CONTENTS

PART I

ALGEBRA

CHAPTER 1

MATRIX OPERATIONS

MATRIX NOTATION

1. Suppose that we have a set of m quantities $y_1, \ldots, y_m$, each of which is a linear function of the n quantities $x_1, \ldots, x_n$. We say that there is a linear transformation expressed by the equations

$$\begin{matrix} a_{11}x_1 + \ldots + a_{1n}x_n = y_1 \\ \vdots \qquad\qquad \vdots \quad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = y_m \end{matrix} \tag{1}$$

The $m \times n$ quantities a_{ij} are said to form the matrix of order $m \times n$ of this transformation, and each of the a_{ij} is called an element of the matrix. It is conventional to write each element as

$$\textit{name of matrix}_{\textit{row number, column number}}$$

We can economise a little by writing the matrix elements separately from the vector of quantities $x_1, \ldots, x_n$

$$\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \tag{2}$$

When this notation is used it is understood that each element in the first column of the factor on the left is multiplied by the element in the first row of that on the right, the second column by the second row, and so on, and that the answers are added to give equations of the form (1). We now need to write each x_i once, not m times.

2. We may wish to refer to eqns. (1) several times, treating them as a whole and not singling out any particular part of them. It then becomes laborious to write them in the form of eqns. (1) or (2) and we need a shorter form. We get this by writing

$$A \equiv \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \quad \mathbf{x} \equiv \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \quad \mathbf{y} \equiv \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \tag{3}$$

The symbol A is said to represent the matrix of order $m \times n$, i.e. m rows and n columns, and $\mathbf{x}$ and $\mathbf{y}$ are called column vectors. We can then rewrite eqns. (1) as

$$A\mathbf{x} = \mathbf{y} \tag{4}$$

This is called matrix algebraic notation and is so concise that we can write many sets of equations in the time otherwise needed for one set. We shall now consider how we can manipulate this notation, and it will appear that the rules are conveniently similar to those of elementary algebra, with certain additions and exceptions.

MATRIX ADDITION

3. If eqns. (1) hold and also

$$\begin{matrix} b_{11}x_1 + \ldots + b_{1n}x_n = z_1 \\ \vdots \qquad\qquad \vdots \qquad \vdots \\ b_{m1}x_1 + \ldots + b_{mn}x_n = z_m \end{matrix} \tag{5}$$

$$\begin{matrix} y_1 + z_1 = w_1 \\ \vdots \qquad \vdots \\ y_m + z_m = w_m \end{matrix} \tag{6}$$

so that

$$\begin{matrix} c_{11}x_1 + \ldots + c_{1n}x_n = w_1 \\ \vdots \qquad\qquad \vdots \qquad \vdots \\ c_{m1}x_1 + \ldots + c_{mn}x_n = w_m \end{matrix} \tag{7}$$

We wish to be able to write

$$A\mathbf{x} + B\mathbf{x} = (A+B)\mathbf{x} = C\mathbf{x} = \mathbf{w} \tag{8}$$

and so

$$A + B = C \tag{9}$$

The operation of forming the elements c_{ij} is called adding the matrices A and B. Evidently we have

$$\begin{matrix} (a_{11}+b_{11})x_1 + \ldots + (a_{1n}+b_{1n})x_n = w_1 \\ \vdots \qquad\qquad\qquad \vdots \qquad\qquad \vdots \\ (a_{m1}+b_{m1})x_1 + \ldots + (a_{mn}+b_{mn})x_n = w_m \end{matrix} \tag{10}$$

so that the rule for doing this is simply

$$c_{ij} = a_{ij} + b_{ij} \tag{11}$$

This is called an element-by-element operation, because each element of C depends on one element only of A and one of B. For any such operation it is obvious that both matrices must have the same number of rows and of columns. When matrices satisfy the conditions for an operation they are said to conform.

4. In Sect. 3 we defined matrix addition. Clearly we can regard $-A$ as the matrix with elements $-a_{ij}$, then we can define matrix subtraction so that if

$$C = A - B \tag{12}$$

then

$$c_{ij} = a_{ij} - b_{ij} \tag{13}$$

We can write

$$A + B + \ldots + T = Z \tag{14}$$

where

$$a_{ij} + b_{ij} + \ldots + t_{ij} = z_{ij} \tag{15}$$

and

$$A - B + \ldots - T = Z \tag{16}$$

where

$$a_{ij} - b_{ij} + \ldots - t_{ij} = z_{ij} \tag{17}$$

We can also define the null matrix

$$0 = A - A \tag{18}$$

as the matrix with every element zero, and we have

$$\begin{aligned} A &= 0 + A \\ B + A &= A + B \end{aligned} \tag{19}$$

Evidently the rules for matrix addition are just those for ordinary addition, with the requirement that the matrices must be of the same order.

