0151.21
E602
197801004
197801004
外文书库

An Introduction to the Theory of Canonical Matrices

BY

H. W. TURNBULL, M.A., F.R.S.
Regius Professor of Mathematics in the United College, University of St. Andrews. Formerly Scholar of Trinity College, Cambridge, and Fereday Fellow of St. John's College, Oxford

AND

A. C. AITKEN, D.Sc.
Lecturer on Statistics and Actuarial Mathematics in the University of Edinburgh

暨南大学
数学系資料

PREFACE

This book has been written with the object of giving an account of the various ways in which matrices of finite order can be reduced to canonical form under different important types of transformation. While the work has been planned to serve as a sequel to a former publication, *The Theory of Determinants, Matrices, and Invariants* (1928), circumstances have allowed us to make it practically independent and self-contained, with the least possible overlapping of material in the two books. A certain knowledge of the elementary theory of determinants is presupposed, but no previous acquaintance with matrices.

The volume on *Invariants*—as it will be referred to in subsequent pages—in giving an introductory account of matrices and determinants, treated only of such properties as belonged to the general linear transformation; for these are the properties which have the most direct bearing on the projective invariant theory, to which the later chapters were devoted. In the nomenclature of the work before us, the treatment was confined to the diagonal case of the classical canonical form, in which the elementary divisors are necessarily linear.

In the present work we return to consider, in close detail, those important cases in which the elementary divisors are no longer restricted to be linear, but may be of general degree. To adopt a geometrical mode of speaking, it is as if we had formerly been concerned purely with the projective properties of quadrics in general position, but had now returned to the consideration of all possible distinctions between quadrics under certain prescribed conditions; such distinctions, for example, as those which persist through all projective transformations, or again through all rotations, and so on.

The subject-matter of the canonical reduction of matrices, which has numerous and important applications, has received attention in several treatises and a large number of original papers. The historical notes which we have appended to each chapter are intended to give a brief review of what has been done on each topic, to apportion due credit to pioneers, and to stimulate the student to further reading. (We would warn him, however, to make sure at the outset, in reading any work on groups or matrices, whether the author means AB or what we have denoted by BA when he writes a product.) The most complete accounts of the theory available are those of Muth (*Elementarteiler*, 1899) and Cullis (*Matrices and Determinoids*, Vols. I, II, III, 1913, 1918, 1925). We have preferred to follow the lead of Cullis, who develops the theory in terms of the structure and properties of matrices—in matrix idiom, as it were, rather than in terms of bilinear and quadratic forms, or of linear substitutions.

We take the opportunity of acknowledging our indebtedness to the work of those writers who have given a sustained account of the theory, in one guise or another; in particular to Muth, as above, to Bromwich (Cambridge Tract on *Quadratic Forms*, 1906, and various papers), to Bôcher (*Higher Algebra*, 1907), Hilton (*Linear Substitutions*, 1914), Cullis (Vol. III of *Matrices and Determinoids*, 1925), and Dickson (*Modern Algebraic Theories*, 1926).

While we have tried to include all the principal features of the theory and have sought to make the sequence of argument reasonably fluent, even allowing ourselves moderate latitude in digression and explanation, we have, at the same time, aimed at a certain compactness in the formulæ and demonstrations. This has been achieved in the first place by a systematic use of the matrix notation, to which we shall again refer; in the second place, by confining the contents of each chapter almost entirely to general theorems, and by relegating corollaries and applications to the interspersed sets of examples. These examples are intended to serve not so much as exercises, many being quite easy, but rather as points of relaxation. and running commentary; they will, however, be found to contain many well-known and important theorems, which the notation establishes in the minimum of space.

