The Carus Mathematical Monographs

NUMBER SEVEN

VECTORS AND MATRICES

CYRUS COLTON MACDUFFEE

Professor of Mathematics
University of Wisconsin

Published by

THE MATHEMATICAL ASSOCIATION OF AMERICA

INTRODUCTION

The theory of matrices had its origin in the theory of determinants, and the latter had its origin in the theory of systems of equations. From Vandermonde and Laplace to Cayley, determinants were cultivated in a purely formal manner. The early algebraists never successfully explained what a determinant was, and indeed they were not interested in exact definitions.

It was Cayley who seems first to have noticed that "the idea of matrix precedes that of determinant." More explicitly, we can say that the relation of determinant to matrix is that of the absolute value of a complex number to the complex number itself, and it is no more possible to define determinant without the previous concept of matrix or its equivalent than it is to have the feline grin without the Cheshire cat.

In fact, the importance of the concept of determinant has been, and currently is, vastly over-estimated. Systems of equations can be solved as easily and neatly without determinants as with, as is illustrated in Chapter I of this Monograph. In fact, perhaps ninety per cent of matric theory can be developed without mentioning a determinant. The concept is necessary in some places, however, and is very useful in many others, so one should not push this point too far.

In the middle of the last century matrices were approached from several different points of view. The paper of Hamilton (1853) on "Linear and vector functions" is considered by Wedderburn to contain the beginnings of the theory. After developing some properties of "linear transformations" in earlier papers, Cayley finally wrote "A Memoir on the Theory of Matrices" in 1858 in which a matrix is considered as a single mathematical quantity. This paper gives Cayley considerable claim to the honor of introducing the modern concept of matrix, although the name is due to Sylvester (1850).

In 1867 there appeared the beautiful paper of Laguerre entitled "Sur le calcul des systèmes linéaires" in which matrices were treated almost in the modern manner. It attracted little attention at the time of its publication. Frobenius, in his fundamental paper "Ueber lineare Substitutionen und bilineare Formen" of 1878, approached matric theory through the composition of quadratic forms.

In fact, Hamilton, Cayley, Laguerre and Frobenius seem to have worked without the knowledge of each others' results. Frobenius, however, very soon became aware of these earlier papers and eventually adopted the term "matrix."

One of the central problems in matric theory is that of similarity. This problem was first solved for the complex field by means of the elementary divisor theory of Weierstrass and for other rings by H. J. S. Smith and Frobenius.

In the present century a number of writers have made direct attacks upon the problem of the rational reduction of a matrix by means of similarity transformations. S. Lattès in 1914 and G. Kowalewski in 1916 were among the pioneers, Kowalewski stating that his inspiration came from Sophus Lie. Since that time many versions of the rational reduction have been published by Dickson, Turnbull and Aitken, van der Waerden, Menge, Wedderburn, Ingraham, and Schreier and Sperner.

The history of these rational reductions has been interesting and not without precedent in the field of mathematical research. The early reductions were short, requiring only a few pages. It is not prudent to say that any of the early papers is incorrect, for certainly a correct result was obtained in each case, but some of them contained arguments which were convincing only to their authors. The exposition in places was certainly too brief. Later writers subjected these difficult passages to closer scrutiny, as well as to the fierce fire of generalization, with the result that an adequate treatment was found to take many pages. The book of Schreier and Sperner, to which the present writer acknowledges indebtedness, contains 133 pages.

A large part of the profit which has come from this mathematical Odyssey has been the by-products. In attempting to justify certain steps in the proof, basic theorems on vectors and matrices were uncovered, theorems which had not previously come to notice. Of this origin are the theorems on the polynomial factors of the rank equation of a matrix—facts which should have been known long ago but which for some peculiar reason escaped discovery.

The present book is an attempt to set forth the new technique in matric theory which the writers on the rational reduction have developed. The long proofs have been broken down into simpler components, and these components have been proved as preliminary theorems in as great generality as appeared possible. With the background developed in the first five chapters, the rational reduction of Chapter VI does not seem difficult or unnatural.

