SIAM J. MATRIX ANAL. APPL.
Vol. 16, No. 1, pp. 93–112, January 1995
© 1995 Society for Industrial and Applied Mathematics
007

ON THE SENSITIVITY OF SOLUTION COMPONENTS IN LINEAR SYSTEMS OF EQUATIONS*

S. CHANDRASEKARAN† AND I. C. F. IPSEN†

Abstract. Expressions are presented for the errors in *individual* components of the solution to systems of linear equations and linear least squares problems. No assumptions about the structure or distribution of the perturbations are made.

The resulting "componentwise condition numbers" measure the sensitivity of each solution component to perturbations. It is shown that any linear system has at least one solution component whose sensitivity to perturbations is proportional to the condition number of the matrix; but there may exist many components that are much better conditioned. Unless the perturbations are restricted, no norm-based relative error bound can predict the presence of well-conditioned components, so these componentwise condition numbers are essential.

For the class of componentwise perturbations, necessary and sufficient conditions are given under which Skeel's condition numbers are informative, and it is shown that these conditions are similar to conditions where componentwise condition numbers are useful. Numerical experiments not only confirm that these circumstances do occur frequently, they also illustrate that for many classes of matrices the ill conditioning of the matrix is due to a few rows of the inverse only. This means that many of the solution components are computed more accurately than current analyses predict.

Key words. condition number, diagonal scaling, forward error, linear system, least squares, perturbation theory

AMS subject classifications. 65F05, 65F20, 65F35, 15A06, 15A09, 15A12, 15A45, 15A60

1. Introduction.

Certain problems in statistics [33], combustion [26], and molecular conformation [10] require the solution of systems of linear equations whose individual solution components have physical significance; knowledge about the accuracy in the computation of the solution components is important. For the solution of problems involving Markov chains, for instance, it turns out that all solution components exhibit essentially the same sensitivity to perturbations in the data [25]. In [8] it is necessary to analyse individual solution components to demonstrate the convergence of inverse iteration in finite precision.

1.1. Motivation.

Consider the solution of a system of linear equations $Ax = b$ with nonsingular coefficient matrix A. The computed solution $\bar{x}$, which is usually different from the true solution x, can be viewed as the true solution to a perturbed system $(A + F)\bar{x} = b + f$.

So far, little work has dealt with trying to assess the error in individual solution components of a linear system; exceptions are the stability analyses of algorithms for solving particular structured linear systems, e.g., [3], [20], [22], [23]. The conventional way of estimating the error in $\bar{x}$, as compared to the true solution x, is to estimate an upper bound on the norm-based relative error $\|\bar{x} - x\|/\|x\|$. The most commonly used first-order bound is

$$\frac{\|\bar{x} - x\|}{\|x\|} \dot{\leq} \kappa(A)(\rho_A + \rho_b),$$

* Received by the editors May 18, 1992; accepted for publication (in revised form) by J. Bunch, June 29, 1993. The work presented in this paper was supported by Defense Advanced Research Projects Agency contract N00014-88-K-0573 and by National Science Foundation grant CCR-9102853.

† Department of Computer Science, Yale University, New Haven, Connecticut 06520 (chandras@math.ncsu.edu and ipsen@math.ncsu.edu).

where the condition number $\kappa(A) = \|A\| \, \|A^{-1}\| \geq 1$ acts as an amplifier for the relative perturbations in the data $\rho_A = \|F\|/\|A\|$ and $\rho_b = \|f\|/\|b\|$.

In many situations this type of error assessment is just fine unless, however, the individual components of the solution have physical significance. The example of the matrix of order four below, which represents a special case of a class of matrices discussed in §7, illustrates that the condition number $\kappa(A)$ can severely overestimate the error in some components,

$$A = \begin{pmatrix} 0.4919 & 0.1112 & -0.6234 & -0.6228 \\ -0.5050 & -0.6239 & 0.0589 & 0.0595 \\ 0.5728 & -0.0843 & 0.7480 & 0.7483 \\ -0.4181 & 0.7689 & 0.2200 & 0.2204 \end{pmatrix}, \qquad b = \begin{pmatrix} 0.4351 \\ -0.1929 \\ 0.6165 \\ -0.8022 \end{pmatrix}.$$

The first three columns of A are nearly orthogonal while the last two columns are almost identical. Both the two-norm condition number $\kappa_2(A)$ and Skeel's condition number [31] are larger than 10^3 (note that the matrix is not ill scaled). But the "componentwise condition numbers" that we introduce in this paper turn out to be

$$< 1.1, \quad < 1.1, \quad > 10^3, \quad > 10^3 .$$

This means that the first two components of x are well conditioned, regardless of the perturbations, and the remaining two are ill conditioned. To illustrate this, compare the "exact" solution x computed with 16-digit arithmetic with the solution $\bar{x}$ computed with 4-digit arithmetic, which can be viewed as the solution to a perturbed problem,

$$x = \begin{pmatrix} 1.000075414240576 \\ -.5000879795933286 \\ -.0242511388797165 \\ .02624513955005858 \end{pmatrix}, \qquad \bar{x} = \begin{pmatrix} 1.000 \\ -.5003 \\ -.0589 \\ .06090 \end{pmatrix}.$$

As predicted by our componentwise condition numbers, the first two components are accurate to almost four digits, whereas the last two have no accuracy whatsoever. As far as we know no other existing condition numbers can predict the well conditioning of the first two components of this system.

1.2. Overview. Given a linear system $Ax = b$ of full column rank and a perturbed system $(A + F)\bar{x} = b + f$, we derive expressions for the error in individual components of the computed solution $\bar{x}$ (§2). Our work is more general than that of Skeel [31] on componentwise perturbations and that of Stewart [34] on stochastic perturbations because we make no assumptions about the perturbations F and f, their size, structure, or distribution.

We associate with a linear system $Ax = b$ not a single condition number but a set of "componentwise condition numbers," one for each solution component. These condition numbers provide a clear separation of the three factors responsible for the loss of accuracy in the computed solution: relative magnitude of the solution components, matrix condition, and relationship between matrix and right-hand side.

We show that there is at least one component of the solution vector whose sensitivity to relative perturbations is proportional to the condition number of the matrix; but there may exist components that are much better conditioned. Consequently, unless the perturbations are restricted, *no* norm-based relative error bound can ever

predict the presence of well-conditioned components in x. Therefore, our componentwise condition numbers are essential.

Along the way, we comment on the tightness of norm-based error bounds (§3), and we clarify some results of Chan and Foulser [6] regarding the influence of the right-hand side on the sensitivity of the solution to perturbations (§4).

We also provide a geometric interpretation (§5) of our condition numbers, which in turn leads to a geometric interpretation of rank-revealing QR factorisations. Unlike traditional condition numbers, our componentwise condition numbers are able to indicate how linearly dependent individual matrix columns are on other columns. They can therefore be considered a continuation of Stewart's work on collinearity in regression problems [33].

We further show that the relative errors in individual components of a linear system are reduced by column scaling only if column scaling manages to reduce the perturbations (§6). Two simple examples are given where our componentwise condition numbers are significantly more accurate than the norm-based condition numbers (§7). We extend the results for linear systems to the solution of linear least squares problems $\min_y \|Ay - b\|$ of full column rank (§8).

For the class of componentwise perturbations, we give necessary and sufficient conditions under which Skeel's condition numbers are informative, and we show that these conditions are similar to those where componentwise condition numbers are useful (§9). Numerical experiments not only confirm that these circumstances do occur frequently, they also illustrate that for many classes of matrices the ill conditioning of the matrix is due to *a few* rows of the inverse only (§11). This means that many of the solution components are computed more accurately than current analyses would lead us to believe. Finally we demonstrate that a componentwise error bound for componentwise perturbations can be significantly better than the norm-based error bounds.

Existing software can be used to compute or estimate componentwise condition numbers (§10). We also prove that the problem of estimating componentwise condition numbers for triangular matrices by means of the comparison matrix is well conditioned.

2. Condition numbers for linear systems. This section presents expressions for errors in individual solution components of linear systems with full column rank and defines condition numbers for each component.