We attach the greatest importance to the choice of notation. Inferring from perusal of Cullis that the emphasis laid since the time of Cayley on the square matrix might well be removed, we resolved to continue the plan adopted in *Invariants* by making the fullest use of rectangular matrices and submatrices, and of partitioned matrices, by insisting on the condition that the non-commutative rules of product order hold without exception, and by distinguishing always between a matrix of a single row and one of a single column. When this is done, all the systems which appear, whether scalars, vectors, or matrices, can be regarded as rectangular matrices or products of rectangular matrices, and the theory is thus greatly unified. We would draw special attention to the notation $x'Ay$ for the bilinear form, $x'Ax$ for the quadratic, and $\bar{x}'Ax$ for the Hermitian form, believing that these notations will enable the linear transformations and the bilinear, quadratic, and Hermitian forms which are fundamental, for example, in analytical geometry, dynamics, or mathematical statistics, to be manipulated with ease.

Through considerations of space we have not been able to include many applications to geometry, but the results are readily adaptable: nor to the theory of Groups, where, as Schur has shown, partitioned matrices can be used with elegance and advantage.

The reader already familiar with the theory will also observe that certain established methods of dealing with the subject have hardly been touched upon, notably the methods of Weierstrass and Darboux, the theory of regular minors of determinants and the treatment of quadratic forms by the methods of Kronecker. We have, in fact, allowed ourselves a free hand in dealing with the results of earlier writers, in the belief that the outcome would prove to be an easier approach to a subject that has often failed to win affection; and the methods of H. J. S. Smith, Sylvester, Frobenius, and Dickson proved in themselves quite adequate without the inclusion of other parallel theories. A thorough assimilation of the algebraic implications of Euclid's H.C.F. process, and of the notion of linear dependence, furnishes the clue to many passages. Our tribute to Kronecker finds expression in Chapter IX, which is an essay towards giving a fresh derivation of his classical results concerning singular pencils; we have treated this by rational methods, and we trust that an intricate argument has been materially simplified.

Our best thanks are due to Dr. E. T. Copson and Mr. D. E. Rutherford at St. Andrews University, who have taken an interest in the progress of the work, and have offered valuable suggestions at the proof-reading stage; and especially to Dr. John Dougall, for his critical vigilance and expert mathematical and technical help during the passage of the work through the press.

H. W. T.
A. C. A.

ST. ANDREWS
EDINBURGH } *December*, 1931.

此为试读，需要完整PDF请访问： www.ertongbook.com

CONTENTS

CHAPTER I

DEFINITIONS AND FUNDAMENTAL PROPERTIES OF MATRICES

CHAPTER II

ELEMENTARY TRANSFORMATIONS. BILINEAR AND QUADRATIC FORMS

CHAPTER III

THE CANONICAL REDUCTION OF EQUIVALENT MATRICES

CHAPTER IV

SUBGROUPS OF THE GROUP OF EQUIVALENT TRANSFORMATIONS

CHAPTER V

A RATIONAL CANONICAL FORM FOR THE COLLINEATORY GROUP

CONTENTS

CHAPTER VI

THE CLASSICAL CANONICAL FORM FOR THE COLLINEATORY GROUP

CHAPTER VII

CONGRUENT AND CONJUNCTIVE TRANSFORMATIONS: QUADRATIC AND HERMITIAN FORMS

CHAPTER VIII

CANONICAL REDUCTION BY UNITARY AND ORTHOGONAL TRANSFORMATIONS

CONTENTS

CHAPTER IX

THE CANONICAL REDUCTION OF PENCILS OF MATRICES

CHAPTER X

APPLICATIONS OF CANONICAL FORMS TO SOLUTION OF LINEAR MATRIX EQUATIONS. COMMUTANTS AND INVARIANTS

CONTENTS

CHAPTER XI

PRACTICAL APPLICATIONS OF CANONICAL REDUCTION

The Theory of Canonical Matrices

CHAPTER I

DEFINITIONS AND FUNDAMENTAL PROPERTIES OF MATRICES

1. Introductory.

The theory of canonical matrices is concerned with the systematic investigation of types of transformation which reduce matrices to the simplest and most convenient shape. The formulation of these various types is not merely useful as a preliminary to the deeper study of the properties of matrices themselves; it serves also to render the theory of matrices more immediately available for numerous applications to geometry, differential equations, analytical dynamics, and the like. Quite early, for example, in co-ordinate geometry, when the equation of a general conic is simplified by reference to principal axes, or again when two general conics are referred to their common self-conjugate triangle, the procedure involved is really equivalent to the canonical reduction of a matrix.

