Linear Algebra and Group Theory

Academician V. I. Smirnov's

LINEAR ALGEBRA AND GROUP THEORY

Revised, Adapted, and Edited by

RICHARD A. SILVERMAN

McGRAW-HILL BOOK COMPANY, INC.
1961 New York Toronto London

LINEAR ALGEBRA AND GROUP THEORY

 Library of Congress Catalog Card Number 61-11939

THE MAPLE PRESS COMPANY, YORK, PA.

57380

Preface

This book represents a selection of material from Prof. Vladimir I. Smirnov's encyclopedic six-volume "Course of Higher Mathematics." The present volume, entitled "Linear Algebra and Group Theory," has several unusual features that especially recommend it to the attention of the English-language readership. Unlike many algebra texts, this book does not delve into mathematical byways remote from the applications; instead, the approach adopted is consistently "concrete." Moreover, the book strives for the maximum coverage compatible with its length. Thus, in addition to a detailed treatment of linear algebra, it also contains an excellent introduction to group theory and an extensive discussion of group representations, a topic usually reserved for the specialized treatise. Under the heading of material rarely encountered in first courses on higher algebra, one should also mention Chap. 5 on infinite-dimensional spaces and Chap. 9 on continuous groups. It is apparent that the author's intention was to write an algebra text emphasizing those topics of greatest importance in applied mathematics and theoretical physics. Despite this, there is nothing in the volume that the pure mathematician can afford to ignore.

Because of the great difference between stylistic norms in English and Russian, as well as the absence of grammatical categories in one language that are present in the other, I have felt obliged to apply appropriate "smoothing operations" to ensure the continuity and readability of the translation. In doing so, I have not hesitated to add transitional sentences where I thought they were called for, make theorems out of some propositions not originally labeled as such, or clarify points that I found obscure. Nor have I hesitated to introduce an additional chapter heading not present in the original, renumber the equations in a way that appeared to me more convenient, redraw two of the figures to improve their perspective, and generally make the book conform to what I regard as the needs of its prospective audience. But the two most substantive changes were the following:

1. Chapter 6 of the present volume was originally the Appendix of Russian Vol. 3, Part II, while the Appendix of the present volume was originally Secs. 63 and 93 of Russian Vol. 3, Part II. In making these changes, it was necessary to supply the proof of the Hamilton-Cayley theorem (in the text) and explore the properties of the exponential of a matrix (in the problems).

2. There are no problems in the original, perhaps due to the Russian predilection for the use of special problem collections. It was thought that the addition of copious problems would greatly enhance the value of the English-language edition, as well as permit the incorporation of some important topics not discussed in the text. With this in mind, I have asked Prof. Allen L. Shields of the University of Michigan, Prof. John S. Lomont of the Polytechnic Institute of Brooklyn, and Prof. Jacob T. Schwartz of New York University to assist in preparing and selecting problems for this volume. Their contributions are identified more explicitly in the appropriate places. In addition, we have culled many problems from I. V. Proskuryakov's "A Collection of Problems on Linear Algebra," Moscow, 1957, and from D. K. Faddeyev and I. S. Sominski's "A Collection of Problems on Higher Algebra," Moscow, 1954. Occasional use was also made of problem collections by N. M. Gyunter and R. O. Kuzmin and by V. A. Krechmar. The net result has been to equip this book with over four hundred pertinent problems. Answers to about half of the computational problems and hints for the solution of the less obvious problems involving proofs are given at the end of the book. I have also listed several books for collateral or supplementary reading. There has been no attempt to make this list complete; in fact, it has been confined to books in English.

Finally, two observations should be made: (1) In a preface to the Russian original, Prof. Smirnov thanks D. K. Faddeyev for helping him with the group theory part of the book, in particular for writing Secs. 76, 87, 93, 94, 95 and 96; (2) To make the book self-contained, I have suppressed some references to other volumes of the six-volume course, and I have occasionally replaced others by the phrase "in an earlier volume" or "in a later volume." In every case, the reference was either an allusion to things to come or to elementary material to be found in any good text on advanced calculus, knowledge of which could be presupposed on the part of any reader of this book.