As for notation, $\|\cdot\|$ represents the two-norm and e_i stands for the ith column of the identity matrix I. Let A be an $n \times m$ matrix A of rank m. Its condition number is $\kappa(A) = \|A\| \, \|A^\dagger\|$ and the rows of its left-inverse $A^\dagger$ are denoted by r_i^T.

Regarding perturbations in the right-hand side, the treatment of linear systems and least squares problems can be combined. Suppose the exact solution $x \neq 0$ solves $\min_y \|Ay - b\|$, while the computed solution $\bar{x}$ solves $\min_y \|Ay - (b + f)\|$. Let β_i be the angle between r_i and b, and ψ_i the angle between r_i and f. If $x_i \neq 0$ and $\epsilon_b = \|f\|/\|b\|$ then

$$\text{(RE1)} \qquad \frac{\bar{x}_i - x_i}{x_i} = \frac{1}{\cos\beta_i} \, \epsilon_b \cos\psi_i = \frac{\|b\|}{\|A\| \, \|x\|} \, \frac{\|x\|}{x_i} \, \|A\| \, \|r_i\| \, \epsilon_b \cos\psi_i.$$

Regarding perturbations in the matrix of a linear system, suppose the exact solution $x \neq 0$ solves $Ax = b$, while the computed solution $\bar{x} \neq 0$ solves $(A + F)\bar{x} = b$.

Denote by ψ_i the angle between r_i and $F\bar{x}$. If $x_i \neq 0$ and $\epsilon_A = \frac{\|F\bar{x}\|}{\|A\|\,\|\bar{x}\|}$ then

$$\text{(RE2)} \qquad \frac{\bar{x}_i - x_i}{x_i} = -\frac{1}{\cos\beta_i}\frac{\|F\bar{x}\|}{\|b\|}\cos\psi_i = -\frac{\|\bar{x}\|}{x_i}\,\|A\|\,\|r_i\|\,\epsilon_A\cos\psi_i.$$

The perturbations in the first expressions for (RE1) and (RE2) are amplified by $1/\cos\beta_i$. Hence the relative error in $\bar{x}_i$ is likely to increase with increasing orthogonality of r_i and b.

The second expressions in (RE1) and (RE2) have two amplification factors in common: the magnitude of x_i relative to $\|x\|$, and the matrix condition $\|A\|\,\|r_i\| \leq \kappa(A)$. The term

$$\frac{\|b\|}{\|A\|\,\|x\|} \geq \frac{1}{\kappa(A)}$$

in (RE1) occurs in the error expressions for all $\bar{x}_i$ and describes the relation between matrix and right-hand side. In the case of linear systems $Ax = b$ it has the upper bound

$$\frac{\|b\|}{\|A\|\,\|x\|} = \frac{\|Ax\|}{\|A\|\,\|x\|} \leq 1.$$

The expressions (RE1) and (RE2) provide a clear separation of the three factors responsible for the loss of accuracy in the computed solution: relative magnitude of the solution components, matrix condition, and relationship between matrix and right-hand side.

Now we determine when the amplification factors are maximal. If $\|r_{\max}\| = \max_k \|r_k\|$ is the row of largest norm in $A^\dagger$ then

$$\text{(CN)} \qquad \|A\|\,\|r_{\max}\| \leq \kappa(A) \leq \sqrt{m}\|A\|\,\|r_{\max}\|.$$

Applying inequalities (CN) to the componentwise relative errors (RE1) and (RE2) shows that there must exist a component $\bar{x}_k$ for which

$$\frac{|\bar{x}_k - x_k|}{|x_k|} \geq \frac{1}{\sqrt{m}}\frac{\|b\|}{\|A\|\,\|x\|}\,\kappa(A)\,\frac{\|x\|}{|x_k|}\,\epsilon_b|\cos\psi_k|$$

and

$$\frac{|\bar{x}_k - x_k|}{|x_k|} \geq \frac{1}{\sqrt{m}}\kappa(A)\,\frac{\|x\|}{|x_k|}\,\epsilon_A|\cos\psi_k|.$$

Therefore, the sensitivity of x_k to matrix perturbations is proportional to the condition number of A, and is proportional to right-hand side perturbations only when the right-hand side has an appropriate direction, that is, whenever $\frac{\|b\|}{\|A\|\,\|x\|}$ is not too small.

DEFINITION 1. *Let $x \neq 0$ solve the linear system $Ax = b$ with $n \times m$ matrix A of rank m, and let $\bar{x} \neq 0$ be the computed solution. If $r_i^T = e_i^T A^\dagger$, then the quantities*

$$\frac{\|\bar{x}\|}{|x_i|}, \quad \|A\|\,\|r_i\|, \quad 1 \leq i \leq m,$$

are called componentwise condition numbers for the linear system *or* condition numbers for x_i.

Support for this kind of definition comes from earlier work of Stewart [33] who introduces the "collinearity indices" $\kappa_i = \|a_i\| \, \|r_i\|$ that represent the scaling-invariant version of $\|A\| \, \|r_i\|$. The main difference between Stewart's condition numbers and ours is that the collinearity indices are designed to reflect the linear dependence of the matrix columns, while our componentwise condition numbers measure the conditioning of the linear system: matrix plus right-hand side.

In 1970 van der Sluis [38], [39] realised the need to distinguish the conditioning of individual components of x and the fact that the conditioning depends on the relative size of a component. He introduced the notion of "ith column condition number of A," $\|A^{-1}\| \, \|a_i\|$, and derived the similar-looking normwise relative error bound (here $f = 0$)

$$\frac{\|\bar{x} - x\|}{\|x\|} \leq \frac{\|F\|}{\|A\|} \sum_i \|A^{-1}\| \, \|a_i\| \frac{|x_i|}{\|x\|}.$$

3. Conventional error bounds. This section argues that for any linear system there exist perturbations for which the norm-based bounds on the relative error are as tight as possible. We also justify our particular representation of the matrix perturbations.

It follows from (RE1) and (CN) that for perturbations of the right-hand side,

$$\frac{1}{\sqrt{m}} \kappa(A) \frac{\|b\|}{\|A\| \, \|x\|} \epsilon_b \mu \leq \frac{\|\bar{x} - x\|}{\|x\|} \leq \sqrt{m} \kappa(A) \frac{\|b\|}{\|A\| \, \|x\|} \epsilon_b,$$

where

$$\epsilon_b = \frac{\|f\|}{\|b\|}, \qquad \mu = \frac{\max_i \{\|r_i\| \, |\cos \psi_i|\}}{\max_k \|r_k\|}.$$

As for perturbations of the matrix,

$$\frac{1}{\sqrt{m}} \kappa(A) \frac{\|\bar{x}\|}{\|x\|} \epsilon_A \mu \leq \frac{\|\bar{x} - x\|}{\|x\|} \leq \sqrt{m} \kappa(A) \frac{\|\bar{x}\|}{\|x\|} \epsilon_A,$$

where $\epsilon_A = \frac{\|F\bar{x}\|}{\|A\| \, \|\bar{x}\|}$.

In the absence of knowledge about the values of $\cos \psi_i$, we must assume the worst case $\mu = 1$, which implies that the norm-based error bounds are tight. Thus the conventional upper bounds are as good as possible given that one has chosen to measure a *norm*-based error. As a consequence, if the normwise bounds give unsatisfying information, it is not because the bounds are loose, but rather because an unsatisfying way of measuring the error was adopted in the first place.

The upper bounds for nonsingular linear systems commonly found in the literature are of the form

$$\frac{\|\bar{x} - x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \kappa(A)\rho_A} (\rho_A + \epsilon_b), \qquad \|A^{-1}F\| < 1,$$

e.g., §III.2.3 in [35], where the matrix perturbations are represented by $\rho_A = \|F\|/\|A\|$. In contrast, our representation of the matrix perturbations is ϵ_A. This is a sensible measure because ϵ_A represents the smallest possible matrix perturbation, as we now show.

