STUDENT'S TEXT

UNIT NO. 23

INTRODUCTION TO MATRIX ALGEBRA

SCHOOL MATHEMATICS STUDY GROUP

YALE UNIVERSITY PRESS

School Mathematics Study Group

Introduction to Matrix Algebra

Unit 23

Introduction to Matrix Algebra

Student's Text

Prepared under the supervision of
the Panel on Sample Textbooks
of the School Mathematics Study Group:

Frank B. Allen	Lyons Township High School
Edwin C. Douglas	Taft School
Donald E. Richmond	Williams College
Charles E. Rickart	Yale University
Henry Swain	New Trier Township High School
Robert J. Walker	Cornell University

New Haven and London, Yale University Press

Financial support for the School Mathematics
Study Group has been provided by the National
Science Foundation.

FOREWORD

The increasing contribution of mathematics to the culture of the modern world, as well as its importance as a vital part of scientific and humanistic education, has made it essential that the mathematics in our schools be both well selected and well taught.

With this in mind, the various mathematical organizations in the United States cooperated in the formation of the School Mathematics Study Group (SMSG). SMSG includes college and university mathematicians, teachers of mathematics at all levels, experts in education, and representatives of science and technology. The general objective of SMSG is the improvement of the teaching of mathematics in the schools of this country. The National Science Foundation has provided substantial funds for the support of this endeavor.

One of the prerequisites for the improvement of the teaching of mathematics in our schools is an improved curriculum—one which takes account of the increasing use of mathematics in science and technology and in other areas of knowledge and at the same time one which reflects recent advances in mathematics itself. One of the first projects undertaken by SMSG was to enlist a group of outstanding mathematicians and mathematics teachers to prepare a series of textbooks which would illustrate such an improved curriculum.

The professional mathematicians in SMSG believe that the mathematics presented in this text is valuable for all well-educated citizens in our society to know and that it is important for the precollege student to learn in preparation for advanced work in the field. At the same time, teachers in SMSG believe that it is presented in such a form that it can be readily grasped by students.

In most instances the material will have a familiar note, but the presentation and the point of view will be different. Some material will be entirely new to the traditional curriculum. This is as it should be, for mathematics is a living and an ever-growing subject, and not a dead and frozen product of antiquity. This healthy fusion of the old and the new should lead students to a better understanding of the basic concepts and structure of mathematics and provide a firmer foundation for understanding and use of mathematics in a scientific society.

It is not intended that this book be regarded as the only definitive way of presenting good mathematics to students at this level. Instead, it should be thought of as a sample of the kind of improved curriculum that we need and as a source of suggestions for the authors of commercial textbooks. It is sincerely hoped that these texts will lead the way toward inspiring a more meaningful teaching of Mathematics, the Queen and Servant of the Sciences.

Below are listed the names of all those who participated in any of the writing sessions at which the following SMSG texts were prepared: First Course in Algebra, Geometry, Intermediate Mathematics, Elementary Functions, and Introduction to Matrix Algebra.

H.W. Alexander, Earlham College
F.B. Allen, Lyons Township High School, La Grange, Illinois
Alexander Beck, Olney High School, Philadelphia, Pennsylvania
E.F. Beckenbach, University of California at Los Angeles
E.G. Begle, School Mathematics Study Group, Yale University
Paul Berg, Stanford University
Emil Berger, Monroe High School, St. Paul, Minnesota
Arthur Bernhart, University of Oklahoma
R.H. Bing, University of Wisconsin
A.L. Blakers, University of Western Australia
A.A. Blank, New York University
Shirley Boselly, Franklin High School, Seattle, Washington
K.E. Brown, Department of Health, Education, and Welfare, Washington, D.C.
J.M. Calloway, Carleton College
Hope Chipman, University High School, Ann Arbor, Michigan
R.R. Christian, University of British Columbia
R.J. Clark, St. Paul's School, Concord, New Hampshire
P.H. Daus, University of California at Los Angeles
R.B. Davis, Syracuse University
Charles DePrima, California Institute of Technology
Mary Dolciani, Hunter College
Edwin C. Douglas, The Taft School, Watertown, Connecticut
Floyd Downs, East High School, Denver, Colorado
E.A. Dudley, North Haven High School, North Haven, Connecticut
Lincoln Durst, The Rice Institute
Florence Elder, West Hempstead High School, West Hempstead, New York
W.E. Ferguson, Newton High School, Newtonville, Massachusetts
N.J. Fine, University of Pennsylvania
Joyce D. Fontaine, North Haven High School, North Haven, Connecticut
F.L. Friedman, Massachusetts Institute of Technology
Esther Gassett, Claremore High School, Claremore, Oklahoma
R.K. Getoor, University of Washington
V.H. Haag, Franklin and Marshall College
R.R. Hartman, Edina-Morningside Senior High School, Edina, Minnesota
M.H. Heins, University of Illinois
Edwin Hewitt, University of Washington
Martha Hildebrandt, Proviso Township High School, Maywood, Illinois
R.C. Jurgensen, Culver Military Academy, Culver, Indiana
Joseph Lehner, Michigan State University
Marguerite Lehr, Bryn Mawr College
Kenneth Leisenring, University of Michigan
Howard Levi, Columbia University
Eunice Lewis, Laboratory High School, University of Oklahoma
M.A. Linton, William Penn Charter School, Philadelphia, Pennsylvania
A.E. Livingston, University of Washington
L.H. Loomis, Harvard University
R.V. Lynch, Phillips Exeter Academy, Exeter, New Hampshire
W.K. McNabb, Hockaday School, Dallas, Texas
K.G. Michaels, North Haven High School, North Haven, Connecticut
E.E. Moise, University of Michigan
E.P. Northrop, University of Chicago
O.J. Peterson, Kansas State Teachers College, Emporia, Kansas
B.J. Pettis, University of North Carolina
R.S. Pieters, Phillips Academy, Andover, Massachusetts
H.O. Pollak, Bell Telephone Laboratories
Walter Prenowitz, Brooklyn College
G.B. Price, University of Kansas
A.L. Putnam, University of Chicago
Persis O. Redgrave, Norwich Free Academy, Norwich, Connecticut
Mina Rees, Hunter College
D.E. Richmond, Williams College
C.E. Rickart, Yale University
Harry Ruderman, Hunter College High School, New York City
J.T. Schwartz, New York University
O.E. Stanaitis, St. Olaf College
Robert Starkey, Cubberley High Schools, Palo Alto, California
Phillip Stucky, Roosevelt High School, Seattle, Washington
Henry Swain, New Trier Township High School, Winnetka, Illinois
Henry Syer, Kent School, Kent, Connecticut
G.B. Thomas, Massachusetts Institute of Technology
A.W. Tucker, Princeton University
H.E. Vaughan, University of Illinois
John Wagner, University of Texas
R.J. Walker, Cornell University
A.D. Wallace, Tulane University
E.L. Walters, William Penn Senior High School, York, Pennsylvania
Warren White, North High School, Sheboygan, Wisconsin
D.V. Widder, Harvard University
William Wooton, Pierce Junior College, Woodland Hills, California
J.H. Zant, Oklahoma State University

CONTENTS

Chapter

PREFACE

The present volume is an experimental edition for a high-school course in the theory of matrices and vectors. In selecting material for the text, the School Mathematics Study Group has been mindful of the fact that this is the last mathematics course in secondary school, the terminal course for many students. As citizens, they should have a sound idea of the nature of mathematics. This point of view has been emphasized in the Harvard report, "General Education in a Free Society," Harvard University Press, Cambridge, 1945, which states: "Mathematics may be defined as the science of abstract form. The discernment of structure is essential, no less to the appreciation of a painting or symphony than in the behaviour of a physical system; no less in economics than in astronomy. Mathematics studies order, abstracted from the particular objects and phenomena which exhibit it, and in a generalized form."

One of our basic aims is thus to demonstrate the structure of mathematics. We shall not be concerned, however, with structure merely as such. Rather, we shall exhibit some rich mathematics that is totally new to the student and demonstrate structure as we proceed. To make abstract form a topic unto itself often leads to a barren presentation; to discuss the structure of the already-familiar arithmetic and algebra seems forced and repetitive to the boy or girl who is dreaming of a place in a jet age, even in a space age.

It is important to give the student some "new" mathematics that has considerable vigor and vitality. Until very recently, the high-school curriculum has been almost entirely concerned with ideas that were developed during or before the sixteenth and seventeenth centuries. Computers and electronic brains are front-page news. In order to appeal to the imagination of the student and to expose some mathematics that is very much alive, the material must be new, different, and bold.

Another criterion is to provide some tools that will be eminently useful in the student's transition from school to college, tools that will help bridge the gap from the manipulative spirit of high-school mathematics to the abstract viewpoint of modern algebraic studies. Yet this material must not come from the usual sequential courses.

A unit on matrix algebra will satisfy the foregoing criteria. As one operation after another is defined, the structure of mathematics can be repeatedly emphasized. Terms like group, ring, field, and isomorphism will be introduced when meaningful and needed for unifying concepts. Thus they will be met in a new, appropriate, and substantial context; they will not be applied to shopworn material. Introduced by Cayley in 1858, recognized by Heisenberg in 1925 as exactly the tool he needed to develop his revolutionary work in quantum mechanics, employed today in such diverse ways as providing a language for atomic physics, measuring the air flow over the wing of an airplane, and keeping the parts inventory at a minimum in a factory, matrices can put the student close to the frontiers of mathematics and provide striking examples of patterns that arise in the most varied circumstances. Moreover, the student meets some mathematics emancipated from the familiar rules of arithmetic, and he learns that it is within his capacities to "invent" some of his own. If this study can make mathematics more alive, then here indeed is a promising path.

Our study of matrix algebra will involve the investigation of a significant postulational system, which will reflect the vigor of abstract mathematics. This is a unit in "hard" mathematics that has power and beauty. It will provide an effective language and some dynamic concepts that will enhance the student's ability to handle his first college courses yet not duplicate material.

Lastly, with the objective that the intellectually vigorous students may, in some small part, obtain an idea of what constitutes "mathematical research," there is appended a set of "Research Exercises." These are by no means overnight homework and any one of them may well constitute a project to be executed by several students. Such team operations are conducive to stimulating discourse and critical thinking.

Chapter 1

MATRIX OPERATIONS

1-1. Introduction

As we have studied more and more sophisticated mathematics, we have had occasion to use more and more sophisticated kinds of "numbers." We began with the set of counting numbers, 1, 2, 3,.... Then, in order to make subtractions like 3 - 7 possible, the system was extended to the entire set of integers, 0, ± 1, ± 2, ± 3,.... Next, in order to make it possible to divide any number by any nonzero number, rational numbers like 1/2, -2/3, -157/321, and 4/2 were invented. This did not bring us to the end of our story, for, in order that every positive number should have a square root, a cube root, a logarithm, etc., it was necessary to invent still more numbers: the infinite decimals or real numbers, such as 1.4142..., 3.1415928..., and 0.13131313.... Finally, in order that negative numbers should also have square roots, and that such quadratic equations as

$$x^2 + x + 1 = 0$$

should have solutions, it was necessary to invent complex numbers like $3 + 2i$, $1 + \pi i$, $-1/2 + (1/37)i$, and $3 + 0i$.

Whenever there has seemed to be a good reason to do so, we have invented new sets of "numbers." For instance, in inventing complex quantities, we began not with the quantities themselves but with a purpose: to find a system of numbers each of which has a square root. When we have made one such invention, it is not hard to realize that there is no reason to stop inventing. Why should we not hope to invent many kinds of new numbers?

It is easy to invent things that do not work, but hard to invent things that do work — easy to invent things that are useless, but hard to invent things that are useful. The same is true of the invention of new kinds of numbers. The hard thing is to invent useful kinds of numbers, and kinds of numbers "that work." Nevertheless, several more or less successful new kinds of numbers have been invented by mathematicians. In this book, we are going to study one of the most successful of these new kinds of numbers: the matrices.

Before we tell you what matrices are, it is well for us to emphasize their importance. They are useful in almost every branch of science and engineering.

A great number of the operations performed by the giant "electronic brains" are computations with matrices. Many problems in statistics are expressed in terms of matrices. Matrices come up in the mathematical problems of economics. They are extremely important in the study of atomic physics; indeed, atomic physicists express almost all their problems in terms of matrices, and it would not be an exaggeration to say that the algebra of matrices is the language of atomic physics. Many other kinds of algebra, such as complex-number algebra and vector algebra, which some of you may already have studied, can be explained very easily in terms of matrices. So, in studying matrices, you will be studying one of the newest and most important, as well as one of the most interesting, branches of mathematics.

Let us look at a few simple examples.

Many a baseball fan, when he first opens the newspaper, refers to a tabulation similar to the following:

	G	AB	R	H
Aaron	68	280	52	109
Williams	52	194	29	60
Mantle	60	228	51	70
Lopez	63	241	38	72

If he is a Mantle fan, he looks at the entry in the third row and fourth column of numbers in order to learn how many hits Mantle has thus far obtained during the season.

You will note that we have said "row" in speaking of a horizontal array, and "column" in speaking of a vertical array. Thus, the third row is

60 228 51 70,

and the fourth column is

109
60
70
72

An assembler of TV sets might have before him a table of the following sort:

	Model A	Model B	Model C
Number of tubes	13	18	20
Number of speakers	2	3	4

This table indicates the number of tubes and the number of speakers used in assembling a set of each model.

Omitting the row and column headings, let us focus our attention on the arrays of numbers in the last two examples:

$$\begin{matrix} 68 & 280 & 54 & 109 \\ 52 & 194 & 29 & 60 \\ 60 & 228 & 51 & 70 \\ 63 & 241 & 38 & 72 \end{matrix} \qquad \begin{matrix} 13 & 18 & 20 \\ 2 & 3 & 4 \end{matrix}$$

Such arrays of entries are called matrices (singular: matrix). Thus a matrix is a rectangular array of entries appearing in rows and columns. Actually, the entries may be complex numbers, functions, and in appropriate circumstances even matrices themselves; however, with a few exceptions that will be clearly indicated, we shall confine our attention to the real numbers with which we are already familiar.

Some examples of matrices are the following:

$$\begin{bmatrix} 2 & 3 & 4 \\ 1 & 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 0 & \sqrt{2} \\ 3.14 & 1 \end{bmatrix}, \quad \begin{bmatrix} 4 \\ 2 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 1/2 & 1/4 & 1/8 \end{bmatrix}. \tag{1}$$

You will note here how square brackets $[\quad]$ are used in the mathematical designation of matrices.

A great advantage of this notation is the fact that we can use it in handling large sets of numbers as single entities, thus simplifying the statement of complicated relationships.

1-2. The Order of a Matrix

The order of a matrix is given by stating first the number of rows and then the number of columns in the matrix. Thus the orders of the matrices in the foregoing examples (1) are respectively 2×3 (read "2 by 3"), 2×2, 4×1, and 1×3. Generally, a matrix that has m rows and n columns is called an $m \times n$ (read "m by n") matrix, or a matrix of order $m \times n$.

If the number of rows is the same as the number of columns, as in the second example above, then the matrix is square. Thus, given two linear equations in two unknowns,

$$2x + 3y = 7,$$
$$1x - 2y = 0,$$

we observe that the coefficients of x and y constitute a square matrix:

$$\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix} .$$

When speaking of a square $n \times n$ matrix, we often refer to its order as n rather than $n \times n$. For example, the 2×2 matrix

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}$$

is a square matrix of order 2, and the 3×3 matrix

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

is a square matrix of order 3.

If the number of rows is 1, as in the fourth example in (1), above, the matrix is sometimes called a row matrix or a row vector. For example, in terms of rectangular coordinates, a point in a plane might be designated by the row matrix $\begin{bmatrix} 2 & 3 \end{bmatrix}$, or a point in space by the row matrix $\begin{bmatrix} 2 & 3 & -1 \end{bmatrix}$.

Similarly, a column matrix or column vector is a matrix having just one column. Thus, the foregoing points can equally well be designated by column matrices,

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \text{or} \quad \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix} ;$$

and the number of men, women, and children in a family might be denoted by

$$\begin{bmatrix} 1 \\ 1 \\ 3 \end{bmatrix}.$$

Capital letters are often used to denote general matrices, and the corresponding small letters with appropriate subscripts are then employed to designate entries. Thus, we might have

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix}.$$

In these examples, the entries located at the intersection of the 2nd row and 3rd column are denoted by a_{23} and b_{23}, respectively.

Generally, the entry located at the intersection of the i-th row and j-th column of matrix A is denoted by a_{ij}. An $m \times n$ matrix can be denoted compactly as $\left[a_{ij}\right]_{m\times n}$. Thus the foregoing matrices A and B are

$$A = \left[a_{ij}\right]_{3\times 3} \quad \text{and} \quad B = \left[b_{ij}\right]_{2\times 3}.$$

If the order is clear from the context or is arbitrary, the notation might be reduced to

$$A = \left[a_{ij}\right] \qquad \text{and} \quad B = \left[b_{ij}\right].$$

Associated with each matrix is another matrix called its <u>transpose</u>, which is often convenient to use and has interesting theoretical properties. The transpose A^t of a matrix A is formed by interchanging its rows and columns. For example, if

$$A = \begin{bmatrix} 1 & 2 & 2 \\ 3 & -1 & 0 \end{bmatrix}, \text{ then } A^t = \begin{bmatrix} 1 & 3 \\ 2 & -1 \\ 2 & 0 \end{bmatrix}.$$

<u>Definition 1-1</u>. If $A = \left[a_{ij}\right]$ is an $m \times n$ matrix, then the <u>transpose</u> A^t of A is the $n \times m$ matrix $B = \left[b_{ij}\right]$ with $b_{ij} = a_{ji}$ for each

i, j ($i = 1,2,\ldots,n$; $j = 1,2,\ldots,m$).

Exercises 1-2

1. (a) Obtain from a newspaper or other similar source six examples of information presented in matrix form.

   (b) In each of your examples, state the order of the matrix.

   (c) In each of the examples, suggest an alternative method (not in matrix form) of presenting the same information.

2. A row vector with three entries can be used to tabulate a person's age, height, and weight.

   (a) Give a row vector that lists your age, height, and weight.

   (b) Suggest when it might be useful to employ such a vector.

3. Let

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 8 & 10 & 12 & 14 & 16 \\ -1 & -3 & -5 & 6 & 3 \\ 0 & 3 & -7 & 8 & 7 \end{bmatrix} .$$

   (a) What is the order of A?

   (b) Name the entries in the 4th row.

   (c) Name the entries in the 3rd column.

   (d) Name the entry a_{43}.

   (e) Name the entry a_{14}.

   (f) Name the entry a_{41}.

   (g) Write the transpose A^t.

4. Let

$$B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} .$$

   (a) What is the order of B?

(b) Name the entries in the 3rd row.

(c) Name the entries in the 3rd column.

(d) Name the entry b_{12}.

(e) For what values i, j is $b_{ij} \neq 0$?

(f) For what values i, j is $b_{ij} = 0$?

(g) Write the transpose B^t.

5\. (a) Write a 3×3 matrix all of whose entries are whole numbers.

(b) Write a 3×4 matrix none of whose entries are whole numbers.

(c) Write a 5×5 matrix having all entries in its first two rows positive, and all entries in its last three rows negative.

6\. (a) How many entries are there in a 2×2 matrix?

(b) In a 4×3 matrix?

(c) In an $n \times n$ matrix?

(d) In an $m \times n$ matrix?

1-3. Equality of Matrices

Two matrices are <u>equal</u> provided they are of the same order and each entry in the first is equal to the corresponding entry in the second. For example,

$$\begin{bmatrix} 1 & 4 & 0 \\ 2 & 8 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \times 2 & 2 - 2 \\ 4/2 & 16/2 & 8/2 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix} = \begin{bmatrix} 2^1 \\ 2^2 \\ 2^3 \end{bmatrix}, \quad \begin{bmatrix} x^2 - 1 \\ x \end{bmatrix} = \begin{bmatrix} (x-1)(x+1) \\ x \end{bmatrix};$$

but

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \neq \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \end{bmatrix} \neq \begin{bmatrix} 0 \end{bmatrix}.$$

<u>Definition 1-2</u>. Two matrices A and B are <u>equal</u>, $A = B$, if and only if they are of the same order and their corresponding entries are equal.

Thus,

$$\left[a_{ij}\right]_{m \times n} = \left[b_{ij}\right]_{m \times n}$$

if and only if $a_{ij} = b_{ij}$ for each i,j ($i = 1,2,\ldots,m$; $j = 1,2,\ldots,n$).

Using the foregoing definition of equality, we can express certain relationships more compactly. For example, the following equation between 2×1 matrices,

$$\begin{bmatrix} 2x + 3y \\ 3x - y \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \end{bmatrix},$$

can be employed instead of the two separate equations

$$\begin{aligned} 2x + 3y &= 7, \\ 3x - y &= 2; \end{aligned}$$

and

$$\begin{bmatrix} x + y & a + b \\ x - y & a - b \end{bmatrix} = \begin{bmatrix} 5 & -1 \\ 1 & 3 \end{bmatrix}$$

can be written in place of the four equations

$$\begin{aligned} x + y &= 5, & a + b &= -1, \\ x - y &= 1, & a - b &= 3. \end{aligned}$$

Exercises 1-3

1. Solve the following equations:

(a) $$\begin{bmatrix} x + 2 \\ 3 - y \end{bmatrix} = \begin{bmatrix} 1 \\ 4 \end{bmatrix},$$

(b) $$\begin{bmatrix} x - 2y \\ x + y \end{bmatrix} = \begin{bmatrix} -3 \\ 3 \end{bmatrix},$$

(c) $$\begin{bmatrix} x^2 & y \\ x & y^2 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

2. From the matrix equalities $A = B$ and $B = C$, would you conclude that $A = C$? Why?

3. Write the matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

if

$$a_{ij} = 2i + 3j - 4.$$

4. Write the matrix whose entries are the sums of the corresponding entries of the matrices

$$\begin{bmatrix} 1 & 0 \\ 2 & -1 \\ -3 & 4 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 2 \\ -3 & 4 \\ 2 & 1 \\ 0 & 0 \end{bmatrix}.$$

5. Write the matrix whose entries are the differences (first minus second) of the corresponding entries of the matrices in Exercise 4.

1-4. Addition of Matrices

We have now defined matrices and studied some of their most elementary properties. But we have not really made them work. To do this, we must give rules for adding and multiplying matrices, just as was done, for example, with complex numbers. If these numbers were defined bluntly as expressions of the form $a + bi$, without the operations of addition and multiplication, and without relation to the solution of such equations as

$$x^2 + x + 1 = 0,$$

they would be of relatively little interest. What gives life to complex numbers is the fact that we are able to define addition and multiplication for them in such a way that we have a whole algebra of complex numbers, which is indeed useful and interesting.

The same remark applies to matrices. To give the study of matrices its real content, we must define "sum" and "product" for matrices. In this section, we define and study sums of matrices. Products will be considered later.

You will recall that when two complex numbers are added, for example $3 + 5i$ and $-2 + 4i$, the two real components and the two imaginary components are added separately. Thus,

$$(3 + 5i) + (-2 + 4i) = (3 + (-2)) + (5 + 4)i = 1 + 9i.$$

If we represent the complex numbers as column vectors, we find their sum by adding corresponding entries; thus,

$$\begin{bmatrix} 3 \\ 5 \end{bmatrix} + \begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 9 \end{bmatrix} .$$

This suggests the pattern used in adding matrices of the same order. The sum of two such matrices is obtained by adding the individual entries in corresponding positions. For example,

$$\begin{bmatrix} 2 & 3 & 1 \\ -1 & 0 & 4 \end{bmatrix} + \begin{bmatrix} -4 & 2 & 1 \\ 1 & 3 & -2 \end{bmatrix} = \begin{bmatrix} -2 & 5 & 2 \\ 0 & 3 & 2 \end{bmatrix} .$$

Since we shall not even give a rule by which matrices of different orders could be added, we shall add two matrices only if they are of the same order. Accordingly, two matrices that have the same order are sometimes said to be conformable for addition. The sum has the same order as the two addends.

Definition 1-3. The sum $A + B$ of two $m \times n$ matrices A and B is the $m \times n$ matrix C such that the entry c_{ij} in the i-th row and j-th column of C is equal to the sum $a_{ij} + b_{ij}$ of the entries a_{ij} and b_{ij} in the i-th row and j-th column of A and B, respectively.

Thus,

$$\left[a_{ij}\right]_{m\times n} + \left[b_{ij}\right]_{m\times n} = \left[a_{ij} + b_{ij}\right]_{m\times n}.$$

For instance,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \\ a_{31} + b_{31} & a_{32} + b_{32} \end{bmatrix} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \end{bmatrix} .$$

If we consider all $m \times n$ matrices, with m and n fixed, as constituting a set $S_{m,n}$, and if A and B are elements of $S_{m,n}$, then $A + B$ is also an element of this set. That is, if $A \in S_{m,n}$ (read "A is an element of $S_{m,n}$") and $B \in S_{m,n}$, then $(A + B) \in S_{m,n}$.

In the algebra of real numbers R, the equation

$$a + 0 = a$$

is satisfied for all $a \in R$ (this time, read "for all $a \in R$" as "for all elements a of R"). Accordingly, we say that 0 is <u>the identity element for addition</u> in R. In the algebra of matrices, the matrices all of whose entries are 0 play a corresponding role. Thus,

$$\begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 + 0 & 3 + 0 \\ -1 + 0 & 4 + 0 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ -1 & 4 \end{bmatrix} .$$

Such a matrix is called a <u>zero matrix</u> and is denoted by $\underline{0}$. If the order $m \times n$ is significant we write $0_{m \times n}$; or, if the matrix is square, we might write 0_n, where n indicates the order of the matrix. Thus,

$$0_{1\times 2} = \begin{bmatrix} 0 & 0 \end{bmatrix} , \quad 0_{2\times 3} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} , \quad 0_3 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} .$$

The equation

$$A_{m\times n} + 0_{m\times n} = A_{m\times n}$$

clearly is valid for all $A_{m\times n}$.

The addition of matrices is a commutative operation, as we can readily verify. Thus,

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} .$$

In particular, the sum of the two matrices on the left is a matrix having $a_{12} + b_{12}$ as element in the first row and second column, and the corresponding element of the sum on the right is $b_{12} + a_{12}$. But

$$a_{12} + b_{12} = b_{12} + a_{12},$$

by the commutative law for the addition of real numbers.

The foregoing observation holds generally, of course, so that we have the following result:

Theorem 1-1. If the matrices A and B are conformable for addition, then they satisfy the commutative law for addition:

$$A + B = B + A.$$

Proof. We have

$$\begin{aligned} A + B &= \left[a_{ij}\right] + \left[b_{ij}\right] \\ &= \left[a_{ij} + b_{ij}\right] \\ &= \left[b_{ij} + a_{ij}\right] \\ &= \left[b_{ij}\right] + \left[a_{ij}\right] \\ &= B + A. \end{aligned}$$

Thus, in terms of our usual notation, the entry in the i-th row and j-th column of the sum on the left is $a_{ij} + b_{ij}$, and the corresponding entry of the sum on the right is $b_{ij} + a_{ij}$. But

$$a_{ij} + b_{ij} = b_{ij} + a_{ij},$$

by the commutative law for the addition of real numbers; hence the theorem follows from the definition (Definition 1-2) of the equality of two matrices.

The addition of conformable matrices is also associative; that is,

$$A + (B + C) = (A + B) + C.$$

For example,

$$\begin{bmatrix} 2 & 3 & 1 \\ -4 & 0 & 6 \end{bmatrix} + \left(\begin{bmatrix} -1 & 2 & 0 \\ -2 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 4 \\ 5 & 1 & 2 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 2 & 3 & 1 \\ -4 & 0 & 6 \end{bmatrix} + \begin{bmatrix} 0 & 2 & 4 \\ 3 & 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 5 \\ -1 & 1 & 9 \end{bmatrix},$$

and also

$$\left(\begin{bmatrix} 2 & 3 & 1 \\ -4 & 0 & 6 \end{bmatrix} + \begin{bmatrix} -1 & 2 & 0 \\ -2 & 0 & 1 \end{bmatrix}\right) + \begin{bmatrix} 1 & 0 & 4 \\ 5 & 1 & 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 5 & 1 \\ -6 & 0 & 7 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 4 \\ 5 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 5 & 5 \\ -1 & 1 & 9 \end{bmatrix}.$$

We can state the associative property as a theorem and prove it, as follows:

<u>Theorem 1-2</u>. If the matrices A, B, and C are conformable for addition, then they satisfy the associative law for addition:

$$A + (B + C) = (A + B) + C.$$

<u>Proof</u>. We note that, in terms of our usual notation, the entry in the i-th row and j-th column of the sum on the left is $a_{ij} + (b_{ij} + c_{ij})$, and the corresponding entry of the sum on the right is $(a_{ij} + b_{ij}) + c_{ij}$. But

$$a_{ij} + (b_{ij} + c_{ij}) = (a_{ij} + b_{ij}) + c_{ij}.$$

You can complete the proof of Theorem 1-2 by telling why this last equality is valid for all real numbers a_{ij}, b_{ij}, and c_{ij}, and why this equality implies the matrix equality

$$A + (B + C) = (A + B) + C.$$

Since it is immaterial in which order the matrices are added, we write $A + B + C$ for either expression:

$$A + (B + C) = (A + B) + C = A + B + C.$$

Once we know how to add numbers, it is usual to consider subtraction. You will recall that the <u>negative</u>, which we might call the <u>additive inverse</u>, of the real number a is denoted by $-a$. It satisfies the equation

$$a + (-a) = 0.$$

Subtraction of matrices arises in a similar manner.

<u>Definition 1-4</u>. Let A be an $m \times n$ matrix. Then the <u>negative</u> of A, written $-A$, is the $m \times n$ matrix each of whose entries is the negative of the corresponding entry of A.

<u>Definition 1-5</u>. If A and B are two $m \times n$ matrices, then the <u>difference</u> of A and B, designated by $A - B$, is the sum of A and the negative of B.

Thus, for $A + (-B)$, where A and B are matrices of equal orders, we write $A - B$ and say that the symbols indicate that B is to be subtracted from A. For example,

$$\begin{bmatrix} 2 & 1 & 3 \\ 4 & 0 & -1 \end{bmatrix} - \begin{bmatrix} 0 & 1 & -2 \\ 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 5 \\ 1 & -4 & -2 \end{bmatrix},$$

and

$$\begin{bmatrix} 1 & t-c \\ t+c & 4 \end{bmatrix} - \begin{bmatrix} 1 & t \\ c & 2 \end{bmatrix} = \begin{bmatrix} 0 & -c \\ t & 2 \end{bmatrix} .$$

Now we can easily prove the following theorem:

<u>Theorem 1-3</u>. If A and B are $m \times n$ matrices, and $\underline{0}$ the $m \times n$ zero matrix, then

(a) $A + (-A) = \underline{0}$,

(b) $-(-A) = A$,

(c) $-\underline{0} = \underline{0}$,

(d) $-(A + B) = (-A) + (-B)$.

Proof of Theorem 1-3 (a). The entry in the i-th row and j-th column of $-A$ is, by definition, $-a_{ij}$. Thus the entry in the i-th row and j-th column of $A + (-A)$ is $a_{ij} + (-a_{ij})$. But $a_{ij} + (-a_{ij}) = 0$. Hence every entry of $A + (-A)$ is zero; that is, $A + (-A)$ is the zero matrix.

The proofs of the remaining parts are similar and are left to the student as exercises.

Exercises 1-4

1. Find values x, y, a, and b that satisfy the matrix relationship

$$\begin{bmatrix} x + 3 & 2y - 8 \\ a + 1 & 4x + 6 \\ b - 3 & 3b \end{bmatrix} = \begin{bmatrix} 0 & -6 \\ -3 & 2x \\ 2b + 4 & -21 \end{bmatrix}.$$

2. If

$$A = \begin{bmatrix} 3 & 2 & 1 \\ 4 & -5 & 6 \\ 0 & 8 & -3 \\ 4 & 6 & 8 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -3 & 4 & 8 \\ -2 & 6 & -1 \\ 0 & 2 & 3 \\ 4 & -1 & 8 \end{bmatrix},$$

determine the entry in the sum $A + B$ that is at the intersection of

(a) the 3rd row and 2nd column,

(b) the 1st row and 3rd column,

(c) the 4th row and 1st column.

3. Compute

$$\begin{bmatrix} 1/2 & 1/3 \\ 1/4 & 1/5 \end{bmatrix} - \begin{bmatrix} 1/6 & 1/7 \\ 1/8 & 1/9 \end{bmatrix}.$$

4. Compute

$$\begin{bmatrix} 1/2 & 1/3 & 1/4 \\ 1/5 & 1/6 & 1/7 \\ 1/8 & 1/9 & 1/10 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

5. Compute

$$\begin{bmatrix} x & y & z \\ p & s & t \\ u & v & w \end{bmatrix} + \begin{bmatrix} 1-x & -y & -z \\ -p & 1-s & -t \\ -u & -v & 1-w \end{bmatrix} .$$

6. (a) Does the expression

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 3 & 1 & 2 \end{bmatrix} + 0_2$$

make sense?

(b) Does the expression

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 3 & 2 \\ 3 & 1 & 2 \end{bmatrix} + 0_3$$

make sense?

(c) What is the latter sum?

7. Compute

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ \sqrt{2} & 1 & 0 \end{bmatrix} + \begin{bmatrix} 3 & 2 & 1 \\ 4 & 17 & 8 \\ 9 & 6 & 14 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 3 \\ 14 & 8 & 6 \\ 1+\sqrt{2} & 11 & 11 \end{bmatrix} .$$

8. Compute

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} + \begin{bmatrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 10 & 10 & 10 \\ 10 & 10 & 10 \\ 10 & 10 & 10 \end{bmatrix} .$$

9. Given

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} , \quad B = \begin{bmatrix} 2 & -1 \\ 3 & -2 \\ 0 & 1 \end{bmatrix} , \quad \text{and} \quad C = \begin{bmatrix} 4 & 2 \\ 1 & 0 \\ -2 & -4 \end{bmatrix} ,$$

compute the following:

(a) $A + B$, (d) $A - B$,

(b) $(A + B) + C$, (e) $(A - B) + C$,

(c) $A + (B + C)$, (f) $B - A$.

10. (a) In Exercise 9, consider the answers to parts (b) and (c). What law is illustrated?

(b) In Exercise 9, consider the answers to parts (d) and (f). What conclusion can be drawn?

11. Prove Theorem 1-3 (b).

12. Prove Theorem 1-3 (c).

13. Prove Theorem 1-3 (d).

14. Assuming that A and B are conformable for addition, prove that $A^t + B^t = (A + B)^t$.

1-5. Addition of Matrices (Concluded)

The theorems given in Section 1-4 include exact analogues of all the basic laws of ordinary algebra, insofar as these laws refer to addition and subtraction. We know that all of the more complicated algebraic laws concerning addition and subtraction are consequences of these basic laws. Since the basic laws of the addition and subtraction of matrices are the same as the basic laws of the addition and subtraction of ordinary algebra, all the other laws for the addition and subtraction of matrices must be the same as the corresponding laws for the addition and subtraction of numbers. We can state this as follows:

Insofar as only addition and subtraction are involved, the algebra of matrices is exactly like the ordinary algebra of numbers.

So you do not have to study the algebra of addition and subtraction of matrices - you already know it! But now the algebra that you already know has a new and much richer content. Formerly, it could be applied only to numbers. Now, it can be applied to matrices of any order. Thus, we have obtained a very considerable result with a very small effort, simply by observing that our old algebraic laws of addition and subtraction apply not only to numbers, but also to quite different kinds of things, namely, matrices. This very powerful trick of putting old results in new settings has been used many times, and often with great success, in the most modern mathematics.

A good example of the general principle emphasized above is provided by the following problem. Suppose that A and B are known matrices of the same order. How can we solve the equation

$$X + A = B$$

for the unknown matrix X? The answer is easy. We do exactly what we learned to do with numbers. Add the matrix $-A$ to both sides. This gives

$$X + A + (-A) = B + (-A) = B - A.$$

Since $A + (-A) = \underline{0}$, and $X + \underline{0} = X$, we have

$$X = B - A.$$

This is our solution.

<u>Exercises 1-5</u>

1. Solve the equation

$$X + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

for the matrix X.

2. Solve the equation

$$X + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 & 2 \\ 3 & 2 & 3 \\ 4 & 3 & 4 \end{bmatrix}$$

for the matrix X.

3. If $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} - \begin{bmatrix} -6 & 0 & 2 \end{bmatrix} = \begin{bmatrix} -6 & 2 & -3 \end{bmatrix}$, determine $\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$.

4. If
$$\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \text{determine} \quad \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}.$$

5. If

$$\begin{bmatrix} 2 & -3 \\ 4 & 0 \end{bmatrix} - \begin{bmatrix} x_1 & x_2 \\ y_1 & y_2 \end{bmatrix} = \begin{bmatrix} -3 & 4 \\ 5 & -1 \end{bmatrix},$$

determine x_1, x_2, y_1, and y_2.

6. Prove that if the matrices A, B, and C are conformable for addition, then $(A + C) - (A + B) = C - B$.

7. Is the equation

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 3 \\ 1 & 0 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} - \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}$$

valid?

1-6. Multiplication of a Matrix by a Number

Once we know how to add numbers, it is customary to define $2x$ as the sum $x + x$, $3x$ as the sum $2x + x$, etc. Fractional parts of x are defined by requiring that $(1/2)x + (1/2)x = x$, $(1/3)x + (1/3)x + (1/3)x = x$, etc. All of this can readily be done with matrices. If we add two equal matrices, the sum is clearly a matrix in which each entry is exactly twice the corresponding entry in the two given matrices. Thus

$$\begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 2(2) & 2(3) \\ 2(-1) & 2(0) \end{bmatrix}.$$

Likewise, for three equal matrices we have

$$\begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 9 \\ -3 & 0 \end{bmatrix} = \begin{bmatrix} 3(2) & 3(3) \\ 3(-1) & 3(0) \end{bmatrix}.$$

Each of the above sums may be considered to be the product of a number and a matrix. We write

$$2\begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ -2 & 0 \end{bmatrix},$$

$$3\begin{bmatrix} 2 & 3 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} 6 & 9 \\ -3 & 0 \end{bmatrix}.$$

The equation

$$\frac{1}{2}A + \frac{1}{2}A = A,$$

defining the matrix $(1/2)A$, is clearly satisfied by the matrix each of whose entries is exactly $1/2$ the corresponding entry of A; the equation

$$\frac{1}{3}A + \frac{1}{3}A + \frac{1}{3}A = A,$$

defining the matrix $(1/3)A$, is clearly satisfied by the matrix each of whose entries is exactly $1/3$ the corresponding entry of A.

