermitian, where D is diagonal with positive diagonal elements, and let $H = DM_1D$ be the positive-definite polar factor of A. If $A + E$ is Hermitian and $|E| \leq \epsilon\,|A|$ for some $\epsilon > 0$ then

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \epsilon \|\,|M|\,\|\;\|M_1^{-1}\|.$$

Proof. This is a consequence of (Veselić and Slapničar 1993, Theorem 2.13).

We use variational inequalities to show that the assumptions of Corollary 2.1 are fulfilled. Since D is positive-definite,

$$|x|^*|A|\,|x| = |x|^*D\,|M|\,D|x| \leq \|\,|M|\,\|\, x^*D^2x \qquad \text{for all } x.$$

Variational inequalities imply

$$x^*D^2x \leq \|M_1^{-1}\|\, x^*Hx.$$

Therefore

$$\epsilon\,|x|^*|A|\,|x| \leq \eta\, x^*Hx \qquad \text{for all } x,$$

where $\eta \equiv \epsilon\, \|\,|M|\,\|\; \|M_1^{-1}\|$. Now apply Corollary 2.1. $\square$

The amplifier in the bound, $\|\,|M|\,\|\; \|M_1^{-1}\|$, is almost like the condition number of the absolute value matrix $|M|$ or almost like the condition number of the scaled polar factor M_1. Therefore the relative error in the eigenvalues of a Hermitian matrix is small if the polar factor and the absolute value of the matrix are well-scaled.

The following bound is similar in the sense that it applies to a column scaling of a Hermitian matrix $A = MD$. In contrast to Corollary 2.6, however, the scaled matrix M is in general not Hermitian anymore; and the inverse of the scaled matrix, M^{-1}, now appears in the bound rather than the inverse of the scaled polar factor, M_1^{-1}.

Corollary 2.7 Let $A = MD$ be non-singular Hermitian and D diagonal with positive diagonal elements. If $A + E$ is Hermitian, and $|E| \leq \epsilon\,|A|$ for some $\epsilon > 0$ then

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \epsilon \|\,|M|\,\|\; \|M^{-1}\|.$$

Proof. (Veselić and Slapničar 1993, Theorem 3.16)

First we take care of the scaling matrix D. The technology of previous proofs requires that D appear on both sides of the matrix. That's why we consider $A^2 = A^*A = DM^*MD$. The component-wise perturbation implies

$$|x^*E^2x| \leq \epsilon^2\,|x|^*|A|^2\,|x|.$$

Proceeding as in the proof of Corollary 2.6 gives

$$|x|^*|A|^2\,|x| \leq \|\,|M|\,\|^2\; \|M^{-1}\|^2\; x^*A^2x \qquad \text{for all } x.$$

Hence

$$|x^*E^2x| \leq \eta^2\; x^*A^2x,$$

where $\eta \equiv \epsilon\,\|\,|M|\,\|\; \|M_1^{-1}\|$. Now that we got rid of D, we need to undo the squares. In order to take the positive square-root without losing the monotonicity we need positive-definite matrices under the squares. Polar factors do the job.

If H and H_E are the Hermitian positive-definite polar factors of A and E, respectively then

$$x^*E^2x = x^*H_E^2x, \qquad x^*A^2x = x^*H^2x.$$

Therefore

$$x^*H_E^2x \leq \eta^2\; x^*H^2x \qquad \text{for all } x.$$

Now comes the trick. Because H_E and H are Hermitian positive-definite we can apply the fact that the square-root is operator-monotone (Bhatia 1997, Proposition V.I.8) and conclude

$$x^* H_E x \leq \eta \, x^* H x.$$

Since $|x^* E x| \leq x^* H_E x$, Theorem 2.10 applies. □

The next bound, the last one in this section, applies to general perturbations. Compared to other bounds it severely constrains the size of the perturbation by forcing it to be smaller than any eigenvalue of any principal submatrix.

Theorem 2.11 Let $A = DMD$ and $A + E = D(M+F)D$ be real, symmetric and D a real non-singular diagonal matrix. Among all eigenvalues of principal submatrices of M, let μ be the smallest in magnitude. If $\|F\| < |\mu|$ then

$$-\|F\| \frac{2|\mu| - \|F\|}{|\mu|^2} \leq \frac{\hat{\lambda}_i - \lambda_i}{\lambda_i} \leq \|F\| \, \frac{2|\mu| - \|F\|}{(|\mu| - \|F\|)^2}, \qquad 1 \leq i \leq n.$$

Proof. (Gu and Eisenstat 1993, Corollary 5) □

2.9. Ritz Values

Ritz values are 'optimal' approximations to eigenvalues of Hermitian matrices.

Let A be a Hermitian matrix of order n and Q a matrix with m orthonormal columns. Then $W \equiv Q^* A Q$ is a matrix of order m whose eigenvalues

$$\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_m,$$

are called *Ritz values* of A (Parlett 1980, §11.3). The corresponding residual is $R \equiv AQ - QW$. Ritz values are optimal in the following sense. Given Q, the norm of R can only increase if we replace W by another matrix, i.e. (Parlett 1980, Theorem 11-4-5),

$$\|R\| = \|AQ - QW\| \leq \|AQ - QC\|$$

for all matrices C of order m.

Moreover one can always find m eigenvalues of A that are within absolute distance $\|R\|$ of the Ritz values (Parlett 1980, Theorem 11-5-1),

$$\max_{1 \leq j \leq m} |\lambda_{\tau(j)} - \hat{\lambda}_j| \leq \|R\|$$

for some permutation τ.

Unfortunately the corresponding relative error bounds are not as simple. They are expressed in terms of angles between the relevant subspaces range(Q), range(AQ), and range($A^{-1}Q$). Let $0 \leq \theta_1 \leq \pi/2$ be the maximal principal angle between range(Q) and range(AQ), and $0 \leq \theta_2 \leq \pi/2$ be the maximal principal angle between range(AQ) and range($A^{-1}Q$).

Theorem 2.12 If A is non-singular Hermitian then there exists a permutation τ so that

$$\max_{1 \leq i \leq m} \frac{|\lambda_{\tau(i)} - \hat{\lambda}_i|}{|\lambda_{\tau(i)}|} \leq \sin\theta_1 + \tan\theta_2.$$

Proof. (Drmač 1996a, Theorem 3, Proposition 5)

In order to exhibit the connection to previous results we sketch the idea for the proof. First express the Ritz values as an additive perturbation. To this end define the Hermitian perturbation

$$E \equiv -(RQ^* + QR^*).$$

Then Q is an invariant subspace of $A + E$,

$$(A + E)Q = QW,$$

and the eigenvalues of W are eigenvalues of $A + E$.

Now proceed as in the proof of Corollary 2.7 and look at the squares,

$$x^*E^2x = x^*A^*\ A^{-*}E^*EA^{-1}\ Ax \leq \|EA^{-1}\|^2 x^*A^2x \qquad \text{for all } x.$$

Undo the squares using polar factors and the operator-monotonicity of the square root, and apply Theorem 2.10. Hence the eigenvalues $\mu_1 \geq \ldots \geq \mu_n$ of $A + E$ satisfy

$$\max_{1 \leq i \leq n} \frac{|\lambda_i - \mu_i|}{|\lambda_i|} \leq \|EA^{-1}\|.$$

Let $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_m$ be those μ_i that are also eigenvalues of W; and let τ be a permutation that numbers the eigenvalues of A corresponding to μ_i first. Then

$$\max_{1 \leq i \leq m} \frac{|\lambda_{\tau(i)} - \hat{\lambda}_i|}{|\lambda_{\tau(i)}|} \leq \|EA^{-1}\|.$$

We still have to worry about $\|EA^{-1}\|$. Write

$$-EA^{-1} = (I - QQ^*)AQQ^*A^{-1} + QQ^*(I - AQQ^*A^{-1}).$$

Here QQ^* is the orthogonal projector onto range(Q), while AQQ^*A^{-1} is the oblique projector onto range(AQ) along range(Q^*A^{-1}). This expression for EA^{-1} appears in (Drmač 1996a, Theorem 3). It can be bounded above by $\sin\theta_1 + \tan\theta_2$. □

Therefore the relative error in the Ritz values of $W = Q^*AQ$ is small if both subspace angles θ_1 and θ_2 are small. Things simplify when the matrix A is also positive-definite because there is only one angle to deal with.

Let A be Hermitian positive-definite with Cholesky factorisation $A = LL^*$. Let $0 \leq \theta \leq \pi/2$ be the maximal principal angle between range(L^*Q) and range$(L^{-1}Q)$.

Theorem 2.13 If A is Hermitian positive-definite and if $\sin\theta < 1$ then there exists a permutation τ so that

$$\max_{1\leq i\leq m} \frac{|\lambda_{\tau(i)} - \hat{\lambda}_i|}{|\lambda_{\tau(i)}|} \leq \frac{\sin\theta}{1-\sin\theta}.$$

Proof. (Drmač 1996*a*, Theorem 6) □

Theorem 2.13 can be extended to semi-definite matrices (Drmač and Hari 1997).

3. Additive Perturbations for Singular Values

Let B be a complex matrix. We want to estimate the absolute and the relative errors in the singular values of the perturbed matrix $B + F$. For definiteness we assume that B is tall and skinny, i.e. B is $m \times n$ with $m \geq n$ (if this is not the case just consider B^*).

Perturbation bounds for singular values are usually derived by first converting the singular value problem to an Hermitian eigenvalue problem.

