METHODS OF MATRIX INVERSION

DONALD GREENSPAN, University of Maryland

Introduction. The physical scientist has long been plagued with large linear systems of equations, hopeless differential equations or systems of such equations, or simply enigmatic integral equations. Each of these problems in turn often lends itself to matrix analysis and the approach usually involves matrix inversion. However, it is only with the advent of high speed computation that the approach has entered the realm of practicality.

The purpose of this paper is to present some of the existing methods of matrix inversion in a fashion requiring a minimal mathematical background. ("Minimal mathematical background" implies knowing the definition of a matrix, how to add and multiply matrices, how to apply matrix notation to systems of linear equations, and how to evaluate a determinant. Most of this material is discussed in [27].)

In each of the following discussions, matrix A will be assumed non-singular and will be represented by:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \cdots a_{1n} \\ a_{21} & a_{22} & a_{23} \cdots a_{2n} \\ a_{31} & a_{32} & a_{33} \cdots a_{3n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{n1} & a_{n2} & a_{n3} \cdots a_{nn} \end{bmatrix}$$

The main diagonal of A will be that part of the matrix consisting of the elements a_{ii}, $i=1, 2, \cdots, n$; and the trace $[A]$ will be the sum of these elements. The unit matrix is one such that $a_{ii}=1$; $i=1, 2, \cdots, n$; and $a_{ij}=0$; $i\neq j$; $i, j=1, 2, 3, \cdots, n$. A is called symmetric if and only if $a_{ij}=a_{ji}$ for all i and j.

Method I—Elimination method. Multiply row 1 by $1/a_{11}$. Then multiply row 1 by $-a_{i1}$ and add to row i for $i=2, 3, 4, \cdots, n$. This sequence of operations has the effect of creating a 1 in the first position of the main diagonal and zeros everywhere else in the first column of the matrix. Apply the same technique to the second row, thus creating a 1 in the second position on the main diagonal and zeros everywhere else in the second column. Apply the same technique to the third row, again resulting in a column of $(n-1)$ zeros and a 1 in the main diagonal position. Keep up this procedure until the given matrix is reduced to the unit matrix. By performing the very same sequence of operations on the unit

matrix as were performed on A, using the very same numbers necessary when operating on A, the unit matrix is transformed into A^{-1}.
Example 1: Let

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then we write down:

$$\left[\begin{array}{rrr|rrr} 2 & -2 & 4 & 1 & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 1 \end{array}\right].$$

We multiply row 1 by $\frac{1}{2}$ and secure:

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 2 & \frac{1}{2} & 0 & 0 \\ 2 & 3 & 2 & 0 & 1 & 0 \\ -1 & 1 & -1 & 0 & 0 & 1 \end{array}\right].$$

By multiplying row 1 by -2 and adding to row 2 and multiplying row 1 by 1 and adding to row 3 we have:

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 5 & -2 & -1 & 1 & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right].$$

Multiply row 2 by $\frac{1}{5}$. Hence:

$$\left[\begin{array}{rrr|rrr} 1 & -1 & 2 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & -\frac{2}{5} & -\frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right].$$

The sequence of operations: multiply row 2 by 1 and add to row 1; multiply row 2 by 0 and add to row 3; multiply row 3 by 1; multiply row 3 by $-\frac{8}{5}$ and add to row 1; multiply row 3 by $\frac{2}{5}$ and add to row 2, yields:

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & 1 & 0 & 0 & \frac{1}{5} & \frac{2}{5} \\ 0 & 0 & 1 & \frac{1}{2} & 0 & 1 \end{array}\right].$$

Hence the inverse of

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} \quad \text{is} \quad \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

The only problem which can arise, assuming no loss of significance to make the matrix at any stage singular, is that one of the a_{ii} becomes 0. In this case, one need only interchange two rows so that the zero on the main diagonal is removed from its choice position. In order to get A^{-1} then, one need only take the final result and interchange in it the same numbered columns as those of the rows which were interchanged.

A technique like this is called an elimination technique. Application of it to a large scale matrix should be accompanied by an examination of [**30**].