That the vector technique will have other applications in matric theory than to the problem which brought it forth is quite certain. The Weyr theory for a general field was easily established (§55) once the key theorem (Corollary 57) was known. The orthogonal reduction (Chapter VIII) surrendered without a struggle.

The author wishes to express his appreciation of the kindness of Professors Richard Brauer, Marguerite Darkow, Mark Ingraham, and Saunders MacLane, who have read the manuscript and offered valuable suggestions. While no attempt has been made to credit ideas to their discoverers, it should not be out of place to state that the author has been greatly influenced by the work, much of it unpublished, of his former colleague, Mark Ingraham.

CYRUS COLTON MACDUFFEE

HUNTER COLLEGE OF THE CITY OF NEW YORK
September 1, 1942

TABLE OF CONTENTS

VECTORS AND MATRICES

CHAPTER I

SYSTEMS OF LINEAR EQUATIONS

1. Graphs. A *solution* of the equation

$$2x + 3y - 6 = 0$$

is a pair of numbers (x_1, y_1) such that

$$2x_1 + 3y_1 - 6 = 0.$$

There are infinitely many such solutions. A solution of the system of equations

$$(1) \qquad \begin{aligned} 2x + 3y - 6 &= 0, \\ 4x - 3y - 6 &= 0 \end{aligned}$$

is a pair of numbers (x_1, y_1) which is a solution of both equations. There exists just one such solution, namely $(2, 2/3)$.

If we picture (x, y) as a point on the Cartesian plane, the infinitely many solutions of the equation

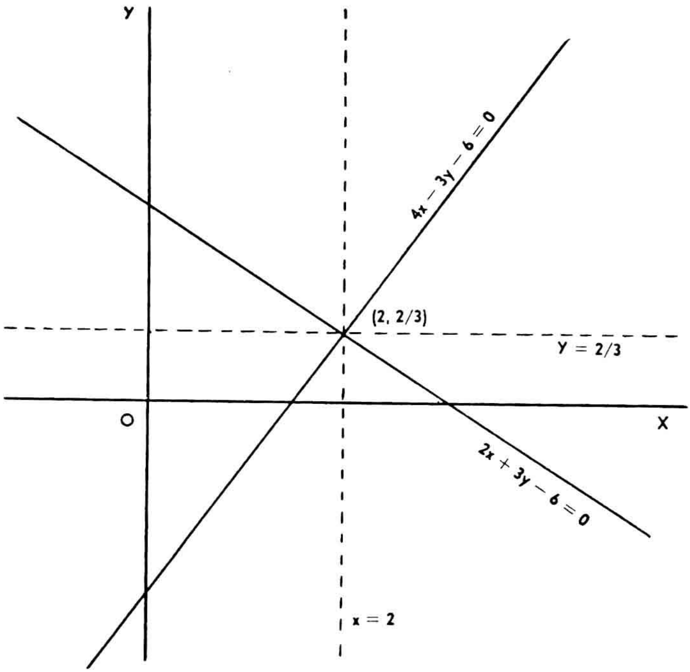
$$2x + 3y - 6 = 0$$

are the points of a straight line l_1, known as the *graph* of the equation. The second equation

$$4x - 3y - 6 = 0$$

also has a graph l_2 which is a straight line. The point of intersection of the two lines, namely $(2, 2/3)$, is the solution of the system of the two equations.

The point $(2, 2/3)$ is evidently the point of intersection of the line $x = 2$ with the line $y = 2/3$. Thus the problem of solving the system of equations (1) is equivalent to the problem of finding the vertical line and the horizontal line which pass through the intersection point of their graphs.