3.1. Converting Singular Values to Eigenvalues

The singular value decomposition of a $m \times n$ matrix B, $m \geq n$, is

$$B = U \begin{pmatrix} \Sigma \\ 0 \end{pmatrix} V^*,$$

where the left singular vector matrix U and the right singular vector matrix V are unitary matrices of order m and n, respectively. The non-negative diagonal matrix Σ of order n contains the singular values σ_i of B,

$$\Sigma = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{pmatrix},$$

where

$$\sigma_1 \geq \ldots \geq \sigma_n \geq 0.$$

There are two popular ways to convert a singular value problem to an eigenvalue problem.

- The eigenvalues of

$$A \equiv \begin{array}{c c} & \begin{array}{cc} m & n \end{array} \\ \begin{array}{c} m \\ n \end{array} & \begin{pmatrix} 0 & B \\ B^* & 0 \end{pmatrix} \end{array}$$

are

$$\sigma_1, \ldots, \sigma_n, -\sigma_1, \ldots, -\sigma_n, \underbrace{0, \ldots, 0}_{m-n}.$$

Therefore the singular values of B are the n largest eigenvalues of A (Horn and Johnson 1985, Theorem 7.3.7).

- The eigenvalues of B^*B are

$$\sigma_1^2, \ldots, \sigma_n^2.$$

Therefore the singular values of B are the positive square-roots of the eigenvalues of B^*B (Horn and Johnson 1985, Lemma 7.3.1).

Since singular values are eigenvalues of a Hermitian matrix, they are well-conditioned in the absolute sense.

3.2. Hoffman-Wielandt-Type Bounds

We bound the sum of squares of all distances between the ith exact and perturbed singular values in terms of the Frobenius norm.

The singular values of B and $B+F$ are, respectively,

$$\sigma_1 \geq \ldots \geq \sigma_n \geq 0, \qquad \hat{\sigma}_1 \geq \ldots \geq \hat{\sigma}_n \geq 0.$$

Converting the singular value problem to an eigenvalue problem à la §3.1 and applying the Hoffman-Wielandt Theorem for Hermitian matrices (2.3) leads immediately to the absolute error bound

$$\sqrt{\sum_{i=1}^{n} |\sigma_i - \hat{\sigma}_i|^2} \leq \|F\|_F.$$

The relative bound below requires in addition that both matrices be non-singular.

Theorem 3.1 Let B and $B+F$ be non-singular. If $\|FB^{-1}\| < 1$ then

$$\sqrt{\sum_{i=1}^{n} \frac{|\sigma_i - \hat{\sigma}_i|^2}{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2} \|(I + FB^{-1})^* - (I + FB^{-1})^{-1}\|_F.$$

Therefore the error in the singular values is small if $I + FB^{-1}$ is close to being unitary (or orthogonal). This is case when $B + F = (I + FB^{-1})\, B$ is more or less a unitary transformation away from B.

Proof. (Li 1994*a*, Theorem 4.3). □

3.3. Weyl-Type Bounds

We bound the worst-case distance between the ith exact and perturbed singular values in terms of the two-norm.

The absolute error bound is an immediate consequence of Weyl's Perturbation Theorem (2.4)

$$|\sigma_i - \hat{\sigma}_i| \le \|F\|, \qquad 1 \le i \le n.$$

The corresponding relative bound below restricts the range of F but not its size. Here $B^\dagger$ is the Moore-Penrose inverse of B.

Theorem 3.2 If range$(B+F) \subset$ range(B) then

$$|\sigma_i - \hat{\sigma}_i| \le \sigma_i \, \|B^\dagger F\|, \qquad 1 \le i \le n.$$

If range$((B+F)^*) \subset$ range(B^*) then

$$|\sigma_i - \hat{\sigma}_i| \le \sigma_i \, \|F B^\dagger\|, \qquad 1 \le i \le n.$$

Proof. (Di Lena, Peluso and Piazza 1993, Theorem 1.1)

We prove the first bound, the proof for the second one is similar.

Life would be easy if we could pull B out of F, say if $F = BC$ for some matrix C. Then we could write $B + F = B(I + C)$ and apply the sum and product inequalities for singular values (Horn and Johnson 1985, page 423) to get the relative bound

$$|\sigma_i - \hat{\sigma}_i| \le \sigma_i \, \|C\|.$$

It turns out that the range condition is exactly what is needed to pull B out of F. This is because range$(B+F) \subset$ range(B) implies $F = BC_1$ for some C_1. This allows us to write

$$F = BC_1 = BB^\dagger B\, C_1 = BB^\dagger F.$$

Consequently, setting $C \equiv B^\dagger F$ gives the desired result. □

When B has full column rank the second range condition in Theorem 3.2 is automatically satisfied.

Corollary 3.1 If B has full column rank then

$$|\sigma_i - \hat{\sigma}_i| \le \sigma_i \, \|F B^\dagger\|, \qquad 1 \le i \le n.$$

Proof. (Di Lena et al. 1993, Remark 1.1)

If B has full column rank n then its rows span n-space. Hence range$((B+F)^*) \subset$ range(B^*) for any F, and the second relative bound in Theorem 3.2 holds. □

Therefore singular values of full-rank matrices are well-conditioned in the absolute as well as relative sense. This may sound implausible at first, in particular when B has full rank while $B+F$ is rank-deficient. In this case all singular values of B are non-zero while at least one singular value of $B+F$ is zero. Hence a zero singular value of $B+F$ must have relative error equal to one. How can the singular values of B be well-conditioned? The answer is that in this case the relative perturbation $\|B^\dagger F\|$ is large. The following example illustrates that $\|B^\dagger F\|$ is large when B and $B+F$ differ in rank.

Example 3.1 Let

$$B = \begin{pmatrix} 2 & 0 \\ 0 & \eta \\ 0 & 0 \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & 0 \\ 0 & -\eta \\ 0 & 0 \end{pmatrix},$$

where $\eta \neq 0$. The rank of B is two, while the rank of $B+F$ is one.

The relative error in the singular value $\hat{\sigma} = 0$ of $B+F$ is equal to one because $|\eta - 0|/|\eta| = 1$. Since

$$B^\dagger = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{\eta} & 0 \end{pmatrix}, \qquad B^\dagger F = \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix},$$

we get $\|B^\dagger F\| = 1$. Corollary 3.1 gives

$$\min_i \frac{|\sigma_i - \hat{\sigma}|}{\sigma_i} \leq 1.$$

Therefore Corollary 3.1 is tight for the zero singular values of $B+F$. □

Corollary 3.1 extends (Demmel and Veselić 1992, Lemma 2.12) to matrices that do not necessarily have full rank (Di Lena et al. 1993, Remark 1.2). When B is non-singular Corollary 3.1 implies that both range conditions in Theorem 3.2 hold automatically.

Corollary 3.2 If B is non-singular then

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \leq \min\{\|B^{-1}F\|, \|FB^{-1}\|\}.$$

The following bound, for a different error measure, is similar to the Frobenius norm bound Theorem 3.1. It requires that both B and $B+F$ be non-singular.

Theorem 3.3 Let B and $B+F$ be non-singular. Then

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2} \, \|(I + FB^{-1})^* - (I + FB^{-1})^{-1}\|.$$

Proof. (Li 1994*a*, Theorem 4.3). □

As in Theorem 3.1, the error in the singular values is small if $I + FB^{-1}$ is close to being unitary (or orthogonal). This is case when $B + F = (I + FB^{-1})\, B$ is more or less a unitary transformation away from B.

3.4. Congruence Transformations

We start with one-sided grading of matrices B with full-column rank. That is, $B = CD$ or $B = DC$, where D is non-singular. In the event D is diagonal, CD represents a column scaling while DC represents a row scaling.

All relative singular value bounds presented so far are invariant under one-sided grading from the appropriate side. That's because the bounds contain terms of the form $B^\dagger F$ or $FB^\dagger$. Consider $B^\dagger F$, for instance. Grading from the right is sandwiched in the middle, between $B^\dagger$ and F, and therefore cancels out.

Let's first look at the Hoffman-Wielandt-type bound for graded matrices, which follows directly from Theorem 3.1.

Corollary 3.3 Let $B = CD$ and $B + F = (C + G)D$ be non-singular. If $\|GC^{-1}\| < 1$ then

$$\sqrt{\sum_{i=1}^{n} \frac{|\sigma_i - \hat{\sigma}_i|^2}{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2} \, \|(I + GC^{-1})^* - (I + GC^{-1})^{-1}\|_F.$$

Proof. (Li 1994*a*, Theorem 4.3)

The grading is sandwiched in the middle of the relative perturbation FB^{-1}, and cancels out,

$$FB^{-1} = (GD)\,(CD)^{-1} = G\,DD^{-1}\,C^{-1} = GC^{-1}.$$

□

Moving right along to the two-norm, we see that Corollary 3.1 is invariant under grading from the right.

Corollary 3.4 If $B = CD$ has full column rank, and if $B + F = (C + G)D$ then

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \, \|GC^\dagger\|, \qquad 1 \leq i \leq n.$$

Proof. (Di Lena et al. 1993, Remark 1.2)

Full column rank is needed to extract the grading matrix from the inverse,

$$B^\dagger = (CD)^\dagger = D^\dagger C^\dagger = D^{-1} C^\dagger.$$

□

Therefore the relative error in the singular values of $B + F$ is small if there is a grading matrix D that causes the relative perturbation of the graded matrices $\|GC^\dagger\|$ to be small. For instance, suppose $B = CD$ has columns whose norms vary widely while the columns of C are almost orthonormal. If the perturbed matrix is scaled in the same way then the error bound in Corollary 3.4 ignores the scaling and acts as if it saw the well-behaved matrices C and $C+G$. Corollary 3.4 extends (Demmel and Veselić 1992, Theorem 2.14) to a larger class of matrices.

The other two-norm bound, Theorem 3.3, is also invariant under grading.