For given $Ax = b$ and $\bar{x}$, let $F_{\min}$ be the perturbation of smallest Frobenius norm among all perturbations F that satisfy $(A+F)\bar{x} = b$ ($F_{\min}$ also has smallest two-norm

among all such perturbations). From Theorem III.2.16 in [35], and also [27], it follows that $F_{\min}$ satisfies

$$(A + F_{\min})\bar{x} = b, \qquad \epsilon_{\min} = \frac{\|F_{\min}\|}{\|A\|} = \frac{\|F\bar{x}\|}{\|A\|\,\|\bar{x}\|},$$

which is exactly the matrix perturbation ϵ_A in the relative error (RE2).

4. Special right-hand sides for linear systems. This section analyses error bounds for linear systems $Ax = b$ whose right-hand side b is a singular vector associated with the smallest singular value σ_m of A. We show that in this case *all* solution components are sensitive to perturbations.

In this case

$$\|A^\dagger b\|/\|b\| = 1/\sigma_m = \|A^\dagger\|$$

and

$$\frac{\|b\|}{\|A\|\,\|x\|} = \frac{1}{\kappa(A)}, \qquad \frac{\|b\|}{\|A\|\,\|x\|}\,\|A\|\,\|r_i\| = \frac{\|r_i\|}{\|A^\dagger\|} \le 1.$$

This implies together with (RE1) that the relative sensitvity of all solution components to right-hand side perturbations is solely determined by their relative magnitude.

According to §2, the norm-based error satisfies

$$\frac{1}{\sqrt{m}}\epsilon_b\mu \le \frac{\|\bar{x} - x\|}{\|x\|} \le \sqrt{m}\epsilon_b.$$

This means the norm-based relative error is about the same magnitude as the perturbation in the right-hand side and does not depend on the condition number of A. This was already observed in [6].

Chan and Foulser [6] try to incorporate a potential relationship between right-hand side and matrix by modifying the conventional bound

$$\frac{\|\bar{x} - x\|}{\|x\|} \le \frac{\kappa(A)}{1 - \kappa(A)\rho_A}(\rho_A + \epsilon_b), \qquad \rho_A = \frac{\|F\|}{\|A\|}.$$

Let

$$A = U\Sigma V^T, \qquad \text{where} \quad U = (\,u_1 \quad \dots \quad u_n\,), \qquad \sigma_1 \ge \sigma_2 \ge \cdots \ge \sigma_n > 0,$$

be the singular value decomposition (SVD) of a nonsingular matrix A with singular values σ_i and right singular vectors u_i. According to Theorem 1 in [6], if $A\bar{x} = b + f$ and P_k is the orthogonal projection onto the space spanned by $u_{n-k+1}, \dots, u_n$,

$$\frac{\|\bar{x} - x\|}{\|x\|} \le \frac{\sigma_{n-k+1}}{\sigma_n}\left(\frac{\|P_k b\|}{\|b\|}\right)^{-1}\epsilon_b.$$

They conclude that if, for some k, a large fraction of b lies in the space spanned by $u_{n-k+1}, \dots, u_n$, and if $\sigma_{n-k+1} \approx \sigma_n$, then x "is relatively insensitive to perturbations in b." For instance, if $b = u_n$ then $P_1 b = b$,

$$\frac{\|\bar{x} - x\|}{\|x\|} \le \epsilon_b,$$

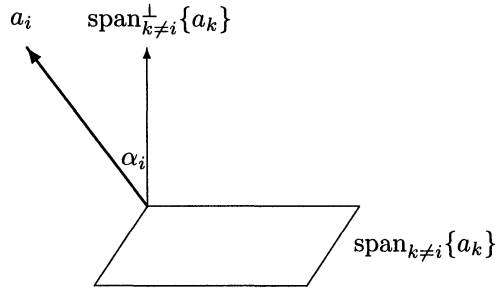

FIG. 1. *Angles associated with columns.*

and we conclude that x is insensitive to perturbations in b.

The interpretation of Theorem 1 given in [6] is valid if f represents the *input* error in the data b. However, we do not agree with the application of Theorem 1 in the case when f represents a backward error chosen to satisfy $A\bar{x} = b + f$ because f depends on the size of $\bar{x}$. Since $F_{\min} = -f\bar{x}^T/\bar{x}^T\bar{x}$ is the perturbation of smallest two-norm and Frobenius norm satisfying $(A + F_{\min})\bar{x} = b$, Theorem III.2.16 in [35], we obtain from the first expression in (RE1)

$$\frac{|\bar{x}_i - x_i|}{|x_i|} \geq \frac{\|A\|\,\|\bar{x}\|}{\|b\|}\,\epsilon_{\min}|\cos\psi_i|.$$

When $b = u_n$, the common term $\|A\|\,\|\bar{x}\|/\|b\|$ is approximately σ_1/σ_n and the sensitivity of *all* solution components is proportional to the condition number. A slightly different argument based on the use of the perturbations

$$\epsilon_{\min} = \frac{\|F_{\min}\|}{\|A\|} = \frac{\|b\|}{\|A\|\,\|\bar{x}\|}\epsilon_b$$

implies that for $b = u_n$ we have $\epsilon_b \approx \kappa(A)\epsilon_{\min}$ and the ill conditioning is merely hidden in the perturbation ϵ_b. Consequently, all components of x are extremely sensitive to perturbations if A is ill conditioned, which disagrees with the interpretation in [6].

5. Geometric interpretation. This section gives a geometric interpretation of the componentwise condition numbers. It is shown that $\|r_i\|$ reflects the linear dependence of column i of A on all other columns. This, in turn, leads to a geometric justification for rank-revealing QR factorisations.

First of all, the size of the $\|r_i\|$ reflects the linear dependence of the ith column of A on all others because

$$\|r_i\| = \frac{1}{\|a_i\|\cos\alpha_i},$$

where α_i is the angle between r_i and a_i. This follows from the expression $1 = r_i^T a_i = \|r_i\|\,\|a_i\|\cos\alpha_i$ for the ith diagonal element of $I = A^\dagger A$, which also implies that $\cos\alpha_i > 0$, so $-\frac{1}{2}\pi < \alpha_i < \frac{1}{2}\pi$. Because $e_i^T = r_i^T A$, r_i is orthogonal to all columns of A except for a_i, see Fig. 1.

To obtain a geometric meaning for r_1, partition $A = (\,a_1 \quad A_1\,)$, where a_1 represents the first column of A and A_1 represents the remaining columns. Let $-\hat{a}_1$ be the residual in the least squares approximation of a_1 by the columns of A_1,

$\|\hat{a}_1\| = \min_y \|A_1 y - a_1\|$ and let $-\check{A}_1$ be the residual in the least squares approximation of the columns of A_1 by a_1, $\|\check{A}_1\| = \min_y \|a_1 y^T - A_1\|$. As in the derivation of the formulae for partial correlation coefficients in [11] one can now show that

$$A^\dagger = (A^T A)^{-1} A^T = \begin{pmatrix} (\hat{a}_1^T a_1)^{-1} \hat{a}_1^T \\ \\ (\check{A}_1^T A_1)^{-1} \check{A}_1^T \end{pmatrix}.$$

It follows that the first row r_1^T of $A^\dagger$ lies in the same direction as the residual $-\hat{a}_1$ in the least squares approximation of column a_1 by the remaining columns. The residual, in turn, is just the projection of a_1 onto the orthogonal complement of the range of A_1. Hence, $\|r_1\| = 1/\|\hat{a}_1\|$, which means that increasing linear dependence of a_1 on the other columns leads to larger $\|r_1\|$. Analogous statements hold for the other rows r_i^T of $A^\dagger$.

Already in [33] Stewart used a different argument to show that

$$\|\hat{a}_i\| = \min_y \|A_i y - a_i\| = 1/\|r_i\|.$$

Here we provide more justification for the choice of r_i as an indicator of sensitivity: because r_i is a multiple of the residual $\hat{a}_i$, the residual is inherent in A and thus represents a most natural choice for sensitivity measure.

Angles between subspaces spanned by different columns of a matrix also occur in the context of nonsymmetric eigenvalue problems [12], [29].

