AN INTRODUCTION TO THE THEORY OF

LINEAR SPACES

GEORGI E. SHILOV

Professor of Mathematics
Moscow State University

Translated from the Russian by
Richard A. Silverman

PRENTICE-HALL, INC.
Englewood Cliffs, New Jersey

Current printing (last digit):

15 14 13 12 11 10 9 8

Library of Congress Catalog Number 61–13845

Printed in the United States of America
49896C

AUTHOR'S PREFACE

This book came about as a result of reworking material presented by the author in lecture courses and seminars given in recent years at Moscow and Kiev Universities. Stated briefly, the basic idea of the book is the following: It is more useful to treat algebra, geometry, and analysis as parts of a connected whole than as separate subjects. Thus, geometric notions help to clarify and often anticipate facts from algebra or analysis, just as algebraic methods often suggest the proper approach to be taken in a geometric or analytic context. Of course, this is not a new idea, and in fact, we can trace its influence on many generations of mathematicians from Descartes to Hilbert.

In the present book, the idea of the unity of algebra, geometry, and analysis is pursued in connection with elementary topics, accessible to students of mathematics and physics, even on the undergraduate level. Moreover, the book contains considerable material which these students are required to study in one form or another, and, in my opinion, the approach adopted here is the most accessible, pleasant, and useful way to master this material. In the U.S.S.R., this book is also used by students and candidates in technical institutes, by staff-workers in computation centers (despite the book's lack of explicit computational procedures) and by engineers in various professional-advancement courses. I also feel that the book is suitable for self-study.

To some extent, the problems given here are intended to help the student acquire technique, but most of them serve to illustrate and develop further the basic subject matter of the text. Some of the problems stem from elementary seminars, others from more advanced seminars, in which a certain degree of enthusiasm on the part of the participants is taken for granted.

I am happy to have the opportunity to express my appreciation to M. A. Krasnosyelski and M. G. Krein for numerous valuable suggestions, and to N. V. Efimov and D. A. Raikov for their careful reading of the manuscript before its publication; in many instances, their critical remarks allowed me to improve the presentation.

I am especially grateful to the Prentice-Hall Publishing Company, and in particular to Dr. Richard A. Silverman, for undertaking to make this book available in an English-language edition.

G. E. S.

TRANSLATOR'S PREFACE

The present volume is the first in a new series of translations of outstanding Russian textbooks and monographs in the fields of mathematics, physics, and engineering. It is my privilege to serve as Editor of the series. It is hoped that this book by Professor G. E. Shilov will set the standard for the volumes to follow.

The translation is a faithful one, to the extent that this is compatible with the syntactic and stylistic differences between Russian and English. However, I have occasionally made slight changes, and I have attempted to detect and correct all typographical errors. I have also added a Bibliography, containing suggestions for collateral and supplementary reading. Finally, it should be noted that sections marked with asterisks contain material of a more advanced nature, which can be omitted without loss of continuity.

R. A. S.

CONTENTS

1

DETERMINANTS

1. Systems of Linear Equations

In this and the next two chapters, we shall study systems of linear equations. In the most general case, such a system has the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \cdots \quad \cdots \quad \cdots \quad \cdots \quad &\quad . \\ a_{k1}x_1 + a_{k2}x_2 + \cdots + a_{kn}x_n &= b_k. \end{aligned} \tag{1}$$

Here $x_1, x_2, \ldots, x_n$ denote the unknowns which are to be determined. (Note that we do not necessarily assume that the number of unknowns equals the number of equations.) The quantities $a_{11}, a_{12}, \ldots, a_{kn}$ are called the *coefficients* of the system. The first index of a coefficient indicates the number of the equation in which the coefficient appears, while the second index indicates the number of the unknown with which the coefficient is associated.[1] The quantities $b_1, b_2, \ldots, b_k$ appearing in the right-hand side of (1) are called the *constant terms* of the system; like the coefficients, they are assumed to be known. By a *solution* of the system (1) we mean any set of numbers $c_1, c_2, \ldots, c_n$ which when substituted for the unknowns $x_1, x_2, \ldots, x_n$, turns all the equations of the system into identities.[2]

---

[1] Thus, for example, the symbol a_{34} should be read as "*a* three four" and not as "*a* thirty-four."

[2] We emphasize that the set of numbers $c_1, c_2, \ldots, c_n$ represents *one* solution of the system and not *n* solutions.