These considerations lead us to make the following general definition.

Definition 1-6. The product $cA = Ac$ of a number c and an $m \times n$ matrix A is the $m \times n$ matrix B such that the entry b_{ij} in the i-th row and j-th column of B is equal to the product ca_{ij} of the number c and the entry a_{ij} in the i-th row and j-th column of A.

Thus,

$$c\begin{bmatrix} a_{ij} \end{bmatrix}_{m \times n} = \begin{bmatrix} a_{ij} \end{bmatrix}_{m \times n} c = \begin{bmatrix} ca_{ij} \end{bmatrix}_{m \times n}.$$

For example,

$$c\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} \\ ca_{21} & ca_{22} \\ ca_{31} & ca_{32} \end{bmatrix}.$$

Note that here we have defined the product of a matrix by a number, not the product of two matrices. It is possible also to define the product of two matrices; this will be done in Section 1-7.

Now we may state the following theorem about products of matrices by numbers.

<u>Theorem 1-4</u>. If A and B are $m \times n$ matrices, and x and y are numbers, then

$$\begin{aligned}
&\text{(a)} \quad x(yA) = (xy)A,\\
&\text{(b)} \quad (x + y)A = xA + yA,\\
&\text{(c)} \quad (-1)A = -A,\\
&\text{(d)} \quad x(A + B) = xA + xB,\\
&\text{(e)} \quad x\underline{0} = \underline{0},\\
&\text{(f)} \quad 0A = \underline{0}.
\end{aligned}$$

Part (e) states that the product of a number and the zero matrix is the zero matrix, and part (f) states that the product of the zero number and any matrix is the zero matrix.

<u>Proof of Theorem 1-4 (d)</u>. The entry in the i-th row and j-th column of the matrix $A + B$ is $a_{ij} + b_{ij}$. The entry in the i-th row and j-th column of matrix $x(A + B)$ is therefore, by definition, $x(a_{ij} + b_{ij})$. Now the entry in the i-th row and j-th column of the matrix xA is xa_{ij}; that in the i-th row and j-th column of the matrix xB is xb_{ij}. Thus the entry in the i-th row and j-th column of the matrix $xA + xB$ is $xa_{ij} + xb_{ij}$. Since the entries are numbers and, for all numbers, $a(b + c) = ab + ac$, we have

$$x(a_{ij} + b_{ij}) = xa_{ij} + xb_{ij},$$

so that each entry in the matrix $x(A + B)$ is the same as the corresponding entry of the matrix $xA + xB$. Hence,

$$x(A + B) = xA + xB.$$

The other parts of Theorem 1.4 may be proved in a similar way.

When we studied the laws governing the addition and subtraction of matrices, we saw that they were parallel to the laws governing addition and subtraction in ordinary algebra. The situation when we come to the multiplication of matrices by numbers is <u>rather</u> similar, but not exactly the same. The various parts of Theorem 1-4 resemble the basic algebraic laws for multiplication very closely. Thus, many of the more complicated ordinary algebraic laws and procedures governing multiplication still remain correct for expressions involving the multiplication of matrices by numbers. The difference is that while the product of a number by a number is a number, <u>the product of a matrix by a number is not a number but a matrix</u>.

We are now able to solve some matrix equations involving addition, subtraction, and multiplication by a number. Let us look at an example.

Suppose we want to solve the equation

$$-2\left(X + \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}\right) = 3X + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We first perform the indicated multiplication by -2, in accordance with part (d) of the above theorem, to get

$$-2X + \begin{bmatrix} -2 & -4 & -6 \\ 0 & -2 & -4 \\ 0 & 0 & -2 \end{bmatrix} = 3X + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then we add $2X$ to both sides of the equation to obtain

$$\begin{bmatrix} -2 & -4 & -6 \\ 0 & -2 & -4 \\ 0 & 0 & -2 \end{bmatrix} = 3X + 2X + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Next we use part (b) of the theorem to find that $3X + 2X = 5X$, so that

$$\begin{bmatrix} -2 & -4 & -6 \\ 0 & -2 & -4 \\ 0 & 0 & -2 \end{bmatrix} = 5X + \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Adding

$$-\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

to both sides, we find that

$$\begin{bmatrix} -3 & -4 & -6 \\ 0 & -2 & -4 \\ 0 & 0 & -3 \end{bmatrix} = 5X.$$

Multiplying both sides of this last equation by $1/5$, we see by part (a) of the theorem that

$$X = \begin{bmatrix} -3/5 & -4/5 & -6/5 \\ 0 & -2/5 & -4/5 \\ 0 & 0 & -3/5 \end{bmatrix}.$$

This is our solution.

Exercises 1-6

1. For

$$A = \begin{bmatrix} 2 & 1 & -3 \\ 1 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 0 & 5 \\ 6 & 9 & -1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 5 & -1 & 0 \\ 7 & 8 & -1 \end{bmatrix},$$

determine the result of the following operations:

(a) $2A - B + C$, (c) $7A - 2(B - C)$,

(b) $3A - 4B - 2C$, (d) $3(A - 2B + 3C)$.

2. For

$$A = \begin{bmatrix} 2 & 2 & 2 \\ 2 & 1 & -3 \\ 1 & 0 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 3 & 3 \\ 3 & 0 & 5 \\ 6 & 9 & -1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 4 & 4 & 4 \\ 5 & -1 & 0 \\ 7 & 8 & -1 \end{bmatrix},$$

determine the result of the following operations:

(a) $2A - B + C$, (c) $7A - 2(B - C)$,

(b) $3A - 6B + 9C$, (d) $3(A - 2B + 3C)$.

3. Let A, B, and C be the matrices of Exercise 2. Solve the equation

$$\frac{1}{2}(X + A) = 3(X + (2X + B)) + C,$$

giving all the steps in detail, and justifying each step.

4. Let A, B, and C be the matrices of Exercise 2. Solve the equation

$$2(X + B) = 3(X + \tfrac{1}{2}(X + A)) + C.$$

5. Prove Theorem 1-4 (a).

6. Prove Theorem 1-4 (b).

1-7. Multiplication of Matrices

Thus far, we have defined and studied the addition and subtraction of matrices and the multiplication of a matrix by a number. We still have not defined the product of two matrices. Since the formal definition is somewhat complicated and may at first seem odd, let us look at a simple practical problem that will lead us to operate with two matrices in the way that we shall ultimately call multiplication.

In Section 1-1, the number of tubes and the number of speakers used in assembling TV sets of three different models were specified by a table:

	Model A	Model B	Model C
Number of tubes	13	18	20
Number of speakers	2	3	4

This array will be called the parts-per-set matrix.

Suppose orders were received in January for 12 sets of model A, 24 sets of model B, and 12 sets of model C; and in February for 6 sets of model A, 12 of model B, and 9 of model C. We can arrange the information in the form of a matrix:

	January	February
Model A	12	6
Model B	24	12
Model C	12	9

This will be called the sets-per-month matrix.

To determine the number of tubes and speakers required in each of the months for these orders, it is clear that we must use both sets of information. For instance, to compute the number of tubes needed in January, we multiply each entry in the 1st row of the parts-per-set matrix by the corresponding entry in the 1st column of the sets-per-month matrix, and then add the three products. Thus, the number of tubes required in January is

$$13(12) + 18(24) + 20(12) = 828.$$

To compute the number of speakers needed in January, we multiply each entry in the 2nd row of the parts-per-set matrix by the corresponding entry in the 1st column of the sets-per-month matrix and then add the products. Thus, the number of speakers for January is

$$2(12) + 3(24) + 4(12) = 144.$$

For February, first we multiply the entries from the 1st row of the parts-per-set matrix by the corresponding entries from the 2nd column of the sets-per-month matrix and add to determine the number of tubes; secondly, we multiply the entries from the 2nd row of the parts-per-set matrix by the corresponding entries from the 2nd column of the sets-per-month matrix and add to determine the number of speakers. Thus the numbers of tubes and speakers for February are, respectively,

$$13(6) + 18(12) + 20(9) = 474,$$

and

$$2(6) + 3(12) + 4(9) = 84.$$

We can arrange the four sums in an array, which we shall call the parts-per-month matrix:

	January	February
Number of tubes	828	474
Number of speakers	144	84

Can we now represent our "operation" in equation form? Let us try:

$$\begin{bmatrix} 13 & 18 & 20 \\ 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 12 & 6 \\ 24 & 12 \\ 12 & 9 \end{bmatrix} = \begin{bmatrix} 828 & 474 \\ 144 & 84 \end{bmatrix} . \qquad (1)$$

We have "multiplied" the parts-per-set matrix by the sets-per-month matrix to get just what should be expected, the parts-per-month matrix!

Note that, in Equation (1), 828 equals the sum of the products of the entries in the 1st row of the left-hand factor by the corresponding entries in the 1st column of the right-hand factor. Likewise, 474 equals the sum of the products of the entries in the 1st row of the left-hand factor by the corresponding entries in the 2nd column of the right-hand factor, and so on. Consider the "product" matrix

$$\begin{bmatrix} 828 & 474 \\ 144 & 84 \end{bmatrix}$$

in the symbolic form,

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} .$$

The subscripts indicate the row and column in which the entry appears; they also indicate the row and the column of the two factor matrices that are combined to get that entry. Thus, the entry a_{21} in the 2nd row and 1st column is found by adding the products formed when the entries in the 2nd row of the left-hand factor are multiplied by the corresponding entries in the 1st column of the right-hand factor. The most concise description of the process is: "<u>Multiply row by column</u>."

The description, "<u>Multiply row by column</u>," of the pattern in the foregoing simple practical problem serves as our guide in establishing the general rule for the multiplication of two matrices. Very simply the rule is to multiply entries of a row by corresponding entries of a column and then add the products. Thus, given two matrices A and B, to find the entry in the i-th row and j-th column of the <u>product matrix</u> AB, multiply each entry in the i-th row of the left-hand factor A by the corresponding entry in the j-th column of the right-hand factor B, and then add all the resulting terms. Since there must be an entry in each row of the left-hand factor to match with each entry in a column of the right-hand factor, and conversely, it follows that the product is not defined unless the number of columns in the left-hand factor is equal to the number of rows in the right-hand factor. When the number of columns in the left-hand factor equals the number of rows in the right-hand factor, the matrices are conformable for multiplication.

A diagram can aid understanding; see Figure 1-1.

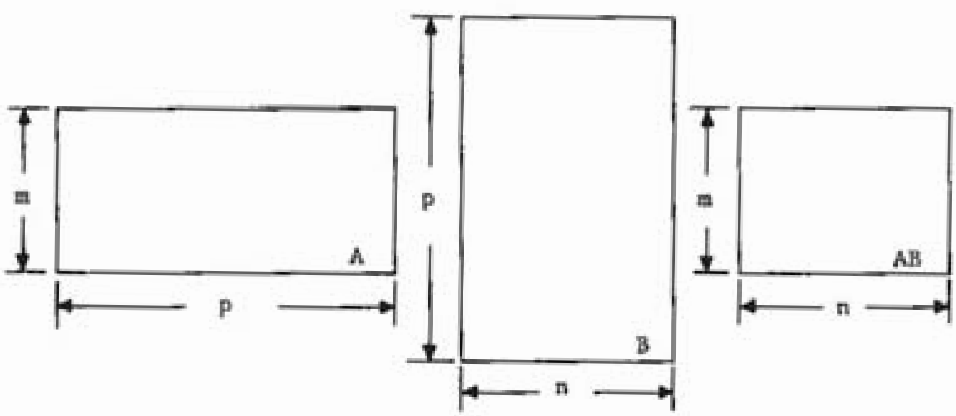


Figure 1-1. Matrices A and B that are conformable for multiplication. The number of columns of A must be equal to the number of rows of B. Then the product AB has the same number of rows as A and the same number of columns as B.

An entry in the product AB is found by multiplying each of the p entries in a row of A by the corresponding one of the p entries in the column of B and taking the sum; see Figure 1-2.

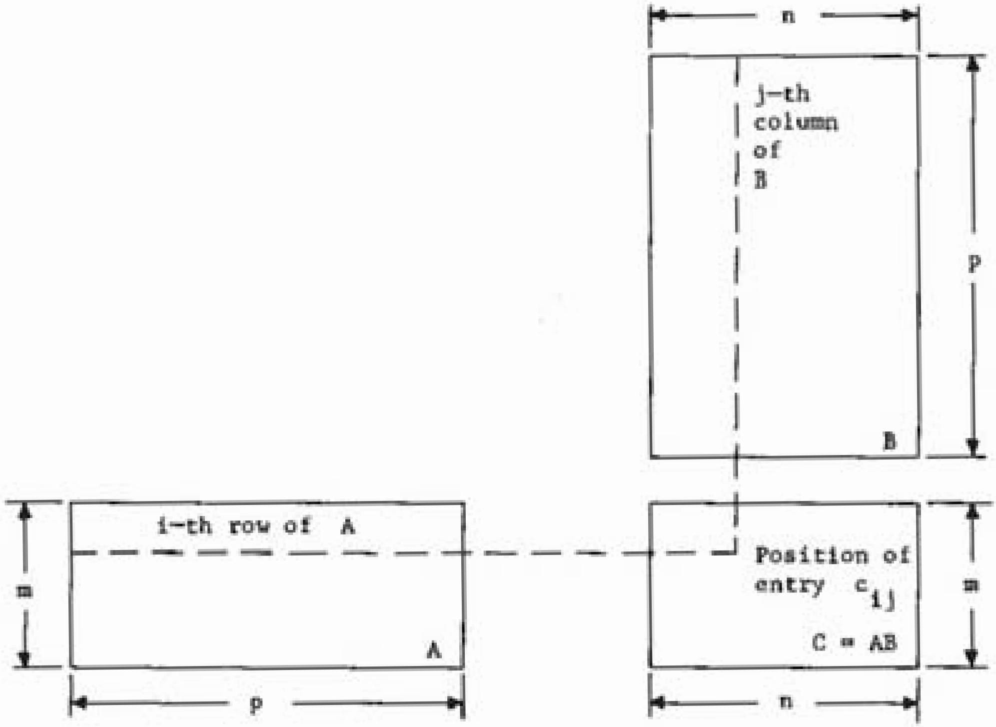
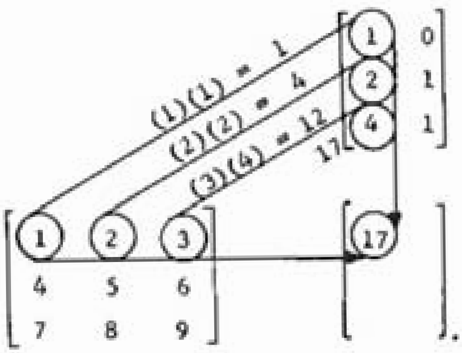


Figure 1-2. Determination of an entry in the product AB of matrices A and B that are conformable for multiplication.

Thus, for the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 4 & 1 \end{bmatrix},$$

to form the product AB, we compute as follows:



Determining one entry of the product after another in this way, we finally obtain the complete answer for the product AB:

$$AB = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 17 & 5 \\ 38 & 11 \\ 59 & 17 \end{bmatrix}.$$

(Check each of the entries of the answer yourself!) To get the answer, 18 multiplications and 12 additions of pairs of numbers are necessary. It might help to think of the first matrix in terms of its rows,

$$A = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix},$$

and the second in terms of its columns,

$$B = \begin{bmatrix} C_1 & C_2 \end{bmatrix}.$$

Then the product appears as

$$AB = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix} \begin{bmatrix} C_1 & C_2 \end{bmatrix} = \begin{bmatrix} R_1C_1 & R_1C_2 \\ R_2C_1 & R_2C_2 \\ R_3C_1 & R_3C_2 \end{bmatrix}.$$

Here, of course, R_1C_1 (for example) represents not a product but a sum of products:

$$R_1C_1 = (1)(1) + (2)(2) + (3)(4) = 17.$$

The symbol R_3C_2 represents the entry in the 3rd row and 2nd column, and you can decide for yourself what symbol represents the entry in the i-th row and j-the column of the product.

Here are some more examples:

(a) $$\begin{bmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1(1) + 2(4) & 1(2) + 2(0) & 1(3) + 2(1) \\ 3(1) + 1(4) & 3(2) + 1(0) & 3(3) + 1(1) \\ -1(1) + 2(4) & -1(2) + 2(0) & -1(3) + 2(1) \end{bmatrix}$$

$$= \begin{bmatrix} 9 & 2 & 5 \\ 7 & 6 & 10 \\ 7 & -2 & -1 \end{bmatrix},$$

(b) $$\begin{bmatrix} 1 & 7 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} = \begin{bmatrix} 1(2) + 7(4) + 3(1) \end{bmatrix} = \begin{bmatrix} 33 \end{bmatrix},$$

(c) $$\begin{bmatrix} 2 \\ 4 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 7 & 3 \end{bmatrix} = \begin{bmatrix} 2(1) & 2(7) & 2(3) \\ 4(1) & 4(7) & 4(3) \\ 1(1) & 1(7) & 1(3) \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 14 & 6 \\ 4 & 28 & 12 \\ 1 & 7 & 3 \end{bmatrix}.$$

Let us now proceed to define multiplication formally.

Definition 1-7. Let

$$A = \begin{bmatrix} a_{ij} \end{bmatrix}_{m\times p} \quad \text{and} \quad B = \begin{bmatrix} b_{jk} \end{bmatrix}_{p\times n}$$

be matrices of order $m \times p$ and $p \times n$, respectively. The product AB is the matrix of order $m \times n$, of which the entry in the i-th row and the j-th column is the sum of the products formed by multiplying entries of the i-th row of A by corresponding entries of the j-th column of B.

The definition of the product of two matrices can be expressed in terms of the "$\sum$ notation" for sums. Recall that, in the "$\sum$ notation," we write the sum

$$S = X_1 + X_2 + \cdots + X_p$$

of p numbers as

$$S = \sum_{j=1}^{p} X_j.$$

For example,

$$\sum_{j=1}^{5} j^2 = 1^2 + 2^2 + 3^2 + 4^2 + 5^2 = 55.$$

Again, the familiar formula for the sum of the first p positive integers,

$$1 + 2 + \cdots + p = \frac{p(p + 1)}{2},$$

can be expressed as

$$\sum_{j=1}^{p} j = \frac{p(p + 1)}{2}.$$

In this notation, the sum

$$a_{31}b_{14} + a_{32}b_{24} + \cdots + a_{3p}b_{p4}$$

is expressed as

$$\sum_{j=1}^{p} a_{3j}b_{j4}.$$

You will recognize this as the element in the third row and fourth column of the matrix AB. More generally,

$$a_{i1}b_{1k} + a_{i2}b_{2k} + \cdots + a_{ip}b_{pk}$$

is expressed as

$$\sum_{j=1}^{p} a_{ij}b_{jk};$$

this sum is the element in the i-th row and k-th column of AB. Thus we can express Definition 1-7 more compactly as follows:

<u>Definition 1-7'</u>. Let

$$A = \left[a_{ij}\right]_{m \times p} \quad \text{and} \quad B = \left[b_{jk}\right]_{p \times n}$$

be matrices of order $m \times p$ and $p \times n$, respectively. The <u>product</u> AB is the matrix of order $m \times n$, given by

$$AB = \left[a_{ij}\right]_{m \times p} \left[b_{jk}\right]_{p \times n} = \left[\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right)\right]_{m \times n} = \left[c_{ik}\right]_{m \times n}.$$

Note that we have defined the product of two matrices only when the number of columns of the left-hand factor is the same as the number of rows of the right-hand factor. Also note that the number of rows in the product is the same as the number of rows in the left-hand factor, and that the number of columns in the product is the same as the number of columns in the right-hand factor.

<u>Exercises 1-7</u>

1. Let

$$A = \begin{bmatrix} 2 & 0 & 3 \\ -1 & 4 & 9 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 0 & 11 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 1 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad D = \begin{bmatrix} -1 & -1 \\ 2 & 2 \\ -3 & -3 \end{bmatrix}.$$

State the orders of each of the following matrices:

(a) AB,
(b) DA,
(c) AD,
(d) CB,
(e) BD,
(f) $D(AB)$,
(g) $(CB)(DA)$,
(h) $B(DA)$.

2. Perform the following matrix multiplication, where possible:

(a) $\begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$,

(b) $\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \end{bmatrix}$,

(c) $\begin{bmatrix} 2 & 3 & 4 \\ -1 & -2 & 0 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 6 & 1 \\ 3 & 5 \end{bmatrix}$,

(d) $\begin{bmatrix} 4 & 2 \\ 6 & 1 \\ 3 & 5 \end{bmatrix} \begin{bmatrix} 2 & 3 & 4 \\ -1 & -2 & 0 \end{bmatrix}$,

(e) $\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 6 \end{bmatrix}$,

(f) $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 6 \end{bmatrix} \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$.

3. Let $X = \begin{bmatrix} 2 & -2 & 4 \end{bmatrix}$, $Y = \begin{bmatrix} 0 & 1 & 2 \end{bmatrix}$,

$$U = \begin{bmatrix} 5 \\ 0 \\ 1 \end{bmatrix}, \text{ and } W = \begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}.$$

Compute the following:

(a) $5UX$,

(b) $(5W)(3Y)$,

(c) $5XU - (2X - Y)W$,

(d) $XU + YW$,

(e) $(U - W)(X + Y)$,

(f) $(X + Y)(U - W)$.

4. Perform the following matrix multiplications:

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 6 & 4 \\ 2 & 0 \end{bmatrix}$,

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$,

(c) $\begin{bmatrix} r_1 & r_2 & r_3 \\ s_1 & s_2 & s_3 \\ t_1 & t_2 & t_3 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(d) $\begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \alpha_2 & 0 \\ 0 & 0 & \alpha_3 \end{bmatrix} \begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$,

(e) $\begin{bmatrix} 0 & 0 & 0 \\ 0 & b_1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \\ z_1 & z_2 & z_3 \end{bmatrix}$,

(f) $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(g) $\begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 1 & 2 \\ 1 & 3 & 1 \end{bmatrix}$.

5. For the matrices

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & 0 \\ 4 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix},$$

test the rule that $(AB)C = A(BC)$.

6. Let

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ -2 & 0 & 1 \end{bmatrix}.$$

Compute

(a) AB,

(b) AB^t,

(c) BB^t,

(d) $(AB)B^t$,

(e) $A(BB^t)$,

(f) $A(B + B^t)$,

(g) $A(B - B^t)$,

(h) $AB - AB^t$,

(i) $AA - BB + B^tB^t$,

(j) $(AA)A$.

7. Let I denote the identity matrix of order 3 (see page 46):

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Let A and B be as in Exercise 6. Compute AI, BI, and B^tI. Compute $(AI)I$ and $((AI)I)I$.

8. Let

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 1 & 2 & 1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}.$$

Find $(AB)^t$ and $B^t A^t$.

9. For a certain manufacturing plant, the following information is given:

	Part 1	Part 2	Part 3
Cost	2	3	5

	Subassembly 1	Subassembly 2
Part 1	4	1
Part 2	3	5
Part 3	7	2

	Model 1	Model 2	Model 3
Subassembly 1	2	1	2
Subassembly 2	3	4	5

	Day 1	Day 2	Day 3
Model 1	7	8	8
Model 2	3	4	5
Model 3	3	5	6

Determine the parts-per-model matrix and the cost-per-day matrix.

1-8. Properties of Matrix Multiplication

We have learned that insofar as only addition and subtraction are involved, the algebra of matrices is exactly like the ordinary algebra of numbers; see Section 1-5. At this moment, we might be concerned about multiplication since the definition seems a bit unusual. Is the algebra of matrices like the ordinary algebra of numbers insofar as multiplication is concerned?

Let us consider an example that will yield an answer to the foregoing question. Let

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

If we compute AB, we find

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Now, if we reverse the order of the factors and compute BA, we find

$$BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Thus AB and BA are different matrices!

For another example, let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1(1) + 2(4) & 1(2) + 2(0) & 1(3) + 2(1) \\ 3(1) + 1(4) & 3(2) + 1(0) & 3(3) + 1(1) \\ -1(1) + 2(4) & -1(2) + 2(0) & -1(3) + 2(1) \end{bmatrix} = \begin{bmatrix} 9 & 2 & 5 \\ 7 & 6 & 10 \\ 7 & -2 & -1 \end{bmatrix},$$

while

$$BA = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1(1) + 2(3) + 3(-1) & 1(2) + 2(1) + 3(2) \\ 4(1) + 0(3) + 1(-1) & 4(2) + 0(1) + 1(2) \end{bmatrix} = \begin{bmatrix} 4 & 10 \\ 3 & 10 \end{bmatrix}.$$

Again AB and BA are different matrices; they are not even of the same order!

Thus we have a first difference between matrix algebra and ordinary algebra, and a very significant difference it is indeed. When we multiply real numbers, we can rearrange factors since the commutative law holds: For all $x \in R$ and $y \in R$, we have $xy = yx$. When we multiply matrices, we have no such law and we must consequently be careful to take the factors in the order given. We must consequently distinguish between the result of multiplying B on the right by A to get BA, and the result of multiplying B on the left by A to get AB. In the algebra of numbers, these two operations of "right multiplication" and "left multiplication" are the same; in matrix algebra, they are not necessarily the same.

Let us explore some more differences! Let

$$A = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix}.$$

Patently, $A \neq \underline{0}$ and $B \neq \underline{0}$. But if we compute AB, we obtain

$$AB = \begin{bmatrix} 3 & 1 \\ 6 & 2 \end{bmatrix} \begin{bmatrix} -1 & 3 \\ 3 & -9 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix};$$

thus, we find $AB = \underline{0}$. Again, let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The second major difference between ordinary algebra and matrix algebra is that the product of two matrices can be a zero matrix without either factor being a zero matrix.

The breakdown for matrix algebra of the law that $xy = yx$ and of the law that $xy = 0$ only if either x or y is zero causes additional differences.

For instance, for real numbers we know that if $ab = ac$, and $a \neq 0$, then $b = c$. This property is called the <u>cancellation law for multiplication</u>.

<u>Proof</u>. We divided the proof into simple steps:

(a) $ab = ac$,

(b) $ab - ac = 0$,

(c) $a(b - c) = 0$,

(d) $b - c = 0$,

(e) $b = c$.

For matrices, the above step from (c) to (d) fails and the proof is not valid. In fact, AB can be equal to AC, with $A \neq \underline{0}$, yet $B \neq C$. Thus, let

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 2 & 2 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

Then

$$AB = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \\ 3 & 2 & -7 \end{bmatrix} = AC,$$

and

$$A \neq \underline{0},$$

but

$$B \neq C.$$

Let us consider another difference. We know that a real number a can have at most two square roots; that is, there are at most two roots of the equation $xx = a$.

<u>Proof</u>. Again, we give the simple steps of the proof:

(a) Suppose that $yy = a$; then

(b) $xx = yy$,

(c) $xx - yy = 0$,

(d) $(x - y)(x + y) = xx + (-yx + xy) - yy$,

(e) $yx = xy$.

(f) From (d) and (e), $(x - y)(x + y) = xx - yy$.

(g) From (c) and (f), $(x - y)(x + y) = 0$.

(h) Therefore, either $x - y = 0$ or $x + y = 0$.

(i) Therefore, either $x = y$ or $x = -y$.

For matrices, statement (e) is false, and therefore the steps to (f) and (g) are invalid. Even if (g) were valid, the step from (g) to (h) fails.

Therefore, the foregoing proof is invalid if we try to apply it to matrices. In fact, it is false that a matrix can have at most two square roots: We have

$$\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},$$

$$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},$$

$$\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},$$

$$\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}\begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

Thus the matrix

$$I = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

has the <u>four different</u> square roots

$$I = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}, \quad J = \begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}, \quad K = \begin{bmatrix}-1 & 0\\0 & 1\end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix}-1 & 0\\0 & -1\end{bmatrix}.$$

There are more! Given any number $x \neq 0$, we have

$$\begin{bmatrix}0 & x\\1/x & 0\end{bmatrix}\begin{bmatrix}0 & x\\1/x & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}.$$

By giving x any one of an infinity of different real values, we obtain an infinity of different square roots of the matrix I:

$$\begin{bmatrix}0 & 2\\1/2 & 0\end{bmatrix}, \quad \begin{bmatrix}0 & 1/3\\3 & 0\end{bmatrix}, \quad \begin{bmatrix}0 & -4\\-1/4 & 0\end{bmatrix}, \quad \text{etc.}$$

Thus the very simple 2×2 matrix I has infinitely many distinct square roots! You can see, then, that the fact that a real or complex <u>number</u> has at most two square roots is by no means trivial.

Exercises 1-8

1. Let

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix}.$$

Calculate:

(a) AB,
(b) BA,
(c) $(AB)A$,
(d) $(BA)A$,
(e) $(BA)B$,
(f) $B(BA)$,
(g) $A(AB)$,
(h) $((BA)A)B$,
(i) $((AB)A)B$.

2. Make the calculations of Exercise 1 for the matrices

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}.$$

3. Let A and B be as in Exercise 2, and let

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Calculate AI, IA, BI, IB, and $(AI)B$.

4. Let

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 1 & 2 \end{bmatrix}.$$

Show by computation that

(a) $(A + B)(A + B) \neq A^2 + 2AB + B^2$,

(b) $(A + B)(A - B) \neq A^2 - B^2$,

where A^2 and B^2 denote AA and BB, respectively.

5. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Calculate A^2, A^3, B^2, B^3, AB^3, A^2B, where A^3 denotes $A(AA)$.

6. Find at least 8 square roots of the matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

7. Show that the matrix

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

satisfies the equation $A^2 = \underline{0}$. How many 2×2 matrices satisfying this equation can you find?

8. Show that the matrix

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

satisfies the equation $A^3 = \underline{0}$.

1-9. Properties of Matrix Multiplication (Concluded)

We have seen that two basic laws governing multiplication in the algebra of ordinary numbers break down when it comes to matrices. The commutative law and the cancellation law do not hold. At this point, you might fear a total collapse of all the other familiar laws. This is not the case. Aside from the two laws mentioned, and the fact that, as we shall see later, many matrices do not have multiplicative inverses (reciprocals), the other basic laws of ordinary algebra generally remain valid for matrices. The associative law holds for the multiplication of matrices and there are distributive laws that unite addition and multiplication.

A few examples will aid us in understanding the laws.

Let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}.$$

Then

$$A(BC) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\left(\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -2 & 4 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ 0 & 7 \end{bmatrix},$$

and

$$(AB)C = \left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\right)\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 4 \\ 0 & 7 \end{bmatrix}.$$

Thus, in this case,

$$A(BC) = (AB)C.$$

Again,

$$A(B + C) = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\left(\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix},$$

and

$$AB + AC = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix},$$

so that also

$$A(B + C) = AB + AC. \qquad (1)$$

Since multiplication is not commutative, we cannot conclude from Equation (1) that the distributive principle is valid with the factor A on the right-hand side of $B + C$. Having illustrated the left-hand distributive law, we now illustrate the right-hand distributive law with the following example: We have

$$(B + C)A = \left(\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 2 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 6 & 2 \end{bmatrix}$$

and

$$BA + CA = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 0 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 6 & 2 \end{bmatrix}.$$

Thus,

$$(B + C)A = BA + CA.$$

You might note, in passing, that, in the above example,

$$A(B + C) \neq (B + C)A.$$

These properties of matrix multiplication can be expressed as theorems, as follows.

<u>Theorem 1-5</u>. If

$$A = \left[a_{ij}\right]_{m \times p}, \quad B = \left[b_{jk}\right]_{p \times n}, \quad \text{and} \quad C = \left[c_{kh}\right]_{n \times q},$$

then

$$(AB)C = A(BC).$$

<u>Proof</u>. (Optional.) We have

$$AB = \left[\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right)\right]_{m\times n},$$

$$(AB)C = \left[\sum_{k=1}^{n}\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right) c_{kh}\right]_{m\times q};$$

$$BC = \left[\left(\sum_{k=1}^{n} b_{jk}c_{kh}\right)\right]_{p\times q},$$

$$A(BC) = \left[\sum_{j=1}^{p} a_{ij}\left(\sum_{k=1}^{n} b_{jk}c_{kh}\right)\right]_{m\times q}.$$

Since the order of addition is arbitrary, we know that

$$\left[\sum_{k=1}^{n}\left(\sum_{j=1}^{p} a_{ij}b_{jk}\right) c_{kh}\right]_{m\times q} = \left[\sum_{j=1}^{p} a_{ij}\left(\sum_{k=1}^{n} b_{jk}c_{kh}\right)\right]_{m\times q}.$$

Hence,

$$(AB)C = A(BC).$$

<u>Theorem 1-6</u>. If

$$A = \left[a_{ij}\right]_{m\times p}, \quad B = \left[b_{jk}\right]_{p\times n}, \quad \text{and} \quad C = \left[c_{jk}\right]_{p\times n},$$

then $A(B + C) = AB + BC$.

<u>Proof</u>. (Optional.) We have

$$(B + C) = \left[b_{jk} + c_{jk}\right]_{p\times n},$$

$$A(B + C) = \left[\sum_{j=1}^{p} a_{ij}\left(b_{jk} + c_{jk}\right)\right]_{m\times n}$$

$$= \left[\sum_{j=1}^{p} a_{ij}b_{jk} + \sum_{j=1}^{p} a_{ij}c_{jk}\right]_{m\times n}$$

$$= \left[\sum_{j=1}^{p} a_{ij}b_{jk}\right]_{m\times n} + \left[\sum_{j=1}^{p} a_{ij}c_{jk}\right]_{m\times n}$$

$$= AB + AC.$$

Hence,

$$A(B + C) = AB + AC.$$

Theorem 1-7. If

$$B = \left[b_{jk}\right]_{p\times n}, \quad C = \left[c_{jk}\right]_{p\times n}, \quad \text{and} \quad A = \left[a_{ki}\right]_{n\times q},$$

then $(B + C)A = BA + CA$.

Proof. The proof is similar to that of Theorem 1-6 and will be left as an exercise for the student.

It should be noted that if the commutative law held for matrices, it would be unnecessary to prove Theorems 1-6 and 1-7 separately, since the two statements

$$A(B + C) = AB + AC$$

and

$$(B + C)A = BA + CA$$

would be equivalent. For matrices, however, the two statements are not equivalent, even though both are true. The order of factors is most important, since statements like

$$A(B + C) = AB + CA$$

and

$$(B + C)A = AB + CA$$

can be false for matrices.

Earlier we defined the zero matrix of order $m \times n$ and showed that it is the identity element for matrix addition:

$$A + \underline{0} = A,$$

where A is any matrix of order $m \times n$. This zero matrix plays the same role in the multiplication of matrices as the number zero does in the multiplication of real numbers. For example, we have

$$\begin{bmatrix} 2 & 0 & 3 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0_2.$$

Theorem 1-8. For any matrix

$$A_{p \times n} = \begin{bmatrix} a_{ij} \end{bmatrix}_{p \times n},$$

we have

$$0_{m \times p} A_{p \times n} = 0_{m \times n} \quad \text{and} \quad A_{p \times n} 0_{n \times q} = 0_{p \times q}.$$

The proof is easy and is left to the student.

Now we may be wondering if there is an identity element for the multiplication of matrices, namely a matrix that plays the same role as the number 1 does in the multiplication of real numbers. (For all real numbers a, $1a = a = a1$.) There is such a matrix, called the unit matrix, or the identity matrix for multiplication, and denoted by the symbol I. The matrix I_2, namely,

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

is called the unit matrix of order 2. The matrix

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is called the unit matrix of order 3. In general, the unit matrix of order n is the square matrix $\begin{bmatrix} e_{ij} \end{bmatrix}_{n \times n}$ such that $e_{ij} = 1$ for all $i = j$ and

$e_{ij} = 0$ for all $i \neq j$ $(i = 1,2,\ldots,n;\ j = 1,2,\ldots,n)$. We now state the important property of the unit matrix as a theorem.

Theorem 1-9. If A is an $m \times n$ matrix, then $AI_n = A$ and $I_mA = A$.

Proof. By definition, the entry in the i-th row and j-th column of the product AI_n is the sum $a_{i1}e_{1j} + a_{i2}e_{2j} + \cdots + a_{in}e_{nj}$. Since $e_{kj} = 0$ for all $k \neq j$, all terms but one in this sum are zero and drop out. We are left with $a_{ij}e_{jj} = a_{ij}$. Thus the entry in the i-th row and j-th column of the product is the same as the corresponding entry in A. Hence $AI_n = A$. The equality $I_mA = A$ may be proved the same way. In most situations, it is not necessary to specify the order of the unit matrix since the order is inferred from the context. Thus, for

$$I_mA = A = AI_n,$$

we write

$$IA = A = AI.$$

For example, we have

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$$

and

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}.$$

Exercises 1-9

1. Let

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Test the formulas

$$A(B + C) = AB + AC,$$

$$(B + C)A = BA + CA,$$

$$A(B + C) = AB + CA,$$

$$A(B + C) = BA + CA.$$

Which are correct, and which are false?

2. Let

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Show that $AB \neq \underline{0}$, but $BA = \underline{0}$.

3. Show that for all matrices A and B of the form

$$A = \begin{bmatrix} a & b \\ -b & a \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} c & d \\ -d & c \end{bmatrix},$$

we have

$$AB = BA.$$

Illustrate by assigning numerical values to a, b, c, and d, with a, b, c, and d integers.

4. Find the value of x for which the following product is I:

$$\begin{bmatrix} 2 & 0 & 7 \\ 0 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix} \begin{bmatrix} -x & -14x & 7x \\ 0 & 1 & 0 \\ x & 4x & -2x \end{bmatrix}.$$

5. For the matrices

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 0 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix},$$

show that $AB = BA$, that $AC = CA$, and that $BC = CB$.

6. Show that the matrix

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

satisfies the equation $A^3 = I$. Find at least one more solution of this equation.

7. Show that for all matrices A of the form

$$A = \begin{bmatrix} ab & b^2 \\ -a^2 & -ab \end{bmatrix},$$

we have

$$A^2 = \underline{0}.$$

Illustrate by assigning numerical values to a and b.

8. Let

$$D = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

Compute the following:

(a) DE,
(b) DF,
(c) EF,
(d) ED,
(e) FD,
(f) FE.