Method II. Consider the matrix B such that:

$$B = \left[\begin{array}{c|c} I & A \\ \hline 0 & I \end{array}\right],$$

where A is the matrix to be inverted, I is the identity matrix, and 0 the matrix composed of all zero elements. Consider the following operations:

a) multiply any row or any column by a non-zero constant.
b) multiply any row (column) by a non-zero constant and add to another row (column).
c) interchange any two rows or any two columns.

Using these operations, we reduce

$$B \to \left[\begin{array}{c|c} P & I \\ \hline 0 & Q \end{array}\right].$$

Then $A^{-1} = QP$.
Example 2:

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then

$$B = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & -2 & 4 \\ 0 & 1 & 0 & 2 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & -1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 0 & 0 & -1 & 1 & -1 & 1 \\ 0 & 1 & 0 & 2 & 3 & 2 \\ 1 & 0 & 0 & 2 & -2 & 4 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 2 & 5 & 0 \\ 1 & 0 & 0 & 2 & 0 & 2 \\ \hline 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \rightarrow \left[\begin{array}{ccc|ccc} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & 0 & 5 & 0 \\ 1 & 0 & 2 & 0 & 0 & 2 \\ \hline 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\rightarrow \left[\begin{array}{ccc|ccc} 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & 0 & 1 & 0 \\ \frac{1}{2} & 0 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right].$$

Hence

$$P = \begin{bmatrix} 0 & 0 & -1 \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}, \quad Q = \begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$A^{-1} = QP = \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & \frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

Let us note first that method I is a special case of method II but that method I is worthy of special discussion, for although Andree [2] claims method II is the faster of the two, method I is more readily coded for high speed machines.

Method III. Given matrix A, consider matrix B such that $AB = I$. Setting the product AB equal to I term by term yields n^2 equations in n^2 unknowns which one tries to solve. These actually form n distinct sets of linear systems.

Example 3: Given

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}, \qquad \text{let } B = \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix}.$$

Then:

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{bmatrix} = \begin{bmatrix} 2x_1-2x_2+4x_3 & 2y_1-2y_2+4y_3 & 2z_1-2z_2+4z_3 \\ 2x_1+3x_2+2x_3 & 2y_1+3y_2+2y_3 & 2z_1+3z_2+2z_3 \\ -x_1+x_2-x_3 & -y_1+y_2-y_3 & -z_1+z_2-z_3 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Hence:

$$\begin{array}{lll} 2x_1 - 2x_2 + 4x_3 = 1 & 2y_1 - 2y_2 + 4y_3 = 0 & 2z_1 - 2z_2 + 4z_3 = 0 \\ 2x_1 + 3x_2 + 2x_3 = 0 & 2y_1 + 3y_2 + 2y_3 = 1 & 2z_1 + 3z_2 + 2z_3 = 0 \\ -x_1 + x_2 - x_3 = 0 & -y_1 + y_2 - y_3 = 0 & -z_1 + z_2 - z_3 = 1, \end{array}$$

the solutions of which are

$$\begin{array}{lll} x_1 = -\frac{1}{2}, & x_2 = 0, & x_3 = \frac{1}{2} \\ y_1 = \frac{1}{5}, & y_2 = \frac{1}{5}, & y_3 = 0 \\ z_1 = -\frac{8}{5}, & z_2 = \frac{2}{5}, & z_3 = 1. \end{array}$$

Therefore:

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

It will be noted that the problem of matrix inversion has here been converted to one of solving linear systems of equations. Much research has been done on this latter problem and hence any technique associated with solving linear systems can be applied to matrix inversion. Besides the usual methods of: a) Cramer's rule, b) Gauss elimination method, c) Gauss-Jordan method, and d) Crout's method, more modern techniques are described in: [**5**], [**8**], [**12**], [**16**], [**21**], [**23**], [**26**].

Method IV. Adjoint method.

Let A_{11} = determinant formed from matrix A by deleting row 1 and column 1 and multiplying by $(-1)^{1+1}$.

Let A_{21} = determinant formed from matrix A by deleting row 2 and column 1 and multiplying by $(-1)^{2+1}$.

