Gantmacher, F. R.: "The Theory of Matrices", Vol 1. Chelsea, New York (1959).

CHAPTER II

THE ALGORITHM OF GAUSS AND SOME OF ITS APPLICATIONS

§ 1. Gauss's Elimination Method

1. Let

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = y_2 \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = y_n \end{array}\right\} \tag{1}$$

be a system of n linear equations in n unknowns $x_1, x_2, \ldots, x_n$ with right-hand sides $y_1, y_2, \ldots, y_n$.

In matrix form this system may be written as

$$Ax = y. \tag{1'}$$

Here $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ are columns and $A = \| a_{ik} \|_1^n$ is the square coefficient matrix.

If A is non-singular, then we can rewrite this as

$$x = A^{-1}y, \tag{2}$$

or in explicit form:

$$x_i = \sum_{k=1}^{n} a_{ik}^{(-1)} y_k \qquad (i = 1, 2, \ldots, n). \tag{2'}$$

Thus, the task of computing the elements of the inverse matrix $A^{-1} = \| a_{ik}^{(-1)} \|_1^n$ is equivalent to the task of solving the system of equations (1) for *arbitrary* right-hand sides $y_1, y_2, \ldots, y_n$. The elements of the inverse matrix are determined by the formulas (25) of Chapter I. However, the actual computation of the elements of A^{-1} by these formulas is very tedious for large n. Therefore, effective methods of computing the elements of an inverse matrix—and hence of solving a system of linear equations—are of great practical value.[1]

[1] For a detailed account of these methods, we refer the reader to the book by Faddeev [15] and the group of papers that appeared in *Uspehi Mat. Nauk*, Vol. 5, 3 (1950).

In the present chapter we expound the theoretical basis of some of these methods; they are variants of Gauss's elimination method, whose acquaintance the reader first made in his algebra course at school.

2. Suppose that in the system of equations (1) we have $a_{11} \neq 0$. We eliminate x_1 from all the equations beginning with the second by adding to the second equation the first multipled by $-\frac{a_{21}}{a_{11}}$, to the third the first multiplied by $-\frac{a_{31}}{a_{11}}$, and so on. The system (1) has now been replaced by the equivalent system

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = y_1 \\ a_{22}^{(1)}x_2 + \cdots + a_{2n}^{(1)}x_n = y_2^{(1)} \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n2}^{(1)}x_2 + \cdots + a_{nn}^{(1)}x_n = y_n^{(1)} \end{array}\right\} \qquad (3)$$

The coefficients of the unknowns and the constant terms of the last $n-1$ equations are given by the formulas

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}, \qquad y_i^{(1)} = y_i - \frac{a_{i1}}{a_{11}} y_1 \qquad (i, j = 2, \ldots, n). \qquad (3')$$

Suppose that $a_{22}^{(1)} \neq 0$. Then we eliminate x_2 in the same way from the last $n-2$ equations of the system (3) and obtain the system

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n = y_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2n}^{(1)}x_n = y_2^{(1)} \\ a_{33}^{(2)}x_3 + \cdots + a_{3n}^{(2)}x_n = y_3^{(2)} \\ \cdots\cdots\cdots\cdots\cdots \\ a_{n3}^{(2)}x_3 + \cdots + a_{nn}^{(2)}x_n = y_n^{(2)}. \end{array}\right\} \qquad (4)$$

The new coefficients and the new right-hand sides are connected with the preceding ones by the formulas:

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} a_{2j}^{(1)}, \quad y_i^{(2)} = y_i^{(1)} - \frac{a_{i2}^{(1)}}{a_{22}^{(1)}} y_2^{(1)} \qquad (i, j = 3, \ldots, n). \qquad (5)$$

Continuing the algorithm, we go in $n-1$ steps from the original system (1) to the triangular recurrent system

$$\left.\begin{array}{r} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n \quad = y_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \cdots + a_{2n}^{(1)}x_n \quad = y_2^{(1)} \\ a_{33}^{(2)}a_3 + \cdots + a_{3n}^{(2)}x_n \quad = y_3^{(2)} \\ \cdots\cdots\cdots\cdots\cdots \\ a_{nn}^{(n-1)}x_n = y_n^{(n-1)}. \end{array}\right\} \qquad (6)$$

This reduction can be carried out if and only if in the process all the numbers $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{n-1,n-1}^{(n-2)}$ turn out to be different from zero.

This *algorithm of Gauss* consists of operations of a simple type such as can easily be carried out by present-day computing machines.