If $AB = -BA$, A and B are said to be <u>anticommutative</u>. What conclusions can be drawn concerning D, E, and F?

9. Show that the matrix $A = \begin{bmatrix} 3 & 1 \\ -1 & 2 \end{bmatrix}$ is a solution of the equation $A^2 - 5A + 7I = \underline{0}$.

10. Explain why, in matrix algebra,

$$(A + B)\,(A - B) \neq A^2 - B^2$$

except in special cases. Can you devise two matrices A and B that will illustrate the inequality? Can you devise two matrices A and B

that will illustrate the special case? (Hint: Use square matrices of order 2.)

11. Show that if V and W are $n \times 1$ column vectors, then

$$V^t W = W^t V.$$

12. Prove that $(AB)^t = B^t A^t$, assuming that A and B are conformable for multiplication.

13. Using $\sum$ notation, prove the right-hand distributive law (Theorem 1.7).

1-10. Summary

In this introductory chapter we have defined several operations on matrices, such as addition and multiplication. These operations differ from those of elementary algebra in that they cannot always be performed. Thus, we do not add a 2×2 matrix to a 3×3 matrix; again, though a 4×3 matrix and a 3×4 matrix can be multiplied together, the product is neither 4×3 nor 3×4. More importantly, the commutative law for multiplication and the cancellation law do not hold.

There is a third significant difference that we shall explore more fully in later chapters but shall introduce now. Recall that the operation of subtraction was closely associated with that of addition. In order to solve equations of the form

$$A + X = B,$$

it is convenient to employ the additive inverse, or negative, $-A$. Thus, if the foregoing equation holds, then we have

$$A + X + (-A) = B + (-A),$$

$$X + A + (-A) = B + (-A),$$

$$X + \underset{\sim}{0} = B - A,$$

$$X = B - A.$$

As you know, every matrix A has a negative $-A$. Now "division" is closely associated with multiplication in a parallel manner. In order to solve equations of the form

$$AX = B,$$

we would analogously employ multiplicative inverse (or reciprocal), which is denoted by the symbol A^{-1}. The defining property is $A^{-1}A = I = AA^{-1}$. This enables us to solve equations of the form

$$AX = B.$$

Thus if the foregoing equation holds, and if A has a multiplicative inverse A^{-1}, then

$$A^{-1}(AX) = A^{-1}B,$$

$$(A^{-1}A)X = A^{-1}B,$$

$$IX = A^{-1}B,$$

$$X = A^{-1}B.$$

Now, many matrices other than the zero matrix $\underline{0}$ do not possess multiplicative inverses; for instance,

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 & -3 \\ -2 & 3 \end{bmatrix}$$

are matrices of this sort. This fact constitutes a very significant difference between the algebra of matrices and the algebra of real numbers. In the next two chapters, we shall explore the problem of matrix inversion in depth.

Before closing this chapter, we should note that matrices arising in scientific and industrial applications are much larger and their entries much more complicated than has been the case in this chapter. As you can imagine, the computations involved when dealing with larger matrices (of order 10 or more), which is usual in applied work, are so extensive as to discourage their use in hand computations. Fortunately, the recent development of high-speed electronic computers has largely overcome this difficulty and thereby has made it more feasible to apply matrix methods in many areas of endeavor.

Chapter 2

THE ALGEBRA OF 2×2 MATRICES

2-1. Introduction

In Chapter 1, we considered the elementary operations of addition and multiplication for rectangular matrices. This algebra is similar in many respects to the algebra of real numbers, although there are important differences. Specifically, we noted that the commutative law and the cancellation law do not hold for matrix multiplication, and that division is not always possible.

With matrices, the whole problem of division is a very complex one; it is centered around the existence of a multiplicative inverse. Let us ask a question: If you were given the matrix equation

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 8 & 9 & 0 & -1 \\ 4 & 5 & 6 & 5 \\ 0 & 4 & 2 & 0 \end{bmatrix} \begin{bmatrix} X_{11} & \cdots & X_{14} \\ . & & . \\ . & & . \\ X_{41} & \cdots & X_{44} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 2 \\ 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \end{bmatrix},$$

could you solve it for the unknown 4×4 matrix X? Do not be dismayed if your answer is "No." Eventually, we shall learn methods of solving this equation, but the problem is complex and lengthy. In order to understand this problem in depth and at the same time comprehend the full significance of the algebra we have developed so far, we shall largely confine our attention in this chapter to a special subset of the set of all rectangular matrices; namely, we shall consider the set of 2×2 square matrices.

When one stands back and takes a broad view of the many different kinds of numbers that have been studied, one sees recurring patterns. For instance, let us look at the rational numbers for a moment. Here is a set of numbers that we can add and multiply. This statement is so simple that we almost take it for granted. But it is not true of all sets, so let us give a name to the notion that is involved.

Definition 2-1. A set S is said to be closed under an operation R on a first member a of S and a second member b of S if

(i) the operation can be performed on each a and b of S,

(ii) for each a and b of S, the result of the operation is a member of S.

For example, the set of positive integers is not closed under the operation of division since for some positive integers a and b the ratio a/b is not a positive integer; neither is the set of rational numbers closed under division, since the operation cannot be performed if $b = 0$; but the set of positive rational numbers is closed under division since the quotient of two positive rational numbers is a positive rational number.

Under addition and multiplication, the set of rational numbers satisfies the following postulates:

The set is closed under addition.

Addition is commutative.

Addition is associative.

There is an identity member (0) for addition.

There is an additive inverse member $-a$ for each member a.

. . .

The set is closed under multiplication.

Multiplication is commutative.

Multiplication is associative.

There is an identity member (1) for multiplication.

There is a multiplicative inverse member a^{-1} for each member a, other than 0.

. . .

Multiplication is distributive over addition.

Since there exists a rational multiplicative inverse for each rational number except 0, division (except by 0) is always possible in the algebra of rational numbers. In other words, all equations of the form

$$ax = b,$$

where a and b are rational numbers and $a \neq 0$, can be solved for x in the algebra of rational numbers. For example: In order to solve the equation

$$-\frac{2}{3}x = \frac{1}{2},$$

we multiply both sides of the equation by $-3/2$, the multiplicative inverse of $-2/3$. Thus we obtain

$$\left(-\frac{3}{2}\right)\left(-\frac{2}{3}\right)x = \left(-\frac{3}{2}\right)\left(\frac{1}{2}\right),$$

or

$$x = -\frac{3}{4},$$

which is a rational number.

The foregoing set of postulates is satisfied also by the set of real numbers. Any set that satisfies such a set of postulates is called a field. Both the set of real numbers and the set of rationals, which is a subset of the set of real numbers, are fields under addition and multiplication. There are many systems that have this same pattern. In each of these systems, division (except by 0) is always possible.

Now our immediate concern is to explore the problem of division in the set of matrices. There is no blanket answer that can readily be reached, although there is an answer that we can find by proceeding stepwise. At first, let us limit our discussion to the set of 2×2 matrices. We do this not only to consider division in a smaller domain, but also to study in detail the algebra associated with this subset. A more general problem of matrix division will be considered in Chapter 3.

Exercises 2-1

1. Determine which of the following sets are closed under the stated operation:

   (a) the set of integers under addition,

   (b) the set of even numbers under multiplication,

   (c) the set $\{1\}$ under multiplication,

   (d) the set of positive irrational numbers under division,

(e) the set of integers under the operation of squaring,

(f) the set of numbers $A = \{x: x \geq 3\}$ under addition.

2. Determine which of the following statements are true, and state which of the indicated operations are commutative:

(a) $2 - 3 = 3 - 2$,

(b) $4 \div 2 = 2 \div 4$,

(c) $3 + 2 = 2 + 3$,

(d) $\sqrt{a} + \sqrt{b} = \sqrt{b} + \sqrt{a}$, a and b positive,

(e) $a - b = b - a$, a and b real,

(f) $pq = qp$, p and q real,

(g) $\sqrt{-1} + 2 = 2 + \sqrt{-1}$.

3. Determine which of the following operations $\ddagger$, defined for positive integers in terms of addition and multiplication, are commutative:

(a) $x \ddagger y = x + 2y$ (for example, $2 \ddagger 3 = 2 + 6 = 8$),

(b) $x \ddagger y = 2xy$,

(c) $x \ddagger y = 2x + 2y$,

(d) $x \ddagger y = xy^2$,

(e) $x \ddagger y = x^y$,

(f) $x \ddagger y = x + y + 1$.

4. Determine which of the following operations $*$, defined for positive integers in terms of addition and multiplication, are associative:

(a) $x * y = x + 2y$ (for example, $(2 * 3) * 4 = 8 * 4 = 16$),

(b) $x * y = x + y$,

(c) $x * y = xy^2$,

(d) $x * y = x$,

(e) $x * y = \sqrt{xy}$,

(f) $x * y = xy + 1$.

5. Determine whether the operation $*$ is distributive over the operation $\ddagger$, that is, determine whether $x * (y \ddagger z) = (x * y) \ddagger (x * z)$ and

$(y \mp z) * x = (y * x) \mp (z * x)$, where the operations $\mp$ and $*$ are defined for positive integers in terms of addition and multiplication of real numbers:

(a) $x \mp y = x + y$, $x * y = xy$;

(b) $x \mp y = 2x + 2y$, $x * y = \frac{1}{2} xy$;

(c) $x \mp y = x + y + 1$, $x * y = xy$.

Why is the answer the same in each case for left-hand distribution as it is for right-hand distribution?

6. In each of the following examples, determine if the specified set, under addition and multiplication, constitutes a field:

   (a) the set of all positive numbers,

   (b) the set of all rational numbers,

   (c) the set of all real numbers of the form $a + b\sqrt{2}$, where a and b are integers,

   (d) the set of all complex numbers of the form $a + bi$, where a and b are real numbers and $i = \sqrt{-1}$.

2-2. The Ring of 2 × 2 Matrices

Since we are confining our attention to the subset of 2×2 matrices, it is very convenient to have a symbol for this subset. We let M denote the set of all 2×2 matrices. If A is a member, or element, of this set, we express this membership symbolically by $A \in M$. Since all elements of M are matrices, our general definitions of addition and multiplication hold for this subset.

The set M is not a field, as defined in Section 2-1, since M does not have all the properties of a field; for example, you saw in Chapter 1 that multiplication is not commutative in M. Thus, for

$$A = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

we have

$$AB = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{while} \quad BA = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Let us now consider a less restrictive sort of mathematical system known as a ring; this name is usually attributed to David Hilbert (1862–1943).

Definition 2-2. A ring is a set with two operations, called addition and multiplication, that possesses the following properties under addition and multiplication:

The set is closed under addition.

Addition is commutative.

Addition is associative.

There is an identity element for addition.

There is an additive inverse for each element.

. . .

The set is closed under multiplication.

Multiplication is associative.

. . .

Multiplication is distributive over addition.

Does the set M satisfy these properties? It seems clear that it does, but the answer is not quite obvious. Consider the set of all real numbers. This set is a field because there exists, among other things, an additive inverse for each number in this set. Now the positive integers are a subset of the real numbers. Does this subset contain an additive inverse for each element? Since we do not have negative integers in the set under consideration, the answer is "No"; therefore, the set of positive integers is not a field. Thus a subset does not necessarily have the same properties as the complete set.

To be certain that the set M is a ring, we must systematically make sure that each ring criterion is satisfied. For the most part, our proof will be a reiteration of the material in Chapter 1, since the general properties of matrices will be valid for the subset M of 2×2 matrices. The sum of two 2×2 matrices is a 2×2 matrix; that is, the set is closed under addition. For example,

$$\begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} -1 & 2 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 5 & 2 \end{bmatrix}.$$

The general proofs of commutativity and associativity are valid. The unit matrix is

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix};$$

the zero matrix is

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix};$$

and the additive inverse of the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is

$$\begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}.$$

When we consider the multiplication of 2×2 matrices, we must first verify that the product is an element of this set, namely a 2×2 matrix. Recall that the number of rows in the product is equal to the number of rows in the left-hand factor, and the number of columns is equal to the number of columns in the right-hand factor. Thus, the product of two elements of the set M must be an element of this set, namely a 2×2 matrix; accordingly, the set is closed under multiplication. For example,

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix}.$$

The general proof of associativity is valid for elements of M, since it is valid for rectangular matrices. Also, both of the distributive laws hold for elements of M by the same reasoning. For example, to illustrate the associative law for multiplication, we have

$$\left(\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}\right)\begin{bmatrix}0 & 2\\1 & 0\end{bmatrix} = \begin{bmatrix}-1 & 1\\-1 & 1\end{bmatrix}\begin{bmatrix}0 & 2\\1 & 0\end{bmatrix} = \begin{bmatrix}1 & -2\\1 & -2\end{bmatrix},$$

and also

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\left(\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix}\begin{bmatrix}0 & 2\\1 & 0\end{bmatrix}\right) = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}-1 & 2\\1 & -2\end{bmatrix} = \begin{bmatrix}1 & -2\\1 & -2\end{bmatrix};$$

and to illustrate the left-hand distributive law, we have

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\left(\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix} + \begin{bmatrix}0 & 2\\1 & 0\end{bmatrix}\right) = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & 1\\0 & 1\end{bmatrix} = \begin{bmatrix}1 & 3\\3 & 7\end{bmatrix}$$

and also

$$\begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}1 & -1\\-1 & 1\end{bmatrix} + \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}\begin{bmatrix}0 & 2\\1 & 0\end{bmatrix} = \begin{bmatrix}-1 & 1\\-1 & 1\end{bmatrix} + \begin{bmatrix}2 & 2\\4 & 6\end{bmatrix} = \begin{bmatrix}1 & 3\\3 & 7\end{bmatrix}.$$

Since we have checked that each of the ring postulates is fulfilled, we have shown that the set M of 2×2 matrices is a ring under addition and multiplication. We state this result formally as a theorem.

Theorem 2-1. The set M of 2×2 matrices is a ring under addition and multiplication.

Since the list of defining properties for a field contains all the defining properties for a ring, it follows that every field is a ring. But the converse statement is not true; for example, we now know that the set M of 2×2 matrices is a ring but not a field. The set M has one more of the field properties, namely there is an identity element

$$I = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$$

for multiplication in M; that is, for each $A \in M$ we have

$$IA = A = AI.$$

Thus the set M is a ring with an identity element.

At this time, we should verify that the commutative law for multiplication and the cancellation law are not valid in M by giving counterexamples. Thus we have

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix},$$

but

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -2 & -2 \\ 2 & 2 \end{bmatrix},$$

so that the commutative law for multiplication does not hold. Also,

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 3 & 0 \end{bmatrix},$$

so that the cancellation law does not hold.

Exercises 2-2

1. Determine if the set of all integers is a ring under the operations of addition and multiplication.

2. Determine which of the following sets are rings under addition and multiplication:

   (a) the set of numbers of the form $a + b\sqrt{2}$, where a and b are integers;

   (b) the set of four fourth roots of unity, namely, $+1$, -1, i, and $-i$;

   (c) the set of numbers $a/2$, where a is an integer.

3. Determine if the set of all matrices of the form $\begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, with $a \in R$, forms a ring under addition and multiplication as defined for matrices.

4. Determine if the set of all matrices of the form $\begin{bmatrix} a & 0 \\ 0 & a^2 \end{bmatrix}$, with $a \in R$, forms a ring under addition and multiplication as defined for matrices.

2-3. The Uniqueness of the Multiplicative Inverse

Once again we turn our attention to the problem of matrix division. As we have seen, this problem arises when we seek to solve a matrix equation of the form

$$AX = C.$$

Let us look at a parallel equation concerning real numbers,

$$ax = c.$$

Each nonzero number a has a reciprocal $1/a$, which is often designated a^{-1}. Its defining property is $aa^{-1} = 1$. Since multiplication of real numbers is commutative, it follows that $a^{-1}a = 1$. Hence if a is a nonzero number, then there is a number b, called the multiplicative inverse of a, such that

$$ab = 1 = ba \qquad (b = a^{-1}).$$

Given an equation $ax = c$, where $a \neq 0$, the multiplicative inverse b enables us to find a solution for x; thus,

$$b(ax) = bc,$$

$$(ba)x = bc,$$

$$1x = bc,$$

$$x = bc.$$

Now our question concerning division by matrices can be put in another way. If $A \in M$, is there a $B \in M$ for which the equation

$$AB = I = BA$$

is satisfied? We shall employ the more suggestive notation A^{-1} for the inverse, so that our question can be restated: Is there an element $A^{-1} \in M$ for which the equation

$$AA^{-1} = I = A^{-1}A$$

is satisfied? Since we shall often be using this defining property, let us state it formally as a definition.

<u>Definition 2-3</u>. If $A \in M$, then an element A^{-1} of M is an <u>inverse</u> of A provided

$$AA^{-1} = I = A^{-1}A.$$

If there were an element B corresponding to each element $A \in M$ such that

$$AB = I = BA,$$

then we could solve all equations of the form

$$AX = C,$$

since we would have

$$B(AX) = BC,$$

$$(BA)X = BC,$$

$$IX = BC,$$

$$X = BC,$$

and clearly this value satisfies the original equation.

From the fact that there is a multiplicative inverse for every real number except zero, we might wrongly infer a parallel conclusion for matrices. As stated in Chapter 1, not all matrices have inverses. Our knowledge that 0 has no inverse suggests that the zero matrix $\underline{0}$ has no inverse. This is true, since we have

$$\underline{0}X = \underline{0}$$

for all $X \in M$, so that there cannot be any $X \in M$ such that

$$\underline{0}X = I.$$

But there is a more fundamental difficulty than this. Let us take the nonzero matrix

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

and try to solve the equation

$$AX = I, \qquad \text{for } X \in M.$$

If we let

$$X = \begin{bmatrix} p & q \\ r & s \end{bmatrix},$$

then we find that

$$AX = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} p & q \\ 0 & 0 \end{bmatrix}.$$

Hence, no matter what entries we take for X, we cannot have

$$AX = I$$

since the entry in the lower right-hand corner of AX is zero, and the entry in the lower right-hand corner of I is 1.

At this point, you might be thinking that no matrix has an inverse. Do not move too fast! Note that

$$I \cdot I = I = I \cdot I.$$

This means that I is its own inverse, just as 1 is its own inverse among the real numbers.

Also, let us note that

$$\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Thus the matrix

$$A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

has the inverse

$$A^{-1} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}.$$

Consequently, the equation

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} X = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

may be solved by multiplying both sides by A^{-1}, thus:

$$\begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} X = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} X = \begin{bmatrix} 1/2 & 1 \\ 3/2 & 2 \end{bmatrix},$$

$$X = \begin{bmatrix} 1/2 & 1 \\ 3/2 & 2 \end{bmatrix}.$$

This is a specific illustration of a general pattern. Let a be any nonzero number. Now

$$\begin{aligned} I &= 1I \\ &= aa^{-1}I \\ &= aa^{-1}(I)(I). \end{aligned}$$

Since the multiplication of real numbers and matrices is associative and commutative, it follows that for all real numbers a and b, and all 2×2 matrices X and Y, we have

$$abXY = (aX)(bY).$$

In particular, then,

$$I = (aI)(a^{-1}I).$$

Since $aa^{-1} = a^{-1}a$, we can also state that

$$I = (a^{-1}I)(aI).$$

This result enables us to enumerate a large number of matrices and their inverses. Thus, let $A = aI$; then $A^{-1} = a^{-1}I$. For example, if $a = 3$ then

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} 1/3 & 0 \\ 0 & 1/3 \end{bmatrix}.$$

If $a = 0.2$, then

$$A = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix} \quad \text{and} \quad A^{-1} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}.$$

At least we know that there are a great many matrices A with the property that there is a corresponding matrix B such that

$$AB = I = BA.$$

Before turning to the problem of finding those matrices that have inverses, let us show first that if a matrix has an inverse, it has only one inverse; that is, this inverse is unique. For instance, in the example directly above, we saw that

$$A^{-1} = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} \quad \text{if} \quad A = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.2 \end{bmatrix}.$$

We wish to show that there is no other inverse. Suppose that we have elements A, B, and C of M such that

$$AB = I = BA,$$

and

$$AC = I = CA;$$

that is, A has an inverse B, and A also has an inverse C. Multiply the first of these two equations on the left by C. Then

$$C(AB) = CI,$$

or

$$(CA)B = C,$$

since multiplication is associative and I is the unit matrix. But $CA = I$. Hence

$$IB = C,$$

or

$$B = C.$$

This result is so important that we call it a theorem and state it formally:

<u>Theorem 2-2</u>. If $A \in M$ and if there exists A^{-1}, $A^{-1} \in M$, such that

$$AA^{-1} = I = A^{-1}A,$$

then A^{-1} is unique; that is, there is no other solution X of the equations

$$AX = I = XA.$$

Now we can readily show that A is the inverse of A^{-1} if we know that A^{-1} is the inverse of A. This may seem a bit trivial, but it is important enough to state formally and prove.

<u>Theorem 2-3</u>. If $A \in M$ and if A has an inverse A^{-1}, then A^{-1} also has an inverse; namely, A is the inverse of A^{-1}.

<u>Proof</u>. Since A^{-1} is the inverse of A, this means, by definition, that

$$AA^{-1} = I = A^{-1}A.$$

However, the statement of equality can be given in reverse order:

$$A^{-1}A = I = AA^{-1}.$$

This, by definition, is the statement that A is the inverse of A^{-1}.

Exercises 2-3

1. Show that each of the following matrices does not have a multiplicative inverse:

   (a) $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$, (c) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} 0 & 0 \\ -3 & 0 \end{bmatrix}$.

2. Which of the following pairs of elements of M are inverses of one another?

   (a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$,

   (b) $\begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix}$ and $\begin{bmatrix} 3 & -1 \\ 2 & -1 \end{bmatrix}$,

   (c) $\begin{bmatrix} 2 & 4 \\ 6 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & -4 \\ -6 & 2 \end{bmatrix}$,

   (d) $\begin{bmatrix} -5 & 7 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$,

   (e) $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.

3. Use the argument in the text to show that, since

$$\begin{bmatrix} 1 & -2 \\ -3 & 6 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \underline{0},$$

   neither of the matrices in the product is invertible (has an inverse).

4. Show that if $a^2 + bc = 0$, then

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix}^2 = \underline{0},$$

and hence that in this case

$$\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$$

has no inverse.

5. Show that if $A \in M$, $B \in M$, $B \neq \underline{0}$, and $AB = \underline{0}$, then A cannot have an inverse. Can B have an inverse?

6. Show that if $A \in M$, and $A^2 - 4A = \underline{0}$, then either $A = 4I$ or A has no inverse. (Hint: Factor the left-hand side and note Exercise 5.)

7. Show that if $A \in M$, $B \in M$, $C \in M$, and $AB = I = CA$, then $B = C$.

8. Show by direct computation that the matrix

$$A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}$$

satisfies the equation

$$A^2 - 2A - 3I = \underline{0},$$

that is,

$$\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix}^2 - 2\begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} - 3\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

9. The matrices

$$\begin{bmatrix} -1 & 3 \\ 2 & -5 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 & 3 \\ 2 & 1 \end{bmatrix}$$

are inverses of one another. Are their squares also inverses? Their transposes?

10. Since

$$\begin{aligned} A^2 &= A \cdot A, \\ A^3 &= A \cdot A^2 = A^2 \cdot A, \\ A^4 &= A \cdot A^3 = A \cdot A^3, \end{aligned}$$

and so on, we can readily demonstrate that A^{n-1} is the inverse of A if $A^n = I$. Using this information, compute the inverse of each of the following matrices:

(a) $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$,

(b) $\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$,

(c) $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

11. Let

$$B = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

and compute B^2 and B^3 if $\theta = 120^\circ$.

12. If

$$A = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix},$$

verify that

$$A^2 - 2A + I = \underset{\sim}{0}.$$

Does the transpose of A also satisfy this same equation?

13. Prove that if $A \in M$, if p, q, and r are numbers, and if

$$pA^2 + qA + rI = \underset{\sim}{0}$$

with $r \neq 0$, then A has an inverse. (Hint: Subtract the "constant term" rI from both members of the equation and factor the remaining terms in the left member.)

14. Prove by direct substitution that if

$$X = \begin{bmatrix} p & q \\ r & s \end{bmatrix}$$

then

$$X^2 - (p + s)X + (ps - qr)I = \underline{0}.$$

Show that X has an inverse if and only if $ps - qr \neq 0$. (Hint: Use Exercise 13.)

15. Use the result of Exercise 14 to show that if $X^2 = 0$ then $ps - qr = 0$ and $p + s = 0$. (Perhaps you may have to consider several cases in the proof.)

2-4. The Inverse of a Matrix of Order 2

At this point, we have proved that the inverse of a 2×2 matrix, if it exists, is unique. Also, we know that there are some 2×2 matrices that have inverses and there are some that do not have inverses. But we have not yet developed any general methods of attacking the problem. Certainly our algebra will lack power unless general methods are developed. We are in a situation similar to that in which a student finds himself when he has not yet learned the quadratic formula or the general procedure for deriving it. He can find the roots of many quadratic equations by trial, but he has no means for solving all these equations.

It is our purpose now to develop a general method of determining the inverse of a 2×2 matrix when it exists. We shall begin with a matrix whose entries are specific numbers and then duplicate our procedure with a matrix whose entries are more general. To start, we shall consider the matrix

$$A = \begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}$$

and determine whether there is an inverse B such that $AB = I = BA$. If we let

$$B = \begin{bmatrix} p & q \\ r & s \end{bmatrix},$$

then

$$\begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

or

$$\begin{bmatrix} 3p - r & 3q - s \\ 5p - 2r & 5q - 2s \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If these two matrices are equal, the respective entries are equal. Thus we have four equations,

$$3p - r = 1, \quad (1) \qquad 3q - s = 0, \quad (3)$$

$$5p - 2r = 0, \quad (2) \qquad 5q - 2s = 1. \quad (4)$$

After multiplying Equation (1) by 2, we subtract Equation (2) from Equation (1) and obtain

$$p = 2.$$

By substituting this value of p in either Equation (1) or Equation (2), we obtain

$$r = 5.$$

Equations (3) and (4) can be solved similarly, yielding

$$q = -1 \quad \text{and} \quad s = -3.$$

Now if we substitute these values for p, q, r, and s, we obtain

$$B = \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix}.$$

To demonstrate that B is the inverse of A, we must show that $AB = I = BA$. This is easy:

$$\begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix}\begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}.$$

Using the notation for the inverse of a matrix introduced earlier, we may write

$$\begin{bmatrix} 3 & -1 \\ 5 & -2 \end{bmatrix}^{-1} = \begin{bmatrix} 2 & -1 \\ 5 & -3 \end{bmatrix}.$$

In our next step, we shall follow the same pattern as above; but now we shall use a general notation for our matrix A. Instead of having specific real numbers for entries, we let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

As before, we represent the inverse, if it exists, as

$$B = \begin{bmatrix} p & q \\ r & s \end{bmatrix}.$$

Assuming $AB = I$, we have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} p & q \\ r & s \end{bmatrix} = \begin{bmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This matrix equation may be written as four equations,

$$ap + br = 1, \quad (5) \qquad aq + bs = 0, \quad (7)$$

$$cp + dr = 0, \quad (6) \qquad cq + ds = 1. \quad (8)$$

Since we wish to find values for p, q, r, and s, in terms of the real numbers a, b, c, and d, we multiply Equation (5) by d, Equation (6) by b, and then subtract. We obtain

$$adp - bcp = d,$$

or

$$(ad - bc)p = d.$$

Repeating this process, using appropriate pairs of equations, we obtain

$$(ad - bc)q = -b, \qquad (ad - bc)r = -c, \qquad (ad - bc)s = a.$$

Should it happen that $ad - bc = 0$, then it follows from the four equations, above, that $a = b = c = d = 0$, so that $A = \underline{0}$.

We have seen in Section 2.3 that the zero matrix does not have an inverse.

Hence if $ad - bc = 0$ we have a contradiction of the assumption that the matrix A has an inverse B. In other words, if A has an inverse, then $ad - bc \neq 0$.

Temporarily, let us denote the number $ad - bc$ by h. Now if $h \neq 0$, we may write

$$p = \frac{d}{h}, \quad q = -\frac{b}{h}, \quad r = -\frac{c}{h}, \quad s = \frac{a}{h}.$$

Substituting these values appropriately in B, we obtain

$$B = \begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix} = \frac{1}{h}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

In order to show that this matrix is the inverse of A, we check:

$$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix} = \begin{bmatrix} \frac{ad-bc}{h} & \frac{-ab+ab}{h} \\ \frac{cd-cd}{h} & \frac{-bc+ad}{h} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

We must also make sure that $BA = I$, thus:

$$BA = \begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} \frac{ad-bc}{h} & \frac{bd-bd}{h} \\ \frac{-ac+ac}{h} & \frac{-bc+ad}{h} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I.$$

The fact that the relationship $BA = I$ follows from the relationship $AB = I$ is quite significant. While the definition of the inverse demands the existence and equality of what are called left and right inverses, we have shown that for 2×2 matrices the existence of one implies the existence of the other and that if they exist then they are, in fact, the same. Since the multiplication of matrices is not generally commutative, we might have expected otherwise.

We shall state our result formally as a theorem.

Theorem 2-4. If the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse, then $h = ad - bc \neq 0$ and

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix}.$$

Also, we state the converse of this result concerning h:

Theorem 2-5. If $h = ad - bc \neq 0$, then the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse, which is

$$\begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix}.$$

Proof. Direct multiplication shows that

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

You will note that Theorems 2.4 and 2.5 together state that the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ has an inverse if and only if $h \neq 0$. That is, the condition $h \neq 0$ is both necessary and sufficient for the matrix to have an inverse. You should remember the formula

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \begin{bmatrix} \frac{d}{h} & -\frac{b}{h} \\ -\frac{c}{h} & \frac{a}{h} \end{bmatrix}, \quad h = ad - bc \neq 0.$$

Exercises 2-4

1. For each of the following matrices, determine whether the inverse exists; if it does exist, find it:

   (a) $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$, (c) $\begin{bmatrix} -3 & 7 \\ 9 & 21 \end{bmatrix}$,

   (b) $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} 4 & 2 \\ 1 & 1 \end{bmatrix}$,

(e) $\begin{bmatrix} -2 & 0 \\ 3 & 4 \end{bmatrix}$, (g) $\begin{bmatrix} 2 & -6 \\ -1 & 3 \end{bmatrix}$.

(f) $\begin{bmatrix} 2 & a \\ 0 & -7 \end{bmatrix}$,

2. Each of the following matrices is actually a function in the sense that it depends on the value assigned to x, where $x \in R$. Determine those values of x for which the matrix has no inverse.

(a) $\begin{bmatrix} x^2 & 1 \\ 1 & x \end{bmatrix}$, (c) $\begin{bmatrix} x+2 & 0 \\ x^4 & x-1 \end{bmatrix}$,

(b) $\begin{bmatrix} x^3 & x \\ 0 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} x^2 & x-1 \\ 2 & 3 \end{bmatrix}$.

3. Show that each matrix of the form

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$$

has an inverse and find it. Show that the product of two such matrices (different values of θ) is again such a matrix. (Hint: Use the addition formulas from trigonometry.)

4. Show that if $A \in M$ then A has an inverse if and only if its transpose has an inverse. If A has an inverse show that

$$\text{transpose } (A^{-1}) = (\text{transpose } A)^{-1}.$$

5. Prove Theorem 2–3 by first computing A^{-1} by Theorem 2–4 and then using Theorem 2–5 to compute the inverse of A^{-1}.

6. Under the assumption that the element A of M has an inverse, show how to solve the equation $AX = B$, with B a 2×1 matrix. Apply this to solve the following equations:

(a) $2x + 3z = 9,$
$-x + 4z = 10;$

(b) $3x + z = 0,$
$-2x + z = 1;$

(c) $2y + 3w = 0,$
$-y + 4w = 0;$

(d) $3y + w = 1,$
$-2y + w = 0.$

2–5. The Determinant Function

We have seen that the criterion for the existence of an inverse for the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

involves the value of the expression $ad - bc$. If $ad - bc \neq 0$, the inverse does exist; if $ad - bc = 0$, the inverse does not exist. Each 2×2 matrix determines one real value for $ad - bc$. For example,

$$\text{if } A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \text{then } ad - bc = 1(1) - 0(0) = 1;$$

$$\text{if } A = \begin{bmatrix} 2 & 3 \\ 4 & 6 \end{bmatrix}, \qquad \text{then } ad - bc = 2(6) - 3(4) = 0;$$

$$\text{if } A = \begin{bmatrix} 0.5 & 3 \\ 4 & 0.6 \end{bmatrix}, \qquad \text{then } ad - bc = 0.5(0.6) - 3(4) = -11.7.$$

(Note that the second matrix does not have an inverse.) With each matrix of M there is thus associated one value, a real number determined by the entries. It is convenient to give a name to this number, the value of the expression $ad - bc$, which is associated with the matrix

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Definition 2–4. If

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then $\delta(X) = ad - bc$ is called the determinant of X.

Thus δ assigns to each member X of M a real number $\delta(X)$, read "delta of X." It is appropriate to regard this assignment or mapping as a function from the set M of 2×2 matrices

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

to the set R of real numbers

$$x = ad - bc.$$

We indicate this as follows:

$$\delta : \begin{bmatrix} a & b \\ c & d \end{bmatrix} \longrightarrow ad - bc,$$

that is,

$$\delta : X \longrightarrow \delta(X).$$

The function δ has interesting properties, some of which we shall demonstrate.

First let us compute the values $\delta(X)$ for a few products:

(a) If

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix},$$

then

$$\delta(A) = 3(2) - 2(1) = \ 4,$$

$$\delta(B) = 0(1) - 3(2) = -\ 6,$$

$$AB = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 11 \\ 4 & 5 \end{bmatrix},$$

$$\delta(AB) = 4(5) - 11(4) = -\ 24.$$

(b) If

$$A = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 8 & 2 \\ 3 & 1 \end{bmatrix},$$

then

$$\delta(A) = -\ 1(3) - 2(0) = -\ 3,$$

$$\delta(B) = 8(1) - 2(3) = 2,$$

$$AB = \begin{bmatrix} -1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 8 & 2 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 9 & 3 \end{bmatrix},$$

$$\delta(AB) = -2(3) - 0(9) = -6.$$

We might suspect that $\delta(AB) = \delta(A)\,\delta(B)$! This is true and we shall now prove it.

<u>Theorem 2-6</u>. If $A \in M$ and $B \in M$, then

$$\delta(AB) = \delta(A)\,\delta(B).$$

<u>Proof</u>. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad B = \begin{bmatrix} p & q \\ r & s \end{bmatrix};$$

then

$$AB = \begin{bmatrix} ap + br & aq + bs \\ cp + dr & cq + ds \end{bmatrix},$$

$$\begin{aligned} \delta(AB) &= (ap + br)(cq + ds) - (aq + bs)(cp + dr) \\ &= apcq + apds + brcq + brds \\ &\qquad - aqcp - aqdr - bscp - bsdr \\ &= apds + brcq - aqdr - bscp, \end{aligned} \tag{1}$$

$$\begin{aligned} \delta(A) &= ad - bc, \\ \delta(B) &= ps - qr, \\ \delta(A)\delta(B) &= (ad - bc)(ps - qr) \\ &= adps - adqr - bcps + bcqr. \end{aligned} \tag{2}$$

By rearranging the terms in expressions (1) and (2), we see that

$$\delta(AB) = \delta(A)\,\delta(B). \qquad \underline{q}.\underline{e}.\underline{d}.$$

Let us look at more examples; let

$$A = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 0 & 3 \\ 2 & 1 \end{bmatrix}.$$

In Section 2–4 we learned that if

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

then

$$X^{-1} = \frac{1}{\delta(X)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Hence

$$A^{-1} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{3}{4} \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} -\frac{1}{6} & \frac{1}{2} \\ \frac{1}{3} & 0 \end{bmatrix}.$$

Further,

$$\delta(A) = 3(2) - 2(1) = 4,$$

$$\delta(B) = 0(1) - 3(2) = -6,$$

$$\delta(A^{-1}) = \frac{1}{2}\,\frac{3}{4} - \left(-\frac{1}{2}\right)\left(-\frac{1}{4}\right) = \frac{1}{4},$$

$$\delta(B^{-1}) = -\frac{1}{6}\,(0) - \frac{1}{2}\left(\frac{1}{3}\right) = -\frac{1}{6}.$$

Theorem 2–7. If A is a 2×2 matrix, and A has a multiplicative inverse, then

$$\delta(A^{-1}) = \frac{1}{\delta(A)}.$$

Proof. We have

$$AA^{-1} = I,$$

$$\delta(AA^{-1}) = \delta(I).$$

But by computing $\delta(I)$, we see that

$$\delta(I) = 1(1) - 0(0) = 1,$$

whence

$$\delta(AA^{-1}) = 1,$$

so that by Theorem 2–6,

$$\delta(A)\ \delta(A^{-1}) = 1,$$

or

$$\delta(A^{-1}) = \frac{1}{\delta(A)} .$$

We shall now prove quite a significant theorem, which will give us the power to decide when a product AB has an inverse and what the inverse is.

<u>Theorem 2–8</u>. If A and B are 2×2 matrices, then AB has an inverse if and only if A and B both have inverses. Further, if these matrices have inverses, then

$$(AB)^{-1} = B^{-1}A^{-1}.$$

<u>Proof</u>. Since

$$\delta(A)\ \delta(B) = \delta(AB),$$

it follows that

$$\delta(AB) \neq 0$$

if and only if

$$\delta(A) \neq 0 \quad \text{and} \quad \delta(B) \neq 0.$$

Then by Theorems 2.4 and 2.5 we see that AB has an inverse if and only if A and B both have inverses.

To complete the proof of Theorem 2-8, we assume that A and B have inverses A^{-1} and B^{-1}, respectively, and then have only to exhibit a matrix X such that

$$ABX = I = XAB.$$

Let

$$X = B^{-1}A^{-1}.$$

Then

$$\begin{aligned} ABX &= ABB^{-1}A^{-1} \\ &= A(BB^{-1})A^{-1} \\ &= A(I)A^{-1} \\ &= AA^{-1} \\ &= I. \end{aligned}$$

Hence $B^{-1}A^{-1}$ is a right inverse. Similarly, we can show that

$$B^{-1}A^{-1}AB = I.$$

Thus $B^{-1}A^{-1}$ is the inverse of AB. This completes the proof.

