A Survey of Componentwise Perturbation Theory

Higham, Nicholas J.

1994

MIMS EPrint: **2006.404**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: http://eprints.maths.manchester.ac.uk/
And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

Proceedings of Symposia in Applied Mathematics

A Survey of Componentwise Perturbation Theory in Numerical Linear Algebra

Nicholas J. Higham

Abstract. Perturbation bounds in numerical linear algebra are traditionally derived and expressed using norms. Norm bounds cannot reflect the scaling or sparsity of a problem and its perturbation, and so can be unduly weak. If the problem data and its perturbation are measured componentwise, much smaller and more revealing bounds can be obtained. A survey is given of componentwise perturbation theory in numerical linear algebra, covering linear systems, the matrix inverse, matrix factorizations, the least squares problem, and the eigenvalue and singular value problems. Most of the results described have been published in the last five years.

Our hero is the intrepid, yet sensitive matrix A.
Our villain is E, who keeps perturbing A.
When A is perturbed he puts on a crumpled hat: $\widetilde{A} = A + E$.
G. W. Stewart and J.-G. Sun, *Matrix Perturbation Theory* (1990)

1. Introduction

Matrix analysis would not have developed into the vast subject it is today without the concept of representing a matrix by a single symbol. Similarly, perturbation theory would not be such a rich and useful area if it were not for norms. A norm compresses the mn numbers in an $m \times n$ matrix into a single scalar measure of size, enabling a perturbation result to be presented in a form that is easy to interpret and gives insight. In the recent book *Matrix Perturbation Theory* [**71**] by Stewart and Sun, norms are exploited throughout.

1991 *Mathematics Subject Classification.* Primary 65F35, 65G05.

Key words and phrases. Componentwise perturbation bounds, componentwise backward error, linear systems, matrix inverse, matrix factorizations, least squares problem, underdetermined system, eigenvalue problem, singular value problem.

This work was supported by Science and Engineering Research Council grant GR/H52139.

This paper is in final form and no version of it will be submitted for publication elsewhere.

Citation details: N. J. Higham. A survey of componentwise perturbation theory in numerical linear algebra. In W. Gautschi, editor, *Mathematics of Computation 1943–1993: A Half Century of Computational Mathematics*, volume 48 of *Proceedings of Symposia in Applied Mathematics*, pages 49–77. American Mathematical Society, Providence, RI, USA, 1994.

Although a great deal can be achieved with norms, they have two main failings. First, norms ignore structure in the form of both scaling and sparsity. Norms can tell us the overall size of a perturbation but not how that size is distributed among the elements it perturbs, and this information can be important when the data is badly scaled or contains many zeros. The second failing is that norm bounds can lack sharpness (particularly in a rounding error analysis), because they are often the result of several applications of the triangle inequality and the submultiplicative inequality, in not all of which there can be equality simultaneously.

It is for these reasons that an alternative style of analysis has become popular, in which the use of norms is kept to a minimum (but not abandoned altogether); it is known as *componentwise analysis*. The idea is not new. Its history can be very roughly outlined as follows (more details are given in the following sections).

Wilkinson frequently obtained componentwise backward error results as intermediate steps in an error analysis, but usually stated the final result in normwise form. Some of the earliest examples are contained in his 1960 paper [**79**] and 1963 book [**80**]. In 1964, Oettli and Prager [**54**] obtained a formula for the componentwise backward error of an approximate solution to a linear system. Bauer [**8**] applied componentwise perturbation analysis to linear systems and to matrix inversion in 1966. Stoer and Bulirsch gave a componentwise error analysis of Gaussian elimination in the original 1972 German edition of *Introduction to Numerical Analysis* [**72**]. In a series of papers Skeel explored componentwise perturbation analysis for linear systems and componentwise error analysis for Gaussian elimination, obtaining new results about the stability of Gaussian elimination and the behaviour of iterative refinement [**63**], [**64**], [**65**]. Skeel's papers did not attract immediate attention in the numerical linear algebra community, but gradually came to be appreciated during the 1980s. Since the mid 1980s componentwise analysis has been widely used for linear systems and linear system solvers, and componentwise perturbation theory and error analysis has been developed for eigenproblems and eigensolvers. Perhaps the best indication of the utility of componentwise analysis is that it has been fully exploited in the development of LAPACK, the state-of-the-art package of Fortran programs for solving linear equation and eigenvalue problems [**1**].

Although the componentwise approach is now well known, at least among numerical analysts, it is hardly represented in current textbooks. For example, Golub and Van Loan [**34**] and Stewart and Sun [**71**] give componentwise error analysis for linear equations but not for eigenvalue problems, while Horn and Johnson [**47**], [**48**] give no componentwise perturbation results at all.

In this paper we survey componentwise perturbation theory in numerical linear algebra. We do not consider rounding error analysis, except briefly when it helps to illuminate the perturbation theory. No survey of componentwise rounding error analysis is currently available, but some relevant references for standard problems, which themselves include further references, are as follows:

- [**41**] for substitution for triangular systems,
- [**43**] for Gaussian elimination,
- [**19**] for Cholesky factorization,
- [**25**] for matrix inversion,
- [**46**] for stationary iterative methods,
- [**44**] for QR factorization and its application to least squares problems,
- [**22**] for underdetermined system solvers,

- [**23**] for the QR algorithm for the bidiagonal SVD problem,
- [**24**] for Jacobi's method for the symmetric eigenproblem and SVD.

In §2 we consider backward and forward perturbation bounds for square linear systems. The matrix inverse is considered in §3. Section 4 is concerned with the least squares problem, for which both normwise and componentwise results have recently been obtained; underdetermined systems are also discussed in this section. Matrix factorizations are discussed in §5. The active area of componentwise perturbation theory for eigenvalue and singular value problems is described in §6. Finally, some concluding remarks are given in §7.

It is possible to develop probabilistic perturbation theory for all the problems considered here, by making assumptions about the statistical distribution of the perturbations. We do not consider this approach, but refer the interested reader to the papers by Fletcher [**29**] and Stewart [**69**].

2. Linear Systems

Throughout this section we are concerned with a linear system $Ax = b$, where $A \in \mathbb{R}^{n\times n}$. In the context of uncertain data or inexact arithmetic there are two important questions:

(1) How much do we have to perturb the data A and b for an approximate solution y to be the exact solution of the perturbed system—in other words, what is the backward error of y?

(2) How much does x change if we perturb A and b, that is, how sensitive is the solution to perturbations in the data?

When we use norms to measure size, the answers are given by the following classical results. We denote by $\|\cdot\|$ any vector norm and the corresponding subordinate matrix norm, and $\kappa(A) = \|A\|\|A^{-1}\|$ is the matrix condition number. The matrix E and the vector f have nonnegative entries and represent tolerances against which the perturbations are measured (their role becomes clear when we consider componentwise results).

The first theorem is due to Rigal and Gaches [**58**], and was also given by Kovarik [**52**].

THEOREM 2.1 (Rigal and Gaches). *The* normwise backward error

$$\eta_{E,f}(y) := \min\{\epsilon : (A+\Delta A)y = b + \Delta b,\ \|\Delta A\| \le \epsilon\|E\|,\ \|\Delta b\| \le \epsilon\|f\|\} \tag{2.1}$$

is given by

$$\eta_{E,f}(y) = \frac{\|r\|}{\|E\|\|y\| + \|f\|}, \tag{2.2}$$

where $r = b - Ay$.

PROOF. It is straightforward to show that the right-hand side of (2.2) is a lower bound for $\eta(y)$. This lower bound is attained for the perturbations

$$\Delta A_{\min} = \frac{\|E\|\|y\|}{\|E\|\,\|y\| + \|f\|} r z^T, \qquad \Delta b_{\min} = -\frac{\|f\|}{\|E\|\,\|y\| + \|f\|} r, \tag{2.3}$$

where z is a vector dual to y, that is,

$$z^T y = \|z\|_D\|y\| = 1, \qquad \text{where} \qquad \|z\|_D = \max_{v\ne 0} \frac{|z^T v|}{\|v\|}. \quad \square$$

For the particular choice $E = |A|$ and $f = |b|$, the quantity $\eta_{E,f}(y)$ is called the *normwise relative backward error*.

The next result measures the sensitivity of the system.

THEOREM 2.2. *Let $Ax = b$ and $(A+\Delta A)y = b+\Delta b$, where $\|\Delta A\| \le \epsilon\|E\|$ and $\|\Delta b\| \le \epsilon\|f\|$, and assume that $\epsilon\|A^{-1}\|\|E\| < 1$. Then*

$$\frac{\|x-y\|}{\|x\|} \le \frac{\epsilon}{1-\epsilon\|A^{-1}\|\|E\|}\left(\frac{\|A^{-1}\|\|f\|}{\|x\|} + \|A^{-1}\|\|E\|\right), \tag{2.4}$$

and this bound is attainable to first order in ϵ.

PROOF. The bound (2.4) is straightforward to derive. It is attained to first order in ϵ for $\Delta A = \epsilon\|E\|\|x\|wv^T$ and $\Delta b = -\epsilon\|f\|w$, where $\|w\| = 1$, $\|A^{-1}w\| = \|A^{-1}\|$, and v is a vector dual to x. □

Associated with the way of measuring perturbations used in these two theorems is the normwise condition number

$$\kappa_{E,f}(A,x) := \lim_{\epsilon\to 0}\sup\Big\{ \frac{\|\Delta x\|}{\epsilon\|x\|} : (A+\Delta A)(x+\Delta x) = b+\Delta b, \quad \|\Delta A\| \le \epsilon\|E\|, \quad \|\Delta b\| \le \epsilon\|f\| \Big\}.$$

Because the bound of Theorem 2.2 is sharp, it follows that

$$\kappa_{E,f}(A,x) = \frac{\|A^{-1}\|\|f\|}{\|x\|} + \|A^{-1}\|\|E\|.$$

For the choice $E = |A|$ and $f = |b|$ (where $|A|$ denotes the matrix $(|a_{ij}|)$), we have $\kappa(A) \le \kappa_{E,f}(A,x) \le 2\kappa(A)$, and the bound (2.4) can be weakened slightly to yield the familiar form

$$\frac{\|x-y\|}{\|x\|} \le \frac{2\epsilon\kappa(A)}{1-\epsilon\kappa(A)}.$$

A numerical example illustrates the above results. Let A be the 8×8 Vandermonde matrix with (i,j) element $j^{2(i-1)}$, and let $b = e_1$, the first unit vector, so that x is the first column of A^{-1}. We take y to be the approximate solution to $Ax = b$ computed by Gaussian elimination with partial pivoting. All our experiments are performed in Matlab, for which the unit roundoff $u \approx 1.1\times 10^{-16}$. We find that $\eta_\infty(y) = 3.05\times 10^{-21}$ for $E = |A|$ and $f = |b|$, and $\kappa_\infty(A) = 1.68\times 10^{13}$. This is an admirably small backward error, but it may be uninformative for two reasons. First, the elements of A vary over 12 orders of magnitude, so while our backward error perturbations are small compared with the large elements of A, we may be making large perturbations in the small elements (indeed we *are* in this particular example). Second, we are perturbing the zero elements of b (as can be seen from (2.3) together with the fact that for this example the residual r has no zero entries); this is unsatisfactory if we wish to regard y as the first column of the inverse of a perturbed matrix.