5.1. Application. Our geometric interpretation of the rows of the left-inverse explains certain algorithms for rank-revealing QR factorisations. These factorisations appeared first in [15], [4], [16], [18], and are further analysed and refined in [32], [13], [5], [33], [9]. In the simplest case, the goal of a rank-revealing QR factorisation is to determine the most linearly dependent column of a matrix A.

The idea [9], [32] is based on the existence of a row of $A^\dagger$ that approximates $\|A^\dagger\|$ well. Perform a QR factorisation $AP = QR$, where Q has orthonormal columns, R is upper triangular, and the permutation matrix P is chosen so as to minimise the magnitude of the trailing diagonal element $(R)_{mm}$ of R. Then the inverse of this element, $1/|(R)_{mm}| = \|e_m^T R^{-1}\| = \|r_m^T\|$, is as large as possible, and the residual $1/\|r_m^T\|$ is as small as possible. Therefore the last column of AP is the column that can be best approximated by all other columns and so is the most linearly dependent among all columns.

6. Implications for column scaling. This section shows that the componentwise relative error decreases under column scaling only if column scaling actually reduces the perturbations.

A diagonal column scaling D of the least squares problem $\min_y \|Ay - b\|$ to $\min_z \|(AD)z - b\|$, where D is a nonsingular diagonal matrix, changes only the lengths of the columns but not the angles. In case of a column equilibrated matrix AD, [17, §3.5.2], and [37], [38], where the diagonal matrix D is chosen so that all columns of AD have identical length, the condition number of AD comes from the largest angle of A,

$$\frac{1}{\cos(\max_i \alpha_i)} \leq \|AD\| \, \|(AD)^\dagger\| \leq \frac{\sqrt{m}}{\cos(\max_i \alpha_i)}.$$

This bound already appeared in a different form in [33].

In [37], van der Sluis showed that a column equilibrated matrix A of order n has a condition number that is at most a factor $\sqrt{n}$ away from the lowest condition number among all matrices of the form AD. This would suggest that one could solve only linear systems and least squares problems with column equilibrated matrices so as to minimise the condition number in

$$\frac{\|\bar{x} - x\|}{\|x\|} \leq \sqrt{m}\kappa(A) \frac{\|b\|}{\|A\| \, \|x\|} \epsilon_b.$$

However, note that the condition number occurs in an upper bound.

Based on the expressions for the componentwise errors (RE1) and (RE2) we come to the following conclusions. In contrast to the norm-based condition numbers, the amplification factors $1/\cos\beta_i$ are preserved when the columns of A are multiplied by nonzero scalars. The computed solution $\bar{z}$ of the system $(AD)z = b$, where $z = D^{-1}x$, satisfies a perturbed system $AD\bar{z} = b + g$. Postmultiplication of A by D corresponds to premultiplication of $A^\dagger$ by D^{-1}, which changes only the lengths of the rows r_i^T in $A^\dagger$ but preserves the angles β_i between b and r_i. Hence the amplification factors $1/\cos\beta_i$ remain invariant under column scaling. Therefore the componentwise relative error decreases under column scaling only if column scaling manages to reduce the perturbations.

7. Example. This section contains two examples that illustrate the previous results. The first example represents a generalisation of the example from §1.1 and demonstrates that even a very ill-conditioned matrix may have robust solution components.

Consider a 4×4 orthogonal matrix $A = (\, a_1 \quad a_2 \quad a_3 \quad a_4 \,)$ and define a one-parameter family of matrices by

$$A(\lambda) = \left(a_1 \quad a_2 \quad a_3 \quad \tfrac{1}{\sqrt{1+\lambda^2}}(\lambda a_3 + a_4) \right).$$

Obviously $A(0) = A$ is a well-conditioned matrix and $A(\infty)$ is a singular matrix. For all λ, $\|A(\lambda)\| \leq 2$. When $\lambda < \infty$, the inverse is given by

$$[A(\lambda)]^{-1} = \begin{pmatrix} a_1^T \\ a_2^T \\ a_3^T - \lambda a_4^T \\ \sqrt{1+\lambda^2}\, a_4^T \end{pmatrix},$$

from which one computes

$$\cos\alpha_1 = \|a_1\| \cos\alpha_1 = \cos\alpha_2 = \|a_2\| \cos\alpha_2 = 1,$$
$$\cos\alpha_3 = \|a_3\| \cos(\alpha_3) = \cos\alpha_4 = \|a_4\| \cos(\alpha_4) = \frac{1}{\sqrt{1+\lambda^2}}.$$

Thus as $\lambda \to \infty$ the matrix $A(\lambda)$ becomes increasingly singular. Its condition number behaves like $O(\lambda)$. Note that the matrix $A(\lambda)$ is column equilibrated (and not necessarily row ill scaled) so the ill conditioning is a result of small angles rather than short columns.

Consider a linear system $A(\lambda)x(\lambda) = b$, where the right-hand side is independent of λ and can be represented as $b = \tau_1 a_1 + \tau_2 a_2 + \tau_3 a_3 + \tau_4 a_4$. Then

$$\cos\beta_1 = \frac{\tau_1}{\|b\|}, \quad \cos\beta_2 = \frac{\tau_2}{\|b\|}, \quad \cos\beta_3 = \frac{\tau_3 - \lambda\tau_4}{\|b\|\sqrt{1+\lambda^2}}, \quad \cos\beta_4 = \frac{\tau_4}{\|b\|}.$$

The solution vector is given by

$$x(\lambda) = (\,\tau_1 \quad \tau_2 \quad \tau_3 - \lambda\tau_4 \quad \sqrt{1+\lambda^2}\tau_4\,)^T .$$

The values of x_1 and x_2 are independent of λ, and so are $\|a_j\| \cos\alpha_j$ and $\cos\beta_j$ for $j = 1, 2$. So the sensitivity of the components x_1 and x_2 depends solely on their size relative to x. If, for instance, $|x_1| \gg |x_i|$ for $i \neq 1$ then, according to (RE), the error in x_1 is not amplified independent of the values of λ and the condition number of $A(\lambda)$.

The second example shows that *all* solution components can be sensitive to perturbations when the choice of right-hand side is unfortunate. In [25] we show that systems with uniformly sensitive solution components also occur in ill-conditioned Markov problems.

The coefficient matrix of the linear system $Ax = b$ is the Hilbert matrix with elements $1/(i+j-1)$ of order 4,

$$A = \begin{pmatrix} 1.00000000000000 & 0.50000000000000 & 0.33333333333333 & 0.25000000000000 \\ 0.50000000000000 & 0.33333333333333 & 0.25000000000000 & 0.20000000000000 \\ 0.33333333333333 & 0.25000000000000 & 0.20000000000000 & 0.16666666666667 \\ 0.25000000000000 & 0.20000000000000 & 0.16666666666667 & 0.14285714285714 \end{pmatrix}$$

while the right-hand side

$$b^T = (-0.02919332316479\ 0.32871205576319\ -0.79141114583313\ 0.51455274999716)$$

is a left singular vector corresponding to the smallest singular value of A. The condition number of A is at least 10^4. If the error matrix F has norm $\|F\| = 10^{-3}\|A\|$, then the solution of the system $(A+F)\bar{x} = b$ contains at least one component that has no accurate digits. We choose the following random matrix with norm $10^{-3}\|A\|$,

$$F = \begin{pmatrix} 0.00057208543036 & 0.00017162562911 & 0.00038139028691 & 0.00038139028691 \\ 0.00019069514345 & 0.00057208543036 & 0.00019069514345 & 0.00057208543036 \\ 0.00019069514345 & 0.00057208543036 & 0.00019069514345 & 0.00057208543036 \\ 0.00038139028691 & 0.00038139028691 & 0.00057208543036 & 0.00005720854304 \end{pmatrix}.$$

Computing x and $\bar{x}$ in 16-digit arithmetic gives

$$x = \begin{pmatrix} -301.88859986174430 \\ 3399.21637943995029 \\ -8183.99472310610599 \\ 5320.99783141589251 \end{pmatrix}, \qquad \bar{x} = \begin{pmatrix} 81.63154025985811 \\ -1310.35333852711346 \\ 3649.17285297454328 \\ -2572.42993839543533 \end{pmatrix}.$$

The components of $\bar{x}$ do not even have the correct sign, let alone any accurate digits. So all solution components of this system are sensitive to perturbations.