For example, let

$$A = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}.$$

Then

$$A^{-1} = \begin{bmatrix} -\frac{3}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad B^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}.$$

Now

$$AB = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 7 & 19 \end{bmatrix},$$

whence

$$(AB)^{-1} = \begin{bmatrix} 1 & 3 \\ 7 & 19 \end{bmatrix}^{-1} = \begin{bmatrix} -\frac{19}{2} & \frac{3}{2} \\ \frac{7}{2} & -\frac{1}{2} \end{bmatrix}.$$

But also,

$$B^{-1} A^{-1} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} -\frac{3}{2} & \frac{1}{2} \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -\frac{19}{2} & \frac{3}{2} \\ \frac{7}{2} & -\frac{1}{2} \end{bmatrix}.$$

Thus, for our example we have $(AB)^{-1} = B^{-1} A^{-1}$.

There are many other theorems that can be developed from the concept of a determinant function. A few of these will be included in the exercises that follow. It is worth noting, though we shall not prove it, that there is a determinant function associated with the other sets of square matrices, that is, with those of order 1, 3, 4,..., and that similar theorems hold for them. Specifically, there is a determinant function associated with each square matrix, and its nonvanishing is a necessary and sufficient condition for the matrix to have an inverse.

Exercises 2-5

1. Verify Theorem 2-6 for the matrices

(a) $A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 4 & 3 \end{bmatrix};$

(b) $A = \begin{bmatrix} t^2 & 1 \\ -1 & t \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix};$

(c) $A = \begin{bmatrix} x & x^2 \\ x^3 & x^4 \end{bmatrix}, \quad B = \begin{bmatrix} x & -x \\ 3 & 4 \end{bmatrix}.$

2. Show that

$$\delta(tA) = t^2\delta(A)$$

for any $A \in M$ and any $t \in R$.

3. For A and t as in Exercise 2, show that $\delta(A)$ is the constant term in the polynomial $\delta(A - tI)$.

4. If

$$A = \begin{bmatrix} x & 1 \\ x^2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 1 \\ -5 & -2 \end{bmatrix},$$

find $\delta(A)$ and $\delta(B^{-1}AB)$ and show that they are equal.

5. Show that if $A \in M$, $B \in M$, and B is invertible, then

$$\delta(B^{-1}AB) = \delta(A).$$

6. Show that if $A \in M$ and A^t is the transpose of A, then

$$\delta(A) = \delta(A^t),$$

and conclude that

$$\delta(AA^t) \geq 0$$

for any $A \in M$.

7. The expression $\delta(A - tI)$ is a polynomial in t. For each of the following matrices A, expand this polynomial and find its zeros:

(a) $\begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix}$,

(b) $\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$,

(c) $\begin{bmatrix} t & 0 \\ -t & 1 \end{bmatrix}$,

(d) $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$.

8. Let

$$A = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

and expand the polynomial $\delta(AA^t - xI)$. Is this the same as the polynomial $\delta(A^tA - xI)$? Are these two polynomials the same for every matrix $A \in M$?

2-6. The Group of Invertible Matrices

In this chapter, we have been restricting our attention to the set M of 2×2 matrices. This set is, itself, a subset of the set of all rectangular matrices. Now this set M can be separated into interesting subsets. In the preceding section, we have divided M into two complementary subsets, the set of 2×2 matrices that do not have inverses and the set of 2×2 matrices that do have inverses. In this section, we shall confine our attention principally to the set of invertible 2×2 matrices. It is convenient to denote this set by the symbol M_i.

Let us summarize certain facts about the set M_i of invertible matrices:

(a) If $A \in M_i$, and $B \in M_i$, then $AB \in M_i$.

(b) If $A \in M_i$, $B \in M_i$, and $C \in M_i$, then $A(BC) = (AB)C$.

(c) In M_i, there is an identity element I, that is, an element I such that $AI = A = IA$ for each $A \in M_i$.

(d) If $A \in M_i$, then A has an inverse $A^{-1} \in M_i$, that is, an element A^{-1} such that $AA^{-1} = I = A^{-1}A$.

Not only does the set M_i satisfy each of these conditions, but there are many subsets of M_i that satisfy conditions analogous to them. Any set S of matrices that satisfies conditions (a), (b), (c), and (d), with S in place of M_i, will be called a group. The concept of a group is fundamental and extremely important in mathematics. More generally, any set of elements $A,B,C,\ldots,$ not necessarily matrices, satisfying the foregoing properties relative to an operation (not necessarily matrix multiplication) is defined to be a group. You will note that only one operation is involved in the group properties. Although we shall later introduce a few examples of the more general concept, for the moment let us consider some examples of groups of invertible matrices.

The smallest set of invertible matrices that constitutes a group is the set whose one element is the unit matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Since $(I)(I) = I$, condition (a) is satisfied; and condition (b) is automatically fulfilled by any set of square matrices. Certainly I is a member of the set, so that condition (c) is satisfied. For condition (d), there must be an inverse for every element;

this condition is satisfied since in our present set the only element I is its own inverse.

All quite simple, isn't it? Was it obvious?

Another set that constitutes a group is the set $\{I, -I\}$. Again conditions (b) and (c) obviously are satisfied. Since

$$(I)(-I) = (-I)(I) = -I$$

and

$$(I)(I) = (-I)(-I) = I,$$

conditions (a) and (d) also are satisfied.

The third set that we shall show to be a group is a bit more significant. The set of all elements $A \in M$ such that $\delta(A) = 1$ is a group. The proof is a bit more difficult, and we must check carefully each one of the defining properties. To provide a language that will be helpful, let us denote this set by W, thus:

$$W = \{A\colon A \in M \text{ and } \delta(A) = 1\}.$$

Let us verify first that condition (a) is satisfied. If $A \in W$ and $B \in W$, then $\delta(A) = 1$ and $\delta(B) = 1$. Since $\delta(AB) = \delta(A)\,\delta(B)$ by Theorem 2–6, we have

$$\delta(AB) = \delta(A)\,\delta(B) = (1)(1) = 1,$$

and thus $AB \in W$.

Property (b) holds automatically.

For property (c), since $\delta(I) = 1$, it is clear that $I \in W$.

To demonstrate that condition (d) is satisfied, we must show not only that each element of W has an inverse but also that the inverse is an element of W. Now, if $A \in W$, then $\delta(A) = 1$. Since $\delta(A) \neq 0$, A has an inverse A^{-1}, by Theorem 2–5. By Theorem 2–7,

$$\delta(A^{-1}) = \frac{1}{\delta(A)} = \frac{1}{1} = 1.$$

Hence $A^{-1} \in W$, and we have now demonstrated that W is a group.

In our last example, we shall discuss all matrices of the form

$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix} \qquad (x,y = \text{real numbers})$$

and denote this set by Z, $Z \subset M$. (Read $Z \subset M$ as, "The set Z is contained in the set M.")

We observe first that the product of any two members of this set Z is also a member of Z. We have, indeed,

$$\begin{bmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{bmatrix} \begin{bmatrix} x_2 & y_2 \\ -y_2 & x_2 \end{bmatrix} = \begin{bmatrix} x_1x_2 - y_1y_2 & x_1y_2 + y_1x_2 \\ -(x_2y_1 + x_1y_2) & -y_1y_2 + x_1x_2 \end{bmatrix}.$$

Condition (b) is automatically satisfied; and I is a member of Z, with $x = 1$, $y = 0$, so that condition (c) is satisfied.

In considering condition (d), we run into trouble. The zero matrix is an element of this set, but the zero matrix does not have an inverse. The set of all matrices of the form

$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

does <u>not</u> form a group.

Although the set Z does not satisfy the four conditions, a subset Z_1 of Z, defined by

$$Z_1 = \{A\colon A \in Z \text{ and } \delta(A) = 1\},$$

does satisfy the conditions and is therefore a group.

The demonstration is easy. Let $A \in Z_1$ and $B \in Z_1$. We know that $AB \in Z$, as already shown; and, since $\delta(A) = 1$ and $\delta(B) = 1$, we know that $\delta(AB) = 1$. Hence $AB \in Z_1$, and therefore condition (a) is satisfied. Obviously, condition (b) also is satisfied. We know that $I \in Z$ and that $\delta(I) = 1$; hence, $I \in Z_1$, so that condition (c) is satisfied. Finally, for condition (d), we must show that if $A \in Z_1$ then there is an inverse A^{-1} such that $A^{-1} \in Z_1$. We follow the pattern set in an earlier illustration. Since $\delta(A) = 1$, there is an inverse. Then, using the fact that $\delta(A^{-1}) = 1/\delta(A)$, we proceed to show that $\delta(A^{-1}) = 1$, which means that $A^{-1} \in Z_1$. Having demonstrated that the four groups postulates are satisfied, we conclude that we have a group.

Before considering the more general concept of a group, we shall demonstrate a fruitful correspondence between the elements of Z_1 and the points on a unit circle, which will let us examine a geometric interpretation of Z_1.

If

$$A = \begin{bmatrix} x & y \\ -y & x \end{bmatrix}$$

is any element of Z_1, we have $\delta(A) = 1$; that is, we have

$$x^2 + y^2 = 1.$$

Now, if we let x and y be coordinates of a point (x,y), we are able to establish a one-to-one correspondence between the elements of Z_1 and the points on a unit circle:

$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix} \longleftrightarrow (x, y).$$

The set of matrices is thus mapped onto the set of points in such a way that to each matrix there corresponds exactly one point of the set, and to each point of the set there corresponds exactly one matrix.

The point (x, y) is on the circle of radius 1 with center at the origin, as shown in Figure 2-1.

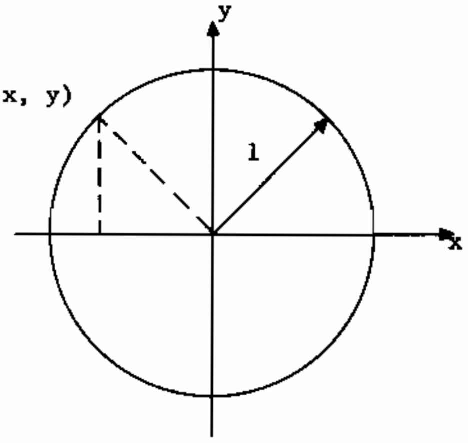


Figure 2-1. The unit circle.

Let us call this circle <u>the unit circle</u> and denote it by Q.

Thus

$$Q = \left\{(x, y): x \in R, \quad y \in R, \quad \text{and} \quad x^2 + y^2 = 1\right\}.$$

A geometrical meaning can be assigned to the inverse of any element of Z_1. If

$$A = \begin{bmatrix} x & y \\ -y & x \end{bmatrix},$$

then we can readily compute A^{-1} by Theorem 2-5, to obtain

$$A^{-1} = \begin{bmatrix} x & -y \\ y & x \end{bmatrix}.$$

Recalling the one-to-one correspondence between the matrices of Z_1 and the points of Q (the unit circle),

$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix} \longleftrightarrow (x, y),$$

we see, by examining Figure 2-2, that the correspondent of A^{-1} is the reflection in the x axis of the correspondent of A.

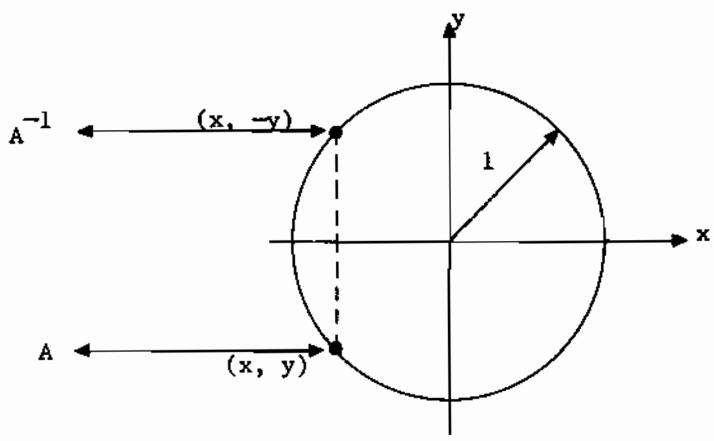


Figure 2-2. Geometric representation of inverse matrices A and $A^{-1} \in Z_1$.

Although a full discussion of the general notion of a group would be too extensive for this book, a few words are in order. The definition of an abstract group is stated somewhat differently from the definition given on page 85, although the abstract definition implies the latter.

Definition 2-5. A group is a set G of elements, a,b,c,..., on which a binary operation o (read "circle") is defined, such that the following properties are satisfied:

(a) If $a \in G$ and $b \in G$, then $a \circ b \in G$. (Closure property.)

(b) If $a \in G$, $b \in G$, and $c \in G$, then $a \circ (b \circ c) = (a \circ b) \circ c$. (Associative property.)

(c) There exists a unique element i, $i \in G$, such that $i \circ a = a = a \circ i$ for all $a \in G$. (Identity property.)

(d) For each $a \in G$, there exists an element a^{-1}, $a^{-1} \in G$, such that $a^{-1} \circ a = i = a \circ a^{-1}$. (Inverse property.)

If, in addition, the following condition is fulfilled, the group is said to be commutative or abelian:

(e) For each $a \in G$ and each $b \in G$, $a \circ b = b \circ a$. (Commutative property.)

Although the operations we are most concerned with in mathematics are addition and multiplication, we are not restricted to these in the foregoing abstract definition. For instance, a very helpful exercise, not only for understanding the notion of a group but also for comprehending a finite number system, is the addition associated with a clock face; see Figure 2-3. This furnishes us with a group. The set of elements is 1,2,...,12. The operation is clockwise

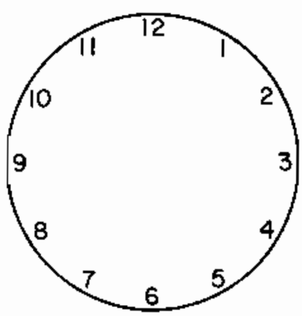


Figure 2-3. A clock face. The addition associated with it gives us a group.

addition of hours. Each defining property of an abstract group is satisfied, as we shall now illustrate. First, the "sum" of any two elements is another element. For example, we have

$$\begin{aligned} 1 + 6 &= 7, \\ 8 + 4 &= 12, \\ 11 + 2 &= 1, \\ 3 + 12 &= 3. \end{aligned}$$

Secondly since, for example,

$$(8 + 2) + 3 = 1 \quad \text{and} \quad 8 + (2 + 3) = 1,$$

we see that the associative property holds. Thirdly, a full clock rotation, an advance of 12 hours, gives the same time, so that 12 is our unique identity element; thus,

$$12 + 2 = 2 = 2 + 12.$$

Finally, to each of the elements, 1,2,...,12, there corresponds a number we can "add" to obtain 12. Thus

$$\begin{aligned} 4 + 8 &= 12 = 8 + 4, \\ 10 + 2 &= 12 = 2 + 10, \\ 12 + 12 &= 12 = 12 + 12. \end{aligned}$$

One of the most elegant examples of a group consists of the three cube roots of 1, namely

$$1, \quad \frac{-1 + i\sqrt{3}}{2}, \quad \frac{-1 - i\sqrt{3}}{2},$$

under multiplication. The demonstration is left to the student as an exercise.

Interestingly enough, although the integers are the most commonly used system that has a group structure (under the operation of addition), they were not the first to have their group structure analyzed. The first groups to be studied extensively were finite groups such as the two examples given above. These groups were found during a study of the theory of equations by Evariste

Galois (1811-1832), to whom is credited the origin of the systematic study of Group Theory. Unfortunately, Galois was killed in a duel at the age of 21, immediately after recording some of his most notable theorems.

Exercises 2-6

1. Determine whether the following sets are groups under multiplication:

(a) $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$;

(b) I, $-I$, K, $-K$,

where

$$K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

2. Show that the set of all elements of M of the form

$$\begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix}, \text{ where } t \in R \text{ and } t \neq 0,$$

constitutes a group under multiplication.

3. Show that the set of all elements of M of the form

$$\begin{bmatrix} t & s \\ s & t \end{bmatrix}, \text{ where } t \in R, \ s \in R, \text{ and } t^2 - s^2 = 1,$$

constitutes a group under multiplication.

4. If

$$A = \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{bmatrix},$$

show that the set

$$\{A, A^2, A^3\}$$

is a group under multiplication. Plot the corresponding points in the plane.

5. Let

$$T = \begin{bmatrix} 1 & -2 \\ 6 & 1 \end{bmatrix} \quad \text{and} \quad K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Show that the set

$$\{TIT^{-1},\ T(-I)T^{-1},\ TKT^{-1},\ T(-K)T^{-1}\}$$

is a group under multiplication. Is this true if T is <u>any</u> invertible matrix?

6. Show that the set of all elements of the form

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}, \quad \text{with} \quad a \in R, \quad b \in R, \quad \text{and} \quad ab = 1,$$

is a group under multiplication. If you plot all of the points (a,b), with a and b as above, what sort of a curve do you get?

7. Let

$$K = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

and let H be the set of all matrices of the form

$$xI + yK, \quad \text{with} \quad x \in R \quad \text{and} \quad y \in R.$$

Prove the following:

(a) The product of two elements of H is also an element of H.

(b) The element $xI + yK$ is invertible if and only if

$$x^2 - y^2 \neq 0.$$

(c) The set of all elements $xI + yK$ with $x^2 - y^2 = 1$ is a group under multiplication.

8. If a set G of 2×2 matrices is a group under multiplication, show that each of the following sets are groups under multiplication:

   (a) $\{A^t: A \in G\}$, where A^t = transpose of A;

   (b) $\{B^{-1}AB: A \in G\}$, where B is a fixed invertible element of M.

9. If a set G of 2×2 matrices is a group under multiplication, show that

   (a) $G = \{A^{-1}: A \in G\}$,

   (b) $G = \{BA: A \in G\}$, where B is any fixed element of G.

10. Using the definition of an abstract group, demonstrate whether or not each of the following sets under the indicated operation is a group:

    (a) the set of odd integers under addition;

    (b) the set R^+ of positive real numbers under multiplication;

    (c) the set of the four fourth roots of 1, $\{1, -1, i, -i\}$, under multiplication;

    (d) the set of all integers of the form $3m$, where m is an integer, under addition.

11. By proper application of the four defining postulates of an abstract group, prove that if a, b, and c are elements in a group and $a \circ b = a \circ c$, then $b = c$.

2-7. An Isomorphism between Complex Numbers and Matrices

It is true that very many different kinds of algebraic systems can be expressed in terms of special collections of matrices. Many theorems of this nature have been proved in modern higher algebra. Without attempting any such proof, we shall aim in the present section to demonstrate how the system of complex numbers can be expressed in terms of matrices.

In the preceding section, several subsets of the set of all 2×2 matrices were displayed. In particular, the set Z of all matrices of the form

$$\begin{bmatrix} x & y \\ -y & x \end{bmatrix}, \quad x \in R \text{ and } y \in R,$$

was considered. We shall exhibit a one-to-one correspondence between the set of all complex numbers, which we denote by C, and the set Z. This one-to-one correspondence would not be particularly significant if it did not preserve algebraic properties — that is, if the sum of two complex numbers did not correspond to the sum of the corresponding two matrices and the product of two complex numbers did not correspond to the product of the corresponding two matrices. There are other algebraic properties that are preserved in this sense.

Usually a complex number is expressed in the form

$$x + yi,$$

where $i = \sqrt{-1}$, $x \in R$, and $y \in R$. Thus, if c is an element of C, the set of all complex numbers, we may write

$$c = x(1) + y(i).$$

The numeral 1 is introduced in order to make the correspondence more apparent. In order to exhibit an element of Z in similar form, we must introduce the special matrix

$$J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Note that

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = -1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -I;$$

thus

$$J^2 = -I.$$

The matrix J corresponds to the number i, which satisfies the analogous equation

$$i^2 = -1.$$

This enables us to verify that

$$
\begin{aligned}
xI + yJ &= x\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + y\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} + \begin{bmatrix} 0 & y \\ -y & 0 \end{bmatrix} \\
&= \begin{bmatrix} x & y \\ -y & x \end{bmatrix},
\end{aligned}
$$

which indicates that any element of Z may be written in the form

$$xI + yJ.$$

For example, we have

$$
\begin{aligned}
2I + 3J &= 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 3\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 3 \\ -3 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix},
\end{aligned}
$$

and

$$
\begin{aligned}
0I + 5J &= 0\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + 5\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 5 \\ -5 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 5 \\ -5 & 0 \end{bmatrix}.
\end{aligned}
$$

Now we can establish a correspondence between C, the set of complex numbers, and Z, the set of matrices:

$$x1 + yi \longleftrightarrow xI + yJ.$$

Since each element of C is matched with one element of Z, and each element of Z is matched with one element of C, we call the correspondence one-to-one. Several special correspondences are notable:

$$0 = 0 \cdot 1 + 0 \cdot i \longleftrightarrow 0 \cdot I + 0 \cdot J = \underline{O}$$

$$1 = 1 \cdot 1 + 0 \cdot i \longleftrightarrow 1 \cdot I + 0 \cdot J = I$$

$$i = 0 \cdot 1 + 1 \cdot i \longleftrightarrow 0 \cdot I + 1 \cdot J = J$$

As stated earlier, it is interesting that there is a correspondence between the complex numbers and 2×2 matrices, but the correspondence is not particularly significant unless the one-to-one matching is preserved in the operations, especially in addition and multiplication. We shall now follow the correspondence in these operations and demonstrate that the one-to-one property is preserved under the operations.

When two complex numbers are added, the real components are added, and the imaginary components are added. Also, remember that the multiplication of a matrix by a number is distributive; thus, for $a \in R$, $b \in R$, and $A \in M$, we have

$$(a + b)A = aA + bA.$$

Hence we are able to show our one-to-one correspondence under addition:

$$\begin{array}{lcr} c_1 + c_2 & & Z_1 + Z_2 = \\ = (x_1 + iy_1) + (x_2 + iy_2) & & (x_1 I + y_1 J) + (x_2 I + y_2 J) = \\ = (x_1 + x_2) + (y_1 + y_2)i & \longleftrightarrow & (x_1 + x_2)I + (y_1 + y_2)J. \end{array}$$

For example, we have

$$\begin{array}{lcr} (2 - 3i) + (4 + 1i) & & (2I - 3J) + (4I + 1J) = \\ = 6 - 2i & \longleftrightarrow & 6I - 2J. \end{array}$$

and

$$\begin{array}{lcr} (3 - 2i) + (2 + 0i) & & (3I - 2J) + (2I + 0J) = \\ = 5 - 2i & \longleftrightarrow & 5I - 2J. \end{array}$$

Before demonstrating that the correspondence is preserved under multiplication, let us review for a moment. An example will suffice:

$$(2 + 4i)(3 - 2i) = 6 - 4i + 12i - 8i^2$$
$$= 6 - 4i + 12i - 8(-1)$$
$$= 6 + 8(1) + (-4 + 12)i$$
$$= 14 + 8i;$$

$$(2I + 4J)(3I - 2J) = 6I^2 - 4IJ + 12JI - 8J^2$$
$$= 6I \ - 4J \ + 12J \ - 8(-I)$$
$$= 6I \ + 8I \ + (-4 + 12)J$$
$$= 14I + 8J.$$

Generally, for multiplication, we have

$$\begin{array}{lll} c_1c_2 & & Z_1Z_2 \ = \\ = (x_1 + y_1i)(x_2 + y_2i) & & (x_1I + y_1J)(x_2I + y_2J) \ = \\ = x_1x_2 + y_1y_2i^2 + x_1y_2i + y_1x_2i & & x_1x_2I^2 + y_1y_2J^2 + x_1y_2IJ + y_1x_2JI \ = \\ = (x_1x_2 - y_1y_2) + (x_1y_2 + x_2y_1)i & \longleftrightarrow & (x_1x_2 - y_1y_2)I + (x_1y_2 + x_2y_1)J. \end{array}$$

If we represent a complex number

$$a + bi$$

as a matrix,

$$a + bi \longleftrightarrow \begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

we do have a significant correspondence! Not only is there a one-to-one correspondence between the elements of the two sets, but also the correspondence is one-to-one under the operations of addition and multiplication.

The additive and multiplicative identity elements are, respectively,

$$0 = 0 + 0i \longleftrightarrow \begin{bmatrix} 0 & 0 \\ -0 & 0 \end{bmatrix} = \underline{0}$$

and

$$1 = 1 + 0i \longleftrightarrow \begin{bmatrix} 1 & 0 \\ -0 & 1 \end{bmatrix} = I;$$

and for the additive inverse of

$$a + bi \longleftrightarrow \begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

we have

$$-a - bi \longleftrightarrow \begin{bmatrix} -a & -b \\ b & -a \end{bmatrix}.$$

Let us now examine how the multiplicative inverses, or reciprocals, can be matched. We have seen that any member of the set of 2×2 matrices has a multiplicative inverse if and only if for it the determinant function does not equal zero. Accordingly, if $A \in Z$ then there exists A^{-1} if and only if $x^2 + y^2 \neq 0$, since $\delta(A) = x^2 + y^2$ for $A = xI + yJ$. Now we know that any complex number has a multiplicative inverse, or reciprocal, if and only if the complex number is not zero. That is, if $c = x + yi$, then there exists a multiplicative inverse of c if and only if $x + yi \neq 0$, which means that x and y are not both 0. This is equivalent to saying that $x^2 + y^2 \neq 0$, since $x \in R$ and $y \in R$. For multiplicative inverses, if

$$x^2 + y^2 \neq 0,$$

our correspondence yields

$$c_1 = x + yi \qquad \longleftrightarrow \qquad xI + yJ = Z_1$$

$$\frac{1}{c_1} = \frac{1}{x^2 + y^2}(x - yi) \longleftrightarrow \frac{1}{x^2 + y^2}(xI - yJ) = Z_1^{-1}.$$

It is now clear that the correspondence between C, the set of complex numbers, and Z, a subset of all 2×2 matrices,

$$x + yi \longleftrightarrow xI + yJ,$$

is preserved under the algebraic operations. All of this may be summed up by saying that C and Z have identical algebraic structures. Another way of expressing this is to say that C and Z are <u>isomorphic</u>. This word is derived from two Greek words and means "of the same form." Two number systems are isomorphic if, first, there is a mapping of one onto the other that is a one-to-one correspondence and, secondly, the mapping preserves sums and products. If two number systems are isomorphic, their structures are the same; it is only their terminology that is different. The world is heavy with examples of isomorphisms, some of them trivial and some quite the opposite. One of the simplest is the isomorphism between the counting numbers and the positive integers, a subset of the integers; another is that between the real numbers and the subset $a + 0i$ of all complex numbers. (We should quickly guess that there is an isomorphism between real numbers a and the set of all matrices of the form $aI + 0J$!)

An example of an isomorphism that is more difficult to understand is that between real numbers and residue classes of polynomials. We won't try to explain this one; but there is one more fundamental concept that can be introduced here, as follows.

We have stated that the real numbers are isomorphic to a subset of the complex numbers. We speak of the algebra of the real numbers as being <u>embedded</u> in the algebra of complex numbers. In this sense, we can say that the algebra of complex numbers is embedded in the algebra of 2×2 matrices. Also, we can speak of the complex numbers as being "richer" than the real numbers, or of the 2×2 matrices as being richer than the complex numbers. The existence of complex numbers gives us solutions to equations such as

$$x^2 + 1 = 0,$$

which have no solution in the domain of real numbers. It is of course clear that Z is a proper subset of M, that is, $Z \subset M$ and $Z \neq M$. Here is a simple example to illustrate the statement that M is "richer" than Z: The equation

$$X^2 - I = \underline{0}$$

has for solution any matrix

$$X = \begin{bmatrix} 0 & t \\ 1/t & 0 \end{bmatrix}, \quad t \in R \text{ and } t \neq 0,$$

as may be seen quickly by easy computation; and there are still other solutions. On the other hand, the equation

$$x^2 - 1 = 0$$

has exactly two solutions among the complex numbers, namely $c = 1$ and $c = -1$.

Exercises 2-7

1. Using the following values, show the correspondence under addition and multiplication between complex numbers of the form $x + yi$ and matrices of the form $xI + yJ$:

   (a) $x_1 = 1$, $y_1 = -1$, $x_2 = 0$, and $y_2 = -2$;

   (b) $x_1 = 3$, $y_1 = -4$, $x_2 = 1$, and $y_2 = 1$;

   (c) $x_1 = 0$, $y_1 = -5$, $x_2 = 3$, and $y_2 = 4$.

2. Carry through, in parallel columns as in the text, the necessary computations to establish an isomorphism between R and the set

$$N = \left\{ \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} : x \in R \right\}$$

   by means of the correspondence

$$x \longleftrightarrow \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix}.$$

3. In the preceding exercise, an isomorphism between R and the sets of matrices

$$\left\{ \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} : x \in R \right\}$$

   was considered. Define a function f by

$$f(x) = \begin{bmatrix} x & 0 \\ 0 & 0 \end{bmatrix}.$$

Determine which of the following statements are correct:

(a) $f(x + y) = f(x) + f(y)$,

(b) $f(xy) = f(x)\,f(y)$,

(c) $f(0) = \underline{0}$,

(d) $f(1) = I$,

(e) $f\left(\frac{1}{x}\right) = (f(x))^{-1}$, $x \neq 0$.

4. Is the set G of matrices

$$\begin{bmatrix} a & b \\ -b & a \end{bmatrix},$$

with a and b rational and $a^2 + b^2 = 1$, a group under multiplication?

2-8. Algebras

The concepts of group, ring, and field are of frequent occurrence in modern algebra. The study of these systems is a study of the structures or patterns that are the framework on which algebraic operations are dependent. In this chapter, we have attempted to demonstrate how these same concepts describe the structure of the set of 2×2 matrices, which is a subset of the set of all rectangular matrices.

Not only have we introduced these embracing concepts, but we have exhibited the "algebra" of the sets. "Algebra" is a generic word that is frequently used in a loose sense. By technical definition, an algebra is a system that has two binary operations, called "addition" and "multiplication," and also has "multiplication by a number," that make it both a ring and a vector space.

Vector spaces will be discussed in Chapter 4, and we shall see then that the set of 2×2 matrices constitutes a vector space under matrix addition and multiplication by a number. Thus the 2×2 matrices form an algebra.

As you yourself might conclude at this time, this algebra is only one of many possible algebras. Some of these algebras are duplicates of one another in the sense that the basic structure of one is the same as the basic structure of another. Superficially, they seem different because of the terminology. When they have the same structure, two algebras are called isomorphic.

Chapter 3

MATRICES AND LINEAR SYSTEMS

3-1. Equivalent Systems

In this chapter, we shall demonstrate the use of matrices in the solution of systems of linear equations. We shall first analyze some of our present algebraic techniques for solving these systems, and then show how the same techniques can be duplicated in terms of matrix operations.

Let us begin by looking at a system of three linear equations:

$$\text{I} \qquad \begin{aligned} x - y + z &= -2, & (1)\\ x - 2y - 2z &= -1, & (2)\\ 2x + y + 3z &= 1. & (3) \end{aligned}$$

In our first step toward finding the solution set of this system, we proceed as follows: Multiply Equation (1) by 1 to obtain Equation (1'); multiply Equation (1) by -1 and add the new equation to Equation (2) to obtain Equation (2'); multiply Equation (1) by -2 and add the new equation to Equation (3) to obtain Equation (3'). This gives the following system:

$$\text{II} \qquad \begin{aligned} x - y + z &= -2, & (1')\\ 0 - y - 3z &= 1, & (2')\\ 0 + 3y + z &= 5. & (3') \end{aligned}$$

Before continuing, we note that what we have done is reversible. In fact, we can obtain System I from System II as follows: Multiply Equation (1') by 1 to obtain Equation (1); add Equation (1') to Equation (2') to obtain Equation (2); multiply Equation (1') by 2 and add to Equation (3') to obtain Equation (3).

Our second step is similar to the first: Retain Equation (1') as Equation (1"); multiply Equation (2') by -1 to obtain Equation (2"); multiply Equation (2') by 3 and add the new equation to Equation (3') to obtain Equation (3").
This gives

$$\text{III}\qquad \begin{aligned} x - y + z &= -2, & (1'')\\ 0 + y + 3z &= -1, & (2'')\\ 0 + 0 - 8z &= 8. & (3'') \end{aligned}$$

Our third step reverses the direction: Multiply Equation (3'') by $-1/8$ to obtain Equation (3'''); multiply Equation (3'') by $3/8$ and add to Equation (2'') to obtain Equation (2'''); multiply Equation (3'') by $1/8$ and add to Equation (1'') to obtain Equation (1'''). We thus get

$$\text{IV}\qquad \begin{aligned} x - y + 0 &= -1, & (1''')\\ 0 + y + 0 &= 2, & (2''')\\ 0 + 0 + z &= -1. & (3''') \end{aligned}$$

Now, by retaining the second and third equations, and adding the second equation to the first, we obtain

$$\text{V}\qquad \begin{aligned} x + 0 + 0 &= 1,\\ 0 + y + 0 &= 2,\\ 0 + 0 + z &= -1, \end{aligned}$$

or, in a more familiar form,

$$\begin{aligned} x &= 1,\\ y &= 2,\\ z &= -1. \end{aligned}$$

In the foregoing procedure, we obtain system II from system I, III from II, IV from III, and V from IV. Thus we know that any set of values that satisfies system I must also satisfy each succeeding system; in particular, from system V we find that any (x, y, z) that satisfies I must be

$$(1, 2, -1).$$

Accordingly, there can be no other solution of the original system I; if there is a solution, then this is it.

But do the values $(1, 2, -1)$ actually satisfy system I? For our systems of linear equations we have already pointed out that system I can be obtained from system II; similarly, II can be obtained from III, III from IV, and IV from V. Thus the solution of V, namely $(1, 2, -1)$, satisfies I.

Of course, you could verify by direct substitution that $(1, 2, -1)$ satisfies the system I, and actually you should do this to guard against computational error. But it is useful to note explicitly that the steps are reversible, so that the systems I, II, III, IV, and V are <u>equivalent</u> in accordance with the following definition:

<u>Definition 3-1</u>. Two systems of linear equations are said to be <u>equivalent</u> if and only if each solution of either system is also a solution of the other.

We know that the foregoing systems I through V are equivalent because the steps we have taken are reversible. In fact, the only operations we have performed have been of the following sorts:

A. Multiply an equation by a nonzero number.

B. Add one equation to another.

Reversing the process, we undo the addition by subtraction and the multiplication by division.

Actually, there is another operation we shall sometimes perform in our systematic solution of systems of linear equations, and it also is reversible:

C. Interchange two equations.

Thus, in solving the system

$$
\begin{aligned}
y + z &= 4,\\
x + 2y + z &= 3,\\
x - y + z &= 1,
\end{aligned}
$$

our first step would be to interchange the first two equations in order to have a leading coefficient differing from zero.

In the present chapter we shall investigate an orderly method of elimination, without regard to the particular values of the coefficients except that we shall avoid division by 0. Our method will be especially useful in dealing with several systems in which corresponding coefficients of the variables are equal while the right-hand members are different — a situation that often occurs in industrial and applied scientific problems.

You might use the procedure, for example, in "programming," i.e., devising a method, or program, for solving a system of linear equations by means of a modern electronic computing machine.

Exercises 3-1

1. Solve the following systems of equations:

(a) $3x + 4y = 4,\quad 5x + 7y = 1;$ (b) $x - 2y = 3,\quad y = 2;$

(c)
$$\begin{aligned} x + y - z &= 3,\\ 2y + z &= 10,\\ 5x - y - 2z &= -3; \end{aligned}$$

(d)
$$\begin{aligned} x - 3y + 2z &= 6,\\ y - z &= -4,\\ z &= 7; \end{aligned}$$

(e)
$$\begin{aligned} x + 2y + z - 3w &= 2,\\ y - 2z - w &= 7,\\ z - 2w &= 0,\\ w &= 3; \end{aligned}$$

(f)
$$\begin{aligned} 1x + 0y + 0z + 0w &= a,\\ 0x + 1y + 0z + 0w &= b,\\ 0x + 0y + 1z + 0w &= c,\\ 0x + 0y + 0z + 1w &= d. \end{aligned}$$

2. Solve by drawing graphs:

(a) $x + y = 2,\quad x - y = 2;$ (b) $3x - y = 11,\quad 5x + 7y = 1.$

3. Perform the following matrix multiplications:

(a) $$\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} u\\ v\\ w \end{bmatrix},$$

(b) $$\begin{bmatrix} 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} a & x\\ b & y\\ c & z \end{bmatrix}.$$

4. Given the following systems A and B, obtain B from A by steps consisting either of multiplying an equation by a nonzero constant or of adding an arbitrary multiple of an equation to another equation:

A:
$$\begin{aligned} x &= 2,\\ y &= 3,\\ z &= -1; \end{aligned}$$

B:
$$\begin{aligned} 2x - 3y + z &= -6,\\ x + 2y - z &= 9,\\ 3x + y + 3z &= 6. \end{aligned}$$

Are the two systems equivalent? Why or why not?

5. The solution set of one of the following systems of linear equations is empty, while the other solution set contains an infinite number of solutions. See if you can determine which is which, and give three particular numerical solutions for the system that does have solutions:

(a) $$\begin{aligned} x + 2y - z &= 3, \\ x - y + z &= 4, \\ 4x - y + 2z &= 14; \end{aligned}$$

(b) $$\begin{aligned} x + 2y - z &= 3, \\ x - y + z &= 4, \\ 4x - y + 2z &= 15. \end{aligned}$$

3-2. Formulation in Terms of Matrices

In applying our method to the solution of the original system of Section 3-1, namely to

$$\begin{aligned} x - y + z &= -2, \\ x - 2y - 2z &= -1, \\ 2x + y + 3z &= 1, \end{aligned}$$

we carried out just two types of algebraic operations in obtaining an equivalent system:

A. Multiplication of an equation by a number other than 0.

B. Addition of an equation to another equation.

We noted that a third type of operation is sometimes required, namely:

C. Interchange of two equations.

This third operation is needed if a coefficient by which we otherwise would divide is 0, and there is a subsequent equation in which the same variable has a nonzero coefficient.

The three foregoing operations can, in effect, be carried out through matrix operations. Before we demonstrate this, we shall see how the matrix notation and operations developed in Chapter 1 can be used to write a system of linear equations in matrix form and to represent the steps in the solution.