2. Definitions and Fundamental Properties.

It will be of advantage to recall briefly the definitions and fundamental properties of matrices. By a matrix A of order n is meant a system of elements, which may be real or complex numbers, arranged in a square formation of n rows and columns,

$$A = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \ldots \quad (1)$$

此为试读，需要完整PDF请访问：www.ertongbook.com

where a_{ij} denotes the element standing in the ith row and the jth column, the (ij)th element, as we shall frequently call it. The determinant having the same elements is denoted by $|A|$, or $|a_{ij}|$, and is naturally called the determinant of the matrix A. We shall also make continual use of *rectangular* matrices, of m rows and n columns, or, as it will be phrased. of order m by n, $m \times n$. Where there is only one row, so that $m = 1$, such a matrix will be termed a *vector of the first kind*. or a *prime*; and it will often be denoted by a single small italic u or v. Thus

$$u = [u_1, u_2. \ldots, u_n]. \qquad (2)$$

On the other hand, a matrix of a single column, of n elements, will be termed a *vector of the second kind*, or a *point*; and to save space it will not be printed vertically but horizontally, and distinguished by brackets $\{\ldots\}$. Thus

$$x = \{x_1, x_2, \ldots, x_n\}. \qquad (3)$$

The accented matrix $A' = [a_{ji}]$, obtained by complete interchange of rows and columns in A, is called the *transposed* of A. The ith row $[a_{i1}, a_{i2}, \ldots, a_{in}]$ of A is identical with the ith column of A'. For vectors we have $u' = [u_j]' = \{u_j\}$, $x' = \{x_i\}' = [x_i]$.

Matrices may be multiplied either by ordinary *scalar* numbers or by matrices. The effect of multiplying a matrix $A = [a_{ij}]$ by a scalar λ is to multiply each element of A by λ. The product is defined by

$$\lambda A = \lambda[a_{ij}] = [\lambda a_{ij}] = A\lambda. \qquad (4)$$

Matrices of the same order are added, or subtracted, by adding, or subtracting, corresponding elements; so that a linear combination of two such matrices A and B, with scalar multipliers λ and μ, is defined by

$$\lambda A + \mu B = [(\lambda a_{ij} + \mu b_{ij})]. \qquad (5)$$

Hence, if $C = \lambda A + \mu B$, then $c_{ij} = \lambda a_{ij} + \mu b_{ij}$: and also $C' = \lambda A' + \mu B'$, for the transposed matrices.

The *null* or *zero matrix*, whether square or rectangular, has all its elements zero, and will often be denoted without ambiguity by an ordinary cipher. The *unit matrix*, I, is necessarily square; it has a unit for each element in the principal diagonal, and the remaining elements all zero. Thus

$$I = [\delta_{ij}], \quad \delta_{ij}\begin{cases} = 0, & i \neq j, \\ = 1, & i = j. \end{cases} \qquad (6)$$

3. Matrix Multiplication.

The multiplication of matrices by matrices, or matrix multiplication, differs in important respects from scalar multiplication. Two matrices can be multiplied together only when the number of columns in the first is equal to the number of rows in the second. Matrices which satisfy this condition will be termed *conformable* matrices; their product AB is defined by

$$AB = [a_{ij}][b_{ij}] = [\sum_{k=1}^{p} a_{ik}b_{kj}] = [c_{ij}] = C, \quad . \quad . \quad . \quad (7)$$

where the orders of A, B, C are $m \times p$, $p \times n$, $m \times n$ respectively. The process of multiplication is thus the same as the *row-by-column* rule for multiplying together determinants of equal order. If the matrices are square and each of order n, then the corresponding relation $|A||B| = |C|$ is true for the determinants $|A|$, $|B|$, $|C|$.