SPECIAL MATRICES

5. We have seen that it is convenient to have a special name for a matrix with several rows and one column, since we called $\mathbf{x}$ in (3) above a column vector. There are a number of special cases of this type. A matrix with one row only is called a row vector. A matrix with one element only is called a

scalar and obeys the rules for an ordinary number. A matrix whose element a_{ij} is zero unless $i = j$ is called a diagonal matrix since it takes the form

$$\begin{bmatrix} x & 0 & 0 & \dots & 0 \\ 0 & x & 0 & \dots & 0 \\ 0 & 0 & x & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \dots & x \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \end{bmatrix} \quad (\text{if } m > n) \qquad (20)$$

A matrix of this type represents a transformation in which each element of $\mathbf{y} = A\mathbf{x}$ depends on the corresponding element of $\mathbf{x}$ only and this is common enough to be worth special consideration. A square matrix has equal numbers of rows and of columns.

SCALAR MULTIPLICATION

6. If eqns. (1), (5) and (7) hold and if

$$\begin{aligned} \lambda y_1 + \mu z_1 &= w_1 \\ \vdots \qquad & \quad \vdots \\ \lambda y_m + \mu z_m &= w_m \end{aligned} \qquad (21)$$

then we can write

$$\lambda A + \mu B = C \qquad (22)$$

to mean

$$\lambda a_{ij} + \mu b_{ij} = c_{ij} \qquad (23)$$

This gives us the rule for multiplication of a matrix A by a scalar λ, and also that for linear combination of matrices.

MATRIX MULTIPLICATION

7. If eqns. (1) hold and if we also have

$$\begin{aligned} b_{11}y_1 + \dots + b_{1m}y_m &= z_1 \\ \vdots \qquad\qquad & \quad \vdots \\ b_{p1}y_1 + \dots + b_{pm}y_m &= z_p \end{aligned} \qquad (24)$$

then it follows by substituting for $y_1, \dots, y_m$ from (1), that

$$\begin{aligned} (b_{11}a_{11} + b_{12}a_{21} + \dots + b_{1m}a_{m1})x_1 + \dots + (b_{11}a_{1n} + \dots + b_{1m}a_{mn})x_n &= z_1 \\ \vdots \qquad\qquad\qquad\qquad & \quad \vdots \\ (b_{p1}a_{11} + b_{p2}a_{21} + \dots + b_{pm}a_{m1})x_1 + \dots + (b_{p1}a_{1n} + \dots + b_{pm}a_{mn})x_n &= z_p \end{aligned} \qquad (25)$$

We can write the transformation of $\mathbf{x}$ into $\mathbf{z}$ in matrix notation as

$$A\mathbf{x} = \mathbf{y} \tag{26}$$

$$B\mathbf{y} = \mathbf{z} \tag{27}$$

We can also find a single transformation which produces $\mathbf{z}$ from $\mathbf{x}$ directly

$$C\mathbf{x} = \mathbf{z} \tag{28}$$

Evidently the effect of the transformation C is the same as that of the transformation A followed by the transformation B and we can write

$$C\mathbf{x} = BA\mathbf{x} \tag{29}$$

or

$$C = BA \tag{30}$$

This indicates that the rule for multiplying matrices is the rule for calculating the effect of successive transformations. From (25)

$$\begin{aligned} c_{11} &= b_{11}a_{11} + b_{12}a_{21} + \ldots + b_{1m}a_{m1} \\ c_{pn} &= b_{p1}a_{1n} + b_{p2}a_{2n} + \ldots + b_{pm}a_{mn} \end{aligned} \tag{31}$$

