

# A Comparison of Several Methods for Inverting Large Symmetric Positive Definite Matrices

By M. H. Lietzke, R. W. Stoughton and Marjorie P. Lietzke

**1. Introduction.** In any least squares fit involving the determination of a large number of parameters the accumulation of round-off error during the course of the computation, resulting in a loss of significance, is one of the most serious restrictions. Since the evaluation of a large number of parameters and their variances by a least squares procedure on a high speed computer involves the inversion of a symmetric, positive-definite matrix it becomes important to choose an inversion scheme in which the effects of the accumulation of round-off error are minimized. This is especially true if limited memory space for storage precludes the use of double precision arithmetic. In this paper a comparison of several direct methods for inverting such matrices is given. An attempt has been made to consider the effects of both condition and order of the matrix to be inverted on the closeness of the computed inverse to the exact inverse.

The matrix inversion methods compared are the Gauss-Jordan [1], Choleski [2], congruent transformation [3], and rank annihilation [4] schemes. To give a fair comparison of the methods each was programmed in IBM 7090 FORTRAN II (Version 2) using only single precision arithmetic (good to about 8 decimal digits). The error indicators were then computed using double precision arithmetic (good to about 16 decimal digits) so that the latter calculation was not a limiting factor. In the following discussion the symbols for matrices will be **underlined** while other symbols will not be.

**2. The Test Matrices.** The first matrix chosen for testing the inversion routines was  $\mathbf{A}^*$  where the elements  $a_{ij}$  of  $\mathbf{A}$  are given by

$$a_{ij} = \begin{cases} 2 & (i = j), \\ -1 & (|i - j| = 1), \\ 0 & (|i - j| > 1). \end{cases}$$

The inverse of  $\mathbf{A}$  is given by  $\mathbf{C}/(n + 1)$  where the elements  $c_{ij}$  of  $\mathbf{C}$  are given by

$$c_{ij} = \begin{cases} i(n - i + 1) & (i = j), \\ c_{i,j-1} - i & (j > i), \\ c_{ji} = c_{ij} & (j < i), \end{cases}$$

and  $n$  is the order of the matrix.  $\mathbf{A}$  is a symmetric positive definite matrix; it has a  $P$ -condition number [5] of approximately  $4n^2/\pi^2$ . In respect to both form and condition the matrix  $\mathbf{A}$  is analogous to a least squares matrix derived from a well designed experiment. In addition to  $\mathbf{A}$  the matrices  $\mathbf{A}^2$  and  $\mathbf{A}^3$  were also inverted. The inverse of  $\mathbf{A}^2$  is given by  $\mathbf{C}^2/(n + 1)^2$ , while the inverse of  $\mathbf{A}^3$  is given by  $\mathbf{C}^3/(n + 1)^3$ .

---

Received August 15, 1963. This work was performed at the Oak Ridge National Laboratory and the Central Data Processing Facility, operated by Union Carbide Corporation for the U. S. Atomic Energy Commission.

\* The negative of this matrix was discussed by Newman and Todd [5].

The latter two matrices have  $P$ -condition numbers of  $6n^4/\pi^4$  and  $64n^6/\pi^6$  respectively and are hence progressively more ill-conditioned. In this respect they correspond to least squares matrices derived from increasingly more poorly designed experiments.

In addition to the three matrices just mentioned two other test matrices were used. They are **B**, whose elements  $b_{ij}$  are given by

$$b_{ij} = \begin{cases} 2 & (i = j), \\ 1 & (i \neq j), \end{cases}$$

and **D**, whose elements  $d_{ij}$  are given by

$$d_{ij} = n - |i - j|.$$

The exact inverse of **B** is given by **BINV** whose elements  $binv_{ij}$  are

$$binv_{ij} = \begin{cases} n/(n + 1) & \text{if } i = j, \\ -1/(n + 1) & \text{if } i \neq j, \end{cases}$$

while the exact inverse of **D** is given by **DINV** whose elements  $dinv_{ij}$  are

$$dinv_{ij} = \begin{cases} (n + 2)/(2n + 2) & \text{if } i = j = 1 \text{ or } n, \\ 1 & \text{if } i = j \neq 1 \text{ or } n, \\ -\frac{1}{2} & \text{if } |i - j| = 1 \text{ and } n \neq 2, \\ -\frac{1}{3} & \text{if } |i - j| = 1 \text{ and } n = 2, \\ 1/(2n + 2) & \text{if } |i - j| = n - 1 = 1, \\ 0 & \text{if } 1 < |i - j| < n - 1. \end{cases}$$