All methods of solving a system of equations such as (1) are but variations of one and the same process. Let k_1 and k_2 be any two numbers not both 0. The equation

$$k_1(2x + 3y - 6) + k_2(4x - 3y - 6) = 0,$$

or

$$(2) \quad (2k_1 + 4k_2)x + (3k_1 - 3k_2)y - 6k_1 - 6k_2 = 0,$$

is clearly the equation of a straight line, for the coefficients of x and y cannot both be 0 unless $k_1 = k_2 = 0$. This line passes through the intersection point of the two given lines; for if (x_1, y_1) is this intersection point, it is true for all values of k_1 and k_2 that

$$k_1(2x_1 + 3y_1 - 6) + k_2(4x_1 - 3y_1 - 6) = k_1 \cdot 0 + k_2 \cdot 0 = 0.$$

Now for various choices of k_1 and k_2, the line (2) represents every line of the plane through (x_1, y_1). This can be proved by showing that, if (x_2, y_2) is an arbitrarily chosen point of the plane different from (x_1, y_1), there is a choice of k_1, k_2 not both zero such that (2) passes through this point. Let k_1, k_2 be unknown, and set

$$k_1(2x_2 + 3y_2 - 6) + k_2(4x_2 - 3y_2 - 6) = 0.$$

We may choose

$$k_1 = 4x_2 - 3y_2 - 6, \qquad k_2 = -2x_2 - 3y_2 + 6.$$

Since (x_2, y_2) is not on both the given lines, not both k_1 and k_2 will be 0.

As the ratio $k_1 : k_2$ varies, the line (2) turns about the point (x_1, y_1). The problem of solving the system (1) is the problem of finding the values of k_1 and k_2 such that (2) is first vertical, then horizontal.

For (2) to be vertical, it is necessary and sufficient that the coefficient of y, namely $3k_1 - 3k_2$, shall be 0. Let $k_1 = k_2 = 1$. Then (2) becomes

$$6x + 0y - 12 = 0,$$

whence $x_1 = 2$. For (2) to be horizontal, it is necessary and sufficient that the coefficient of x, namely $2k_1 + 4k_2$, shall be 0. Let $k_1 = 2$, $k_2 = -1$. Then

$$0x + 9y - 6 = 0,$$

whence $y_1 = 2/3$.

2. Equivalence of systems. The principle illustrated in the above example is of general applicability, but when more than three unknowns are involved, or when the coefficient field is not real, the geometric interpretation becomes artificial. Suppose that we have a system of m equations in n unknowns,

$$
(3) \qquad \begin{aligned}
f_1 &= a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n - c_1 = 0, \\
f_2 &= a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n - c_2 = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
f_m &= a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n - c_m = 0,
\end{aligned}
$$

with coefficients in any field. A *solution* of the equation $f_i = 0$ is a set of numbers $(x_1', x_2', \cdots, x_n')$ such that

$$a_{i1}x_1' + a_{i2}x_2' + \cdots + a_{in}x_n' - c_i = 0.$$

A solution of the system (3) is a set of numbers which is a solution of every equation of the system.

Suppose that there is another system of equations

$$
(4) \qquad \begin{aligned}
g_1 &= b_{11}x_1 + b_{12}x_2 + \cdots + b_{1n}x_n - d_1 = 0, \\
g_2 &= b_{21}x_1 + b_{22}x_2 + \cdots + b_{2n}x_n - d_2 = 0, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
g_k &= b_{k1}x_1 + b_{k2}x_2 + \cdots + b_{kn}x_n - d_k = 0.
\end{aligned}
$$

The two systems (3) and (4) are called *equivalent* if every solution of each is a solution of the other.

The process of *solving* a system of equations is the process of finding an equivalent system of simplest possible form. Thus (3) is equivalent to

$$
\begin{aligned}
x_1 &= h_{11}p_1 + h_{12}p_2 + \cdots + h_{1l}p_l + e_1, \\
x_2 &= h_{21}p_1 + h_{22}p_2 + \cdots + h_{2l}p_l + e_2, \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
x_n &= h_{n1}p_1 + h_{n2}p_2 + \cdots + h_{nl}p_l + e_n
\end{aligned}
$$

where $p_1, p_2, \cdots, p_l$ are parameters which may assume arbitrary values in the coefficient field.