Next, let $b = Ae$, where $e = [1,1,\ldots,1]^T$, and let z be the solution to the perturbed system $(A+\Delta A)z = b+\Delta b$, where $\Delta A = \mathrm{tol}|A|$ and $\Delta b = \mathrm{tol}|b|$, with $\mathrm{tol} = 8u$. We find that

$$\frac{\|x-z\|_\infty}{\|x\|_\infty} = 2.40\times 10^{-12}, \tag{2.5}$$

while the corresponding bound from (2.4) with $\epsilon = \mathrm{tol}$, $E = |A|$ and $f = |b|$ is 3.03×10^{-2}. Thus the normwise forward error bound is extremely pessimistic for this special choice of perturbation.

To obtain a more satisfactory backward error measure and a sharper perturbation bound, we need componentwise analysis.

The *componentwise backward error* is defined as

$$\omega_{E,f}(y) = \min\{\epsilon : (A + \Delta A)y = b + \Delta b, \ \ |\Delta A| \le \epsilon E, \ \ |\Delta b| \le \epsilon f\}, \tag{2.6}$$

where inequalities between matrices hold componentwise. In this definition each element of a perturbation is measured relative to its individual tolerance, so, unlike in the normwise definition, we are making full use of the $n^2 + n$ parameters in E and f.

How should E and f be chosen? The most common choice of tolerances is $E = |A|$ and $f = |b|$, which yields the *componentwise relative backward error*. For this choice

$$a_{ij} = 0 \ \Rightarrow \ \Delta a_{ij} = 0 \quad \text{and} \quad b_i = 0 \ \Rightarrow \ \Delta b_i = 0$$

in (2.6), and so if $\omega_{E,f}(y)$ is small then y solves a problem that is close to the original one in the sense of componentwise relative perturbations and has the same sparsity pattern. Another attractive property of the componentwise relative backward error is that it is insensitive to the scaling of the system: if $Ax = b$ is scaled to $(S_1 A S_2)(S_2^{-1}x) = S_1 b$, where S_1 and S_2 are diagonal, and y is scaled to $S_2^{-1}y$, then ω remains unchanged.

The choice $E = |A|ee^T$, $f = |b|$ gives a row-wise backward error. The constraint $|\Delta A| \le \epsilon E$ is now $|\Delta a_{ij}| \le \epsilon \alpha_i$, where α_i is the 1-norm of the ith row of A, so perturbations to the ith row of A are being measured relative to the norm of that row. A columnwise backward error can be formulated in a similar way.

The third natural choice of tolerances is $E = \|A\|ee^T$ and $f = \|b\|e$, for which $\omega_{E,f}(y)$ is the same as the normwise backward error $\eta_{E,f}(y)$, up to a constant.

As for the normwise backward error in Theorem 2.1, there is a simple formula for $\omega_{E,f}(y)$.

THEOREM 2.3 (Oettli and Prager [**54**]). *The componentwise backward error is given by*

$$\omega_{E,f}(y) = \max_i \frac{|r_i|}{(E|y| + f)_i}, \tag{2.7}$$

where $r = b - Ay$, and $\xi/0$ is interpreted as zero if $\xi = 0$ and infinity otherwise.

PROOF. It is easy to show that the right-hand side of (2.7) is a lower bound for $\omega(y)$, and that this bound is attained for the perturbations

$$\Delta A = D_1 E D_2, \qquad \Delta b = -D_1 f, \tag{2.8}$$

where $D_1 = \mathrm{diag}\big(r_i/(E|y| + f)_i\big)$ and $D_2 = \mathrm{diag}\big(\mathrm{sign}(y_i)\big)$. □

The next result gives a forward error bound corresponding to the componentwise backward error. It is a straightforward generalization of a result of Skeel [**63**, Theorems 2.1 and 2.2]. A monotonic norm is one for which $|x| \le |y|$ implies $\|x\| \le \|y\|$, which can be shown to be equivalent to $\| \, |x| \, \| = \|x\|$ [**71**, p. 52].

THEOREM 2.4. *Let $Ax = b$ and $(A + \Delta A)y = b + \Delta b$, where $|\Delta A| \le \epsilon E$ and $|\Delta b| \le \epsilon f$, and assume that $\epsilon \| \, |A^{-1}|E \, \| < 1$, where $\|\cdot\|$ is a monotonic norm. Then*

$$\frac{\|x - y\|}{\|x\|} \le \frac{\epsilon}{1 - \epsilon \| \, |A^{-1}|E \, \|} \frac{\| \, |A^{-1}|f + |A^{-1}|E|x| \, \|}{\|x\|}, \tag{2.9}$$

and for the ∞-norm this bound is attainable to first order in ϵ.

PROOF. The bound (2.9) is straightforward to derive. For the ∞-norm it is attained, to first order in ϵ, for $\Delta A = \epsilon D_1 E D_2$ and $\Delta b = -\epsilon D_1 f$, where $D_2 = \mathrm{diag}(\mathrm{sign}(x_i))$ and $D_1 = \mathrm{diag}(\xi_j)$, where $\xi_j = \mathrm{sign}(A^{-1})_{kj}$ and $\| \, |A^{-1}|f + |A^{-1}|E|x| \, \|_\infty = \big(|A^{-1}|f + |A^{-1}|E|x|\big)_k$. □

When the components of x vary widely in magnitude, the bound (2.9) can be weak for the smaller components. It is possible to obtain perturbation bounds for individual solution components by refraining from taking norms in the analysis. Chandrasekaran and Ipsen present this type of analysis [**14**]; see also the condition number (2.14) below. Componentwise solution bounds for Markov chains are obtained by Ipsen and Meyer [**49**], who measure the perturbation matrix normwise, and by O'Cinneide [**53**], who measures the perturbation matrix componentwise.

Theorem 2.4 implies that the condition number

$$\begin{aligned} \mathrm{cond}_{E,f}(A,x) \quad := \quad \lim_{\epsilon \to 0} \sup \Big\{ \frac{\|\Delta x\|_\infty}{\epsilon \|x\|_\infty} : (A + \Delta A)(x + \Delta x) = b + \Delta b, \\ |\Delta A| \le \epsilon E, \quad |\Delta b| \le \epsilon f \Big\} \end{aligned}$$

is given by

$$\mathrm{cond}_{E,f}(A,x) = \frac{\| \, |A^{-1}|f + |A^{-1}|E|x| \, \|_\infty}{\|x\|_\infty}. \tag{2.10}$$

For the special case $E = |A|$ and $f = |b|$ we have the condition numbers introduced by Skeel [**63**]:

$$\mathrm{cond}(A,x) := \frac{\| \, |A^{-1}||A||x| \, \|_\infty}{\|x\|_\infty},$$

which differs from $\mathrm{cond}_{|A|,|b|}(A,x)$ by at most a factor 2, and

$$\mathrm{cond}(A) := \mathrm{cond}(A,e) = \| \, |A^{-1}||A| \, \|_\infty \le \kappa_\infty(A). \tag{2.11}$$

Note that $\mathrm{cond}(A)$ is invariant under row scaling $Ax = b \to (DA)x = Db$, where D is diagonal, and for this reason it can be arbitrarily smaller than $\kappa_\infty(A)$. In fact, it is straightforward to show that

$$\min\{ \, \kappa_\infty(DA) : D \text{ diagonal} \, \} = \mathrm{cond}(A), \tag{2.12}$$

where the optimal scaling D equilibrates the rows of A, that is, DA has rows of unit 1-norm.

How does cond compare with κ? Chandrasekaran and Ipsen [**14**] note the following inequalities. First, if D_R equilibrates the rows of A ($D_R|A|e = e$) then

$$\frac{\kappa_\infty(A)}{\kappa_\infty(D_R)} \le \mathrm{cond}(A) \le \kappa_\infty(A)$$

(these inequalities imply (2.12)). Thus $\mathrm{cond}(A)$ can be much smaller than $\kappa_\infty(A)$ only when the rows of A are badly scaled. Second, if D_C equilibrates the columns

of A ($e^T|A|D_C = e^T$) then

$$\frac{\kappa_1(A)}{n\kappa_\infty(D_C)} \min_j \frac{\|A^{-1}e_j\|_\infty}{\|A^{-1}\|_1} \le \mathrm{cond}(A,x) \le \kappa_\infty(A).$$

These inequalities show that $\mathrm{cond}(A,x)$ can be much smaller than $\kappa_\infty(A)$ only when the columns of A are badly scaled or the columns of A^{-1} are badly scaled.

Returning to our numerical example, we find that $\omega_{E,f}(y) = 1.10 \times 10^{-12}$ for $E = |A|$ and $f = |b|$ or $f = 0$. This tells us that for y to be the first column of the inverse of a perturbed matrix we must make relative changes to A four orders of magnitude larger than the unit roundoff. For the perturbed system, Theorem 2.4 with $\epsilon = \mathrm{tol}$, $E = |A|$ and $f = |b|$ gives the bound

$$\frac{\|x - z\|_\infty}{\|x\|_\infty} \le 4.08 \times 10^{-10},$$

which is eight orders of magnitude smaller than the normwise bound from Theorem 2.2, and only a factor 170 larger than the actual forward error (2.5).

A numerical example of Kahan [**51**] is also instructive. Let

$$A = \begin{bmatrix} 2 & -1 & 1 \\ -1 & \epsilon & \epsilon \\ 1 & \epsilon & \epsilon \end{bmatrix}, \quad b = \begin{bmatrix} 2(1+\epsilon) \\ -\epsilon \\ \epsilon \end{bmatrix},$$

where $0 < \epsilon \ll 1$, so that $x = [\epsilon, -1, 1]^T$. The normwise condition number $\kappa_\infty(A) = 2(1+\epsilon^{-1})$, so the system is very sensitive to arbitrary perturbations in A and b. Moreover,

$$|A^{-1}||A| = \begin{bmatrix} 1 & \epsilon & \epsilon \\ \dfrac{2\epsilon+1}{2\epsilon} & 1 & 1 \\ \dfrac{2\epsilon+1}{2\epsilon} & 1 & 1 \end{bmatrix},$$

so $\mathrm{cond}(A) = 3 + (2\epsilon)^{-1}$, which implies that the system is also very sensitive to componentwise perturbations for some right-hand sides. However, $\mathrm{cond}(A,x) = 5/2 + \epsilon$, so *for this particular* b the system is very well-conditioned under componentwise perturbations.