**3. The Error Indicators.** In carrying out the comparison of the various inversion methods the test matrix was generated and inverted by each method to give  $(\mathbf{M})_{\text{approx}}^{-1}$ . Then the exact inverse was generated (using the above formulas and double precision arithmetic) and as a measure of error the quantity  $Q$  defined as

$$(3.1) \quad Q = (1/n^2) \sum_{ij} |(\mathbf{M})_{\text{exact}}^{-1} - (\mathbf{M})_{\text{approx}}^{-1}|_{ij}$$

was computed for each method, where **M** represents the matrix inverted. The computations were performed with matrices of order 10, 15, 20, 25, and 30. In addition to  $Q$  two other quantities, recommended as error measures by Newman and Todd [5],

$$(3.2) \quad a = (1/n^2) \sum_{ij} |r_{ij}|$$

and

$$(3.3) \quad f = (1/n) \left\{ \sum_{ij} r_{ij}^2 \right\}^{1/2}$$

were calculated, where **R** =  $(r_{ij})$  is the error matrix taken as

$$(3.4) \quad R = (\mathbf{M})_{\text{approx}}^{-1}(\mathbf{M}) - \mathbf{I}$$

where **I** is the identity matrix. The number  $Q$  obviously reflects the per term average of the overall difference between the approximate inverse matrix and the exact inverse matrix.

**4. Results of the Computations. The A Matrix.** All the methods tried gave an approximate inverse that was very close to the exact inverse of the **A** matrix for all orders from 10 to 30 except the rank annihilation method, in which case the routine used failed to invert the **A** matrix. The reason for this failure was evidently peculiar to the exact integers in the matrix and the particular computer procedure for the rank annihilation method; the procedure could undoubtedly be rewritten to obviate this particular failure although it was not considered warranted in the current comparative study. Hence any one of the methods would be suitable for a well-conditioned least squares matrix in this range. All elements of the inverse matrix produced by the congruent transformation method were good to one in the eighth significant figure even for order 30. The Gauss-Jordan method and the Choleski method were close behind in that order. Not only did the congruent transformation method produce an inverse closest to the exact inverse of the **A** matrix but the values of *a* and *f* (as defined by Eqs. (3.2) and (3.3)) computed from the approximate inverse were lower than the corresponding values derived from the Gauss-Jordan and Choleski inversion routines as can be seen in Tables IA-ID. All the methods produced inverse matrices, all elements of which were smaller in magnitude than the exact inverse for all orders from 10 to 30. (The reason for this is unknown.)

*The A<sup>2</sup> Matrix.* In the inversion of the **A<sup>2</sup>** matrix the effect of the condition of the matrix starts to become apparent. This matrix is not as well conditioned as the **A** matrix; in fact, its condition number is the square of the condition number for the **A** matrix and hence is proportional to the fourth power of the order.

In the inversion of the **A<sup>2</sup>** matrix for all orders from 10 to 30 the Gauss-Jordan method produced an inverse matrix closest to the exact inverse. The congruent transformation, the Choleski, and the rank annihilation methods gave very similar results and were somewhat inferior to the Gauss-Jordan method. All elements of

TABLE IA  
A Comparison of the *a*, *f*, and *Q* Values as a Function of Matrix Order and Inversion Method  
Gauss Jordan Method