Let us consider System (3). Let $k_1, k_2, \cdots, k_m$ be any numbers, and form the linear equation

$$k_1f_1 + k_2f_2 + \cdots + k_mf_m = 0.$$

Consider the system of equations

$$\begin{aligned} f_1 &= 0, \\ &\cdots\cdots\cdots \\ f_{i-1} &= 0, \\ k_1f_1 + k_2f_2 + \cdots + k_mf_m &= 0, \\ f_{i+1} &= 0, \\ &\cdots\cdots\cdots \\ f_m &= 0. \end{aligned} \tag{5}$$

Clearly every solution of System (3) is a solution of System (5). Conversely, let $(x_1', x_2', \cdots, x_n')$ be any solution of (5). It is evidently a solution of every equation of (3) except possibly $f_i=0$. But the i-th equation of (5) reduces to $k_if_i=0$, so if $k_i \neq 0$, it is also true that $f_i=0$, and every solution of (5) is a solution of (3). Hence we have

THEOREM 1. *If in a system of equations* (3) *the i-th equation is replaced by*

$$k_1f_1 + k_2f_2 + \cdots + k_mf_m = 0, \qquad k_i \neq 0,$$

the new system is equivalent to the given system.

All methods of solving a system of equations, even the method by determinants, employ the above principle.

3. Elementary operations. There are three *elementary operations* which can be performed upon the equations of a system to yield an equivalent system. We shall call these the elementary operations of Types I, II and III. They are:

Type I. The interchange of two equations of the system.

Type II. The multiplication of an equation of the system by a number $k \neq 0$.

Type III.* The addition to any equation of the system of k times any other equation of the system.

The proof that each of these elementary operations when applied to a system of equations yields an equivalent system is now immediate. That an operation of Type I leaves the common solutions unchanged is evident. Operations of Types II and III are special cases of the operation of Theorem 1. Furthermore, the operation of Theorem 1 can be achieved by one operation of Type II followed by $m-1$ operations of Type III. That is, we first replace $f_i=0$ by $k_i f_i=0$ where $k_i \neq 0$, then replace this by $k_1 f_1 + k_i f_i = 0$, and so on.

4. Systems of homogeneous equations. Let us now restrict attention to a system of homogeneous equations

$$
\begin{aligned}
f_1 &= a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n = 0, \\
f_2 &= a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n = 0, \\
&\cdots\cdots\cdots\cdots\cdots \\
f_m &= a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n = 0.
\end{aligned}
\tag{6}
$$

* These operations are not independent, for an operation of Type I can be obtained by a succession of operations of Type III with $k=1$ and operations of Type II.

A system of equations is called *triangular* if the last coefficient of f_{m-1} is 0, the last two coefficients of f_{m-2} are $0, \cdots,$ the last $m-1$ coefficients of f_1 are 0. If $m>n$, this means that all the coefficients of $f_1, f_2, \cdots, f_{m-n}$ are 0 or, as we shall say, that these polynomials vanish.

THEOREM 2. *The system* (6) *of homogeneous equations is equivalent to a triangular system.*

If some coefficient of x_n is not 0, we can by an interchange of equations if necessary insure that $a_{mn} \neq 0$. By adding to the first equation $-a_{1n}/a_{mn}$ times the last equation, we can make the new coefficient in the place of a_{1n} equal to 0. Similarly we can make every coefficient of x_n except a_{mn} equal to 0. If at the start every coefficient of x_n was 0, no reduction was required.

Now ignore the last equation. Unless every coefficient of x_{n-1} (above $a_{m,\, n-1}$) is 0, we can assume that $a_{m-1,\, n-1} \neq 0$ and as before make every other coefficient of x_{n-1} equal to 0. In this way we obtain a system of equations of triangular form equivalent to (6). If $m>n$, the first $m-n$ equations have vanished, each coefficient having become 0.

In every triangular system the number of non-vanishing equations is $m \leqq n$. By filling in with vanishing equations we may assume that $m=n$. In this form the coefficients $a_{11}, a_{22}, \cdots, a_{nn}$ are called the *diagonal coefficients*. If the system is triangular, every coefficient to the right of the diagonal coefficients is 0.

THEOREM 3. *The system* (6) *of homogeneous equations is equivalent to one of triangular form in which every diagonal coefficient is either* 0 *or* 1; *and if the diagonal coefficient in any equation is* 0, *the equation vanishes.*