More generally for any i and j,

$$c_{ij} = b_{i1}a_{1j} + b_{i2}a_{2j} + \ldots + b_{im}a_{mj} \tag{32}$$

since the last element of C is c_{pn}, C is a $p \times n$ matrix, that is a matrix of p rows and n columns, transforming the n quantities $x_1, \ldots, x_n$ into the p quantities $z_1, \ldots, z_p$. We have therefore

$$\begin{matrix} B & \times & A & = & C \end{matrix}$$

$$\begin{bmatrix} b_{11} & & b_{1m} \\ & & \\ b_{i1} & \cdots & b_{im} \\ & & \\ b_{p1} & & b_{pm} \end{bmatrix} \times \begin{bmatrix} a_{11} & a_{1j} & a_{1n} \\ & \vdots & \\ & \vdots & \\ & \vdots & \\ a_{m1} & a_{mj} & a_{mn} \end{bmatrix} = \begin{bmatrix} c_{11} & & c_{1j} & & c_{1n} \\ & & \vdots & & \\ c_{i1} & \cdots & c_{ij} & \cdots & c_{in} \\ & & \vdots & & \\ c_{p1} & & c_{pj} & & c_{pn} \end{bmatrix} \tag{33}$$

The element c_{ij} (ith row and jth column) is formed from the ith row of B and the jth column of A, by summing the products of corresponding elements. For B and A to conform, B must have the same number of columns as A has rows.

8. A simple example illustrates the use of matrix multiplication. Let the coordinates of a point in space, referred to a right-handed set of axes, be

(x_1, x_2, x_3), see Fig. 1.1. Let A represent an anticlockwise rotation of 90° about axis 1, then

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \tag{34}$$

This expresses the fact that this rotation leaves x_1 unchanged, gives a new x_2 equal to minus the old x_3, and a new x_3 equal to the old x_2.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ -x_3 \\ x_2 \end{bmatrix} \tag{35}$$

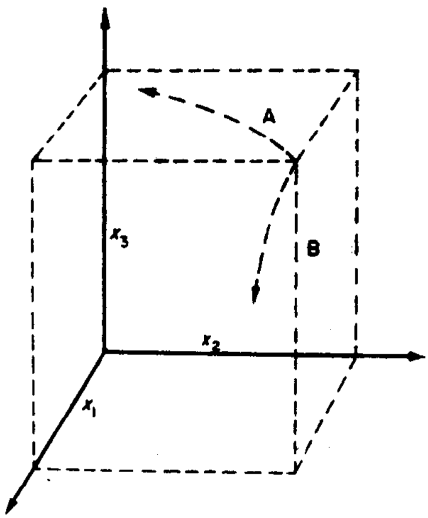

FIG. 1.1.

Let B represent an anticlockwise rotation of 90° about axis 2, then

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \tag{36}$$

If C represents a rotation of 90° about axis 1 followed by a rotation of 90° about axis 2 then

$$C = BA = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix} \tag{37}$$

by the rules for multiplication given in Sect. 7.

Now let F represent the operations done in the reverse order

$$F = AB = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \tag{38}$$

Obviously $F \neq C$, so that $BA \neq AB$, and this will be so unless A and B happen to be specially related. This demonstrates that for 90° rotations about fixed axes at right angles a change in the order of the operations makes a difference to the result. Also we have

$$\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} x_3 \\ x_1 \\ x_2 \end{bmatrix} \tag{39}$$

The effect of B followed by A is to permute the coordinates, and it is well known that this can be done by an anticlockwise rotation of 120° about an axis making equal angles with the original three axes. The relevance of this to the treatment of cubic symmetry will be clear, and we shall see later that matrices can readily be used to express the effects of rotation operations in a quite general way.

SCALAR PRODUCT

9. We saw in Sect. 7 that each element of a product matrix is the result of multiplying together corresponding elements of a row and a column and adding the answers. The result of a single set of operations of this type is called a scalar product, and the calculation of scalar products, in one guise or another, is an extremely common operation in calculations which can be expressed in matrix form. Note that if the elements of the row are the same as those of the column we form the sum of the squares of the elements in the column

$$[x_1\ x_2 \ldots x_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i^2 \tag{40}$$

It will appear that this has a number of applications and we need a sign to represent the operation of turning a column vector into a row vector so that we can write down a scalar product concisely. This is a special case of the matrix operation which we shall discuss in Sect. 10.

TRANSPOSITION

10. It is useful to have a notation for the matrix obtained by writing down each row of a given matrix as the corresponding column of the result. This interchange of rows and columns is called transposition and the result is known as the transpose A^T of the original matrix A. (Many authors write A' for the transpose, but we shall use A^T because the superscript ^{T} is a mnemonic aid.) If we have

$$A \equiv \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \tag{41}$$

then

$$A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix} \tag{42}$$

If A is of order $m \times n$ (m rows and n columns) then A^T is of order $n \times m$.