<i>n</i>		Matrix				
		<b>A</b>	<b>A<sup>2</sup></b>	<b>A<sup>3</sup></b>	<b>B</b>	<b>D</b>
10	<i>a</i>	$1.3 \times 10^{-8}$	$9.6 \times 10^{-6}$	$6.7 \times 10^{-4}$	$2.2 \times 10^{-8}$	$7.3 \times 10^{-8}$
	<i>f</i>	$1.9 \times 10^{-8}$	$1.4 \times 10^{-5}$	$1.1 \times 10^{-3}$	$2.7 \times 10^{-8}$	$9.1 \times 10^{-8}$
	<i>Q</i>	$4.8 \times 10^{-8}$	$2.7 \times 10^{-5}$	$2.1 \times 10^{-3}$	$5.5 \times 10^{-9}$	$3.3 \times 10^{-8}$
15	<i>a</i>	$2.3 \times 10^{-8}$	$3.6 \times 10^{-5}$	$3.0 \times 10^{-2}$	$3.0 \times 10^{-8}$	$1.5 \times 10^{-7}$
	<i>f</i>	$3.6 \times 10^{-8}$	$7.1 \times 10^{-5}$	$4.7 \times 10^{-2}$	$3.8 \times 10^{-8}$	$1.9 \times 10^{-7}$
	<i>Q</i>	$7.8 \times 10^{-8}$	$8.3 \times 10^{-5}$	$2.6 \times 10^{-2}$	$6.4 \times 10^{-9}$	$4.8 \times 10^{-8}$
20	<i>a</i>	$3.7 \times 10^{-8}$	$1.1 \times 10^{-4}$	$5.0 \times 10^{-1}$	$4.0 \times 10^{-8}$	$2.2 \times 10^{-7}$
	<i>f</i>	$6.1 \times 10^{-8}$	$1.6 \times 10^{-4}$	$8.1 \times 10^{-1}$	$5.3 \times 10^{-8}$	$2.8 \times 10^{-7}$
	<i>Q</i>	$1.2 \times 10^{-7}$	$1.1 \times 10^{-3}$	3.4	$6.6 \times 10^{-9}$	$7.4 \times 10^{-7}$
25	<i>a</i>	$4.3 \times 10^{-8}$	$4.3 \times 10^{-4}$	3.4	$5.0 \times 10^{-8}$	$3.7 \times 10^{-7}$
	<i>f</i>	$7.1 \times 10^{-8}$	$7.2 \times 10^{-4}$	5.8	$6.4 \times 10^{-8}$	$4.8 \times 10^{-7}$
	<i>Q</i>	$2.0 \times 10^{-7}$	$5.2 \times 10^{-3}$	27	$6.6 \times 10^{-9}$	$9.2 \times 10^{-8}$
30	<i>a</i>	$5.1 \times 10^{-8}$	$1.3 \times 10^{-3}$	13	$6.1 \times 10^{-8}$	$5.3 \times 10^{-7}$
	<i>f</i>	$8.2 \times 10^{-8}$	$2.0 \times 10^{-3}$	23.7	$7.8 \times 10^{-8}$	$6.9 \times 10^{-7}$
	<i>Q</i>	$3.9 \times 10^{-7}$	$1.9 \times 10^{-2}$	69	$6.5 \times 10^{-9}$	$1.2 \times 10^{-7}$

TABLE IB

*A Comparison of the a, f, and Q Values as a Function of Matrix Order and Inversion Method  
Congruent Transformation Method*

n		Matrix				
		A	A <sup>2</sup>	A <sup>3</sup>	B	D
10	a	1.0 × 10 <sup>-8</sup>	1.2 × 10 <sup>-6</sup>	5.6 × 10 <sup>-5</sup>	1.5 × 10 <sup>-8</sup>	6.8 × 10 <sup>-8</sup>
	f	1.5 × 10 <sup>-8</sup>	1.4 × 10 <sup>-6</sup>	7.2 × 10 <sup>-5</sup>	2.5 × 10 <sup>-8</sup>	1.1 × 10 <sup>-7</sup>
	Q	5.9 × 10 <sup>-9</sup>	1.0 × 10 <sup>-4</sup>	1.3 × 10 <sup>-2</sup>	7.9 × 10 <sup>-9</sup>	2.6 × 10 <sup>-8</sup>
15	a	0.0	4.3 × 10 <sup>-6</sup>	5.2 × 10 <sup>-4</sup>	1.8 × 10 <sup>-8</sup>	1.1 × 10 <sup>-7</sup>
	f	0.0	5.6 × 10 <sup>-6</sup>	7.1 × 10 <sup>-4</sup>	3.0 × 10 <sup>-8</sup>	1.7 × 10 <sup>-7</sup>
	Q	0.0	1.2 × 10 <sup>-3</sup>	1.3	9.2 × 10 <sup>-9</sup>	2.9 × 10 <sup>-8</sup>
20	a	1.6 × 10 <sup>-8</sup>	9.1 × 10 <sup>-6</sup>	2.3 × 10 <sup>-3</sup>	2.1 × 10 <sup>-8</sup>	1.4 × 10 <sup>-7</sup>
	f	2.6 × 10 <sup>-8</sup>	1.2 × 10 <sup>-5</sup>	3.2 × 10 <sup>-3</sup>	3.9 × 10 <sup>-8</sup>	2.5 × 10 <sup>-7</sup>
	Q	1.0 × 10 <sup>-8</sup>	7.4 × 10 <sup>-3</sup>	29	1.0 × 10 <sup>-8</sup>	5.4 × 10 <sup>-8</sup>
25	a	1.8 × 10 <sup>-8</sup>	1.5 × 10 <sup>-5</sup>	5.7 × 10 <sup>-3</sup>	2.3 × 10 <sup>-8</sup>	1.7 × 10 <sup>-7</sup>
	f	3.1 × 10 <sup>-8</sup>	2.0 × 10 <sup>-5</sup>	7.5 × 10 <sup>-3</sup>	4.3 × 10 <sup>-8</sup>	3.1 × 10 <sup>-7</sup>
	Q	6.2 × 10 <sup>-8</sup>	2.7 × 10 <sup>-2</sup>	229	1.0 × 10 <sup>-8</sup>	6.4 × 10 <sup>-8</sup>
30	a	1.9 × 10 <sup>-8</sup>	2.9 × 10 <sup>-5</sup>	1.5 × 10 <sup>-2</sup>	2.4 × 10 <sup>-8</sup>	2.1 × 10 <sup>-7</sup>
	f	3.2 × 10 <sup>-8</sup>	4.0 × 10 <sup>-5</sup>	2.3 × 10 <sup>-2</sup>	4.6 × 10 <sup>-8</sup>	4.3 × 10 <sup>-7</sup>
	Q	2.2 × 10 <sup>-8</sup>	1.1 × 10 <sup>-1</sup>	1885	1.0 × 10 <sup>-8</sup>	7.3 × 10 <sup>-8</sup>