8. Condition numbers for least squares problems. This section presents expressions for componentwise errors in the solution of least squares problems of full column rank. The treatment in §2 on perturbations of the right-hand side is now extended to also allow perturbations in the matrix.

Suppose $x \neq 0$ solves

$$\min_y \|Ay - b\|, \qquad \text{where} \qquad r = b - Ax,$$

and $\bar{x} \neq 0$ solves

$$\min_y \|(A+F)y-(b+f)\|, \qquad \text{where} \qquad \bar{r} = b+f-(A+F)\bar{x} \neq 0.$$

Let $q_i^T = e_i^T(A^TA)^{-1}$ and define the following error angles: $\psi_{F,i}$ is the angle between r_i and $F\bar{x}$, $\psi_{f,i}$ is the angle between r_i and f, ω_i is the angle between r_i and $\bar{r}$, and $\omega_{q,i}$ is the angle between q_i and $F^T\bar{r}$. By applying (RE1) to the associated augmented nonsingular system one can show the following.

If $x_i \neq 0$ and $\epsilon_{A,r} = \frac{\|F^T\bar{r}\|}{\|A^T\|\,\|\bar{r}\|}$ then

$$\text{(LS)} \quad \frac{\bar{x}_i - x_i}{x_i} = \text{RE} + \frac{\|\bar{r}\|}{\|b\|}\frac{1}{\cos\beta_i}\cos\omega_i = \text{RE} + \frac{\|\bar{r}\|}{\|A\|\,\|\bar{x}\|}\frac{\|\bar{x}\|}{x_i}\,\|q_i\|\,\|A\|^2\,\epsilon_{A,r}\cos\omega_{q,i},$$

where

$$\text{RE} = -\frac{\|\bar{x}\|}{x_i}\,\|A\|\,\|r_i\|\left[\epsilon_A\cos\psi_{F,i} - \frac{\|b\|}{\|A\|\,\|\bar{x}\|}\epsilon_b\cos\psi_{f,i}\right]$$

is the componentwise relative error in the solution of a linear system solution.

Equations (LS) contain two different expressions that account for the least squares nature. The perturbation in the first expression is amplified by $1/\cos\beta_i$, which reflects how linearly dependent b is on the space spanned by a_k, $k \neq i$; and it is invariant under column scaling.

The relative perturbation $\epsilon_{A,r}\cos\omega_{q,i}$ in the second expression is amplified by three factors. The first factor represents, as in the error for linear system solution, the size of the component x_i relative to $\|\bar{x}\|$. The second factor $\|q_i\|\,\|A\|^2$ has the bounds

$$(\|r_i\|\,\|A\|)^2 \leq \|q_i\|\,\|A\|^2 \leq \|(A^TA)^{-1}\|\,\|A\|^2 = \kappa^2(A),$$

as a result of $\|q_i\| \geq \|r_i\|^2$. Since there exists a row r_k of $A^\dagger$ whose norm approximates $\|A^\dagger\|$ to a factor of $\sqrt{m}$, there must exist at least one component x_k for which

$$\|q_k\|\,\|A\|^2 \geq \frac{1}{m}\kappa^2(A).$$

The third factor $\frac{\|\bar{r}\|}{\|A\|\,\|\bar{x}\|}$ describes the relationship between matrix and right-hand side. If θ is the angle between b and the range of A, then the exact residual r satisfies

$$\frac{1}{\kappa(A)}\tan\theta \leq \frac{\|r\|}{\|A\|\,\|x\|} \leq \tan\theta$$

and for some x_k

$$\frac{1}{m}\kappa(A)\tan\theta \leq \frac{\|r\|}{\|A\|\,\|x\|}\,\|q_k\|\,\|A\|^2 \leq \kappa^2(A)\tan\theta.$$

Consequently, least squares problems are always more sensitive to ill conditioning than linear systems and, depending on the angle between b and the range of A, their sensitivity may be as high as the square of the condition number.

DEFINITION 2. *Let $x \neq 0$ solve the least squares problem $\min_y \|Ay-b\|$ with $n \times m$ matrix A of* rank m, *and let $\bar{x} \neq 0$ be the computed solution with residual $\bar{r} \neq 0$. If $q_i = e_i^T(A^TA)^{-1}$ and $r_i^T = e_i^TA^\dagger$ then the quantities*

$$\frac{\|\bar{x}\|}{|x_i|}, \quad \|A\|\,\|r_i\|, \quad \frac{\|\bar{r}\|}{\|A\|\,\|\bar{x}\|}\,\|A\|^2\|q_i\|, \quad 1 \leq i \leq m,$$

are called componentwise condition numbers for the least squares problem.

The condition numbers for linear systems from [38] and [39] are extended to least squares problems in [14].

9. A special class of perturbations. Unlike the previous sections, which assumed no knowledge about the perturbations, this section analyses the reduction in error bounds brought about by the special structure of perturbations resulting from floating point computations. This issue was first investigated by Skeel in [31] for the case of "componentwise perturbations." We provide necessary and sufficient conditions under which Skeel's condition numbers are useful, and we show that these conditions are similar to those where componentwise condition numbers are useful. The experiments in §11 illustrate that these conditions indeed occur frequently.

For $Ax = b$ and $(A + F)\bar{x} = b + f$ the perturbations F and f are called *componentwise perturbations* if the inequalities

$$|F| \leq \epsilon |A|, \qquad |f| \leq \epsilon |b|$$

hold componentwise for some $\epsilon \geq 0$.

In [31] Skeel defines a condition number that exploits componentwise perturbations. Theorem 2.1 in [31] shows that

$$\frac{\|\bar{x} - x\|_\infty}{\|x\|_\infty} \leq \epsilon \frac{\| \, |A^{-1}| \, |A| \, |x| \, + \, |A^{-1}| \, |b| \, \|_\infty}{(1 - \epsilon \| \, |A^{-1}| \, |A| \, \|_\infty) \, \|x\|_\infty},$$

and Skeel uses

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

as the condition number for the linear system $Ax = b$. He also introduces

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

as an upper bound for $\mathrm{cond}(A, x)$. A componentwise version of Skeel's condition number

$$e_i^T (|A^{-1}| \, |A| \, |x| \, + \, |A^{-1}| \, |b|)/|x_i|$$

is advocated in [28]; and [1] introduces condition numbers similar to the one used by Skeel for matrix inversion, least squares problems, and the solution of Vandermonde-like linear systems.

Skeel's condition number is invariant under row-scaling. Therefore, $\mathrm{cond}(A)$ may be much lower than the traditional condition number $\kappa_\infty(A) = \|A^{-1}\|_\infty \|A\|_\infty$ when the rows of A are ill scaled, i.e., when the norms of the rows of A differ widely. But the less known fact is that $\mathrm{cond}(A)$ can be much lower than $\kappa_\infty(A)$ *only* when the rows of A are ill scaled. The reasoning is as follows. Let e be the vector of all ones and D_R a nonsingular diagonal matrix with $D_R|A|e = e$, that is, the diagonal elements of D_R are the inverse row norms of A. Then

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

This means, if $\kappa_\infty(D) \approx 1$ then the rows of A are not ill scaled and $\mathrm{cond}(A) \approx \kappa_\infty(A)$, which limits the applicability of $\mathrm{cond}(A)$.