Let us again consider the system we worked with in Section 3-1:

$$\begin{aligned} x - \ y + \ z &= -2, \\ x - 2y - 2z &= -1, \\ 2x + \ y + 3z &= 1. \end{aligned}$$

We may display the detached coefficients of x, y, and z as a matrix A, namely

$$A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}.$$

Next, let us consider the matrices

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix};$$

the entries of X are the variables x, y, z, and of B are the right-hand members of the equations we are considering. By the definition of matrix multiplication, we have

$$AX = \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - \ y + \ z \\ x - 2y - 2z \\ 2x + \ y + 3z \end{bmatrix},$$

which is a 3×1 matrix having as entries the left-hand members of the equations of our linear system.

Now the equation

$$AX = B \tag{1}$$

that is,

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix},$$

is equivalent, by the definition of equality of matrices, to the entire system of linear equations. It is an achievement not to be taken modestly that we are able to consider and work with a large system of equations in terms of such a simple representation as is exhibited in Equation (1). Can you see the pattern that is emerging?

In passing, let us note that there is an interesting way of viewing the matrix equation

$$AX = Y, \tag{2}$$

where A is a given 3×3 matrix and X and Y are variable 3×1 column matrices. We recall that equations such as

$$y = ax + b$$

and

$$y = \sin x$$

define functions, where numbers x of the domain are mapped into numbers y of the range. We may also consider Equation (2) as defining a function, but in this case the domain consists of column matrices X and the range consists of column matrices Y; thus we have a matrix function of a matrix variable! For example, the matrix equation

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - y + z \\ x - 2y - 2z \\ 2x + y + 3z \end{bmatrix} = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

defines a function with a domain of 3×1 matrices

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{3}$$

and a range of 3×1 matrices

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} x - y + z \\ x - 2y - 2z \\ 2x + y + 3z \end{bmatrix}. \tag{4}$$

Thus with any column matrix of the form (3), the equation associates a column matrix of the form (4).

Looking again at the equation

$$AX = B,$$

where A is a given 3×3 matrix, B is a given 3×1 matrix, and X is a variable 3×1 matrix, we note that here we have an inverse question: What matrices X (if any) are mapped onto the particular B? For the case we have been considering,

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix},$$

we found in Section 3-1 that the unique solution is

$$X = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.$$

We shall consider some geometric aspects of the matrix-function point of view in Chapters 4 and 5.

We are now ready to look again at the procedure of Section 3-1, and restate each system in terms of matrices:

$$\begin{aligned}
\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \quad &\Leftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -3 \\ 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \\
&\Leftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -8 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 8 \end{bmatrix} \\
&\Leftrightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix} \\
&\Leftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}.
\end{aligned}$$

The two-headed arrows $\Leftrightarrow$ indicate that, as we saw in Section 3-1, the matrix equations are equivalent.

In order to see a little more clearly what has been done, let us look only at the first and last equations. The first is

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} ;$$

the last is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} .$$

Our process has converted the coefficient matrix A into the identity matrix I. Recall that

$$A^{-1}A = I.$$

Thus to bring about this change we must have completed operations equivalent to multiplying on the left by A^{-1}.

In brief, our procedure amounts to multiplying each member of

$$AX = B$$

on the left by A^{-1},

$$A^{-1}AX = A^{-1}B,$$

to obtain

$$X = A^{-1}B.$$

Let us note how far we have come. By introducing arrays of numbers as objects in their own right, and by defining suitable algebraic operations, we are able to write complicated systems of equations in the simple form

$$AX = B.$$

This condensed notation, so similar to the long-familiar

$$ax = b,$$

which we learned to solve "by division," indicates for us our solution process; namely, multiply both sides on the left by the reciprocal or inverse of A, obtaining the solution

$$X = A^{-1}B.$$

This similarity between

$$ax = b$$

and

$$AX = B$$

should not be pushed too far, however. There is only one real number, 0, that has no reciprocal; as we already know, there are many matrices that have no multiplicative inverses. Nevertheless, we have succeeded in our aim, which is perhaps the general aim of mathematics: to make the complicated simple by discovering its pattern.

Exercises 3-2

1. Write in matrix form:

(a) $$\begin{aligned} 4x - 2y + 7z &= 2, \\ 3x + y + 5z &= -1, \\ 6y - z &= 3; \end{aligned}$$

(b) $$\begin{aligned} x + y &= 2, \\ x - y &= 2. \end{aligned}$$

2. Determine the systems of algebraic equations to which the following matrix equations are equivalent:

(a) $$\begin{bmatrix} 3 & 4 & 5 \\ 1 & 2 & 3 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix},$$

(b) $$\begin{bmatrix} 3 & 2 & -2 \\ 2 & -1 & -4 \\ -1 & 1 & 5 \end{bmatrix}\begin{bmatrix} x & u \\ y & v \\ z & w \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 1 \end{bmatrix}.$$

3. Solve the following system of equations; then restate your work in matrix form:

$$\begin{aligned} x + y + z - w &= 1, \\ x - y + 3z + 2w &= 2, \\ 2x + y + 3z + w &= -2, \\ x - 2y + z + 3w &= 10. \end{aligned}$$

4\. (a) Onto what vector Y does the function defined by

$$Y = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

map the vector $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$? (b) What vector $\begin{bmatrix} x \\ y \end{bmatrix}$ does it map onto the vector $Y = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$?

5\. Let $A = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \end{bmatrix}$, $Y = \begin{bmatrix} y_1 \end{bmatrix}$, $a_i, x_i, y_1 \in R$. Discuss the domain of the function defined by

$$AX = Y.$$

Define the inverse function, if any.

3-3. Inverse of a Matrix

In Section 3-2 we wrote a system of linear equations in matrix form,

$$AX = B,$$

and saw that solving the system amounts to determining the inverse matrix A^{-1}, if it exists, since then we have

$$A^{-1}AX = B,$$

whence

$$X = A^{-1}B.$$

Our work involved a series of algebraic operations; let us learn how to duplicate this work by a series of matrix alterations. To do this, we suppress the column matrices in the scheme shown in Section 3-2 and look at the coefficient matrices on the left:

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -3 \\ 0 & 3 & 1 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -8 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Longleftrightarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Observe what happens if we substitute "row" for "equation" in the procedure of Section 3-1 and repeat our steps. In the first step, we multiply row 1 by -1 and add the result to row 2; multiply row 1 by -2 and add to row 3. Second, we multiply row 2 by 3 and add to row 3; multiply row 2 by -1. Third, we multiply row 3 by $3/8$ and add to row 2; multiply row 3 by $1/8$ and add to row 1; multiply row 3 by $-1/8$. Last, we add row 2 to row 1. Through "row operations" we have duplicated the changes in the coefficient matrix as we proceed through the series of equivalent systems.

The three algebraic operations described in Section 3-2 are paralleled by three matrix row operations:

Definition 3-2. The three row operations,

Interchange of any two rows,
Multiplication of all elements of any row by a nonzero constant,
Addition of an arbitrary multiple of any row to any other row,

are called elementary row operations on a matrix.

In Section 3-5, the exact relationship between row operations and the operations developed in Chapter 1 will be demonstrated. Earlier we defined equivalent systems of linear equations; in a corresponding way, we now define equivalent matrices.

Definition 3-3. Two matrices are said to be row equivalent if and only if each can be transformed into the other by means of elementary row operations.

We now turn our attention to the right-hand member of the equation

$$AX = B.$$

At the moment, the right-hand member consists solely of the matrix B, which we wish temporarily to suppress just as we temporarily suppressed the matrix X in considering the left-hand member. Accordingly, we need a coefficient matrix for the right-hand member. To obtain this, we use the identity matrix to write the equivalent equation

$$AX = IB. \tag{1}$$

Now our process converts this to

$$X = A^{-1}B,$$

which can be expressed equivalently as

$$IX = A^{-1}B. \tag{2}$$

When we compare Equation (1) and Equation (2) we notice that as A goes to I, I goes to A^{-1}. Might it be that the row operations, which convert A into I when applied to the left-hand member, will convert I into A^{-1} when applied to the right-hand member? Let us try. For convenience, we do this in parallel columns, giving an indication in parentheses of how each new row is obtained:

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} \Bigg| \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix};$$

$(R_1;\ -1R_1 + R_2;\ -2R_1 + R_3)$:

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & -1 & -3 \\ 0 & 3 & 1 \end{bmatrix} \Bigg| \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix};$$

$(R_1;\ -1R_2;\ 3R_2 + R_3)$:

$$\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & -8 \end{bmatrix} \Bigg| \begin{bmatrix} 1 & 0 & 0 \\ 1 & -1 & 0 \\ -5 & 3 & 1 \end{bmatrix};$$

$(\frac{1}{8} R_3 + R_1;\ \frac{3}{8} R_3 + R_2;\ -\frac{1}{8} R_3)$:

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Bigg| \begin{bmatrix} \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \\ -\frac{7}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix};$$

$(1R_2 + R_1;\ R_2;\ R_3)$:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \Bigg| \quad \begin{bmatrix} -\frac{4}{8} & \frac{4}{8} & \frac{4}{8} \\ -\frac{7}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix}.$$

To demonstrate that

$$B = \begin{bmatrix} -\frac{4}{8} & \frac{4}{8} & \frac{4}{8} \\ -\frac{7}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix}$$

is a left-hand inverse for A, it is necessary to show that

$$BA = \begin{bmatrix} -\frac{4}{8} & \frac{4}{8} & \frac{4}{8} \\ -\frac{7}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix} = I.$$

You are asked to verify this as an exercise, and also to check that $AB = I$, thus demonstrating that B is the inverse A^{-1} of A. In Section 3-5 we shall see why it is that $AB = I$ follows from $BA = I$ for matrices B that are obtained in this way as products of elementary matrices.

We now have the following rule for computing the inverse A^{-1} of a matrix A, if there is one: Find a series of elementary row operations that convert A into the identity matrix I; the same series of elementary row operations will convert the identity matrix I into the inverse A^{-1}.

When we start the process we may not know if the inverse exists. This need not concern us at the moment. If the application of the rule is successful, that is, if A is converted into I, then we know that A^{-1} exists. In subsequent sections, we shall discuss what happens when A^{-1} does not exist and we shall also demonstrate that the row operations can be brought about by means of the matrix operations that were developed in Chapter 1.

Exercises 3-3

1. Determine the inverse of each of the following matrices through row operations. (Check your answers.)

(a) $\begin{bmatrix} 2 & 3 \\ -1 & 2 \end{bmatrix}$, (b) $\begin{bmatrix} 0 & 3 \\ 4 & -2 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & 0 & 3 \\ 2 & 1 & 0 \\ -2 & 0 & 0 \end{bmatrix}$, (d) $\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 3 & 0 & 1 \end{bmatrix}$.

2. Determine the inverse, if any, of each of the following matrices:

(a) $\begin{bmatrix} 2 & 2 \\ -1 & \sqrt{2} \end{bmatrix}$, (b) $\begin{bmatrix} 4 & -2 \\ 6 & -3 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & 0 & 3 \\ 2 & 3 & 4 \\ 1 & 3 & 1 \end{bmatrix}$, (d) $\begin{bmatrix} -1 & 2 & 1 \\ 0 & 1 & -2 \\ 1 & 3 & 1 \end{bmatrix}$,

(e) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 3 & 6 \\ 5 & 1 & -1 \end{bmatrix}$.

3. Solve each of the following matrix equations:

(a) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 3 & 6 \\ 5 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ 7 \end{bmatrix}$, (b) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 3 & 6 \\ 5 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 3 & 6 \\ 5 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ -4 \end{bmatrix}$, (d) $\begin{bmatrix} 1 & 4 & 7 \\ 2 & 3 & 6 \\ 5 & 1 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$.

4. Solve

$$\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x & u & m & r \\ y & v & n & s \\ z & w & p & t \end{bmatrix} = \begin{bmatrix} 3 & 1 & -6 & \frac{3}{2} \\ 0 & -1 & 1 & 0 \\ 4 & 3 & -9 & \frac{11}{2} \end{bmatrix}.$$

5. Perform the multiplications

(a)
$$\begin{bmatrix} 3 & 2 & -2 \\ 2 & -1 & -4 \\ -1 & 1 & 5 \end{bmatrix} \begin{bmatrix} \frac{1}{17} & \frac{12}{17} & \frac{10}{17} \\ \frac{6}{17} & -\frac{13}{17} & -\frac{8}{17} \\ -\frac{1}{17} & \frac{5}{17} & \frac{7}{17} \end{bmatrix},$$

(b)
$$\begin{bmatrix} \frac{1}{17} & \frac{12}{17} & \frac{10}{17} \\ \frac{6}{17} & -\frac{13}{17} & -\frac{8}{17} \\ -\frac{1}{17} & \frac{5}{17} & \frac{7}{17} \end{bmatrix} \begin{bmatrix} 3 & 2 & -2 \\ 2 & 1 & -4 \\ -1 & 1 & 5 \end{bmatrix}.$$

6. Multiply both members of the matrix equation

$$\begin{bmatrix} 1 & 12 & 10 \\ 6 & -13 & -8 \\ -1 & 5 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 17 \\ 0 \\ 0 \end{bmatrix}$$

on the left by

$$\begin{bmatrix} 3 & 2 & -2 \\ 2 & -1 & -4 \\ -1 & 1 & 5 \end{bmatrix}$$

and use the result to solve the equation.

7. Solve:

$$\begin{aligned} 2x + y + 2z - 3w &= 0, \\ 4x + y + z + w &= 15, \\ 6x - y - z - w &= 5, \\ 4x - 2y + 3z - w &= 2. \end{aligned}$$

8. Solve:

$$\begin{aligned} 9x - y &= 37, \\ 8y - 2z &= -4, \\ 7z - 3w &= -17, \\ 2x + 6w &= 14. \end{aligned}$$

3-4. Linear Systems of Equations

In Sections 3-2 and 3-3, a procedure for solving a system of linear equations,

$$AX = B,$$

was presented. The method produces the multiplicative inverse A^{-1} if it exists.

In the present section we shall consider systems of linear equations in general.

Let us begin by looking at a simple illustration.

Example 1. Consider the system of equations,

$$\begin{aligned} 2x - 3y + 4z &= 5, \\ 2x + 7y - 2z &= 1, \\ 2x + 2y + z &= 3. \end{aligned}$$

We start in parallel columns, thus:

$$\begin{bmatrix} 2 & -3 & 4 \\ 2 & 7 & -2 \\ 2 & 2 & 1 \end{bmatrix} \quad \Bigg| \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Proceeding, we arrive after three steps at the following:

$$\begin{bmatrix} 1 & 0 & \frac{11}{10} \\ 0 & 1 & -\frac{3}{5} \\ 0 & 0 & 0 \end{bmatrix} \quad \Bigg| \quad \begin{bmatrix} \frac{7}{20} & \frac{3}{20} & 0 \\ -\frac{1}{10} & \frac{1}{10} & 0 \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{bmatrix}$$

If we multiply these two matrices on the right by the matrices

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix}, \quad \text{respectively,}$$

we obtain the system

$$x + \frac{11}{10} z = \frac{19}{10} ,$$

$$y - \frac{3}{5} z = - \frac{2}{5} ,$$

$$0 = 0.$$

There is no mathematical loss in dropping the equation $0 = 0$ from the system, which then can be written equivalently as

$$x = \frac{19}{10} - \frac{11}{10} z,$$

$$y = - \frac{2}{5} + \frac{3}{5} z.$$

Whatever value is given to z, this value and the corresponding values of x and y determined by these equations satisfy the original system. For example, a few solutions are shown in the following table:

z	x	y
-2	$\frac{41}{10}$	$-\frac{8}{5}$
1	$\frac{4}{5}$	$\frac{1}{5}$
0	$\frac{19}{10}$	$-\frac{2}{5}$
2	$-\frac{3}{10}$	$\frac{4}{5}$

<u>Example 2</u>. Now consider the systems

$$\begin{aligned} x + 2y - z &= 3, \\ x - y + z &= 4, \\ 4x - y + 2z &= 14. \end{aligned}$$

If we start in parallel columns and proceed as before, we obtain (as you should verify)

$$\begin{bmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 0 \end{bmatrix} \qquad \Bigg| \qquad \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & 0 \\ \frac{1}{3} & -\frac{1}{3} & 0 \\ -1 & -3 & 1 \end{bmatrix}$$

Multiplying these matrices on the right by the column matrices

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 3 \\ 4 \\ 14 \end{bmatrix},$$

we obtain the systems

$$x + \frac{1}{3}z = \frac{11}{3},$$

$$y - \frac{2}{3}z = -\frac{1}{3},$$

$$0 = -1.$$

If there were a solution of the original system, it would also be a solution of this last system, which is equivalent to the first. But the latter system contains the contradiction $0 = -1$; hence there is no solution of either system.

Do you have an intuitive geometric notion of what might be going on in each of the above systems? Relative to a 3-dimensional rectangular coordinate system, each of the equations represents a plane. Each pair of these planes actually intersect in a line. We might hope that the three lines of intersection (in each system) would intersect in a point. In the first system, however, the three lines are coincident; there is an entire "line" of solutions. On the other hand, in the second system, the three lines are parallel; there is no point that lies on all three planes. In the example worked out in Sections 3-2 and 3-3, the three lines intersect in a single point.

How many possible configurations, as regards intersections, can you list for 3 planes, not necessarily distinct from one another? They might, for example, have exactly one point in common; or two might be coincident and the third distinct from but parallel to them; and so on. There are systems of linear equations that correspond to each of these geometric situations.

Here are two additional systems that even more obviously than the system in Example 2, above, have no solutions:

$$x = 2, \qquad x + y + z = 2,$$

$$x = 3; \qquad x + y + z = 3.$$

Thus you see that the number of variables as compared with the number of equations does not determine whether or not there is a solution.

The method we have developed for solving linear systems is routine and can be applied to systems of any number of linear equations in any number of variables. Let us examine the general case and see what can happen. Suppose we have a system of m linear equations in n variables:

$$\begin{array}{c} a_1x_1 + a_2x_2 + \cdots + a_nx_n = \alpha_1 \\ \vdots \\ m_1x_1 + m_2x_2 + \cdots + m_nx_n = \alpha_m \end{array}$$

If any variable occurs with coefficient 0 in every equation, we drop it. If a coefficient needed as a divisor is zero, we interchange our equations; we might even interchange some of the terms in all the equations for the convenience of getting a nonzero coefficient in leading position. Otherwise we can proceed in the customary manner until there remain no equations in which any of the variables have nonzero coefficients. Eventually we have a system like this:

$$\begin{array}{llr} x_1 + & \text{linear terms in variables other than } x_1 \ldots x_k & = \beta_1, \\ \quad x_2 + & \qquad\qquad " & = \beta_2, \\ \qquad \ddots & & \\ \qquad\quad x_k + & \qquad\qquad " & = \beta_k, \end{array}$$

and (possibly) other equations in which no variable appears.

Either all of these other equations (if there are any) are of the form

$$0 = 0,$$

in which case they can be disregarded, or there is at least one of them of the form

$$0 = b, \qquad b \neq 0,$$

in which case there is a contradiction.

If there is a contradiction, the contradiction was present in the original system, which means there is no solution set.

In case there are no equations of the form

$$0 = b, \qquad\qquad b \neq 0,$$

we have two possibilities. Either there are no variables other than $x_1,\ldots,x_k$ which means that the system reduces to

$$\begin{aligned} x_1 &= \beta_1, \\ &\vdots \\ x_k &= \beta_k, \end{aligned}$$

and has a unique solution, or there really are variables other than $x_1,\ldots,x_k$ to which we can assign arbitrary values and obtain a family of solutions as in Example 1, above.

Exercises 3-4

1. (a) List all possible configurations, as regards intersections, for 3 distinct planes.

   (b) List also the additional possible configurations if the planes are allowed to be coincident.

2. Find the solutions, if any, of each of the following systems of equations:

(a)
$$\begin{aligned} x + y + 2z &= 1, \\ 2x + z &= 3, \\ 3x + 2y + 4z &= 4; \end{aligned}$$

(c)
$$\begin{aligned} x - 2y + z &= 1, \\ 2x + y - z &= 1, \\ x + 2y + 2z &= 2; \end{aligned}$$

(b)
$$\begin{aligned} x + y + z &= 6, \\ x + y + 2z &= 7, \\ y + z &= 1; \end{aligned}$$

(d)
$$\begin{aligned} 2v + x + y + z &= 0, \\ v - x + 2y + z &= 0, \\ 4v - x + 5y + 3z &= 1, \\ v - x + y - z &= 2; \end{aligned}$$

(e) $2x + y + z + w = 2,$
$x + 2y + z - w = -1,$
$4x + 5y + 3z - w = 0.$

3-5. Elementary Row Operations

In Section 3-1, three types of algebraic operations were listed as being involved in the solution of a linear system of equations; in Section 3-4, three types of row operations were used when we duplicated the procedure employing matrices. In this section we shall show how matrix multiplication underlies this earlier work, — in fact, how this basic operation can duplicate the row operations.

Let us start by looking at what happens if we perform any one of the three row operations on the identity matrix I of order 3. First if we multiply a row of I by a nonzero number n, we have a matrix of the form

$$\begin{bmatrix} n & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{or} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & n \end{bmatrix}.$$

Let J represent any matrix of this type. Second, if we add one row of I to another, we obtain

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

or one of three other similar matrices. (What are they?) Let K stand for any matrix of this type. Third, if we interchange two rows, we form matrices (denoted by L) of the form

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Matrices of these three types are called elementary matrices.

Definition 3-4. An elementary matrix is any square matrix obtained by performing a single elementary row operation on the identity matrix.

Each elementary matrix E (that is, each J, K, or L) has an inverse, that is, a matrix E^{-1} such that

$$E^{-1}E = I = E\,E^{-1}.$$

For example, the inverses of the elementary matrices

$$J = \begin{bmatrix} 1 & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad K = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad L = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

are the elementary matrices .

$$J^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{n} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad K^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}, \quad \text{and} \quad L^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

respectively, as you can verify by performing the multiplications involved.

An elementary matrix is related to its inverse in a very simple way. For example, J was obtained by <u>multiplying</u> a row of the identity matrix by n; J^{-1} is formed by <u>dividing</u> the same row of the identity matrix by n. In a sense J^{-1} "undoes" whatever was done to obtain J from the identity matrix; conversely, J will undo J^{-1}. Hence

$$J^{-1}J = I = JJ^{-1}.$$

The product of two elementary matrices also has an inverse, as the following theorem indicates.

<u>Theorem 3-1</u>. If square matrices A and B of order n have inverses A^{-1} and B^{-1}, then AB has an inverse $(AB)^{-1}$, namely

$$(AB)^{-1} = B^{-1}\,A^{-1}.$$

<u>Proof</u>. We have

$$(AB)\,(B^{-1}A^{-1}) = A(B\,B^{-1})A^{-1} = A\,A^{-1} = I$$

and

$$(B^{-1}A^{-1})\,(AB) = B^{-1}(A^{-1}A)B = B^{-1}B = I,$$

so that $B^{-1}A^{-1}$ is the inverse of AB by the definition of inverse.

You will recall that, for 2×2 matrices, this same proof was used in establishing Theorem 2.8.

Corollary 3-1-1. If square matrices $A, B, \ldots, K$ of order n have inverses A^{-1}, $B^{-1}, \ldots, K^{-1}$, then the product $AB \cdots K$ has an inverse $(AB \cdots K)^{-1}$, namely

$$(AB \cdots K)^{-1} = K^{-1} \cdots B^{-1}A^{-1}.$$

The proof by mathematical induction is left as an exercise.

Corollary 3-1-2. If $E_1, E_2, \ldots, E_k$ are elementary matrices of order n, then the matrix

$$B = E_1E_2 \cdots E_k$$

has an inverse, namely

$$B^{-1} = E_k^{-1} \cdots E_2^{-1}\ E_1^{-1}.$$

This follows immediately from Corollary 3-1-1 and the fact that each elementary matrix has an inverse.

The primary importance of an elementary matrix rests on the following property. If an $m \times n$ matrix A is multiplied on the left by an $m \times n$ elementary matrix E, then the product is the matrix obtained from A by the row operation by which the elementary matrix was formed from I. For example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ nd & ne & nf \\ g & h & i \end{bmatrix},$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \\ e & f \end{bmatrix} = \begin{bmatrix} a & b \\ c + e & d + f \\ e & f \end{bmatrix}.$$

You should verify these and other similar left multiplications by elementary matrices to familiarize yourself with the patterns.

Theorem 3-2. Any elementary row operation can be performed on an $m \times n$ matrix A by multiplying A on the left by the corresponding elementary matrix of order m.

The proof is left as an exercise.

Multiplications by elementary matrices can be combined. For example, to add -1 times the first row to the second row, we would multiply A on the left by the product of elementary matrices of the type $J^{-1}KJ$:

$$J^{-1}KJ = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that J multiplies the first row by -1; it is necessary to multiply by J^{-1} in order to change the first row back again to its original form after adding the first row to the second. Similarly, to add -2 times the first row to the third, we would multiply A on the left by

$$\begin{bmatrix} -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

To perform both of the above steps at the same time, we would multiply A on the left by

$$M_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}.$$

Since M_1 is the product of two matrices that are themselves products of elementary matrices, M_1 is the product of elementary matrices. By Corollary 3-1-1, the inverse of M_1 is the product, in reverse order, of the corresponding inverses of the elementary matrices.

Now our first step in the solution of the system of linear equations at the start of this chapter corresponds precisely to multiplying on the left by the above matrix M_1. If we multiply both sides of System I on page 103 on the left by M_1,

$$\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ 1 & -2 & -2 \\ 2 & 1 & 3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}\begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix},$$

we obtain System II. For the second, third, and fourth steps, the corresponding matrix multipliers must be respectively

$$M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 3 & 1 \end{bmatrix}, \quad M_3 = \begin{bmatrix} 1 & 0 & \frac{1}{8} \\ 0 & 1 & \frac{3}{8} \\ 0 & 0 & -\frac{1}{8} \end{bmatrix}, \quad M_4 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus multiplying on the left by M_2 has the effect of leaving the first row unaltered, multiplying the second row by -1, and adding 3 times the second row to the third.

Let us now take advantage of the associative law for the multiplication of matrices to form the product

$$M = M_4M_3M_2M_1 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & \frac{1}{8} \\ 0 & 1 & \frac{3}{8} \\ 0 & 0 & \frac{-1}{8} \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -\frac{4}{8} & \frac{4}{8} & \frac{4}{8} \\ -\frac{7}{8} & \frac{1}{8} & \frac{3}{8} \\ \frac{5}{8} & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix}.$$

We recognize the inverse of the original coefficient matrix, as determined in Section 3-3.

<u>Theorem 3-3</u>. If

$$BA = I,$$

where B is a product of elementary matrices, then

$$AB = I,$$

so that B is the inverse of A.

<u>Proof</u>. By Corollary 3-1-2, B has an inverse B^{-1}. Now from

$$BA = I$$

we get

$$B^{-1}BA = B^{-1}I,$$

whence

$$A = B^{-1},$$

so that

$$AB = B^{-1}B = I$$

and

$$A^{-1} = (B^{-1})^{-1} = B.$$

Exercises 3-5

1. Find matrices A, B, and C such that

(a) $$A \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \\ 1 & -1 & -1 \end{bmatrix},$$

(b) $$B \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 3 & 4 \\ 1 & 0 & -2 \\ 1 & -1 & 1 \end{bmatrix},$$

(c) $$C \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 2 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2. Express each of the following matrices as a product of elementary matrices:

(a) $$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & -\frac{3}{8} & -\frac{1}{8} \end{bmatrix},$$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

3. Using your answers to Exercise 2, form the inverse of each of the given matrices.

4. Find three 4×4 matrices, one each of the types J, K, and L, that will accomplish elementary row transformations when applied as a left multiplier to a 4×2 matrix.

5. Solve the following system of equations by means of elementary matrices:

$$\begin{aligned} x - \ \ y - 2z &= 3, \\ y + 3z &= 5, \\ 2x + 2y - 3z &= 15. \end{aligned}$$

6. (a) Find the inverse of the matrix

$$\begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 2 \\ 2 & 0 & 1 \end{bmatrix}.$$

(b) Express the inverse a product of elementary matrices.

(c) Do you think the answer to Exercise 6(b) is unique? Why or why not? Compare, in class, your answer with the answers found by other members of your class.

7. Give a proof of Corollary 3-1-1 by mathematical induction.

8. Perform each of the following multiplications:

(a) $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(b) $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$,

(c) $\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

9. State a general conjecture you can make on the basis of your solution of Exercise 8.

3-6. Summary

In this chapter, we have discussed the role of matrices in finding the solution set of a system of linear equations. We began by looking at the familiar algebraic methods of obtaining such solutions and then learned to duplicate this procedure using matrix notation and row operations. The intimate connection between row operations and the fundamental operation of matrix multiplication was demonstrated; any elementary operation is the result of left multiplication by an elementary matrix. We can obtain the solution set, if it exists, to a system of linear equations either by algebraic methods, or by row operations, or by multiplication by elementary matrices. Each of these three procedures involves steps that are reversible, a condition that assures the equivalence of the systems.

Our work with systems of linear equations led us to a method for producing the inverse of a matrix when it exists. The identical sequence of row operations that converts a matrix A into the identity matrix will convert the identity matrix into the inverse of A, namely A^{-1}. The inverse is particularly helpful when we need to solve many systems of linear equations, each possessing the same coefficient matrix A but different right-hand column matrices B.

Since the matrix procedure 'diagonalized' the coefficient matrix, the method is often called the "diagonalization method." Although we have not previously mentioned it, there is an alternative method for solving a linear system that is often more useful when dealing with a single system. In this alternative procedure, we apply elementary matrices to reduce the system to the form

$$\begin{bmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}, \qquad (1)$$

as in System III in Section 3-1, from which the value for z can readily be obtained. This value of z can then be substituted in the next to last equation to determine the value of y, and so on. An examination of the coefficient matrix displayed above shows clearly why this procedure is called the "triangularization method."

On the other hand, the diagonalization method can be speeded up to bypass the triangularized matrix of coefficients in (1), or in System III of Section 3-1, altogether. Thus, after pivoting on the coefficient of x in the System I,

$$\text{I} \qquad \begin{aligned} x - y + z &= -2, \\ x - 2y - 2z &= -1, \\ 2x + y + 3z &= 1, \end{aligned}$$

to obtain the system

$$\text{II} \qquad \begin{aligned} x - y + z &= -2, \\ 0 - y - 3z &= 1, \\ 0 + 3y + z &= 5, \end{aligned}$$

we could next pivot completely on the coefficient of y to obtain the system

$$\text{III}' \qquad \begin{aligned} x + 0 + 4z &= -3, \\ 0 + y + 3z &= -1, \\ 0 + 0 - 8z &= 8, \end{aligned}$$

and then on the coefficient of z to get the system

$$\text{IV}' \qquad \begin{aligned} x + 0 + 0 &= 1, \\ 0 + y + 0 &= 2, \\ 0 + 0 + z &= -1, \end{aligned}$$

which you will recognize as the system V of Section 3-1.

It should now be plain that the routine diagonalization and triangularization methods can be applied to systems of any number of equations in any number of variables and that the methods can be adapted to machine computations.

Chapter 4

REPRESENTATION OF COLUMN MATRICES AS GEOMETRIC VECTORS

4-1. The Algebra of Vectors

In the present chapter, we shall develop a simple geometric representation for a special class of matrices — namely, the set of column matrices $\begin{bmatrix} a \\ b \end{bmatrix}$ with two entries each. The familiar algebraic operations on this set of matrices will be reviewed and also given geometric interpretation, which will lead to a deeper understanding of the meaning and implications of the algebraic concepts.

By definition, a column vector of order 2 is a 2×1 matrix. Consequently, using the rules of Chapter 1, we can add two such vectors or multiply any one of them by a number. The set of column vectors of order 2 has, in fact, an algebraic structure with properties that were largely explored in our study of the rules of operation with matrices.

In the following pair of theorems, we summarize what we already know concerning the algebra of these vectors, and in the next section we shall begin the interpretation of that algebra in geometric terms.

Theorem 4-1. Let V and W be column vectors of order 2, let r be a number, and let A be a square matrix of order 2. Then

$$V + W, \ rV, \ \text{and} \ AV$$

are each column vectors of order 2.

For example, if

$$V = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad W = \begin{bmatrix} 0 \\ -2 \end{bmatrix}, \quad r = 4, \quad \text{and} \quad A = \begin{bmatrix} 0 & 1 \\ 3 & -2 \end{bmatrix},$$

then

$$V + W = \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 0 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad rV = 4\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 4 \\ 8 \end{bmatrix},$$

and

$$AV = \begin{bmatrix} 0 & 1 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \end{bmatrix}.$$

<u>Theorem 4-2</u>. Let V, W, and U be column vectors of order 2, let r and s be numbers, and let A and B be square matrices of order 2. Then all the following laws are valid.

I. Laws for the addition of vectors:

(a) $V + W = W + V$,

(b) $(V + W) + U = V + (W + U)$,

(c) $V + \underline{0} = V$,

(d) $V + (-V) = \underline{0}$.

II. Laws for the numerical multiplication of vectors:

(a) $r(V + W) = rV + rW$,

(b) $r(sV) = (rs)V$,

(c) $(r + s)V = rV + sV$,

(d) $0V = \underline{0}$,

(e) $1V = V$,

(f) $r\underline{0} = \underline{0}$.

III. Laws for the multiplication of vectors by matrices:

(a) $A(V + W) = AV + AW$,

(b) $(A + B)V = AV + BV$,

(c) $A(BV) = (AB)V$,

(d) $0_2 V = \underline{0}$,

(e) $IV = V$,

(f) $A(rV) = (rA)V = r(AV)$.

In Theorem 4-2, $\underline{0}$ denotes the column vector of order 2, and 0_2 the square matrix of order 2, all of whose entries are 0.

Both of the preceding theorems have already been proved for matrices. Since column vectors are merely special types of matrices, the theorems as stated must likewise be true. They would also be true, of course, if 2 were replaced by 3 or by a general n, throughout, with the understanding that a <u>column vector of order</u> n is a matrix of order $n \times 1$.

Exercises 4-1

1. Let

$$V = \begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad W = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \text{and} \quad U = \begin{bmatrix} -4 \\ 2 \end{bmatrix};$$

let

$$r = 2 \quad \text{and} \quad s = -1;$$

and let

$$A = \begin{bmatrix} 3 & 0 \\ 2 & -1 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 1 \\ -2 & 1 \end{bmatrix}.$$

Verify each of the laws stated in Theorem 4-2 for this choice of values for the variables.

2. Determine the vector V such that $AV - AW = AW + BW$, where

$$A = \begin{bmatrix} 5 & 1 \\ 4 & -2 \end{bmatrix}, \quad W = \begin{bmatrix} 3 \\ 9 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ -4 & 2 \end{bmatrix}.$$

3. Determine the vector V such that $2V + 2W = AV + BV$, if

$$W = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} -1 & -2 \\ 1 & -1 \end{bmatrix}.$$

4. Find V, if

$$A = \begin{bmatrix} 2/3 & -1/3 \\ -1/3 & 2/3 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \quad \text{and} \quad A(3V) = A(BV).$$

5. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}.$$

Evaluate

(a) $A \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and (b) $A \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

(c) Using your answers to parts (a) and (b), determine the entries of A if, for every vector V of order 2,

$$AV = \underline{0}.$$

(d) State your result as a theorem.

6. Restate the theorem obtained in Exercise 5 if A is a square matrix of order n and V stands for any column vector of order n. Prove the new theorem for n = 3. Try to prove the theorem for all n.

7. Using your answers to parts (a) and (b) of Exercise 5, determine the entries of A if, for <u>every</u> vector V of order 2,

$$AV = V.$$

State your result as a theorem.

8. Restate the theorem obtained in Exercise 7 if A is a square matrix of order n and V stands for any column vector of order n. Prove this theorem for n = 3. Try to prove the theorem for all n.

9. Theorems 4-1 and 4-2 summarize the properties of the algebra of column vectors with 2 entries. State two analogous theorems summarizing the properties of the algebra of row vectors with 2 entries. Show that the two algebraic structures are isomorphic.

4-2. Vectors and Their Geometric Representation

The notion of a vector occurs frequently in the physical sciences, where a vector is often referred to as a quantity having both length and direction and accordingly is represented by an arrow. Thus force, velocity, and even displacement are vector quantities.

Confining our attention to the coordinate plane, let us investigate this intuitive notion of vector and see how these physical or geometric vectors are related to the algebraic column and row vectors of Section 4-1.

An <u>arrow</u> in the plane is determined when the coordinates of its <u>tail</u> and the coordinates of its <u>head</u> are given. Thus the arrow A_1 from (1,2) to (5,4) is shown in Figure 4-1. Such an arrow, in a given position, is called a <u>located</u>

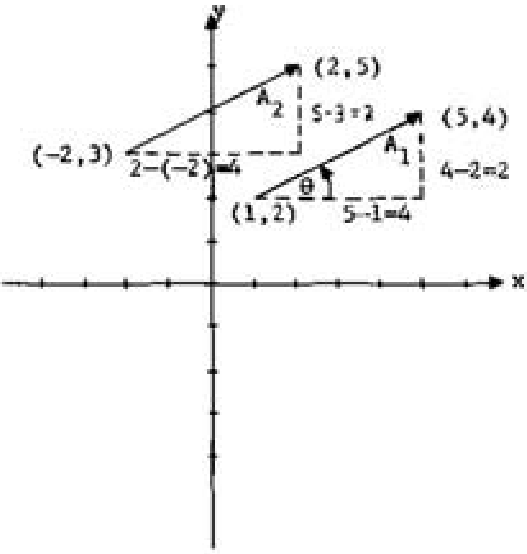


Figure 4-1. Arrows in the plane.

vector; its tail is called the initial point, and its head the terminal point (or end point), of the located vector. A second located vector A_2, with initial point (-2,3) and terminal point (2,5), is also shown in Figure 4-1.

A located vector A may be described briefly by giving first the coordinates of its initial point and then the coordinates of its terminal point. Thus the located vectors in Figure 4-1 are

$$A_1\text{: }(1,2)(5,4) \quad \text{and} \quad A_2\text{: }(-2,3)(2,5).$$

The horizontal run, or x component, of A_1 is

$$5 - 1 = 4,$$

and its vertical rise, or y component, is

$$4 - 2 = 2.$$

Accordingly, by the Pythagorean theorem, its length is

$$\sqrt{(5-1)^2 + (4-2)^2} = \sqrt{4^2 + 2^2} = 2\sqrt{5} .$$

Its direction is determined by the angle θ that it makes with the positive x direction:

$$\cos \theta = \frac{4}{2\sqrt{5}} = \frac{2\sqrt{5}}{5}, \quad \sin \theta = \frac{2}{2\sqrt{5}} = \frac{\sqrt{5}}{5} .$$

Since $\sin \theta$ is the cosine of the angle that A_1 makes with the y direction, $\cos \theta$ and $\sin \theta$ are called the direction cosines of A_1.