The transpose of a vector is a special case of this and we can rewrite eqn. (40) as

$$\mathbf{x}^T.\mathbf{x} = \sum_{i=1}^{n} x_i^2$$

This gives us a notation for a row vector and for a scalar product. Notice that the transpose of a scalar is equal to the scalar. Thus $\mathbf{x}^T.\mathbf{y} = \mathbf{y}^T.\mathbf{x}$ and so on.

11. The transpose of the product of two matrices can be written in terms of the transposed matrices. The i, j element of $(BA)^T$ is

$$(BA)_{ij}^T = (BA)_{ji} = b_{j1}a_{1i} + \dots + b_{jm}a_{mi} \tag{43}$$

where B and A are defined by eqns. (5) and (1). The i, j element of A^TB^T is

$$(A^TB^T)_{ij} = a_{1i}b_{j1} + \dots + a_{mi}b_{jm} \tag{44}$$

since $(A^T)_{ij} = (A)_{ji}$. Hence we have

$$(BA)^T = A^TB^T \tag{45}$$

and this can readily be generalised to

$$(Z \dots BA)^T = A^TB^T \dots Z^T \tag{46}$$

FORMS

12. From a given matrix it is possible to generate scalar quantities by means of multiplications with vectors. For any matrix A of order $m \times n$ we can calculate

$$\mathbf{y}^T A\mathbf{x} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} y_i x_j \tag{47}$$

where $\mathbf{y}$ and $\mathbf{x}$ are column vectors of order m and n respectively. This is called a bilinear form. If A is square and equal to its own transpose A^T, we can calculate

$$\mathbf{x}^T A\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j \tag{48}$$

which is called a quadratic form.

Notice that $\mathbf{y}^T A\mathbf{x}$ = a scalar = $\mathbf{x}^T A^T \mathbf{y}$.

NORMS

13. The norm of a vector $\mathbf{x}$ is usually written $\| \mathbf{x} \|$ and defined as the positive square root of $\mathbf{x}^T.\mathbf{x}$. There are various norms in use for matrices. The Euclidean norm is

$$\| A \|_e = (\sum_i \sum_j a_{ij}^2)^{\frac{1}{2}} \tag{49}$$

Another commonly used norm is the maximum value of the norm of the vector $A\mathbf{x}$ for all $\mathbf{x}$ with $\| \mathbf{x} \| = 1$ so that

$$\| A \| = \max_{\mathbf{x}} [\| A\mathbf{x} \| / \| \mathbf{x} \|] \tag{50}$$

A discussion of norms is beyond our scope here, and we merely remark that they can be used to express the overall magnitude of the elements of the matrix, for such purposes as analyses of the effects of rounding errors on numerical processes.

PARTITIONED MATRICES

14. For various purposes it may be convenient to make a distinction between different parts of a matrix. It may happen that we cannot handle the whole matrix because it is too large for the store of a computer, so that we deal with several sections in turn. We may know that certain parts of the matrix take a special form (e.g. that all elements in one area are zero) and wish to express this fact. These needs can be met by writing a partitioned matrix

$$F = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \tag{51}$$

Each part A, B, C, D of the matrix F is called a submatrix. Evidently, submatrices in the same row must have equal numbers of rows, and similarly for columns. There is no other limit on the number of submatrices we may have (up to mn, each of order 1×1). Partitioned matrices may be added and multiplied, provided that the appropriate submatrices have dimensions which conform. We can write

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]\left[\begin{array}{c|c} E & F \\ \hline G & H \end{array}\right]=\left[\begin{array}{c|c} (AE+BG) & (AF+BH) \\ \hline (CE+DG) & (CF+DH) \end{array}\right] \tag{52}$$

provided that the products AE, BG, . . . , DH can be formed.

CHAPTER 2

MATRIX INVERSION AND SOLUTION OF EQUATIONS

THE INVERSE TRANSFORMATION

1. Chapter 1 began with the linear transformation of a set of quantities $x_1, \ldots, x_n$ into another set $y_1, \ldots, y_m$,

$$\begin{matrix} a_{11}x_1 + \ldots + a_{1n}x_n = y_1 \\ \vdots \qquad\qquad \vdots \qquad \vdots \\ a_{m1}x_1 + \ldots + a_{mn}x_n = y_m \end{matrix} \tag{1}$$