Remember that our componentwise condition numbers are useful if there are large differences among the $\|r_i\|$. It turns out that something similar holds for $\mathrm{cond}(A,x)$: $\mathrm{cond}(A,x)$ is useful only when the norms of the columns of A^{-1} differ widely in magnitude. Denote the columns of A by a_i and the columns of A^{-1} by p_i. If $j(i)$ is the index of the largest element in column i, $\|a_i\|_\infty = |a_{i,j(i)}|$, then

$$\| \, |A^{-1}| \, |A| \, |x| \, \|_\infty \geq \|p_{j(i)}\|_\infty \|a_i\|_\infty |x_i| = \|A\|_1 \|p_{j(i)}\|_\infty |x_i| \frac{\|a_i\|_\infty}{\|A\|_1}.$$

Choosing i such that $|x_i| = \|x\|_\infty$ and defining D_C as the diagonal matrix that equilibrates the columns of A, $e^T|A|D_C = e^T$, gives

$$\begin{aligned} \mathrm{cond}(A,x) = \frac{\| \, |A^{-1}| \, |A| \, |x| \, \|_\infty}{\|x\|_\infty} &\geq \|p_{j(i)}\|_\infty \|A\|_1 \frac{1}{n\kappa_\infty(D_C)} \\ &\geq \frac{\min_i \|p_i\|_\infty}{\|p_j\|_\infty} \kappa_1(A) \frac{1}{n^2\kappa_\infty(D_C)} \end{aligned}$$

for some column p_j of A^{-1}. This means, $\mathrm{cond}(A,x) \approx \kappa_1(A)$ for all x whenever A^{-1} is not badly column-scaled. Therefore the conditions under which $\mathrm{cond}(A,x)$ is useful are quite similar to those for our componentwise condition numbers.

It is possible to profitably combine Skeel's analysis with our componentwise errors because componentwise perturbations induce upper bounds on the cosines. If $A\bar{x} = b + f$ and $|f| \leq \epsilon|b|$ then

$$|\cos\psi_i| = \frac{|r_i^T f|}{\|r_i\| \, \|f\|} \leq \epsilon \frac{|r_i^T| \, |b|}{\|r_i\| \, \|f\|},$$

implies

$$\frac{|\bar{x}_i - x_i|}{|x_i|} \leq \epsilon \frac{|r_i^T| \, |b|}{|r_i^T b|} \tag{CRE1}$$

for error (RE1). Similarly, if $(A+F)\bar{x} = b$ and $|F| \leq \epsilon|A|$ then the upper bound for (RE2) simplifies to

$$\frac{|\bar{x}_i - x_i|}{|x_i|} \leq \epsilon \frac{|r_i^T| \, |A| \, |\bar{x}|}{|x_i|}. \tag{CRE2}$$

This last inequality illustrates that componentwise perturbations in our error expressions lead to a componentwise version of Skeel's condition number $\mathrm{cond}(A,x)$. Although these expressions already exist implicitly in Skeel's work, it is the observation that we lose a lot by taking norms that is important. Because the rows of the inverse may differ significantly in size, the difference between our bounds and $\mathrm{cond}(A,x)$ may be arbitrarily large as shown in the following example.

Let $\epsilon > 0$ and

$$A = \begin{pmatrix} \epsilon & \frac{1}{\epsilon} \\ & \frac{1}{\epsilon} \end{pmatrix}, \qquad b = \begin{pmatrix} \epsilon^2 + \frac{1}{\epsilon} \\ \frac{1}{\epsilon} \end{pmatrix},$$

so that

$$A^{-1} = \begin{pmatrix} \frac{1}{\epsilon} & -\frac{1}{\epsilon} \\ & \epsilon \end{pmatrix}, \qquad x = \begin{pmatrix} \epsilon \\ 1 \end{pmatrix}.$$

Hence

$$|A^{-1}|\,|A|\,|x| = |A^{-1}|\,|b| = \begin{pmatrix} \epsilon + \frac{2}{\epsilon^2} \\ 1 \end{pmatrix}.$$

Skeel's condition number $\mathrm{cond}(A,x)$ is unbounded as ϵ becomes small. In fact, it has the same order of magnitude $O(1/\epsilon^2)$ as the traditional condition number $\kappa_\infty(A)$, although the columns of A are badly scaled (this is because $|x|$ lies almost in the singular direction corresponding to the largest singular value of $|A^{-1}|\,|A|$). In contrast, the amplifier in our error bound for x_2, which is the largest component of x, equals $|r_2^T|\,|A|\,|x|/|x_2| = 1$.

Because our error expressions represent a componentwise version of Skeel's condition number, we get the same componentwise error bounds as appear in the literature. For instance, when $|A^{-1}|\,|b| = |x|$, as is the case for certain Vandermonde systems [20] and M-matrices with positive right-hand sides, the term amplifying ϵ in (CRE1) equals one. So the individual solution components are insensitive to perturbations in the right-hand side (an algorithm for such systems that gives rise to a small componentwise backward error f is called "weakly stable" in [23]).

For triangular M-matrices A with positive right-hand side b, it is shown in [22] that

$$|A^{-1}|\,|A|\,|x| \leq (2n-1)|x|,$$

which implies

$$|r_i^T|\,|A|\,|\bar{x}| = e_i^T A^{-1}|A|\bar{x} \leq (2n-1)|\bar{x}_i|.$$

Hence the term amplifying ϵ in (CRE2) is essentially bounded above by $2n-1$. This is true in particular if b is the vector of all ones. Thus, estimating the componentwise condition numbers of a triangular matrix by solving a linear system involving the comparison matrix, as in [21] and §10, is a well-conditioned problem.

10. Estimation of componentwise condition numbers. This section shows that componentwise condition numbers can be efficiently estimated with existing software.

For a $n \times m$ matrix A, bounds for $\|A\|$ can be determined in $O(mn)$ operations, and $\|\bar{x}\|/|x_i|$ and $\|\bar{r}\|$ can be estimated a posteriori in $O(mn)$ operations. This leaves the computation of $\|r_i\|$ and $\|q_i\|$. Numerical issues in the computation of the $\|r_i\|$, due to the potential ill conditioning of A, are addressed in [32] and in the context of statistical errors in [33]. If a factorization of A is available, then upper bounds on $\|r_i\|$ can be determined in $O(n^2)$ additional operations and an estimate of $\|q_i\|$ can be obtained by making use of the inequality $\|q_i\| \geq \|r_i\|^2$.

For instance, suppose the QR factorization

$$A = Q \begin{pmatrix} R \\ 0 \end{pmatrix}$$

is available, where Q is a $n \times n$ orthogonal matrix, and R is an $m \times m$ nonsingular upper triangular matrix. To compute $\|r_i\|$ and $\|q_i\|$, it suffices to work with R instead of A. From

$$q_i^T = e_i^T (A^T A)^{-1} = e_i^T R^{-1} R^{-T} = v_i^T R^{-T}, \qquad v_i = R^{-T} e_i,$$

it follows that q_i is the solution of the triangular system $Rq_i = v_i$ and $\|r_i\| = \|v_i\|$.

As for the actual computation of $\|q_i\|$ and $\|r_i\|$, observe that $v_m = R^{-T}e_m = \frac{1}{\rho}e_m$, where ρ is the element of R in position (m,m). Hence $\|r_m\| = 1/|\rho|$ and $Rq_m = \frac{1}{\rho}e_m$. Therefore, if a QR decomposition of A is available, $\|r_m\|$ is available right away and the computation of q_m requires $O(m^2)$ operations. This process can be carried out for all i, and is described in [32] for the computation of $\|r_i\|$ by permuting the columns of A. Gragg and Stewart [18] show how to efficiently "update" the QR factorisation from one permutation to the next in $O(m^2)$ operations; see also [17, §12.6].

Next, we indicate how the condition number estimators for triangular matrices in [21] can be used to compute upper bounds for the $\|r_i\|$ in $O(n^2)$ operations. Since A is triangular, $(A^{-1})_{ii} = 1/a_{ii}$ and $1/|a_{ii}| \leq \|r_i\| \leq \|r_i\|_1$. Replace A by its *comparison matrix* $C(A) = (c_{ij})$ of A [2], which is defined as

$$c_{ij} = \begin{cases} |a_{ii}| & \text{if } i = j, \\ -|a_{ij}| & \text{if } i \neq j, \end{cases}$$

and satisfies the componentwise inequalities

$$C(A)^{-1} \geq 0, \qquad |A^{-1}| \leq C(A)^{-1}$$

because it is an M-matrix [40]. The first inequality implies that the ith element of $C(A)^{-T}e$ equals $\|C(A)^{-T}e_i\|_1$, where e is the vector of all ones, while the second one implies $\|r_i\| \leq \|r_i\|_1 \leq \|C(A)^{-T}e_i\|_1$. Hence all $\|C(A)^{-T}e_i\|_1$ can be computed with a total of $O(n^2)$ operations by solving the system $C(A)^T y = e$. Since $C(A)$ is an M-matrix, so is $C(A)^T$. According to §9, the solution of linear systems with triangular M-matrices and positive right-hand side produces a small componentwise error. Hence, the estimation of componentwise condition numbers from the solution of $C(A)^T y = e$ is a well-conditioned problem.