TABLE IC

*A Comparison of the a, f, and Q Values as a Function of Matrix Order and Inversion Method  
Choleski Method*

n		Matrix				
		A	A <sup>2</sup>	A <sup>3</sup>	B	D
10	a	2.0 × 10 <sup>-8</sup>	1.6 × 10 <sup>-6</sup>	8.7 × 10 <sup>-5</sup>	2.7 × 10 <sup>-8</sup>	8.0 × 10 <sup>-8</sup>
	f	3.1 × 10 <sup>-8</sup>	2.2 × 10 <sup>-6</sup>	1.1 × 10 <sup>-4</sup>	3.7 × 10 <sup>-8</sup>	1.0 × 10 <sup>-7</sup>
	Q	1.3 × 10 <sup>-7</sup>	1.1 × 10 <sup>-4</sup>	2.1 × 10 <sup>-2</sup>	7.8 × 10 <sup>-9</sup>	4.5 × 10 <sup>-8</sup>
15	a	4.0 × 10 <sup>-8</sup>	8.9 × 10 <sup>-6</sup>	1.6 × 10 <sup>-3</sup>	3.7 × 10 <sup>-8</sup>	1.6 × 10 <sup>-7</sup>
	f	5.8 × 10 <sup>-8</sup>	1.4 × 10 <sup>-5</sup>	2.2 × 10 <sup>-3</sup>	4.8 × 10 <sup>-8</sup>	2.2 × 10 <sup>-7</sup>
	Q	2.4 × 10 <sup>-7</sup>	1.7 × 10 <sup>-3</sup>	1.3 × 10 <sup>-1</sup>	9.2 × 10 <sup>-9</sup>	4.6 × 10 <sup>-8</sup>
20	a	5.7 × 10 <sup>-8</sup>	1.8 × 10 <sup>-5</sup>	1.4 × 10 <sup>-2</sup>	4.3 × 10 <sup>-8</sup>	3.3 × 10 <sup>-7</sup>
	f	8.8 × 10 <sup>-8</sup>	2.7 × 10 <sup>-5</sup>	1.9 × 10 <sup>-2</sup>	5.8 × 10 <sup>-8</sup>	4.5 × 10 <sup>-7</sup>
	Q	2.4 × 10 <sup>-7</sup>	1.1 × 10 <sup>-2</sup>	8.3	9.9 × 10 <sup>-9</sup>	8.4 × 10 <sup>-8</sup>
25	a	7.4 × 10 <sup>-8</sup>	4.0 × 10 <sup>-5</sup>	6.4 × 10 <sup>-2</sup>	5.2 × 10 <sup>-8</sup>	5.9 × 10 <sup>-7</sup>
	f	1.2 × 10 <sup>-7</sup>	5.5 × 10 <sup>-5</sup>	9.6 × 10 <sup>-2</sup>	6.8 × 10 <sup>-8</sup>	7.9 × 10 <sup>-7</sup>
	Q	3.7 × 10 <sup>-7</sup>	4.5 × 10 <sup>-2</sup>	130	1.2 × 10 <sup>-8</sup>	1.3 × 10 <sup>-7</sup>
30	a	8.9 × 10 <sup>-8</sup>	1.2 × 10 <sup>-4</sup>	2.2 × 10 <sup>-1</sup>	6.1 × 10 <sup>-8</sup>	9.6 × 10 <sup>-7</sup>
	f	1.4 × 10 <sup>-7</sup>	1.7 × 10 <sup>-4</sup>	3.7 × 10 <sup>-1</sup>	7.9 × 10 <sup>-8</sup>	1.2 × 10 <sup>-6</sup>
	Q	6.4 × 10 <sup>-7</sup>	1.4 × 10 <sup>-1</sup>	751	1.5 × 10 <sup>-8</sup>	1.6 × 10 <sup>-7</sup>