We shall now consider the reverse process of obtaining from this transformation one which enables us to derive $\mathbf{x}$ from $\mathbf{y}$. It is well known that no unique exact solution can exist for arbitrary a_{ij} and y_i unless we have $m = n$ and we shall for the rest of this chapter concern ourselves with this case only. We rewrite (1) as

$$A\mathbf{x} = \mathbf{y} \tag{2}$$

where $\mathbf{x}$ and $\mathbf{y}$ are both vectors of order n and A is a square matrix of order $n \times n$, usually called a matrix of order n. If the transformation of $\mathbf{y}$ into $\mathbf{x}$ exists, we can write it

$$B\mathbf{y} = \mathbf{x} \tag{3}$$

It follows that

$$\begin{aligned} BA\mathbf{x} &= B\mathbf{y} = \mathbf{x} \\ AB\mathbf{y} &= A\mathbf{x} = \mathbf{y} \end{aligned} \tag{4}$$

The product matrix BA is equal to the product AB in this special case and is the matrix which leaves a vector unaltered. This matrix is called the identity matrix I and is

$$I \equiv \begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix} \tag{5}$$

The matrix B is called the inverse or reciprocal of the matrix A and can be written as A^{-1}, so that we have

$$\begin{aligned} A\mathbf{x} &= \mathbf{y} \\ A^{-1}\mathbf{y} &= \mathbf{x} \\ AA^{-1} &= A^{-1}A = I \\ AA^{-1}\mathbf{x} &= I\mathbf{x} = \mathbf{x} = A^{-1}A\mathbf{x} \end{aligned} \tag{6}$$

We note that a factor I can be inserted or suppressed anywhere that we please in a matrix product, provided that we understand it to mean the identity matrix of appropriate order.

2. We saw in Sect. 1 that the transformations (1) can be treated as equations for $\mathbf{x}$ which can be solved provided that the inverse matrix A^{-1} exists. The condition for the existence of A^{-1} is the same as the condition that the equations can be solved and we shall now consider what this is. If we have a pair of equations

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= y_1 \\ a_{21}x_1 + a_{22}x_2 &= y_2 \end{aligned} \tag{7}$$

we find by simple elimination that

$$\begin{aligned} x_1 &= (y_1a_{22} - y_2a_{12})/(a_{11}a_{22} - a_{12}a_{21}) \\ x_2 &= (y_2a_{11} - y_1a_{21})/(a_{11}a_{22} - a_{12}a_{21}) \end{aligned} \tag{8}$$

The solution can be obtained unless the denominator $a_{11}a_{22} - a_{12}a_{21}$ turns out to be zero. This quantity is called the determinant of the 2×2 matrix of the equations and the determinant of a square matrix A is written $|A|$ or $\det A$. For a matrix of order 3 we can write

$$\det A = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \tag{9}$$

This rule can be extended to give the expression for the determinant of a square matrix of any order and the solution of a set of simultaneous equations $A\mathbf{x} = \mathbf{y}$ of any order can be written in terms of quotients in which $\det A$ is the denominator. It follows that A^{-1} exists for any square matrix A provided that $\det A$ is not zero. If $\det A$ is zero, A is said to be singular, otherwise it is said to be non-singular. We state without proof that for a product

$$\det(AB\ldots Z) = \det A . \det B . \ldots \det Z \tag{10}$$

A singular transformation $A\mathbf{x} = \mathbf{y}$ produces a vector $\mathbf{y}$ in which one or more elements obey some law, whatever the vector with which we start. This law might be

$$y_3 = 0, \text{ all } y_i = 0,$$

$$y_4 = y_5 + y_7, \text{ or } y_n = \sum_{i=1}^{n-1} y_i$$

and so on. In all such cases many initial vectors transform into the same vector and we obviously cannot decide from the result which initial vector was concerned. No unique inverse transformation can be found in these circumstances.

3. We can add to the operations of matrix addition, subtraction and multiplication that of matrix division. If A is non-singular we may pre-divide B by A by forming $A^{-1}B$ and we may post-divide B by A by forming BA^{-1}, if B and A conform. The two results will not usually be the same.

4. We can form the reciprocal of a product by a reversal rule similar to that for the transpose obtained in Chapt. 1, Sect. 11. We have

$$(AB)(AB)^{-1} = I = ABB^{-1}A^{-1} \tag{11}$$

because we can cancel BB^{-1} ($= I$), and then AA^{-1}($= I$). Hence

$$(AB)^{-1} = B^{-1}A^{-1} \tag{12}$$

This can readily be extended to

$$(AB\ldots Z)^{-1} = Z^{-1}\ldots B^{-1}A^{-1} \tag{13}$$

5. For a square matrix of order n we can extract a sub-matrix of order r by deleting all except r rows and r columns (not necessarily the corresponding rows and columns). Such a matrix is called a minor. We can then form the determinant of the result. If one square sub-matrix of order r is non-singular but all of order $r+1$ are singular, then the matrix is said to be of rank r and nullity $n-r$.