In [7] we fit the linear-time algorithms in [19] for computing $\|A^{-1}\|_\infty$ for bi or tridiagonal matrices A to the computation of $\|r_i\|$.

We are currently investigating techniques based on appropriate rank-revealing QR decompositions that estimate componentwise condition numbers in $O(n^2)$ operations.

11. Numerical experiments. This section presents numerical experiments that reveal the existence of large classes of matrices for which the componentwise matrix condition numbers vary widely. For these matrices, componentwise condition numbers can therefore predict the sensitivity of individual solution components much more accurately than norm-based or Skeel's condition numbers.

Here we consider only nonsingular linear systems $Ax = b$. The componentwise condition numbers consist of two parts: the relative magnitude $\|\bar{x}\|/x_i$ of the solution component and the associated matrix condition $\|A\|\,\|r_i\|$ where $r_i = e_i^T A^{-1}$. We consider only matrices for which $\|r_i\|$ differ widely in size because they exhibit a large difference between $\|r_i\|$ and $\|A^{-1}\|$, as well as between $|r_i^T|\,|A|$ and $\mathrm{cond}(A)$. According to inequalities (CN), at least one $\|r_k\|$ approximates $\|A^{-1}\|$ to within a factor of $\sqrt{n}$, where n is the order of A. The potential for deviation of other $\|r_i\|$ from $\|A^{-1}\|$ increases, of course, with increasing ill conditioning of A. Below we present examples where some $\|r_i\|$ are orders of magnitude smaller than $\|A^{-1}\|$.

All experiments were performed in CLAM, version 2.00 [30], on a SPARCstation 1. The tests involved more than twenty classes of matrices, most of them from [24], their orders ranging up to $n = 500$. Among these, only the Minij and Pei matrices have r_i that are essentially identical in size. A group of matrices with a little more variation in

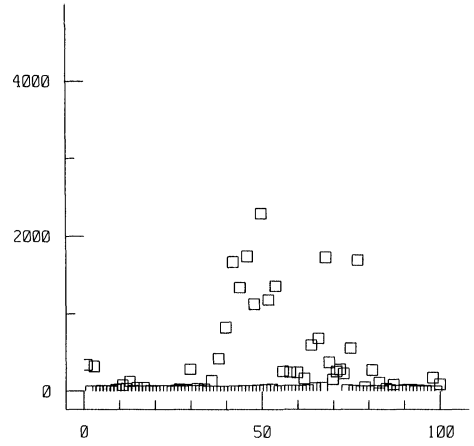

FIG. 2. *Random tridiagonal matrix A with $n = 100$, $\|A^{-1}\| = 4752$, $\min_i \|r_i\| = 1.1$.*

the $\|r_i\|$ are the highly ill-conditioned Pascal, Cauchy, Hilbert, and Lotkin matrices. In the group of matrices that comprises random symmetric and nonsymmetric matrices, random Toeplitz and Vandermonde matrices, at least half of the $\|r_i\|$ differ from $\|A^{-1}\|$ by a small multiple of ten. This means, at least half of the components of x are 1–2 digits more accurate than predicted by $\|A^{-1}\|$ (assuming the components are not too small). The group of matrices with the widest variation in the $\|r_i\|$ includes random tridiagonal matrices, Jordan matrices, Chebyshev–Vandermonde matrices, and triangular comparison matrices.

The surprising outcome of our experiments is that often only a few rows of A^{-1} are responsible for $\|A^{-1}\|$, while most of the remaining rows are small in size. This is more pronounced for ill-conditioned matrices. It also comes out in the plots in Figs. 2–7, where we plot $\|r_i\|$ against i, $1 \leq i \leq n$, for matrices from the last group. In case of high ill conditioning, the difference among the $\|r_i\|$ can be as high as 10^{15} for matrices of order $n = 100$. In addition, preliminary statistical analyses show that for these matrices usually more than half of the $\|r_i\|$ are small. Therefore, although a norm-based error bound would predict a total loss of accuracy, many components could actually be computed to a significant number of correct digits.

Figures 2–4 contain plots of three typical random tridiagonal matrices of order $n = 100$. The differences in the $\|r_i\|$ for each matrix are illustrated in Table 1.

TABLE 1

Figure	$\|A^{-1}\|$	$\min_i \|r_i\|$
2	4752	1.1
3	678	1.2
4	2577	1.1

Figure 5 shows the $\|r_i\|$ for a random Chebyshev–Vandermonde matrix of order $n = 10$, for which $\|A^{-1}\| = 11922$ and $\min \|r_i\| = .72$. Figures 6 and 7 plot the $\|r_i\|$ on a logarithmic scale for a random Jordan and a random unit upper triangular comparison matrix, respectively, both of order $n = 100$. The Jordan matrix has $\|A^{-1}\| = 4 \cdot 10^{14}$ and $\min \|r_i\| = 1.4$, while the triangular matrix has $\|A^{-1}\| = 2 \cdot 10^{17}$ and $\min \|r_i\| = 1.1$. Similar observations about the ill conditioning of random unit-triangular matrices are made in [36].

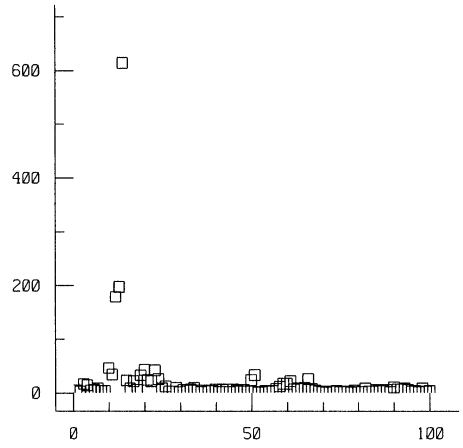

FIG. 3. *Random tridiagonal matrix A with* $n = 100$, $\|A^{-1}\| = 678$, $\min_i \|r_i\| = 1.2$.

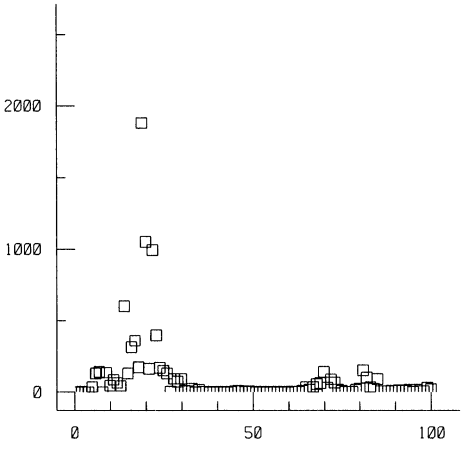

FIG. 4. *Random tridiagonal matrix A with* $n = 100$, $\|A^{-1}\| = 2577$, $\min_i \|r_i\| = 1.1$.

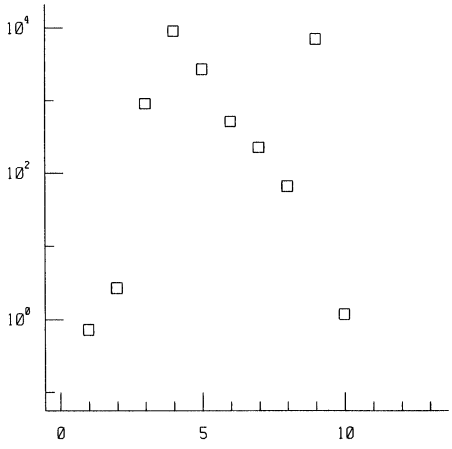

FIG. 5. *Random Chebyshev–Vandermonde matrix A with* $n = 10$, $\|A^{-1}\| = 11922$, $\min_i \|r_i\| = .72$.

