

# How to Determine the Accuracy of the Output of a Matrix Inversion Program

Morris Newman

Institute for Basic Standards, National Bureau of Standards, Washington, D.C. 20234

(January 3, 1974)

If  $A$  is a given nonsingular matrix and  $X$  an approximate inverse of  $A$  such that  $N(R) < 1$ , where  $R = I - AX$  and  $N$  is any matrix norm, then it is shown that

$$\frac{N(XR)}{1+N(R)} \leq N(A^{-1}-X) \leq \frac{N(XR)}{1-N(R)}.$$

This inequality provides the means for checking the output of a matrix inversion program. Methods for checking the solution of a single linear system are also discussed.

Key words: Approximate inverses; approximate solutions; error bounds; high speed digital computation; matrix norms.

## 1. Introduction

Although numerous excellent programs exist for the inversion of a matrix  $A$  or the solution of a linear system  $Ax = b$ , not enough attention has been given to the question of guaranteeing the accuracy of the output of such a program. In fact the problem of certifying the results of computation in general is a basic one and has been largely neglected, although it should be of pressing concern to any user of a high speed digital computer.

The purpose of this article (which is mainly expository) is to point out that for matrix inversion there is a satisfactory solution to this question, and to derive the necessary estimates. It is the author's hope that this article will influence users of large machines to incorporate the procedures suggested (or some version of them) into their programs for matrix inversion.

## 2. Matrix Norms

The principal tool for providing such error estimates is the matrix norm. A good theoretical discussion of this topic may be found for example in Householder's book [3],<sup>1</sup> chapter 2. We depart somewhat from the usual definition in what follows. For our purposes a matrix norm  $N$  will be a real-valued function defined for all rectangular matrices over the complex numbers  $C$  such that for any matrices  $A, B$

- (1)  $N(A) \geq 0$ ;  $N(A) = 0$  if and only if  $A = 0$ ,
- (2)  $N(A+B) \leq N(A) + N(B)$ , provided that  $A$  and  $B$  may be added,

AMS Subject Classification: 65F35; 15A60.

<sup>1</sup> Figures in brackets indicate the literature references at the end of this paper.

(3)  $N(cA) = |c|N(A)$ , where  $c$  is any complex number,

(4)  $N(AB) \leq N(A)N(B)$ , provided that  $A$  and  $B$  may be multiplied in the order shown.

Properties (1), (2), (3) are the usual properties of a distance function, but (4) is peculiar to matrices.

Let  $A = (a_{ij})$  be any  $m \times n$  matrix over  $C$ . Then the most important examples of norms are

$$(5) \quad N(A) = F(A) = \left\{ \sum_{i,j} |a_{ij}|^2 \right\}^{1/2} \text{ (the Euclidean or Frobenius norm),}$$

$$(6) \quad N(A) = M(A) = \sqrt{mn} \max_{i,j} |a_{ij}| \text{ (the maximum element norm),}$$

$$(7) \quad N(A) = S(A) = \max_i \sum_j |a_{ij}| \text{ (the maximum row sum norm).}$$

It is a simple exercise to show that these are indeed norms. The reader may wish to consult [4] where it is proved that (5) satisfies property (4).

The Euclidean norm has the property that it is a unitary invariant; that is, if  $U, V$  are unitary matrices of respective orders  $m, n$ , then

$$F(UAV) = F(A).$$

This is most easily seen from the observation that

$$F(A)^2 = \text{tr}(AA^*),$$

where  $A^*$  is the conjugate transpose of  $A$ .

This may be used to deduce that if  $A$  is square of order  $n$  and has eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_n$ , then

$$(8) \quad |\lambda_1|^2 + |\lambda_2|^2 + \dots + |\lambda_n|^2 \leq F(A)^2,$$

with equality if and only if  $A$  is normal (that is, if and only if  $A$  is unitarily equivalent to a diagonal matrix). This happens if and only if  $AA^* = A^*A$ . Inequality (8) is known as Schur's inequality.