We mention in passing another interesting role of the matrix $|A^{-1}||A|$. Recall from (2.12) that $\min_D \kappa_\infty(DA) = \| \, |A^{-1}||A| \, \|_\infty$. Bauer [**7**] shows that for two-sided scalings

$$\min\{ \kappa_\infty(D_1 A D_2) : D_1,\, D_2 \text{ diagonal} \} \ge \rho(|A^{-1}||A|), \tag{2.13}$$

where ρ is the spectral radius, and he characterizes the minimizing D_1 and D_2 as $\mathrm{diag}(x_1)^{-1}$ and $\mathrm{diag}(x_2)^{-1}$, where x_1 and x_2 are Perron vectors of $|A||A^{-1}|$ and $|A^{-1}||A|$, respectively (thus the minimum is achieved whenever these Perron vectors have positive entries, which is guaranteed if A is irreducible—see Businger [**13**]). For the Kahan example,

$$\rho(|A^{-1}||A|) \approx 2.62 + 1.79\epsilon \ll 3 + (2\epsilon)^{-1} = \| \, |A^{-1}||A| \, \|_\infty,$$

and, in fact, $\kappa_\infty(DAD) = 3$ for $D = \mathrm{diag}(\epsilon^{1/2}, \epsilon^{-1/2}, \epsilon^{-1/2})$, so a symmetric two-sided scaling is nearly optimal in this case.

2.1. Further notes and references. It is clear from Bauer's comments in [**8**] that the bound (2.9), with $E = |A|$ and $f = |b|$, was known to him, though he does not state the bound. This is the earliest reference we know in which componentwise analysis is used to derive forward perturbation bounds (for more details of Bauer's paper see §3).

Theorems 2.1 and 2.3 both remain valid when A is rectangular. Componentwise backward error for rectangular A was considered by Oettli, Prager and Wilkinson [**55**], but their results are subsumed by those of Oettli and Prager [**54**] and Rigal and Gaches [**58**].

The componentwise analyses can be generalized in three main ways:

(1) We can use more general measures of size for the data and the solution. Higham and Higham [**39**] measure ΔA, Δb and Δx by $\nu_p([(\Delta a_{ij}/e_{ij})\ (\Delta b_i/f_i)])$ and $\nu_p((\Delta x_i/g_i))$, where $\nu_p(A) = \left(\sum_{i,j} |a_{ij}|^p\right)^{1/p}$, $1 \le p \le \infty$, and the e_{ij}, f_i and g_i are tolerances. In [**39**] the corresponding backward error is shown to be given by the explicit formula

$$\left\| \left(\frac{r_j}{\| D_j \left[\begin{smallmatrix} y \\ -1 \end{smallmatrix} \right] \|_q} \right) \right\|_p,$$

where $r = b - Ay$, $D_j = \operatorname{diag}(e_{j1}, \ldots, e_{jn}, f_j)$ and $p^{-1} + q^{-1} = 1$; bounds for the corresponding condition number are also obtained. Theorem 2.3, and Theorem 2.4 with the ∞-norm, correspond to $p = \infty$ and $g_i \equiv \|x\|_\infty$. If we take $p = \infty$ and $g = |x|$, we are measuring the change in the solution in a componentwise relative sense, and the condition number is [**39**]

$$\| \operatorname{diag}(|x_i|)^{-1} |A^{-1}| (E|x| + f) \|_\infty. \tag{2.14}$$

This latter case has also been considered by Rohn [**59**] and Gohberg and Koltracht [**32**].

(2) The analysis can be extended to systems with multiple right-hand sides [**39**]. For the general ν_p measure described in (1), the backward error can be computed by finding the minimum p-norm solutions to n underdetermined linear systems.

(3) Structure in A and b can be preserved in the analysis. For example, if A is symmetric or Toeplitz then its perturbation can be forced to be symmetric or Toeplitz too, while still using componentwise measures. References include Higham and Higham [**38**] and Gohberg and Koltracht [**32**] for linear structure, and Bartels and D. J. Higham [**6**] for Vandermonde structure. A symmetry-preserving normwise backward error is explored by Bunch, Demmel and Van Loan [**12**], and symmetry-preserving normwise condition numbers are considered by D. J. Higham [**37**].

Practical issues associated with the use of componentwise backward error and perturbation bounds in the context of sparse matrices are considered by Arioli, Demmel and Duff [**2**]. In particular, they show how to estimate the condition number (2.10) cheaply, given a factorization of A. Their approach avoids explicit computation of A^{-1} by manipulating the condition number into a form that can be estimated using a matrix norm estimator of Hager [**36**] and Higham [**40**]. This approach can be adapted to estimate virtually any form of componentwise condition number for a linear system.

Finally, we mention how the analysis of this section is reflected in LAPACK. The LAPACK expert driver routine `xGESVX`[1] solves a linear system $Ax = b$ by

[1]The leading '`x`' in routine names stands for the Fortran data type: single precision (`S`), double precision (`D`), complex (`C`) or double precision complex (`Z`).

Gaussian elimination with partial pivoting and then applies iterative refinement in fixed precision. The main termination criterion for the refinement is $\omega_{|A|,|b|}(\widehat{x}) \le u$, where $\widehat{x}$ is the current iterate and u is the unit roundoff. This routine returns as a forward error estimate the upper bound in

$$\frac{\|x-\widehat{x}\|_\infty}{\|x\|_\infty} \le \frac{\| \, |A^{-1}|(|\widehat{r}|+\xi) \, \|_\infty}{\|x\|_\infty},$$

where $\widehat{r} = fl(b - A\widehat{x})$ and $\xi = (n+1)u(|A||\widehat{x}| + |b|)$ bounds the rounding errors in forming $\widehat{r}$. This a posteriori bound is used in preference to (2.9) because it is expected to be smaller in general, since less inequalities are used in its derivation.

3. The Matrix Inverse

Much of the analysis in §2 can be adapted to the matrix inverse. If $A \in \mathbb{R}^{n\times n}$ and $A + \Delta A$ are nonsingular then, writing $X \equiv A^{-1}$ and $X + \Delta X \equiv (A + \Delta A)^{-1}$, we have $\Delta X = -A^{-1}\Delta A A^{-1} - A^{-1}\Delta A \Delta X$. Therefore derivation of perturbation bounds and condition numbers involves consideration of the first-order term $A^{-1}\Delta A A^{-1}$. For normwise measures it is well known that the condition number for inversion is the usual matrix condition number $\kappa(A)$. This result can be stated in a precise and general way as follows [**37**]: for the mixed subordinate norm $\|A\|_{\alpha,\beta} = \max \|Ax\|_\beta / \|x\|_\alpha$,

$$\mu_{\alpha,\beta}(A) := \lim_{\epsilon\to 0} \sup\Big\{ \frac{\|\Delta X\|_{\beta,\alpha}}{\epsilon \|X\|_{\beta,\alpha}} : \|\Delta A\|_{\alpha,\beta} \le \epsilon \|A\|_{\alpha,\beta} \Big\} = \|A\|_{\alpha,\beta}\|A^{-1}\|_{\beta,\alpha}.$$

Note that this definition uses the $\|\cdot\|_{\alpha,\beta}$ norm on the data space and the $\|\cdot\|_{\beta,\alpha}$ norm on the solution space, as is natural. A componentwise condition number is

$$\mu_E(A) := \lim_{\epsilon\to 0} \sup\Big\{ \frac{\|\Delta X\|_\infty}{\epsilon \|X\|_\infty} : |\Delta A| \le \epsilon E \Big\} \le \frac{\| \, |A^{-1}|E|A^{-1}| \, \|_\infty}{\|A^{-1}\|_\infty}.$$

In general, the inequality is strict, but there is equality when $|A^{-1}| = D_1 A^{-1} D_2$ for D_i of the form $\mathrm{diag}(\pm 1)$, [**30**, Theorem 1.10], [**31**]. For perturbations satisfying $|\Delta A| \le \epsilon |A|$, Bauer [**8**, p. 413] obtains the bound

$$|\Delta X| \le (I - \epsilon |A^{-1}||A|)^{-1}|A^{-1}||A||A^{-1}|\epsilon,$$

and makes various deductions from it.

Measuring ΔX componentwise relative to X, and taking $E = |A|$, gives the condition number

$$\mu_2(A) := \lim_{\epsilon\to 0} \sup\Big\{ \max_{i,j} \frac{|\Delta x_{ij}|}{|x_{ij}|} : |\Delta A| \le \epsilon |A| \Big\} = \max_{i,j} \frac{\big(|A^{-1}||A||A^{-1}|\big)_{ij}}{\big(|A^{-1}|\big)_{ij}}.$$

This condition number appears in a perturbation bound of Bauer [**8**, p. 413] and was derived by Rohn [**59**].

The natural definition of backward error of an approximate inverse Y of A is a smallest ΔA such that $(A + \Delta A)Y = I$. If Y is nonsingular there is a unique ΔA satisfying this equation and the problem of computing the backward error is trivial.

It is well known that the normwise matrix condition number is the reciprocal of the relative distance to singularity. Specifically, Kahan [**51**, pp. 775–76] shows that

$$\min\Big\{ \frac{\|\Delta A\|_{\alpha,\beta}}{\|A\|_{\alpha,\beta}} : A + \Delta A \text{ singular} \Big\} = \mu_{\alpha,\beta}^{-1}.$$

The componentwise distance to singularity,

$$d_E(A) = \min\{\, \epsilon : A + \Delta A \text{ singular}, \ |\Delta A| \le \epsilon E \,\},$$

has been characterized by Rohn [**60**], [**61**] as

$$d_E(A) = \frac{1}{\max_{S_1,S_2} \rho_0(S_1 A^{-1} S_2 E)},$$

where the maximum is taken over all signature matrices $S_i = \mathrm{diag}(\pm 1)$ and where

$$\rho_0(X) = \max\{\, |\lambda| : \lambda \text{ is a real eigenvalue of } A \,\}.$$

This formula involves 4^n eigenproblems, so is computationally intractable (in fact it is NP-hard [**57**]).

Demmel [**20**] shows by complexity arguments that there can be no simple relationship between $d_E(A)$ and the quantity $\| \, |A^{-1}|E \, \|_\infty$, which is an upper bound for $\mu_E(A)$. He also presents evidence for the conjecture that

$$\frac{1}{\rho(|A^{-1}||A|)} \le d_{|A|}(A) \le \frac{\gamma_n}{\rho(|A^{-1}||A|)}$$

for a constant γ_n. The lower bound always holds and Demmel identifies several classes of matrices for which the upper bound holds. This conjecture is both plausible and aesthetically pleasing because $d_{|A|}(A)$ is invariant under two-sided diagonal scalings of A and $\rho(|A^{-1}||A|)$ is the minimum condition number achievable by such scalings, as shown by (2.13).

4. The Least Squares Problem

In this section we consider the least squares (LS) problem $\min_x \|Ax - b\|_2$, where $A \in \mathbb{R}^{m\times n}$ $(m \ge n)$ has full rank. Normwise perturbation theory for the LS problem has been developed by various authors. The earliest bounds were obtained by Golub and Wilkinson [**33**]. The following theorem of Wedin is taken from [**78**, Theorem 5.1]. Here, $\kappa_2(A) = \|A\|_2\|A^+\|_2$, where A^+ is the pseudo-inverse.