Richard A. Silverman
Translator and Editor

Contents

PART I

Determinants and Systems of Equations

Chapter 1

Determinants and Their Properties

1. The Concept of a Determinant. We begin this section with a simple algebraic problem, i.e., solving a system of linear equations. Investigation of this problem will lead us to the important concept of a determinant.

We begin by considering the simplest special cases. First we take a system of two equations in two unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1, \\ a_{21}x_1 + a_{22}x_2 &= b_2. \end{aligned}$$

Here, the coefficients a_{ik} of the unknowns are provided with two indices; the first index shows the equation in which the coefficient occurs, and the second index shows the unknown with which the coefficient is associated. As is well known, the solution of this system has the form

$$x_1 = \frac{b_1a_{22} - a_{12}b_2}{a_{11}a_{22} - a_{12}a_{21}}, \qquad x_2 = \frac{a_{11}b_2 - b_1a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Next we take three equations in three unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2, \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3, \end{aligned}$$

where we use the previous notation for the coefficients. We rewrite the first two equations in the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 &= b_1 - a_{13}x_3, \\ a_{21}x_1 + a_{22}x_2 &= b_2 - a_{23}x_3. \end{aligned}$$

Solving these equations with respect to the unknowns x_1 and x_2 by using the previous formulas, we have

$$x_1 = \frac{(b_1 - a_{13}x_3)a_{22} - a_{12}(b_2 - a_{23}x_3)}{a_{11}a_{22} - a_{12}a_{21}},$$

$$x_2 = \frac{a_{11}(b_2 - a_{23}x_3) - (b_1 - a_{13}x_3)a_{21}}{a_{11}a_{22} - a_{12}a_{21}}.$$

Substituting these expressions into the last equation of the system, we obtain an equation determining the unknown x_3; when this equation is finally solved, we find

$$x_3 = \frac{a_{11}a_{22}b_3 + a_{12}b_2a_{31} + b_1a_{21}a_{32} - a_{11}b_2a_{32} - a_{12}a_{21}b_3 - b_1a_{22}a_{31}}{a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}}. \tag{1}$$

We now examine in detail the construction of the formula (1). First of all, we note that its numerator can be obtained from its denominator by substituting the constant terms b_i for the coefficients a_{i3} of the unknown x_3. Thus, the problem is to explain the law of formation of the denominator, which does not contain any terms b_i, but rather is made up exclusively of the coefficients of the system. We write these coefficients in the form of a square array

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Vmatrix}, \tag{2}$$

preserving the order in which they appear in the system itself. This array has three rows and three columns, and the numbers a_{ik} are called its *elements*. The first index gives the row in which the element appears, and the second index gives the column. Writing out the denominator

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \tag{3}$$

of the ratio (1), we see that it consists of six terms. Each term is a product of three elements from the array (2), and in fact each term contains an element from every row and every column. The general form of these products is

$$a_{1p}a_{2q}a_{3r}, \tag{4}$$

where p, q, r are the integers 1, 2, 3 arranged in some definite order, i.e., the first and second indices are both integers (from 1 to 3) and each product (4) contains an element from every row and every column. To obtain all the terms of the expression (3), the second indices p, q, r in the product (4) have to be taken in all possible orders. Clearly, there are six possible *permutations*† of the second indices, namely,

$$1, 2, 3; \quad 2, 3, 1; \quad 3, 1, 2; \quad 1, 3, 2; \quad 2, 1, 3; \quad 3, 2, 1; \tag{5}$$

this gives us all six terms of (3). However, we see that some of the products (4) appear in (3) with a plus sign, while others appear with a minus sign. Thus, it remains only to explain the rule by which the sign is to be chosen. As we see, the products (4) whose second indices are the permutations

$$1, 2, 3; \quad 2, 3, 1; \quad 3, 1, 2 \tag{5a}$$

† See Sec. 2.

appear with a *plus* sign, while the products whose second indices are the permutations

$$1, 3, 2; \qquad 2, 1, 3; \qquad 3, 2, 1 \tag{5b}$$

appear with a *minus* sign.