Let A_{37} = determinant formed from matrix A by deleting row 3 and column 7 and multiplying by $(-1)^{3+7}$.

Let A_{ij} = determinant formed from matrix A by deleting row i and column j and multiplying by $(-1)^{i+j}$.

Let C be the matrix:

$$C = \begin{bmatrix} A_{11} & A_{21} & A_{31} \cdots A_{n1} \\ A_{12} & A_{22} & A_{32} \cdots A_{n2} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ A_{1n} & A_{2n} & \cdots A_{nn} \end{bmatrix}.$$

Then $A^{-1} = (1/|A|)C$, where $|A|$ is the determinant of matrix A.

Example 4. Let:

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then:

$$A_{11} = (-1)^2 \begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} = -5 \qquad A_{21} = (-1)^3 \begin{vmatrix} -2 & 4 \\ 1 & -1 \end{vmatrix} = 2$$

$$A_{12} = (-1)^3 \begin{vmatrix} 2 & 2 \\ -1 & -1 \end{vmatrix} = 0 \qquad A_{22} = (-1)^4 \begin{vmatrix} 2 & 4 \\ -1 & -1 \end{vmatrix} = 2$$

$$A_{13} = (-1)^4 \begin{vmatrix} 2 & 3 \\ -1 & 1 \end{vmatrix} = 5 \qquad A_{23} = (-1)^5 \begin{vmatrix} 2 & -2 \\ -1 & 1 \end{vmatrix} = 0$$

$$A_{31} = (-1)^4 \begin{vmatrix} -2 & 4 \\ 3 & 2 \end{vmatrix} = -16$$

$$A_{32} = (-1)^5 \begin{vmatrix} 2 & 4 \\ 2 & 2 \end{vmatrix} = 4$$

$$A_{33} = (-1)^6 \begin{vmatrix} 2 & -2 \\ 2 & 3 \end{vmatrix} = 10.$$

Also:

$$|A| = \begin{vmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{vmatrix} = 10;$$

therefore

$$A^{-1} = \frac{1}{10}\begin{bmatrix} -5 & 2 & -16 \\ 0 & 2 & 4 \\ 5 & 0 & 10 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & \frac{1}{10} \end{bmatrix}.$$

Method V—Method of Partitions. Given matrix A, let the sequence of matrices $S_1, S_2, \cdots, S_N$, be defined by:

$$S_1 = [a_{11}],\ S_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \left[\begin{array}{c|c} S_1 & a_{12} \\ \hline a_{21} & a_{22} \end{array}\right],\ S_3 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \left[\begin{array}{c|c} S_2 & \begin{matrix} a_{13} \\ a_{23} \end{matrix} \\ \hline a_{31}a_{32} & a_{33} \end{array}\right], \cdots$$

$$S_N = A = \begin{bmatrix} a_{11} & a_{12} \cdots a_{1N} \\ a_{21} & a_{22} \cdots a_{2N} \\ \cdot & \\ \cdot & \\ \cdot & \\ a_{N1} & a_{N2} \cdots a_{NN} \end{bmatrix} = \left[\begin{array}{c|c} S_{N-1} & \begin{matrix} a_{1N} \\ a_{2N} \\ \cdot \\ \cdot \\ \cdot \end{matrix} \\ \hline a_{N1}a_{N2}\cdots & a_{NN} \end{array}\right].$$

In general then, we shall want to partition a matrix into 4 sub-matrices which we write as: (See [11], pp. 112–115)

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}.$$

For example, for S_3, we have:

$$\alpha_{11} = S_2, \qquad \alpha_{12} = \begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}, \qquad \alpha_{21} = [a_{31}\ a_{32}], \qquad \alpha_{22} = [a_{33}].$$

If we denote the inverse of A by B and let:

$$B = \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix},$$

then we have:

$$\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \cdot \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} = I.$$

Multiplying out and setting the resulting matrix equal to unity yields four matrix equations. Solving these matrix equations for β_{11}, β_{12}, β_{21}, and β_{22}, we have:

$$\begin{aligned} \beta_{22} &= D \\ \beta_{21} &= -D(\alpha_{21}\alpha_{11}^{-1}) \\ \beta_{12} &= -(\alpha_{11}^{-1}\alpha_{12})D \\ \beta_{11} &= \alpha_{11}^{-1} + (\alpha_{11}^{-1}\alpha_{12})D\alpha_{21}\alpha_{11}^{-1}, \end{aligned}$$

where $D=(\alpha_{22}-\alpha_{21}\alpha_{11}^{-1}\alpha_{12})^{-1}$.