Corollary 3.5 Let $B = CD$ and $B + F = (C+G)D$ be non-singular. If $\|GC^{-1}\| < 1$ then

$$\max_{1\le i\le n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt{\sigma_i \hat{\sigma}_i}} \le \frac{1}{2}\,\|(I+GC^{-1})^* - (I+GC^{-1})^{-1}\|.$$

Proof. (Li 1994*a*, Theorem 4.3) □

Finally we present the only bound that is invariant under grading from both sides. It requires that the grading matrices be real diagonal; and it restricts the size of the perturbation more severely than the other bounds.

Theorem 3.4 Let $B = D_l C D_r$ and $B + F = D_l(C+G)D_r$ be real symmetric, where D_l and D_r are real non-singular diagonal matrices. Among the singular values of all square submatrices of B, let θ be the smallest one. If $\|G\| < \theta$ then

$$-\|G\|\,\frac{2\theta - \|G\|}{\theta^2} \le \frac{\hat{\sigma}_i - \sigma_i}{\sigma_i} \le \|G\|\,\frac{2\theta - \|G\|}{(\theta - \|G\|)^2} \qquad 1 \le i \le n.$$

Proof. (Gu and Eisenstat 1993, Corollary 10) □

4. Some Applications of Additive Perturbations

We discuss Jacobi's method for computing singular values and eigenvalues, and deflation of triangular and bidiagonal matrices.

4.1. Jacobi's Method for Singular Values

Jacobi's method is generally viewed as a method that computes eigenvalues and singular values to optimal accuracy. It was Jacobi's method that first attracted attention to invariance of eigenvalue and singular values error bounds under congruence (Demmel and Veselić 1992, Mathias 1995*a*, Rosanoff et al. 1968). We give a very intuitive plausibility argument, shoving many subtleties under the rug, to explain the high accuracy and invariance under grading of Jacobi's method. Our discussion runs along the lines of (Demmel 1997, §5.4.3) and (Mathias 1995*a*, §2, 3). Other detailed accounts can be found in (Demmel and Veselić 1992), (Drmač 1996*b*). An attempt at a geometric interpretion of Jacobi's high accuracy is made in (Rosanoff et al. 1968, pages 1045-6).

A one-sided Jacobi method computes the singular values of a tall and skinny matrix by applying a sequence of orthogonal transformations on the right side of the matrix. The duty of each orthogonal transformation is to orthogonalise two columns of the matrix. The method stops once all columns are sufficiently orthogonal to each other. At this point the singular values are approximated by the column norms, i.e. the Euclidian lengths of the columns.

For simplicity assume that B is a real non-singular matrix of order n. Let D be a row scaling of B, i.e. $B = DC$, where D is diagonal. We show that the one-sided Jacobi method ignores the row scaling. When Jacobi applies an orthogonal transformation Q to B the outcome in floating point arithmetic is $BQ + F$. Corollary 3.2 implies that the singular values $\hat{\sigma}_i$ of $BQ + F$ satisfy

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \leq \|C^{-1}G\| \leq \|C^{-1}\| \, \|G\|,$$

where $F = DG$.

Let's bound the squared error $\|G\|$. Round-off error analysis tells us that the error in the ith row is

$$\|e_i^T F\| \leq \epsilon\beta_i \, \|e_i^T B\| + O(\epsilon^2),$$

where β_i depends on the matrix size n and $\epsilon > 0$ reflects the machine accuracy. Now the crucial observation is that the orthogonal transformations happen on one side of the matrix and the scaling on the other side. Because Q operates on columns it does not mix up different rows and therefore preserves the row-scaling. This means we can pull the ith diagonal element of D out of $e_i^T F$,

$$\|e_i^T F\| \leq \epsilon\beta_i \, \|e_i^T B\| = \epsilon\beta_i \, \|e_i^T (DC)\| = |d_{ii}| \, \epsilon\beta_i \, \|e_i^T C\|.$$

This gives a bound for the ith row of G,

$$\|e_i^T G\| = |d_{ii}|^{-1} \, \|e_i^T F\| \leq \epsilon\beta_i \, \|e_i^T C\| + O(\epsilon^2).$$

The total error is therefore bounded by

$$\|G\| \leq \epsilon\beta \, \|C\| + O(\epsilon^2),$$

where β depends on n. Therefore the error bound for the singular values of $BQ + F$ is independent of the row-scaling,

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \leq \epsilon\beta \, \kappa(C) + O(\epsilon^2).$$

This means Jacobi's method produces singular values of B but acts as if it saw C instead. That's good, particularly if D manages to pull out all the grading. Then all singular values of C have about the same magnitude and $\kappa(C)$ is close to one. Therefore the above bound $\kappa(C)\,\beta\epsilon$ tends to be on the order of machine accuracy ϵ, implying that the relative error in the singular values is on the order of machine accuracy.

The argument is more complicated when the orthogonal transformations are applied on the same side as the scaling matrix. Fortunately the resulting error bounds do not tend to be much weaker (Mathias 1995*a*, §4).

4.2. Jacobi's Method for Eigenvalues

A two-sided Jacobi method computes eigenvalues of a real symmetric positive-definite matrix by applying a sequence of orthogonal similarity transformations to the matrix. An orthogonal similarity transformation operates on two rows, i and j, and two columns, i and j, to zero out elements (i,j) and (j,i). The method stops once all off-diagonal elements are sufficiently small. At this point the eigenvalues are approximated by the diagonal elements.

Let A be real symmetric positive-definite of order n, and $A = DMD$, where D is a non-singular diagonal matrix. The Jacobi method computes the eigenvalues of a matrix $A+E$. According to Corollary 2.2, the error bound for the eigenvalues $\hat{\lambda}_i$ of $A+E$ is

$$\max_{1\le i\le n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \le \|M^{-1/2} F M^{-1/2}\| \le \|M^{-1}\|\,\|F\|, \tag{4.1}$$

where $E = DFD$. One can show that the error is bounded by

$$\|F\| \le \alpha\epsilon\,\|M\| + O(\epsilon^2),$$

where α depends on n and $\epsilon > 0$ reflects the machine accuracy. Therefore

$$\max_{1\le i\le n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \le \alpha\epsilon\,\kappa(M) + O(\epsilon^2).$$

This means, the relative error in the eigenvalues is small, provided the amplifier $\kappa(M)$ is small.

The amplifier $\kappa(M)$ can be minimised via an appropriate choice of the scaling matrix D. If

$$D = \begin{pmatrix} \sqrt{a_{11}} & & \\ & \ddots & \\ & & \sqrt{a_{nn}} \end{pmatrix}$$

then all diagonal elements of M are equal to one. Therefore (van der Sluis 1969, Theorem 4.1)

$$\kappa(M) \le n \min_S \kappa(SAS),$$

where the minimum ranges over all non-singular diagonal matrices S. This means, a diagonal scaling that makes all diagonal elements the same gives the minimal condition number among all diagonal scalings (up to a factor of matrix size n).

We claimed above that the error $\|F\|$ in (4.1) is small. Let's examine in more detail why. The error F comes about because of floating point arithmetic and because of the fact that eigenvalues are approximated by diagonal elements when the off-diagonal elements are small but not necessarily zero. Let's ignore the round-off error, and let's ask why ignoring small off-diagonal elements results in a small $\|F\|$. The trick here is to be clever about what it means to be 'small enough'.

Suppose the Jacobi method has produced a matrix A whose off-diagonal elements are small compared to the corresponding diagonal elements,

$$|a_{ij}| \leq \epsilon\sqrt{a_{ii}a_{jj}}.$$

This implies $|m_{ij}| \leq \epsilon$ where m_{ij} are the elements of the graded matrix $M = D^{-1}AD^{-1}$ and D is the above grading matrix with $\sqrt{a_{ii}}$ on the diagonal. Since the diagonal elements of M are equal to one, we can write $M = I + F$, where F contains all the off-diagonal elements of M and

$$\|F\| \leq (n-1)\epsilon.$$

Therefore $\|F\|$ is small, and (4.1) implies that the error in the eigenvalues is bounded by

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \|M^{-1}\| \, (n-1)\epsilon.$$

Furthermore one can bound $\|M^{-1}\|$ in terms of ϵ,

$$\|M^{-1}\| = \|(I+F)^{-1}\| \leq \frac{1}{1-\|F\|} \leq \frac{1}{1-(n-1)\epsilon}.$$

Replacing this in the error bound gives

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{|\lambda_i|} \leq \frac{(n-1)\epsilon}{1-(n-1)\epsilon}.$$

Therefore ignoring small off-diagonal elements produces a small relative error.

The preceding arguments illustrate that Jacobi's method views a matrix in the best possible light, i.e. in its optimally scaled version. Therefore eigenvalues produced by Jacobi's method tend to have relative accuracy close to machine precision. This is as accurate as it gets. In this sense Jacobi's method is considered optimally accurate.

One way to implement a two-sided Jacobi method is to apply a one-sided method to a Cholesky factor (Barlow and Demmel 1990, Mathias 1996). Let A be a Hermitian positive-definite matrix with Cholesky decomposition $A = L^*L$. The squares of the singular values of L are the eigenvalues of A. The singular values of L can be computed by the one-sided Jacobi method from §4.1. The preliminary Cholesky factorisation does not harm the accuracy of the eigenvalues. Here is why. The computed Cholesky factor is the exact Cholesky factor of a matrix $A+E$, where (Mathias 1995*a*, Lemma 2.6)

$$|e_{ij}| \leq \gamma\epsilon\sqrt{a_{ii}a_{jj}},$$

and γ depends on n. These perturbations have the same form as the ones above, hence lead to a small relative error. A similar argument shows that the squares of the diagonal elements of L are often good approximations to the eigenvalues of A (Mathias 1996).