Not every system of linear equations of the form (1) has a solution. For example, the system

$$\begin{aligned} 2x_1 + 3x_2 &= 5, \\ 2x_1 + 3x_2 &= 6 \end{aligned} \tag{2}$$

obviously has no solution at all. Indeed, whatever numbers c_1, c_2 we substitute in place of the unknowns x_1, x_2, the left-hand sides of the equations of the system (2) are the same, while the right-hand sides are different. Therefore, no such substitution can simultaneously convert both equations of the system into identities.

A system of equations of the form (1) which has at least one solution is called *compatible*; a system which does not have solutions is called *incompatible*. A compatible system can have one solution or several solutions. In the latter case, we distinguish the solutions by indicating the number of the solution by a superscript in parentheses; for example, the first solution will be denoted by $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$, the second solution by $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$, and so on. The solutions $c_1^{(1)}, c_2^{(1)}, \ldots, c_n^{(1)}$ and $c_1^{(2)}, c_2^{(2)}, \ldots, c_n^{(2)}$ are considered to be *distinct* if at least one of the numbers $c_i^{(1)}$ does not coincide with the corresponding numbers $c_i^{(2)}$ $(i = 1, 2, \ldots, n)$. For example, the system

$$\begin{aligned} 2x_1 + 3x_2 &= 0, \\ 4x_1 + 6x_2 &= 0 \end{aligned} \tag{3}$$

has the distinct solutions

$$c_1^{(1)} = c_2^{(1)} = 0 \quad \text{and} \quad c_1^{(2)} = 3,\ c_2^{(2)} = -2$$

(and also infinitely many other solutions). If a compatible system has a unique solution, the system is called *determinate*; if a compatible system has at least two different solutions, it is called *indeterminate*.

We can now formulate the basic problems which arise in studying the system (1):

1. *To ascertain whether the system* (1) *is compatible or incompatible;*
2. *If the system* (1) *is compatible, to ascertain whether it is determinate;*
3. *If the system* (1) *is compatible and determinate, to find its unique solution;*
4. *If the system* (1) *is compatible and indeterminate, to describe the set of all its solutions.*

The basic mathematical tool for studying linear systems is *the theory of determinants*, which we consider next.

2. Determinants of Order n

2.1 Suppose that we are given a *square matrix*, i.e., an array of n^2 numbers a_{ij} $(i, j = 1, 2, \ldots, n)$:

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \cdot & \cdot & \ldots & \cdot \\ \cdot & \cdot & \ldots & \cdot \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{array} \right\|. \tag{4}$$

The number of rows and columns of the matrix (4) is called its *order*. The numbers a_{ij} are called the *elements* of the matrix. The first index indicates the row and the second index the column in which a_{ij} appears.

Consider any product of n elements which appear in different rows and different columns of the matrix (4), i.e., a product containing *just one element from each row and each column*. Such a product can be written in the form

$$a_{\alpha_1 1} a_{\alpha_2 2} \ldots a_{\alpha_n n}. \tag{5}$$

Actually, for the first factor we can always choose the element appearing in the first column of the matrix (4); then, if we denote by α_1 the number of the row in which the element appears, the indices of the element will be $\alpha_1, 1$. Similarly, for the second factor we can choose the element appearing in the second column; then its indices will be $\alpha_2, 2$, where α_2 is the number of the row in which the element appears, and so on. Thus, the indices $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the numbers of the rows in which the factors of the product (5) appear, when we agree to write the column indices in increasing order. Since, by hypothesis, the elements $a_{\alpha_1 1}, a_{\alpha_2 2}, \ldots, a_{\alpha_n n}$ appear in *different* rows of the matrix (4), one from each row, then the numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ are all different and represent some permutation of the numbers $1, 2, \ldots, n$.

By an *inversion* in the sequence $\alpha_1, \alpha_2, \ldots, \alpha_n$, we mean an arrangement of two indices such that the larger index comes before the smaller index. The total number of inversions will be denoted by $N(\alpha_1, \alpha_2, \ldots, \alpha_n)$. For example, in the permutation 2, 1, 4, 3, there are two inversions (2 before 1, 4 before 3), so that

$$N(2, 1, 4, 3) = 2.$$

In the permutation 4, 3, 1, 2, there are five inversions (4 before 3, 4 before 1, 4 before 2, 3 before 1, 3 before 2), so that

$$N(4, 3, 1, 2) = 5.$$

If the number of inversions in the sequence $\alpha_1, \alpha_2, \ldots, \alpha_n$ is even, we put a plus sign before the product (5); if the number is odd, we put a minus sign before the product. In other words, we agree to write in front of each product of the form (5) the sign determined by the expression

$$(-1)^{N(\alpha_1, \alpha_2, \ldots, \alpha_n)}.$$

The total number of products of the form (5) which can be formed from the elements of a given matrix of order n is equal to the total number of permutations of the numbers $1, 2, \ldots, n$. As is well known, this number is equal to $n!$.