We now explain how the permutations (5a) differ from the permutations (5b). We shall call the fact that a larger number comes before a smaller number in a permutation an *inversion*, and we shall calculate the number of inversions in the different permutations (5a). In the first of these permutations, there are no inversions at all, i.e., the number of inversions equals zero. Next, we consider the second permutation and compare the size of each number appearing in it with the numbers that follow. We see that there are two inversions here, since the numbers 2 and 3 come before the number 1. Similarly, it can easily be seen that the third of the permutations (5a) also contains two inversions. We can summarize this situation by saying that each of the permutations (5a) contains an *even* number of inversions. We are now in a position to formulate the following sign rule for the expression (3): The products (4) for which the number of inversions in the permutations formed by the second indices is an *even* number appear in (3) without any change. However, the products (4) for which the permutations formed by the second indices contain an *odd* number of inversions appear in (3) with a *minus* sign. The expression (3) is called the *determinant* (of order 3) corresponding to the array (2). We can now easily generalize these considerations and define a determinant of any order.

Suppose that we are given n^2 numbers arranged in the form of a square array

$$\left\| \begin{matrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{matrix} \right\| \tag{6}$$

with n rows and n columns. The elements a_{ik} of this array are certain complex numbers, and the indices i and k indicate the row and column in which the number a_{ik} appears. We now take the elements of the array (6) and form all possible products containing one (and only one) element from each row and column. These products have the form

$$a_{1p_1}a_{2p_2} \cdots a_{np_n}, \tag{7}$$

where $p_1, p_2, \ldots, p_n$ is some arrangement of the numbers $1, 2, \ldots, n$. To obtain all possible products of the form (7), we must take all possible permutations of the second indices. As is well known from elementary algebra (see also Sec. 2), the total number of such permutations equals n

factorial, i.e., $1 \cdot 2 \cdots n = n!$ Each permutation has a certain number of inversions as compared with the basic arrangement $1, 2, \ldots, n$. We now ascribe signs to the products (7) by the following rule: We put a plus sign before products whose second indices form a permutation with an even number of inversions, and a minus sign before products whose second indices form a permutation with an odd number of inversions. The sum of all these products, with the appropriate signs, is called the *determinant of order n* corresponding with the array (6). Clearly, this sum contains $n!$ terms.

It is not hard to give this definition in terms of a formula. Let $p_1, p_2, \ldots, p_n$ be a permutation of the numbers $1, 2, \ldots, n$, and denote the number of inversions in this permutation by the symbol

$$[p_1, p_2, \ldots, p_n].$$

Then, the definition just given can be written as the formula

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{1p_1} a_{2p_2} \cdots a_{np_n}, \tag{8}$$

where the summation extends over all possible permutations $p_1, p_2, \ldots, p_n$ of the second indices, and the determinant of an array is indicated by writing the array between vertical lines. (When talking about an array as such and not about its determinant, we put the array between *double* vertical lines.)

We note that in the expression (3), the factors in each product are arranged in such a way that the *first* indices occur in increasing order. Thus, so far, we have been concerned with the permutations formed by the second indices. However, the factors in each product can just as well be arranged in such a way that the *second* indices occur in increasing order. Then (3) becomes

$$a_{11}a_{22}a_{33} + a_{31}a_{12}a_{23} + a_{21}a_{32}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33} - a_{31}a_{22}a_{13}. \tag{9}$$

Here, the first indices form all possible permutations p, q, r of the integers 1, 2, 3, and it is easily verified that the sign rule for the terms of the expression (9) can be formulated just as before, but with respect to the first indices instead. This leads us to consider not only the sum (8) but also the similar sum

$$\sum_{(p_1, p_2, \ldots, p_n)} (-1)^{[p_1, p_2, \ldots, p_n]} a_{p_1 1} a_{p_2 2} \cdots a_{p_n n}. \tag{10}$$

It is clear that the sum (10) consists of the same products as the sum (8).