4.3. Deflation of Block Triangular Matrices

When a matrix is tall and skinny, or short and fat, one can save operations by first converting it to a skinny, short matrix before computing singular values. This can be accomplished by applying a QR decomposition and then computing the singular values of the resulting triangular matrix (Chan 1982). If done properly, the relative accuracy of the singular values is preserved (Mathias 1995*a*, Theorem 3.2).

Suppose we compute the singular values of a triangular matrix by reducing the matrix to diagonal form, say by a Jacobi or QR method. Partition the triangular matrix as

$$B = \begin{pmatrix} B_{11} & B_{12} \\ & B_{22} \end{pmatrix}.$$

If the off-diagonal block B_{12} were zero then the problem of finding the singular values of B could be split into the two smaller, independent subproblems of finding the singular values of B_{11} and of B_{22}. However if the off-diagonal block B_{12} is not zero, we want to know when it can be thrown away without causing too much harm to the singular values of B. The process of discarding information in a matrix to reduce the problem complexity is called 'deflation'.

The deflated matrix and the perturbation are, respectively,

$$B+F = \begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix}, \qquad F = \begin{pmatrix} 0 & -B_{12} \\ 0 & 0 \end{pmatrix}.$$

Corollary 3.2 implies the following relative bound for the singular values $\hat{\sigma}_i$ of the deflated matrix $B+F$.

Corollary 4.1 If B is non-singular then

$$\max_{1\leq i\leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \leq \min\{\|B_{11}^{-1}B_{12}\|,\ \|B_{12}B_{22}^{-1}\|\}.$$

Proof. (Di Lena et al. 1993, Theorem 2.1) □

This means the singular values of the deflated matrix have small relative error when the off-diagonal block is small compared to one of the diagonal blocks.

Now let's suppose a preliminary ordering of the singular values has already taken place. Say, the large singular values have floated to the top of the matrix B while the smaller ones have sunk to the bottom. The bound below is useful when the singular values of the top diagonal block are well separated from the singular values of the bottom block.

Theorem 4.1 If

$$\sigma_{min}(B_{11}) \geq \alpha > \beta \geq \sigma_{max}(B_{22})$$

then

$$\max_{1 \leq i \leq n} |\sigma_i - \hat{\sigma}_i| \leq \sigma_i \, \frac{\|B_{12}\|^2}{\alpha^2 - \beta^2}.$$

Proof. (Di Lena et al. 1993, Theorem 2.2) □

This means the singular values of the deflated matrix have small relative error if the off-diagonal block is small compared to the separation between the singular values of the two diagonal blocks. Theorem 4.1 is an extension of (Demmel and Kahan 1990, Theorem 5). Other bounds that profit from a strong singular value separation appear in (Chandrasekaran and Ipsen 1995, Theorem 5.2.1), (Eisenstat and Ipsen 1995, §5), and (Mathias and Stewart 1993, Theorem 3.1).

4.4. Deflation of Bidiagonal Matrices

Triangular matrices are often further reduced to bidiagonal form before singular values are computed. A bidiagonal matrix is of the form

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ & \ddots & \ddots & \\ & & \alpha_{n-1} & \beta_{n-1} \\ & & & \alpha_n \end{pmatrix}.$$

Bidiagonal matrices can also arise when one computes the vibrational frequencies of a linear mass-spring system (Demmel et al. 1997, §12.1).

There are several algorithms for computing singular values of a bidiagonal matrix to high relative accuracy (Demmel and Kahan 1990, Fernando and Parlett 1994). Because such algorithms apply a sequence of transformations to reduce B to diagonal form, they need to decide when an off-diagonal element β_j is small enough to be neglected without severely harming the singular values.

Suppose we are contemplating the removal of a single off-diagonal element. Here $B + F$ is equal to B, except for the off-diagonal element in row j and column $j+1$, which is equal to zero. Then $F = \beta_j e_j e_{j+1}^T$ and $\|F\| = |\beta_j|$. Corollary 3.2 implies the following bound.

Corollary 4.2 If B is non-singular bidiagonal then

$$\max_{1 \le i \le n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \le |\beta_j| \ \min\{\|B^{-1}e_j\|, \|B^{-*}e_{j+1}\|\}.$$

Proof. (Di Lena et al. 1993, §3) □

This means if we remove a small element from row j and column $j+1$ of a bidiagonal matrix then the relative error in the singular values of the deflated matrix is small if column j or row $j+1$ of B^{-1} are small in norm. Similar bounds, but for a different error measure, appear in (Deift, Demmel, Li and Tomei 1991, Theorem 4.7) and (Demmel and Kahan 1990, Theorem 4).

Corollary 4.2 justifies the use of Convergence Criterion 1 (Demmel and Kahan 1990, §2), (Deift et al. 1991, §4) in the zero-shift Golub-Kahan algorithm for computing singular values of bidiagonal matrices. The practical usefulness of this bound also derives from the fact that it can be computed via the simple recursion below.

Corollary 4.3 If B is non-singular bidiagonal then

$$\max_{1 \le i \le n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sigma_i} \le \min\{\sqrt{r_j}, \sqrt{c_j}\},$$

where

$$r_1 = \frac{\beta_1^2}{\alpha_1^2}, \qquad r_j = \frac{\beta_j^2}{\alpha_j^2}(1 + r_{j-1}), \quad 2 \le j \le n,$$

and

$$c_{n-1} = \frac{\beta_{n-1}^2}{\alpha_n^2}, \qquad c_j = \frac{\beta_j^2}{\alpha_{j+1}^2}(1 + c_{j+1}), \quad n-2 \ge j \ge 1.$$

Proof. (Di Lena et al. 1993, Theorems 3.1, 3.2) □

When the shift in the Golub-Kahan algorithm or the qd algorithm is non-zero, it can be incorporated into the perturbation bounds (Eisenstat and Ipsen 1995, Theorem 5.7), (Fernando and Parlett 1994).

5. Multiplicative Perturbations for Eigenvalues

We shift gears and represent the perturbed matrix from now on as D_1AD_2 where D_1 and D_2 are non-singular. When $D_2 = D_1^{-1}$ this is just a similarity transformation, which means that A and D_1AD_2 have the same eigenvalues. When $D_2 = D_1^*$ this is a congruence transformation, which means that A and D_1AD_2 have the same inertia when A is Hermitian. Since the non-singularity of D_1 and D_2 forces A and D_1AD_2 to have the same rank, multiplicative perturbations are less powerful than additive perturbations.

Then why are multiplicative perturbations useful? It turns out that it is sometimes easier to express a component-wise relative perturbation of a sparse matrix as a multiplicative perturbation than as an additive perturbation. The following example illustrates how natural multiplicative perturbations can be, especially for bidiagonal and tridiagonal matrices.

Example 5.1 (Barlow and Demmel 1990, p 770), (Eisenstat and Ipsen 1995, Corollary 4.1)

Consider the real, symmetric tridiagonal matrix

$$A = \begin{pmatrix} 0 & \alpha_1 & & & & \\ \alpha_1 & 0 & \alpha_2 & & & \\ & \alpha_2 & 0 & \alpha_3 & & \\ & & \alpha_3 & 0 & \alpha_4 & \\ & & & \alpha_4 & 0 & \alpha_5 \\ & & & & \alpha_5 & 0 \end{pmatrix}.$$

Such a matrix occurs, for instance, when one converts the singular value problem of a bidiagonal matrix to an eigenvalue problem (see §3.1). A component-wise relative perturbation of a single off-diagonal pair in A produces the perturbed matrix

$$\hat{A} = \begin{pmatrix} 0 & \alpha_1 & & & & \\ \alpha_1 & 0 & \alpha_2 & & & \\ & \alpha_2 & 0 & \beta\alpha_3 & & \\ & & \beta\alpha_3 & 0 & \alpha_4 & \\ & & & \alpha_4 & 0 & \alpha_5 \\ & & & & \alpha_5 & 0 \end{pmatrix},$$

where $\beta \neq 0$. For instance, β could be of the form $\beta = 1 + \epsilon$, where $|\epsilon|$ does not exceed machine epsilon. The perturbed matrix $\hat{A}$ can be represented as a multiplicative perturbation $\hat{A} = D^T A D$, where

$$D = \begin{pmatrix} \sqrt{\beta} & & & & & \\ & \frac{1}{\sqrt{\beta}} & & & & \\ & & \sqrt{\beta} & & & \\ & & & \sqrt{\beta} & & \\ & & & & \frac{1}{\sqrt{\beta}} & \\ & & & & & \sqrt{\beta} \end{pmatrix}.$$

In this case a component-wise relative perturbation of an off-diagonal pair can be represented as a multiplicative perturbation. □

The simple-minded approach of disguising a multiplicative perturbation as an additive perturbation, like so

$$D_1 A D_2 = A + E, \qquad \text{where} \quad E = D_1 A D_2 - A,$$

produces a perturbation matrix E that may not be small or meaningful.

There are different techniques for deriving multiplicative perturbation bounds, and some of them are compared in (Li and Mathias 1997, §4.2). Here we start from absolute perturbation bounds and show that they imply many of relative bounds.

5.1. Bauer-Fike-Type Bounds

Again we start with a diagonalisable matrix, and we bound the distance of a perturbed eigenvalue to a closest exact eigenvalue in terms of the two-norm.

Let $A = X \Lambda X^{-1}$ be an eigendecomposition of A, where

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix},$$

and λ_i are the eigenvalues of A. Let $\hat{\lambda}$ be an eigenvalue of the perturbed matrix $D_1 A D_2$ and $\hat{x} \neq 0$ a corresponding unit eigenvector,

$$(D_1 A D_2)\, \hat{x} = \hat{\lambda} \hat{x}, \qquad \|\hat{x}\| = 1,$$

with residual

$$r \equiv A\hat{x} - \hat{\lambda}\hat{x}.$$

This time we use the Bauer-Fike Theorem with residual bound (Bauer and Fike 1960, Theorem IIIa),

$$\min_{1 \leq i \leq n} |\lambda_i - \hat{\lambda}| \leq \kappa(X)\, \|r\|. \tag{5.1}$$

The relative error bound below for the eigenvalue $\hat{\lambda}$ of the perturbed matrix $D_1 A D_2$ measures the error relative to the perturbed eigenvalue rather than an exact eigenvalue.