Matrices are regarded as equal only when they are element for element identical. Therefore, since a row-by-column rule will in general give different elements from a column-by-row rule, the product BA, if it exists at all, is usually different from AB. (AB and BA, it may be observed, can coexist only if $m = n$.) We must therefore distinguish always between *premultiplication*, as when B, premultiplied by A, yields the product AB, and *postmultiplication*, as when B, postmultiplied by A, yields the product BA. If $AB = BA$ the matrices A and B are said to *commute*, or to be *permutable*, and one of the applications of the theory of canonical matrices is to find the general matrix X permutable with a given matrix A. Except for the non-commutative law of multiplication (and therefore of division, defined as the inverse operation) all the ordinary laws of algebra apply to matrices, very much as they do in the elementary theory of vectors. Of particular importance is the associative law $(AB)C = A(BC)$, which allows us to dispense with brackets and to write ABC without ambiguity, since the double summation $\sum_k \sum_l a_{ik}b_{kl}c_{lj}$ can be carried out in either of the orders indicated. Similarly for the sum $A + B + C$.

The above remarks are restricted to the case of matrices of finite order; for the associative law of multiplication does not necessarily hold when any of the matrices involved has one or both of the orders m, n infinite.

The integers m, n, p which appear in (7) may take any positive value. One extreme case, when $m = n = 1$, yields the *inner product* of the vectors u and x. Thus

$$ux = u_1x_1 + u_2x_2 + \ldots + u_px_p = \sum_k u_kx_k = x'u'. \quad . \quad (8)$$

The product here is a scalar. On the other hand, the product xu, which exemplifies (7) with $p = 1$, $m = n$, is a square matrix of order n, having $x_i u_j$ for its (ij)th element, namely

$$xu = [x_i u_j] = (u'x')'. \quad \quad (9)$$

4. Reciprocal of a Non-Singular Matrix.

When the determinant $|A| = |a_{ij}|$ of a square matrix A does not vanish A is said to be *non-singular*, and possesses a *reciprocal* or *inverse* matrix R such that

$$AR = RA = I.$$

The reciprocal R is unique, as will be seen, and is readily obtained, from the theory of determinants. If A_{ij} denotes the co-factor of a_{ij} in $|A|$, the matrix $[A_{ji}]$ is called the *adjoint* of A, and exists whether A is singular or not. (The determinant $|A_{ij}|$ is the *adjugate* of $|A|$.) It follows that

$$[a_{ij}][A_{ji}] = [\sum_k a_{ik} A_{jk}] = [\,|A|\,\delta_{ij}] = |A|\,I. \quad \quad (10)$$

Thus the product of A and its adjoint is that special type of diagonal matrix called a *scalar matrix*; each diagonal element $(i = j)$ is equal to the determinant $|A|$, and the rest are zero. If $|A| \neq 0$, we may divide throughout by the scalar $|A|$, obtaining at once the required form of R. The (ij)th element of R is therefore $A_{ji}|A|^{-1}$, or, let us say, a^{ji}, where the *reversed* order of upper indices must be carefully noted. Writing now A^{-1} instead of R, we have

$$A^{-1} = [a^{ji}] = [A_{ji}|A|^{-1}], \qquad |A| \neq 0. \quad \quad (11)$$

By actual multiplication $AA^{-1} = A^{-1}A = I$; so that the name reciprocal and the notation A^{-1} are justified. It may be observed in passing that in products of matrices the unit factor I may be introduced or suppressed at pleasure, like the unit factor of scalar algebra.