THEOREM 4.1 (Wedin). *Let $A \in \mathbb{R}^{m\times n}$ $(m \ge n)$ and $A + \Delta A$ both be of full rank, and let*

$$\begin{aligned} \|Ax - b\|_2 = \min, \quad r = b - Ax, \\ \|(A + \Delta A)y - (b + \Delta b)\|_2 = \min, \\ \|\Delta A\|_2 \le \epsilon \|A\|_2, \quad \|\Delta b\|_2 \le \epsilon \|b\|_2. \end{aligned}$$

Provided that $\kappa_2(A)\epsilon < 1$,

$$\frac{\|x - y\|_2}{\|x\|_2} \le \frac{\kappa_2(A)\epsilon}{1 - \kappa_2(A)\epsilon} \left(1 + \frac{\|b\|_2}{\|A\|_2\|x\|_2} + \kappa_2(A) \frac{\|r\|_2}{\|A\|_2\|x\|_2} \right). \quad \square \tag{4.1}$$

Bounds can also be given for the change in the residual, but we do not consider them here.

The bound (4.1) is usually interpreted as saying that the sensitivity of the LS problem is measured by $\kappa_2(A)$ when the residual is small or zero and by $\kappa_2(A)^2$ otherwise. For further discussion see standard textbooks (for example, [**34**, pp. 230–231]).

Surprisingly, it is easier to derive componentwise perturbation bounds than normwise ones for the LS problem. The key idea is to express the LS solution and its residual as the solution of the augmented system

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} r \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}, \tag{4.2}$$

which is simply another way of writing the normal equations, $A^T Ax = A^T b$. This is a square nonsingular system, so standard techniques can be applied. The perturbed system of interest is

$$\begin{bmatrix} I & A+\Delta A \\ (A+\Delta A)^T & 0 \end{bmatrix} \begin{bmatrix} s \\ y \end{bmatrix} = \begin{bmatrix} b+\Delta b \\ 0 \end{bmatrix}, \tag{4.3}$$

where we assume that

$$|\Delta A| \le \epsilon E, \quad |\Delta b| \le \epsilon f. \tag{4.4}$$

From (4.2) and (4.3) we obtain

$$\begin{bmatrix} I & A \\ A^T & 0 \end{bmatrix} \begin{bmatrix} s-r \\ y-x \end{bmatrix} = \begin{bmatrix} \Delta b - \Delta A y \\ -\Delta A^T s \end{bmatrix}.$$

Premultiplying by the inverse of the matrix on the left gives

$$\begin{bmatrix} s-r \\ y-x \end{bmatrix} = \begin{bmatrix} I - AA^+ & (A^+)^T \\ A^+ & -(A^T A)^{-1} \end{bmatrix} \begin{bmatrix} \Delta b - \Delta A y \\ -\Delta A^T s \end{bmatrix}.$$

From the second block component we obtain

$$|y-x| \le \epsilon(|A^+|(f+E|y|) + |(A^T A)^{-1}|E^T|s|). \tag{4.5}$$

This yields the following theorem.

THEOREM 4.2. *Let $A \in \mathbb{R}^{m\times n}$ $(m \ge n)$ and $A+\Delta A$ be of full rank. For the perturbed LS problem described by (4.3) and (4.4) we have*

$$\frac{\|x-y\|}{\|x\|} \le \epsilon \frac{\| \, |A^{-1}|(f+E|y|) \, \| + \| \, |(A^T A)^{-1}|E^T|s| \, \|}{\|x\|}, \tag{4.6}$$

for any monotonic norm. □

For a square system, $s = 0$, and we essentially recover Theorem 2.4. Note, however, that the right-hand side of (4.6) contains the perturbed vectors y and s. For theoretical analysis it may be preferable to use an alternative bound in which x and r replace y and s and there is an extra factor

$$\left(1 - \epsilon \left\| \begin{bmatrix} |I - AA^+| & |(A^+)^T| \\ |A^+| & |(A^T A)^{-1}| \end{bmatrix} \begin{bmatrix} 0 & E \\ E^T & 0 \end{bmatrix} \right\| \right)^{-1},$$

where the term in parentheses is assumed to be positive. For practical computation (4.6) is unsatisfactory because we do not know $s = b + \Delta b - (A+\Delta A)y$. However, as Stewart and Sun observe [**71**, p. 159], $\widehat{r} = b - Ay$ is computable and

$$|s| \le |\widehat{r}| + \epsilon(f + E|y|),$$

and using this bound in (4.6) makes only a second-order change.

A componentwise bound of the form (4.6) was first derived by Björck in 1988 and variations of it have been given by Arioli, Duff and de Rijk [**3**], Björck [**11**] and Higham [**42**].

Apart from its increased sharpness over (4.1), the bound (4.6) has better scaling properties. It is not invariant under row or column scalings, but it is less sensitive to these scalings than (4.1). In [**42**] we examined the famous Longley test problem—a regression problem which has a notoriously ill-conditioned 16×7 coefficient matrix with $\kappa_2(A) \approx 5 \times 10^9$. We found that (4.5) gives tight bounds for the effect of random componentwise relative perturbations of the problem generated in experiments of Beaton, Rubin and Barone [**10**]. Thus componentwise perturbation bounds are potentially useful in regression analysis, as an alternative to the existing statistically-based techniques.

Although Wilkinson proved in the 1960s that a specific method for solving the LS problem, namely the QR factorization method, yields a small normwise backward error [**81**], it has long been an open problem to obtain a formula for the backward error of an arbitrary approximate solution. Stewart [**68**] discusses the problem and offers some backward perturbations that are candidates for being of minimal norm. Little progress had been made towards solving this problem until very recently when Waldén, Karlson and Sun [**77**] found an extremely elegant solution. We will denote by $\lambda_{\min}$ and $\sigma_{\min}$ the smallest eigenvalue of a symmetric matrix and the smallest singular value of a general matrix, respectively.

THEOREM 4.3 (Waldén, Karlson and Sun). *Let $A \in \mathbb{R}^{m\times n}$ $(m \ge n)$, $b \in \mathbb{R}^m$ and $r = b - Ay$. The normwise backward error*

$$\eta_F(y) = \min\{\, \|[\Delta A,\, \theta\Delta b]\|_F : \|(A+\Delta A)y - (b+\Delta b)\|_2 = \min \,\} \tag{4.7}$$

is given by

$$\eta_F(y) = \begin{cases} \dfrac{\|r\|_2}{\|y\|_2}\sqrt{\gamma}, & \lambda_* \ge 0, \\ \left(\dfrac{\|r\|_2^2}{\|y\|_2^2}\gamma + \lambda_*\right)^{1/2}, & \lambda_* < 0, \end{cases}$$

where

$$\lambda_* = \lambda_{\min}\Big(AA^T - \gamma\frac{rr^T}{\|y\|_2^2}\Big), \quad \gamma = \frac{\theta^2\|y\|_2^2}{1+\theta^2\|y\|_2^2}. \quad \square$$

The backward error (4.7) is not a direct generalization of the one in (2.1) because it minimizes $\|[\Delta A,\, \theta\Delta b]\|_F$ instead of $\max\{\|\Delta A\|_2/\|E\|_2, \|\Delta b\|_2/\|f\|_2\}$. However, the parameter θ allows us some flexibility: taking the limit $\theta \to \infty$ gives the case where only A is perturbed.

Theorem 4.3 can be interpreted as saying that if $\lambda_* \ge 0$ then the backward error is essentially that given by Theorem 2.1 for a consistent system. If $\lambda_* < 0$, however, the nearest perturbed system of which y is the LS solution is inconsistent. A sufficient condition for $\lambda_* < 0$ is $r \notin \mathrm{range}(A)$ (assuming $\gamma \ne 0$).

The formulae given in Theorem 4.3 are unsuitable for computation because they can suffer from catastrophic cancellation when $\lambda_* < 0$. Instead, the following alternative formula derived in [**77**] should be used:

$$\eta_F(y) = \min\left\{ \frac{\|r\|_2}{\|y\|_2}\sqrt{\gamma},\, \sigma_{\min}\left(\left[\, A, \quad \sqrt{\gamma}\frac{\|r\|_2}{\|y\|_2}\left(I - \frac{r_y {r_y}^T}{\|r\|_2^2}\right)\,\right]\right)\right\}.$$

To illustrate Theorem 4.3, we consider an LS problem with a 25×15 Vandermonde matrix $A = (p_i^{j-1})$, where the p_i are equally spaced on $[0,1]$, and a right-hand side b with elements equally spaced on $[0,1]$. The condition number $\kappa_2(A) = 1.47 \times 10^9$. We solved the LS problem in Matlab in two different ways:

by using the normal equations with Cholesky factorization, to give $\widehat{x}_{NE}$, and via Householder QR factorization, obtaining $\widehat{x}_{QR}$. We found that

$$\frac{\eta_F(\widehat{x}_{NE})|_{\theta=1}}{\|[A,b]\|_F} = 2.37 \times 10^{-10}, \qquad \frac{\eta_F(\widehat{x}_{NE})|_{\theta=\infty}}{\|A\|_F} = 2.56 \times 10^{-10},$$
$$\frac{\eta_F(\widehat{x}_{QR})|_{\theta=1}}{\|[A,b]\|_F} = 2.97 \times 10^{-17}, \qquad \frac{\eta_F(\widehat{x}_{QR})|_{\theta=\infty}}{\|A\|_F} = 3.22 \times 10^{-17}.$$

Thus, as would be expected from known backward error analysis, $\widehat{x}_{QR}$ is a backward stable solution but $\widehat{x}_{NE}$ is not. In this example it makes little difference whether or not we perturb b.

Componentwise backward error for the LS problem has been investigated by Arioli, Duff and de Rijk [**3**], Björck [**11**] and Higham [**42**]. The simplest approach is to apply the componentwise backward error $\omega_{E,f}(y)$ of (2.6) to the augmented system (4.2), setting

$$E = \begin{bmatrix} 0 & E_A \\ E_A^T & 0 \end{bmatrix}$$

so as not to perturb the diagonal blocks I and 0 of the augmented system coefficient matrix. However, this approach allows A and A^T to undergo different perturbations ΔA_1 and ΔA_2 with $\Delta A_1 \neq \Delta A_2^T$, so does not give a true backward error. This problem can be overcome by using the structured componentwise backward error of [**38**] to force symmetry of the perturbations; see [**38**] for details. One problem remains: as far as the backward error of y is concerned, the vector r in the augmented system is a vector of free parameters, so to obtain the true componentwise backward error we have to minimize the structure-preserving componentwise backward error over all r. This is a nonlinear optimization problem to which no closed-form solution is known. Experiments in [**38**] and [**44**] show that when y is a computed LS solution, $r = b - Ay$ is often a good approximation to the minimizing r.

4.1. Underdetermined systems. Perturbation theory is also available for the minimum 2-norm solution to an underdetermined system. Wedin's result, Theorem 4.1, is still valid when $m < n$ and x and y are interpreted as minimum 2-norm solutions of $Ax = b$ and $(A+\Delta A)y = b+\Delta b$. Of course, $r = 0$, so there is no $\kappa_2(A)^2$ term in the bound. A componentwise bound is given by Demmel and Higham [**22**].