Theorem 5.1 If A is diagonalisable then

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

Proof. (Eisenstat and Ipsen 1996, Theorem 6.1)

The idea is to concoct a residual that contains the factor $\hat{\lambda}$ and then to use the absolute bound (5.1).

From $(D_1 A D_2)\hat{x} = \hat{\lambda}\hat{x}$ follows

$$A\, z = \hat{\lambda}\, D_1^{-1} D_2^{-1}\, z, \qquad \text{where} \quad z \equiv D_2\hat{x}/\|D_2\hat{x}\|.$$

The residual for $\hat{\lambda}$ and z is

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

and it contains $\hat{\lambda}$ as a factor. Now apply the absolute bound (5.1) to f. $\square$

The perturbed matrix D_1AD_2 is not required to be diagonalisable. As in the case of additive perturbations, $\kappa(X)$ can be interpreted as a condition number. The factor $\|I - D_1^{-1}D_2^{-1}\|$ represents a relative deviation from similarity, because

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

represents a difference relative to D_2.

There are two cases in which the bound in Theorem 5.1 is guaranteed to be zero and hence tight. First, when $D_1 = D_2^{-1}$, because similar matrices have the same eigenvalues. Second, when $\hat{\lambda} = 0$, because A and D_1AD_2 are singular and both have a zero eigenvalue.

5.2. Hoffman-Wielandt-Type Bounds for Diagonalisable Matrices

Based on a one-to-one correspondence between exact and perturbed eigenvalues, we bound the sum (of squares) of all distances between exact and perturbed eigenvalues in terms of the Frobenius norm. In contrast to the previous section, the perturbed matrix must now also be diagonalisable.

Let A and D_1AD_2 be diagonalisable with respective eigendecompositions

$$A = X\Lambda X^{-1}, \qquad D_1AD_2 = \hat{X}\hat{\Lambda}\hat{X}^{-1}.$$

The eigenvalues are

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}, \qquad \hat{\Lambda} = \begin{pmatrix} \hat{\lambda}_1 & & \\ & \ddots & \\ & & \hat{\lambda}_n \end{pmatrix}.$$

Theorem 5.2 If A and D_1AD_2 are non-singular and diagonalisable then there exists a permutation τ so that

$$\sqrt{\sum_{i=1}^{n} \left(\frac{|\lambda_i - \hat{\lambda}_{\tau(i)}|}{|\lambda_i|} \right)^2} \le \kappa(X)\kappa(\hat{X})\, \|D_2\| \|D_1 - D_2^{-1}\|_F.$$

There also exists a permutation σ so that

$$\sqrt{\sum_{i=1}^{n} \left(\frac{|\lambda_i - \hat{\lambda}_{\sigma(i)}|}{|\lambda_i|} \right)^2} \le \kappa(X)\kappa(\hat{X})\, \|D_1\| \|D_1^{-1} - D_2\|_F.$$

Proof. The first bound is (Li 1997, Theorem 2.1), while the second one is (Li 1997, Theorem 2.1$'$). $\square$

Therefore the relative error bound can only be small if D_1 and D_2 are close to being a similarity transformation. A similar bound in (Li and Mathias 1997, Proposition 3.5) applies to a one-sided perturbation and matrices with positive eigenvalues and holds in any unitarily invariant norm.

5.3. Hoffman-Wielandt-Type Bounds for Hermitian Matrices

Hoffman-Wielandt-type bounds for Hermitian matrices require that the perturbed matrix also be Hermitian. This means the perturbed matrix better be of the form DAD^*, where D is non-singular. Since the perturbed matrix is congruent to A, it has the same inertia as A. Number the eigenvalues of A and DAD^* so that

$$\lambda_n \leq \ldots \leq \lambda_1, \qquad \hat{\lambda}_n \leq \ldots \leq \hat{\lambda}_1.$$

Theorem 5.3 If A and DAD^* are Hermitian and non-singular then

$$\sqrt{\sum_{i=1}^{n} \frac{|\lambda_i - \hat{\lambda}_i|^2}{|\lambda_i \hat{\lambda}_i|}} \leq \|D^* - D^{-1}\|_F.$$

Proof. (Li and Mathias 1997, Corollary 3.2')

A proof for the special case of positive-definite matrices appears in (Li 1994*a*, Theorem 3.1). □

Therefore, the relative error in the eigenvalues of DAD^* is small if D is close to a unitary (or orthogonal) matrix. The bound (Li 1997, Theorem 2.2) is weaker than Theorem 5.3 (Li and Mathias 1997, §4.1). Majorisation theory can deliver bounds that are stronger than Theorem 5.3 and hold for any unitarily invariant norm (Li and Mathias 1997, Proposition 3.2, (3.8)).

5.4. Ostrowski-Type Bounds

In 1959 Ostrowski presented the first relative perturbation bounds for eigenvalues. He created a multiplicative perturbation DAD^* of a Hermitian matrix A, where D is non-singular; and he bounded the ratio of exact and perturbed eigenvalues in terms of the smallest and largest eigenvalues of DD^* (Ostrowski 1959), (Horn and Johnson 1985, Theorem 4.5.9),

$$\lambda_{min}(DD^*)\, \lambda_i \leq \hat{\lambda}_i \leq \lambda_i\, \lambda_{max}(DD^*). \tag{5.2}$$

Ostrowski's theorem can also be phrased in terms of absolute values of eigenvalues.

Theorem 5.4 If A and DAD^* are Hermitian then

$$\frac{|\lambda_i|}{\|(D^*D)^{-1}\|} \leq |\hat{\lambda}_i| \leq |\lambda_i|\, \|D^*D\|.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 2.1) □

This bound is tight, for instance, when D is a multiple of an orthogonal matrix. The following example illustrates what the bound looks like in the case of tridiagonal matrices.

Example 5.2 (Eisenstat and Ipsen 1995, Corollary 4.1)
Return to the symmetric tridiagonal matrix with zero diagonal and its single-element perturbation in Example 5.1. In this case the bound in Theorem 5.4 amounts to (Kahan 1966, p 49ff), (Demmel and Kahan 1990, Theorem 2)

$$\frac{1}{\eta}|\lambda_i| \leq |\hat{\lambda}_i| \leq \eta\,|\lambda_i|,$$

where $\eta \equiv \max\{|\beta|, 1/|\beta|\}$. Therefore the ratio between perturbed and exact eigenvalues is close to one if the perturbation $|\beta|$ is close to one.
This bound can be extended to the perturbation of any number of off-diagonal pairs of a real symmetric tridiagonal matrix with zero diagonal (Demmel and Kahan 1990, Corollary 1). □

Ostrowski's Theorem (5.2) can also be extended to products of eigenvalue ratios (Li and Mathias 1997, Theorem 2.3).

5.5. Weyl-Type Bounds

Ostrowski's theorem leads to a relative Weyl-type bound for multiplicative perturbations.

Theorem 5.5 If A and DAD^* are Hermitian then

$$|\lambda_i - \hat{\lambda}_i| \leq |\lambda_i|\,\|DD^* - I\|, \qquad 1 \leq i \leq n.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 2.1)
The proof is similar to that of Theorem 2.7 for additive perturbations. Fix an index i. Since 0 is the ith eigenvalue of $A - \lambda_i I$, Sylvester's Law of Inertia (Horn and Johnson 1985, Theorem 4.5.8) implies that 0 is the ith eigenvalue of $D(A - \lambda_i I)D^*$. Write

$$D(A - \lambda_i I)D^* = DAD^* - \lambda_i DD^* = \bar{A} + \bar{E},$$

where

$$\bar{A} \equiv DAD^* - \lambda_i I, \qquad \bar{E} \equiv \lambda_i(I - DD^*).$$

Applying Weyl's absolute bound (2.4) to $\bar{A}$ and $\bar{A} + \bar{E}$ gives

$$\max_{1\leq j\leq n} |\lambda_j(\bar{A}) - \lambda_j(\bar{A} + \bar{E})| \leq \|\bar{E}\| = |\lambda_i|\,\|DD^* - I\|.$$

In particular, for $j = i$,

$$|0 - (\hat{\lambda}_i - \lambda_i)| \leq |\lambda_i|\,\|DD^* - I\|.$$

□

The bound above holds even for zero eigenvalues. The factor $\|DD^* - I\|$ represents the deviation of the congruence transformation from similarity.

This means, if the perturbed matrix is a congruence transformation of the original matrix, then the relative error in the perturbed eigenvalues is small if the congruence transformation is close to similarity.

Example 5.3 We can apply Theorem 5.5 to the symmetric tridiagonal matrix with zero diagonal in Example 5.1. If we assume that the multiplicative perturbation α is of the form $\alpha = 1 + \epsilon$ then

$$|\lambda_i - \hat{\lambda}_i| \leq |\lambda_i|\, |\epsilon|.$$

Therefore a small relative error in a pair of off-diagonal elements causes only a small relative error in the eigenvalues.

The bound below is similar in spirit to Theorem 5.5 but applies to a different error measure.

Theorem 5.6 If A and DAD^* are non-singular Hermitian then

$$\max_{1\leq i\leq n} \frac{|\lambda_i - \hat{\lambda}_i|}{\sqrt{\lambda_i \hat{\lambda}_i}} \leq \|D^{-1} - D^*\|.$$

Proof. (Li and Mathias 1997, Proposition 3.2') □

This bound was first derived for the special case of positive-definite matrices (Li 1994*a*, Theorem 3.1).

