A UNITED STATES
DEPARTMENT OF
COMMERCE
PUBLICATION

NOAA Technical Report NOS 56

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration
National Ocean Survey

Cholesky Factorization and Matrix Inversion

ERWIN SCHMID

ROCKVILLE, MD.
March 1973

NOAA TECHNICAL REPORTS

National Ocean Survey Series

The National Ocean Survey (NOS) provides charts and related information for the safe navigation of marine and air commerce. The survey also furnishes other earth science data--from geodetic, hydrographic, oceanographic, geomagnetic, seismologic, gravimetric, and astronomic surveys, observations, investigations, and measurements--to protect life and property and to meet the needs of engineering, scientific, defense, commercial, and industrial interests.

Because many of these reports deal with new practices and techniques, the views expressed are those of the authors and do not necessarily represent final survey policy. NOS series NOAA Technical Reports is a continuation of, and retains the consecutive numbering sequence of, the former series, Environmental Science Services Administration (ESSA) Technical Reports Coast and Geodetic Survey (C&GS), and the earlier series, C&GS Technical Bulletins.

Those publications marked by an asterisk are out of print. The others are available through the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C. 20402. Price as indicated. Beginning with 39, microfiche is available at the National Technical Information Service (NTIS), U.S. Department of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, Va. 22151. Price $0.95. When given, order by accession number in parentheses.

COAST AND GEODETIC SURVEY TECHNICAL BULLETINS

*No. 17 On the Time Interval Between Two Consecutive Earthquakes. Tokuji Utsu, February 1962.

*No. 18 Submarine Physiography of the U.S. Continental Margins. G. F. Jordan, March 1962.

*No. 19 Analytic Absolute Orientation in Photogrammetry. G. C. Tewinkel, March 1962.

*No. 20 The Earth as Viewed from a Satellite. Erwin Schmid, April 1962.

*No. 21 Analytic Aerotriangulation. W. D. Harris, G. C. Tewinkel, and C. A. Whitten, July 1962. (Corrected July 1963.)

*No. 22 Tidal Current Surveys by Photogrammetric Methods. Morton Keller, October 1963.

*No. 23 Aerotriangulation Strip Adjustment. M. Keller and G. C. Tewinkel, August 1964.

*No. 24 Satellite Triangulation in the Coast and Geodetic Survey. February 1965.

*No. 25 Aerotriangulation: Image Coordinate Refinement. M. Keller and G. C. Tewinkel, March 1965.

*No. 26 Instrumented Telemetering Deep Sea Buoys. H. W. Straub, J. M. Arthaber, A. L. Copeland, and D. T. Theodore, June 1965.

*No. 27 Survey of the Boundary Between Arizona and California. Lansing G. Simmons, August 1965.

*No. 28 Marine Geology of the Northeastern Gulf of Maine. R. J. Malloy and R. N. Harbison, February 1966.

*No. 29 Three-Photo Aerotriangulation. M. Keller and G. C. Tewinkel, February 1966.

*No. 30 Cable Length Determinations for Deep-Sea Oceanographic Operations. Robert C. DARLING, June 1966.

*No. 31 The Automatic Standard Magnetic Observatory. L. R. Alldredge and I. Saldukas, June 1966.

ESSA TECHNICAL REPORTS

*C&GS 32 Space Resection in Photogrammetry. M. Keller and G. G. Tewinkel, September 1966.

*C&GS 33 The Tsunami of March 28, 1964, as Recorded at Tide Stations. M. G. Spaeth and S. C. Berkman, July 1967.

*C&GS 34 Aerotriangulation: Transformation of Surveying and Mapping Coordinate Systems. Melvin J. Umbach, August 1967.

*C&GS 35 Block Analytic Aerotriangulation. M. Keller and G. C. Tewinkel, November 1967.

(Continued on inside back cover)

U.S. DEPARTMENT OF COMMERCE
Frederick B. Dent, Secretary

NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION
Robert M. White, Administrator

NATIONAL OCEAN SURVEY
Allen L. Powell, Director

NOAA Technical Report NOS 56

Cholesky Factorization and Matrix Inversion

Erwin Schmid

ROCKVILLE, MD.
MARCH 1973

UDC 519.281.2:528.14

519	Combinatorial and statistical analysis
.2	Calculus of probability
.281.2	Method of least squares
528	Geodesy
.14	Adjustment by method of least squares

For sale by the Superintendent of Documents, U.S. Government Printing Office
Washington, D.C. 20402. Price: 40 cents, domestic postpaid; 30 cents, GPO Bookstore
Stock No. 0317-00114

Contents

	Page
Abstract	1
1. Factorization of the normal equations matrix	1
2. Inversion of a triangular matrix	3
3. Improving the precision of the inverse	6
4. Comparison with the back solution	8
5. Application to inversion of nonsymmetric matrix	9
6. Equivalence of the symmetric solution with the least squares postulate	9
A. Observation equations	9
B. Condition equations	9
7. The Gaussian algorithm for symmetric matrices	10
8. The Gaussian algorithm for nonsymmetric matrices	11
9. Square root factorization	12
References	13

Cholesky Factorization and Matrix Inversion

ERWIN SCHMID [1]

ABSTRACT. The Cholesky square root algorithm used in the solution of linear equations with a positive definite matrix of coefficients is developed by elementary matrix algebra, independent of the Gaussian elimination from which it was originally derived. The Cholesky factorization leads to a simple inversion procedure for the given matrix. A simple transformation makes the inversion applicable to nonsymmetric matrices. The least squares hypothesis is shown to be the simplest and most general unique solution of a system of linear equations with a nonsquare matrix of coefficients. The method of proof is extended to develop the Gaussian elimination algorithm in a readily comprehensible procedure.

1. FACTORIZATION OF THE NORMAL EQUATIONS MATRIX

The Cholesky algorithm for solving a set of normal equations, in the sense as used in geodesy, follows quite readily from the earlier Doolittle solution, now known as Gauss-Doolittle, which in turn is developed in the textbooks as a special case of Gaussian elimination applicable to a general set of linear equations. Although Doolittle's only publication (U.S. Coast and Geodetic Survey, 1881) on the subject is merely a presentation of the procedure without proof or reference to source, there is little doubt that his algorithm is based directly, or possibly by way of a translation such as Bertrand (1855), on a method for solving (symmetric) normal equations which Gauss (1811a) proposed and proved some time before he developed the general "Gaussian elimination."

In any case, the Cholesky algorithm can be readily established with elementary matrix algebra. Since this algorithm is well documented, we limit ourselves to a heuristic approach with a 4×4 matrix which can readily be generalized to an $n\times n$.

