ENLARGEMENT METHODS FOR COMPUTING THE INVERSE MATRIX

BY LOUIS GUTTMAN

Cornell University

1. Summary. The enlargement principle provides techniques for inverting any nonsingular matrix by building the inverse upon the inverses of successively larger submatrices. The computing routines are relatively easily learned since they are repetitive. Three different enlargement routines are outlined: first-order, second-order, and geometric. None of the procedures requires much more work than is involved in squaring the matrix.

2. Introduction. A set of methods is presented here for computing the inverse matrix, based on what we shall call an *enlargement principle.* The principle is to build the inverse upon the inverses of successively larger submatrices. This leads to simple repetitive routines that are not unlike iterative steps, but afford a direct solution.

The basis for such routines has also been noticed before,[1] but does not seem to have attracted the attention it merits. A possible reason for this lack of attention may be the belief that the methods apply only to a restricted class of matrices. We establish a simple lemma in this paper which shows that the enlargement methods apply to *all* nonsingular matrices, so that their use is perfectly general.

The enlargement principle may be considered an opposite of the "condensation" principle that governs Gauss' method of elimination and its variants such as the Doolittle procedure and Aitken's "pivotal condensation."[2] It is interesting that the same formula upon which the enlargement methods are based can also serve as a foundation for the condensation methods, as is shown in section 7 below.

The enlargement methods have the following characteristics:

(1) The first-order procedure outlined in the next section has been learned by statistical clerks in about ten minutes. People who calculate inverses only occasionally and forget the process between times should find the method as economical as those who must constantly compute inverses.

(2) They are direct methods, and yield an exact answer with not much more work than is involved in squaring the matrix.

(3) They can be adapted to electric punch-card systems, which will be efficient when very large matrices are to be inverted.

---

[1] It has appeared earlier in [2]. Waugh's recent note [10] also rediscovers the basic formula although only a specialized use is suggested there. Professor Harold Hotelling has called my attention to reference [1], which overlaps substantially with the present paper, and to a use of an enlargement approach to computing latent roots and vectors [9]. I am also indebted to Professor Hotelling for other helpful comments on the present paper.

[2] For an excellent summary and bibliography of direct and iterative methods for computing the inverse matrix see ([5], [6]).

(4) A sequence of inverses is yielded. Exact inverses of successively larger submatrices are computed in the routines, and these inverses are often themselves of interest. For correlation problems, this means that a sequence of sets of successively higher order multiple correlation constants is produced routinely.

(5) The general formula upon which the methods are based allows many variations in procedure, so that special adaptations can be easily made for special matrices.

A "first-order" enlargement procedure for computing the inverse matrix will be outlined in the next section. The proof for the method follows from the general formula in section 4. This procedure and formula are also described in [2]. Other enlargement routines are described in subsequent sections. Some additional formulas of relevance are discussed in section 8.

3. First-order enlargement. Let the matrix whose inverse is desired be

$$A_n = \begin{Vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{Vmatrix}.$$

The following sequence of successively larger principal submatrices will be assumed to be nonsingular:

$$A_2 = \begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix}, \quad A_3 = \begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \cdots, A_n.$$

If necessary, the rows and columns of A_n can always be shifted to obtain such a sequence. The following additional notation will be used:

$$B_i = (a_{1,i+1}\, a_{2,i+1} \cdots a_{i,i+1})$$
$$C_i = (a_{i+1,1}\, a_{i+1,2} \cdots a_{i+1,i})$$
$$d_i = a_{i+1,i+1}.$$

Thus, we can write

$$A_{i+1} = \begin{Vmatrix} A_i & B_i' \\ C_i & d_i \end{Vmatrix}, \quad (i = 2, 3, \cdots, n-1).$$

The first-order enlargement procedure is to compute in turn $A_2^{-1}, A_3^{-1}, \cdots, A_n^{-1}$. The inverse of A_2 is computed by the traditional steps:

{1} Compute $\Delta = a_{11}a_{22} - a_{21}a_{12}$, and compute $1/\Delta$.

{2} Then

$$A_2^{-1} = \begin{Vmatrix} \Delta^{-1}a_{22} & -\Delta^{-1}a_{12} \\ -\Delta^{-1}a_{21} & \Delta^{-1}a_{11} \end{Vmatrix}.$$

Remember that $B_2 = (a_{13}\ a_{23})$, $C_2 = (a_{31}\ a_{32})$, and that $d_2 = a_{33}$. The steps for computing A_3^{-1} are as follows:

{3} Compute $E_2' = A_2^{-1}B_2'$.

{4} Compute $f_2 = d_2 - C_2E_2'$.