You might wonder why we did not write simply

$$\tan \theta = \frac{2}{4} = \frac{1}{2}$$

instead of the equations for $\cos \theta$ and $\sin \theta$. The reason is that while the value of $\tan \theta$ determines slope, it does not determine the direction of A_1. Thus if ϕ is the angle for the located vector (5,4)(1,2) opposite to A_1, then

$$\tan \phi = \frac{-2}{-4} = \frac{1}{2} = \tan \theta;$$

but the angles ϕ and θ from the positive x direction cannot be equal since they terminate in directions differing by π.

Now the x and y components of the second located vector A_2: (-2,3) (2,5) in Figure 4-1 are, respectively,

$$2 - (-2) = 4 \quad \text{and} \quad 5 - 3 = 2,$$

so that A_1 and A_2 have equal x components and equal y components; consequently, they have the same length and the same direction. They are not in the same position, so of course they are not the same located vectors; but since in dealing with vectors we are especially interested in length and direction we say that they are equivalent.

Definition 4-1. Two located vectors are said to be equivalent if and only if they have the same length and the same direction.

For any prescribed point P in the plane, there is a located vector equivalent to A_1 (and to A_2) and having P as initial point. To determine the coordinates of the terminal point, you have only to add the components of A_1 to the corresponding coordinates of P. Thus for the initial point P: (3,−7), the terminal point is

$$(3 + 4, -7 + 2) = (7, -5),$$

so that the located vector is

$$A_3: (3, -7)(7, -5).$$

You might plot the initial point and the terminal point of A_3 in order to check that it actually is equivalent to A_1 and to A_2.

In general, we denote the located vector A with initial point (x_1,y_1) and terminal point (x_2,y_2) by

$$A: (x_1,y_1)(x_2,y_2).$$

Its x and y components are, respectively,

$$x_2 - x_1 \quad \text{and} \quad y_2 - y_1.$$

Its length is

$$r = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

If $r \neq 0$, then its direction is determined by the angle that it makes with the x axis:

$$\cos\theta = \frac{x_2 - x_1}{r}, \qquad \sin\theta = \frac{y_2 - y_1}{r}.$$

If $r = 0$, we find it convenient to say that the vector is directed in any direction we please. As you will see, this makes it possible to state several theorems more simply, without having to mention special cases. For example, this <u>null</u> vector is both parallel and perpendicular to every direction in the plane!

A second located vector

$$A': (x_3,y_3)(x_4,y_4)$$

is equivalent to A if and only if it has the same length and the same direction as A, or, what amounts to the same thing, if and only if it has the same components as A:

$$x_4 - x_3 = x_2 - x_1, \quad y_4 - y_3 = y_2 - y_1.$$

For any given point (x_0,y_0), the located vector

$$B: (x_0,y_0)(x_0 + x_2 - x_1, y_0 + y_2 - y_1)$$

is equivalent to A and has (x_0,y_0) as its initial point.

It thus appears that the located vector A is determined except for its position by its components

$$a = x_2 - x_1 \qquad \text{and} \qquad b = y_2 - y_1.$$

These can be considered as the entries of a column vector

$$V = \begin{bmatrix} a \\ b \end{bmatrix}.$$

In this way, any located vector A determines a column vector V. Conversely, for any given point P, the entries of any column vector V can be considered as the components of a located vector A with P as initial point. The locater vector A is said to <u>represent</u> V.

A column vector is a "free vector" in the sense that it determines the components (and therefore the magnitude and direction), but not the position, of any located vector that represents it. In particular, we shall assign to the column vector

$$V = \begin{bmatrix} u \\ v \end{bmatrix}$$

a <u>standard representation</u>

$$\overrightarrow{OP} : (0,0)(u,v)$$

as the located vector from the origin to the point

$$P : (u,v),$$

as illustrated in Figure 4-2; this is the representation to which we shall ordinarily refer unless otherwise stated.

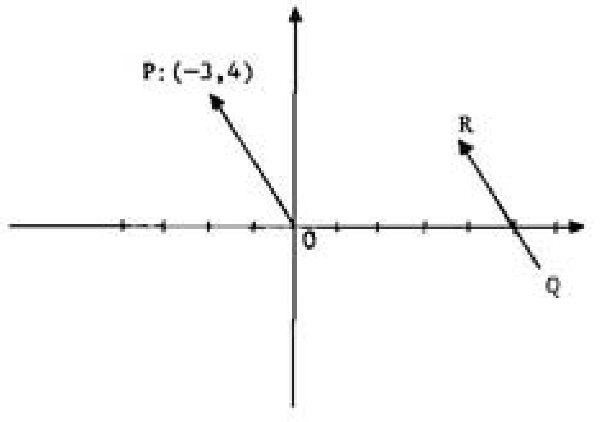


Figure 4-2. Representations of the column vector $\begin{bmatrix} -3 \\ 4 \end{bmatrix}$ as located vectors $\overrightarrow{OP}$ and $\overrightarrow{QR}$.

Similarly, of course, the components of the located vector A can be considered as the entries of a row vector. For the present, however, we shall consider only column vectors and the corresponding geometric vectors; in this chapter, the term "vector" will ordinarily be used to mean "column vector," not "row vector" or "located vector."

The length of the located vector $\overrightarrow{OP}$, to which we have previously referred, is called the <u>length</u> or <u>norm</u> of the column vector

$$V = \begin{bmatrix} u \\ v \end{bmatrix}.$$

Using the symbol $||V||$ to stand for the norm of V, we have

$$||V|| = \sqrt{u^2 + v^2} \quad .$$

Thus, if u and v are not both zero, the direction cosines of $\overrightarrow{OP}$ are

$$\frac{u}{\|V\|} \quad \text{and} \quad \frac{v}{\|V\|},$$

respectively; these are also called the direction cosines of the column vector V.

The association between column or row vectors and directed line segments, introduced in this section, is as applicable to 3-dimensional space as it is to the 2-dimensional plane. The only difference is that the components of a located vector in 3-dimensional space will be the entries of a column or row vector of order 3, not a column or row vector of order 2.

In the rest of this chapter and in Chapter 5, you will see how Theorems 4-1 and 4-2 can be interpreted through geometric operations on located vectors and how the algebra of matrices leads to beautiful geometric results.

Exercises 4-2

1. Of the following pairs of vectors,

(a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 3 \\ 3 \end{bmatrix};$ (f) $\begin{bmatrix} -5 \\ -12 \end{bmatrix}, \quad \begin{bmatrix} 12 \\ 5 \end{bmatrix};$

(b) $\begin{bmatrix} 4 \\ -3 \end{bmatrix}, \quad \begin{bmatrix} -8 \\ 6 \end{bmatrix};$ (g) $\begin{bmatrix} \sqrt{2} \\ -3\sqrt{2} \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -6 \end{bmatrix};$

(c) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ -\sqrt{2} \end{bmatrix};$ (h) $\begin{bmatrix} -8 \\ 15 \end{bmatrix}, \quad \begin{bmatrix} 16 \\ 30 \end{bmatrix};$

(d) $\begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix};$ (i) $\begin{bmatrix} 1 \\ 5 \end{bmatrix}, \quad \begin{bmatrix} -4 \\ 1 \end{bmatrix};$

(e) $\begin{bmatrix} -9 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 2\sqrt{15} \\ -5 \end{bmatrix};$ (j) $\begin{bmatrix} 3 \\ 4 \end{bmatrix}, \quad \begin{bmatrix} 3t \\ 4t \end{bmatrix},$

which have the same length? Which have the same direction?

2. Let $V = t \begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Draw arrows from the origin representing V for

$$t = 1, \quad t = 2, \quad t = 3, \quad t = -1, \quad t = -2, \quad \text{and} \quad t = -3.$$

In each case, compute the length and direction cosines of V.

3. In a rectangular coordinate plane, draw the standard representation for each of the following sets of vectors. Use a different coordinate plane for each set of vectors. Find the length and direction cosines of each vector:

(a) $\begin{bmatrix}1\\0\end{bmatrix}$, $\begin{bmatrix}0\\1\end{bmatrix}$, and $\begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}0\\1\end{bmatrix}$;

(b) $\begin{bmatrix}1\\0\end{bmatrix}$, $\begin{bmatrix}3\\2\end{bmatrix}$, and $\begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}3\\2\end{bmatrix}$;

(c) $\begin{bmatrix}1\\0\end{bmatrix}$, $\begin{bmatrix}-4\\-2\end{bmatrix}$, and $\begin{bmatrix}1\\0\end{bmatrix} + \begin{bmatrix}-4\\-2\end{bmatrix}$;

(d) $\begin{bmatrix}2\\3\end{bmatrix}$, $\begin{bmatrix}3\\-1\end{bmatrix}$, and $\begin{bmatrix}2\\3\end{bmatrix} + \begin{bmatrix}3\\-1\end{bmatrix}$;

(e) $\begin{bmatrix}5\\-4\end{bmatrix}$, $\begin{bmatrix}2\\2\end{bmatrix}$, and $\begin{bmatrix}5\\-4\end{bmatrix} + \begin{bmatrix}2\\2\end{bmatrix}$.

4. Let $V = \begin{bmatrix}x\\y\end{bmatrix} = t\begin{bmatrix}1\\m\end{bmatrix} + \begin{bmatrix}0\\b\end{bmatrix}$.

Draw the line segments $\overrightarrow{OP}$ representing V if $t = 0, \pm 1, \pm 2$, and

(a) $m = 1$, $b = 0$;

(b) $m = 2$, $b = 1$;

(c) $m = -1/2$, $b = 3$.

In each case, verify that the corresponding set of five points (x,y) lies on a line.

5. Two column vectors are called <u>parallel</u> provided their standard geometric representations lie on the same line through the origin. If A and B are nonzero parallel column vectors, determine the two possible relationships between the direction cosines of A and the direction cosines of B.

6. Determine all the vectors of the form $\begin{bmatrix}u\\v\end{bmatrix}$ that are parallel to

(a) $\begin{bmatrix}2\\4\end{bmatrix}$,

(b) $\begin{bmatrix}-1\\3\end{bmatrix}$,

(c) $\begin{bmatrix}4\\5\end{bmatrix} + \begin{bmatrix}1\\10\end{bmatrix}$,

(d) $\begin{bmatrix}9\\-5\end{bmatrix} - \begin{bmatrix}3\\1\end{bmatrix}$,

(e) $\begin{bmatrix}0\\3\end{bmatrix}$.

4-3. Geometric Interpretation of the Multiplication of a Vector by a Number

The geometrical significance of the multiplication of a vector by a number is readily guessed on comparing the geometrical representations of the vectors V, $2V$, and $-2V$ for

$$V = \begin{bmatrix} -3 \\ 4 \end{bmatrix}.$$

By definition,

$$2V = \begin{bmatrix} -6 \\ 8 \end{bmatrix},$$

while

$$-2V = \begin{bmatrix} 6 \\ -8 \end{bmatrix}.$$

Thus, as you can see in Figures 4-3 and 4-4, the standard representations of V and $2V$ have the same direction, while $-2V$ is represented by an arrow pointing in the opposite direction. The length of the arrow associated with V

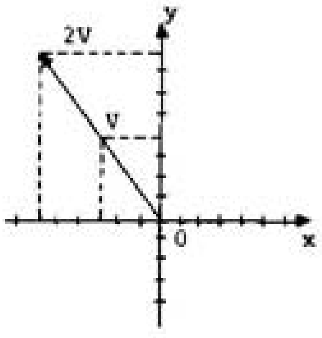


Figure 4-3. The product of a vector and a positive number.

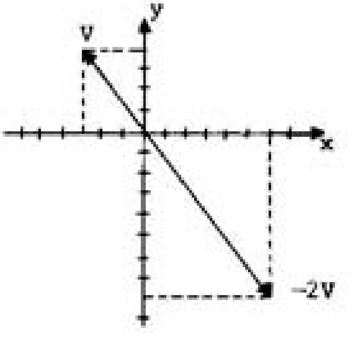


Figure 4-4. The product of a vector and a negative number.

is 5, while those for $2V$ and $-2V$ each have length 10. Thus, multiplying V by 2 produced a stretching of the associated geometric vector to twice its original length while leaving its direction unchanged. Multiplication by -2 not only doubled the length of the arrow but also reversed its direction.

These observations lead us to formulate the following theorem.

Theorem 4-3. Let the directed line segment $\overrightarrow{PQ}$ represent the vector V and let r be a number. Then the vector rV is represented by a directed line having length $|r|$ times the length of $\overrightarrow{PQ}$. If $r \geq 0$, the representation of rV has the same direction as $\overrightarrow{PQ}$; if $r < 0$, the direction of the representation of rV is opposite to that of $\overrightarrow{PQ}$.

Proof. Let V be the vector $\begin{bmatrix} u \\ v \end{bmatrix}$. Then

$$||V|| = \sqrt{u^2 + v^2} .$$

Now,

$$rV = \begin{bmatrix} ru \\ rv \end{bmatrix} ;$$

hence,

$$\begin{aligned} ||rV|| &= \sqrt{(ru)^2 + (rv)^2} \\ &= \sqrt{r^2(u^2 + v^2)} \\ &= |r| \sqrt{u^2 + v^2} \\ &= |r| \ ||V||. \end{aligned}$$

This proves the first part of the theorem.

If

$$r = 0 \quad \text{or} \quad V = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

the second part of the theorem is certainly true.

If

$$r \neq 0 \text{ and } V \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

the direction cosines of $\overrightarrow{PQ}$ are

$$\frac{u}{||V||} \quad \text{and} \quad \frac{v}{||V||},$$

while those of rV are

$$\frac{ru}{|r| \ ||V||} \quad \text{and} \quad \frac{rv}{|r| \ ||V||}.$$

If $r > 0$, we have $|r| = r$, whence it follows that the arrows associated with V and rV have the same direction cosines and, therefore, the same direction. If $r < 0$, we have $|r| = - r$, and the direction cosines of the arrow associated with rV are the negatives of those of $\overrightarrow{PQ}$. Thus, the direction of the representation of rV is opposite to that of $\overrightarrow{PQ}$. This completes the proof of the theorem.

One way of stating part of the theorem just proved is to say that if r is a number and V is a vector, then V and rV are parallel vectors (see Exercise 4-2-6); thus they can be represented by arrows lying on the same line through the origin. On the other hand, if the arrows representing two vectors are parallel, it is easy to show that you can always express one of the vectors as the product of the other vector by a suitably chosen number. Thus, by checking direction cosines, it is easy to verify that

$$\begin{bmatrix} 5 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} 50 \\ -20 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} -10 \\ 4 \end{bmatrix}$$

are parallel vectors, and that

$$\begin{bmatrix} 50 \\ -20 \end{bmatrix} = 10 \begin{bmatrix} 5 \\ -2 \end{bmatrix}, \quad \text{while} \quad \begin{bmatrix} -10 \\ 4 \end{bmatrix} = -2 \begin{bmatrix} 5 \\ -2 \end{bmatrix}.$$

In the exercises that follow, you will be asked to show why the general result illustrated by this example holds true.

Exercises 4-3

1. Let L be the set of all vectors parallel to the vector $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$. Fill in the following blanks so as to produce in each case a true statement:

(a) $\begin{bmatrix} 4 \\ - \end{bmatrix} \in L$; (d) $\begin{bmatrix} 6 \\ - \end{bmatrix} \notin L$;

(b) $\begin{bmatrix} - \\ 9 \end{bmatrix} \in L$; (e) for every real number t, $\begin{bmatrix} 8t \\ - \end{bmatrix} \in L$;

(c) $\begin{bmatrix} -2/3 \\ - \end{bmatrix} \in L$; (f) for every real number t, $\begin{bmatrix} - \\ -12t \end{bmatrix} \in L$;

(g) for every real number $h \neq 0$, $\begin{bmatrix} h \\ - \end{bmatrix} \notin L$.

2. Verify graphically and prove algebraically that the vectors in each of the following pairs are parallel. In each case, express the first vector as the product of the second vector by a number:

(a) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$; (d) $\begin{bmatrix} 2 \\ -32 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -16 \end{bmatrix}$;

(b) $\begin{bmatrix} 5 \\ 4 \end{bmatrix}$, $\begin{bmatrix} 10 \\ 8 \end{bmatrix}$; (e) $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$, $\begin{bmatrix} -8 \\ 4 \end{bmatrix}$;

(c) $\begin{bmatrix} -12 \\ 15 \end{bmatrix}$, $\begin{bmatrix} 4 \\ -5 \end{bmatrix}$; (f) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 9 \end{bmatrix}$.

3. Let V be a vector and W a nonzero vector such that V and W are parallel. Prove that there exists a real number r such that

$$V = rW.$$

4. Prove:

(a) If $rV = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $r \neq 0$, then $V = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

(b) If $rV = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $V \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, then $r = 0$.

5. Show that the vector $V + rV$ has the same direction as V if $r \geq -1$, and the opposite direction to V if $r < -1$. Show also that

$$||V + rV|| = ||V|| \ |1 + r|.$$

4-4. Geometrical Interpretation of the Addition of Two Vectors

If we have two vectors V and W, $V = \begin{bmatrix} a \\ b \end{bmatrix}$ and $W = \begin{bmatrix} c \\ d \end{bmatrix}$, their sum is, by definition,

$$V + W = \begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a + c \\ b + d \end{bmatrix}.$$

To interpret the addition geometrically, let us return momentarily to the concept of a "free" vector. Previously we have associated a column vector

$$W = \begin{bmatrix} c \\ d \end{bmatrix}$$

with some located vector

$$A: (x_1, y_1)(x_2, y_2)$$

such that $c = x_2 - x_1$ and $d = y_2 - y_1$. In particular, we can associate W with the vector

$$A: (a,b)(a + c, \;\; b + d).$$

We use A to represent W and use the standard representation for V and $V + W$ in Figure 4-5.

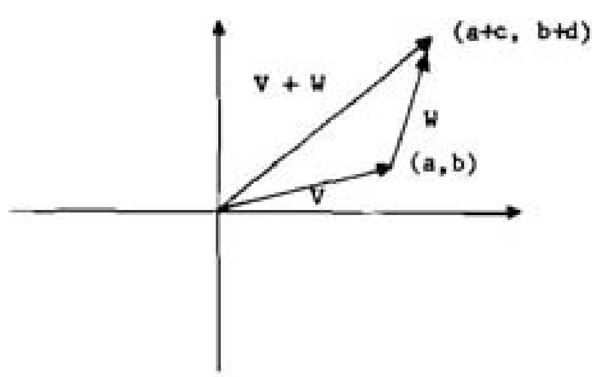


Figure 4-5. Vector addition.

Also, we can represent V as the located vector

$$B: (c,d)(a + c, \;\; b + d)$$

and obtain an alternative representation. If the two possibilities are drawn on one set of coordinate axes, we have a parallelogram in which the diagonal represents the sum; see Figure 4-6.

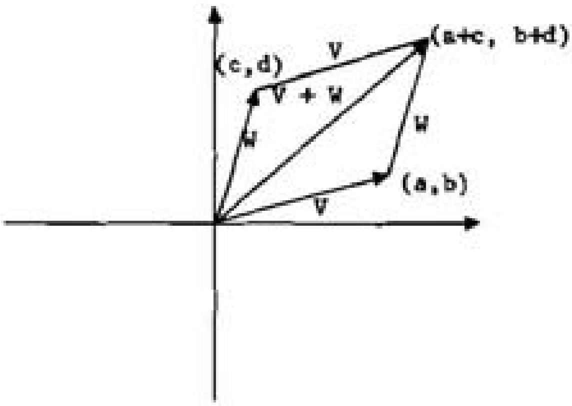


Figure 4-6. Parallelogram rule for addition.

The parallelogram rule is often used in physics to obtain the resultant when two forces are acting from a single point.

Let us consider now the sum of three vectors

$$U = \begin{bmatrix} e \\ f \end{bmatrix}, \quad V = \begin{bmatrix} a \\ b \end{bmatrix}, \quad \text{and} \quad W = \begin{bmatrix} c \\ d \end{bmatrix}.$$

We choose the three located vectors

$$\overrightarrow{OP} : (0,0)(a,b)$$

$$\overrightarrow{PQ} : (a,b)(a + c, \ b + d)$$

and

$$\overrightarrow{QR} : (a + c, \ b + d)(a + c + e, \ b + d + f)$$

to represent V, W, and U respectively; see Figure 4-7.

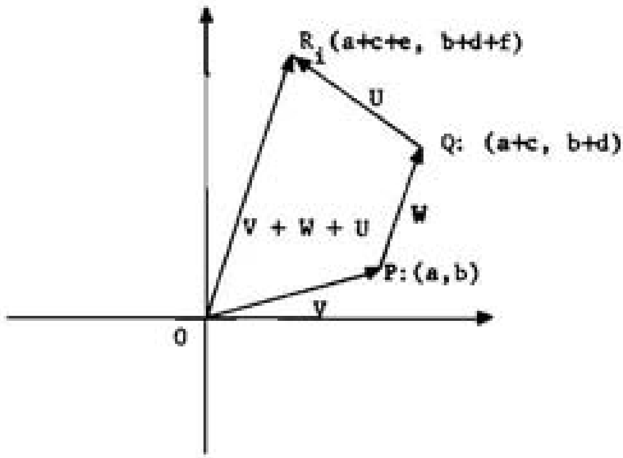


Figure 4–7. The sum V + W + U.

Order of addition does not affect the sum although it does affect the geometric representation, as indicated in Figure 4–8.

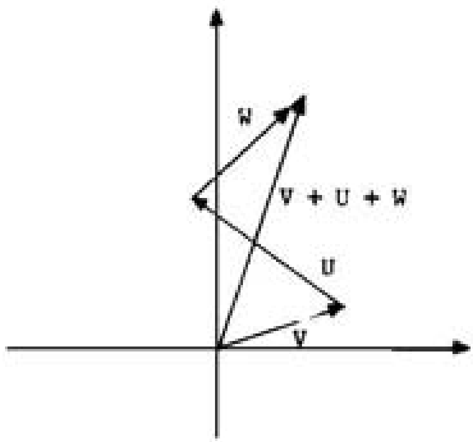


Figure 4–8. The sum V + U + W.

If V and W are parallel, the construction of the proposed representative of V + W is made in the same manner. The details will be left to the student.

<u>Theorem 4–4</u>. If the vectors V and W are represented by the directed

line segments $\overrightarrow{OP}$ and $\overrightarrow{PQ}$, respectively, then $V + W$ is represented by $\overrightarrow{OQ}$.

Since $V - W = V + (-W)$, the operation of subtracting one vector from another offers no essentially new geometric idea, once the construction of $-W$ is understood. Figure 4-9 illustrates the construction of the geometric vector representing $V - W$. It is useful to note, however, that since

$$||V - W|| = \left\|\begin{bmatrix} u - r \\ v - s \end{bmatrix}\right\| = \sqrt{(u - r)^2 + (v - s)^2},$$

the length of the vector $V - W$ equals the distance between the points P: (u, v) and T: (r, s).

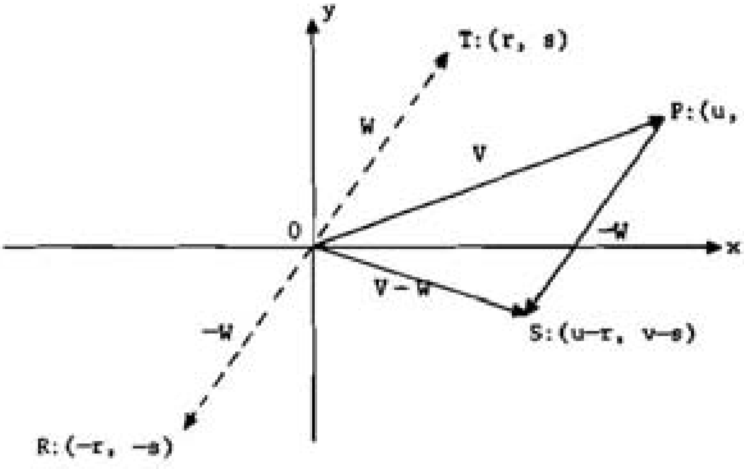


Figure 4-9. The subtraction of vectors, $V - W$.

Exercises 4-4

1. Determine graphically the sum and difference of the following pairs of vectors. Does order matter in constructing the sum? the difference?

   (a) $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$, $\begin{bmatrix} 7 \\ 1 \end{bmatrix}$; (d) $\begin{bmatrix} 3 \\ 6 \end{bmatrix}$, $2\begin{bmatrix} -1 \\ 2 \end{bmatrix}$;

   (b) $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\begin{bmatrix} -5 \\ 2 \end{bmatrix}$; (e) $\begin{bmatrix} 7 \\ 4 \end{bmatrix}$, $\begin{bmatrix} -2 \\ -3 \end{bmatrix}$;

   (c) $\begin{bmatrix} 7 \\ 2 \end{bmatrix}$, $\begin{bmatrix} -2 \\ -3 \end{bmatrix}$; (f) $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

2. Illustrate graphically the associative law:

$$(V + W) + U = V + (W + U).$$

3. Compute each of the following graphically:

(a) $\begin{bmatrix} 3 \\ 7 \end{bmatrix} + \begin{bmatrix} 2 \\ -1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

(b) $\begin{bmatrix} 4 \\ -2 \end{bmatrix} + \begin{bmatrix} -3 \\ 2 \end{bmatrix} + \begin{bmatrix} -1 \\ 0 \end{bmatrix}$,

(c) $\begin{bmatrix} 5 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 4 \\ 1 \end{bmatrix} + \begin{bmatrix} -2 \\ -3 \end{bmatrix}$,

(d) $\begin{bmatrix} -1 \\ 1 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} - \begin{bmatrix} 1 \\ 5 \end{bmatrix}$.

4. State the geometric significance of the following equations:

(a) $V + W = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$,

(b) $V + W + U = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$,

(c) $V + W + U + T = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

5. Complete the proof of both parts of Theorem 4-4.

4-5. The Inner Product of Two Vectors

Thus far in our development, we have investigated a geometrical interpretation for the algebra of vectors. We have represented column vectors of order 2 as arrows in the plane, and have established a one-to-one correspondence between this set of column vectors and the set of directed line segments from the origin of a coordinate plane. The algebraic operations of addition of two vectors and of multiplication of a vector by a number have acquired geometrical significance.

But we can also reverse our point of view and see that the geometry of vectors can lead us to the consideration of additional algebraic structure.

For instance, if you look at the pair of arrows drawn in Figure 4-10, you may comment that they appear to be mutually perpendicular. You have begun to

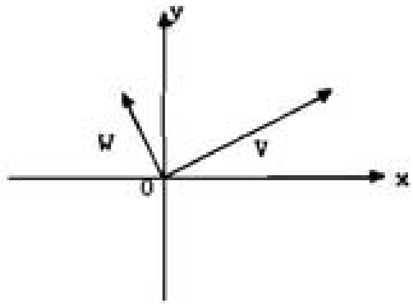


Figure 4-10. Perpendicular vectors.

talk about the angle between the pair of arrows. Since our vectors are located vectors, the following definition is needed.

<u>Definition 4-2</u>. The angle between two vectors is the angle between the standard geometric representations of the vectors.

Let us suppose, in general, that the points P, with coordinates (a,b), and R, with coordinates (c,d), are the terminal points of two geometric vectors with initial points at the origin. Consider the angle POR, which we denote by the Greek letter θ, in the triangle POR of Figure 4-11.

We can compute the cosine of θ by applying the law of cosines to the triangle POR. If $|OP|$, $|OR|$, and $|PR|$ are the lengths of the sides of the triangle, then by the law of cosines we have

$$2|OP|\ |OR| \cos \theta = |OP|^2 + |OR|^2 - |PR|^2.$$

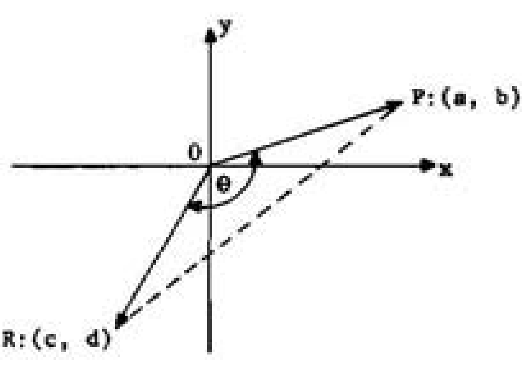


Figure 4-11. The angle between two vectors.

But

$$|OP| = \sqrt{a^2 + b^2},$$

$$|OR| = \sqrt{c^2 + d^2},$$

$$|PR| = \sqrt{(a-c)^2 + (b-d)^2}.$$

Thus,

$$2(\sqrt{a^2 + b^2})(\sqrt{c^2 + d^2}) \cos \theta = (a^2 + b^2) + (c^2 + d^2) - \left\{ (a-c)^2 + (b-d)^2 \right\}$$
$$= 2(ac + bd).$$

Hence,

$$|OP|\ |OR| \cos \theta = ac + bd. \qquad (1)$$

The number on the right-hand side of this equation, although clearly a function of the two vectors, has not heretofore appeared explicitly. Let us give it a name and introduce, thereby, a new binary operation for vectors.

Definition 4-3. The inner product of the vectors

$$\begin{bmatrix} a \\ b \end{bmatrix} \text{ and } \begin{bmatrix} c \\ d \end{bmatrix}, \quad \text{written} \quad \begin{bmatrix} a \\ b \end{bmatrix} \bullet \begin{bmatrix} c \\ d \end{bmatrix},$$

is the algebraic sum of the products of corresponding entries. Symbolically,

$$\begin{bmatrix} a \\ b \end{bmatrix} \bullet \begin{bmatrix} c \\ d \end{bmatrix} = ac + bd.$$

We can similarly define the inner product of two row vectors: $\begin{bmatrix} a & b \end{bmatrix} \bullet \begin{bmatrix} c & d \end{bmatrix} = ac + bd$.

Another name for the inner product of two vectors is the "dot product" of the vectors. You notice that the inner product of a pair of vectors is simply a number. In Chapter 1, you met the product of a row vector by a column vector, say $\begin{bmatrix} a & b \end{bmatrix}$ times $\begin{bmatrix} c \\ d \end{bmatrix}$, and found that

$$\begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} ac + bd \end{bmatrix},$$

the product being a 1×1 matrix. As you can observe, these two kinds of products are closely related; for, if V and W are the respective vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$, we have $V^t = \begin{bmatrix} a & b \end{bmatrix}$ and

$$V^tW = \begin{bmatrix} ac + bd \end{bmatrix} = \begin{bmatrix} V \bullet W \end{bmatrix} .$$

Later we shall exploit this close connection between the two products in order to deduce the algebraic properties of the inner product from the known properties of the matrix product.

Using the notion of the inner product and the formula (1) obtained above, we can state another theorem. We shall speak of the cosine of the angle included between two column (or row) vectors, although we realize that we are actually referring to an angle between associated directed line segments.

Theorem 4-5. The inner product of two vectors equals the product of the lengths of the vectors by the cosine of their included angle. Symbolically,

$$V \bullet W = \|V\| \ \|W\| \cos \theta,$$

where θ is the angle between the vectors V and W.

Theorem 4-5 has been proved in the case in which V and W are not parallel vectors. If we agree to take the measure of the angle between two parallel vectors to be 0° or 180° according as the vectors have the same or opposite directions, the result still holds. Indeed, as you may recall, the law of cosines on which the burden of the proof rests remains valid even when the three vertices of the "triangle" POR are collinear (Figures 4-12 and 4-13).

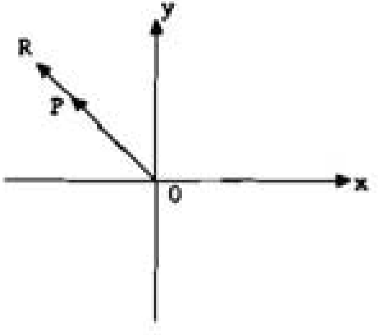


Figure 4-12. Collinear vectors in the same direction.

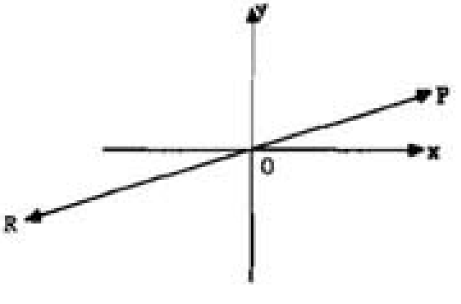


Figure 4-13. Collinear vectors in opposite directions.

Corollary 4-5-1. The relationship

$$V \bullet V = ||V||^2$$

holds for every vector V.

The corollary follows at once from Theorem 4-5 by taking $V = W$, in which case $\theta = 0^\circ$. To be sure, the result also follows immediately from the facts that, for any vector $V = \begin{bmatrix} a \\ b \end{bmatrix}$, we have

$$V \bullet V = a^2 + b^2, \quad \text{while} \quad ||V|| = \sqrt{a^2 + b^2}.$$

Two column vectors V and W are said to be orthogonal if the arrows OP and OR representing them are perpendicular to each other. In particular, the null vector $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is orthogonal to every vector. Since

$$\cos 90^\circ = \cos 270^\circ = 0,$$

we have the following result:

Corollary 4-5-2. The vectors V and W are orthogonal if and only if

$$V \bullet W = 0.$$

You will note that the condition $V \bullet W = 0$ is automatically satisfied if either V or W is the null vector.

We have examined some of the geometrical facets of the inner product of two vectors, but let us now look at some of its algebraic properties. Does it satisfy commutative, associative, or other algebraic laws we have met in studying number systems?

We can show that the commutative law holds, that is,

$$V \bullet W = W \bullet V.$$

For if V and W are any pairs of 2×1 matrices, a computation shows that

$$V^tW = W^tV,$$

But

$$V^tW = \begin{bmatrix} V \bullet W \end{bmatrix}, \quad \text{while} \quad W^tV = \begin{bmatrix} W \bullet V \end{bmatrix}.$$

Hence

$$V \bullet W = W \bullet V.$$

It is equally possible to show that the associative law cannot hold for inner products. Indeed, the products $V \bullet (W \bullet U)$ and $(V \bullet W) \bullet U$ are meaningless. To evaluate $V \bullet (W \bullet U)$, for example, you are asked to find the inner product of the vector V with the number $W \bullet U$. But the inner product is defined for two row vectors or two column vectors and not for a vector and a number. Incidentally, the product $V(W \bullet U)$ should not be confused with the meaningless $V \bullet (W \bullet U)$. The former product has meaning, for it is the product of the vector V by the number $W \bullet U$.

In the exercises that follow, you will be asked to consider some of the other possible properties of the inner product. In particular, you will be asked to establish the following theorem, the first part of which was proved above.

<u>Theorem 4-6</u>. If V, W, and U are column vectors of order 2, and r is a real number, then

(a) $V \bullet W = W \bullet V$,

(b) $(rV) \bullet W = r(V \bullet W)$,

(c) $V \bullet (W + U) = V \bullet W + V \bullet U$,

(d) $V \bullet V \geq 0$; and

(e) if $V \bullet V = 0$, then $V = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

<u>Exercises 4-5</u>

1. Compute the cosine of the angle between the two vectors in each of the following pairs:

(a) $\begin{bmatrix} 3 \\ -2 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$; (c) $\begin{bmatrix} -3 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 6 \\ -2 \end{bmatrix}$;

(b) $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$; (d) $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 3 \\ -1 \end{bmatrix}$;

(e) $\begin{bmatrix} -6 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 12 \end{bmatrix}$; (g) $\begin{bmatrix} 2 \\ 5 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 2 \end{bmatrix}$;

(f) $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$; (h) $\begin{bmatrix} 2t \\ t \end{bmatrix}$, $\begin{bmatrix} -t \\ 2t \end{bmatrix}$.

In which cases, if any, are the vectors orthogonal? In which cases, if any, are the vectors parallel?

2. Let

$$E_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad E_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Show that, for every nonzero vector V,

$$\frac{V \bullet E_1}{||V||} \quad \text{and} \quad \frac{V \bullet E_2}{||V||}$$

are the direction cosines of V.

3. (a) Prove that two vectors V and W are parallel if and only if

$$V \bullet W = \pm ||V|| \ ||W||.$$

Explain the significance of the sign of the right-hand side of this equation.

(b) Prove that

$$(V \bullet W)^2 \leq ||V||^2 \ ||W||^2$$

and write this inequality in terms of the entries of V and W.

(c) Show also that $V \bullet W \leq ||V|| \ ||W||$.

4. Show that if V is the null vector then

$$V \bullet W = 0.$$

5. Fill in the blanks in the following statements so as to make the resulting sentences true:

(a) The vectors $\begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\begin{bmatrix} -10 \\ \_ \end{bmatrix}$ are parallel.

(b) The vectors $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -6 \\ - \end{bmatrix}$ are orthogonal.

(c) The vectors $\begin{bmatrix} 4 \\ -3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ -4 \end{bmatrix}$ are__________.

(d) The vectors $\begin{bmatrix} -18 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} - \\ 12 \end{bmatrix}$ are parallel.

(e) For every positive real number t, the vectors

$$\begin{bmatrix} 3t \\ 2t \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ - \end{bmatrix} \text{ are orthogonal.}$$

(f) For every negative real number t, the vectors

$$\begin{bmatrix} 3t \\ 2t \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ - \end{bmatrix} \text{ are orthogonal.}$$

6. Verify that parts (a) – (d) of Theorem 4–6 are true if

$$U = \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \quad W = \begin{bmatrix} 2 \\ -3 \end{bmatrix}, \quad V = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \text{ and } r = 4.$$

7. Prove Theorem 4–6

(a) by using the definition of the inner product of two vectors;

(b) by using the fact that the matrix product V^tW satisfies the equation

$$V^tW = \begin{bmatrix} V \bullet W \end{bmatrix}.$$

8. Prove that $||V + W||^2 = (V + W) \bullet (V + W) = ||V||^2 + 2\, V \bullet W + ||W||^2$ for every pair of vectors V and W.