We now introduce the following definition:

By the determinant D of the matrix (4) *is meant the algebraic sum of the $n!$ products of the form* (5), *each of which is preceded by the sign determined by the rule just given, i.e.,*

$$D = \sum (-1)^{N(\alpha_1, \alpha_2, \ldots, \alpha_n)} a_{\alpha_1 1} a_{\alpha_2 2} \ldots a_{\alpha_n n}. \tag{6}$$

Henceforth, the products of the form (5) will be called the *terms* of the determinant. The elements a_{ij} of the matrix (4) will be called the *elements* of the determinant. We denote the determinant corresponding to the matrix (4) by one of the following symbols:

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \det \|a_{ij}\|. \tag{7}$$

For example, we obtain the following expressions for the determinants of orders two and three:

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

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

We now indicate the role of determinants in solving systems of linear equations, by considering the example of a system of two equations in two unknowns:

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

Eliminating one of the unknowns in the usual way, we can easily obtain the formulas

$$x_1 = \frac{b_1a_{22} - b_2a_{12}}{a_{11}a_{22} - a_{21}a_{12}}, \qquad x_2 = \frac{a_{11}b_2 - a_{21}b_1}{a_{11}a_{22} - a_{21}a_{12}},$$

assuming that these ratios have nonvanishing denominators. The numerators and denominators of the ratios can be represented by the second-order determinants

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

$$b_1a_{22} - b_2a_{12} = \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix},$$

$$a_{11}b_2 - a_{21}b_1 = \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}.$$

It turns out that similar formulas hold for the solutions of systems with an arbitrary number of unknowns (Sec. 7).

2.2. The rule for determining the sign of a given term of a determinant can be formulated somewhat differently, in geometric terms. Corresponding to the enumeration of elements in the matrix (4), we can distinguish two natural positive directions: from left to right along the rows, and from top to bottom along the columns. Moreover, the slanting lines joining any two elements of the matrix can be furnished with a direction: we shall say that the line segment joining the element a_{ij} with the element a_{kl} has *positive slope* if its right endpoint lies lower than its left endpoint, and that it has *negative slope* if its right endpoint lies higher than its left endpoint.[3] Now imagine that in the matrix (4) we draw all the segments with *negative* slope joining pairs of elements $a_{\alpha_1 1}, a_{\alpha_2 2}, \ldots, a_{\alpha_n n}$ of the product (5). Then we put a plus sign before the product (5) if the number of all such segments is even, and a minus sign if the number is odd.

For example, in the case of a fourth-order matrix, a plus sign must be put before the product $a_{21}a_{12}a_{43}a_{34}$, since there are two segments of negative slope joining the elements of this product:

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

However, a minus sign must be put before the product $a_{41}a_{32}a_{13}a_{24}$, since in the matrix there are five segments of negative slope joining these elements:

$$\begin{Vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{Vmatrix}.$$

[3] This definition of "slope" is not to be confused with the geometric notion with the same name. In fact, the sign convention adopted here is the opposite of that used in geometry (*Translator*).

In these examples, the number of segments of negative slope joining the elements of a given term equals the number of *inversions* in the order of the first indices of the elements appearing in the term. In the first example, the sequence 2, 1, 4, 3 of first indices has two inversions; in the second example, the sequence 4, 3, 1, 2 of first indices has five inversions.

We now show that *the second definition of the sign of a term in a determinant is equivalent to the first.* To show this, it suffices to prove that the number of inversions in the sequence of first indices of a given term (with the second indices in natural order) is always equal to the number of segments of negative slope joining the elements of the given term in the matrix. But this is almost obvious, since the presence of a segment of negative slope joining the elements $a_{\alpha_i i}$ and $a_{\alpha_j j}$ means that $\alpha_i > \alpha_j$ for $i < j$, i.e., there is an inversion in the order of the first indices.

Problem 1. With what sign do the terms

(a) $a_{23}a_{31}a_{42}a_{56}a_{14}a_{65}$,
(b) $a_{32}a_{43}a_{14}a_{51}a_{66}a_{25}$

appear in the determinant of order 6? *Ans.* (a) +, (b) +.