6. Multiplicative Perturbations for Singular Values

The perturbed matrix is represented as $D_1 B D_2$, where D_1 and D_2 are non-singular diagonal matrices. Such a perturbation can occur, for instance, when a one-sided Jacobi method is applied to B. In this case the computed singular values are exact singular values of a matrix $D_1 B D_2$ where D_1 and D_2 are close to the identity (Demmel 1997, §5.4.3).

Again let B be a tall and skinny matrix, i.e. B is $m \times n$ with $m \geq n$. The singular values of B are

$$\sigma_1 \geq \ldots \geq \sigma_n \geq 0.$$

The perturbed matrix is represented as $D_1 B D_2$, where D_1 and D_2 are non-singular. The singular values of $D_1 B D_2$ are

$$\hat{\sigma}_1 \geq \ldots \geq \hat{\sigma}_n \geq 0.$$

When D_1 is diagonal it represents a row scaling, while a diagonal D_2 represents a column scaling.

6.1. *Ostrowski-Type Bounds*

Let's first determine by which factor the singular values of B change when multiplicative perturbations D_1 and D_2 are applied.

Theorem 6.1

$$\frac{\sigma_i}{\|D_1^{-1}\|\,\|D_2^{-1}\|} \le \hat{\sigma}_i \le \sigma_i\,\|D_1\|\,\|D_2\|.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 3.1)
Convert the problem to an eigenvalue problem as in §3.1 and apply the eigenvalue result Theorem 5.4. □

This means, if D_1 and D_2 are almost unitary then the norms in Theorem 6.1 are almost one, and a perturbed singular value differs from the corresponding exact singular value by a factor close to one.

Theorem 6.1 can reproduce perturbation bounds for component-wise perturbations of bidiagonal matrices from (Barlow and Demmel 1990, Theorem 1), (Deift et al. 1991, Theorem 2.12) and (Demmel and Kahan 1990, Corollary 2). The example below illustrates how.

Example 6.1 (Eisenstat and Ipsen 1995, Corollary 4.2)
Consider the bidiagonal matrix

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ & \alpha_2 & \beta_2 & \\ & & \alpha_3 & \beta_3 \\ & & & \alpha_4 \end{pmatrix}$$

and its component-wise perturbation

$$\hat{B} = \begin{pmatrix} \gamma_1\alpha_1 & \gamma_2\beta_1 & & \\ & \gamma_3\alpha_2 & \gamma_4\beta_2 & \\ & & \gamma_5\alpha_3 & \gamma_6\beta_3 \\ & & & \gamma_7\alpha_4 \end{pmatrix},$$

where $\gamma_j \neq 0$. For instance, if $\gamma_j = 1+\epsilon_j$ for small ϵ_j then $\hat{B}$ is a component-wise relative perturbation of B.

Write $\hat{B} = D_1 B D_2$, where D_1 takes care of the odd-numbered, diagonal perturbations, while D_2 takes care of the even-numbered, off-diagonal perturbations, like so:

$$D_1 = \begin{pmatrix} \gamma_1 & & & \\ & \frac{\gamma_1\gamma_3}{\gamma_2} & & \\ & & \frac{\gamma_1\gamma_3\gamma_5}{\gamma_4} & \\ & & & \frac{\gamma_1\gamma_3\gamma_5\gamma_7}{\gamma_6} \end{pmatrix}, \qquad D_2 = \begin{pmatrix} 1 & & & \\ & \frac{\gamma_2}{\gamma_1} & & \\ & & \frac{\gamma_2\gamma_4}{\gamma_3} & \\ & & & \frac{\gamma_2\gamma_4\gamma_6}{\gamma_5} \end{pmatrix}.$$

The perturbation D_1 operates on rows and D_2 operates on columns. The resulting interference is apparent in the increasing products and the denominators, where each D_i undoes part of the other's action in its own territory.

Application of Theorem 6.1 yields

$$\frac{1}{\eta}\,\sigma_i \le \hat{\sigma}_i \le \eta\,\sigma_i,$$

where $\eta \equiv \prod_{j=1}^{7} \max\{|\gamma_j|, 1/|\gamma_j|\}$. This means, if each factor γ_j is close to one then the ratio of perturbed to exact singular values is close to one. □

These bounds are actually realistic. There are algorithms that deliver singular values of bidiagonal matrices to high relative accuracy: the dqds algorithm (Fernando and Parlett 1994) and, to a large extent, a fine-tuned zero-shift version of the Golub-Kahan algorithm (Demmel and Kahan 1990, Deift et al. 1991).

6.2. Hoffman-Wielandt-Type Bounds

Now let's bound the sum of squares of all relative errors.

Theorem 6.2 If B and D_1BD_2 have full column rank then

$$\sqrt{\sum_{i=1}^{n} \frac{|\sigma_i - \hat{\sigma}_i|^2}{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2}\left(\|D_1^* - D_1^{-1}\|_F + \|D_2^* - D_2^{-1}\|_F\right).$$

Proof. (Li and Mathias 1997, Proposition 3.3') □

The terms $\|D_1^* - D_1^{-1}\|_F$ and $\|D_2^* - D_2^{-1}\|_F$ indicate how far D_1 and D_2, respectively, are from being unitary (or orthogonal). The relative error is small if D_1 and D_2 are close to unitary. A weaker version of Theorem 6.2 appears in (Li 1994*a*, Theorem 4.1). As with Theorem 5.3, majorisation theory yields a bound stronger than Theorem 6.2 that holds in any unitarily invariant norm (Li and Mathias 1997, Proposition 3.3, (3.12)).

The bound below is similar in spirit but applies to a different error measure.

Theorem 6.3 If B and D_1BD_2 have full column rank then

$$\sqrt{\sum_{i=1}^{n} \left(\frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt[p]{\sigma_i^p + \hat{\sigma}_i^p}}\right)^2} \leq \frac{1}{2^{1+1/p}}\left(\|D_1^* - D_1^{-1}\|_F + \|D_2^* - D_2^{-1}\|_F\right),$$

where $1 \leq p \leq \infty$ is an integer.

Proof. (Li 1994*a*, Theorem 4.2)

Follows from Theorem 6.2 because

$$\left(\frac{2}{\sigma_i^p + \hat{\sigma}_i^p}\right)^{1/p} \leq \frac{1}{\sigma_i \hat{\sigma}_i}.$$

□

6.3. Weyl-Type Bounds

At last we bound individual relative errors.

Theorem 6.4

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \ \max\{\|I - D_1^{-1} D_1^{-*}\|, \ \|I - D_2^{-*} D_2^{-1}\|\}.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 3.3)
Convert the singular value problem to a large eigenvalue problem as in §3.1 and apply Theorem 5.5 to the eigenvalue problem. □

The factors $\|I - D_1^{-1} D_1^{-*}\|$ and $\|I - D_2^{-*} D_2^{-1}\|$ represent relative deviations of D_1 and D_2, respectively, from being unitary (or orthogonal). Hence the relative error in the singular values of $D_1 B D_2$ is small if D_1 and D_2 are close to unitary.

The following bound is similar to the Frobenius norm bound in Theorem 6.2.

Theorem 6.5 If B and $D_1 B D_2$ have full column rank then

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2} \left(\|D_1^* - D_1^{-1}\| + \|D_2^* - D_2^{-1}\| \right).$$

Proof. (Li and Mathias 1997, Proposition 3.3') □

A weaker bound appears in (Li 1994*a*, Theorem 4.1).

The following bound is a counterpart of the Frobenius norm bound in Theorem 6.3.

Theorem 6.6 If B and $D_1 B D_2$ have full column rank then

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt[p]{|\sigma_i|^p + |\hat{\sigma}_i|^p}} \leq \frac{1}{2^{1+1/p}} \left(\|D_1^* - D_1^{-1}\| + \|D_2^* - D_2^{-1}\| \right),$$

where $1 \leq p \leq \infty$ is an integer.

Proof. (Li 1994*a*, Theorem 4.2)
Follows from Theorem 6.5 in the same way that Theorem 6.3 follows from Theorem 6.2. □

7. Some Applications of Multiplicative Perturbations

We discuss component-wise perturbations of generalised bidiagonal matrices; deflation of triangular matrices; and rank-revealing decompositions.

7.1. Component-Wise Perturbations of Generalised Bidiagonal Matrices

Example 6.1 illustrates that small relative changes in any bidiagonal matrix cause only small relative changes in its singular values, regardless of the value of the non-zero matrix elements. Are there other matrices with this pleasant property? The answer is not really. The only other matrices with this property are those whose sparsity structure is 'essentially' bidiagonal

(Demmel and Gragg 1993). At first glance this result looks pretty negative. It appears to suggest that we can forget about high relative accuracy for matrices other than bidiagonals. But then again, not all perturbations are component-wise relative perturbations. Just think about the perturbations caused by the deflation of triangular matrices in §§4.3 and 4.4. There is still plenty of room for singular values of all kinds of matrices to have relative accuracy, but mostly not with regard to component-wise relative perturbations*.

So, what are 'essentially' bidiagonal matrices? We define an undirected bipartite graph G of a matrix B as follows. Each row of B is represented by a node r_i, and each column by a node c_i. There is an edge between r_i and c_j if and only if element (i,j) of B is non-zero. The matrix B is called *biacyclic* if its graph G is acyclic (Demmel and Gragg 1993, §1). Examples of biacyclic matrices, in addition to bidiagonal matrices, include the following 'half arrow' matrices (Demmel and Gragg 1993, §5)

$$\begin{pmatrix} * & & & & * \\ & * & & & * \\ & & * & & * \\ & & & * & * \\ & & & & * \end{pmatrix}, \qquad \begin{pmatrix} * & & & * & & \\ & * & & & * & \\ & & * & & & * \\ & & & * & & * \\ & & & & * & * \\ & & & & & * \end{pmatrix}.$$

The nice thing about subjecting biacyclic matrices to component-wise relative perturbations is that we get an Ostrowski-type bound. This means, changing an element of a biacyclic matrix by a factor does not change the singular values by more than this factor, regardless of the value of the non-zero matrix elements. No other matrices possess this property.