Moreover, we shall see below that these products have the same signs as in the sum (8), i.e., the two sums (8) and (10) are the same. (We have just seen that this is the case for $n = 3$.)

Finally, we return to the case $n = 2$. Here, the array has the form

$$\begin{Vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{Vmatrix},$$

and (8) gives the expression

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \tag{11}$$

for the second order determinant corresponding to this array.

The above considerations show that if we wish to understand the properties of determinants, we must become more familiar with the properties of permutations. Thus, we now turn to this subject.

2. Permutations. A set of n elements arranged in a definite order is called a *permutation*. There are $n!$ different permutations that can be formed out of n elements. For $n = 2$, this is obvious, since there are only two possible arrangements of two elements. For $n = 3$, the assertion follows immediately from the enumeration (5); here, the elements are the numbers 1, 2, 3, and it is easily verified that (5) gives all possible permutations of these three elements. For arbitrary n, we shall prove our assertion by mathematical induction, i.e., assuming that the assertion is valid for n elements, we shall show that it is also valid for $n + 1$ elements. Thus, we assume that n elements give $n!$ permutations, and we consider any $n + 1$ elements, which we denote by

$$C_1, C_2, \ldots, C_{n+1}.$$

Consider first the permutations whose first element is C_1. To obtain all possible permutations of this type, we must put C_1 in the first position and then write down all possible permutations of the n remaining elements; by assumption, the number of such permutations equals $n!$ In just the same way, the number of permutations whose first element is C_2 also equals $n!$ Thus, the total number of different permutations of the elements $C_1, C_2, \ldots, C_{n+1}$ equals

$$n!(n + 1) = 1 \cdot 2 \cdots n \cdot (n + 1) = (n + 1)!,$$

as was to be proved. (Of course, we can assume that our elements are the positive integers, and we shall henceforth adhere to this convention.)

The operation consisting of interchanging the positions of two elements in a permutation is called a transposition. It is immediately clear that we can obtain any permutation from any other permutation by performing transpositions. For example, take the two permutations

$$1, 3, 4, 2; \qquad 2, 4, 1, 3$$

of four elements. We can go from the first of these permutations to the second by performing the following transpositions

$$1, 3, 4, 2 \rightarrow 2, 3, 4, 1 \rightarrow 2, 4, 3, 1 \rightarrow 2, 4, 1, 3.$$

Here we needed three transpositions to go from the first permutation to the second. If we had used other transpositions, we could have gone from the first permutation to the second differently. In other words, the number of transpositions needed to go from one permutation to another is not uniquely determined, i.e., we can go from one permutation to another by using different numbers of transpositions. Thus, it is important to show that for any two given permutations, the different numbers of transpositions needed to go from one permutation to the other are either all even or all odd; this can be expressed differently by saying that these numbers always have the same *parity*. To see this, we introduce the concept of an inversion, which was already used in the preceding section. Suppose we are given a permutation of the n numbers $1, 2, \ldots, n$. The permutation

$$1, 2, \ldots, n, \tag{12}$$

where the numbers appear in increasing order, will be called the *basic permutation*. By an *inversion*, we mean the fact that two elements of a permutation do not appear in the order in which they appear in the basic permutation (12), or in other words, that a larger number comes before a smaller number. *Permutations in which the number of inversions is an even number will be called permutations of the first class, and those in which the number of inversions is odd will be called permutations of the second class.* The following theorem is basic for our further work:

A transposition changes the number of inversions by an odd number.

PROOF. Take any permutation

$$a, b, \ldots, k, \ldots, p, \ldots, s, \tag{13}$$

and assume that we carry out a transposition of the elements k and p, i.e., that we interchange the positions of these two elements. After such a transposition, the position of the elements k and p with respect to the elements standing to the left of k and to the right of p remains unchanged. The only thing that changes is the position of the elements k and p with respect to the elements of the permutation between k and p and, of course, the position of the elements k and p with respect to each other. We now calculate the total change in the number of inversions. Suppose