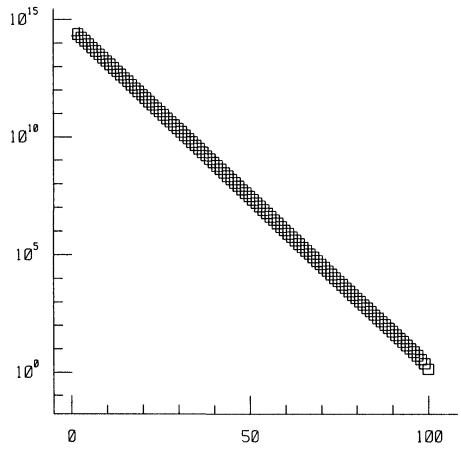

FIG. 6. *Random Jordan matrix A with diagonal element* .7183, $n = 100$, $\|A^{-1}\| = 4 \cdot 10^{14}$, $\min_i \|r_i\| = 1.4$.

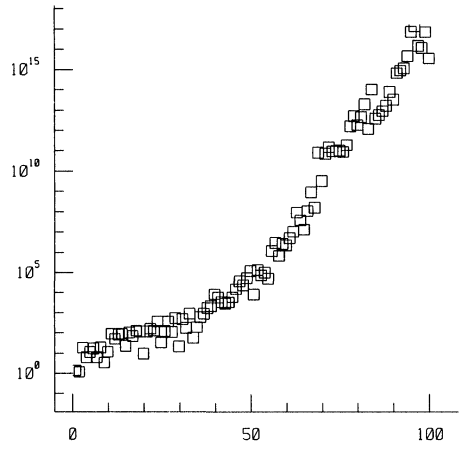

FIG. 7. *Random unit upper triangular comparison matrix A*, $n = 100$, $\|A^{-1}\| = 2 \cdot 10^{17}$, $\min_i \|r_i\| = 1.1$.

Acknowledgments. We would like to thank Stan Eisenstat and Jean-Marc Delosme for helpful discussions, Sham Sao for performing many of the numerical experiments, and Nick Higham for pointing out an error in an earlier version of the manuscript.

REFERENCES

[1] S. BARTELS AND D. HIGHAM, *The structured sensitivity of Vandermonde-like systems*, Numer. Math., 62 (1992), pp. 17–33.

[2] A. BERMAN AND R. PLEMMONS, *Nonnegative Matrices in the Mathematical Sciences*, Academic Press, New York, 1979.

[3] Å. BJÖRCK, *Component-wise perturbation analysis and error bounds for linear least squares solutions*, BIT, 31 (1991), pp. 238–244.

[4] P. BUSINGER AND G. GOLUB, *Linear least squares solutions by Householder transformations*, Numer. Math., 7 (1965), pp. 269–276.

[5] T. CHAN, *Rank revealing QR factorizations*, Linear Algebra Appl., 88/89 (1987), pp. 67–82.

[6] T. CHAN AND D. FOULSER, *Effectively well-conditioned linear systems*, SIAM J. Sci. Statist. Comput., 9 (1988), pp. 963–969.
[7] S. CHANDRASEKARAN AND I. IPSEN, *Perturbation theory for the solution of linear systems of equations and linear least squares problems*, Research Report 866, Department of Computer Science, Yale University, New Haven, CT, 1991.
[8] ———, *Finite precision analysis of inverse iteration*, manuscript.
[9] ———, *On rank-revealing QR factorisations*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 592–622.
[10] T. COLEMAN, D. SHALLOWAY, AND Z. WU, *Parallel packet annealing for molecular conformation* (II): *Simulated annealing and packet approximation*, Tech. Report, Advanced Computing Research Institute, Cornell University, Ithaca, NY, 1991.
[11] J. DELOSME AND I. IPSEN, *From Bareiss' algorithm to the stable computation of partial correlations*, J. Comput. Appl. Math., 27 (1989), pp. 53–91; *Parallel Algorithms for Numerical Linear Algebra* (*Advances in Parallel Computing*, 1), H. van der Vorst and P. van Dooren, eds., North Holland, Amsterdam, 1990.
[12] J. DEMMEL, *The condition number of equivalence transformations that block diagonalize matrix pencils*, SIAM J. Numer. Anal., 20 (1983), pp. 599–610.
[13] L. FOSTER, *Rank and null space calculations using matrix decomposition without column interchanges*, Linear Algebra Appl., 74 (1986), pp. 47–71.
[14] A. GEURTS, *A contribution to the theory of condition*, Numer. Math., 39 (1982), pp. 85–96.
[15] G. GOLUB, *Numerical methods for solving linear least squares problems*, Numer. Math., 7 (1965), pp. 206–216.
[16] G. GOLUB, V. KLEMA, AND G. STEWART, *Rank degeneracy and least squares problems*, Tech. Report STAN-CS-76-559, Computer Science Department, Stanford University, Stanford, CA, 1976.
[17] G. GOLUB AND C. VAN LOAN, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, MD, 1989.
[18] W. GRAGG AND G. STEWART, *A stable variant of the secant method for solving nonlinear equations*, SIAM J. Numer. Anal., 13 (1976), pp. 889–903.
[19] N. HIGHAM, *Efficient algorithms for computing the condition number of a tridiagonal matrix*, SIAM J. Sci. Statist. Comput., 7 (1986), pp. 150–165.
[20] ———, *Error analysis of the Björck–Pereyra algorithms for solving Vandermonde systems*, Numer. Math., 50 (1987), pp. 613–632.
[21] ———, *A survey of condition number estimation for triangular matrices*, SIAM Rev., 29 (1987), pp. 575–596.
[22] ———, *The accuracy of solutions to triangular systems*, SIAM J. Numer. Anal., 26 (1989), pp. 1252–1265.
[23] ———, *Stability analysis of algorithms for solving confluent Vandermonde-like systems*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 23–41.
[24] ———, *Algorithm* 694: *A collection of test matrices in MATLAB*, ACM TOMS, 17 (1991), pp. 289–305.
[25] I. IPSEN AND C. MEYER, *Uniform stability of Markov chains*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 1061–1074.
[26] D. KEYES AND M. SMOOKE, *Flame sheet starting estimates for counterflow diffusion flame problems*, J. Comput. Physics, 73 (1987), pp. 267–288.
[27] J. RIGAL AND J. GACHES, *On the compatibility of a given solution with the data of a linear system*, JACM, 14 (1967), pp. 543–548.
[28] J. ROHN, *New condition numbers for matrices and linear systems*, Computing, 41 (1989), pp. 167–169.
[29] A. RUHE, *Properties of a matrix with a very ill-conditioned eigenproblem*, Numer. Math, 15 (1970), pp. 57–60.
[30] SCIENTIFIC COMPUTING ASSOCIATES, INC., *CLAM User's Guide, The Computational Linear Algebra Machine*, New Haven, CT, 1990.
[31] R. SKEEL, *Scaling for numerical stability in Gaussian elimination*, JACM, 26 (1979), pp. 494–526.
[32] G. STEWART, *Rank degeneracy*, SIAM J. Sci. Statist. Comput., 5 (1984), pp. 403–413.
[33] ———, *Collinearity and least squares regression*, Statist. Sci., 2 (1987), pp. 68–100.
[34] ———, *Stochastic perturbation theory*, SIAM Rev., 32 (1990), pp. 579–610.
[35] G. STEWART AND J. SUN, *Matrix Perturbation Theory*, Academic Press, New York, 1990.
[36] L. TREFETHEN AND R. SCHREIBER, *Average-case stability of Gaussian elimination*, SIAM J. Matrix Anal. Appl., 11 (1990), pp. 335–360.

[37] A. VAN DER SLUIS, *Condition numbers and equilibration of matrices*, Numer. Math., 14 (1969), pp. 14–23.
[38] ———, *Condition, equilibration and pivoting in linear algebraic systems*, Numer. Math., 14 (1970), pp. 74–86.
[39] ———, *Stability of solutions of linear algebraic systems*, Numer. Math., 14 (1970), pp. 246–251.
[40] R. VARGA, *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, NJ, 1962.