The product CC^T of a lower triangular, nonsingular matrix with real coefficients

$$C=\begin{pmatrix} C_{11} & 0 & 0 & 0 \\ C_{12} & C_{22} & 0 & 0 \\ C_{13} & C_{23} & C_{33} & 0 \\ C_{14} & C_{24} & C_{34} & C_{44} \end{pmatrix}$$

and its transpose C^T is (row on row multiplication of C)

$$CC^T=\begin{pmatrix} C_{11}C_{11} & C_{11}C_{12} & C_{11}C_{13} & C_{11}C_{14} \\ C_{11}C_{12} & C_{12}C_{12}+C_{22}C_{22} & C_{12}C_{13}+C_{22}C_{23} & C_{12}C_{14}+C_{22}C_{24} \\ C_{11}C_{13} & C_{12}C_{13}+C_{22}C_{23} & C_{13}C_{13}+C_{23}C_{23}+C_{33}C_{33} & C_{13}C_{14}+C_{23}C_{24}+C_{33}C_{34} \\ C_{11}C_{14} & C_{12}C_{14}+C_{22}C_{24} & C_{13}C_{14}+C_{23}C_{24}+C_{33}C_{34} & C_{14}C_{14}+C_{24}C_{24}+C_{34}C_{34}+C_{44}C_{44} \end{pmatrix}$$

[1] National Ocean Survey, National Oceanic and Atmospheric Administration.

a symmetric matrix which is positive definite, since $\det(CC^T) = \det(C) \cdot \det(C^T) = [\det(C)]^2$. Arranging this product matrix, as is customary, in alternate rows labeled (1), (2), (3), and (4), and deleting the redundant terms below the diagonal, gives the scheme:

$$
\begin{array}{lllll}
(1) & C_{11}C_{11} & C_{11}C_{12} & C_{11}C_{13} & C_{11}C_{14} \\
(1') & C_{11} & C_{12} & C_{13} & C_{14} \\
(2) & & C_{12}C_{12}+C_{22}C_{22} & C_{12}C_{13}+C_{22}C_{23} & C_{12}C_{14}+C_{22}C_{24} \\
(2') & & C_{22} & C_{23} & C_{24} \\
(3) & & & C_{13}C_{13}+C_{23}C_{23}+C_{33}C_{33} & C_{13}C_{14}+C_{23}C_{24}+C_{33}C_{34} \\
(3') & & & C_{33} & C_{34} \\
(4) & & & & C_{14}C_{14}+C_{24}C_{24}+C_{34}C_{34}+C_{44}C_{44} \\
(4') & & & & C_{44}
\end{array}
\tag{1}
$$

From (1) the matrix C can be reconstructed in rows (1′), (2′), (3′), and (4′) in that order. This reversal of the multiplication procedure results in the Cholesky algorithm. We now assume the matrix CC^T given as a symmetric positive definite matrix N with entries n_{ik}. In row (1′) are developed the entries of the first column of C. The first term of (1′) is evidently the square root of the first term of row (1), and the remaining terms are obtained by dividing the corresponding term of row (1) by the first term of row (1′). In row (2) the first term is "reduced" by the product of C_{12}, the term immediately above it, multiplied by itself; and the root of the difference gives C_{22}. The other terms in that row are reduced by the product of C_{12} and the corresponding term of the pertinent column; then the remainder is divided by C_{22}. This completes row (2′) which is the second column of C.

The element n_{ik} in the ith row and kth column ($i \leq k$) of the given matrix is

$$n_{ik} = C_{1i}C_{1k} + C_{2i}C_{2k} + \ldots + C_{ii}C_{ik} = \sum_{r=1}^{i} C_{ri}C_{rk} \tag{2}$$

as is easily verified by multiplication of the ith row of C with the kth. Writing the expression (2) in the form

$$C_{ii}C_{ik} = n_{ik} - \sum_{r=1}^{i-1} C_{ri}C_{rk} \tag{3}$$

displays the complete algorithm in a single formula. The first factor C_{ri} in the summation represents all the entries in the column of the diagonal term situated above this term and previously reduced. The second factor C_{rk} represents similar terms in the column of the term n_{ik} being reduced. For the diagonal term ($k=i$), which is computed first in a given row, the indicated reduction on the right-hand side of (3) results in $C_{ii}C_{ii}$ which requires a square root extraction. The other terms C_{ik} in the row ($k>i$) are obtained by division with C_{ii}.

A simple numerical example will illustrate the algorithm and point out some computational characteristics that are difficult to formalize algebraically.

Example 1:

Given is the positive definite symmetric matrix

$$N = \begin{pmatrix} 729 & 432 & 621 & 405 \\ 432 & 1856 & 1928 & 560 \\ 621 & 1928 & 2054 & 685 \\ 405 & 560 & 685 & 741 \end{pmatrix}$$

to be factored into the product of a lower triangular matrix C times its transpose C^T. The entries of each row, beginning with the diagonal term, are written below in alternate rows as in (1).

$$
\begin{array}{rrrr}
729 & 432 & 621 & 405 \\
27 & 16 & 23 & 15 \\
\hline
 & 1856 & 1928 & 560 \\
 & 40 & 39 & 8 \\
\hline
 & & 2054 & 685 \\
 & & 2 & 14 \\
\hline
 & & & 741 \\
 & & & 16 \\
\hline
\end{array}
$$

The entries of the triangular matrix to be computed are written directly below the corresponding term of the given matrix. For example, $n_{34}=685$ of the given matrix is reduced by $23\times 15+39\times 8=657$, $685-657=28$ which, divided by the diagonal term 2, gives the reduced term $C_{34}=14$. A complete numerical check on the computations consists in multiplying the computed triangular matrix by itself, column by column, since it is presented in the above scheme in its transposed position.

In order to make the algorithm readily comprehensible, the above example was designed so that all the numerical operations result in exact integers. This hides the effect of error accumulation in the general case. Extra significant figures must be carried in all the computations because all the entries a_{ik} of the answer are, in accordance with (3), the result of a difference of two numbers of roughly the same magnitude. The situation is aggravated when a reduced diagonal term is small relative to those previously reduced. If, for example, due to error accumulation, $a_{23}=39$ in the example were increased to 39.05 the reduced diagonal term in the next row would become $(2054-23^2-39.05^2)^{1/2}=0.31$ instead of 2, a completely erroneous figure which would falsify all subsequent results, particularly the entries in that row.

By changing the diagonal terms of the given matrix N of the numerical example very slightly, say by adding 1 to each of these diagonal terms and factoring the resulting matrix, we should obtain numbers close to those obtained before but now no longer exact or rational. Operating with floating decimal point to four significant figures (not decimals), since this is the largest number of digits given in the problem, and comparing the result with that obtained with a larger number of significant digits, it will be found that:

1. The results are correct to roughly four figures in the first two rows, i.e., as long as the *reduced* diagonal terms are of the same decimal magnitude.
2. The diagonal term of the third row again reduces to a number which is no greater than 1/10 of the two previously reduced diagonals, i.e., roughly one magnitude smaller. The figures in this row are found to be good to two digits only, and the degradation of accuracy is carried into all subsequent computations.
3. With six-figure floating point precision the factorization will prove correct to at least four digits in all the numbers of the result.

By constructing a problem in which a diagonal term reduces to a number two magnitudes smaller than the previously reduced diagonals, it will be found that four additional significant digits are needed. In general we may conclude that if the ratio of the largest reduced diagonal to the smallest is on the order of 10^k, the solution requires at least $n+2k$ significant digits to approximate n-figure accuracy in all the entries of the reduced matrix. Computing with less precision may result in the small diagonal term reducing to zero or a small quantity which represents merely computer "noise," and consequently a completely erroneous result.

There is no practical advantage in formulating such criteria more rigorously because the reduced diagonals are not even approximately known beforehand, but are developed in the solution. They can only indicate the precision of the solution when it is complete (unless, of course, it breaks down before that) and the increase in needed precision if a repetition of the factorization seems indicated. The check on the solution, mentioned above, of multiplying C by C^T and comparing the result with the given matrix N will spot blunders but is not sensitive to this type of error accumulation, any more than substitution of approximated roots back into an algebraic equation for example. In both cases error compensation masks the location and the amount of error.

From equation (3) it follows that the Cholesky factorization is unique except for the ambiguity in sign introduced when $i=k$ and the consequent root extraction to find C_{ii} from C_{ii}^2. However, a second and equally valid and useful factorization of a normal matrix can be found, analogous to the development above, by postulating a product D^TD, where D^T is an upper triangular matrix, instead of CC^T. Writing this product in terms of the elements d_{ik} of d, it will be readily apparent that the corresponding algorithm starts at the lower right-hand corner of the given matrix N and proceeds up the last column (or row) of the matrix and, in sequence, through the matrix from right to left (or upward).

