SIAM J. SCI. STAT. COMPUT.
Vol. 1, No. 2, June 1980

0196-5204/80/0102-0005 $1.00/0

ESTIMATING MATRIX CONDITION NUMBERS*

DIANNE PROST O'LEARY†

Abstract. In this note we study certain estimators for the condition number of a matrix which, given an LU factorization of a matrix, are easily calculated. The main observations are that the choice of estimator is very norm-dependent, and that although some simple estimators are consistently bad, none is consistently best. These theoretical conclusions are confirmed by experimental data, and recommendations are made for the one and infinity norms.

Key words. Matrix condition number

1. Introduction. Cline, Moler, Stewart, and Wilkinson [1] give an excellent exposition of various methods for estimating the condition number of a matrix

$$\kappa(A)=\|A\|\|A^{-1}\|,$$

where $\|\cdot\|$ is some matrix norm compatible with a vector norm. One of the applications considered is that of estimating κ given a LU factorization of a matrix formed using partial pivoting:

$$A=LU,$$

where L is unit lower triangular, U is upper triangular, and all elements of L are bounded in absolute value by 1. The strategy suggested is to solve two linear systems,

$$A^Tx=e,$$

$$Ay=x,$$

and to use $\|y\|/\|x\|$ as the estimate for $\|A^{-1}\|$. Here the vector e is chosen during the first step of the solution procedure, finding z such that $U^Tz=e$. Each element e_i is ± 1, with sign to promote growth in the subsequent components of z. $\|A\|_1$ or $\|A\|_\infty$ can easily be calculated exactly, and $\|A\|_2$ can be estimated using, for example, the power method.

The experiments in [1] show this to be a good algorithm for the one-norm, but the strategy is norm-dependent as suggested by the following example.

Example. Let A_k be the Hadamard matrix of dimension $n=2^k$ defined by

$$A_1\equiv\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ 0 & -2\end{bmatrix}\equiv L_1U_1,$$

$$A_k\equiv\begin{bmatrix}A_{k-1} & A_{k-1}\\ A_{k-1} & -A_{k-1}\end{bmatrix}=\begin{bmatrix}L_{k-1} & 0\\ L_{k-1} & L_{k-1}\end{bmatrix}\begin{bmatrix}U_{k-1} & U_{k-1}\\ 0 & -2U_{k-1}\end{bmatrix}\equiv L_kU_k,\quad k>1.$$

It is easy to see that to estimate $\|A_k^{-1}\|$, the choice of e will be such that $A_k^Te_n=e$,

*Received by the editors September 26, 1979.
This work was supported by the National Bureau of Standards and the U.S. Office of Naval Research under Grant N00014-76-C-0391.

†Computer Science Department and Institute for Physical Science and Technology, University of Maryland, College Park, Maryland, and National Bureau of Standards, Gaithersburg, Maryland.

where e_n is the nth unit vector. Then $y=A_k^{-1}e_n=e/n$, and we obtain the estimates

$$\|A^{-1}\|_1 \simeq 1, \qquad \|A^{-1}\|_1 = 1,$$

$$\|A^{-1}\|_2 \simeq 1/\sqrt{n}\,, \qquad \|A^{-1}\|_2 = 1,$$

$$\|A^{-1}\|_\infty \simeq 1/n, \qquad \|A^{-1}\|_\infty = 1.$$

Although the one-norm estimate is exact, the infinity norm estimate is off by a factor of n.

2. Methods. To study the behavior of norm estimates, we develop some basic relations. Recall that the one-norm of a matrix is the maximum absolute column sum: if a matrix B has columns b_j with components b_{ij} then

$$\|B\|_1 = \max_{x\neq 0} \frac{\|Bx\|_1}{\|x\|_1} = \max_{j=1,\ldots,n} \sum_{i=1}^{n} |b_{ij}| = \max_{j=1,\ldots,n} b_j^T \tilde{e}_j,$$

where $\tilde{e}_j$ is a vector with components ± 1, with signs chosen to match those of b_j. Also recall that $\|B\|_\infty = \|B^T\|_1$. The strategy defined above is designed to produce a vector x which is close to being a maximizer for $\|Bx\|/\|x\|$, where $B=A^{-1}$. In the course of computing the vector x, another estimate, based on the sizes of the vectors e and $x=B^Te$, is available for the condition number. Let