Theorem 7.1 Let B be a matrix of order n. The following two conditions are equivalent:

- B is biacyclic.
- If $\hat{B}$ is equal to B except for element (k,l) which is multiplied by $\gamma \neq 0$, then the singular values $\hat{\sigma}_i$ of $\hat{B}$ satisfy

$$\min\{|\gamma|, |\gamma^{-1}|\}\, \sigma_i \leq \hat{\sigma}_i \leq \sigma_i \, \max\{|\gamma|, |\gamma^{-1}|\}, \qquad 1 \leq i \leq n.$$

Proof. (Demmel and Gragg 1993, Theorem 1)

The result also follows directly from Theorem 6.1 (Eisenstat and Ipsen 1995, Corollary 4.3). □

* However, if we are willing to restrict the values of the matrix elements and impose signs on the non-zero entries of a matrix so as to forestall cancellation in the computation of certain quantities, then one can also obtain high relative accuracy (Demmel et al. 1997).

Theorem 7.1 implies that a small relative perturbation in a matrix element causes small relative changes in the singular values, regardless of the values of the non-zero matrix elements, if and only if the matrix has an acyclic graph. To see this consider $\gamma = 1 + \epsilon$ for some $\epsilon > 0$. Theorem 7.1 implies the relative error bound

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \epsilon, \qquad 1 \leq i \leq n.$$

Theorem 7.1 can be extended to the case where all elements of a biacyclic matrix are multiplied by non-zero factors. This gives a bound similar to the one in Example 6.1 (Demmel and Gragg 1993, page 206), (Eisenstat and Ipsen 1995, Corollary 4.3).

7.2. Deflation of Block Triangular Matrices, Again

First we prove an auxiliary bound for a special multiplicative perturbation which is useful for modelling deflation in triangular matrices. Let

$$B = \begin{pmatrix} B_{11} & B_{12} \\ & B_{22} \end{pmatrix}$$

be a block triangular matrix of order n; and let the perturbed matrix be DB or BD, where

$$D = \begin{pmatrix} I & X \\ & I \end{pmatrix}$$

is partitioned commensurately with B.

Theorem 7.2 The singular values $\hat{\sigma}_i$ of DB, or BD, satisfy

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \, \|X\|, \qquad 1 \leq i \leq n.$$

If in addition B is non-singular then

$$\max_{1 \leq i \leq n} \frac{|\sigma_i - \hat{\sigma}_i|}{\sqrt{\sigma_i \hat{\sigma}_i}} \leq \frac{1}{2}\|X\|.$$

Proof. The first inequality (Eisenstat and Ipsen 1995, Lemma 5.1) follows from Theorem 6.1,

$$\frac{\sigma_i}{\|D^{-1}\|} \leq \hat{\sigma}_i \leq \|D\|,$$

and from

$$\|D\| = \|D^{-1}\| \leq 1 + \|X\|.$$

The second inequality (Li 1994*a*, Corollary 4.1) follows from Theorem 6.5. □

Now look at a perturbed matrix that is a deflated version of the block triangular matrix B,

$$\hat{B} = \begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix}.$$

The following bound is the same as the one in Corollary 4.1 which was derived in the context of additive pertubations.

Theorem 7.3 If B_{11} or B_{22} are non-singular then the singular values of B and $\hat{B}$ satisfy

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \, \min\{\|B_{11}^{-1}B_{12}\|, \, \|B_{12}B_{22}^{-1}\|\}.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 5.2)
Follows directly from Theorem 7.2. □

Let's see what happens for bidiagonal matrices. We write the bidiagonal matrix and it's perturbation so as to highlight the action,

$$B = \begin{pmatrix} B_{11} & \beta_j e_j e_1^T \\ & B_{22} \end{pmatrix}, \qquad \hat{B} = \begin{pmatrix} B_{11} & \\ & B_{22} \end{pmatrix}.$$

Both matrices are bidiagonal, and $\hat{B}$ is equal to B, except for the off-diagonal element in row j and column $j+1$, which is equal to zero. Application of Theorem 7.3 to B and $\hat{B}$ produces a bound similar to the one in Corollary 4.2.

Theorem 7.4 If B and $\hat{B}$ are non-singular then

$$\frac{1}{1+\eta} \leq \frac{\hat{\sigma}_i}{\sigma_i} \leq 1+\eta,$$

where

$$\eta \equiv |\beta_j| \, \min\{\|B_{11}^{-1}e_j\|, \, \|B_{22}^{-*}e_1\|\}.$$

Proof. (Eisenstat and Ipsen 1995, Theorem 5.5) □

This bound, like Corollary 4.2 in the context of additive perturbations, can be used to justify Convergence Criterion 1 in the Golub-Kahan algorithm with zero shift (Eisenstat and Ipsen 1995, Corollary 5.6).

7.3. Rank-Revealing Decompositions

A rank-revealing decomposition is a cheap imitation of a singular value decomposition. It can serve as an intermediate step in the high-accuracy computation of a singular value decomposition. Passing through a rank-revealing decomposition on the way to a singular value decomposition allows one to represent all errors in terms of multiplicative perturbations (Demmel et al. 1997, §3).

Consider a $m \times n$ matrix B, $m \geq n$, of rank r. Decompose $B = XDY^*$ where X is $m \times r$, Y is $n \times r$, and D is diagonal of order r and $r \leq n$. This means D is non-singular, and X and Y have full column rank. The decomposition $B = XDY^*$ is a *rank-revealing decomposition of* A if X and Y are well-conditioned, that is, if

$$\kappa(X) \equiv \|X\| \, \|X^\dagger\| \qquad \text{and} \qquad \kappa(Y) \equiv \|Y\| \, \|Y^\dagger\|$$

are close to one (Demmel et al. 1997, Definition 2.1).

A singular value decomposition qualifies as the luxury edition of a rank revealing decomposition because X and Y are orthogonal, hence perfectly conditioned. Gaussian elimination with complete pivoting may be a more affordable model. Here $X = P_l L$ and $Y = UP_u$, where P_l and P_u are permutation matrices, L is unit lower triangular and U is unit upper triangular. Because all entries of L and U are bounded by one in absolute value, X and Y tend to be well-conditioned.

Suppose the computed version of our rank revealing decomposition is $\hat{B} = \hat{X}\hat{D}\hat{Y}^*$. If the elements of the diagonal matrix $\hat{D}$ have high relative accuracy, and $\hat{X}$ and $\hat{Y}$ have high norm-wise accuracy then the singular values $\hat{\sigma}_i$ of $\hat{B}$ have small relative error.

Theorem 7.5 Let

$$\hat{D} = D + \Delta, \qquad \hat{X} = X + E, \qquad \hat{Y} = Y + F.$$

If for some $0 \leq \epsilon < 1$

$$\frac{|\Delta_{ii}|}{|D_{ii}|} \leq \epsilon, \qquad \frac{\|E\|}{\|X\|} \leq \epsilon, \qquad \frac{\|F\|}{\|Y\|} \leq \epsilon$$

then

$$|\sigma_i - \hat{\sigma}_i| \leq \sigma_i \, (2\eta + \eta^2),$$

where

$$\eta \equiv \epsilon(2 + \epsilon) \, \max\{\kappa(X), \kappa(Y)\}.$$

Proof. (Demmel et al. 1997, Theorem 2.1)

The idea is to express Δ, E and F as multiplicative perturbations. Since X has full column rank, it has a left-inverse $X^\dagger$ and we can write

$$\hat{B} = (X + E)\hat{D}\hat{Y}^* = (I + EX^\dagger)X\hat{D}\hat{Y}^* = D_1 \, X\hat{D}\hat{Y}^*,$$

where $D_1 \equiv I + EX^\dagger$ is a multiplicative perturbation and

$$\|D_1\| \leq 1 + \|E\| \, \|X^\dagger\| \leq 1 + \epsilon\kappa(X).$$

Similarly one can show $\hat{B} = D_1 B D_2$, where

$$\|D_2\| \leq 1 + \kappa(Y) \, (2\epsilon + \epsilon^2).$$

Application of Theorem 6.1 gives the desired bound. □

Therefore if X and Y are well-conditioned the relative error in the singular values of $\hat{B}$ is proportional to the accuracy ϵ of the rank-revealing decomposition. Note that the error bound depends on $\kappa(X)$ and $\kappa(Y)$ but not on $\kappa(D)$. That's because of the stricter requirement for the perturbation Δ of D, which must be a component-wise relative perturbation.

8. The End

We have seen that many absolute perturbation bounds imply relative bounds. Examples include the bounds by Bauer-Fike, Hoffman-Wielandt and Weyl. So there is no question of existence. Relative error bounds always exist, for any matrix and for any perturbation.

Like absolute bounds, relative bounds become stronger when the matrices have structure. A Weyl-type bound for Hermitian positive-definite matrices, for instance, is stronger than a Bauer-Fike-type bound for diagonalisable matrices. In contrast to absolute bounds, though, relative bounds can impose more stringent conditions on the matrices to achieve the corresponding bound. For example, most relative bounds for additive perturbations require that the original matrix be non-singular.