2. INVERSION OF A TRIANGULAR MATRIX

One of the most useful applications of the Cholesky factorization lies in the direct inversion of a symmetric positive definite matrix N, which in the present context we designate a normal matrix, such as is encountered in the normal equations of least squares theory. The method is easy to comprehend, economical in computing space and time, and capable of optimal refinement of precision to a specified number of digits. Basically all that is required is the inversion of the triangular matrix C obtained in the Cholesky factorization of N. From $N=CC^T$ it follows that

$$N^{-1}=(C^T)^{-1}C^{-1}=(C^{-1})^TC^{-1} \tag{4}$$

The inversion of a triangular matrix C is a relatively simple algorithm to execute. Consideration of a 3×3 upper triangular matrix

$$C^T=\begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix}$$

should be sufficient to indicate the sequence of

operations and their validity for any size and type of triangular matrix. The nonsingularity and invertibility of a triangular matrix are apparent from the criterion that, for all i, $C_{ii} \neq 0$. We postulate for the inverse an upper triangular matrix $(C^T)^{-1}$ with undetermined coefficients γ_{ik}

$$(C^T)^{-1} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ 0 & \gamma_{22} & \gamma_{23} \\ 0 & 0 & \gamma_{33} \end{pmatrix}$$

and set the condition $C^T(C^T)^{-1} = I$, the unit matrix of order 3. In order to make the algorithm more convenient for visual presentation and hand operation, we write the inverse in transposed form C^{-1} underneath the given matrix C^T so that the matrix multiplication can be performed row by row:

$$C^T = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix} \begin{matrix} (1) \\ (2) \\ (3) \end{matrix} \qquad (5)$$

$$C^{-1} = \begin{pmatrix} \gamma_{11} & 0 & 0 \\ \gamma_{12} & \gamma_{22} & 0 \\ \gamma_{13} & \gamma_{23} & \gamma_{33} \end{pmatrix} \begin{matrix} (1)^{-1} \\ (2)^{-1} \\ (3)^{-1} \end{matrix}$$

The numbers on the right designate the respective rows of the given matrix and of the desired inverse. Since the product must equal I, each row multiplied by the corresponding primed row $= 1$, and $= 0$ otherwise. In the arrangement of the matrices according to (5) it is convenient to start with the last row $(3)^{-1}$ of C^{-1} and multiply it in turn with row (3), (2), and (1), setting the products equal to 1, 0, and 0 respectively. Each multiplication yields a new γ entry. Thus the first three conditions read

$$\begin{cases} \gamma_{33}C_{33} & = 1 \\ \gamma_{33}C_{23} + \gamma_{23}C_{22} & = 0 \\ \gamma_{33}C_{13} + \gamma_{23}C_{12} + \gamma_{13}C_{11} & = 0 \end{cases} \qquad (6)$$

With an $n \times n$ matrix C^T there would have been n such equations. From the first of equations (5) we get the diagonal term $\gamma_{33} = 1/C_{33}$, the reciprocal of the diagonal term of the given matrix, a relation which holds for all the diagonal terms. Substituting this value in the second of equations (6) gives $\gamma_{23} = \gamma_{33}C_{23}/-C_{22}$, and with both γ_{33} and γ_{23} in the third of equations (6) $\gamma_{13} = (\gamma_{33}C_{13} + \gamma_{23}C_{12})/-C_{11}$. This completes row $(3)^{-1}$, and we proceed to evaluate row $(2)^{-1}$ in similar fashion by multiplying it in order with rows (2) and (1). The multiplication with row (3) is unnecessary, since it imposes no new condition on the coefficients. It merely proves the validity of our original assumption, i.e., that the inverse is the same type of triangular matrix.

Generally, we start with the bottom row of the inverse and compute its entries in turn, from right to left, by accumulating the products of the γ's (which were previously computed and entered) times the corresponding C's in the given matrix and in the row above that used just previously.

Example 2:

A simple numerical example will illustrate that the algorithm is easier to perform than to formulate in words. Given, to invert the triangular matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix}$$

The significant portion of the matrix is written above the solid line, and the transpose of the inverse is developed immediately below in the space occupied by the omitted zeros.

$$\begin{array}{ccc} 1 & 2 & 3 \\ 1 & 4 & 5 \\ -1/2 & 1/4 & 6 \\ \hline -1/12 & -5/24 & 1/6 \end{array}$$

Starting with the diagonal term of the third row, the corresponding term of the inverse is the reciprocal 1/6. The adjacent term is $\dfrac{(1/6)(5)}{-4} = -5/24$, and the next is $\dfrac{(1/6)(3) + (2)(-5/24)}{-1} = -1/12$. The next row from the bottom is computed similarly, starting with the diagonal term 1/4 and continuing with $\dfrac{(1/4)(2)}{-1} = -1/2$.

It should be noted that:

a) The above arrangement is for compactness in hand computing but also indicates the possibility for similar space saving in computer memory. It fits conveniently into the unused space of the preceding Cholesky factorization. The existence of the omitted zero portion of both the given matrix and its inverse must be kept in mind for an understanding of the algorithm.

b) In practice the entries of the inverse are of course carried as decimal fractions.

c) The entries in the inverse are completely independent of all previous rows of the inverse as well as of the corresponding columns of the given triangular matrix. This indicates that the computation could equally well have started with the first single-entry row of the inverse and proceeded downward. It follows, therefore, that if a given triangular matrix is augmented with an additional row or rows, the portion of the inverse already

computed remains unaltered, which is not true for other types of matrices.

d) An independent check on the computations, as well as an alternative first computation, consists of column-by-column multiplication of the two matrices in the above arrangement. For example, the first column vector of the inverse $(1, -1/2, -1/12)$ times (inner product) the 1st, 2d, and 3d column vectors of the given matrix, i.e., (1, 0, 0), (2, 4, 0), and (3, 5, 6) respectively, satisfy the conditions

$$\begin{cases} (1)(1)+(-1/2)(0)+(-1/12)(0)=1 \\ (1)(2)+(-1/2)(4)+(-1/12)(0)=0 \\ (1)(3)+(-1/2)(5)+(-1/12)(6)=0 \end{cases}$$

with similar results for the second and subsequent column vectors of the inverse. This can be interpreted as the result of actually interchanging the role of the two matrices, which is valid because of the postulated reciprocity of the matrix inverse. A summation check is superfluous because each row is independent of the others. A little reflection will show how the column by column multiplication can be used to compute the inverse in the first place.

e) The same arrangement can be used to invert a lower triangular matrix by writing it in its transposed position. The inverse will then appear in its proper form.