$$\mu_1 = \|y\|_1/\|x\|_1,$$

$$\nu_1 = \|x\|_\infty/\|e\|_\infty,$$

$$\mu_\infty = \|y\|_\infty/\|x\|_\infty,$$

$$\nu_\infty = \|x\|_1/\|e\|_1,$$

where the first two estimate $\|A^{-1}\|_1$ and the last two estimate $\|A^{-1}\|_\infty$. Then

$$\nu_1 = \max_{j=1,\ldots,n} |b_j^T e|,$$

$$\nu_\infty = \frac{1}{n}\sum_{j=1}^{n} |b_j^T e|,$$

$$\mu_1 = \frac{\displaystyle\sum_{i=1}^{n}\left|\sum_{j=1}^{n}(b_j^T e)b_{ij}\right|}{\displaystyle\sum_{j=1}^{n}|b_j^T e|},$$

$$\mu_\infty = \frac{\displaystyle\max_{i=1,\ldots,n}\left|\sum_{j=1}^{n}(b_j^T e)b_{ij}\right|}{\displaystyle\max_{i=1,\ldots,n}|b_i^T e|}.$$

From this we obtain the relationships

$$\frac{\nu_1}{\mu_1}=\frac{\sum_{j=1}^{n}|b_j^Te|}{\sum_{i=1}^{n}\left|\sum_{j=1}^{n}\frac{(b_j^Te)}{\nu_1}b_{ij}\right|},$$

$$\nu_1 \geqq \nu_\infty,$$

$$\frac{\nu_\infty}{\mu_\infty}=\frac{\frac{1}{n}\sum_{j=1}^{n}|b_j^Te|}{\max_{i=1,\dots,n}\left|\sum_{j=1}^{n}\frac{(b_j^Te)}{\nu_1}b_{ij}\right|}\leqq\frac{\nu_1}{\mu_1},$$

The third relation says that the improvement gained by solving the second linear system, $Ay=x$, is greater for the infinity norm estimate than for the one-norm estimate. The second relation says that the first one-norm estimate is an upper bound for the first infinity norm estimate. For a symmetric matrix the two norms are equal and so the one-norm estimate is always more accurate. (No such relation exists between the estimates μ_1 and μ_∞.) The first equation gives the relation between the two one-norm estimators.

Unfortunately none of the estimators can be labeled as best. The estimate ν_1 (respectively, ν_∞) is sometimes greater than μ_1 (μ_∞) and sometimes less. Because of this, estimators which use all the information might be more useful:

$$\rho_1=\max(\mu_1,\nu_1),$$

$$\rho_\infty=\max(\mu_\infty,\nu_\infty).$$

To gain a better understanding of the behavior of the condition number estimators, tests were performed on matrices with elements taken from a uniform distribution on $[-1,1]$. The LINPACK [2] routine SGECO, which factors a matrix and returns an estimate of the condition number based on μ_1, was modified to compute all of the estimators. LINPACK's SGEDI was used to compute the inverse so that $\|A^{-1}\|$ could be calculated for comparison. Test results are summarized in Tables 1 and 2. Results for $5\leqq n\leqq 50$ were obtained using 100 matrices of each dimension n. A distribution-free method gave confidence intervals for the medians [3]. From this data we make the following observations:

(1) For small $n(n<20)$, ν_1 produced a better estimate than the LINPACK estimate in over 50% of the trials. For larger $n(n=50)$ the estimate μ_1 was better approximately 80% of the time.

(2) The estimate ρ_1, the maximum of these estimates, was a noticeable improvement over both ν_1 and μ_1 for small n.

(3) For each set of trials, the estimate ν_1 had a higher maximum than μ_1 (except on symmetric matrices of dimension 50), and a lower minimum (except on general matrices of dimension 5). For small matrices ν_1 was often exact ($\nu_1/\|A^{-1}\|_1=1$ in 42 trials out of 100 for $n=5$) and for such matrices, μ_1 was often 30% smaller.

(4) The first estimate of the infinity norm, ν_∞, was unreliable and the second was almost always better.