A consequence of (8) is that if  $\lambda$  is any eigenvalue of  $A$ , then  $|\lambda| \leq F(A)$ . In fact any matrix norm has this useful property, which we state as a theorem:

**THEOREM 1:** *Let  $A$  be a square matrix with arbitrary complex entries. Let  $\lambda$  be any eigenvalue of  $A$ . Then if  $N$  is any matrix norm,*

$$(9) \quad |\lambda| \leq N(A).$$

**PROOF:** Let  $x$  be a nonzero eigenvector corresponding to  $\lambda$ , so that  $Ax = \lambda x$ . Then

$$\begin{aligned} |\lambda|N(x) &= N(\lambda x) \\ &= N(Ax) \\ &\leq N(A)N(x), \end{aligned}$$

so that

$$|\lambda| \leq N(A),$$

since  $N(x) \neq 0$ . This completes the proof.

This result provides a useful criterion for the convergence of a certain power series. Let  $A$  be square. It is well-known that the series  $I + A + A^2 + \dots$  converges if and only if every eigenvalue of  $A$  is of modulus less than 1, in which case  $I - A$  is nonsingular and the sum of the series is  $(I - A)^{-1}$  (a proof may be found in [4]). Thus Theorem 1 has the useful corollary that if  $N(A) < 1$ , where  $N$  is any matrix norm, then  $I - A$  is nonsingular and the series  $I + A + A^2 + \dots$  converges to  $(I - A)^{-1}$ . Since the convergence of this series underlies many iterative schemes used in matrix computations, the corollary is a valuable one.

### 3. Error Bounds for Approximate Inverses

Now suppose that  $A$  is a square nonsingular matrix, and that some computational scheme has been used to determine an approximate inverse  $X$  of  $A$ . Put  $R = I - AX$ . The fact that  $R$  is close to the 0 matrix does not necessarily imply that  $X$  is close to  $A^{-1}$ . A simple example which illustrates this phenomenon is furnished by choosing

$$A = \begin{pmatrix} 1 & 1 - n^{-2} \\ 1 & 1 \end{pmatrix}, \quad X = \begin{pmatrix} n^2 + n & 1 - n^2 \\ -n^2 - n & n^2 \end{pmatrix}$$

where  $n$  is large. Then

$$R = I - AX = \begin{pmatrix} -n^{-1} & 0 \\ 0 & 0 \end{pmatrix}, \quad A^{-1} - X = \begin{pmatrix} -n & 0 \\ n & 0 \end{pmatrix},$$

so that  $R$  is arbitrarily close to the 0 matrix although  $X$  is arbitrarily far from  $A^{-1}$ .

We are entitled to assume that  $R$  (which ideally should be 0) satisfies  $N(R) < 1$  for some matrix norm  $N$ , since otherwise  $X$  cannot even be considered an approximation to  $A^{-1}$ . We now prove the following, which is the main result of this article:

**THEOREM 2:** *Let  $A$  be a square nonsingular matrix and  $X$  an approximate inverse of  $A$  such that  $N(R) < 1$ , where  $R = I - AX$  and  $N$  is some matrix norm. Then*

$$(10) \quad \frac{N(XR)}{1 + N(R)} \leq N(A^{-1} - X) \leq \frac{N(XR)}{1 - N(R)} \leq \frac{N(X)N(R)}{1 - N(R)}.$$

**PROOF:** Since  $N(R) < 1$ ,  $I - R$  is nonsingular and the power series  $I + R + R^2 + \dots$  converges to  $(I - R)^{-1}$ . Thus we have

$$\begin{aligned} AX &= I - R, \\ X^{-1}A^{-1} &= (I - R)^{-1} = I + R + R^2 + \dots \end{aligned}$$

Hence

$$\begin{aligned} A^{-1} - X &= XR + XR^2 + \dots, \\ N(A^{-1} - X) &\leq N(XR) + N(XR^2) + \dots \\ &\leq N(XR) + N(XR)N(R) + \dots \\ &= \frac{N(XR)}{1 - N(R)}. \end{aligned}$$

Thus since  $N(XR) \leq N(X)N(R)$ , the upper bounds given by (10) are valid.