the inverse matrix produced by the latter method were good to one part in 600,000 for order 10 to one part in 20,000 for order 30. The corresponding figures for the other methods were one part in about 150,000 for  $n = 10$  to one part in about 3000 for  $n = 30$ . The values of  $Q$ ,  $a$ , and  $f$  for the  $A^2$  matrix are shown in Table I.

*The A<sup>3</sup> Matrix.* An examination of the values of  $Q$  for the  $A^3$  matrix (Tables IA-ID) indicates that the Gauss-Jordan routine produced an inverse matrix closest

TABLE ID  
*A Comparison of the a, f, and Q Values as a Function of Matrix Order and Inversion Method Rank Annihilation Method*

n		Matrix		
		A <sup>2</sup>	A <sup>3</sup>	B
10	a	2.2 × 10 <sup>-5</sup>	1.5 × 10 <sup>-3</sup>	1.3 × 10 <sup>-8</sup>
	f	2.6 × 10 <sup>-5</sup>	2.0 × 10 <sup>-3</sup>	1.6 × 10 <sup>-8</sup>
	Q	8.8 × 10 <sup>-5</sup>	2.6 × 10 <sup>-1</sup>	0.0
15	a	4.2 × 10 <sup>-5</sup>	6.4 × 10 <sup>-3</sup>	0.0
	f	5.9 × 10 <sup>-5</sup>	1.1 × 10 <sup>-2</sup>	0.0
	Q	1.5 × 10 <sup>-3</sup>	9.1	0.0
20	a	7.5 × 10 <sup>-5</sup>	5.2 × 10 <sup>-2</sup>	3.1 × 10 <sup>-8</sup>
	f	1.0 × 10 <sup>-4</sup>	9.4 × 10 <sup>-2</sup>	4.2 × 10 <sup>-8</sup>
	Q	1.2 × 10 <sup>-2</sup>	97	0.0
25	a	1.9 × 10 <sup>-4</sup>	2.6 × 10 <sup>-1</sup>	3.6 × 10 <sup>-8</sup>
	f	2.7 × 10 <sup>-4</sup>	4.7 × 10 <sup>-1</sup>	5.0 × 10 <sup>-8</sup>
	Q	3.9 × 10 <sup>-2</sup>	440	0.0
30	a	4.0 × 10 <sup>-4</sup>	9.1 × 10 <sup>-1</sup>	4.8 × 10 <sup>-8</sup>
	f	6.4 × 10 <sup>-4</sup>	1.4	7.1 × 10 <sup>-8</sup>
	Q	9.0 × 10 <sup>-2</sup>	1386	0.0

to the exact inverse for all orders from 10 to 30. All elements of the inverse matrix produced by this routine are good to one part in 80,000 for  $n = 10$ , to one part in 1100 for  $n = 20$  and one part in 400 for  $n = 30$ . Close behind the Gauss-Jordan routine for  $n = 10$  is the congruent transformation method; however, for  $n = 15$  to 30 the Choleski method produced a more nearly exact inverse. At  $n = 30$  the rank annihilation method appears slightly better than the congruent transformation method, producing an inverse each element of which is good to one part in 16.