For computer programing it is necessary to have a formula for the general term γ_{ik} of the inverse of an n-dimensional triangular matrix. This follows directly from a consideration of the extension of (5) to n dimensions.

$$C^T=\begin{pmatrix} C_{11} & C_{12} & \cdots & C_{1i} & C_{1,i+1} & \cdots & C_{1,k-1} & C_{1k} & \cdots & C_{1n} \\ 0 & C_{22} & \cdots & C_{2i} & C_{2,i+1} & \cdots & C_{2,k-1} & C_{2k} & \cdots & C_{2n} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & C_{ii} & C_{i,i+1} & \cdots & C_{i,k-1} & C_{ik} & \cdots & C_{in} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & C_{nn} \end{pmatrix}$$

$$C^{-1}=\begin{pmatrix} \gamma_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \gamma_{12} & \gamma_{22} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{1k} & \gamma_{2k} & \cdots & \gamma_{ik} & \gamma_{i+1,k} & \cdots & \gamma_{k-1,k} & \gamma_{kk} & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \gamma_{1n} & \gamma_{2n} & \cdots & \gamma_{in} & \gamma_{i+1,n} & \cdots & \gamma_{k-1,n} & \gamma_{kn} & \cdots & \gamma_{nn} \end{pmatrix}$$

Generalizing equations (6), we have for $i<k$ the condition

$$\gamma_{kk}C_{ik}+\gamma_{k-1,k}C_{i,k-1}+\ldots+\gamma_{i+1,k}C_{i,i+1}+\gamma_{ik}C_{ii}=0$$

from which

$$\gamma_{ik}=-\frac{1}{C_{ii}}(\gamma_{kk}C_{ik}+\gamma_{k-1,k}C_{i,k-1}+\ldots+\gamma_{i+1,k}C_{i,i+1})$$

where all the γ's on the right have been computed in the preceding steps. This formula can be written as

$$\gamma_{ik}=-\frac{1}{C_{ii}}\sum_{r=0}^{r=k-(i+1)}\gamma_{k-r,k}C_{i,k-r} \quad \text{for } i<k$$

or more concisely

$$\gamma_{ik}=-\frac{1}{C_{ii}}\sum_{r=i+1}^{r=k}\gamma_{rk}C_{ir} \qquad \text{for } i<k.$$

In analogy with the first of equations (6) the first γ in the row is $\gamma_{kk}=\frac{1}{C_{kk}}$, while for $i>k$, $\gamma_{ik}=0$.

Having factored the matrix N into CC^T and having inverted C, it is a simple matter to obtain the inverse of N from (4).

The quantity on the right side of (4) results from row on row multiplication of $(C^T)^{-1}$ on itself, where $(C^T)^{-1}$ is by our convention the upper triangular form. If the triangular matrix inverse presents itself in the lower triangular form C^{-1}, then N^{-1} is produced by column on column multiplication of C^{-1} on itself.

Since the solution to a set normal equation

$$Nx=\ell$$

is

$$x=N^{-1}\ell$$

the vector x is found by multiplying N^{-1} with the given vector ℓ. This solution is no more complicated

491-890 OL - 73 - 2

or lengthy, as will be shown, than the conventional back solution, and it contains error theoretical information that only the inverse of N can provide.

3. IMPROVING THE PRECISION OF THE INVERSE

From $N = CC^T$ it follows that

$$C^{-1}N(C^{-1})^T = I. \tag{7}$$

When the computations outlined above are executed rigorously in floating decimal mode, the resulting inverse will be optimal; and the indicated multiplication in (7) will fail to exactly equal the unit matrix I only to the extent that the computer carries too few significant digits for the problem. If the given matrix N is known to be positive definite and one or more of the reduced diagonal terms in the Cholesky factorization (1) reduce to an excessively small number relative to the other diagonals, then the corresponding diagonal term of the inverse N^{-1} will be excessively large, indicating that the mean error of the variable associated with this diagonal term is so large that the determination of this particular variable is meaningless from the standpoint of least squares theory. Such a near-singularity in the N matrix is a direct consequence of a poorly conceived phase of the measuring process and can be corrected only at the source.

Loss of precision can, however, be considerable in hand computation or some other form involving a fixed decimal point. In such a case the multiplication on the left side of (7) produces a matrix I^* which is symmetric but only approximately diagonal:

$$C^{-1}N(C^{-1})^T = I^* \tag{8}$$

The Cholesky factorization and the inverse can now be improved to match the precision of floating point computation, be extended to a larger number of significant digits, or corrected for possible blunders by the following procedure.

The matrix I^* in (8) is well conditioned and can be factored very precisely by the Cholesky algorithm into $I^* = C^*(C^*)^T$ so that (8) becomes

$$C^{-1}N(C^{-1})^T = C^*(C^*)^T \tag{9}$$

Inverting C^*, a process which is again capable of high precision since C^* is strongly diagonal, (9) becomes

$$[(C^*)^{-1}C^{-1}]N[(C^*)^{-1}C^{-1}]^T = I \tag{10}$$

where now, it will be found, the identity with the unit matrix is good to the number of significant digits used in computing C^* and $(C^*)^{-1}$, less the inevitable degradation caused by the variation in magnitude of the reduced diagonal terms. The quantity $(C^*)^{-1}C^{-1}$ inside the brackets in (10) is a corrected value C_f^{-1} for C^{-1} and will satisfy the condition (7) optimally. The corrected inverse of N will be $N^{-1} = (C_f^{-1})^T C_f^{-1}$.

Example 3:

The matrix

$$N = \begin{pmatrix} 730 & 432 & 621 & 405 \\ 432 & 1857 & 1928 & 560 \\ 621 & 1928 & 2055 & 685 \\ 405 & 560 & 685 & 742 \end{pmatrix}$$

factors by the Cholesky algorithm into CC^T where

$$C = \begin{pmatrix} 27.02 & 0 & 0 & 0 \\ 15.90 & 40.02 & 0 & 0 \\ 22.98 & 39.00 & 2.455 & 0 \\ 14.99 & 8.005 & 11.54 & 17.89 \end{pmatrix}$$

This result is correct to four significant digits and can be obtained by floating point computation, carrying six digits throughout since the ratio of the largest to the smallest reduced diagonal is 40.02/2.455, roughly one magnitude. A small blunder is included: the first entry of the second row should read 15.99.

A rough inversion of C produces

$$C^{-1} = \begin{pmatrix} .03701 & 0 & 0 & 0 \\ -.01471 & .02499 & 0 & 0 \\ -.1128 & -.3969 & .4073 & 0 \\ .04838 & .2449 & -.2628 & .05590 \end{pmatrix}$$

The multiplication $C^{-1}N(C^{-1})^T$ produces the symmetric matrix I^* of (7).

$$I^* = \begin{pmatrix} .999910273 & .0021236338 & -.032254215 & .020536849 \\ \text{———} & 1.0000481731 & .000362250 & -.0005391296 \\ & \text{———} & 1.00058520 & -.000923670 \\ & & \text{———} & 1.000893582 \end{pmatrix}$$

The factorization $I^* = C^*(C^*)^T$ yields

$$C^* = \begin{pmatrix} .9999551355 & 0 & 0 & 0 \\ .0021237291 & 1.0000218312 & 0 & 0 \\ -.0322556621 & .0004307429 & .9997722674 & 0 \\ .0205377704 & -.0005827335 & -.0002610191 & 1.0002356594 \end{pmatrix}$$

The matrix I^* is so nearly a unit matrix that this factorization and the subsequent inversion of C^* can be computed precisely without the aid of floating decimal to whatever number of significant digits the computer can handle. In this case we have used 10-digit accuracy.