5. The Reversal Law in Transposed and Reciprocal Products.

A fundamental consequence of the non-commutative law of matrix multiplication is the *reversal* law, exemplified in transposing and reciprocating a continued product of matrices. Thus

$$(AB)' = B'A', \quad (ABC)' = C'B'A'; \quad \quad (12)$$

and, if $|A| \neq 0$, $|B| \neq 0$, $|C| \neq 0$,

$$(AB)^{-1} = B^{-1}A^{-1}, \quad (ABC)^{-1} = C^{-1}B^{-1}A^{-1}. \quad \quad (13)$$

EXAMPLES

1. Prove that the reciprocal of a non-singular matrix is unique. [If $AR = I$, and also $AS = I$, then $AR - AS = 0$, the null matrix. By the distributive law $A(R-S) = 0$, and hence $A^{-1}A(R-S) = 0$. Thus $I(R-S) = R-S = 0$; and so $R = S$. All solutions X of the equation $AX = I$ are therefore equal. But A^{-1} is a solution and is therefore the unique solution.]

2. Verify (13) by premultiplying by B, A, or C, B, A in turn.

3. Prove that $[a^{ij}][b^{ij}] = [c^{ij}]$, where $i, j = 1, 2, \ldots, n$, provided that $c^{ij} = \sum_{k=1}^{n} a^{ik}b^{kj}$.

4. If A is a square matrix of order n, while u and x are vectors of the row and column kinds respectively, then uA denotes a row vector while Ax denotes a column vector. The products Au, xA are undefined if $n > 1$.

5. What do $A'u'$, $x'A'$ represent? [Column vector, row vector.]

6. Matrices partitioned into Submatrices.

It is convenient to extend the use of the fundamental laws of combination for matrices to the case where a matrix is regarded as constructed not so much from elements as from submatrices, or minor matrices, of elements. (Cf. *Invariants*, p. 38.) For example, the matrix

$$A = \left[\begin{array}{cc:c} 1 & 2 & 3 \\ 4 & 5 & 6 \\ \hdashline 7 & 8 & 9 \end{array}\right]$$

can be written

$$A = \begin{bmatrix} P & Q \\ R & S \end{bmatrix},$$

where

$$P = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad Q = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad R = [7, 8], \quad S = [9].$$

Here the diagonal submatrices P and S are square, and the partitioning is diagonally symmetrical. In the general case there may evidently be n or fewer partitions row-wise or column-wise. Let B be a second square matrix of the third order similarly partitioned:

$$B = \begin{bmatrix} P_1 & Q_1 \\ R_1 & S_1 \end{bmatrix} = \left[\begin{array}{cc:c} 2 & . & 1 \\ 3 & 1 & 2 \\ \hdashline 1 & 2 & . \end{array}\right];$$

then by addition and multiplication we have

$$A+B = \begin{bmatrix} P+P_1 & Q+Q_1 \\ R+R_1 & S+S_1 \end{bmatrix}, \quad AB = \begin{bmatrix} PP_1+QR_1 & PQ_1+QS_1 \\ RP_1+SR_1 & RQ_1+SS_1 \end{bmatrix}, \qquad (14)$$

as may readily be verified. In each case the resulting matrix is of the same order, and is partitioned in the same way, as the original matrix factors. For example, in AB the first element, $PP_1 + QR_1$, stands for a square submatrix of two rows and columns: and this is possible since, by definition, both products PP_1 and QR_1 consist of two rows and two columns. Similar remarks apply to the other submatrix " elements ". Thus

$$PQ_1 + QS_1 = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ 6 \end{bmatrix} [0] = \begin{bmatrix} 5 \\ 14 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 14 \end{bmatrix},$$

giving the proper rectangular shape for the upper right-hand minor.

It was observed earlier that a rectangular matrix B could be pre-multiplied by another rectangular matrix A, provided that the number of rows in B were equal to the number of columns in A. If A and B are both partitioned into submatrices such that the grouping of columns in A agrees exactly with the grouping of rows in B, it is not difficult to show that the product AB can equally well be obtained by treating the submatrices as elements and proceeding according to rule.