In addition to the above observations concerning the closeness of the approximate inverse of the  $A^3$  matrix to the exact inverse it should be noted that the Gauss-Jordan and congruent transformation routines produced inverse matrices whose elements were in all cases smaller in magnitude than those of the exact inverse, while the inverse matrices produced by the rank annihilation method were for all orders greater than the exact inverse. The Choleski routine produced an inverse matrix at orders 10 and 15 with elements smaller than those of the exact inverse, while at orders 20, 25 and 30 the inverse matrix elements were greater than those of the exact inverse. The low value of  $Q$  at  $n = 15$  for the Choleski method reflects the fact that the elements of the approximate inverse are passing through a transition from all being too small to all being too large.

*The B Matrix.* The B matrix is apparently so well conditioned that all four methods produced inverse matrices very close to the true inverse as can be seen in Table I. In fact the inverse matrix produced by the rank annihilation routine varied at most by one in the eighth significant figure from the exact inverse, and the values of  $Q$  computed were in all cases 0.0 to eight decimal digits.

*The D Matrix.* The D matrix is also well conditioned and all the methods except the rank annihilation method produced inverse matrices close to the exact inverse. The rank annihilation routine failed to invert properly the D matrix, undoubtedly

because of the peculiar form of this matrix and the peculiarities of the computer program as written. The values of  $a$ ,  $f$  and  $Q$  for the  $\mathbf{D}$  matrix are shown in Tables IA-ID.

**5. A Comparison of the Indices of Error  $a$ ,  $f$  and  $Q$ .** The following observations may be made concerning the results presented in Tables IA-ID:

(1) The value of  $a$  is in all cases lower than the value of  $f$ .

(2) The values of  $a$  and  $f$  are not, by themselves, reliable estimates of the closeness of the computed inverse to the exact inverse. For example, if the  $a$  and  $f$  values alone were used to compare the approximate inverses of the  $\mathbf{A}^3$  matrix produced by the various inversion methods then the congruent transformation method would appear to be as good as the Gauss-Jordan method. However, the Gauss-Jordan routine produced an inverse for  $n = 30$  good to one part in 400 while the inverse produced by the congruent transformation method was good only to one part in 15. Thus the  $Q$  values (which follow this pattern) reflect the observed relation between the exact and approximate inverse matrices while the  $a$  and  $f$  values may not. A possible explanation in the latter case may be in the fact that the  $\mathbf{A}^3$  matrix contains elements of alternating sign while all the elements of the inverse

TABLE II  
Values of the Two Sides of the Inequality  
 $\| \mathbf{A}^{-1} - \mathbf{C} \| \leq \| \mathbf{CH} \| / (1 - \| \mathbf{H} \|)$

$n$	Matrix	A	A <sup>2</sup>	A <sup>3</sup>	B	D
<i>Gauss-Jordan</i>						
10	L	$8.3 \times 10^{-7}$	$4.1 \times 10^{-4}$	$3.1 \times 10^{-2}$	$1.2 \times 10^{-7}$	$5.4 \times 10^{-7}$
	R	$9.6 \times 10^{-7}$	$5.8 \times 10^{-4}$	$4.7 \times 10^{-1}$	$2.1 \times 10^{-7}$	$8.9 \times 10^{-7}$
15	L	$2.5 \times 10^{-6}$	$2.4 \times 10^{-3}$	$6.7 \times 10^{-1}$	$2.0 \times 10^{-7}$	$1.5 \times 10^{-6}$
	R	$3.7 \times 10^{-6}$	$9.9 \times 10^{-3}$	34.5	$3.2 \times 10^{-7}$	$1.9 \times 10^{-6}$
20	L	$4.7 \times 10^{-6}$	$3.5 \times 10^{-2}$	103	$2.5 \times 10^{-7}$	$2.8 \times 10^{-6}$
	R	$8.7 \times 10^{-6}$	$6.7 \times 10^{-2}$	871	$4.9 \times 10^{-7}$	$4.1 \times 10^{-6}$
25	L	$7.9 \times 10^{-6}$	$2.0 \times 10^{-1}$	1034	$3.5 \times 10^{-7}$	$4.4 \times 10^{-6}$
	R	$2.0 \times 10^{-5}$	$3.9 \times 10^{-1}$	31107	$6.4 \times 10^{-7}$	$6.4 \times 10^{-6}$
30	L	$1.7 \times 10^{-5}$	$8.6 \times 10^{-1}$	—	$4.0 \times 10^{-7}$	$6.2 \times 10^{-6}$
	R	$3.4 \times 10^{-5}$	2.0	—	$7.7 \times 10^{-7}$	$8.6 \times 10^{-6}$
<i>Choleski</i>						
10	L	$1.9 \times 10^{-6}$	$1.6 \times 10^{-3}$	$3.0 \times 10^{-1}$	$1.2 \times 10^{-7}$	$7.9 \times 10^{-7}$
	R	$1.2 \times 10^{-6}$	$1.5 \times 10^{-3}$	$6.1 \times 10^{-1}$	$3.1 \times 10^{-7}$	$9.2 \times 10^{-7}$
15	L	$6.2 \times 10^{-6}$	$4.0 \times 10^{-2}$	2.9	$2.3 \times 10^{-7}$	$9.8 \times 10^{-7}$
	R	$5.2 \times 10^{-6}$	$4.4 \times 10^{-2}$	107.5	$4.9 \times 10^{-7}$	$2.8 \times 10^{-6}$
20	L	$8.5 \times 10^{-6}$	$3.4 \times 10^{-1}$	249	$3.4 \times 10^{-7}$	$2.3 \times 10^{-6}$
	R	$1.4 \times 10^{-5}$	$3.4 \times 10^{-1}$	25966	$7.0 \times 10^{-7}$	$5.9 \times 10^{-6}$
25	L	$1.5 \times 10^{-5}$	1.7	—	$4.8 \times 10^{-7}$	$6.0 \times 10^{-6}$
	R	$2.9 \times 10^{-5}$	1.7	—	$9.1 \times 10^{-7}$	$1.1 \times 10^{-5}$
30	L	$3.5 \times 10^{-5}$	6.5	—	$6.7 \times 10^{-7}$	$7.6 \times 10^{-6}$
	R	$5.9 \times 10^{-5}$	7.6	—	$1.2 \times 10^{-6}$	$1.6 \times 10^{-5}$