{5} Compute $1/f_2$.

{6} Compute $G_2' = f_2^{-1}E_2'$, and compute $H_2 = f_2^{-1}C_2A_2^{-1}$.

{7} To each element in A_2^{-1} add the product of the corresponding elements in E_2 and H_2 to form $K_2 = A_2^{-1} + E_2'H_2$.

Then the third order inverse is

$$A_3^{-1} = \left\| \begin{matrix} K_2 & -G_2' \\ -H_2 & 1/f_2 \end{matrix} \right\|.$$

In general, to obtain A_{i+1}^{-1} from A_i^{-1}, $(i = 2, 3, \cdots, n - 1)$, imitate[3] steps {3} through {7}:

{3′} Compute $E_i' = A_i^{-1}B_i'$.

{4′} Compute $f_i = d_i - C_iE_i'$,

{5′} Compute f_i^{-1}.

{6′} Compute $G_i' = f_i^{-1}E_i'$, and compute $H_i = f_i^{-1}C_iA_i^{-1}$.

{7′} Compute $K_i = A_i^{-1} + E_i'H_i$. Then

$$A_{i+1}^{-1} = \left\| \begin{matrix} K_i & -G_i' \\ -H_i & 1/f_i \end{matrix} \right\|.$$

By repeated applications of steps {3′} through {7′} to the successively larger A_i^{-1}, A_n^{-1} is attained.

If A_n is symmetric, then almost half the work is saved, for then $B_i = C_i$, $G_i = H_i$, and K_i is symmetric, $(i = 2, 3, \cdots, n - 1)$.

To help gauge the amount of work needed to arrive at A_n^{-1}, let us compare it with the work that would be needed to square A_n. For the general asymmetric case, n^2 product sums of n terms each are required for A_n^2, a total of n^3 multiplications. With calculating machines, the sums of the products are accumulated, so that no separate process of addition is involved. To reach A_n^{-1} by the above enlargement method, $n^3 - n$ multiplications are required. Most of the addition is accomplished in the process by accumulative multiplication, but an additional $\frac{n(n-1)(2n-1)}{6} + n - 3$ terms have to be added otherwise. Furthermore, $n - 1$ reciprocal numbers are needed. Thus, A_n^{-1} involves somewhat less multiplications than does A_n^2, but needs more additions, as well as some reciprocal numbers.

---

[3] Actually, these steps could be used immediately in place of steps {1} and {2} to compute A_2^{-1}, by letting $i = 1$, and letting $A_1 = a_{11}$ (which may be assumed different from zero). The traditional method, however, is quicker for the 2x2 matrix.

In linear multiple correlation problems, if A_{i+1} is the correlation matrix of the first $i + 1$ variates, then E_i consists of the regression coefficients of the first i variates for predicting the $(i + 1)$th variate, and f_i is the square of the multiple correlation coefficient for this regression.

4. A lemma and the general formula. The enlargement procedure just outlined is one of many possible routines which can be developed from a general formula for the inverse matrix in partitioned form. This formula seems to have appeared first in [2], where it is stated that the method applies only to the cases where $f_i \neq 0$ in step {4}. We shall establish here a lemma that shows that this is no restriction, for the submatrix in step {4} is always nonsingular. Our lemma proves that the enlargement methods will invert *any* nonsingular matrix.

Let A_n be a nonsingular matrix of order n, partitioned in the form

$$A_n = \begin{Vmatrix} A & B' \\ C & D \end{Vmatrix}. \tag{1}$$

where A is of order m, $(1 \leq m < n)$, and will be assumed nonsingular. B and C are of $n - m$ rows and m columns, and D is of order $n - m$.

The following lemma is needed to show that enlargement methods will invert any nonsingular matrix:

LEMMA. *If in (1), both A_n and A are nonsingular, then the matrix*

$$F = D - CA^{-1}B' \tag{2}$$

is nonsingular.

For the proof, postmultiply the first submatric column of A_n by $A^{-1}B'$ and subtract from the second, leaving

$$M = \begin{Vmatrix} A & 0 \\ C & F \end{Vmatrix}.$$

M differs from A_n only by an elementary transformation; hence its rank is that of A_n. But clearly the rank of M is the sum of the ranks of A and F. Therefore, the rank of F is $n - m$, and F is nonsingular.

The inversion formula itself is the following identity:

$$\begin{Vmatrix} A & B' \\ C & D \end{Vmatrix}^{-1} = \begin{Vmatrix} A^{-1} + A^{-1}B'F^{-1}CA^{-1} & -A^{-1}B'F^{-1} \\ -F^{-1}CA^{-1} & F^{-1} \end{Vmatrix}. \tag{3}$$

A direct verification that the identity holds can be obtained by multiplying the right member in either direction by the right member of (1), yielding the unit matrix.