For the lower bound, we have

$$\begin{aligned}
R &= I - AX, \\
XR &= X - XAX \\
&= (A^{-1} - X) AX \\
&= (A^{-1} - X)(I - R) \\
&= (A^{-1} - X) - (A^{-1} - X)R,
\end{aligned}$$

so that

$$\begin{aligned}
N(XR) &= N\{(A^{-1} - X) - (A^{-1} - X)R\} \\
&\leq N(A^{-1} - X) + N\{(A^{-1} - X)R\} \\
&\leq N(A^{-1} - X) + N(A^{-1} - X)N(R) \\
&= N(A^{-1} - X) \{1 + N(R)\}, \\
\frac{N(XR)}{1 + N(R)} &\leq N(A^{-1} - X).
\end{aligned}$$

This completes the proof.

It follows that for all practical purposes  $N(A^{-1} - X)$  may be replaced by  $N(XR)$ .

Results elaborating on Theorem 2, and applications to the evaluation of programs for matrix inversion, may be found in [1] and [2].

Theorem 2 provides the theoretical basis for inserting absolute error checks into any matrix inversion program, using only the information supplied to the computer and the output of the program. In programs which incorporate an "iterative improvement" feature there is no time loss at all in doing so, since the new approximant  $X'$  is computed from the odd approximant  $X$  and the residual  $R = I - AX$  by means of the formula

$$X' = X + XR,$$

so that  $XR$  must be calculated in any case. For matrices of moderate order, say at most 50, the time involved in computing  $XR$  is negligible and the sharp error bound should be used. For matrices of large order one multiplication may be saved by computing only  $R$  and using the weaker upper bound. We note parenthetically that the computation of  $XR$  must generally be carried out in double or higher precision, if it is to be used for iterative improvement.

#### 4. Error Bounds for Approximate Solutions

The corresponding question for the solution of a linear system  $Ax = b$ , where the output is a single vector which is an approximate solution of the system, does not seem capable of being answered so simply and completely. However, a very good relative check exists in this situation which is easily incorporated into any program to find  $x = A^{-1}b$ .

Let  $\epsilon$  be the vector  $(1, 1, \dots, 1)^T$ , so that  $r = A\epsilon$  is the vector whose coordinates are the row sums of  $A$  (and so available from the data). Adjoin  $r$  to the system and solve  $A(x, y) = (b, r)$ , the exact solution of which is  $(x, y) = (A^{-1}b, A^{-1}r) = (A^{-1}b, \epsilon)$ . Then the maximum deviation from 1 in the coordinates of the computed vector  $y$  is an excellent indication of the maximum deviation from  $A^{-1}b$  in the computed vector  $x$ . This checking procedure is both simple and effective and is readily incorporated into any program which finds  $x = A^{-1}b$ .

#### 5. References

- [1] Fitzgerald, K. E., Error estimates for the solution of linear algebraic systems. J. Res. Nat. Bur. Stand. (U.S.), **74B** (Math. Sci.), No. 4, 251-310 (Oct.-Dec. 1970).
- [2] Fitzgerald, K. E., Comparison of some FORTRAN programs for matrix inversion, J. Res. Nat. Bur. Stand. (U.S.), **78B** (Math. Sci.) No. 1, (Jan.-Mar. 1974).
- [3] Householder, A. S., The theory of matrices in numerical analysis. (Blaisdell, New York, 1964).
- [4] Newman, M., Matrix computations, Survey of numerical analysis, J. Todd, editor, (McGraw-Hill, New York, 1962).

(Paper 78B2-399)