9. Show that, in each of the following sets of vectors, V and W are orthogonal, V and T are parallel, and T and W are orthogonal:

(a) $V = \begin{bmatrix} 1 \\ -2 \end{bmatrix}, \quad T = \begin{bmatrix} -4 \\ 8 \end{bmatrix}, \quad W = \begin{bmatrix} -6 \\ -3 \end{bmatrix};$

(b) $V = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \quad T = \begin{bmatrix} 14 \\ 21 \end{bmatrix}, \quad W = \begin{bmatrix} -3 \\ 2 \end{bmatrix}.$

Do the same relationships hold for the set

$$V = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad T = \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \quad W = \begin{bmatrix} 2 \\ 7 \end{bmatrix}?$$

10. Let V be a nonzero vector. Suppose that W and V are orthogonal, while T and V are parallel. Show that W and T are then orthogonal.

11. Show that, for every set of real numbers r, s, and t, the vectors

$$\begin{bmatrix} r \\ s \end{bmatrix} \text{ and } t\begin{bmatrix} -s \\ r \end{bmatrix} \text{ are orthogonal.}$$

12. Let $V = \begin{bmatrix} u \\ v \end{bmatrix}$, where V is not the zero vector. Show that if W and V are orthogonal, there exists a real number t such that

$$W = t\begin{bmatrix} -v \\ u \end{bmatrix}.$$

13. Show that the vectors V and W are orthogonal if and only if

$$||V + W||^2 - ||V - W||^2 = 0.$$

14. Show that if $A = \begin{bmatrix} a \\ b \end{bmatrix}$ and $B = \begin{bmatrix} c \\ d \end{bmatrix}$, then

$$||A||^2 \, ||B||^2 - (A \bullet B)^2 = (ad - bc)^2.$$

15. Show that the vectors V and W are orthogonal if and only if

$$(V + W) \bullet (V + W) = V \bullet V + W \bullet W.$$

16. Show that the equation

$$(V + W) \bullet (V - W) = V \bullet V - W \bullet W$$

holds for all vectors V and W.

17. Show that the inequality

$$||V + W|| \leq ||V|| + ||W||$$

holds for all vectors V and W.

4-6. Geometric Considerations

In Section 4-4, we saw that two parallel vectors determine a parallelogram. That is, if

$$A = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} c \\ d \end{bmatrix}$$

are two nonparallel vectors with initial points at the origin, then the points P:(a,b), O:(0,0), R:(c,d) and S:(a + c, b + d) are the vertices of a parallelogram (Figure 4-14.) A reasonable question to ask is: "How can we determine the area of the parallelogram PORS?"

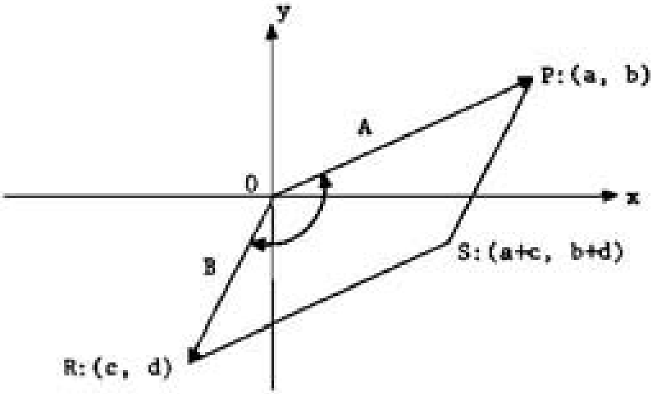


Figure 4-14. A parallelogram determined by vectors.

As you recall, the area of a parallelogram equals the product of the lengths of its base and its altitude. Thus, in Figure 4-15, the area of the parallelogram KLMN is b_1h, where b_1 is the length of side NM and h is the length of the altitude KD.

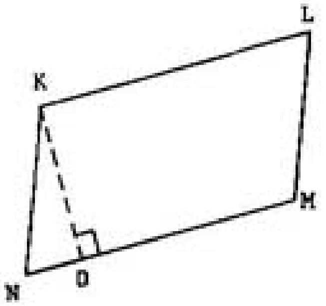


Figure 4-15. Determination of the area of a parallelogram.

But if b_2 is the length of side NK, and θ is the measure of either angle NKL or angle KNM, we have

$$h = b_2 \ |\sin \theta| .$$

Hence, the area of the parallelogram equals $b_1 b_2 \ |\sin \theta|$.

Returning to Figure 4-14 and letting θ be the angle between the vectors A and B, we can now say that if G is the area of parallelogram PQRS, then

$$G^2 = ||A||^2 \ ||B||^2 \sin^2 \theta.$$

Now

$$\sin^2 \theta = 1 - \cos^2 \theta.$$

It follows from Theorem 4-5 that

$$\cos^2 \theta = \frac{(A \bullet B)^2}{||A||^2 \ ||B||^2} ;$$

therefore,

$$\sin^2 \theta = \frac{||A||^2 \ ||B||^2 - (A \bullet B)^2}{||A||^2 \ ||B||^2} .$$

Thus, we have

$$G^2 = ||A||^2 \ ||B||^2 - (A \bullet B)^2.$$

It follows from the result of Exercise 14 of the preceding section that

$$G^2 = (ad - bc)^2.$$

Therefore,

$$G = |ad - bc| .$$

But $ad - bc$ is the value of the determinant $\delta(D)$, where D is the matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$. For easy reference, let us write our result in the form of a theorem.

Theorem 4-7. The area of the parallelogram determined by the standard representation of the vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ equals $|\delta(D)|$, where $D = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$.

Corollary 4-7-1. The vectors $\begin{bmatrix} a \\ b \end{bmatrix}$ and $\begin{bmatrix} c \\ d \end{bmatrix}$ are parallel if and only if $\delta(D) = 0$.

The argument proving the corollary is left as an exercise for the student.

You notice that we have been led to the determinant of a 2×2 matrix in examining a geometrical interpretation of vectors. The role of matrices in this interpretation will be further investigated in Chapter 5.

From geometric considerations, you know that the cosine of an angle cannot exceed 1 in absolute value,

$$|\cos \theta| \leq 1,$$

and that the length of any side of a triangle cannot exceed the sum of the lengths of the other two sides,

$$OQ \leq OP + PQ.$$

Accordingly, by the geometric interpretation of column vectors, for any

$$V = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} c \\ d \end{bmatrix}$$

we must have

$$\left| \frac{V \cdot W}{(V \cdot V)^{1/2} \, (W \cdot W)^{1/2}} \right| \leq 1 \tag{1}$$

and

$$\|V + W\| \leq \|V\| + \|W\|. \tag{2}$$

But can these inequalities be established algebraically? Let us see.

The inequality (1) is equivalent to

$$(V \bullet W)^2 \leq (V \bullet V)(W \bullet W),$$

that is,

$$(ac + bd)^2 \leq (a^2 + b^2)(c^2 + d^2),$$

or — as we see when we multiply out and simplify — to

$$2\ abcd \leq a^2d^2 + b^2c^2.$$

But this can be written as

$$0 \leq (ad - bc)^2,$$

which certainly is valid since the square of any real number is nonnegative.

Since $ad - bc = \delta(D)$, where

$$D = \begin{bmatrix} a & c \\ b & d \end{bmatrix},$$

you can see that the foregoing result is consistent with Corollary 4-7-1, above; that is, the sign of equality holds in (1) if and only if the vectors V and W are parallel.

As for the inequality (2), it can be written as

$$\sqrt{(a + c)^2 + (b + d)^2} \leq \sqrt{a^2 + b^2} + \sqrt{c^2 + d^2},$$

which simplifies to

$$2\ ac + 2\ bd \leq 2\sqrt{a^2 + b^2}\ \sqrt{c^2 + d^2}\ ,$$

which again is valid since

$$0 \leq (ad - bc)^2.$$

This time, the sign of equality holds if and only if the vectors V and W are parallel and

$$ac + bd \geq 0,$$

that is, if and only if the vectors V and W are parallel and in the same direction.

If you would like to look further into the study of inequalities and their applications, you might consult the SMSG Monograph, "An Introduction to Inequalities," by E. F. Beckenbach and R. Bellman.

Exercises 4-6

1. Let $\overrightarrow{OP}$ represent the vector A, and $\overrightarrow{OT}$ the vector B. Determine the area of triangle TOP if

   (a) $A = \begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\quad B = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$;

   (b) $A = \begin{bmatrix} -1 \\ 4 \end{bmatrix}$, $\quad B = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$;

   (c) $A = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\quad B = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$.

2. Compute the area of the triangle with vertices:

   (a) $(0,0)$, $(1,3)$, and $(-3,1)$;

   (b) $(0,0)$, $(5,2)$, and $(-10,-4)$;

   (c) $(1,0)$, $(0,1)$, and $(2,3)$;

   (d) $(1,1)$, $(2,2)$, and $(0,5)$;

   (e) $(1,2)$, $(-1,3)$, and $(1,0)$.

3. Verify the inequalities

$$(V \bullet W)^2 \leq (V \bullet V)(W \bullet W)$$

   and

$$||V + W|| \leq ||V|| + ||W||$$

   for the vectors

(a) $V = (3,4)$ and $W = (5,12)$,

(b) $V = (2,1)$ and $W = (4,2)$,

(c) $V = (-2,-1)$ and $W = (4,2)$.

4-7. Vector Spaces and Subspaces

Thus far our discussion of vectors has been concerned essentially with individual vectors and operations on them. In this section we shall take a broader point of view.

It will be convenient to have a symbol for the set of 2×1 matrices. Thus we let

$$H = \left\{ \begin{bmatrix} u \\ v \end{bmatrix} : \quad u \in R \quad \text{and} \quad v \in R \right\},$$

where R is the set of real numbers. The set H together with the operations of addition of vectors and of multiplication of a vector by a real number is an example of an algebraic system called a vector space.

Definition 4-4. A set of elements is a vector space over the set R of real numbers provided the following conditions are satisfied:

(a) The sum of any two elements of the set is also an element of the set.

(b) The product of any element of the set by a real number is also an element of the set.

(c) The laws I and II of Theorem 4-2 hold.

In applying laws I and II, $\underline{0}$ will denote the zero element of the vector space. Let us emphasize, however, that the elements of a vector space are not necessarily vectors in the sense thus far discussed in this chapter; for example, the set of 2×2 matrices, together with ordinary matrix addition and multiplication by a number, forms a vector space.

Since a vector space consists of a set together with the operations of addition of elements and of multiplication of elements by real numbers, strictly speaking we should not use the same symbol for the set of elements and for the vector space. But the practice is not likely to cause confusion and will be followed.

A completely trivial example of a vector space over R is the set consisting of the zero vector $\begin{bmatrix}0\\0\end{bmatrix}$ alone. Another vector space over R is the set of vectors parallel to $\begin{bmatrix}2\\3\end{bmatrix}$, that is, the set

$$\left\{ r \begin{bmatrix}2\\3\end{bmatrix} : r \in R \right\}.$$

It is evident that we are concerned with subsets of H in these two examples. Actually, these subsets are <u>subspaces</u>, in accordance with the following definition.

<u>Definition 4-5</u>. Any nonempty subset F of H is a <u>subspace</u> of H provided the following conditions are satisfied:

The sum of any two elements of F is also an element of F.

The product of any element of F by a real number is an element of F.

By definition, a subspace must contain at least one element V and also must contain each of the products rV for real numbers r. Every subspace of H therefore has the zero vector as an element, since for r = 0 we have

$$0V = \begin{bmatrix}0\\0\end{bmatrix}.$$

It is easy to see that the set consisting of the zero vector alone is a subspace. We can also verify that the set of all vectors parallel to any given nonzero vector is a subspace. Other than H itself, subsets of these two types are the only subspaces of H.

<u>Theorem 4-8</u>. Every subspace of H consists of exactly one of the following: the zero vector; the set of vectors parallel to a nonzero vector; the space H itself.

<u>Proof</u>. If F is a subspace containing only one vector, then

$$F = \left\{ \begin{bmatrix}0\\0\end{bmatrix} \right\},$$

since the zero vector belongs to every subspace.

If F contains a nonzero vector V, then F contains all vectors rV for real r. Accordingly, if all vectors of F are parallel to V, it follows that

$$F = \left\{rV: r \in R\right\}.$$

If F also contains a vector W not parallel to V, then F is actually equal to H, as we shall now prove.

Let

$$V = \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} c \\ d \end{bmatrix}$$

be nonparallel vectors in the subspace F, and let

$$Z = \begin{bmatrix} r \\ s \end{bmatrix}$$

be any other vector of H. We shall show that Z is a member of F.

By the definition of subspace, this will be the case if there are numbers x and y such that

$$xV + yW = Z, \tag{1}$$

that is,

$$x \begin{bmatrix} a \\ b \end{bmatrix} + y \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} r \\ s \end{bmatrix}.$$

These can be found if we can solve the system

$$ax + cy = r,$$
$$bx + dy = s,$$

for x and y in terms of the known numbers, a, b, c, d, r, and s. But since V and W are not parallel, it follows (see Corollary 4-7-1) that $ad - bc \neq 0$ and therefore the equations have the solution

$$x = \frac{dr - cs}{ad - bc} \quad \text{and} \quad y = \frac{as - br}{ad - bc}.$$

Since F is a subspace that contains V and W, it contains xV, yW, and their sum Z. Thus every vector Z of H must belong to F; that is, H is a subset of F. But F is given to be a subset of H. Accordingly, F = H.

Using the ideas of Section 4-4, we can give a geometric interpretation to Equation (1). Let the nonparallel vectors V and W in the subspace F have standard representations $\overrightarrow{OP}$ and $\overrightarrow{OR}$, respectively; see Figure 4-16. Let Z be represented by $\overrightarrow{OT}$. Since $\overrightarrow{OP}$ and $\overrightarrow{OR}$ are not parallel, any line parallel to one of them must intersect the line containing the other. Draw the lines through T parallel to $\overrightarrow{OP}$ and $\overrightarrow{OR}$, and let S and Q be the points in which these lines intersect the lines containing $\overrightarrow{OR}$ and $\overrightarrow{OP}$ respectively. Then

$$\overrightarrow{OT} = \overrightarrow{OQ} + \overrightarrow{OS}.$$

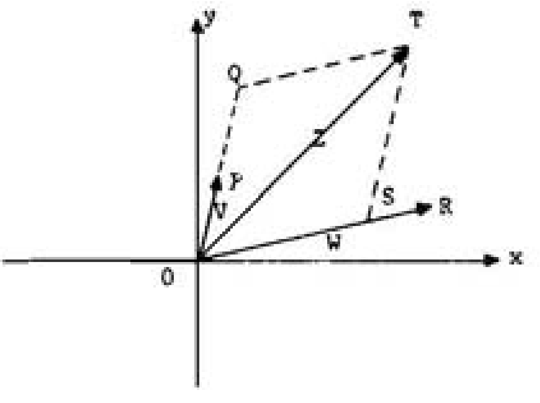


Figure 4-16. Representation of an arbitrary vector Z as a linear combination of a given pair of nonparallel vectors V and W.

But $\overrightarrow{OQ}$ is parallel to $\overrightarrow{OP}$ and $\overrightarrow{OS}$ to $\overrightarrow{OR}$. Therefore, there are real numbers x and y such that

$$\overrightarrow{OQ} = x\overrightarrow{OP} \quad \text{and} \quad \overrightarrow{OS} = y\overrightarrow{OR}.$$

Hence,

$$Z = xV + yW. \tag{1}$$

This ends our discussion of Theorem 4-7 and introduces the important concept of a <u>linear combination</u>:

<u>Definition 4-6</u>. If a vector Z can be expressed in the form $xV + yW$, where x and y are real numbers and V and W are vectors, then Z is called a <u>linear combination</u> of V and W.

Further, we have incidentally established the useful facts stated in the following theorems:

<u>Theorem 4-9</u>. A subspace F contains every linear combination of each pair of vectors in F.

<u>Theorem 4-10</u>. Each vector of H can be expressed as a linear combination of any given pair of nonparallel vectors in H.

For example, to express

$$Z = \begin{bmatrix} 5 \\ 10 \end{bmatrix}$$

as a linear combination of

$$V = \begin{bmatrix} 4 \\ 3 \end{bmatrix} \quad \text{and} \quad W = \begin{bmatrix} -3 \\ 4 \end{bmatrix},$$

we must determine real numbers x and y such that

$$\begin{bmatrix} 5 \\ 10 \end{bmatrix} = x \begin{bmatrix} 4 \\ 3 \end{bmatrix} + y \begin{bmatrix} -3 \\ 4 \end{bmatrix}$$
$$= \begin{bmatrix} 4x - 3y \\ 3x + 4y \end{bmatrix}.$$

Thus, we must solve the set of equations

$$5 = 4x - 3y,$$
$$10 = 3x + 4y.$$

We find the unique solution $x = 2$ and $y = 1$; that is, we have

$$Z = 2V + W.$$

If you observe that the given vectors V and W in the foregoing example are orthogonal, that is, $V \bullet W = 0$ (see Corollary 4-5-2 on page 156) then a second method of solution may occur to you. For if

$$Z = aV + bW,$$

then for the products $Z \bullet V$ and $Z \bullet W$ you have

$$Z \bullet V = a\|V\|^2 \quad \text{and} \quad Z \bullet W = b\|W\|^2.$$

But

$$Z \bullet V = 50,\ Z \bullet W = 25,\ \|V\|^2 = 25, \text{ and } \|W\|^2 = 25.$$

Hence,

$$50 = 25a \quad \text{and} \quad 25 = 25b\ ,$$

and accordingly

$$a = 2 \quad \text{and} \quad b = 1.$$

It is worth noting that the representation of a vector Z as a linear combination of two given nonparallel vectors is unique; that is, if the vectors V and W are not parallel, then for each vector Z the coefficients x and y can be chosen in exactly one way (Exercise 4-7-11, below) so that

$$Z = xV + yW.$$

The pair of nonparallel vectors V and W is called a <u>basis</u> for H, while the real numbers x and y are called the coordinates of Z relative to that basis. In the example above, the vector $\begin{bmatrix} 5 \\ 10 \end{bmatrix}$ has coordinates 2 and 1 relative to the basis $\begin{bmatrix} 4 \\ 3 \end{bmatrix}$ and $\begin{bmatrix} -3 \\ 4 \end{bmatrix}$.

In particular, the pair of vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is called the <u>natural basis</u> for H. This basis allows us to employ the coordinates of the point (u,v) associated with the vector $V = \begin{bmatrix} u \\ v \end{bmatrix}$ as the coordinates relative to the basis; thus,

$$V = \begin{bmatrix} u \\ v \end{bmatrix} = u \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Since every vector of H can be expressed as a linear combination of any pair of basic vectors, the basis vectors are said to <u>span</u> the vector space. The minimal number of vectors that span a vector space is called the <u>dimension</u> of the particular space.

For example, the dimension of the vector space H is 2. In the same sense, the set F

$$F = \left\{ r \begin{bmatrix} 2 \\ 3 \end{bmatrix} : r \in R \right\}$$

is a subspace of dimension 1. Note that neither $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ nor $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is a basis for this subspace. (What is?)

In a 2-space, that is, a vector space of dimension 2, it is necessary that any set of basis vectors be <u>linearly independent</u>.

<u>Definition 4-7</u>. Two vectors V and W are <u>linearly independent</u> if and only if, for all real numbers x and y, the equation

$$xV + yW = \underline{0}$$

implies $x = y = 0$. Otherwise, V and W are said to be <u>linearly dependent</u>.

For example, let $V = \begin{bmatrix} \sqrt{2} \\ -3 \end{bmatrix}$ and $W = \begin{bmatrix} -\frac{2}{3} \\ \sqrt{2} \end{bmatrix}$. Since

$$\sqrt{2} \begin{bmatrix} \sqrt{2} \\ -3 \end{bmatrix} + 3 \begin{bmatrix} -\frac{2}{3} \\ \sqrt{2} \end{bmatrix} = \underline{0},$$

V and W are linearly dependent. Note that

$$W = -\frac{\sqrt{2}V}{3},$$

which indicates that V and W are parallel.

Exercise 4-7

1. Express each of the following vectors as linear combinations of $\begin{bmatrix}2\\1\end{bmatrix}$ and $\begin{bmatrix}4\\3\end{bmatrix}$, and illustrate your answer graphically:

(a) $\begin{bmatrix}-2\\1\end{bmatrix}$, (d) $\begin{bmatrix}1\\0\end{bmatrix}$, (g) $\begin{bmatrix}-8\\-6\end{bmatrix}$,

(b) $\begin{bmatrix}1\\-2\end{bmatrix}$, (e) $\begin{bmatrix}0\\1\end{bmatrix}$, (h) $\begin{bmatrix}0\\0\end{bmatrix}$,

(c) $\begin{bmatrix}0\\-3\end{bmatrix}$, (f) $\begin{bmatrix}4\\2\end{bmatrix}$, (i) $\begin{bmatrix}-3\\4\end{bmatrix}$.

2. In parts (a) through (i) of Exercise 1, determine the coordinates of each of the vectors relative to the basis $\begin{bmatrix}2\\-1\end{bmatrix}$ and $\begin{bmatrix}2\\4\end{bmatrix}$.

3. Prove that the following set is a subspace of H:

$$S = \left\{ r\begin{bmatrix}2\\3\end{bmatrix} : r \in R \right\}.$$

4. Prove that, for any given vector W, the set $\{rW : r \in R\}$ is a subspace of H.

5. For

$$V = \begin{bmatrix}u\\v\end{bmatrix},$$

determine which of the following subsets of H are subspaces:

(a) all V with $u = 0$,
(b) all V with v equal to an integer,
(c) all V with u rational,
(d) all V with $2u - v = 0$,
(e) all V with $u + v = 2$,
(f) all V with $uv = 0$.

6. Prove that F is a subspace of H if and only if F contains every linear combination of two vectors in F.

7. Show that $\begin{bmatrix}3\\1\end{bmatrix}$ cannot be expressed as a linear combination of the vectors

$$\begin{bmatrix}-2\\5\end{bmatrix} \text{ and } \begin{bmatrix}6\\-15\end{bmatrix}.$$

8. Describe the set of all linear combinations of two given parallel vectors.

9. Let F_1 and F_2 be subspaces of H. Prove that the set F of all vectors belonging to both F_1 and F_2 is also a subspace.

10. In proving Theorem 4-10, we showed that if V and W are not parallel vectors, then each vector of H can be expressed as a linear combination of V and W. Prove the converse: If each vector of H has a representation as a linear combination of V and W, then V and W are not parallel.

11. Prove that if V and W are not parallel, then the representation of any vector Z in the form aV + bW is unique; that is, the coefficients a and b can be chosen in exactly one way.

12. Show that any vector

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

can be expressed uniquely as a linear combination of the basis vectors

$$\begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} -2 \\ -4 \\ 5 \end{bmatrix}.$$

4-8. Summary

In this chapter, we have developed a geometrical representation — namely, directed line segments — for 2×1 matrices, or column vectors. Guided by the definition of the algebraic operation of addition of vectors, we have found the "parallelogram law of addition" of directed line segments. The multiplication of a vector by a number has been represented by the expansion or contraction of the corresponding directed line segment by a factor equal to the number, with the sign of the factor determining whether or not the direction of the line segment is reversed. Thus, from a set of algebraic elements we have produced a set of geometric elements. Geometrical observations in turn led us back to additional algebraic concepts.

Also in this chapter, we have introduced the important concepts of a vector space and linear independence.

Since the nature of the elements of a vector space is not limited except by the postulates, the vector space does not necessarily consist of a set whose elements are 2×1 column vectors; thus its elements might be $n \times n$ matrices, real numbers, and so on.

For example let us look at the set P of linear and constant polynomials with real coefficients, that is, the set

$$P = \left\{p: p(x) = ax + b, \quad a,b \in R\right\},$$

under ordinary addition and multiplication by a number. The sum of two such polynomials,

$$(a_1x + b_1) + (a_2x + b_2) = (a_1 + a_2)x + (b_1 + b_2),$$

is an element of the set since the sums $a_1 + a_2$ and $b_1 + b_2$ are real numbers. The product of any element of P by a real number,

$$c(ax + b) = acx + bc,$$

is also a member of P since the products ac and bc are real numbers. We can similarly show that addition is commutative and associative, that there is an identity for addition, and that each element has an additive inverse; thus, the laws I of Theorem 4.2 are valid. In like fashion, we can demonstrate that laws II are satisfied: both distributive laws hold; the multiplication of an element by two real numbers is associative; the product of any element p by the real number 1 is p itself; the product of 0 and any element p is the zero element; and the product of any real number and the zero element is the zero element.

We have outlined the proof that the set P of linear and constant polynomials is a vector space. Thus the expression, "the vector, $ax + b$," is meaningful when we are speaking of the vector space P.

The mathematics to which our algebra has led us forms the beginnings of a discipline called "vector analysis," which is an important tool in classical and modern physics, as well as in geometry. The "free" vectors that you meet in physics, namely, forces, velocities, etc., can be represented by our geometric vectors. The study in which we are engaged is consequently of vital importance for physicists, engineers, and other applied scientists, as well as for mathematicians.

Chapter 5

TRANSFORMATIONS OF THE PLANE

5-1. Functions and Geometric Transformations

You have discovered that one of the most fundamental concepts in your study of mathematics is the notion of a function. In geometry the function concept appears in the idea of a transformation. It is the aim of this chapter to recall what we mean by a function, to define geometric transformation, and to explore the role of matrices in the study of a significant class of these transformations.

You recall that a function from set A to set B is a correspondence or mapping from the elements of the set A to those of the set B such that with each element of A there is associated exactly one element of B. The set A is the domain of the function and the subset of B onto which A is mapped is the range of the function. In your previous work, the functions you met generally had sets of real numbers both for domain and for range. Thus the function symbolized in the form

$$x \longrightarrow x^2$$

is likely to be interpreted as associating the nonnegative real number x^2 with the real number x. Here you have a simple example of a "real function" of a "real variable."

In Chapter 4, however, you met a function $V \longrightarrow \|V\|$ having for its domain the vector space H, and for its range the set of nonnegative real numbers.

In the present chapter, we shall consider functions that have their range as well as their domain in H. Specifically, we want to find a geometric interpretation for these "vector functions" of a "vector variable"; this is a continuation of the discussion started in Chapter 3. All vectors will now be considered in their standard representations $\overrightarrow{OP}$, so that they will be parallel if and only if represented by collinear geometric vectors.

Such a vector function will associate, with the point P having coordinates (x,y), a point P' with coordinates (x',y'). Or we may say that it maps the geometric vector $\overrightarrow{OP}$ onto the geometric vector $\overrightarrow{OP'}$. The function can, therefore, be viewed as a process that associates with each point P of the plane some point P' of this plane. We shall call this process a transformation of the plane into itself or a geometric transformation. As a matter of fact, these transformations are often called "point transformations" in contrast to more general mappings in which a point may be carried into a line, a circle, or some other geometric configuration. For us, a geometric transformation is a helpful means of visualizing a vector function of a vector variable. As a matter of convenient terminology, we shall call the vector that such a function associates with a given vector V the image of V; furthermore, we shall say that the function maps V onto its image.

Let us look at the simple function

$$V \longrightarrow 2V, \quad V \in H.$$

This function maps each vector V onto the vector that has the same direction as V, but that is twice as long as V. Another way of asserting this is to say that the function associates with each point P of the plane a point P' such that P and P' lie on the same ray from the origin, but

$$||\overrightarrow{OP'}|| = 2||\overrightarrow{OP}||;$$

see Figure 5-1. You may therefore think of the function in this example as uniformly stretching the plane by a factor 2 in all directions from the origin. (Under this mapping, what is the point onto which the origin is mapped?)

As a second example, consider the function

$$V \longrightarrow -V, \quad V \in H.$$

This time, each vector is mapped onto the vector having length equal and direction opposite to that of the given vector. Viewed as a point transformation, the function associates with any point P its "reflection" in the origin; see Figure 5-2.

The function

$$V \longrightarrow -2V$$

combines both of the effects of the preceding functions, so that the vector associated with V is twice as long as V, but has the opposite direction to that of V.

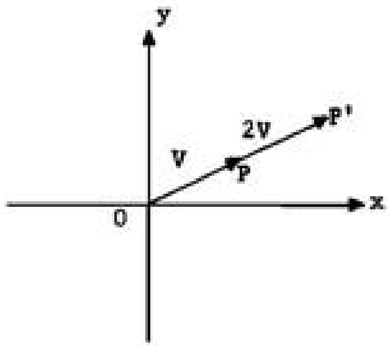


Figure 5-1. The transformation $V \longrightarrow 2V$.

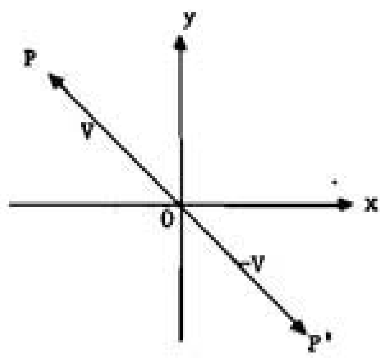


Figure 5-2. The transformation $V \longrightarrow -V$.

Now, let us look at the function

$$V \longrightarrow \|V\|\, V.$$

As in our first example, each vector is mapped by the function onto a vector having the same direction as the given vector. Indeed, every vector of length 1 is its own image. But if $\|V\| > 1$, then the image of V has a length greater than that of V, with the expansion factor increasing with the length of V itself. Thus, the vector

$$\begin{bmatrix} 2 \\ 0 \end{bmatrix},$$

having length 2, is mapped onto

$$\begin{bmatrix} 4 \\ 0 \end{bmatrix},$$

which is twice as long. The vector

$$\begin{bmatrix} 5 \\ 12 \end{bmatrix},$$

whose length is 13, has the image

$$\begin{bmatrix} 65 \\ 156 \end{bmatrix},$$

with length 169. On the other hand, for nonzero vectors of length less than 1, we obtain image vectors of shorter length, the contraction factor decreasing with decreasing length of the original vector. Thus,

$$\begin{bmatrix} -\frac{1}{2} \\ 0 \end{bmatrix} \quad \text{is mapped onto} \quad \begin{bmatrix} -\frac{1}{4} \\ 0 \end{bmatrix},$$

the image being half as long as the given vector. Again, the vector

$$\begin{bmatrix} -\frac{4}{7} \\ -\frac{3}{7} \end{bmatrix} \quad \text{is mapped onto} \quad \begin{bmatrix} -\frac{20}{49} \\ -\frac{15}{49} \end{bmatrix},$$

the length of the first vector being $5/7$, while the length of its image is only $(5/7)^2$, or $25/49$. Although we may try to think of this mapping as a kind of stretching of the plane in all directions from the origin, so that any point and its image are collinear with the origin, this mental picture has also to take into account the fact that the amount of expansion varies with the distance of a given point from the origin, and that for points within the circle of radius 1 about the origin the so-called stretching is actually a compression.

We have been considering transformations of individual vectors; let us look at certain transformations of the square ORST, determined by the basis vectors $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$. As shown in Figure 5-3, the function

$$V \longrightarrow -V$$

maps

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

respectively, onto

$$\begin{bmatrix} -1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ -1 \end{bmatrix}.$$

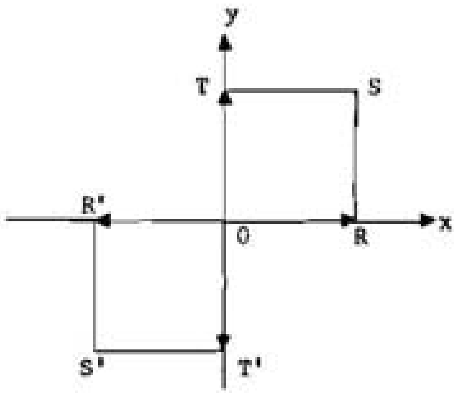


Figure 5-3. Reflection in the origin.

Another transformation that is readily visualized is the reflection in the x axis. For this mapping, the point (x,y) goes into the point (x,−y); see Figure 5-4.

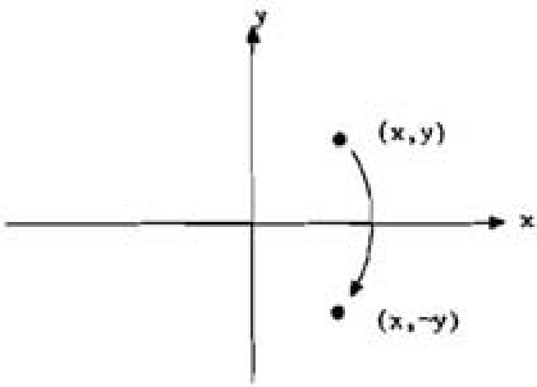


Figure 5-4. Reflection in the x axis.

That is, the map is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x \\ -y \end{bmatrix}.$$

Using a matrix, we may rewrite this result in the form

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

as is easily verified.

Now this transformation, applied to the square ORST, leaves the point (1,0) unchanged; thus the vector $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is mapped onto itself. The vector $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ is mapped onto $\begin{bmatrix} 0 \\ -1 \end{bmatrix}$.

Reflection in the y axis, or a rotation of 180° in space about the y axis, can be expressed similarly:

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} -x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

as shown in Figure 5-5.

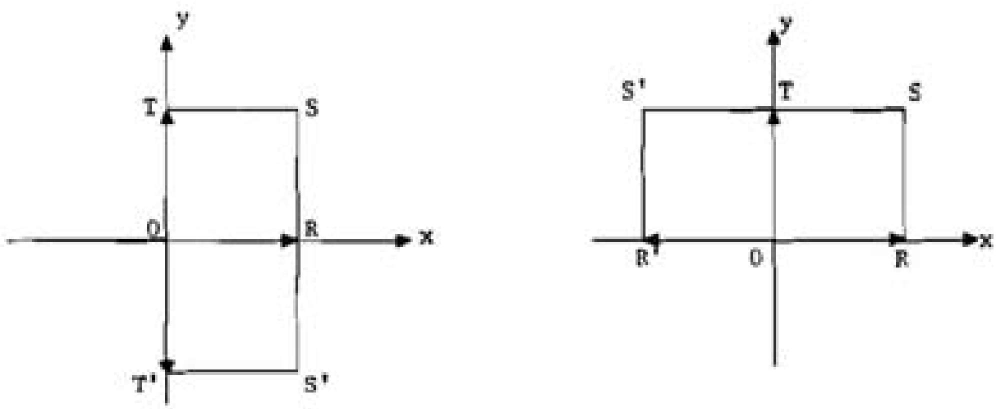


Figure 5-5. Rotation of 180° about the axes.

Casual observation (see Figures 5-5 and 5-6) may lead you to assume that a 180° rotation about the y axis and 90° rotation about the origin in the (x,y) plane are equivalent; they are not. The first transformation leaves the point (0,1) unchanged, whereas the second transformation maps (0,1) onto (-1,0). As a vector function, the 90° rotation with respect to the origin is expressed by

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} -y \\ x \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

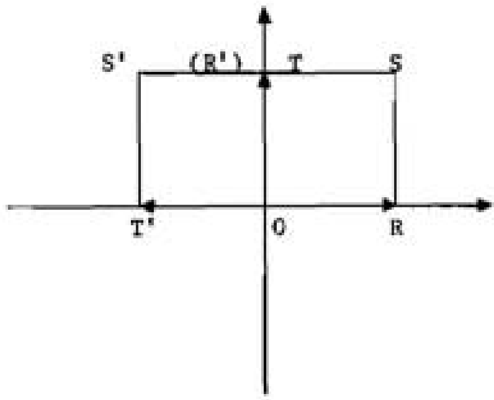


Figure 5-6. Rotation of 90° about the origin.

The transformations of Figures 5-2 through 5-6 have altered neither the size nor the shape of the square. The "stretching" function of Figure 5-1,

$$V \longrightarrow 2V,$$

does alter size. A "shear" that moves each point parallel to the x axis through a distance equal to twice the ordinate of the point alters shape. Consider the transformation

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x + 2y \\ y \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

which maps the basis vectors onto

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \text{ respectively.}$$

The result is a shearing that transforms the square into a parallelogram; see Figure 5-7. What does the stretching do to shape? The shear to size?

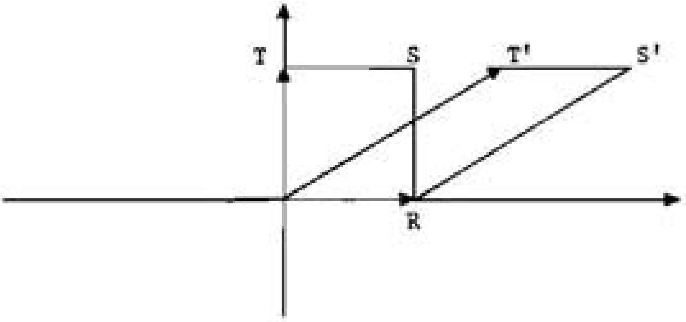


Figure 5-7. Shearing.

Another type of transformation involves a displacement or "translation" in the direction of a fixed vector. The mapping

$$V \longrightarrow V + U, \quad \text{where} \quad U = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

can be written in the form

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x + 3 \\ y + 2 \end{bmatrix}.$$

One way of visualizing this function is to regard it as translating the plane in the direction of the vector U through a distance equal to the length of U.

Transformation of two different types can be combined in one function. For instance, the mapping

$$V \longrightarrow \tfrac{1}{2}(V + U), \quad \text{where} \quad U = \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

involves a translation and then a compression. When the function is expressed in the form

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} \dfrac{x + 3}{2} \\ \dfrac{y + 2}{2} \end{bmatrix},$$

we recognize more easily that every point P is mapped onto the midpoint of the line segment joining P to the point $(3,2)$; see Figure 5-8.

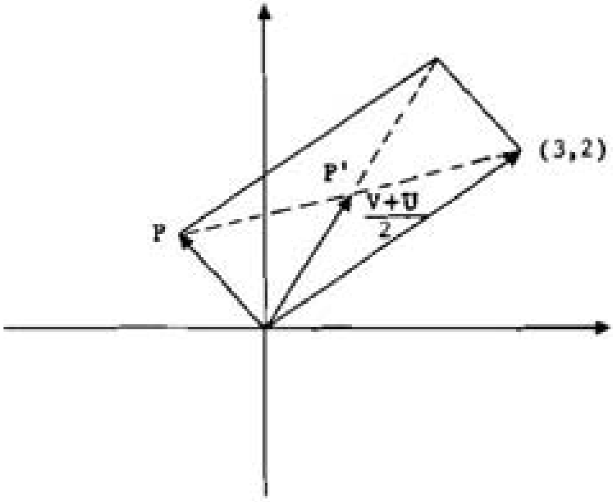


Figure 5-8. The transformation $V \longrightarrow \frac{1}{2}(V + U)$.