The case of square matrices of the same order, similarly and *symmetrically* partitioned, is important. Let A and B be two such matrices, and let A_{ij} henceforth denote the (ij)th submatrix in the partitioned form of A. (There will be little further occasion in this book to refer to determinantal co-factors, and the notation A_{ij} is well suited to the new concept.) Then if p, r are the orders of A_{ik}, those of B_{ki} are r, p, and those of another minor with the same k, as B_{kj}, will be r, q with the same r. For each value of k the product $A_{ik}B_{kj}$ is thus a submatrix of orders p, q; the sum $\sum_k A_{ik}B_{kj}$ can therefore be formed, and gives the (ij)th submatrix of the product AB, where the latter is in partitioned form similar to A and B. We have then, for matrices $A = [a_{ij}]$, $B = [b_{ij}]$, similarly and symmetrically partitioned,

$$AB = [\sum_k A_{ik}B_{kj}] = [C_{ij}], \quad \ldots \ldots \quad (15)$$

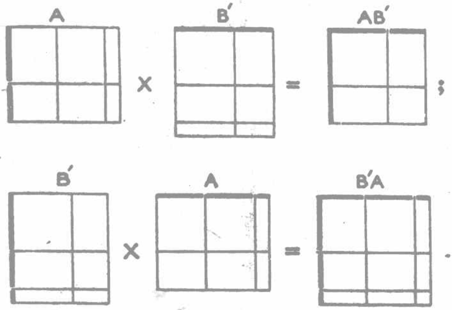
where, of course, in each term of C_{ij} the order A, B is preserved.

Similarly but *unsymmetrically* partitioned square matrices A and B cannot be multiplied together by a rule of this kind; each can, however, be multiplied by the *transposed* matrix of the other, for then the partitioning of the column-groups of the multiplier agrees with that of the row-groups of the multiplicand. The transposed matrix A' is readily seen to be $A' = [A'_{ji}]$, where the minor matrix A'_{ji} is itself transposed, and for matrices A, B similarly but unsymmetrically partitioned,

$$AB' \doteq [\sum_k A_{ik} B_{jk}]. \qquad (16)$$

EXAMPLES

1. If A and B are similarly but unsymmetrically partitioned, with μ partitions into row-groups and ν into column-groups, show that the product AB' is symmetrically partitioned according to the μ row-partitions of A; and that $B'A$ is symmetrically partitioned according to the ν column-groups of A.



2. Distinguish by examples between a symmetrical matrix and a symmetrically partitioned matrix.

3. If

$$C = \begin{bmatrix} A & x \\ u & \cdot \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & x_1 \\ a_{21} & a_{22} & a_{23} & x_2 \\ a_{31} & a_{32} & a_{33} & x_3 \\ u_1 & u_2 & u_3 & \cdot \end{bmatrix},$$

prove that $C^2 = \begin{bmatrix} A^2 + xu & Ax \\ uA & ux \end{bmatrix}$.

[Note that xu is a square matrix of order three, while ux is a scalar of the first order.]

4. If

$$A = \begin{bmatrix} \lambda & 1 & \cdot & \cdot & \cdot \\ \cdot & \lambda & \cdot & \cdot & \cdot \\ \cdot & \cdot & \mu & 1 & \cdot \\ \cdot & \cdot & \cdot & \mu & \cdot \\ \cdot & \cdot & \cdot & \cdot & \nu \end{bmatrix} = \begin{bmatrix} L & \cdot & \cdot \\ \cdot & M & \cdot \\ \cdot & \cdot & N \end{bmatrix},$$

where $L = \begin{bmatrix} \lambda & 1 \\ \cdot & \lambda \end{bmatrix}$, $M = \begin{bmatrix} \mu & 1 \\ \cdot & \mu \end{bmatrix}$, $N = \nu$, $(\lambda, \mu, \nu \neq 0)$, find the values of A^2, A^p, A^{-1}, and of any rational function $f(A)$ in terms of L, M, N.

[In general $f(A) = \begin{bmatrix} f(L) & \cdot & \cdot \\ \cdot & f(M) & \cdot \\ \cdot & \cdot & f(N) \end{bmatrix}$.]

此为试读，需要完整PDF请访问：www.ertongbook.com