In section 3, the formula exhibited for A_{i+1}^{-1} at step {7′} is easily identified as a special case of formula (3) where $n = i + 1$, $m = i$. F corresponds to f_i, which is a scalar number; hence F^{-1} is easily computed in this case.

5. Second order enlargement. In formula (3), once A^{-1} is given, the rest of the work is essentially straightforward matric multiplication, except for computing F^{-1}. In section 3, F was easily inverted since it was of order unity. F can also be easily inverted if it is of order two, so that a *second order* enlargement procedure is feasible, computing A_{i+2}^{-1} from A_i^{-1}. The steps are similar to those in section 3 but involve larger matrices.

Letting A_i have the same meaning as in section 3, define now B_i, C_i, and D_i according to the partitioning

$$A_{i+2} = \begin{Vmatrix} A_i & B_i' \\ C_i & D_i \end{Vmatrix}.$$

Then B_i and C_i are of two rows and i columns, and D_i is of order two. Compute A_2^{-1} as in section 3. From then on, to compute A_{i+2}^{-1} from A_i, the steps are:

{3″} Compute $E_i' = A_i^{-1}B_i$.
{4″} Compute $F_i = D_i - C_iE_i'$.
{5″} Compute F_i^{-1} by steps [1] and [2] of section 3.
{6″} Compute $G_i' = F_i^{-1}E_i'$, and compute $H_i = F_i^{-1}C_iA_i^{-1}$.
{7″} Compute $K_i = A_i^{-1} + E_i'H_i$.

Then

$$A_{i+2}^{-1} = \begin{Vmatrix} K_i & -G_i \\ -H_i & F_i^{-1} \end{Vmatrix}.$$

If n is even, successive enlargements will lead A_n^{-1}. If n is odd, then A_{n-1}^{-1} is attained, from which A_n^{-1} can be computed according to section 3.

The number of multiplications and additions for this procedure is the same as for section 2. However, less writing is involved since only about half as many A_i are inverted. A disadvantage is that it is more complicated at each stage than is the procedure of section 3.

6. Geometric enlargement. Another routine is that which may be called *geometric* enlargement. Here, A_{2i}^{-1} is computed from A_i^{-1}. The steps may be described as follows. Letting A_i have the same meaning as previously, redefine B_i, C_i, and D_i according to the partitioning

$$A_{2i} = \begin{Vmatrix} A_i & B_i' \\ C_i & D_i \end{Vmatrix}.$$

Then B_i, C_i, and D_i are all, like A_i, square matrices of order i. Compute A_2^{-1} according to steps {1} and {2}, and compute A_4^{-1} according to steps {3″} through {7″}. From then on, to compute A_{2i}^{-1} from A_i^{-1}, the steps are formally the same as before, with a complication in step {5‴}:

{3'''} Compute $E_i' = A_i^{-1}B_i'$.
{4'''} Compute $F_i = D_i - C_iE_i'$.
{5'''} Compute F_i^{-1} by *geometric enlargement* in the same way as A_i^{-1}.
{6'''} Compute $G_i' = F_i^{-1}E_i'$, and compute $H_i = F_i^{-1}C_iA_i^{-1}$.
{7'''} Compute $K_i = A_i^{-1} + E_i'H_i$.

Then,

$$A_{2i}^{-1} = \left\| \begin{matrix} K_i & -G_i \\ -H_i & F_i^{-1} \end{matrix} \right\|.$$

This method involves less writing than the others, but is more complicated.

7. Condensation methods; special cases. Formula (3) also affords a basis for condensation methods by "back solution." For example, let A be of order m, where m is one or two so that A is easily inverted. Then F is of order $n - m$, and we will denote it by F_{n-m}. Partition F_{n-m} into the form

$$F_{n-m} = \left\| \begin{matrix} A^{(2)} & B'^{(2)} \\ C^{(2)} & D^{(2)} \end{matrix} \right\|$$

where $A^{(2)}$ is again of order m, defining F_{n-2m} of order $n - 2m$. Continue the process until an F_i is reached which is easily inverted, and solve backwards to reach F_{n-m}^{-1}, and then A_n^{-1}, by repeated use of (3).

Formula (3) is of great help in those special cases where A is large but easily inverted, such as a diagonal matrix, orthogonal matrix, etc. The labor can then be focussed on inverting an F which is much smaller than A_n.