The inversion routine gives

$$(C^*)^{-1} = \begin{pmatrix} 1.0000448665 & 0 & 0 & 0 \\ -.0021237780 & .9999781693 & 0 & 0 \\ .0322653720 & -.0004308316 & 1.0002277845 & 0 \\ -.0205266703 & .0005824711 & .0002610170 & .9997643961 \end{pmatrix}$$

and the product $(C^*)^{-1}C^{-1}$ gives the corrected inverse C_f^{-1}:

$$C_f^{-1} = \begin{pmatrix} .03701166051 & 0 & 0 & 0 \\ -.01478827989 & .02498945445 & 0 & 0 \\ -.1116252151 & -.3970011742 & .4073927766 & 0 \\ .04757089854 & .2447532589 & -.2626317711 & .05588682974 \end{pmatrix}$$

from which the inverse of the given matrix N can be computed, correct to at least eight significant figures in all the entries by performing the matrix multiplication $N^{-1} = (C_f^{-1})^T C_f^{-1}$, the result being

$$N^{-1} = \begin{pmatrix} \underline{.01631173527} & .05558892286 & -.05796893565 & .002658586707 \\ & \underline{.2181385629} & -.2260153925 & .01367848371 \\ & & \underline{.2349443216} & -.01467765708 \\ & & & \underline{.003123337738} \end{pmatrix}$$

It is important to note that the matrix C_f^{-1} is not the precise inverse of the approximate matrix C from the first Cholesky factorization of N. It is, rather, the inverse of the factor C which would have been obtained in such a factorization if precision to a larger number of significant digits had been available. This factor can, in fact, be obtained without such a more precise factorization by inverting the inverse C_f^{-1}, i.e., by computing $(C_f^{-1})^{-1}$. For this numerical example the result of such an inversion yields

$$(C_f^{-1})^{-1} = \begin{pmatrix} 27.01851217 & 0 & 0 & 0 \\ 15.98903733 & 40.01688000 & 0 & 0 \\ 22.98424117 & 38.99614637 & 2.45463361 & 0 \\ 14.98972250 & 8.00484118 & 11.53518233 & 17.89330339 \end{pmatrix}$$

This triangular matrix multiplied by its transpose reproduces the given matrix N to at least eight significant figures in all the entries.

4. COMPARISON WITH THE BACK SOLUTION

Gauss developed his algorithm for the solution of linear equations from the standpoint of obtaining an equivalent set of equations in each of which an additional variable has been eliminated. He also proved that the algorithm operating on the coefficients of the unknowns can be extended to the column of constant terms to produce the corresponding set of constants for the new equations. He could then solve for the unknowns in order, starting with the last equation which contains only one unknown, the so-called back solution. This type of solution is still being used to some extent. When the Cholesky algorithm is viewed as a simple variation of the Gauss-Doolittle algorithm, the analogous treatment of the constant column follows directly without further proof. However, with a development of this algorithm from the standpoint of matrix algebra, independent of Gauss, justification of the validity of extending the reduction to the column of constants is necessary and can be demonstrated as follows:

Consider a set of, say, four homogeneous linear equations, the matrix of whose coefficients N is nonsingular, symmetric, and positive definite:

$$\begin{cases} n_{11}x_1 + n_{12}x_2 + n_{13}x_3 + n_{14}x_4 = 0 \\ n_{12}x_1 + n_{22}x_2 + n_{23}x_3 + n_{24}x_4 = 0 \\ n_{13}x_1 + n_{23}x_2 + n_{33}x_3 + n_{34}x_4 = 0 \\ n_{14}x_1 + n_{33}x_3 + n_{34}x_3 + n_{44}x_4 = 0 \end{cases} \tag{11}$$

or $Nx = 0$. Factoring N we obtain the equivalent set $CC^Tx = 0$ and, on multiplying both sides by C^{-1},

$$Cx = 0$$

which, written out in full, is

$$\begin{cases} C_{11}x_1 + C_{12}x_2 + C_{13}x_3 + C_{14}x_4 = 0 \\ \qquad\quad C_{22}x_2 + C_{23}x_3 + C_{24}x_4 = 0 \\ \qquad\qquad\quad C_{33}x_3 + C_{34}x_4 = 0 \\ \qquad\qquad\qquad\quad C_{44}x_4 = 0 \end{cases} \tag{12}$$

where the C's are derived in the Cholesky factorization of N. The set of conditions (12) is completely equivalent to (11). Furthermore, the first three equations of (11) are equivalent to the first three of (12). This follows because in producing the coefficients for the first three equations of (12) the last row of the matrix N has not yet been considered, and these three equations must therefore be independent of the condition expressed by the fourth equation (11). These two equivalent sets of three equations each have four unknowns, and one unknown is therefore a free parameter. Setting $x_4 = 1$ in both sets of three equations we have the equivalence of

$$\begin{cases} n_{11}x_1 + n_{12}x_2 + n_{13}x_3 + n_{14} = 0 \\ n_{12}x_1 + n_{22}x_2 + n_{23}x_3 + n_{24} = 0 \\ n_{13}x_1 + n_{23}x_2 + n_{33}x_3 + n_{34} = 0 \end{cases} \tag{13}$$

and

$$\begin{cases} C_{11}x_1 + C_{12}x_2 + C_{13}x_3 + C_{14} = 0 \\ \qquad\quad C_{22}x_2 + C_{23}x_3 + C_{24} = 0 \\ \qquad\qquad\quad C_{33}x_3 + C_{34} = 0 \end{cases} \tag{14}$$

where (13) is typical of the nonhomogeneous, symmetric, linear, normal equations of least squares theory and (14) the corresponding set of Cholesky-reduced equations that can be solved with a back solution. The extension of the above demonstration from four to n equations involves no essential difficulties. It is customary and convenient to designate the coefficients in the last column of (13) and (14), i.e., the constants in the equations, by symbols different from the symbols for the coefficients of the unknowns x and occasionally to transfer these constants to the other side of the equations. Thus equations (14) can be written in the conventional form

$$\begin{cases} C_{11}x_1 + C_{12}x_2 + C_{13}x_3 = \ell_1 \\ \qquad\quad C_{22}x_2 + C_{23}x_3 = \ell_2 \\ \qquad\qquad\quad C_{33}x_3 = \ell_3 \end{cases}$$

or

$$C^Tx = \ell \tag{15}$$

where C^T is upper triangular, and the vector ℓ has components that are the negatives of the constant terms in (14). Computing the inverse of $(C^T)^{-1}$ and multiplying it into both sides of (15) gives the solution for x:

$$x = (C^T)^{-1}\ell$$

By going through these computations it can be seen that they involve the identical operations used in the conventional back solution. Given that the inverse can be improved, if necessary, as shown in section 3, there is no doubt that this approach is at least as good as the conventional type of back solution. Furthermore, having computed $(C^T)^{-1}$ it is merely necessary to multiply this matrix by its transpose to obtain the complete inverse N^{-1} of N, which

a) solves the equations (13) directly, with the option of refining the solution by improving the inverse, and

b) as a covariance matrix permits the statistical interpretation of the solution and of subsequent computations with these results.

5. APPLICATION TO INVERSION OF NONSYMMETRIC MATRIX

Although the method of inversion described above applies to the symmetric, positive definite matrices associated with the normal equations of geodesy, it can also be used to invert a nonsymmetric matrix with real coefficients. To solve the equations

$$Ax = \ell \tag{16}$$

where A is such a nonsymmetric matrix, premultiply both sides of (16) by A^T:

$$A^T A x = A^T \ell \tag{17}$$

The product $A^T A$ is of the type which we have designated by N and which can be factored into CC^T and inverted. Premultiplying (17) with N^{-1} found in this manner gives

$$x = N^{-1} A^T \ell \tag{18}$$

as the solution to (16) and shows that the inverse of the matrix A, if it exists, is

$$A^{-1} = N^{-1} A^T \tag{19}$$

6. EQUIVALENCE OF THE SYMMETRIC SOLUTION WITH THE LEAST SQUARES POSTULATE

Since an inverse is defined only for square nonsingular matrices, the assumption is implicit in (16) that this is a set of independent linear equations with an equal number of unknowns x which has therefore a unique solution.

The process of symmetrization used to form (17), when applied to a nonsquare matrix A, leads to some interesting and rather unexpected results.

If A in (16) is not square, i.e., if there are more equations than unknowns, or vice versa, then A is not invertible, corresponding to the well-known fact from linear algebra that no set of x's or an infinity of such sets will satisfy the equations. This raises the question of what legitimate operation on the equations (16) can produce a form with an invertible matrix for the coefficients of x. The problem is analogous to the purely formal device of introducing an integrating factor into a differential equation or, more basically, of multiplying the algebraic equation $ax = b$ by the reciprocal of a.