The technique of inversion will be to apply the above formulas in a recursive fashion by first finding the inverse of S_1, then of S_2, then of S_3, $\cdots$, finally of S_N and this will be A^{-1}. One interchanges two rows to avoid trouble when a submatrix is singular and the process cannot be continued directly. This necessitates only changing the same two columns in the result to find A^{-1}.

Example 5:

$$A=\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then $S_1=[2]$, $S_1^{-1}=[\frac{1}{2}]$. Hence:

$$S_2=\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}=\begin{bmatrix} S_1 & a_{12} \\ a_{21} & a_{22} \end{bmatrix}=\begin{bmatrix} 2 & -2 \\ 2 & 3 \end{bmatrix}.$$

Using the formulas for β_{ij}, it follows that: $\beta_{11}=\frac{3}{10}$, $\beta_{12}=\frac{1}{5}$, $\beta_{21}=-\frac{1}{5}$, $\beta_{22}=\frac{1}{5}$. Hence

$$S_2^{-1}=\begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix}=\begin{bmatrix} \frac{3}{10} & \frac{1}{5} \\ -\frac{1}{5} & \frac{1}{5} \end{bmatrix}.$$

Now:

$$S_3=\left[\begin{array}{c|c} S_2 & \begin{matrix} a_{13} \\ a_{23} \end{matrix} \\ \hline a_{31}a_{32} & a_{33} \end{array}\right]=\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix}=\left[\begin{array}{cc|c} 2 & -2 & 4 \\ 2 & 3 & 2 \\ \hline -1 & 1 & -1 \end{array}\right].$$

Then:

$$\alpha_{11}=S_2=\begin{bmatrix} 2 & -2 \\ 2 & 3 \end{bmatrix}, \qquad \alpha_{12}=\begin{bmatrix} a_{13} \\ a_{23} \end{bmatrix}=\begin{bmatrix} 4 \\ 2 \end{bmatrix},$$

$$\alpha_{21}=[a_{31}\ a_{32}]=[-1,\ 1], \qquad \alpha_{22}=[a_{33}]=[-1].$$

Note that $S_2^{-1}=\alpha_{11}^{-1}$ is known from the previous calculation. Again using the formulas for the β_{ij}, it follows that:

$$\beta_{11}=\begin{bmatrix} -\frac{1}{2} & \frac{1}{5} \\ 0 & \frac{1}{5} \end{bmatrix}, \qquad \beta_{12}=\begin{bmatrix} -\frac{8}{5} \\ \frac{2}{5} \end{bmatrix}, \qquad \beta_{21}=[\tfrac{1}{2}\ 0], \qquad \beta_{22}=[1].$$

Therefore:

$$A^{-1}=S_3^{-1}=\begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

Concerning experiments with this method, one should consult [17].

Method VI. Given matrix A, consider the matrix: $B = A - \lambda I$, where I is the unit matrix and λ is a constant. If we set the determinant of B equal to zero, *i.e.*, $|B| = 0$, we see that this is nothing more than a polynomial equation in λ of degree n. The values $\lambda_1, \lambda_2, \cdots, \lambda_n$ which satisfy this equation are called the eigenvalues of the matrix A and the equation itself is called the characteristic equation of the matrix A. For example, if:

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2-\lambda & -2 & 4 \\ 2 & 3-\lambda & 2 \\ -1 & 1 & -1-\lambda \end{bmatrix},$$

the characteristic equation is:

$$\lambda^3 - 4\lambda^2 + 7\lambda - 10 = 0.$$

A fundamental theorem of matrix theory says that every matrix satisfies its characteristic equation. By the example above, we explain this to mean:

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}^3 - 4\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}^2 + 7\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}$$
$$- 10\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and this may be readily verified by performing the indicated operations.