Therefore relative error bounds are not necessarily stronger than absolute error bounds. They just rely for their accuracy on different perturbations. Consider eigenvalues of normal matrices, for instance. A small absolute perturbation E guarantees a small absolute error, while a small relative perturbation, such as

$$\|A^{-1/2}EA^{-1/2}\| \qquad \text{or} \qquad \|I - D_1^{-1}D_2^{-1}\|,$$

guarantees a small relative error. This means, before requesting high relative accuracy you'd better be sure to have a small relative perturbation.

Several theses have been written on the subject of relative error bounds in the context of Jacobi methods for computing singular values (Drmač 1994), eigendecompositions of Hermitian matrices (Slapničar 1992), and eigenvalues of skew-symmetric matrices (Pietzsch 1993), as well as fast algorithms for computing eigendecompositions of real symmetric tridiagonal matrices (Dhillon 1997).

We have omitted the following issues in our discussion of relative error bounds:

- generalised eigenvalue problems (Barlow and Demmel 1990, Hari and Drmač 1997, Li 1994*a*, Veselić and Slapničar 1993),
- sensitivity of eigenvalues and singular values to perturbations in the factors of a matrix (Dhillon 1997, Demmel et al. 1997, Parlett 1997, Veselić and Slapničar 1993),
- relative errors in the form of derivatives when the matrix elements

depend smoothly on a parameter (Deift et al. 1991, §2), (Parlett 1997, Theorem 1).

It is also possible to derive relative perturbation bounds for invariant subspaces and singular vector spaces. These are generally bounds on the angle between an exact and perturbed invariant subspace in terms of a relative eigenvalue separation as opposed to an absolute eigenvalue separation. Many of the papers cited here also discuss bounds for subspaces. Papers solely dealing with subspaces include among others (Eisenstat and Ipsen 1994, Li 1994*b*, Mathias 1995*b*, Mathias and Veselić 1995, Slapničar and Veselič 1992, Truhar and Slapničar 1997).

Acknowledgement

I thank Stan Eisenstat for carefully reading the manuscript and for many helpful suggestions.

REFERENCES

J. Barlow and J. Demmel (1990), 'Computing accurate eigensystems of scaled diagonally dominant matrices', *SIAM J. Numer. Anal.* **27**(3), 762–91.

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

R. Bhatia (1997), *Matrix Analysis*, Springer Verlag, New York.

T. Chan (1982), 'An improved algorithm for computing the singular value decomposition', *ACM Trans. Math. Soft.* **8**(1), 72–83.

S. Chandrasekaran and I. Ipsen (1995), 'Analysis of a QR algorithm for computing singular values', *SIAM J. Matrix Anal. Appl.* **16**(2), 520–35.

P. Deift, J. Demmel, L. Li and C. Tomei (1991), 'The bidiagonal singular value decomposition and Hamiltonian mechanics', *SIAM J. Numer. Anal.* **28**(5), 1463–1516.

J. Demmel (1997), *Applied Numerical Linear Algebra*, SIAM, Philadelphia.

J. Demmel and W. Gragg (1993), 'On computing accurate singular values and eigenvalues of matrices with acyclic graphs', *Linear Algebra Appl.* **185**, 203–17.

J. Demmel and W. Kahan (1990), 'Accurate singular values of bidiagonal matrices', *SIAM J. Sci. Stat. Comput.* **11**(5), 873–912.

J. Demmel and K. Veselič (1992), 'Jacobi's method is more accurate than QR', *SIAM J. Matrix Anal. Appl.* **13**(4), 1204–45.

J. Demmel, M. Gu, S. Eisenstat, I. Slapničar, K. Veselič and Z. Drmač (1997), Computing the singular value decomposition with high relative accuracy, Technical report, Computer Science Division, University of California, Berkeley, CA.

I. Dhillon (1997), A New $O(n^2)$ Algorithm for the Symmetric Tridiagonal Eigenvalue/Eigenvector Problem, PhD thesis, University of California, Berkeley, California.

I. Dhillon, G. Fann and B. Parlett (1997), Application of a new algorithm for the symmetric eigenproblem to computational quantum chemistry, in *Proceedings of the Eighth SIAM Conference on Parallel Processing for Scientific Computing*, SIAM, Philadelphia.

G. Di Lena, R. Peluso and G. Piazza (1993), 'Results on the relative perturbation of the singular values of a matrix', *BIT* **33**, 647–53.

Z. Drmač (1994), Computing the Singular and the Generalized Singular Values, PhD thesis, Fachbereich Mathematik, Fernuniversität Gesamthochschule Hagen, Germany.

Z. Drmač (1996*a*), 'On relative residual bounds for the eigenvalues of a Hermitian matrix', *Linear Algebra Appl.* **244**, 155–64.

Z. Drmač (1996*b*), 'On the condition behaviour in the Jacobi method', *SIAM J. Matrix Anal. Appl.* **17**(3), 509–14.

Z. Drmač and V. Hari (1997), 'Relative residual bounds for the eigenvalues of a Hermitian semidefinite matrix', *SIAM J. Matrix Anal. Appl.* **18**(1), 21–9.

S. Eisenstat and I. Ipsen (1994), Relative perturbation bounds for eigenspaces and singular vector subspaces, in *Applied Linear Algebra*, SIAM, Philadelphia, pp. 62–5.

S. Eisenstat and I. Ipsen (1995), 'Relative perturbation techniques for singular value problems', *SIAM J. Numer. Anal.* **32**(6), 1972–88.

S. Eisenstat and I. Ipsen (1996), Relative perturbation results for eigenvalues and eigenvectors of diagonalisable matrices, Technical Report CRSC-TR96-6, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University.

S. Eisenstat and I. Ipsen (1997), Absolute perturbation bounds for matrix eigenvalues imply relative bounds, Technical Report CRSC-TR97-16, Center for Research in Scientific Computation, Department of Mathematics, North Carolina State University.

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

K. Fernando and B. Parlett (1994), 'Accurate singular values and differential qd algorithms', *Numer. Math.* **67**(2), 191–229.

M. Gu and S. Eisenstat (1993), Relative perturbation theory for eigenproblems, Research Report YALEU/DCS/RR-934, Department of Computer Science, Yale University.

V. Hari and Z. Drmač (1997), 'On scaled almost diagonal Hermitian matrix pairs', *SIAM J. Matrix Anal. Appl.*

R. Horn and C. Johnson (1985), *Matrix Analysis*, Cambridge University Press.

W. Kahan (1966), Accurate eigenvalues of a symmetric tri-diagonal matrix, Technical Report CS41, Computer Science Department, Stanford University. Revised June 1968.

C. Li and R. Mathias (1997), On the Lidskii-Mirsky-Wielandt theorem, Technical report, Department of Mathematics, College of William and Mary, Williamsburg, VA.

R. Li (1994*a*), Relative perturbation theory: (I) eigenvalue variations, LAPACK working note 84, Computer Science Department, University of Tennessee, Knoxville. Revised May 1997.

R. Li (1994*b*), Relative perturbation theory: (II) eigenspace variations, LAPACK working note 85, Computer Science Department, University of Tennessee, Knoxville. Revised May 1997.

R. Li (1997), 'Relative perturbation theory: (III) more bounds on eigenvalue variation', *Linear Algebra Appl.* to appear.

K. Löwner (1934), 'Über monotone Matrix Funktionen', *Math. Z.* **38**, 177–216.

R. Mathias (1994), Spectral perturbation bounds for positive definite matrices, Technical report, Department of Mathematics, College of William and Mary, Williamsburg, VA. Revised May 1997.

R. Mathias (1995*a*), 'Accurate eigensystem computations by Jacobi methods', *SIAM J. Matrix Anal. Appl.* **16**(3), 977–1003.

R. Mathias (1995*b*), 'A bound for matrix square root with application to eigenvector perturbation', *SIAM J. Matrix Anal. Appl.* to appear.

R. Mathias (1996), 'Fast accurate eigenvalue methods for graded positive-definite matrices', *Numer. Math.* **74**, 85–103.

R. Mathias and G. Stewart (1993), 'A block QR algorithm and the singular value decomposition', *Linear Algebra Appl.* **182**, 91–100.

R. Mathias and K. Veselić (1995), A relative perturbation bound for positive-definite matrices, revised December 1996.

A. Ostrowski (1959), 'A quantitative formulation of Sylvester's law of inertia', *Proc. Nat. Acad,. Sci.* **45**, 740–4.

B. Parlett (1980), *The Symmetric Eigenvalue Problem*, Prentice Hall.

B. Parlett (1995), The new qd algorithms, in *Acta Numerica 1995*, Cambridge University Press, pp. 459–91.

B. Parlett (1997), Spectral sensitivity of products of bidiagonals.

E. Pietzsch (1993), Genaue Eigenwertberechnung Nichtsingulärer Schiefsymmetrischer Matrizen, PhD thesis, Fachbereich Mathematik, Fernuniversität Gesamthochschule Hagen, Germany.

R. Rosanoff, J. Glouderman and S. Levy (1968), Numerical conditions of stiffness matrix formulations for frame structures, in *Proceedings of the Conference on Matrix Methods in Structural Mechanics*, AFFDL-TR-68-150, Wright-Patterson Air Force Base, Ohio, pp. 1029–60.

I. Slapničar (1992), Accurate Symmetric Eigenreduction by a Jacobi Method, PhD thesis, Fernuniversität Hagen, Germany.

I. Slapničar and K. Veselič (1992), 'Perturbations of the eigenprojections of a factorised Hermitian matrix', *Linear Algebra Appl.* to appear.

N. Truhar and I. Slapničar (1997), Relative perturbation bounds for invariant subspaces of indefinite Hermitian matrices.

A. van der Sluis (1969), 'Condition, equilibration, and pivoting in linear algebraic systems', *Numer. Math.* **15**, 74–86.

K. Veselić and I. Slapničar (1993), 'Floating-point perturbations of Hermitian matrices', *Linear Algebra Appl.* **195**, 81–116.