THEOREM 4.4 (Demmel and Higham). *Let $A \in \mathbb{R}^{m\times n}$ $(m \leq n)$ be of full rank and $0 \neq b \in \mathbb{R}^m$. Suppose*

$$|\Delta A| \leq \epsilon E, \quad |\Delta b| \leq \epsilon f,$$

where $\epsilon\|E\|_2 < \sigma_{\min}(A)$. If x and y are the minimum norm solutions to $Ax = b$ and $(A + \Delta A)y = b + \Delta b$, respectively, then

$$\frac{\|x-y\|_2}{\|x\|_2} \leq \left(\| \, |I - A^+A| \cdot E^T \cdot |A^{+T}x| \, \|_2 + \| \, |A^+| \cdot (f + E|x|) \, \|_2 \right) \frac{\epsilon}{\|x\|_2} + O(\epsilon^2).$$

The bound is attainable to within a constant factor depending on n. □

Note that $\|I - A^+A\|_2 = \min(1, n-m)$, so again there is no $\kappa_2(A)^2$ effect in the bound. This theorem is used in [**22**], with different choices of E and f, to investigate the stability of standard methods for computing the minimal 2-norm solution to an underdetermined system.

5. Matrix Factorizations

Perturbation theory for matrix factorizations has been an active area of research in recent years. The aim of such analysis is to determine bounds for the changes in the factors of a matrix when the matrix is perturbed. Most of the existing results are norm-based. One of the first results was that of Stewart [**67**]. He showed that if $A \in \mathbb{R}^{m\times n}$ has rank n and

$$A = QR \quad \text{and} \quad A + \Delta A = (Q + \Delta Q)(R + \Delta R)$$

are QR factorizations, then, for sufficiently small ΔA,

$$\begin{aligned} \frac{\|\Delta R\|_F}{\|R\|_F} &\leq c_n \kappa_F(A) \frac{\|\Delta A\|_F}{\|A\|_F}, \\ \|\Delta Q\|_F &\leq c_n \kappa_F(A) \frac{\|\Delta A\|_F}{\|A\|_F}, \end{aligned}$$

where c_n is a modest constant.

Similar analyses for Cholesky, $\mathrm{LDL}^{\mathrm{T}}$, LU and QR factorizations are given by Barrlund [**5**], Stewart [**70**], and Sun [**73**].

Componentwise analyses have been given by Zha [**82**] for the QR factorization and Sun [**74**] for the Cholesky, $\mathrm{LDL}^{\mathrm{T}}$, LU and QR factorizations. Sun's results are rather complicated to state and interpret. Zha's bounds can be summarized as follows, with the same assumptions and notation as for Stewart's result above. Let $|\Delta A| \leq \epsilon G|A|$, where G is nonnegative with $|g_{ij}| \leq 1$. Then, for sufficiently small ϵ,

$$\begin{aligned} \frac{\|\Delta R\|_\infty}{\|R\|_\infty} &\leq c_{m,n}\epsilon\big(\mathrm{cond}(R^{-1}) + \mathrm{cond}(R^T)\big) + O(\epsilon^2), \\ \|\Delta Q\|_\infty &\leq c_{m,n}\epsilon\big(\mathrm{cond}(R^{-1}) + \mathrm{cond}(R^T)\big) + O(\epsilon^2), \end{aligned}$$

where cond is defined in (2.11). The quantity $\phi(A) = \mathrm{cond}(R^{-1}) + \mathrm{cond}(R^T)$ can therefore be thought of as a condition number for the QR factorization under the columnwise class of perturbations considered. Note that ϕ is independent of the column scaling of A. Zha [**82**] shows how this perturbation result can be used together with a componentwise rounding error analysis to bound the difference $Q - \widehat{Q}$ between the exact QR factor Q and the computed one from Householder QR factorization.

6. Eigenvalue and Singular Value Problems

Perturbation theory for eigenvalue problems is inherently more difficult than for linear equations because the mapping from data to solution is no longer a rational map. Normwise perturbation theory for eigenproblems is well developed (see, for example, Stewart and Sun's comprehensive presentation [**71**]) but componentwise results have been obtained only in the last five years or so.

Existing results are not easy to present in an organized fashion because they cover different problems (standard or generalized, symmetric positive definite or indefinite), make different assumptions, and measure perturbations in different ways. We give a selected summary, concentrating on the standard eigenvalue and singular value problems and noting which results have counterparts for the generalized problem. We do not necessarily state results in their most general form, or use their authors' original notation. In particular, the roles of A and H are reversed between

our presentation and those of most of the references, because it seems natural to denote the original matrix, whose eigenvalues or singular values we seek, by A.

Most of the eigenvalue results are for symmetric matrices, but we begin with the nonsymmetric case. It is a standard result that if $A \in \mathbb{R}^{n\times n}$ has a simple eigenvalue λ with corresponding right and left eigenvectors x and y, then for sufficiently small $\|\Delta A\|$ there is an eigenvalue μ of $A + \Delta A$ with

$$\mu = \lambda + \frac{y^*\Delta A x}{y^* x} + O(\|\Delta A\|^2). \tag{6.1}$$

(See, for example, [**71**, Theorem 4.2.3] or [**81**, Chapter 2]). If we assume that $\|\Delta A\| \le \epsilon\|E\|$, then we have the perturbation bound

$$|\mu - \lambda| \le \epsilon\|E\|\frac{\|y\|_D\|x\|}{|y^* x|} + O(\epsilon^2), \tag{6.2}$$

which is sharp to first order (equality is attained for ΔA an appropriate multiple of uv^T, where u is dual to y and v is dual to x). On the other hand, if we assume that $|\Delta A| \le \epsilon E$, we obtain the componentwise perturbation bound

$$|\mu - \lambda| \le \epsilon\frac{|y^*|E|x|}{|y^* x|} + O(\epsilon^2) \tag{6.3}$$

of Geurts [**31**], for which equality to first order is attained for

$$\Delta A = \epsilon\,\mathrm{diag}(\mathrm{sign}(y))E\,\mathrm{diag}(\mathrm{sign}(x)),$$

where, for possibly complex α, $\mathrm{sign}(\alpha) = \overline{\alpha}/|\alpha|$ (or 1 if $\alpha = 0$). The first-order terms in the bounds (6.2) and (6.3) yield condition numbers under the respective classes of perturbations. One interesting implication of (6.3) is that the Perron root of an irreducible nonnegative matrix is perfectly conditioned under componentwise relative perturbations ($|\Delta A| \le \epsilon E = \epsilon A$), for the eigenvectors corresponding to the Perron root have positive elements and so the bound is $|\mu-\lambda| \le \epsilon\lambda+O(\epsilon^2)$. In fact, this bound holds for any $\epsilon \le 1$ without the second-order term [**27**, Theorem 1].

Corresponding eigenvector perturbation results can also be obtained. We state condition numbers in the form given by Chatelin [**15**] and Frayssé [**30**], whose analysis extends earlier analysis by Wilkinson [**81**], Stewart [**66**] and Geurts [**31**]; see also the book by Saad [**62**, pp. 95–97]. With the same notation as above (with λ simple), a condition number of x under perturbations measured normwise by $\|\Delta A\| \le \epsilon\|E\|$ is

$$\|E\|\|\mathcal{P}\|,$$

where $\mathcal{P} = X(B - \lambda I)^{-1}Y^*$ is a partial inverse for Y in the space $\mathrm{range}(y)^\perp$, and where

$$Z^{-1}AZ = \begin{bmatrix} \lambda & c^T \\ 0 & B \end{bmatrix}, \quad Z = [\,x \quad X\,], \quad Z^{-1} = \begin{bmatrix} y^* \\ Y^* \end{bmatrix}.$$

Note that $\|\mathcal{P}\|$ depends on the particular choice of X and Y. This freedom stems from the nonuniqueness of an eigenvector up to scalar multiplication, which means that there is not a unique way to measure the perturbation in an eigenvector. What measure is appropriate depends on the application (see [**71**, p. 241] for a discussion and example). The minimum condition number in the 2-norm is obtained when X and Y are orthonormal and is

$$\|\mathcal{P}\|_2 = \|(B - \lambda I)^{-1}\|_2 = \mathrm{sep}(B, \lambda)^{-1}.$$

A condition number for perturbations measured componentwise by $|\Delta A| \le \epsilon E$, with the ∞-norm on the output space, is

$$\frac{\| \, |\mathcal{P}|E|x| \, \|_\infty}{\|x\|_\infty}.$$

In the special case where A is diagonalizable we can take B to be diagonal and X and Y to comprise right and left eigenvectors. Then it is clear that the size of $\mathcal{P}$ depends on both the separation of λ from the other eigenvalues and on the individual eigenvalue sensitivities.

Another classical eigenvalue perturbation result is a theorem of Bauer and Fike [**9**, Theorem 3a], which says that if μ is an eigenvalue of $A + \Delta A \in \mathbb{R}^{n\times n}$ and $X^{-1}AX = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ then

$$\min_{\lambda\in\lambda(A)} |\lambda - \mu| \le \kappa_p(X)\|\Delta A\|_p,$$

for any Hölder p-norm. As Deif [**17**] shows, it is trivial to strengthen this result to

$$\min_{\lambda\in\lambda(A)} |\lambda - \mu| \le \| \, |X^{-1}||\Delta A||X| \, \|,$$

for any monotonic norm.

In the rest of this section we describe the new-style componentwise perturbation results, developed initially by Demmel and his co-workers. For the eigenvalue problem, most of these results require that A be symmetric positive definite. We therefore let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and consider a symmetric perturbation $A+\Delta A$. Let the eigenvalues of A and $A+\Delta A$ be ordered $\lambda_n \le \cdots \le \lambda_1$ and $\lambda'_n \le \cdots \le \lambda'_1$, respectively. From classical perturbation results for the symmetric eigenproblem we know that

$$|\lambda_i - \lambda'_i| \le \|\Delta A\|_2$$

(specifically, this follows from Weyl's inequality [**71**, p. 203], which itself is a consequence of the Courant–Fischer minimax theorem [**71**, p. 201]). For $\lambda_1 = \|A\|_2$ this bound says that $|\lambda_1 - \lambda'_1|/\lambda_1 \le \|\Delta A\|_2/\|A\|_2$, so this largest eigenvalue undergoes a relative change no larger than that in A. But eigenvalues with $\lambda_i \ll \lambda_1$ can undergo a large relative change, and the best relative perturbation bound that can be deduced for all i is

$$\frac{|\lambda_i - \lambda'_i|}{\lambda_i} \le \frac{\|\Delta A\|_2}{\lambda_i} \le \frac{\|\Delta A\|_2}{\lambda_n} = \kappa_2(A)\frac{\|\Delta A\|_2}{\|A\|_2}. \tag{6.4}$$

This bound is sharp, in that for any A there is a perturbation ΔA for which equality is attained with $i = n$. However, by restricting the class of perturbations, tighter bounds can be obtained. Demmel and Veselić [**24**, Theorems 2, 3] derive a potentially much smaller bound for the case of componentwise relative perturbations.