A matrix may satisfy many different equations besides its characteristic equation. It can be shown that there exists an equation of least degree which the matrix satisfies and this degree may be less than that of the characteristic equation. This equation of least degree is unique to within a multiplicative constant and is appropriately called the minimal equation associated with the matrix. If $C(x) = 0$ is the characteristic equation associated with a matrix and $\mathrm{M}(x) = 0$ is the minimal equation which the matrix satisfies, then it can be shown that $\mathrm{M}(x)$ is a factor of $C(x)$.

Now, suppose some equation has been found which the matrix satisfies, the characteristic equation usually sufficing. Then if the equation is:

$$a_0 x^n + a_1 x^{n-1} + a_2 x^{n-2} + \cdots + a_{n-1} x + a_n = 0, \qquad a_n \neq 0,$$

and if A is the matrix:

$$a_0A^n + a_1A^{n-1} + \cdots + a_{n-1}A + a_nI = 0.$$

Multiplying through by A^{-1}, we have:

$$a_0A^{n-1} + a_1A^{n-2} + \cdots + a_{n-1}I + a_nA^{-1} = 0,$$

or:

$$A^{-1} = \frac{1}{a_n}\left\{- a_0A^{n-1} - a_1A^{n-2} - \cdots - a_{n-1}I\right\}.$$

Example 6:

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

It has been shown that A has the characteristic equation:

$$\lambda^3 - 4\lambda^2 + 7\lambda - 10 = 0.$$

Then:

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}^3 - 4\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}^2 + 7\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}$$

$$- 10\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and:

$$\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}^2 - 4\begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix} + 7\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - 10A^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

or:

$$\begin{bmatrix} -5 & 2 & -16 \\ 0 & 2 & 4 \\ 5 & 0 & 10 \end{bmatrix} - 10A^{-1} = 0,$$

and finally,

$$A^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

The case where $a_n=0$ offers no exorbitant trouble, for then the matrix has no inverse.

Method VII—Frame's method (See [10].) Given matrix A, let $A_0=I$. Using:

$$(1) \quad c_k = \frac{1}{k} \text{ trace } [AA_{k-1}]$$

$$(2) \quad A_k = AA_{k-1} - c_k I,$$

we have:

$$A^{-1} = \frac{A_{n-1}}{c_n}, \qquad c_n \neq 0.$$

Example 7. Let

$$A = \begin{bmatrix} 2 & -2 & 4 \\ 2 & 3 & 2 \\ -1 & 1 & -1 \end{bmatrix}.$$

Then:

$$c_1 = \tfrac{1}{1} \text{ trace } [AA_0] = 4, \qquad A_1 = AA_{k-1} - 4I = AA_0 - 4I = \begin{bmatrix} -2 & -2 & 4 \\ 2 & -1 & 2 \\ -1 & 1 & -5 \end{bmatrix},$$

$$c_2 = \tfrac{1}{2} \text{ trace } [AA_1] = -7, \quad A_2 = AA_1 + 7I = \begin{bmatrix} -5 & 2 & -16 \\ 0 & 2 & 4 \\ 5 & 0 & 10 \end{bmatrix},$$

$$c_3 = \tfrac{1}{3} \text{ trace } [AA_2] = \tfrac{1}{3}(-10 + 0 + 20 + 4 + 6 + 0 + 16 + 4 - 10) = 10.$$

Finally:

$$A^{-1} = \frac{A_2}{c_3} = \begin{bmatrix} -\frac{1}{2} & \frac{1}{5} & -\frac{8}{5} \\ 0 & \frac{1}{5} & \frac{2}{5} \\ \frac{1}{2} & 0 & 1 \end{bmatrix}.$$

Method VIII. Let a_k, $k=0, 1, 2, \cdots, n-1$, be n constants. Consider the special matrix:

$$B = \begin{bmatrix} a_0 & a_{n-1} & a_{n-2} & a_{n-3} \cdots a_2 & a_1 \\ a_1 & a_0 & a_{n-1} & a_{n-2} \cdots a_3 & a_2 \\ a_2 & a_1 & a_0 & a_{n-1} \cdots a_4 & a_3 \\ \vdots & & & & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & a_{n-4} \cdots a_1 & a_0 \end{bmatrix}.$$