TABLE II—Continued  
Congruent Transformation

10	<i>L</i>	$1.2 \times 10^{-7}$	$1.4 \times 10^{-3}$	$1.9 \times 10^{-1}$	$1.4 \times 10^{-7}$	$4.6 \times 10^{-7}$
	<i>R</i>	$4.9 \times 10^{-7}$	$1.1 \times 10^{-3}$	$3.2 \times 10^{-1}$	$1.9 \times 10^{-7}$	$9.1 \times 10^{-7}$
15	<i>L</i>	0.0	$2.7 \times 10^{-2}$	27.7	$2.2 \times 10^{-7}$	$5.4 \times 10^{-7}$
	<i>R</i>	0.0	$3.0 \times 10^{-2}$	41.9	$3.2 \times 10^{-7}$	$2.1 \times 10^{-6}$
20	<i>L</i>	$4.4 \times 10^{-7}$	$2.2 \times 10^{-1}$	873	$3.4 \times 10^{-7}$	$1.9 \times 10^{-6}$
	<i>R</i>	$4.6 \times 10^{-6}$	$2.3 \times 10^{-1}$	1106	$4.8 \times 10^{-7}$	$3.8 \times 10^{-6}$
25	<i>L</i>	$2.8 \times 10^{-6}$	1.0	8678	$4.5 \times 10^{-7}$	$3.0 \times 10^{-6}$
	<i>R</i>	$6.5 \times 10^{-6}$	1.1	15416	$6.5 \times 10^{-7}$	$5.4 \times 10^{-6}$
30	<i>L</i>	$1.6 \times 10^{-6}$	5.2	—	$5.6 \times 10^{-7}$	$4.0 \times 10^{-6}$
	<i>R</i>	$8.7 \times 10^{-6}$	5.8	—	$7.9 \times 10^{-7}$	$8.6 \times 10^{-6}$

Rank Annihilation

10	<i>L</i>	—	$1.3 \times 10^{-3}$	3.8	0.0	—
	<i>R</i>	—	$1.3 \times 10^{-2}$	15.1	$9.1 \times 10^{-8}$	—
15	<i>L</i>	—	$3.6 \times 10^{-2}$	202	0.0	—
	<i>R</i>	—	$2.4 \times 10^{-1}$	2729	0.0	—
20	<i>L</i>	—	$3.6 \times 10^{-1}$	—	0.0	—
	<i>R</i>	—	1.4	—	$3.8 \times 10^{-7}$	—
25	<i>L</i>	—	1.5	—	0.0	—
	<i>R</i>	—	5.1	—	$5.0 \times 10^{-7}$	—
30	<i>L</i>	—	4.2	—	0.0	—
	<i>R</i>	—	19.6	—	$7.1 \times 10^{-7}$	—

matrix are positive. Hence there may be some cancellation of errors in forming the  $(\mathbf{M})_{\text{approx}}^{-1} \mathbf{M}$  product in Eq. (3.4), this effect becoming more pronounced as the matrix becomes more ill-conditioned. Also the  $\mathbf{A}^3$  matrix contains many zeros (in fact a higher proportion the larger the order) and hence the errors in some of the inverse matrix elements do not show up in the *a* and *f* values.