TABLE 1

Results for matrices with elements from a uniform distribution[1]

GENERAL MATRICES

n	Average $\nu_1/\|A^{-1}\|_1$	Average $\mu_1/\|A^{-1}\|_1$	Average $\rho_1/\|A^{-1}\|_1$	% of Trials $\nu_1>\mu_1$	Average $\nu_\infty/\|A^{-1}\|_\infty$	Average $\mu_\infty/\|A^{-1}\|_\infty$	Average $\rho_\infty/\|A^{-1}\|_\infty$	% of Trials $\nu_\infty>\mu_\infty$
5	.84	.69	.86	80	.45	.66	.67	3
10	.70	.60	.74	70	.31	.60	.60	0
20	.49	.52	.57	48	.19	.53	.53	0
30	.43	.48	.52	40	.15	.48	.48	0
40	.31	.43	.45	21	.11	.44	.44	0
50	.31	.45	.46	19	.10	.45	.45	0

SYMMETRIC MATRICES

n	Average $\nu_1/\|A^{-1}\|_1$	Average $\mu_1/\|A^{-1}\|_1$	Average $\rho_1/\|A^{-1}\|_1$	% of Trials $\nu_1>\mu_1$	Average $\nu_\infty/\|A^{-1}\|_\infty$	Average $\mu_\infty/\|A^{-1}\|_\infty$	Average $\rho_\infty/\|A^{-1}\|_\infty$	% of Trials $\nu_\infty>\mu_\infty$
5	.83	.66	.83	88	.46	.66	.66	2
10	.66	.58	.71	69	.29	.55	.55	0
20	.53	.50	.59	58	.20	.49	.49	0
30	.38	.44	.48	33	.14	.44	.44	0
40	.35	.45	.47	31	.12	.44	.44	0
50	.30	.40	.42	27	.10	.41	.41	0

[1]*Note.* As part of the refereeing process for this paper, the results reported in Table 1 were substantiated through similar tests by Curtis Hunt at the University of New Mexico.

TABLE 2
99% confidence intervals for the medians for trials on general matrices

n	$\nu_1/\|A^{-1}\|_1$	$\mu_1/\|A^{-1}\|_1$	$\rho_1/\|A^{-1}\|_1$
5	.80 – 1.00	.67 – .73	.83 – 1.00
10	.60 – .80	.57 – .65	.67 – .80
20	.42 – .55	.50 – .57	.54 – .61
30	.33 – .50	.46 – .52	.48 – .56
40	.23 – .34	.41 – .49	.41 – .50
50	.23 – .33	.43 – .49	.44 – .50

(5) The infinity norm estimates had consistently larger relative error than the one-norm estimates.
(6) Results for matrices with elements randomly 1, 0, or -1 were similar to those tabulated.

3. Conclusion. An inexpensive algorithm to estimate the one-norm of A^{-1} using an LU factorization of A, is:
(a) Solve $A^Tx=e$ where e is chosen as described above. Let $\nu_1=\|x\|_\infty$.
(b) Solve $Ay=x$ and let $\mu_1=\|y\|_1/\|x\|_1$.
(c) Estimate $\|A^{-1}\|_1$ by $\rho_1=\max(\nu_1,\mu_1)$.
This costs only n comparisons more than the LINPACK algorithm and gives more reliable results for small matrices.

To estimate the infinity norm of A^{-1}, the above algorithm should be applied to the matrix A^T. This gives better results than the procedure formed by interchanging the roles of the one and infinity norms in the three steps above.

Acknowledgment. This work benefited from helpful comments of G. W. Stewart.

REFERENCES

[1] A. K. CLINE, C. B. MOLER, G. W. STEWART, AND J. H. WILKINSON, *An estimate for the condition number of a matrix*, SIAM J. Numer. Anal., 16 (1979), pp. 368–375.
[2] J. J. DONGARRA, C. B. MOLER, J. R. BUNCH, AND G. W. STEWART, *LINPACK Users' Guide*, Society for Industrial and Applied Mathematics, Philadelphia, 1979.
[3] WILLIAM J. MACKINNON, *Table for both the sign test and distribution-free confidence intervals of the median for sample sizes to* 1000, J. Amer. Statist. Assoc. (1964), pp. 935–956.