Such special matrices are called circulant matrices. To invert these we use the procedure (see [13]):

1) Compute:

$$w = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}\cdot$$

2) Calculate the eigenvalues of the given matrix by:

$$e_s(B) = \sum_{r=0}^{n-1} a_r w^{rs}, \qquad s = 0, 1, \cdots, n-1.$$

3) Calculate the inverses of these eigenvalues and denote these by e_s^{-1}.

4) Calculate the numbers:

$$b_s = \frac{1}{n}\sum_{r=0}^{n-1}(e_r^{-1})w^{-rs}, \qquad s = 0, 1, \cdots, n-1,$$

5)

$$B^{-1} = \begin{bmatrix} b_0 & b_{n-1} & \cdots & b_2 & b_1 \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ b_{n-2} & b_{n-3} & & b_0 & b_{n-1} \\ b_{n-1} & b_{n-2} & & b_1 & b_0 \end{bmatrix}.$$

Example 8. Let

$$B = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ -2 & 3 & 1 \end{bmatrix},$$

which is a circulant matrix where $a_0 = 1$, $a_1 = 3$, $a_2 = 2$.

1) $w = -\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i$

2) $e_0(A) = 6$

$$e_1(A) = 3\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) + 2\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) + 1 = -\frac{3}{2} + \frac{\sqrt{3}}{2}i$$

$$e_2(A) = 3\left(-\frac{1}{2} - \frac{\sqrt{3}}{2}i\right) + 2\left(-\frac{1}{2} + \frac{\sqrt{3}}{2}i\right) + 1 = -\frac{3}{2} - \frac{\sqrt{3}}{2}i$$

3) $e_0^{-1} = \dfrac{1}{6}$

$$e_1^{-1} = -\frac{1}{2} - \frac{\sqrt{3}}{6}i$$

$$e_2^{-1} = -\frac{1}{2} + \frac{\sqrt{3}}{6}i$$

4) $$b_0 = -\frac{5}{18}$$

$$b_1 = \frac{1}{3}(e_0^{-1} + e_1^{-1}w^{-1} + e_2^{-1}w^{-2}) = \frac{1}{18}$$

$$b_2 = \frac{1}{3}(e_0^{-1} + e_1^{-1}w^{-2} + e_2^{-1}w^{-4}) = \frac{7}{18}$$

5)
$$B^{-1} = \begin{bmatrix} -\frac{5}{18} & \frac{7}{18} & \frac{1}{18} \\ \frac{1}{18} & -\frac{5}{18} & \frac{7}{18} \\ \frac{7}{18} & \frac{1}{18} & -\frac{5}{18} \end{bmatrix}.$$

The method is a good one in that it works readily on matrices with complex components. If B is real and symmetrical, [**13**] gives simpler formulas for the computation of the eigenvalues.

Method IX—Newton's formula. In general, given a number W, suppose one wishes to calculate $1/W$. Let $x=1/W$. Then by Newton's formula, setting $f(x)=W-1/x$,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \qquad f'(x_n) \neq 0,$$

$$= x_n - \frac{W - \frac{1}{x_n}}{\frac{1}{x_n^2}}$$

$$= x_n(2 - Wx_n).$$

If we let A be a matrix whose inverse is desired, and we let X_0 be a reasonable approximation to A^{-1}, then the above formula, written in terms of matrices is:

$$X_{n+1} = X_n(2I - AX_n).$$

The following example is selected from [**27**]:

Example 9. Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 4 \end{bmatrix}.$$

Then: choose

$$X_0 = \begin{bmatrix} -3.0 & 4.1 & -0.9 \\ 4.1 & -5.1 & 1.9 \\ -0.9 & 1.9 & -1.1 \end{bmatrix}.$$