Assuming $\underset{mn}{A}$ to have dimensions $m \times n$, with $m > n$, then by matrix algebra, if $\underset{mn}{A}$ is premultiplied by a matrix having n rows and m columns the resulting product will be a square matrix which is, with certain known exceptions, nonsingular and hence invertible. The obvious choice for such an "inversion factor" for the equations (16) is A^T, the transpose of A, since it introduces a minimum of extraneous information into the problem—less than, for example, an arbitrary matrix M with dimensions $n \times m$. Premultiplying the equation

$$\underset{mn}{A}\ \underset{n1}{x} = \underset{m1}{\ell} \tag{20}$$

on both sides by $\underset{nm}{A^T}$ we obtain

$$\underset{n\ n}{(A^T A)}\ \underset{n1}{x} = \underset{n1}{(A^T \ell)}$$

with the unique solution

$$x = (A^T A)^{-1} A^T \ell \tag{21}$$

obtained by purely formalistic considerations and with a minimum of additional assumptions.

A. Observation Equations

In the calculus of observations of directly measured functions of linearized variables we are faced with the identical problem of solving the so-called *observation equations* or error equations, of the form (20), linear in the unknowns or corrections to unknowns x, whose number n is exceeded by the number of equations (observations) to be satisfied. The adjustment of triangulation by variation of coordinates is an example of this type of computation. The interpretation of the individual quantities $a_i x - \ell_i$ in each equation (20) to be a residual $v_i = a_i x - \ell_i$ for the measured function corresponding to fixed and sufficiently close values of the unknowns x in all these equations, together with the condition that Σv^2 be a minimum, also leads to the solution (21). We can conclude, therefore, that the purely formalistic considerations leading to (21) are equivalent to the least squares hypothesis which was in no way implicit in our assumptions. This shows the least squares postulate to be an irreducible hypothesis.

B. Condition Equations

Similar conclusions are reached in the alternative and equivalent method of adjustment by indirect observations or with so-called condition equations. The typical set of equations to be solved in this type of adjustment is

$$\underset{nm}{B}\ \underset{m1}{v} = \underset{n1}{\ell} \tag{22}$$

with $m > n$ and again subject to the condition $Ev^2 =$ minimum. Clearly the equations (22) by themselves are not sufficiently restrictive to yield an unambiguous solution, since $m - n$ independent conditions could be added to the set (22) before a solution for the v's becomes unique. Seeking the simplest formalistic solution for this case without postulating the least squares condition, we see that premultiplication of (22) by B^T will not work because the product $B^T B$ with $m > n$ will be necessarily

singular and not invertible. However, BB^T will be of dimension $n \times n$ and will possess an inverse $(BB^T)^{-1}$ if the conditions (22) are independent. It is not difficult to see that the simplest way to introduce B^T as a factor after B is to make a legitimate transformation of the variable v:

$$\underset{m1}{v} = \underset{mn}{B^T} \; \underset{n1}{k}$$

resulting in the conditions

$$BB^T k = \ell \tag{23}$$

equivalent to (22) and having the unique solution

$$k = (BB^T)^{-1}\ell$$

so that

$$v = B^T k = B^T(BB^T)^{-1}\ell$$

This, likewise, is the Gaussian least squares solution for "condition" equations.

7. THE GAUSSIAN ALGORITHM FOR SYMMETRIC MATRICES

Before the advent of electronic calculators and computers the labor of root extraction prevented the Cholesky factorization, with its advantages due to symmetry, from displacing in practice the standard Gauss-Doolittle solution for normal equations. For comparison, we show the simple relation between the two.

From the classical development of the Cholesky algorithm from Gauss-Doolittle, it is known that Gauss divides each reduced equation by the corresponding diagonal term, thus making each divided and reduced diagonal term equal to unity. Cholesky, on the other hand, divides by the square root of these diagonals. The matrix of coefficients of the undivided reduced equations is in each case the same upper triangular matrix with diagonal terms $d_1, d_2, \ldots d_n$. If the diagonal matrix whose entries are these d's is designated D, a corresponding diagonal matrix consisting of entries $\sqrt{d_1}$, $\sqrt{d_2}$, $\ldots \sqrt{d_n}$ can be designated $D^{1/2}$. The relation between the divided Gaussian upper triangular matrix G^T, with diagonal terms each equal to 1, and the corresponding Cholesky matrix C^T can then be written as

$$C^T = D^{1/2}G^T \tag{24}$$

because of the theorem that premultiplication with a diagonal matrix multiplies all entries of a row of the matrix being multiplied with the corresponding entry of the diagonal matrix. From (24) follows $C = G(D^{1/2})^T = GD^{1/2}$ and

$$N = CC^T = GD^{1/2}D^{1/2}G^T = (GD)G^T \tag{25}$$

This, together with (24), gives the Gaussian factorization in terms of the Cholesky factors C and C^T.

The Gaussian factorization algorithm can also be established, independent of the Cholesky factorization, by matrix algebra. Like the method of section 1 we postulate a given symmetric matrix N to be, according to (25), the product GDG^T, where G is the lower triangular matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ g_{12} & 1 & 0 & \cdot & \cdot & 0 \\ g_{13} & g_{23} & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 0 \\ g_{1n} & g_{2n} & g_{3n} & \cdot & \cdot & 1 \end{pmatrix}$$

and D the diagonal matrix

$$D = \begin{pmatrix} d_{11} & & & & & 0 \\ & d_{22} & & & & \\ & & d_{33} & & & \\ & & & & & \\ & & & & & \\ 0 & & & & & d_{nn} \end{pmatrix}$$

and G^T the transpose of G.

By actual multiplication the product $N = GDG^T$ is found to be the symmetrical matrix

$$N = \begin{pmatrix} \underline{d_{11}} & d_{11}g_{12} & d_{11}g_{13} & \cdot & d_{11}g_{1n} \\ & \underline{d_{11}g_{12}g_{12} + d_{23}} & d_{11}g_{13}g_{12} + d_{22}g_{23} & \cdot & d_{11}g_{1n} + d_{22}g_{2n} \\ & & \underline{d_{11}g_{13}g_{13} + d_{22}g_{23}g_{23} + d_{33}} & \cdot & d_{11}g_{1n}g_{1n} + d_{22}g_{2n}g_{2n} + d_{33}g_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ & & & & \underline{\qquad\qquad} \end{pmatrix} \tag{26}$$

Conversely, if a symmetric matrix N with elements n_{ik} is given we can find, by the algebraic method of undetermined coefficients, the d_{ii} and g_{ik} in (26) in sequence, computing each row in turn. This approach leads to the same sequence of operations and results specified by Gauss (1811a)* and codified by Doolittle (U.S. Coast and Geodetic Survey, 1881). Formulas analogous to (2) and (3) of section 1 are somewhat more cumbersome than for the Cholesky factorization and are not given here.

From (25) the inverse of N is

$$N^{-1} = (G^T)^{-1} D^{-1} G^{-1}$$

a product which requires the inversion of a triangular matrix G and the simple inverse of the diagonal matrix D.