6. It is possible to construct matrices A which are such that $AA^T = AA^{-1} = I$ so that $A^T = A^{-1}$. A matrix of this type is said to be orthogonal for this reason. If the vector $\mathbf{x}$ represents the coordinates of a point referred to equal orthogonal axes then $\mathbf{x}^T\mathbf{x}$ is the square of the distance from the origin. Any rotation of axes about the origin leaves this unaltered provided that the axes remain orthogonal. After a transformation A the new (origin distance)^2 is $\mathbf{x}^TA^TA\mathbf{x}$. If $A^TA = I$ this is the same as $\mathbf{x}^T\mathbf{x}$ for all $\mathbf{x}$. The matrix A then represents a rotation of orthogonal axes.

7. For a set of quantities $x_1, \ldots, x_n$ there may be a necessary relation

$$c_1 x_1 + c_2 x_2 + \ldots + c_n x_n = 0 \tag{14}$$

where the c_i are constants not all of which are zero.

If this is so then $x_1, \ldots, x_n$ are said to be linearly dependent. If no such relation exists they are linearly independent. This concept can be extended to vectors so that $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are linearly dependent if and only if there exists some relation

$$c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \ldots + c_n \mathbf{x}_n = 0 \tag{15}$$

where the c_i are scalar constants at least one of which is not zero, as for (14) above. If we write X as the matrix whose ith row is $\mathbf{x}_i^T$ then we can rewrite (15) as

$$\mathbf{c}^T X = \mathbf{0}^T, \ \mathbf{c}^T \neq 0 \tag{16}$$

where $\mathbf{c}^T$ is the row vector whose ith element is c_i. We state without proof that a square matrix is singular if and only if its rows (and its columns) are linearly dependent.

SOLUTION OF SIMULTANEOUS EQUATIONS

8. We shall now consider practical methods of computing the solution for a set of linear simultaneous equations taking into account the number of items to be stored, the number of operations and the precision which can be expected. We shall deal with equations $A\mathbf{x} = \mathbf{b}$ where A is a square matrix of order n and $\mathbf{x}$, $\mathbf{b}$ are vectors of order n. In this chapter we shall deal with direct methods only, that is with methods which produce the answer in a specified number of operations. Such methods can be classified as pivotal condensations and matrix decompositions. We have mentioned in Sect. 2 that the solution can be obtained by calculation of determinants. This is uneconomic in practice unless the order of the equations is very low.

9. Pivotal condensation is the process of eliminating the unknowns one by one from successive equations until the system is transformed into $U\mathbf{x} = \mathbf{c}$, where $U_{ij} = 0$ if $i>j$ so that U is upper triangular. (This means that all elements of U below the leading diagonal are zero.) We can then find x_n directly and " back substitute " for $x_{n-1}, \ldots, x_1$ in turn, since at each step one term alone in the equation concerned contains an unknown x_i.

10. Matrix decomposition involves the " factorisation " of A, which is the determination of two non-singular matrices L and U which are such that $A = LU$ where L is lower triangular (all elements above the diagonal zero) and U is upper triangular. We can then write $U\mathbf{x} = \mathbf{y}$, so that $A\mathbf{x} = LU\mathbf{x} = L\mathbf{y} = \mathbf{b}$ and so determine $y_1, y_2, \ldots, y_n$ in turn. We then have $U\mathbf{x} = \mathbf{y}$ and we can find $x_n, x_{n-1}, \ldots, x_1$ in turn.

GAUSS ELIMINATION

11. There are various ways of carrying out pivotal condensation and it would be pointless to try to explain all of them. Instead we will describe the single method of Gauss elimination, which can be used for any non-singular matrix. The algorithm is

1. Set a counter $i = 1$, to count rows of the matrix;
2. Set $j = i$, to count the elements of column i, go to 18 if $i = n$;
3. Set a register l to zero, to store the largest element in column i;
4. Compare l and $|a_{ji}|$, if $|a_{ji}| > l$, set $l = |a_{ji}|$ and $k = j$;