On the Simultaneous Numerical Inversion of a Matrix and All its Leading Submatrices

This note considers the case in which a square matrix A may be factorized

(1) $$A = X'DY,$$

where X and Y are square matrices whose diagonal elements (where $i = j$) are unity, and whose subdiagonal elements (where $i > j$) are zero, whilst D is a diagonal matrix, that is, one whose off-diagonal elements (where $i \neq j$) are zero [1]. Where this factorization exists, it is unique [1]. Where it does not exist, there exists another factorization of the form (1) in which the matrices X and Y differ from the foregoing specification by having their rows in some way interchanged [2].

Let the matrices appearing in relation (1) all be partitioned after the rth row and column, the resultant leading submatrices of order r by r being denoted by A_r, X_r', D_r, Y_r, where $r = 1, 2, \cdots, n - 1$. Then it is easily seen that

$$A_r = X_r'D_rY_r.$$

If D_r is non-singular, so is A_r, and hence

(2) $$A_r^{-1} = Y_r^{-1}D_r^{-1}X_r'^{-1}.$$

This holds also for $r = n$, when suffixes may be omitted. It is, in fact, equivalent to taking $r = n$ and replacing the elements in the last $n - r$ columns and rows of all matrices by zero. This suggests that if a typical element, say b_{ij}, of $A^{-1} = B$ is obtained as the scalar product of the ith row of $Y^{-1}D^{-1}$ and the jth row of X^{-1}, the corresponding element $b_{ij(r)}$ (where $i, j \leqslant r$) of $A_r^{-1} = B_r$, say, is obtained as the rth partial sum of the n products of the row-by-row multiplication. Thus, with a desk calculating machine, on which products can be accumulated, the inverses $B_1, B_2, \cdots$, of $A_1, A_2, \cdots$ are obtained simply as the partial sums of the row-by-row multiplications needed anyhow for the computation of the inverse B of A.

As the factorization is unique any process yielding X', D, and Y (or $Y^{-1}D^{-1}$ and X'^{-1}, etc.) may be used. The following arithmetic check is suggested. Relation (1) gives

$$X'^{-1}A = DY.$$

Comparison of diagonal elements on either side shows that the scalar product of the rth column of X_r^{-1} into the rth column of A_r is equal to d_r, the rth diagonal element of D_r. This relation may be used to check each column of X^{-1} fresh after computation. Similarly it is found that the scalar product of the rth column of Y_r^{-1} into the rth row of A_r is also equal to d_r.

H. Nagler

25 Helenslea Avenue
London N.W.11
England

1. A. M. TURING, "Rounding-off errors in matrix processes," *Quart. Jour. Mech. and Appl. Math.*, v. 1, 1948, p. 289.
2. R. A. FRAZER, W. J. DUNCAN, & A. R. COLLAR, *Elementary Matrices and Some Applications to Dynamics and Differential Equations*, Cambridge Univ. Press, 1938, p. 99.

A Note on Large Linear Systems

Towards the end of 1955 the author of the present note was charged by Svenska Aeroplan Aktiebolaget (SAAB Aircraft Company), with setting up a program for the automatic computer BESK for the solution of large symmetric systems of linear algebraic equations. The reason was that the solution of 4 systems with 214 unknowns was required. The systems had been set up in conjunction with certain calculations for airplane structural analysis. Approximately 40% of the coefficients in the matrices of these systems differed from zero. This fact was taken into account in setting up the program. Nevertheless, the whole system could not be accommodated at one time in the magnetic drum memory of 8192 words associated with BESK, and a splicing method was developed which does not involve excessive loss of time in connection with the input and output of intermediate results. The four systems were solved simultaneously by the BESK. The method used for solving the equations was Gauss' method of elimination. The calculating time was about 2 hours. If p% of the coefficients of the systems differ from zero, and the order of the systems is n, the calculating time will be proportional to p^2n^3.

By formation of residuals and one iteration a solution was obtained which, when inserted into the original systems, proved to satisfy the equations with deviations less than 10^{-6}%, i.e., if the computed solution was $\bar{x}_j$ and the system was $AX = B$, the "error vector" $E = \sum_j a_{ij}\bar{x}_j - b_i$ satisfied $|E| < 10^{-8}|B|$. The systems were badly conditioned since, if the matrix was (a_{ik}), the "condition number" was:

$$R(A) = \frac{|\text{determinant } (a_{ik})|}{\prod_{k=1}^{214}\left(\sum_{i=1}^{214} |a_{ik}|\right)} = 2^{-886}.$$

The matrix is therefore comparable with the $n \times n$ matrix

$$A_n = \begin{bmatrix} 5 & -4 & 1 & \cdot & \cdot & \cdot & & & \\ -4 & 6 & -4 & 1 & \cdot & \cdot & & & \\ 1 & -4 & 6 & -4 & 1 & \cdot & & & \\ \cdot & 1 & -4 & 6 & -4 & 1 & & & \\ & & & & \cdot & & & & \\ & & & & & \cdot & & & \\ & & & & & & \cdot & & \\ & & & & 1 & -4 & 6 & -4 & 1 \\ & & & & \cdot & 1 & -4 & 6 & -4 \\ & & & & \cdot & \cdot & 1 & -4 & 5 \end{bmatrix}$$