Problem 2. Write down all the terms appearing in the determinant of order 4 which have a minus sign and contain the factor a_{23}.

Ans. $a_{11}a_{32}a_{23}a_{44}$, $a_{41}a_{12}a_{23}a_{34}$, $a_{31}a_{42}a_{23}a_{14}$.

Problem 3. With what sign does the term $a_{1n}a_{2,n-1}\dots a_{n1}$ appear in the determinant of order n? *Ans.* $(-1)^{n(n-1)/2}$.

3. Properties of Determinants of Order n

3.1. The transposition operation. The determinant

$$\begin{vmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ \cdot & \cdot & \cdots & \cdot \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{vmatrix} \tag{8}$$

obtained from the determinant (7) by interchanging rows and columns with the same indices is said to be the *transpose* of the determinant (7). We now show that *the transpose of a determinant has the same value as the original determinant.* In fact, the determinants (7) and (8) obviously consist of the same terms; therefore, it is enough for us to show that identical terms in the determinants (7) and (8) have identical signs. Transposition of the matrix of a determinant is clearly the result of rotating it (in space) through 180° about the diagonal $a_{11}a_{22}\dots a_{nn}$. As a result of this rotation, every segment with negative slope (e.g., making an angle $\alpha < 90°$ with the rows of the matrix) again becomes a segment with negative slope (i.e., making the angle

$90^\circ - \alpha$ with the rows of the matrix). Therefore, the number of segments with negative slope joining the elements of a given term does not change after transposition. Consequently, the sign of the term does not change either. Thus, the signs of all the terms are preserved, which means that the value of the determinant remains unchanged.

The property just proved establishes the equivalence of the rows and columns of a determinant. Therefore, further properties of determinants will be stated and proved only for columns.

3.2. The antisymmetry property. By the property of being *antisymmetric with respect to columns*, we mean the fact that a determinant changes sign when two of its columns are interchanged. We consider first the case where two adjacent columns are interchanged, for example columns j and $j + 1$. The determinant which is obtained after these columns are interchanged obviously still consists of the same terms as the original determinant. Consider any of the terms of the original determinant. Such a term contains an element of the jth column and an element of the $(j + 1)$th column. If the segment joining these two elements originally had negative slope, then after the interchange of columns, its slope becomes positive, and conversely. As for the other segments joining pairs of elements of the term in question, each of these segments does not change the character of its slope after the column interchange. Consequently, the number of segments with negative slope joining the elements of the given term changes by one when the two columns are interchanged; therefore, each term of the determinant, and hence the determinant itself, changes sign when the columns are interchanged.

Suppose now that two nonadjacent columns are interchanged, e.g., column j and column k $(j < k)$, where there are m other columns between them. This interchange can be accomplished by successive interchanges of adjacent columns as follows: First column j is interchanged with column $j + 1$, then with columns $j + 2, j + 3, \ldots, k$. Then the column $k - 1$ so obtained (which was formerly column k) is interchanged with columns $k - 2, k - 3, \ldots, j$. In all, $m + 1 + m = 2m + 1$ interchanges of adjacent columns are required, each of which, according to what has been proved, changes the sign of the determinant. Therefore, at the end of the process, the determinant will have a sign opposite to its original sign (since for any integer m, the number $2m + 1$ is odd).

COROLLARY. *A determinant with two identical columns vanishes.*

Proof. Interchanging the columns, we do not change the determinant; on the other hand, by what has been proved, the determinant must change its sign. Thus $D = -D$, which implies that $D = 0$.

Problem. Show that of the $n!$ terms of a determinant, exactly half $(n!/2)$

have a plus sign according to the definition of Sec. 2, while the other half have a minus sign.

Hint. Consider the determinant all of whose elements equal 1.

3.3. The linear property of determinants. This property can be formulated as follows:

If all the elements of the j'th column of a determinant D are "linear combinations" of two columns of numbers, i.e.,

$$a_{ij} = \lambda b_i + \mu c_i \qquad (i = 1, 2, \ldots, n),$$

where λ and μ are fixed numbers, then the determinant D is equal to a linear combination of two determinants, i.e.,

$$D = \lambda D_1 + \mu D_2. \tag{9}$$

Here both determinants have the same columns as the determinant D except for the j'th column; the j'th column of D_1 consists of the numbers b_i, while the j'th column of D_2 consists of the numbers c_i.