3. Let us express the coefficients and the right-hand sides of the reduced system in terms of the coefficients and the right-hand sides of the original system (1). We shall not assume here that in the reduction process all the numbers $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, a_{n-1,n-1}^{(n-2)}$ turn out to be different from zero; we consider the general case, in which the first p of these numbers are different from zero:

$$a_{11} \neq 0, \quad a_{22}^{(1)} \neq 0, \quad \ldots, \quad a_{pp}^{(p-1)} \neq 0 \qquad (p \leq n-1). \tag{7}$$

This enables us (at the p-th step of the reduction) to put the original system of equations into the form

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots\cdots\cdots\cdots + a_{1n}x_n = y_1 \\ \quad a_{22}^{(1)}x_2 + \cdots\cdots\cdots\cdots + a_{2n}^{(1)}x_n = y_2^{(1)} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad a_{pp}^{(p-1)}x_p + \cdots\cdots\cdots + a_{pn}^{(p-1)}x_n = y_p^{(p-1)} \\ \qquad\quad a_{p+1,p+1}^{(p)}x_{p+1} + \cdots + a_{p+1,n}^{(p)}x_n = y_{p+1}^{(p)} \\ \qquad\quad \cdots\cdots\cdots\cdots\cdots\cdots \\ \qquad\quad a_{n,p+1}^{(p)}x_{p+1} + \cdots + a_{nn}^{(p)}x_n = y_n^{(p)}. \end{array}\right\} \tag{8}$$

We denote the coefficient matrix of this system of equations by G_p:

$$G_p = \left\| \begin{array}{cccccccc} a_{11} & a_{12} & \ldots & a_{1p} & a_{1,p+1} & \ldots & a_{1n} \\ 0 & a_{22}^{(1)} & \ldots & a_{2p}^{(1)} & a_{2,p+1}^{(1)} & \ldots & a_{2n}^{(1)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & a_{pp}^{(p-1)} & a_{p,p+1}^{(p-1)} & \ldots & a_{pn}^{(p-1)} \\ 0 & 0 & \ldots & 0 & a_{p+1,p+1}^{(p)} & \ldots & a_{p+1,n}^{(p)} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \ldots & 0 & a_{n,p+1}^{(p)} & \ldots & a_{nn}^{(p)} \end{array} \right\|. \tag{9}$$

The transition from A to G_p is effected as follows: To every row of A in succession from the second to the n-th there are added some preceding rows (from the first p) multiplied by certain factors. Therefore all the minors of order h contained in the first h rows of A and G_p are equal:

$$A\begin{pmatrix} 1 & 2 & \ldots & h \\ k_1 & k_2 & \ldots & k_h \end{pmatrix} = G_p\begin{pmatrix} 1 & 2 & \ldots & h \\ k_1 & k_2 & \ldots & k_h \end{pmatrix} \quad \begin{pmatrix} 1 \leq k_1 < k_2 < \cdots < k_h \leq n \\ h = 1, 2, \ldots, n \end{pmatrix}. \tag{10}$$

rank r. Then, by a suitable permutation of the equations and a renumbering of the unknowns, we can arrange that the following inequalities hold:

$$A\begin{pmatrix}1 & 2 & \dots & j\\ 1 & 2 & \dots & j\end{pmatrix}\neq 0 \qquad (j=1, 2, \dots, r). \tag{17}$$

This enables us to eliminate $x_1, x_2, \dots, x_r$ consecutively and to obtain the system of equations

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots\cdots + a_{1n}x_n = y_1 \\ \quad a_{22}^{(1)}x_2 + \cdots\cdots + a_{2n}^{(1)}x_n = y_2^{(1)} \\ \quad \cdots\cdots\cdots\cdots \\ \qquad a_{rr}^{(r-1)}x_r + \cdots + a_{rn}^{(r-1)}x_n = y_r^{(r-1)} \\ \qquad\quad a_{r+1,r+1}^{(r)}x_{r+1} + \cdots + a_{r+1,n}^{(r)}x_n = y_{r+1}^{(r)} \\ \qquad\quad \cdots\cdots\cdots\cdots \\ \qquad\quad a_{n,r+1}^{(r)}x_{r+1} + \cdots + a_{nn}^{(r)}x_n = y_n^{(r)}. \end{array}\right\} \tag{18}$$

Here the coefficients are determined by the formulas (13). From these formulas it follows, because the rank of the matrix $A=\| a_{ik} \|_1^n$ is equal to r, that

$$a_{ik}^{(r)}=0 \qquad (i, k=r+1, \dots, n). \tag{19}$$

Therefore the last $n-r$ equations (18) reduce to the consistency conditions

$$y_i^{(r)}=0 \qquad (i=r+1, \dots, n). \tag{20}$$

Note that in the elimination algorithm the column of constant terms is subjected to the same transformations as the other columns, of coefficients. Therefore, by supplementing the matrix $A=\| a_{ik} \|_1^n$ with an $(n+1)$-th column of the constant terms we obtain:

$$y_i^{(p)}=\frac{A\begin{pmatrix}1 & \dots & p & i\\ 1 & \dots & p & n+1\end{pmatrix}}{A\begin{pmatrix}1 & \dots & p\\ 1 & \dots & p\end{pmatrix}} \qquad (i=1, 2, \dots, n;\ p=1, 2, \dots, r). \tag{21}$$

In particular, the consistency conditions (20) reduce to the well-known equations

$$A\begin{pmatrix}1 & \dots & r & r+j\\ 1 & \dots & r & n+1\end{pmatrix}=0 \qquad (j=1, 2, \dots, n-r). \tag{22}$$