Using the recursion formula above, it follows that:

$$X_1 = \begin{bmatrix} -4.26 & 4.14 & -0.86 \\ 4.14 & -5.06 & 1.94 \\ -0.86 & 1.94 & -1.06 \end{bmatrix}, \qquad X_2 = \begin{bmatrix} -3.9976 & 3.9864 & -1.0138 \\ 3.9864 & -4.9896 & 2.0104 \\ -1.0136 & 2.0104 & -0.9896 \end{bmatrix}.$$

The exact solution is:

$$A^{-1} = \begin{bmatrix} -4 & 4 & -1 \\ 4 & -5 & 2 \\ -1 & 2 & -1 \end{bmatrix}.$$

Concerning convergence of iteration processes applied to matrices, one should consult [22] and [17].

Method X. On certain matrices, the following method may be applied: Let $A = (I+B)$. Then, $A^{-1} = (I+B)^{-1} = I - B + B^2 - B^3 + \cdots$
Example 10: Let

$$A = \begin{bmatrix} 4 & 0 & 0 \\ -1 & 4 & 0 \\ 0 & -1 & 4 \end{bmatrix}.$$

Then:

$$\begin{aligned} A^{-1} &= \left\{ \begin{bmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \end{bmatrix} \right\}^{-1} \\ &= \tfrac{1}{4} \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \end{bmatrix} \right\}^{-1} \end{aligned}$$

$$= \frac{1}{4}\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \end{bmatrix}^2 - \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \end{bmatrix}^3 + \cdots \right\}$$

$$= \frac{1}{4}\left\{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{16} & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdots \right\}$$

$$= \frac{1}{4}\begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{4} & 1 & 0 \\ \frac{1}{16} & \frac{1}{4} & 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & 0 & 0 \\ \frac{1}{16} & \frac{1}{4} & 0 \\ \frac{1}{64} & \frac{1}{16} & \frac{1}{4} \end{bmatrix}.$$

XI. Other methods and concluding remarks.

A. A newer type method which necessitates knowledge of sampling theory and game theory, concepts beyond the scope of an elementary paper, is the Monte Carlo method, developed by Von Neumann and Ulam, written up by Forsythe and Liebler in [**7**], and played with by Todd in [**28**].

B. Variations of the preceding methods are often useful for special matrices. For example, consider a matrix with the following properties:

1. It is symmetric.
2. The largest element of any row lies on the main diagonal.
3. The matrix has zeros everywhere except possibly on the main diagonal and in the positions just above and just below the main diagonal.

Such matrices occur frequently in physical problems, for example, in vibration. Preceding methods can be applied but a rapid variation is given in [**24**].

C. Lastly, it should be noted that if A can be written as a product of matrices whose inverses are readily found, then A^{-1} is easily produced, for:

$$\text{if:}\quad A = BCDEF, \qquad A^{-1} = F^{-1}E^{-1}D^{-1}C^{-1}B^{-1}.$$

Hence ingenuity must be used to find matrices whose inverses are simply found and whose product is A. Cholesky's method [**27**] is such a scheme.

References and Sources for Further Reading

1. Albert, A. A., A rule for computing the inverse of a matrix, this Monthly, vol. 48, 1941, pp. 198–199.

2. Andree, R. V., Computation of inverse of a matrix, this Monthly, vol. 58, 1951, pp. 87–92.

3. Bingham, M. D., A new method for obtaining the inverse matrix, J. Amer. Stat. Assoc., vol. 36, 1941, pp. 530–534.

4. Collar, A. R., On the reciprocation of certain matrices, Proc. Roy. Soc. Edinburgh, vol. 59, 1939, pp. 195–206.

5. Dwyer, P. S., *Linear Computation*, Wiley, 1951.

6. Forsythe, G. E., *Theory of Selected Methods of Finite Matrix Inversion and Decomposition*, Lectures in Math. 136, notes by D. G. Aronson and K. Iverson, INA Report 52-3, 1951.

7. Forsythe, G. E., and Liebler, R. A., *Matrix Inversion by a Monte Carlo Method*, MTAC 4, at INA, 1950.

8. Forsythe, G. E., Solving linear algebraic equations can be interesting, Bull. Amer. Math.

Soc., vol. 59, 1953, pp. 299–329.