Proof. Every term of the determinant D can be represented in the form

$$\begin{aligned} a_{\alpha_1 1}a_{\alpha_2 2} \ldots a_{\alpha_j j} \ldots a_{\alpha_n n} &= a_{\alpha_1 1}a_{\alpha_2 2} \ldots (\lambda b_{\alpha_j} + \mu c_{\alpha_j}) \ldots a_{\alpha_n n} \\ &= \lambda a_{\alpha_1 1}a_{\alpha_2 2} \ldots b_{\alpha_j} \ldots a_{\alpha_n n} + \mu a_{\alpha_1 1}a_{\alpha_2 2} \ldots c_{\alpha_j} \ldots a_{\alpha_n n}. \end{aligned}$$

Adding up all the first terms (with the signs which the corresponding terms have in the original determinant), we clearly obtain the determinant D_1, multiplied by the number λ. Similarly, adding up all the second terms, we obtain the determinant D_2, multiplied by the number μ. This proves formula (9).

It is convenient to write this formula in a somewhat different form. Let D be an arbitrary fixed determinant. Denote by $D_j(p_i)$ the determinant which is obtained by replacing the elements of the jth column of D by the numbers p_i $(i = 1, 2, \ldots, n)$. Then (9) takes the form

$$D_j(\lambda b_i + \mu c_i) = \lambda D_j(b_i) + \mu D_j(c_i).$$

This linear property of determinants can easily be extended to the case where every element of the jth column is a linear combination not of two terms but of any other number of terms, i.e.

$$a_{ij} = \lambda b_i + \mu c_i + \cdots + \tau f_i.$$

In this case

$$\begin{aligned} D_j(a_{ij}) &= D_j(\lambda b_i + \mu c_i + \cdots + \tau f_i) \\ &= \lambda D_j(b_i) + \mu D_j(c_i) + \cdots + \tau D_j(f_i). \end{aligned} \tag{10}$$

COROLLARY 1. *We can factor out of a determinant any common factor of a column of the determinant.*

Proof. If $a_{ij} = \lambda b_i$, then by (10) we have

$$D_j(a_{ij}) = D_j(\lambda b_i) = \lambda D_j(b_i),$$

as required.

COROLLARY 2. *If a column of a determinant consists entirely of zeros, then the determinant vanishes.*

Proof. Since 0 is a common factor of the elements of one of the columns, we can factor it out of the determinant, obtaining

$$D_j(0) = D_j(0 \cdot 1) = 0 \cdot D_j(1) = 0.$$

Problem. By making a termwise expansion, calculate the determinant

$$\Delta = \begin{vmatrix} am + bp & an + bq \\ cm + dp & cn + dq \end{vmatrix}.$$

Ans. $\Delta = (mq - np)(ad - bc)$.

3.4. Addition of an arbitrary multiple of one column to another column: *We do not change the value of a determinant by adding the elements of one column multiplied by an arbitrary number to the corresponding elements of another column.*

Suppose that we add the kth column multiplied by the number λ to the jth column ($k \neq j$). The jth column of the resulting determinant consists of elements of the form $a_{ij} + \lambda a_{ik}$ ($i = 1, 2, \ldots, n$). By (10) we have

$$D_j(a_{ij} + \lambda a_{ik}) = D_j(a_{ij}) + \lambda D_j(a_{ik}).$$

The jth column of the second determinant consists of the elements a_{ik}, and hence is identical with the kth column. It follows from the corollary on p. 7 that $D_j(a_{ik}) = 0$, so that

$$D_j(a_{ij} + \lambda a_{ik}) = D_j(a_{ij}),$$

as required.

This property can also be formulated more generally:

Suppose we add to the elements of the j'th column of a determinant first the corresponding elements of the k'th column multiplied by λ, next the elements of the l'th column multiplied by μ, etc., and finally the elements of the p'th column multiplied by τ ($k \neq j, l \neq j, p \neq j$). Then the value of the determinant remains unchanged.

Problem. The numbers 20604, 53227, 25755, 20927 and 78421 are divisible by 17. Show that the determinant

$$\begin{vmatrix} 2 & 0 & 6 & 0 & 4 \\ 5 & 3 & 2 & 2 & 7 \\ 2 & 5 & 7 & 5 & 5 \\ 2 & 0 & 9 & 2 & 7 \\ 7 & 8 & 4 & 2 & 1 \end{vmatrix}$$

is also divisible by 17.