THEOREM 6.1 (Demmel and Veselić). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and write $A = DHD$, where $D = \mathrm{diag}(A)^{1/2}$. Let the symmetric perturbation $\Delta A = D\Delta HD$ satisfy $\|\Delta H\|_2 \equiv \epsilon < \lambda_n(H)$. Then*

$$\frac{|\lambda_i - \lambda'_i|}{\lambda_i} \le \kappa_2(H)\epsilon. \tag{6.5}$$

In particular, if $|\Delta A| \le (\epsilon/n)|A|$, so that $|\Delta h_{ij}| \le \epsilon/n$, then $\|\Delta H\|_2 \le \epsilon$ and the bound (6.5) holds if $\epsilon < \lambda_n(H)$. □

The key idea in this theorem is to factor out the diagonal of A to leave a symmetric positive definite matrix H with unit diagonal. A famous result of van der Sluis [**75**, Theorem 4.1] states that

$$\kappa_2(H) \le n \min_{F \text{ diagonal}} \kappa_2(FAF),$$

so D is nearly a condition-minimizing diagonal scaling. It follows that $\kappa_2(H) \le n\kappa_2(A)$ and that $\kappa_2(H) \ll \kappa_2(A)$ is possible if A is badly scaled.

The proof of Theorem 6.1 in [**24**] is quite short and is based on the following lemma.

LEMMA 6.2. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and write $A = YHY$, where Y is any nonsingular matrix. Let the symmetric perturbation $\Delta A = Y\Delta HY$ satisfy $|z^T\Delta Hz| \le \eta\, z^THz$ for all z. Then*

$$1 - \eta \le \frac{\lambda_i'}{\lambda_i} \le 1 + \eta.$$

PROOF. Note first that

$$\frac{|x^T\Delta Ax|}{x^TAx} = \frac{|x^TY\Delta HYx|}{x^TYHYx} \equiv \frac{|z^T\Delta Hz|}{z^THz} \le \eta.$$

From the Courant–Fischer minimax theorem we have

$$\begin{aligned}
\lambda_i' &= \max_{\dim(S)=i} \min_{x\in S} \frac{x^T(A+\Delta A)x}{x^Tx} \\
&= \max_{\dim(S)=i} \min_{x\in S} \frac{x^T(A+\Delta A)x}{x^TAx} \cdot \frac{x^TAx}{x^Tx} \\
&\le (1+\eta) \max_{\dim(S)=i} \min_{x\in S} \frac{x^TAx}{x^Tx} = (1+\eta)\lambda_i.
\end{aligned}$$

The lower bound is proved similarly, starting with the max–min expression for λ_i. □

PROOF OF THEOREM 6.1. We have

$$\frac{|z^T\Delta Hz|}{z^THz} = \frac{|z^T\Delta Hz|}{z^Tz}\frac{z^Tz}{z^THz} \le \frac{\epsilon}{\lambda_n(H)}.$$

We obtain (6.5) on applying Lemma 6.2 with $Y = D$ and $\eta = \epsilon/\lambda_n(H) = \epsilon\|H^{-1}\|_2$, and using the inequality $\|H\|_2 \ge \max_{i,j}|h_{ij}| = 1$. □

A weakness of Theorem 6.1 is that it gives the same relative perturbation bound for each eigenvalue, yet some eigenvalues may be much less sensitive than others. Furthermore, λ_1 is insensitive to perturbations, yet the bound (6.5) can be large even for $i = 1$. Demmel and Veselić note that, under the same conditions as in Theorem 6.1, one can show, using (6.1), that

$$\frac{|\lambda_i - \lambda_i'|}{\lambda_i} \le \epsilon\frac{\|Dv_i\|_2^2}{\lambda_i} + O(\epsilon^2),$$

where $Av_i = \lambda_i v_i$ with $\|v_i\|_2 = 1$, and that this bound is attained, to first order, for $\Delta A = \epsilon D^2$. This bound depends on i, and can be much smaller than (6.5), but it is a first-order bound only.

Demmel and Veselić show, via rounding error analysis, that Jacobi's method (with a suitable stopping criterion), the bisection method, and inverse iteration (the

latter two methods both applied to the original matrix) all compute the eigenvalues of a symmetric positive definite matrix to within the accuracy specified by the bound of Theorem 6.1 (with ϵ the machine precision). By contrast, any method that begins by tridiagonalizing the matrix (for example, the QR algorithm) cannot achieve this level of accuracy.

We turn now to eigenvectors. Standard perturbation theory (for which a classic reference is [**16**]) is expressed in terms of the absolute gap for eigenvalues, defined by

$$\operatorname{absgap}_i = \min_{j\neq i} \frac{|\lambda_i - \lambda_j|}{\|A\|_2}.$$

Let y be a unit eigenvector of $A + \Delta A$ with Rayleigh quotient $\alpha = y^T A y$. Let λ_i be the eigenvalue of A closest to α and x_i the corresponding unit eigenvector. Denoting by $\theta(x_i, y)$ the acute angle between x_i and y, we have (by rewriting the bound of [**56**, Theorem 11-7-1])

$$|\sin\theta(x_i, y)| \leq \frac{4}{\operatorname{absgap}_i} \frac{\|\Delta A\|_2}{\|A\|_2}. \tag{6.6}$$

Demmel and Veselić [**24**, Theorem 2.5] derive a potentially much smaller bound.

THEOREM 6.3 (Demmel and Veselić). *Let $A = DHD$ be as in Theorem 6.1, and define $A(\epsilon) = D(H + \epsilon E)D$, where E with $\|E\|_2 = 1$ is an arbitrary symmetric matrix. Let $\lambda_i(\epsilon)$ be the ith eigenvalue of $A(\epsilon)$ and assume that $\lambda_i(0)$ is simple, so that the corresponding unit eigenvector $v_i(\epsilon)$ is well defined for sufficiently small ϵ. Then*

$$\|v_i(\epsilon) - v_i(0)\|_2 \leq \frac{(n-1)^{1/2}\kappa_2(H)\epsilon}{\operatorname{relgap}_i} + O(\epsilon^2), \tag{6.7}$$

where the relative gap is defined by

$$\operatorname{relgap}_i = \min_{j\neq i} \frac{|\lambda_i - \lambda_j|}{|\lambda_i \lambda_j|^{1/2}}. \quad \square$$

The new bound (6.7) contains an extra factor $\kappa_2(H)$ compared with (6.6), but the important difference is that the relative gap in the denominator of (6.7) is much larger than the absolute gap in (6.6) for a small eigenvalue that is in a cluster of eigenvalues. For example, if $\lambda_1 = 1$, $\lambda_2 = 2 \times 10^{-10}$ and $\lambda_3 = 10^{-10}$, then

$$\begin{aligned} \operatorname{absgap}(1\colon 3) &= 1 - 2\times 10^{-10},\ 10^{-10},\ 10^{-10}, \\ \operatorname{relgap}(1\colon 3) &= 2^{-1/2}(10^5 - 2\times 10^{-5}),\ 2^{-1/2},\ 2^{-1/2}. \end{aligned}$$

How sharp are these bounds? For any A, approximate equality is attained for some i in the bound (6.5) for a perturbation ΔA satisfying the conditions of Theorem 6.1 [**24**, Prop. 2.10]. This implies that the only symmetric positive definite matrices whose eigenvalues are determined to high relative accuracy by the matrix elements are those $A = DHD$ where H is well conditioned. A weaker sharpness result holds for (6.7) [**24**, Prop. 2.11].

We mention another interesting result that can be proved by Rayleigh quotient manipulations similar to those in the proof of Lemma 6.2. This result adds further information to the straightforward inequalities $\lambda_n \leq a_{ii} \leq \lambda_1$, by showing that the eigenvalues of A differ from the diagonal elements by at most a factor $\kappa_2(H)$.

THEOREM 6.4 (Demmel and Veselić [**24**, Prop. 2.7]). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric positive definite and write $A = DHD$, where $D = \operatorname{diag}(A)^{1/2}$. Let $\alpha_{nn} \le \cdots \le \alpha_{11}$ be the diagonal entries of A arranged in decreasing order. Then*

$$\lambda_n(H) \le \frac{\lambda_i}{\alpha_{ii}} \le \lambda_1(H) \le n. \quad \square$$

Analogous results to Theorem 6.1 and 6.3 hold for singular values and singular vectors. Let $A \in \mathbb{R}^{m\times n}$, where there is no restriction on m and n. We denote the singular values of A and $A + \Delta A$ by $\sigma_{\min} \le \cdots \le \sigma_1$ and $\sigma'_{\min} \le \cdots \le \sigma'_1$, respectively.

THEOREM 6.5 (Demmel and Veselić [**24**, Theorem 2.14]). *Let $A \in \mathbb{R}^{m\times n}$ have full rank and write $A = BD$, where D is diagonal and the columns of B have unit 2-norm. Let $\Delta A = \Delta B D$ satisfy $\|\Delta B\|_2 \equiv \epsilon < \sigma_{\min}(B)$. Then*

$$\frac{|\sigma_i - \sigma'_i|}{\sigma_i} \le \kappa_2(B)\epsilon. \tag{6.8}$$

In particular, if $|\Delta A| \le (\epsilon/\sqrt{n})|A|$ then $\|\Delta B\|_2 \le \epsilon$ and the bound (6.8) holds if $\epsilon < \sigma_{\min}(B)$. $\square$

We will not state the corresponding theorem giving a perturbation bound for the singular vectors; it involves a relative gap defined differently than for the eigenvalue case as $\operatorname{relgap}_i = \min_{j\ne i} |\sigma_i - \sigma_j|/(\sigma_i + \sigma_j)$ (see [**18**], [**24**]). As for Theorem 6.1, the bound in Theorem 6.5 is attainable, though not necessarily for a componentwise perturbation of B (that is, approximate equality is attained for a perturbation that satisfies $\|\Delta B\|_2 = \epsilon < \sigma_{\min}(B)$ but does not necessarily satisfy $|\Delta A| \le (\epsilon/\sqrt{n})|A|$) [**24**, Prop. 2.21].

The earliest componentwise eigenvalue perturbation results were obtained by Kahan in 1966 [**50**]. His results were slightly strengthened in [**4**] and [**23**].

THEOREM 6.6 (Kahan). *Let $A \in \mathbb{R}^{n\times n}$ be a symmetric tridiagonal matrix with zero diagonal elements, and let the symmetric perturbation ΔA satisfy $|\Delta A| \le \epsilon|A|$. Then*

$$\frac{\lambda_i}{\tau^{2n-1}} \le \lambda'_i \le \tau^{2n-1}\lambda_i,$$

where $\tau = \max(1+\epsilon, 1/(1-\epsilon))$. $\square$

Thus relative perturbations of size ϵ in the elements of A change the eigenvalues by relative amounts at most $(2n-1)\epsilon + O(\epsilon^2)$.

COROLLARY 6.7. *Let $B \in \mathbb{R}^{n\times n}$ be bidiagonal and let $|\Delta B| \le \epsilon|B|$. Then*

$$\frac{\sigma_i}{\tau^{2n-1}} \le \sigma'_i \le \tau^{2n-1}\sigma_i,$$

where $\tau = \max(1+\epsilon, 1/(1-\epsilon))$. $\square$

PROOF. Use the facts that the eigenvalues of $C = \left[\begin{smallmatrix} 0 & B \\ B^T & 0 \end{smallmatrix}\right]$ are plus and minus the singular values of B, and that C can be permuted to a symmetric tridiagonal matrix with zero diagonal. $\square$

Corollary 6.7 states that the singular values of a bidiagonal matrix are determined to approximately the same relative accuracy as the elements of the matrix. Demmel and Kahan [**23**] use this result, together with rounding error analysis, to show that a zero-shift version of the QR algorithm computes the singular values

of a bidiagonal matrix to high relative accuracy. Fernando and Parlett [**28**] have developed a shifted quotient-difference algorithm that satisfies even smaller error bounds than the algorithm of Demmel and Kahan, and is at least as fast.