8. THE GAUSSIAN ALGORITHM FOR NONSYMMETRIC MATRICES

The term "Gaussian elimination" is commonly reserved for Gauss's method of reducing a system of linear equations with a nonsymmetric square matrix of coefficients to triangular form. To establish simply the procedure to be followed in this reduction, it is again convenient to consider the product AB^T of two triangular matrices A and B^T where A is the lower triangular matrix

$$A = \begin{pmatrix} a_{11} & 0 & 0 & \cdot & \cdot & 0 & \cdot & \cdot & 0 \\ a_{21} & a_{22} & 0 & \cdot & \cdot & 0 & \cdot & \cdot & 0 \\ a_{31} & a_{32} & a_{33} & \cdot & \cdot & 0 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & a_{j3} & \cdot & \cdot & a_{jj} & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} & \cdot & \cdot & a_{nj} & \cdot & \cdot & a_{nn} \end{pmatrix} \tag{27}$$

and B^T an upper triangular matrix

$$B^T = \begin{pmatrix} 1 & b_{12} & b_{13} & \cdot & \cdot & b_{1k} & \cdot & \cdot & b_{1n} \\ 0 & 1 & b_{23} & \cdot & \cdot & b_{2k} & \cdot & \cdot & b_{2n} \\ 0 & 0 & 1 & \cdot & \cdot & b_{3k} & \cdot & \cdot & b_{3n} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & 1 & \cdot & \cdot & b_{kn} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & 0 & \cdot & \cdot & 1 \end{pmatrix}$$

whose diagonal terms equal 1 in conformance with the Gauss equations. Together the matrices (27) contain the necessary n^2 parameters to correspond with those of an arbitrary $n \times n$ matrix. Actual multiplication of AB^T yields the matrix M of (28).

$$M = AB^T = \begin{pmatrix} \underline{a_{11}} & a_{11}\underline{b_{12}} & a_{11}\underline{b_{13}} & \cdot\ a_{11}\underline{b_{ik}} & \cdot\ a_{11}\underline{b_{1n}} \\ \underline{a_{21}} & a_{21}b_{12}+\underline{a_{22}} & a_{21}b_{13}+a_{22}\underline{b_{23}} & \cdot\ a_{21}b_{1k}+a_{22}\underline{b_{2k}} & \cdot\ a_{21}b_{1n}+a_{22}\underline{b_{2n}} \\ \underline{a_{31}} & a_{31}b_{12}+\underline{a_{32}} & a_{31}b_{13}+a_{32}b_{23}+\underline{a_{33}} & \cdot\ a_{31}b_{1k}+a_{32}b_{2k}+a_{33}\underline{b_{3k}} & \cdot\ a_{31}b_{1n}+a_{32}b_{2n}+a_{33}\underline{b_{3n}} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \underline{a_{j1}} & a_{j1}b_{12}+\underline{a_{j2}} & a_{j1}b_{13}+a_{j2}b_{23}+\underline{a_{j3}} & \underline{a_{j1}b_{1k}+a_{j2}b_{2k}+a_{j3}b_{3k}+\cdots} & \cdot\ a_{j1}b_{1n}+a_{j2}b_{2n}+\ldots+\underline{a_{jj}b_{jn}} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \underline{a_{n1}} & a_{n1}b_{12}+\underline{a_{n2}} & a_{n1}b_{13}+a_{n2}b_{23}+\underline{a_{n3}} & \cdot\ a_{n1}b_{1k}+a_{n2}b_{2k}+a_{n3}b_{3k}+\ldots+\underline{a_{nk}} & a_{n1}b_{1n}+a_{n2}b_{2n}+\ldots+\underline{a_{nn}} \end{pmatrix} \tag{28}$$

$$m_{jk} = a_{j1}b_{1k} + a_{j2}\,b_{2k} + a_{j3}b_{3k} + \ldots$$
$$= \sum_{i=1}^{i=\ell} a_{ji}b_{ik} \qquad \text{where } \ell = \text{lesser of } j,\ k$$
$$b_{ii} = 1$$

*This reference contains a printer's error in a very essential formula which Gauss corrected in an addendum (1811b). Bertrand (1855) copies the erroneous formulas, which may explain why Doolittle was the only one in the geodetic community to follow the elegant symmetric approach of the earlier Gauss work.

This is the makeup of the given square matrix to be factored into the triangular matrices A and B^T. The elements b of the B^T matrix are developed in the rows above the diagonal, and the a's of A in the columns on and below the diagonal. The elements should be recorded, as they are computed (28), in their proper relative position as indicated by the nnderlining. The computation sequence is first column and row, second column and row beginning with the diagonal term, etc.

A reduction with a 4×4 matrix should make clear the necessary steps in the reduction which will be found identical with the Gaussian elimination process. In general, when the term m_{jk} in the jth row and kth column, $j \geqslant k$, is being reduced, the set of a's, $k-1$ in number, $a_{j1}, a_{j2} \ldots a_{j,k-1}$ will have been computed and will occupy the spaces in the same row and preceding m_{jk}. Similarly, the column extending above m_{jk} will contain the j-1 b's: b_{1k}, $b_{2k}, \ldots b_{j-1,k}$. The sum of the products of the first $(k-1)$ a's, each multiplied with the corresponding b from the colnmn set, is subtracted from m_{jk}, leaving the answer a_{jk}, since the last term in m_{jk} is $a_{jk}b_{kk}$ and $b_{kk}=1$ by definition. For $k>j$, i.e., in the portion of the matrix above the diagonal, the sum of $j-1$ such products is subtracted from m_{jk}, leaving $a_{jj}b_{jk}$ from which follows b_{jk} by division with a_{jj}, already computed. When the factorization is complete, the matrix multiplication AB^T should equal the given matrix M for a check on the numerical work.

If the equations to be solved by this algorithm are

$$Mx=\ell \tag{29}$$

then considerations similar to those of section 4 will show that extending the algorithm to the constant column ℓ will produce a vector ℓ' satisfying a system

$$B^Tx=\ell' \tag{30}$$

equivalent to (29) and in triangular form, ready for a Gaussian back solution.

9. SQUARE ROOT FACTORIZATION

The Cholesky modification of the Gauss algorithm for solving linear equations with a symmetric positive definite matrix can be readily generalized to parallel tbe general case of Gaussian elimination of section 8. The factorization with real numbers is again possible if the given nonsymmetric matrix is positive definite.

The assumed factors are

$$A_1=\begin{pmatrix} a_{11} & 0 & 0 & \cdot & 0 & \cdot \\ a_{21} & a_{22} & 0 & \cdot & 0 & \cdot \\ a_{31} & a_{32} & a_{33} & \cdot & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{j1} & a_{j2} & a & \cdot & a_{jj} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}$$

$$a_2^T=\begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdot & a_{1k} & \cdot \\ 0 & a_{22} & a_{23} & \cdot & a_{2k} & \cdot \\ 0 & 0 & a_{33} & \cdot & a_{3k} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & a_{kk} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{31}$$

which differ, in essence, from (27) only in that both matrices have identical diagonal entries. The product $A_1A_2^T$ gives

$$M=\begin{pmatrix} a_{11}a_{11} & a_{11}a_{12} & a_{11}a_{13} & \cdot & a_{11}a_{1k} & \\ a_{21}a_{11} & a_{21}a_{12}+a_{22}a_{22} & a_{21}a_{13}+a_{22}a_{23} & \cdot & a_{21}a_{1k}+a_{22}a_{2k} & \\ a_{31}a_{11} & a_{31}a_{12}+a_{32}a_{22} & a_{31}a_{13}+a_{32}a_{23}+a_{33}a_{33} & \cdot & a_{31}a_{1k}+a_{32}a_{2k}+a_{33}a_{3k} & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{j1}a_{11} & a_{j1}a_{12}+a_{j2}a_{22} & a_{j1}a_{13}+a_{j2}a_{23}+a_{j3}a_{33} & \cdot & a_{j1}a_{1k}+a_{j2}a_{2k}+\cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{32}$$

The general entry in the given product matrix is

$$m_{jk} = \sum_{i=1}^{i=n} a_{ji}a_{ik} \tag{33}$$

(where $i=1, 2, \ldots n$, n being the lesser of j, k) as obtained by multiplying the jth row vector of A_1 with the kth column vector of A_2^T. When the individual a's are evaluated in the sequence used in the Gaussian elimination of section 8, all the terms except the last in the sum (33) representing m_{jk} will be known. For $j > k$, m_{jk} ends with $a_{jk}a_{kk}$, with all a's known except a_{jk}.