As shown above, the value of *Q*, defined by Eq. (3.1), reflects the overall difference between the approximate inverse matrix and the exact inverse matrix. This quantity can be computed only if the exact inverse matrix is known, as it is in the case of the test matrices used in this report. In general, however, this will not be the case and yet it is desirable to have some method for estimating at least an upper bound of the difference which does not depend on knowing the exact inverse matrix. Such an upper bound is given by the following inequality [6]

$$(5.1) \quad \|\mathbf{A}^{-1} - \mathbf{C}\| \leq \|\mathbf{CH}\| / (1 - \|\mathbf{H}\|)$$

where  $\mathbf{A}^{-1}$  is the exact inverse of the matrix  $\mathbf{A}$ ,  $\mathbf{C}$  is the computed inverse, and  $\mathbf{H}$  is a matrix defined by

$$\mathbf{H} = \mathbf{I} - \mathbf{AC}$$

where  $\mathbf{I}$  is the identity matrix. (Enclosure within the double vertical lines indicates that the maximum row sum of magnitudes of the elements of the appropriate matrix is taken.)

For comparison purposes values of both the left *L* and right *R* side of the in-

equality (Eq. (5.1)) were computed for all the test matrices used and the results are presented in Table II.

As shown in Table II the right side of the inequality (Eq. (5.1)) is indeed usually greater than the term on the left, which is the maximum row sum of the matrix formed by subtracting the computed inverse from the exact inverse. In the few cases shown in Table II where the inequality does not seem to hold the apparent violation can probably be attributed to the particular method of rounding used in the arithmetic subroutines of the program and the consequent accumulation of round-off error. Hence it is possible to get an estimate at least of an upper bound of the term on the left of the inequality without knowing the exact inverse of a matrix.

**6. Conclusions.** From the foregoing calculations it appears that any one of the matrix inversion methods tried will invert a well-conditioned least squares matrix of order 10 to 30. When the matrix becomes ill-conditioned then the Gauss-Jordan method appears to be clearly superior (at least for the matrices studied). However, if memory capacity must be considered and only a triangular array can be tolerated, then the Choleski and congruent transformation methods appear to offer an advantage for large  $n$ . Only in the case of the relatively poorly conditioned  $\mathbf{A}^3$  matrix did the Choleski method appear to be the better of these two for evaluating the inverse matrix. From a study involving the inversion of several well-conditioned least squares matrices of orders to  $n = 29$  all the methods tried gave  $a$  and  $f$  values about two to three orders of magnitude smaller than those obtained with the  $\mathbf{A}^3$  matrix and consistent with those obtained with the  $\mathbf{A}$  matrix.

**7. Acknowledgment.** The authors wish to express their appreciation to Drs. H. A. Levy and A. S. Householder for interesting and helpful discussions concerning this work.

Chemistry Division, Oak Ridge National Laboratory  
Oak Ridge, Tennessee  
Central Data Processing Facility  
Oak Ridge, Tennessee

1. ANTHONY RALSTON & HERBERT S. WILF, *Mathematical Methods for Digital Computers*, John Wiley and Sons, Inc., New York, 1960, p. 39.
2. *Contributions to the Solution of Systems of Linear Equations and the Determination of Eigenvalues*, U. S. Department of Commerce, National Bureau of Standards Applied Mathematics Series 39, 1954, p. 31.
3. WILLIAM R. BUSING & HENRI A. LEVY, "A procedure for inverting large symmetric matrices," *Comm. ACM*, v. 5, 1962, p. 445.
4. ANTHONY RALSTON & HERBERT S. WILF, loc. cit., p. 73.
5. MORRIS NEWMAN & JOHN TODD, "The evaluation of matrix inversion programs," *J. Soc. Indust. Appl. Math.*, v. 6, 1958, p. 466.
6. ALSTON S. HOUSEHOLDER, *The Theory of Matrices in Numerical Analysis*, Ginn and Co., Boston, in press, Chapter 4.