Barlow and Demmel [**4**] were interested in finding classes of matrices other than those in Theorem 6.6 or Corollary 6.7 that determine their eigenvalues or singular values to high relative accuracy. Their main result shows that such a class is the class of "scaled diagonally dominant" symmetric positive definite matrices A: those for which $\|H - I\|_2 = \gamma < 1$, where $A = DHD$ with $D = \mathrm{diag}(A)^{1/2}$. For this class the relative eigenvalue perturbations are bounded by a multiple of $(1-\gamma)^{-1}$. Barlow and Demmel also obtained an analogous result for the generalized symmetric positive definite eigenproblem. As Demmel and Veselić later showed, the assumption on diagonal dominance is unnecessary: Theorem 6.1 shows that it is only necessary that $\kappa_2(H)$ is not large for the eigenvalues to be accurately determined. Indeed, for a scaled diagonally dominant matrix, $\kappa_2(H) < (1+\gamma)/(1-\gamma)$. Barlow and Demmel also gave perturbation results for the eigenvectors of scaled diagonally dominant matrices; these revealed, for the first time, the role of the eigenvalue relative gap in measuring eigenvector sensitivity.

Demmel and Gragg [**21**] consider the question "for which sparsity patterns do small componentwise relative perturbations to the matrix cause small relative perturbations to the singular values?" From Corollary 6.7 we know that one allowable sparsity pattern is that of a bidiagonal matrix. Demmel and Gragg give a very elegant answer to the question: a sparsity pattern has this property if and only if its associated bipartite graph is acyclic. Examples of such acyclic sparsity patterns are, in addition to the bidiagonal pattern,

$$\begin{bmatrix} \times & & & & & \times \\ & \times & & & & \times \\ & & \times & & & \times \\ & & & \times & & \times \\ & & & & \times & \times \\ & & & & & \times \end{bmatrix}, \qquad \begin{bmatrix} \times & & & \times & & \\ & \times & & & \times & \\ & & \times & & & \times \\ & & & \times & & \times \\ & & & & \times & \times \\ & & & & & \times \end{bmatrix}.$$

Eisenstat and Ipsen [**26**] answer an open question raised in [**21**] by obtaining, for the singular vectors of a matrix with an acyclic sparsity pattern, a perturbation bound that involves a relative gap for the singular values. They also use a theorem of Ostrowski to give new, elegant proofs of Theorem 6.6 and of Demmel and Gragg's componentwise perturbation bound for the singular values of a matrix with an acyclic sparsity pattern.

The results of Demmel and Veselić have been generalized to indefinite symmetric matrices by two pairs of authors. Veselić and Slapničar make use of the polar decomposition $A = US \in \mathbb{R}^{n\times n}$, where U is orthogonal and $S = (A^TA)^{1/2}$ is symmetric positive semidefinite.

THEOREM 6.8 (Veselić and Slapničar [**76**, Theorem 2.13]). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric, with the polar decomposition $A = US$, and write $A = DHD$, where $D = \mathrm{diag}(S)^{1/2}$. Let the symmetric perturbation ΔA satisfy $|\Delta A| \le \epsilon|A|$ and assume that $\nu(A)\epsilon < 1$, where $\nu(A) = \| \, |H| \, \|_2 \|DS^{-1}D\|_2$. Then*

$$|\lambda_i - \lambda_i'| \le \nu(A)\epsilon|\lambda_i|. \quad \square$$

If A is positive definite then $S = A$, so $\kappa_2(H) \le \nu(A) \le \sqrt{n}\kappa_2(H)$, and Theorem 6.8 essentially reduces to Theorem 6.1. Veselić and Slapničar also prove

an eigenvector perturbation result and a generalization of Theorem 6.8 for the symmetric generalized eigenproblem.

Gu and Eisenstat [**35**] give results for indefinite matrices that have a form very different from those of Veselić and Slapničar. To state a sample result we need the notation that $\mu_0(A) = |\lambda|$, where λ is the eigenvalue of smallest absolute value among all eigenvalues of principle submatrices of A.

THEOREM 6.9 (Gu and Eisenstat [**35**, Corollary 5]). *Let $A \in \mathbb{R}^{n\times n}$ be symmetric with $\mu_0(H) > 0$, where $A = DHD$ with D a positive definite diagonal matrix. Let the symmetric perturbation $\Delta A = D\Delta HD$ satisfy $\|\Delta H\|_2 \equiv \epsilon < \mu_0(H)$. Then*

$$|\lambda_i - \lambda_i'| \le |\lambda_i|\epsilon\frac{(2\mu_0(H) - \epsilon)}{(\mu_0(H) - \epsilon)^2}. \quad \square$$

If A is positive definite then $\mu_0(H) = \lambda_n(H)$, and if we take $D = \mathrm{diag}(A)^{1/2}$ the bound is

$$|\lambda_i - \lambda_i'| \le 2\epsilon\lambda_i/\lambda_n(H) + O(\epsilon^2) \le 2\epsilon\lambda_i\kappa_2(H) + O(\epsilon^2),$$

which is asymptotically the same as the bound (6.5) in Theorem 6.1, to within a factor 2. Gu and Eisenstat also give an eigenvector perturbation bound for indefinite A involving the relative gap, and a singular value perturbation result that allows two-sided diagonal scalings.

The componentwise bounds described in this section are reflected in several LAPACK routines. For example, the LAPACK routines `xSTEBZ` and `xPTEQR` for finding the eigenvalues of a symmetric (positive definite) tridiagonal matrix achieve the accuracy shown possible by Theorem 6.1. And the singular values of a bidiagonal matrix are computed to high relative accuracy by `xBDSQR` (cf. Corollary 6.7).

We conclude with a numerical example involving the "ipjfact" matrix from the Test Matrix Toolbox [**45**], for which $a_{ij} = (i+j)!$. This matrix is symmetric positive definite but badly scaled and ill conditioned. For $n = 10$, $\kappa_2(A) = 3.00\times 10^{21}$, while $\kappa_2(H) = 6.04 \times 10^8$ and

$$\mathrm{absgap}_n = 1.33 \times 10^{-20}, \quad \mathrm{relgap}_n = 6.24.$$

Therefore, for perturbations satisfying $|\Delta A| \le \epsilon|A|$, the new bound (6.5) is smaller than the traditional bound (6.4) by a factor of order 10^{12}, and is able to guarantee some unchanged digits in all the eigenvalues if $\epsilon < 10^{-10}$. Similarly, the new eigenvector perturbation bound (6.7) is of the order 10^{12} times smaller than traditional bound (6.6). Finally, the sorted diagonal elements α_{ii} of A and the eigenvalues λ_i of A are shown in Table 1. The maximum value of α_{ii}/λ_i is 3.59×10^3 (occurring for $i = 5$), which is less than $1.30 \times 10^8 = \lambda_n(H)^{-1}$, illustrating Theorem 6.4.

7. Concluding Remarks

Componentwise perturbation results are now available for most of the standard problems in numerical linear algebra. The forward perturbation results vary in their form and underlying assumptions, but most have the following property: when the problem data is badly scaled and the perturbations are measured in a componentwise relative fashion, the bounds can be much smaller than the traditional normwise bounds.

There are several reasons why componentwise bounds do not feature strongly in the literature of the 1960s and 1970s. In the first place, many (dense) problems

α_{ii}	λ_i	α_{ii}/λ_i
2.00e+00	8.12e-04	2.46e+03
2.40e+01	3.32e-02	7.22e+02
7.20e+02	5.72e-01	1.26e+03
4.03e+04	1.28e+01	3.15e+03
3.63e+06	1.01e+03	3.59e+03
4.79e+08	2.21e+05	2.17e+03
8.72e+10	1.12e+08	7.79e+02
2.09e+13	1.25e+11	1.67e+02
6.40e+15	3.23e+14	1.98e+01
2.43e+18	2.44e+18	9.97e−01

TABLE 1. Sorted diagonal elements versus eigenvalues.

that arise in practice are well scaled, and for these, normwise bounds are adequate. Second, badly scaled problems can be equilibrated prior to being solved, and various equilibration algorithms have been developed over the years. However, how best to scale a matrix is a tricky question for which there is no simple answer (see [**68**] for a discussion). A third reason is that normwise bounds tend to be "simpler" than componentwise ones, with the dependence on the matrix A often appearing only through $\kappa(A)$. A fourth reason is that the most impressive applications of componentwise analysis are in proving results that were not contemplated until relatively recently. Examples of such results are that to improve the stability of a solution to a linear system by iterative refinement it suffices to compute the residual in the working precision [**64**], and that the singular values of a bidiagonal matrix can be computed to the precision to which they are determined by the matrix entries [**23**].

Acknowledgements

I thank Des Higham for suggesting perturbations to the manuscript.

References

1. E. Anderson, Z. Bai, C. H. Bischof, J. W. Demmel, J. J. Dongarra, J. J. Du Croz, A. Greenbaum, S. J. Hammarling, A. McKenney, S. Ostrouchov, and D. C. Sorensen, *LAPACK users' guide*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1992.
2. M. Arioli, J. W. Demmel, and I. S. Duff, *Solving sparse linear systems with sparse backward error*, SIAM J. Matrix Anal. Appl. **10** (1989), no. 2, 165–190.
3. M. Arioli, I. S. Duff, and P. P. M. de Rijk, *On the augmented system approach to sparse least-squares problems*, Numer. Math. **55** (1989), 667–684.
4. Jesse L. Barlow and James W. Demmel, *Computing accurate eigensytems of scaled diagonally dominant matrices*, SIAM J. Numer. Anal. **27** (1990), no. 3, 762–791.
5. Anders Barrlund, *Perturbation bounds for the LDL^H and LU decompositions*, BIT **31** (1991), 358–363.
6. Sven G. Bartels and Desmond J. Higham, *The structured sensitivity of Vandermonde-like systems*, Numer. Math. **62** (1992), 17–33.
7. F. L. Bauer, *Optimally scaled matrices*, Numer. Math. **5** (1963), 73–87.
8. ———, *Genauigkeitsfragen bei der Lösung linearer Gleichungssysteme*, Z. Angew. Math. Mech. **46** (1966), no. 7, 409–421.
9. F. L. Bauer and C. T. Fike, *Norms and exclusion theorems*, Numer. Math. **2** (1960), 137–141.