When all the a_{jk} have been computed, the two triangular matrices of (31) will be known and can be inverted individually. From the assumption $M = A_1A_2^T$ follows

$$M^{-1} = (A_2^T)^{-1}A_1^{-1} \tag{34}$$

and the solution of a system of equations $Mx = \ell$ is $x = M^{-1}\ell$. The inverse (34), when computing with fixed decimal point, can be made more precise by a method analogous to that of section 3. In general the relation $A_1^{-1}M(A_2^T)^{-1} = I$ will not be satisfied numerically exactly, but will produce a result

$$A_1^{-1}M(A_2^T)^{-1} = I^* \tag{35}$$

where I^* has nonsymmetric small off-diagonal terms. It is of the same type as M but much more diagonal and can therefore be factored very precisely by the same algorithm used for factoring (32) into $I^* = A_1^*(A_2^*)^T$. Inverting these two triangular matrices and introducing the result in (35) produces a near identity

$$(A_1^*)^{-1}A_1^{-1}M(A_2^T)^{-1}((A_2^*)^T)^{-1} = I$$

or, designating the products $(A_1^*)^{-1}A_1^{-1}$ as A_1^{-1} and $(A_2^T)^{-1}((A_2^*)^T)^{-1}$ as $(A_2^T)_1^{-1}$,

$$A_1^{-1}M(A_2^T)^{-1} = I$$

These improved values of A_1^{-1} and $(A_2^T)^{-1}$ substituted in the right-hand side of (34) result in the improved inverse of M.

This method of factorization and inversion has the same advantages over the classical Gaussian elimination of section 8 that the original Cholesky method has over the Gauss-Doolittle solution for symmetric matrices. On the whole, however, the method of symmetrization described in section 5 seems preferable to either of these two in terms of simplicity, generality, and economy of computation.

REFERENCES

Bertrand, J. *Méthode des moindres carrés. Memoires sur la combinaison des observations*, Par ch.-Fr. Ganss. Mallet-Bachelier, Paris, 1855.

Gauss, C. F. Disqnisitio de elementis ellipticis palladis. *Commentationes Societatis Regiae Scientarium Gottingensis*, Göttingen, 1811a.

Gauss, C. F. (Announcement), *Von Zachs Monatliche Correspondenz zur Beförderung der Erd und Himmelskunde*, vol. 24, p. 462, Gotha, 1811b.

U.S. Coast and Geodetic Survey. *Report of the Superintendent of the U.S. Coast and Geodetic Survey showing the progress of the work during the fiscal year ending with June 1878.* Government Printing Office, Washington, D.C., 1881.

U. S. GOVERNMENT PRINTING OFFICE : 1973 OL - 491-890

(Continued from inside front cover)

*C&GS 36 Geodetic and Grid Angles--State Coordinate Systems. Lansing G. Simmons, January 1968.

*C&GS 37 Precise Echo Sounding in Deep Water. George A. Maul, January 1969.

*C&GS 38 Grid Values of Total Magnetic Intensity IGRF--1965. E. B. Fabiano and N. W. Peddie, April 1969.

C&GS 39 An Advantageous, Alternative Parameterization of Rotations for Analytical Photogrammetry. Allen J. Pope, September 1970. Price $0.30 (COM-71-00077)

C&GS 40 A Comparison of Methods for Computing Gravitational Potential Derivatives. L. J. Gulick, September 1970. Price $0.40 (COM-71-00185)

NOAA TECHNICAL REPORTS

NOS 41 A User's Guide to a Computer Program for Harmonic Analysis of Data at Tidal Frequencies. R. E. Dennis and E. E. Long, July 1971. Price $0.65 (COM-71-50606)

NOS 42 Computational Procedures for the Determination of a Simple Layer Model of the Geopotential From Doppler Observations. Bertold U. Witte, April 1971. Price $0.65 (COM-71-50400)

NOS 43 Phase Correction for Sun-Reflecting Spherical Satellite. Erwin Schmid, August 1971. Price $0.25 (COM-72-50080)

NOS 44 The Determination of Focal Mechanisms Using P- and S-Wave Data. William H. Dillinger, Allen J. Pope, and Samuel T. Harding, July 1971. Price $0.60 (COM-71-50392)

NOS 45 Pacific SEAMAP 1961-70 Data for Area 15524-10: Longitude 155°W to 165°W, Latitude 24°N to 30°N, Bathymetry, Magnetics, and Gravity. J. J. Dowling, E. E. Chiburis, P. Dehlinger, ad M. J. Yellin, January 1972. Price $3.50 (COM-72-51029)

NOS 46 Pacific SEAMAP 1961-70 Data for Area 15530-10: Longitude 155°W to 165°W, Latitude 30°N to 36°N, Bathymetry, Magnetics, and Gravity. J. J. Dowling, E. F. Chiburis, P. Dehlinger, and M. J. Yellin, January 1972. Price $3.50

NOS 47 Pacific SEAMAP 1961-70 Data for Area 15248-14: Longitude 152°W to 166°W, Latitude 48°N to 54°N, Bathymetry, Magnetics, and Gravity. J. J. Dowling, E. F. Chiburis, P. Dehlinger, and M. J. Yellin, April 1972. Price $3.50 (COM-72-51030)

NOS 48 Pacific SEAMAP 1961-70 Data for Area 16648-14: Longitude 166°W to 180°, Latitude 48°N to 54°N, Bathymetry, Magnetics, and Gravity. J. J. Dowling, E. F. Chiburis, P. Dehlinger, and M. J. Yellin, April 1972. Price $3.00 (COM-72-51028)

NOS 49 Pacific SEAMAP 1961-70 Data for Areas 16530-10 and 17530-10: Longitude 165°W to 180°, Latitude 30°N to 36°N, Bathymetry, Magnetics, and Gravity. E. F. Chiburis, J. J. Dowling, P. Dehlinger, and M. J. Yellin, July 1972. Price $4.75

NOS 50 Pacific SEAMAP 1961-70 Data for Areas 16524-10 and 17524-10: Longitude 165°W to 180°, Latitude 24°N to 30°N, Bathymetry, Magnetics, and Gravity. E. F. Chiburis, J. J. Dowling, P. Dehlinger, and M. J. Yellin, July 1972. Price $5.75

NOS 51 Pacific SEAMAP 1961-70 Data for Areas 15636-12, 15642-12, 16836-12, and 16842-12: Longitude 156°W to 180°, Latitude 36°N to 48°N, Bathymetry, Magnetics, and Gravity. E. F. Chiburis, J. J. Dowling, P. Dehlinger, and M. J. Yellin, July 1972.

NOS 52 Pacific SEAMAP 1961-70 Data Evaluation Summary. P. Dehlinger, E. F. Chiburis, and J. J. Dowling, July 1972. Price $0.40

NOS 53 Grid Calibration by Coordinate Transfer. Lawrence W. Fritz, November 1972.

NOS 54 A Cross-Coupling Computer for the Oceanographer's Askania Gravity Meter. Carl A. Pearson and Thomas E. Brown, November 1972.

NOS 55 A Mathematical Model for the Simulation of a Photogrammetric Camera Using Stellar Control. Chester C Slama, December 1972.