9. Fox, L., Practical methods for solutions of linear equations and the inversion of matrices, J. Royal Stat. Soc., Ser. 12, B, 1950, pp. 120–136.

10. Frame, J. S., A simple recursion formula for inverting a matrix, Abstract, Bull. Amer. Math. Soc., vol. 55, 1949, p. 1045.

11. Frazer, R. A., Duncan, W. J., and Collar, A. R., *Elementary Matrices*, Oxford Univ. Press, 1951.

12. Froberg, C. E., Solutions of linear systems of equations on a relay machine, from: *Simultaneous Linear Equations and Determination of Eigenvalues*. U. S. Dept. of Commerce, N.B.S., Applied Math. Series, 29, 1953, pp. 39–43.

13. Good, I. J., On the inversion of circulant matrices, Biometrika, vol. 37, 1950, pp. 185–186.

14. Grosh, L. E., and Usdin, E., A method for evaluating determinants and inverting matrices with arbitrary polynomial elements by IBM punched card methods, Proc., Industrial Corporation Seminar, Sept. 1950, IBM Corp., N. Y., 1951, pp. 99–103.

15. Guttshall, W. D., Practical inversion of matrices of high order, Proc., Computation Seminar, Dec., 1949, IBM Corp., N. Y., 1951, pp. 171–173.

16. Hestenes, M. R., and Stiefel, E., Method of conjugate gradients for solving linear systems, J. Res. of NBS, vol. 49, 1952.

17. Lotkin, M., and Remage, R., Matrix inversion by partitioning, BRL Report 823, Aberdeen Proving Ground, Maryland, 1952.

18. Mitchell, H. F., Inversion of a matrix of order 38, MTAC 3, 1948, pp. 161–166.

19. Ostrowski, A., Sur la Variation de la Matrice Inverse d'une Matrice Donneé, Comptes Rendu Acad. Sci. Paris, vol. 231, 1950, pp. 1019–1021.

20. Ostrowski, A., Taussky, O., and Todd, J., Bibliography on Computational Aspects of Finite Matrix Theory, Part I, NBS, 1949.

21. Petrie, G. W., III, Matrix inversion and solution of simultaneous linear algebraic equations with the IBM 604 electronic calculating punch. Same source as reference [12], pp. 107–112.

22. Plunkett, R., On the convergence of matrix iteration processes, Quart. Appl. Math. 7, 1950, pp. 419–421.

23. Rosser, J. B., Rapidly converging iterative methods for solving linear equations, Same source as reference [12], pp. 59–64.

24. Saibel, E., A rapid method of inversion of certain types of matrices, J. Franklin Inst. 237, 1944, pp. 197–201.

25. Sherman, J., Computations relating to inverse matrices. Same source as reference [12], pp. 123–124.

26. Stiefel, E., Some special methods of relaxation technique. Same source as reference [12], pp. 43–48.

27. Taussky, O., and Todd, J., Systems of equations, matrices, and determinants, Math. Mag., Part I, Sept.–Oct., 1952; Part II, Nov.–Dec., 1952.

28. Todd, J., Matrix inversion by a Monte Carlo method, NBS-CL 50-2, 1951.

29. Todd, J., Experiments on the inversion of a 16×16 matrix. Same source as reference [12], pp. 113–116.

30. Von Neumann, J., and Goldstine, H. H., Numerical inverting of matrices of high order, Bull. Amer. Math. Soc., vol. 53, 1947, pp. 1021–1099.

31. Wagner, H. M., Matrix Inversion on an Automatic Calculator by Row and Column Partitioning, Rand Corp. Publication.

32. Wasow, W. R., A note on inversion of matrices by random walks, MTAC 6, 1952, pp. 78–81.

33. Woodbury, M. A., Inverting modifies matrices, Stat. Res. Group Memorandum Report 42, Princeton, N. J., 14 June, 1950, 4 pages.

34. Zurmühl, R., Zur numerischen Auflösung linearer Gleichungsysteme nach dem Matrizenverfahren von Banachiewicz, Zeitschrf. Angew Math. Mech., vol. 29, 1949, pp. 76–84.