10. Albert E. Beaton, Donald B. Rubin, and John L. Barone, *The acceptability of regression solutions: Another look at computational accuracy*, J. Amer. Stat. Assoc. **71** (1976), no. 353, 158–168.
11. Åke Björck, *Component-wise perturbation analysis and error bounds for linear least squares solutions*, BIT **31** (1991), 238–244.
12. James R. Bunch, James W. Demmel, and Charles F. Van Loan, *The strong stability of algorithms for solving symmetric linear systems*, SIAM J. Matrix Anal. Appl. **10** (1989), no. 4, 494–499.
13. P. A. Businger, *Matrices which can be optimally scaled*, Numer. Math. **12** (1968), 346–348.
14. Shivkumar Chandrasekaran and Ilse C. F. Ipsen, *On rank-revealing factorisations*, SIAM J. Matrix Anal. Appl. **15** (1994), no. 2, 592–622.
15. Françoise Chatelin, *Eigenvalues of matrices*, Wiley, Chichester, UK, 1993.
16. Chandler Davis and W. M. Kahan, *The rotation of eigenvectors by a perturbation. III*, SIAM J. Numer. Anal. **7** (1970), no. 1, 1–46.
17. A. S. Deif, *Realistic a priori and a posteriori error bounds for computed eigenvalues*, IMA J. Numer. Anal. **9** (1990), 323–329.
18. Percy Deift, James W. Demmel, Luen-Chau Li, and Carlos Tomei, *The bidiagonal singular value decomposition and Hamiltonian mechanics*, SIAM J. Numer. Anal. **28** (1991), 1463–1516.
19. James W. Demmel, *On floating point errors in Cholesky*, Technical Report CS-89-87, Department of Computer Science, University of Tennessee, Knoxville, TN, USA, October 1989, LAPACK Working Note 14.
20. ———, *The componentwise distance to the nearest singular matrix*, SIAM J. Matrix Anal. Appl. **13** (1992), no. 1, 10–19.
21. James W. Demmel and William Gragg, *On computing accurate singular values and eigenvalues of matrices with acyclic graphs*, Linear Algebra and Appl. **185** (1993), 203–217.
22. James W. Demmel and Nicholas J. Higham, *Improved error bounds for underdetermined system solvers*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 1, 1–14.
23. James W. Demmel and W. Kahan, *Accurate singular values of bidiagonal matrices*, SIAM J. Sci. Stat. Comput. **11** (1990), no. 5, 873–912.
24. James W. Demmel and Krešimir Veselić, *Jacobi's method is more accurate than QR*, SIAM J. Matrix Anal. Appl. **13** (1992), no. 4, 1204–1245.
25. Jeremy J. Du Croz and Nicholas J. Higham, *Stability of methods for matrix inversion*, IMA J. Numer. Anal. **12** (1992), 1–19.
26. Stanley C. Eisenstat and Ilse C. F. Ipsen, *Relative perturbation techniques for singular value problems*, SIAM J. Numer. Anal. **32** (1995), no. 6, 1972–1988.
27. L. Elsner, I. Koltracht, M. Neumann, and D. Xiao, *On accurate computations of the Perron root*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 2, 456–467.
28. K. Vince Fernando and Beresford N. Parlett, *Accurate singular values and differential qd algorithms*, Numer. Math. **67** (1994), 191–229.
29. R. Fletcher, *Expected conditioning*, IMA J. Numer. Anal. **5** (1985), 247–273.
30. V. Frayssé, *Reliability of computer solutions*, Ph.D. thesis, L'Institut National Polytechnique de Toulouse, Toulouse, France, July 1992, CERFACS Report TH/PA/92/11.
31. A. J. Geurts, *A contribution to the theory of condition*, Numer. Math. **39** (1982), 85–96.
32. I. Gohberg and I. Koltracht, *Mixed, componentwise, and structured condition numbers*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 3, 688–704.
33. G. H. Golub and J. H. Wilkinson, *Note on the iterative refinement of least squares solution*, Numer. Math. **9** (1966), 139–148.
34. Gene H. Golub and Charles F. Van Loan, *Matrix computations*, second ed., Johns Hopkins University Press, Baltimore, MD, USA, 1989.
35. Ming Gu and Stanley C. Eisenstat, *Relative perturbation theory for eigenproblems*, Research Report YALEU/DCS/RR-934, Department of Computer Science, Yale University, February 1993.
36. William W. Hager, *Condition estimates*, SIAM J. Sci. Stat. Comput. **5** (1984), no. 2, 311–316.
37. Desmond J. Higham, *Condition numbers and their condition numbers*, Linear Algebra and Appl. **214** (1995), 193–213.
38. Desmond J. Higham and Nicholas J. Higham, *Backward error and condition of structured linear systems*, SIAM J. Matrix Anal. Appl. **13** (1992), no. 1, 162–175.

39. ———, *Componentwise perturbation theory for linear systems with multiple right-hand sides*, Linear Algebra and Appl. **174** (1992), 111–129.
40. Nicholas J. Higham, *FORTRAN codes for estimating the one-norm of a real or complex matrix, with applications to condition estimation (Algorithm 674)*, ACM Trans. Math. Software **14** (1988), no. 4, 381–396.
41. ———, *The accuracy of solutions to triangular systems*, SIAM J. Numer. Anal. **26** (1989), no. 5, 1252–1265.
42. ———, *Computing error bounds for regression problems*, Statistical Analysis of Measurement Error Models and Applications, Contemporary Mathematics 112 (Philip J. Brown and Wayne A. Fuller, eds.), American Mathematical Society, Providence, RI, USA, 1990, pp. 195–208.
43. ———, *How accurate is Gaussian elimination?*, Numerical Analysis 1989, Proceedings of the 13th Dundee Conference (D. F. Griffiths and G. A. Watson, eds.), Pitman Research Notes in Mathematics, vol. 228, Longman Scientific and Technical, Essex, UK, 1990, pp. 137–154.
44. ———, *Iterative refinement enhances the stability of QR factorization methods for solving linear equations*, BIT **31** (1991), 447–468.
45. ———, *The Test Matrix Toolbox for Matlab*, Numerical Analysis Report No. 237, Manchester Centre for Computational Mathematics, Manchester, England, December 1993.
46. Nicholas J. Higham and Philip A. Knight, *Componentwise error analysis for stationary iterative methods*, Linear Algebra, Markov Chains, and Queueing Models (Carl D. Meyer and Robert J. Plemmons, eds.), IMA Volumes in Mathematics and its Applications, vol. 48, Springer-Verlag, New York, 1993, pp. 29–46.
47. Roger A. Horn and Charles R. Johnson, *Matrix analysis*, Cambridge University Press, 1985.
48. ———, *Topics in matrix analysis*, Cambridge University Press, 1991.
49. Ilse C. F. Ipsen and Carl D. Meyer, *Uniform stability of Markov chains*, SIAM J. Matrix Anal. Appl. **15** (1994), no. 4, 1061–1074.
50. W. Kahan, *Accurate eigenvalues of a symmetric tri-diagonal matrix*, Technical Report No. CS41, Department of Computer Science, Stanford University, 1966, Revised June 1968.
51. ———, *Numerical linear algebra*, Canadian Math. Bulletin **9** (1966), 757–801.
52. Z. V. Kovarik, *Compatibility of approximate solutions of inaccurate linear equations*, Linear Algebra and Appl. **15** (1976), 217–225.
53. Colm Art O'Cinneide, *Entrywise perturbation theory and error analysis for Markov chains*, Numer. Math. **65** (1993), 109–120.
54. W. Oettli and W. Prager, *Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides*, Numer. Math. **6** (1964), 405–409.
55. W. Oettli, W. Prager, and J. H. Wilkinson, *Admissible solutions of linear systems with not sharply defined coefficients*, SIAM J. Numer. Anal. **2** (1965), no. 2, 291–299.
56. Beresford N. Parlett, *The symmetric eigenvalue problem*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1980, Reprinted by Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 1998.
57. Svatopluk Poljak and Jiří Rohn, *Checking robust nonsingularity is NP-hard*, Math. Control Signals Systems **6** (1993), 1–9.
58. J. L. Rigal and J. Gaches, *On the compatibility of a given solution with the data of a linear system*, J. Assoc. Comput. Mach. **14** (1967), no. 3, 543–548.
59. Jiří Rohn, *New condition numbers for matrices and linear systems*, Computing **41** (1989), 167–169.
60. ———, *Systems of linear interval equations*, Linear Algebra and Appl. **126** (1989), 39–78.
61. ———, *Nonsingularity under data rounding*, Linear Algebra and Appl. **139** (1990), 171–174.
62. Youcef Saad, *Numerical methods for large eigenvalue problems*, Manchester University Press, Manchester, and Halsted Press, New York, 1992.
63. Robert D. Skeel, *Scaling for numerical stability in Gaussian elimination*, J. Assoc. Comput. Mach. **26** (1979), no. 3, 494–526.
64. ———, *Iterative refinement implies numerical stability for Gaussian elimination*, Math. Comp. **35** (1980), no. 151, 817–832.
65. ———, *Effect of equilibration on residual size for partial pivoting*, SIAM J. Numer. Anal. **18** (1981), no. 3, 449–454.
66. G. W. Stewart, *Error and perturbation bounds for subspaces associated with certain eigenvalue problems*, SIAM Review **15** (1973), no. 4, 727–764.

67. ———, *Perturbation bounds for the QR factorization of a matrix*, SIAM J. Numer. Anal. **14** (1977), no. 3, 509–518.
68. ———, *Research, development, and LINPACK*, Mathematical Software III (John R. Rice, ed.), Academic Press, New York, 1977, pp. 1–14.
69. ———, *Stochastic perturbation theory*, SIAM Review **32** (1990), no. 4, 579–610.
70. ———, *On the perturbation of LU, Cholesky, and QR factorizations*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 4, 1141–1145.
71. G. W. Stewart and Ji-guang Sun, *Matrix perturbation theory*, Academic Press, London, 1990.
72. J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, Springer-Verlag, New York, 1980.
73. Ji-guang Sun, *Perturbation bounds for the Cholesky and QR factorizations*, BIT **31** (1991), 341–352.
74. ———, *Rounding-error and perturbation bounds for the Cholesky and LDL^T factorizations*, Linear Algebra and Appl. **173** (1992), 77–97.
75. A. van der Sluis, *Condition numbers and equilibration of matrices*, Numer. Math. **14** (1969), 14–23.
76. Krešimar Veselić and Ivan Slapničar, *Floating-point perturbations of Hermitian matrices*, Linear Algebra and Appl. **195** (1993), 81–116.
77. Bertil Waldén, Rune Karlson, and Ji-guang Sun, *Optimal backward perturbation bounds for the linear least squares problem*, Numerical Linear Algebra with Applications **2** (1995), no. 3, 271–286.
78. Per-Åke Wedin, *Perturbation theory for pseudo-inverses*, BIT **13** (1973), 217–232.
79. J. H. Wilkinson, *Error analysis of floating-point computation*, Numer. Math. **2** (1960), 319–340.
80. ———, *Rounding errors in algebraic processes*, Notes on Applied Science No. 32, Her Majesty's Stationery Office, London, 1963, Also published by Prentice-Hall, Englewood Cliffs, NJ, USA. Reprinted by Dover, New York, 1994.
81. ———, *The algebraic eigenvalue problem*, Oxford University Press, 1965.
82. Hongyuan Zha, *A componentwise perturbation analysis of the QR decomposition*, SIAM J. Matrix Anal. Appl. **14** (1993), no. 4, 1124–1131.

Department of Mathematics, University of Manchester, Manchester, M13 9PL, England

E-mail address: na.nhigham@@na-net.ornl.gov