8. Further identities. It is of some interest to exhibit some matric identities relevant to formula (3). Using the notation of section 4, let us seek the inverse of A_n partitioned in the form

$$A_n^{-1} = \left\| \begin{matrix} W & X' \\ Y & Z \end{matrix} \right\|. \tag{4}$$

An equation to be satisfied is

$$\left\| \begin{matrix} W & X' \\ Y & Z \end{matrix} \right\| \cdot \left\| \begin{matrix} A & B' \\ C & D \end{matrix} \right\| = \left\| \begin{matrix} I & 0 \\ 0 & I \end{matrix} \right\|,$$

which yields the equations

$$WA + X'C = I \tag{5}$$

$$WB' + X'D = 0 \tag{6}$$

$$YA + ZC = 0 \tag{7}$$

$$YB' + ZD = I. \tag{8}$$

If A and D are nonsingular, then from (6) and (7),

$$X' = -WB'D^{-1}, \qquad Y = -ZCA^{-1}. \tag{9}$$

Using (9) in (5) and (8), and remembering the lemma of section 4, we obtain

$$W = (A - B'D^{-1}C)^{-1}, \qquad Z = (D - CA^{-1}B')^{-1}. \tag{10}$$

Using (10) in (9) yields

$$X' = -(A - B'D^{-1}C)^{-1}B'D^{-1}, \qquad Y = -(D - CA^{-1}B')^{-1}CA^{-1}. \tag{11}$$

Putting (10) and (11) into (4) completes the formula

$$\left\| \begin{matrix} A & B' \\ C & D \end{matrix} \right\|^{-1} = \left\| \begin{matrix} (A - B'D^{-1}C)^{-1} & -(A - B'D^{-1}C)^{-1}B'D^{-1} \\ -(D - CA^{-1}B')^{-1}CA^{-1} & (D - CA^{-1}B')^{-1} \end{matrix} \right\|. \tag{12}$$

Comparing (3) with (12), we have the identities

$$(A - B'D^{-1}C)^{-1} = A^{-1} + A^{-1}B'(D - CA^{-1}B')^{-1}CA^{-1} \tag{13}$$

$$(A - B'D^{-1}C)^{-1}B'D^{-1} = A^{-1}B'(D - CA^{-1}B')^{-1}, \tag{14}$$

which may of course be verified by direct simplification.

An important feature of each of these identities is that the matrix in parentheses on the left is of order m, while that in parentheses on the right is of order $n - m$.

A special case of (13) was noticed by the writer [3], [4] and of (14) by Ledermann ([7], [8]) and the writer ([3], [4]), in connection with regression problems of factor analysis. In this special case, A is a diagonal matrix and hence easily inverted; $n - m$ is the number of common factors, which is usually small compared with m; the correlation matrix of m observed variates is given factored into the form $A - B'D^{-1}C$; and the work of inverting the correlation matrix of order m is simplified essentially into inverting a much smaller matrix.

It should be noticed that (12), (13), and (14) assume that both A and D are nonsingular, where (3) assumes only that A is nonsingular (since then F must be nonsingular from the lemma of section 4).

REFERENCES

[1] W. J. Duncan, "Some devices for the solution of large sets of simultaneous linear equations," *Phil. Mag.*, Vol. 35 (1944), pp. 660–670.

[2] R. A. Frazer, W. J. Duncan, and R. A. Collar, *Elementary Matrices*, Cambridge Univ. Press, 1938.

[3] Louis Guttman, "Multiple rectilinear prediction and the resolution into components," *Psychometrika*, Vol. 5 (1940), pp. 75–99.

[4] Louis Guttman and Jozef Cohen, "Multiple rectilinear prediction and the resolution into components II," *Psychometrika*, Vol. 8 (1943), pp. 169–183.

[5] Harold Hotelling, "Some new methods in matrix calculation," *Annals of Math. Stat.*, Vol. 14 (1943), pp. 1–34.

[6] HAROLD HOTELLING, "Further points on matrix calculation and simultaneous equations," *Annals. of Math. Stat.*, Vol. 14 (1943), pp. 440–441.
[7] WALTER LEDERMANN, "A shortened method for the estimation of mental factors by regression," *Nature*, Vol. 141 (1938), p. 246.
[8] WALTER LEDERMANN, "On a shortened method for the estimation of mental factors by regression," *Psychometrika*, Vol. 4 (1939), pp. 109–116.
[9] J. MORRIS AND J. W. HEAD, "Lagrangian frequency equations, an 'escalator' method for numerical solution," *Aircraft Eng.* Vol. 14 (1942), pp. 312–314, 316.
[10] F. V. WAUGH, "A note concerning Hotelling's method of inverting a partitioned matrix," *Annals of Math. Stat.*, Vol. 16 (1945), pp. 216–217.