Under this mapping, the square ORST will likewise be translated toward the point (3,2) and then compressed by a factor $\frac{1}{2}$, as shown in Figure 5-9. This figure enables us to see that the points O, R, S, and T are mapped onto midpoints of lines connecting these points to (3,2).

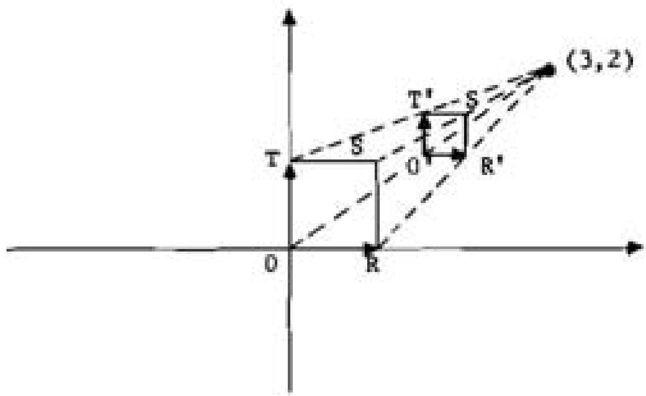


Figure 5-9. Translation and compression.

All the vector functions discussed above map distinct points of the plane onto distinct points. This is not always the case; we can certainly produce functions that do not have this property. Thus, the function

$$V \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

maps every point of the plane onto the origin. On the other hand, the transformation

$$V = \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x \\ 0 \end{bmatrix}$$

maps the point (x,y) onto the point of the x axis that has the same first component as V. For example, every point of the line $x = 3$ is mapped onto the point $(3,0)$. Since the image of each point P can be located by drawing a perpendicular line from P to the x axis, we may think of P as being carried or projected on the x axis by a line perpendicular to this axis. Consequently, this mapping may be described as a perpendicular or orthogonal projection of the plane on the x axis. You notice that these last two functions map H onto subspaces of H.

Since we have met examples of transformations that map distinct points onto distinct points and have also seen transformations under which distinct points may have the same image, it is useful to define a new term to distinguish between these two kinds of vector functions.

Definition 5-1. A transformation from the set H onto the set H is one-to-one provided that the images of distinct vectors are also distinct vectors.

Thus, if f is a function from H onto H and if we write $f(V)$ for the image of V under the transformation f, then Definition 5-1 can be formulated symbolically as follows: The function f is a one-to-one transformation of H provided that, for vectors V and U in H,

$$V \neq U$$

implies

$$f(V) \neq f(U).$$

Exercises 5-1

1. Find the image of the vector V under the mapping

$$V \longrightarrow 3V$$

for each of the following values of V:

(a) $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$, (c) $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$, (e) $\begin{bmatrix} 5 \\ 1 \end{bmatrix} + \begin{bmatrix} 7 \\ -3 \end{bmatrix}$,

(b) $\begin{bmatrix} 2 \\ -1 \end{bmatrix}$, (d) $\begin{bmatrix} 7 \\ -3 \end{bmatrix}$, (f) $\begin{bmatrix} 5 \\ 1 \end{bmatrix} - \begin{bmatrix} 7 \\ -3 \end{bmatrix}$.

2. Find f(V) under the mapping

$$f : V = \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} y \\ 0 \end{bmatrix}$$

for each of the following values of V:

(a) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, (c) $\begin{bmatrix} -1 \\ -3 \end{bmatrix}$, (e) $5 \begin{bmatrix} -1 \\ -3 \end{bmatrix}$,

(b) $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$, (d) $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$, (f) $-2 \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

3. Describe the geometric effect of each of the following transformations of H on the vector $V = \begin{bmatrix} x \\ y \end{bmatrix}$:

(a) $V \longrightarrow V$, (h) $V \longrightarrow \begin{bmatrix} x \\ -y \end{bmatrix}$,

(b) $V \longrightarrow \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, (i) $V \longrightarrow \begin{bmatrix} 2x \\ y \end{bmatrix}$,

(c) $V \longrightarrow aV,\ a > 0$, (j) $V \longrightarrow \begin{bmatrix} 3x \\ 3y \end{bmatrix}$,

(d) $V \longrightarrow -aV,\ a > 0$, (k) $V \longrightarrow \begin{bmatrix} x + y \\ y \end{bmatrix}$,

(e) $V \longrightarrow \begin{bmatrix} 0 \\ y \end{bmatrix}$, (l) $V \longrightarrow \begin{bmatrix} x \\ 2x + y \end{bmatrix}$,

(f) $V \longrightarrow \begin{bmatrix} y \\ y \end{bmatrix}$, (m) $V \longrightarrow \begin{bmatrix} x - 2y \\ y \end{bmatrix}$,

(g) $V \longrightarrow \begin{bmatrix} -x \\ y \end{bmatrix}$, (n) $V \longrightarrow \begin{bmatrix} x \\ y - 3x \end{bmatrix}$.

4. Determine which of the transformations in the preceding exercise are one-to-one.

5. Find expressions of the type $V \longrightarrow V'$ for the transformations of H that map each point P onto the point P' related to P in the ways described below:

(a) P' is one unit to the right of P and four units above P;

(b) P' is the perpendicular projection of P on the horizontal line through (3,2);

(c) P' is the perpendicular projection of P on the vertical line through (−1,−2);

(d) $\overrightarrow{OP}$ and $\overrightarrow{OP}'$ are collinear but opposite in direction, and $||\overrightarrow{OP}'|| = \frac{1}{2}\,||\overrightarrow{OP}||$;

(e) P' is the intersection of the horizontal line through P with the line of slope −1 passing through the origin (horizontal projection on the line $y = -x$);

(f) P' is the intersection of the vertical line through P with the line $y = 2x$ (vertical projection on the line $y = 2x$).

6. Show that the mapping of H into itself that sends each point P into the point of intersection of the line $y = x$ with the line through P having slope 2 is given by

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} 2x - y \\ 2x - y \end{bmatrix}.$$

7. (a) Show that the mapping

$$V = \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x + 2y \\ 4x + 3y \end{bmatrix}$$

can be expressed in the form

$$V \longrightarrow \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} V.$$

(b) Find the image under this transformation of $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(c) Find the image under this transformation of the subspace of vectors collinear with $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

8. Solve parts (b) and (c) of Exercise 7 when $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is replaced by

(a) $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, (c) $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$,

(b) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$, (d) $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$.

9. Under the transformation given in Exercise 7, find by two different methods the image of each of the following vectors:

(a) $\begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, (d) $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$,

(b) $\begin{bmatrix} 5 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 3 \end{bmatrix}$, (e) $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$,

(c) $\begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, (f) $\begin{bmatrix} 5 \\ 3 \end{bmatrix}$.

10. Consider the mapping

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

(a) Find the images under this mapping of the pair of points (5,1) and (1,−2), and show that the distance between the given pair of points equals the distance between their images.

(b) Solve part (a) if the given points are (−2,10) and (6,−5).

(c) Solve part (a) if the given points are (a,b) and (c,d).

5–2. Matrix Transformations

As noted earlier, especially in Chapter 3, the pair of equations

$$a_{11}\, x + a_{12}\, y = b_1,$$
$$a_{21}\, x + a_{22}\, y = b_2,$$

can be written in the form

$$AV = B,$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad V = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}.$$

Consequently, in solving the equations you actually determine all the vectors V that are mapped onto the particular vector B by the function

$$V \longrightarrow AV. \qquad (1)$$

The study of the solution of systems of linear equations thus leads to the consideration of the special class of transformations on H that are expressible in the form (1), where A is any 2×2 matrix with real entries. These matrix transformations constitute a very important class of mappings, having extensive applications in mathematics, statistics, physics, operations research, and engineering.

An important property of matrix transformations is that they are <u>linear mappings</u>; that is, they preserve vector sums and the products of vectors with real numbers.

Let us formulate these ideas explicitly.

<u>Definition 5-2</u>. <u>A linear transformation</u> on H is a function f from H into H such that

(a) for every pair of vectors V and U in H, we have

$$f(V + U) = f(V) + f(U);$$

(b) for every real number r and every vector V in H, we have

$$f(rV) = rf(V).$$

<u>Theorem 5-1</u>. Every matrix transformation is linear.

<u>Proof</u>. Let f be the transformation

$$f : V \longrightarrow AV,$$

where A is any real matrix of order 2. We must show that for any vectors V and U, we have

$$A(V + U) = AV + AU;$$

further, we must show that for any vector V and any real number r we have

$$A(rV) = r(AV).$$

But these equalities hold in virtue of parts III (a) and III (f) of Theorem 4-2 (see page 134).

The linearity property of matrix transformations can be used to derive the following result concerning transformations of the subspaces of H.

__Theorem 5-2__. A matrix A maps every subspace F of H onto a subspace F' of H.

__Proof__. Let F' denote the set of vectors

$$\{AU: U \in F\}.$$

To prove that F' is a subspace of H, we must show that the following statements are true:

(a) For any pair of vectors P', Q' in F', the sum $P' + Q'$ is in F'.

(b) For any vector P' in F' and any real number r, rP' is in F'.

If P' and Q' are in F', then they must be the images of vectors P and Q in F; that is,

$$P' = AP,$$
$$Q' = AQ.$$

It follows that

$$P' + Q' = AP + AQ = A(P + Q),$$

and $P' + Q'$ is the image of the vector $P + Q$ in F. (Can you tell why $P + Q$ is in F?) Hence, $(P' + Q') \in F'$. Similarly,

$$rP' = r(AP) = A(rP),$$

and hence rP' is the image of rP. But $rP \in F$ because F is a subspace. Thus, rP' is the image of a vector in F; therefore, $rP' \in F'$.

__Corollary 5-2-1__. Every matrix maps the plane H onto a subspace of H, either the origin, or a straight line through the origin, or H itself.

For example, to determine the subspaces onto which

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

maps

(a) $F = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : y = -3x \right\}$,

(b) H itself,

we proceed as follows.

For (a), the vectors of F are of the form

$$U = \begin{bmatrix} x \\ -3x \end{bmatrix} = x \begin{bmatrix} 1 \\ -3 \end{bmatrix}, \quad x \in R.$$

Hence,

$$AU = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \left(x \begin{bmatrix} 1 \\ -3 \end{bmatrix} \right) = x \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = x \begin{bmatrix} -2 \\ -1 \end{bmatrix} = -x \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Thus, F is mapped onto F', the set of vectors collinear with $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$; that is,

$$F' = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : y = \frac{1}{2} x \right\}.$$

In other words, A maps the line passing through the origin with slope -3 onto the line through the origin with slope 1/2.

As regards (b), we note that for any vector

$$V = \begin{bmatrix} x \\ y \end{bmatrix} \in H,$$

we have

$$AV = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 4x + 2y \\ 2x + y \end{bmatrix} = (2x + y) \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Since $2x + y$ assumes all real values as x and y run over the set of real numbers, it follows that H is also mapped onto F'; that is, A maps the entire plane onto the line

$$y = \frac{1}{2} x.$$

Exercises 5-2

1. Let $A = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}$. For each of the following values of the vector V,

(a) $V = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, (d) $V = \begin{bmatrix} 5 \\ -1 \end{bmatrix}$,

(b) $V = \begin{bmatrix} -3 \\ -2 \end{bmatrix}$, (e) $V = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$,

(c) $V = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, (f) $V = \begin{bmatrix} -6 \\ 4 \end{bmatrix}$,

determine:

(i) the vector into which A maps V,

(ii) the line onto which A maps the line containing V.

2. A certain matrix maps

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} \text{ into } \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ into } \begin{bmatrix} 4 \\ 5 \end{bmatrix}.$$

Using this information, determine the vector into which the matrix maps each of the following:

(a) $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $\left(\text{Hint: } \begin{bmatrix} 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix}.\right)$,

(b) $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$, (e) $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$,

(c) $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, (f) $\begin{bmatrix} 4 \\ 5 \end{bmatrix}$,

(d) $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$, (g) $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$.

3. Consider the following subspaces of H:

$$F_1 = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}, \qquad F_2 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : y = 2x \right\},$$

$$F_3 = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : y = -2x \right\}, \quad F_4 = H \text{ itself.}$$

Determine the subspaces onto which F_1, F_2, F_3, and F_4 are mapped by each of the following matrices:

(a) $A = \begin{bmatrix} -2 & 1 \\ -2 & 1 \end{bmatrix}$, (B) $B = \begin{bmatrix} 0 & 1 \\ -2 & 3 \end{bmatrix}$, (c) AB, (d) BA.

4. Let $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$.

(a) Calculate AV for

$$V = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} p \\ q \end{bmatrix}.$$

(b) Find the vector V for which

$$AV = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} r \\ s \end{bmatrix}.$$

5. Determine which of the following transformations of H are linear, and justify your answer:

(a) $V = \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} x+1 \\ y \end{bmatrix}$,

(b) $V \longrightarrow \begin{bmatrix} x \\ 1 \end{bmatrix}$,

(c) $V \longrightarrow \begin{bmatrix} x-y \\ x+y \end{bmatrix}$,

(d) $V \longrightarrow \begin{bmatrix} 2x \\ 5y \end{bmatrix}$,

(e) $V \longrightarrow \frac{1}{2}(V + U)$, where $U = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$,

(f) $V \longrightarrow \|V\|\, V$.

6. Prove that the matrix A maps the plane onto the origin if and only if

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

7. Prove that the matrix A maps every vector of the plane onto itself if and only if

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

8. Prove that

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

maps the line $y = 0$ onto itself. Is any point of that line mapped onto itself by this matrix?

9\. (a) Show that each of the matrices

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$

maps H onto the x axis.

(b) Determine the set of all matrices that map H onto the x axis. (Hint: You must determine all possible matrices A such that corresponding to each $V \in H$ there is a real number r for which

$$AV = r \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \tag{1}$$

In particular, (1) must hold for suitable r when V is replaced by $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and by $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.)

10\. Determine the set of all matrices that map H onto the y axis.

11\. (a) Determine the matrix A such that

$$AV = 2V$$

for all V.

(b) The mapping

$$V \longrightarrow aV \qquad (a > 0)$$

multiplies the lengths of all vectors without changing their directions. It amounts to a change of scale. The number a is accordingly called a scale factor or <u>scalar</u>. Find the matrix A that yields only a change of scale:

$$AV = aV.$$

12\. Prove that for every matrix A the set F of all vectors U for which

$$AU = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

is a subspace of H. This subspace is called the kernel of the mapping.

13. Prove that a transformation f of H into itself is linear if and only if

$$f(rV + sU) = r\,f(V) + s\,f(U)$$

for every pair of vectors V and U of H and every pair of real numbers r and s.

5-3. Linear Transformations

In the preceding section, we proved that every matrix represents a linear transformation of H into H. We now prove the converse: Every linear transformation of H into H can be represented by a matrix.

Theorem 5-3. Let f be a linear transformation of H into H. Then, relative to any given basis for H, there exists one and only one matrix A such that, for all $V \in H$,

$$AV = f(V).$$

Proof. We prove first that there cannot be more than one matrix representing f. Suppose that there are two matrices A and B such that, for all $V \in H$,

$$AV = f(V) \quad \text{and} \quad BV = f(V).$$

Then

$$AV - BV = f(V) - f(V)$$

for each V. Hence,

$$(A - B)V = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \text{for all} \quad V \in H.$$

Thus, A − B maps every vector onto the origin. It follows (Exercise 5-2-6) that A − B is the zero matrix; therefore,

$$A = B.$$

Hence, there is at most one matrix representation of f.

Next, we show how to find the matrix representation for the linear transformation f. Let S_1 and S_2 be a pair of noncollinear vectors of H. Let

$$f(S_1) = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} \quad \text{and} \quad f(S_2) = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

be the respective images of S_1 and S_2 under the mapping f. If V is any vector of H, it follows from Theorem 4-10 that there exist real numbers v_1 and v_2 such that $V = v_1S_1 + v_2S_2$. Since f is a linear transformation, we have

$$f(V) = f(v_1S_1 + v_2S_2) = v_1f(S_1) + v_2f(S_2).$$

Accordingly,

$$f(V) = v_1 \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} + v_2 \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} a_{11}v_1 + a_{12}v_2 \\ a_{21}v_1 + a_{22}v_2 \end{bmatrix} .$$

Thus,

$$f(V) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} .$$

It follows that f is represented by the matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

when vectors are expressed in terms of their coordinates relative to the basis S_1, S_2.

You notice that the matrix A is completely determined by the effect of f on the pair of noncollinear vectors used as the basis for H. Thus, once you know that a given transformation on H is linear, you have a matrix representing the mapping when you have the images of the natural basis vectors,

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix} .$$

For example, it can be shown by a geometric argument that the counterclockwise rotation of the plane through an angle of 30° about the origin is a linear transformation. This function maps any point P onto the point P', where the measure of the angle POP' is equal to 30° (Figure 5-10). It is easy to see (Figure 5-11) that

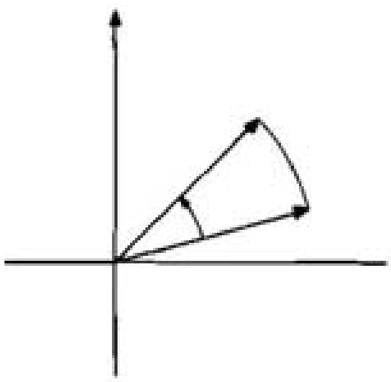


Figure 5-10. A rotation through an angle of 30° about the origin.

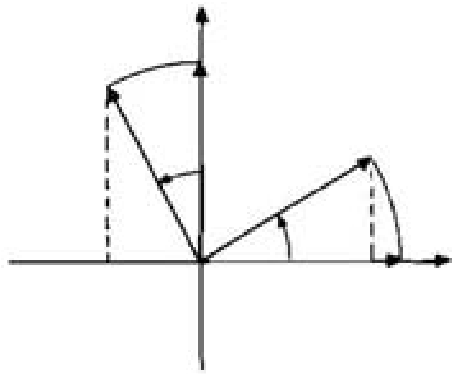


Figure 5-11. The images of the points (1,0) and (0,1) under a rotation of 30° about the origin.

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{is mapped onto} \quad \begin{bmatrix} \cos 30^\circ \\ \sin 30^\circ \end{bmatrix}$$

and

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{is mapped onto} \quad \begin{bmatrix} -\sin 30^\circ \\ \cos 30^\circ \end{bmatrix} .$$

Thus, the matrix representing this rotation is

$$A = \begin{bmatrix} \cos 30^\circ & -\sin 30^\circ \\ \sin 30^\circ & \cos 30^\circ \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} .$$

Note that the first column of A is the vector onto which $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is mapped; the second column of A is the image of $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

The <u>product</u> or <u>composition</u> of two transformations is defined just as you define the composition of two real functions of a real variable.

Definition 5-3. If f and g are transformations on H, then for each vector V in H the composition transformations fg and gf are the transformations such that

$$fg(V) = f(g(V)) \quad \text{and} \quad gf(V) = g(f(V)).$$

Thus, to find the image of V under the transformation fg, you first apply g, and then apply f. Consequently, if g maps V onto U, and if f maps U onto W, then fg maps V onto W.

The following theorem is readily proved (Exercise 5-3-7).

Theorem 5-4. If f is a linear transformation represented by the matrix A, and g is a linear transformation represented by the matrix B, then fg and gf are both linear transformations; fg is represented by AB, while gf is represented by BA.

For example, suppose that in the coordinate plane each position vector is first reflected in the y axis, and then the resulting vector is doubled in length. Let us find a matrix representation of the resulting linear transformation on H. If g is the mapping that transforms each vector into its reflection in the vertical axis, then we have

$$g : \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} -x \\ y \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

If f maps each vector into twice the vector, then we have

$$f : \begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow 2\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Accordingly, the matrix representing fg is

$$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 2 \end{bmatrix}.$$

Exercises 5-3

1. Show that each of the mappings in Exercise 5-1-3 is linear, by determining matrices representing the mappings.

2. Consider the linear transformations,

   p: reflection in the horizontal axis,

   q: horizontal projection on the line $y = -x$ (Exercise 5-1-5e),

   r: rotation counterclockwise through 90°,

   s: shear moving each point vertically through a distance equal to the abscissa of the point,

   of H into H. In each of the following, determine the matrix representing the given transformation:

(a) p,	(f) qp,	(k) s(rs),
(b) q,	(g) pr,	(l) (sr)s,
(c) r,	(h) rp,	(m) p(sq),
(d) s,	(i) qs,	(n) (ps)q,
(e) pq,	(j) sq,	(o) (sp)(rq).

3. Let f be the rotation of the plane counterclockwise through 45° about the origin, and let g be the rotation clockwise through 30°. Determine a matrix representing the rotation counterclockwise through 15° about the origin.

4. (a) Prove that every linear transformation maps the origin onto itself.

   (b) Prove that every linear transformation maps every subspace of H onto a subspace of H.

5. For every two linear transformations f and g on H, define $f + g$ to be the transformation such that, for each $V \in H$,

$$(f + g)(V) = f(V) + g(V).$$

   Without using matrices, prove that $f + g$ is a linear transformation on H.

6. For each linear transformation f on H and each real number a, define af to be the transformation such that

$$af(V) = f(aV).$$

Without using matrices, prove that af is a linear transformation on H.

7. Prove Theorem 5-4.

8. Without using matrices, prove each of the following:

(a) $f(g + h) = fg + fh$,

(b) $(f + g)h = fh + gh$,

(c) $f(ag) = a(fg)$,

where f, g, and h are any linear transformations on H and a is any real number.

5-4. One-to-one Linear Transformations

The reflection of the plane in the x axis clearly maps distinct points onto distinct points; thus, the reflection is a one-to-one linear transformation on H. Moreover, the reflection maps any pair of noncollinear vectors onto a pair of noncollinear vectors. It is easy to show that this property is common to all one-to-one linear transformations of H into itself.

Theorem 5-5. Every one-to-one linear transformation on H maps noncollinear vectors onto noncollinear vectors.

Proof. Let S_1 and S_2 be a pair of noncollinear vectors and let

$$f(S_1) = T_1 \quad \text{and} \quad f(S_2) = T_2$$

be their images under the one-to-one linear mapping f. Since f is one-to-one, we know that T_1 and T_2 are not both the zero vector. We may suppose, therefore, that T_1 is not the zero vector. To show that T_1 and T_2 are not collinear, we shall demonstrate that the assumption that they are collinear leads to a contradiction.

If T_1 and T_2 are collinear, then there exists a real number r such that $T_2 = r\,T_1$. Now, consider the image under f of the vector $r\,S_1$. Since f is linear, we have

$$\begin{aligned} f(r\,S_1) &= r\,f(S_1) \\ &= r\,T_1 \\ &= T_2. \end{aligned}$$

Thus, each of the vectors $r S_1$ and S_2 is mapped onto T_2. Since f is one-to-one, it follows that

$$r S_1 = S_2,$$

and therefore that S_1 and S_2 are collinear vectors. But this contradicts the fact that S_1 and S_2 are not collinear. Hence, the assumption that T_1 and T_2 are collinear must be false. Consequently, f must map noncollinear vectors onto noncollinear vectors.

Corollary 5-5-1. The subspace onto which a one-to-one linear transformation maps H is H itself.

Proof. Since the subspace contains a pair of noncollinear vectors, the corollary follows by use of Theorems 4-9 and 4-10.

The link between one-to-one transformations on H and second-order matrices having inverses is given in the next theorem.

Theorem 5-6. Let f be a linear transformation represented by the matrix A. Then f is one-to-one if and only if A has an inverse.

Proof. Suppose that A has an inverse. Let S_1 and S_2 be vectors in H having the same image under f. Now,

$$f(S_1) = AS_1 \quad \text{and} \quad f(S_2) = AS_2.$$

Thus,

$$AS_1 = AS_2.$$

Hence,

$$A^{-1}(AS_1) = A^{-1}(AS_2),$$

$$(A^{-1}A)S_1 = (A^{-1}A)S_2,$$

$$IS_1 = IS_2,$$

and

$$S_1 = S_2.$$

Thus, f must be a one-to-one transformation.

On the other hand, suppose that f is one-to-one. From Theorem 5-5, it follows that every vector in H is the image of some vector in H. In particular, there are vectors W and U such that

$$f(W) = AW = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

and

$$f(U) = AU = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Accordingly, the matrix having for its first column the vector W, and for its second column the vector U, is the inverse of A.

<u>Corollary 5-6-1</u>. A linear transformation represented by the matrix A is one-to-one if and only if

$$\delta(A) \neq 0.$$

The theory of systems of two linear equations in two variables can now be studied geometrically. Writing the system

$$\begin{aligned} a_{11}\,x + a_{12}\,y &= u, \\ a_{21}\,x + a_{22}\,y &= v, \end{aligned} \qquad (1)$$

in the form

$$AV = U, \qquad (2)$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad V = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \text{and} \quad U = \begin{bmatrix} u \\ v \end{bmatrix},$$

we seek the vectors V that are mapped by the matrix A onto the vector U.

If $\delta(A) \neq 0$, we now know that A represents a one-to-one mapping of H onto H. Therefore, A maps exactly one vector V onto U, namely, $V = A^{-1}U$. Thus, the system (1) — or, equivalently, (2) — has exactly one solution.

If $\delta(A) = 0$, then, in virtue of Corollary 4-7-1, the columns of A must be collinear vectors. Hence, A must have one of the forms

$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & a \\ 0 & b \end{bmatrix}, \quad \text{or} \quad \begin{bmatrix} a & ra \\ b & rb \end{bmatrix},$$

where not both a and b are zero. If A has the first of these forms, then A maps H onto the origin. In the other two cases, A maps H onto the line of vectors collinear with the vector $\begin{bmatrix} a \\ b \end{bmatrix}$. (See Exercise 5-4-7, below.) With these results in mind, you may now complete the discussion of the solution of equation (2).

Exercises 5-4

1. Using Theorem 5-6 or its corollary, determine which of the transformations in Exercise 5-1-3 are one-to-one.

2. Show that a linear transformation is one-to-one if and only if the kernel of the mapping consists only of the zero vector. (See Exercise 5-2-12.)

3. (a) Show that if f is a one-to-one linear transformation on H, then there exists a linear transformation g such that, for all $V \in H$,

$$gf(V) = V$$

and

$$fg(V) = V.$$

The transformation g is called the <u>inverse</u> of f and is usually written $g = f^{-1}$.

(b) Show that the transformation $g = f^{-1}$ in part (a) is a one-to-one transformation on H.

4. Prove that if f and g are one-to-one linear transformations of H, then fg is also a one-to-one transformation of H.

5. Prove that the set of one-to-one linear transformations on H is a group relative to the operation of composition of transformations.

6. Show that if f and g are linear transformations of H such that fg is a one-to-one transformation, then both f and g are one-to-one transformations.

7. (a) Show that if $\delta(A) = 0$, then the matrix A maps H onto a point (the origin) or onto a line.

   (b) Show that if A is the zero matrix and U is the zero vector, then every vector V of H is a solution of the equation $AV = U$.

   (c) Show that if $\delta(A) = 0$, but A is not the zero matrix, then the solution set of the equation

$$AV = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

   is a set of collinear vectors.

   (d) Show that if $\delta(A) = 0$, but A is not the zero matrix, and U is not the zero vector, then the solution set of the equation

$$AV = U$$

   either is empty or consists of all vectors of the form

$$\{V_1 + tV_2 : t \in R\},$$

   where V_1 and V_2 are fixed vectors such that

$$AV_1 = U \quad \text{and} \quad AV_2 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

8. Show that if the equation $AV = U$ has more than one solution for any given U, then A does not have an inverse.

5-5. Characteristic Values and Characteristic Vectors

If we think of a mapping as "carrying" points of the plane onto other points of the plane, we might ask, through curiosity, if there are cases in which the image point under a mapping is the same as the point itself. Such 'fixed' points, or vectors, are of great importance in mathematical analysis.

Let us look at an example. The reflection with respect to the x axis,

$$V \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} V,$$

that is,

$$\begin{bmatrix} x \\ y \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix},$$

has the property of mapping each vector on the x axis onto itself; thus, each of these vectors is fixed under the transformation.

<u>Definition 5-4</u>. If a transformation of H into itself maps a given vector onto itself, then that vector is a <u>fixed vector</u> for the transformation.

More generally, we are interested in any vector that is mapped into a multiple of itself; that is, we seek a vector $V \in H$ and a number $c \in R$ such that

$$AV = cV.$$

Since the equation is automatically satisfied by the zero vector regardless of the value c, this vector is ruled out.

The number c is called a <u>characteristic value</u> (or <u>eigenvalue</u>) of A, and the vector V a <u>characteristic vector</u> of A. These notions are fundamental in atomic physics since the energy levels of atoms and molecules turn out to be given by the eigenvalues of certain matrices. Also the analysis of flutter and vibration phenomena, the stability analysis of an airplane, and many other physical problems require finding the characteristic values and vectors of matrices.

In Section 5-1, we saw that the mapping

$$A = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix}$$

carried the plane H onto the line $y = x/2$. If we consider the set F

$$F = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : y = \frac{1}{2}x \right\}$$

under this same mapping, we see that F is mapped onto F', the set of vectors collinear with F. Note that

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2t \\ t \end{bmatrix} = \begin{bmatrix} 10t \\ 5t \end{bmatrix} = 5 \begin{bmatrix} 2t \\ t \end{bmatrix},$$

and hence that 5 is a characteristic value associated with A, and $\begin{bmatrix} 2t \\ t \end{bmatrix}$ is a characteristic vector for any $t \in R$, $t \neq 0$.

Definition 5-5. Each nonzero vector satisfying the equation

$$AV = cV$$

is called a characteristic vector, corresponding to the characteristic value (or characteristic root) c of A.

Note that, as remarked above, the trivial solution, $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ of the equation is not considered a characteristic vector.

Because of the importance of characteristic values in pure and applied mathematics, we need a method for finding them. We seek nonzero vectors V and real numbers c such that

$$AV = cV. \qquad (1)$$

If I is the identity matrix of order 2, then (1) can be written as

$$AV = (cI)V,$$

or

$$(A - cI)V = \underline{0}. \qquad (2)$$

If we let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \text{ and } V = \begin{bmatrix} x \\ y \end{bmatrix},$$

equation (2) becomes

$$\begin{bmatrix} a_{11} - c & a_{12} \\ a_{21} & a_{22} - c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{3}$$

We know that there is a nonzero vector V satisfying an equation of the form

$$BV = \underline{0}$$

if and only if

$$\delta(B) = 0.$$

Hence equation (2) has a solution other than the zero vector if and only if c is chosen in such a way as to satisfy the equation

$$(a_{11} - c)(a_{22} - c) - a_{12}a_{21} = 0.$$

Rearranged, this equation becomes

$$c^2 - (a_{11} + a_{22})c + \delta(A) = 0, \tag{4}$$

which is called the <u>characteristic equation</u> of the matrix A. Once this quadratic equation is solved for c, the corresponding vectors V satisfying equation (1) can readily be found, as illustrated in the following example.

<u>Example</u>. Determine the fixed lines under the mapping

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}.$$

We must solve the matrix equation

$$\begin{bmatrix} 2 - c & 3 \\ 0 & 1 - c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Since

$$\delta(A - cI) = (2 - c)(1 - c),$$

the characteristic equation is

$$(2 - c)(1 - c) = 0,$$

or

$$c^2 - 3c + 2 = 0,$$

the roots of which are 1 and 2. For $c = 1$, equation (3) becomes

$$\begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

which is equivalent to the system

$$x + 3y = 0,$$

$$0x + 0y = 0.$$

Thus, A maps the line $x + 3y = 0$ onto itself; that is, the set F

$$F = \left\{ r \begin{bmatrix} -3 \\ 1 \end{bmatrix} : \ r \in R \right\}$$

is mapped onto itself. Actually, since $c = 1$, each vector of this subspace is invariant: if V is a characteristic vector, f(V) its image, and $c = 1$, then

$$f(V) = V.$$

For $c = 2$, equation (3) becomes

$$\begin{bmatrix} 0 & 3 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

or

$$3y = 0,$$

$$-1y = 0.$$

Hence,

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

maps the line $y = 0$ onto itself; the set of vectors F

$$F = \left\{ r \begin{bmatrix} 1 \\ 0 \end{bmatrix} : \quad r \in R \right\}$$

is closed under the transformation.

The characteristic equation associated with the matrix

$$A = \begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}$$

is

$$c^2 - 3c + 2 = 0.$$

This equation expresses a real-number function. For a matrix function, the corresponding equation is

$$C^2 - 3C + 2I = \underline{0},$$

where I is the identity matrix of order 2 and $\underline{0}$ is the zero matrix of order 2. If we substitute A in this matrix equation,

$$\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix}^2 - 3\begin{bmatrix} 2 & 3 \\ 0 & 1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \underline{0},$$

we find that A is a root of its characteristic equation. This is true for any 2×2 matrix.

<u>Theorem 5-7</u>. The matrix A

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

is a solution of its characteristic equation

$$A^2 - (a_{11} + a_{22})A + \delta(A)I = \underline{0}.$$

The proof is left as an exercise.

Theorem 5-7 is the case $n = 2$ of a famous theorem called the Cayley-Hamilton Theorem, which states that an analogous result holds for matrices of any order n.

Exercises 5-5

1. Determine the characteristic roots and vectors of each of the following matrices:

   (a) $\begin{bmatrix} 2 & 5 \\ 0 & 3 \end{bmatrix}$, (c) $\begin{bmatrix} 2 & 1 \\ -1 & 0 \end{bmatrix}$,

   (b) $\begin{bmatrix} -3 & 4 \\ -1 & 2 \end{bmatrix}$, (d) $\begin{bmatrix} 0 & 2 \\ 0 & 1 \end{bmatrix}$.

2. Prove that zero is a characteristic root of a matrix A if and only if $\delta(A) = 0$.

3. Show that a linear transformation f is one-to-one if and only if zero is not a characteristic root of the matrix representing f.

4. Determine the invariant subspaces (fixed lines) of the mapping given by

$$\begin{bmatrix} 6 & 2 \\ 2 & 3 \end{bmatrix}.$$

   Show that these lines are mutually perpendicular.

5. Show that the matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is a solution of its characteristic (matrix) equation

$$A^2 - (a_{11} + a_{22})A + \delta(A)I = \underline{0}.$$

6. Show that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an invariant vector of the transformation

$$V \longrightarrow \|V\| \, V,$$

   but that $2\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is not invariant under this mapping.

7. Show that A maps every line through the origin onto itself if and only if

$$A = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$$

for $r \neq 0$.

8. Let $d = (a_{11} - a_{22})^2 + 4\, a_{12}a_{21}$, where a_{11}, a_{12}, a_{21}, and a_{22} are any real numbers. Show that the number of distinct real characteristic roots of the matrix

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

is

$$\begin{aligned} 0 &\text{ if } d < 0, \\ 1 &\text{ if } d = 0, \\ 2 &\text{ if } d > 0. \end{aligned}$$

9. Find a nonzero matrix that leaves no line through the origin fixed.

10. Determine a one-to-one linear transformation that maps exactly one line through the origin onto itself.

11. Show that every matrix of the form $\begin{bmatrix} r & s \\ s & t \end{bmatrix}$ has two distinct characteristic roots if $s \neq 0$.

12. Show that the matrix A and its transpose A^t have the same characteristic roots.

5-6. Rotations and Reflections

Since length is an important property in Euclidean geometry, we shall look for the linear transformations of the plane that leave unchanged the length $\|V\|$ of every vector V. Examples of such transformations are the following:

(a) the reflection of the plane in the x axis,

(b) a rotation of the plane through any given angle about the origin,

(c) a reflection in the x axis followed by a rotation about the origin.

Actually, we can show that any linear transformation that preserves the lengths of all vectors is equivalent to one of these three. The following theorem will be very useful in proving that result.

<u>Theorem 5-8</u>. A linear transformation of H that leaves unchanged the length of every vector also leaves unchanged (a) the inner product of every pair of vectors and (b) the magnitude of the angle between every pair of vectors.

<u>Proof</u>. Let V and U be a pair of vectors in H and let V' and U' be their respective images under the transformation. In virtue of Exercise 4-5-8, we have

$$||V + U||^2 = ||V||^2 + 2V \bullet U + ||U||^2 \qquad (1)$$

and

$$||V' + U'||^2 = ||V'||^2 + 2V' \bullet U' + ||U'||^2. \qquad (2)$$

Since the transformation is linear, for the image of $V + U$ we have

$$(V + U)' = V' + U'.$$

Consequently, (2) can be written as

$$||(V + U)'||^2 = ||V'||^2 + 2V' \bullet U' + ||V'||^2. \qquad (3)$$

But the transformation preserves the length of each vector; thus, we obtain

$$||V'|| = ||V||, \quad ||U'|| = ||U||, \quad \text{and} \quad ||(V + U)'|| = ||V + U||.$$

Making these substitutions in equation (3), we get

$$||V + U||^2 = ||V||^2 + 2V' \bullet U' + ||U||^2. \qquad (4)$$

Comparing equations (1) and (4), you see that we must have

$$V \bullet U = V' \bullet U';$$

that is, the transformation preserves the value of the inner product.

Since the magnitude of the angle between V and U can be expressed in terms of inner products (Theorem 4-5), it follows that the transformation also preserves that magnitude.

<u>Corollary 5-8-1</u>. If a linear transformation preserves the length of every vector, then it maps orthogonal vectors onto orthogonal vectors.

By the definition of orthogonality, this simply means that the geometric vectors are mutually perpendicular.

It is very easy to show the transformations we are considering also preserve the distance between every pair of points in the plane. We state this property formally in the next theorem, the proof of which is left as an exercise.

<u>Theorem 5-9</u>. A linear transformation that preserves the length of every vector leaves unchanged the distance between every pair of points in the plane; that is, if V' and U' are the respective images of the vectors V and U, then

$$||V' - U'|| = ||V - U||.$$

Let us now find a matrix representing any given linear length-preserving transformation of H. All we need to find are the images of the vectors

$$S_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad S_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

under such a transformation. (Why is this so?)

If S_1' and S_2' are the respective images of S_1 and S_2, then we know that both S_1' and S_2' are of length 1 and that they are orthogonal to each other.

Suppose that S_1' forms the angle α (alpha) with the positive half of the x axis (Figure 5-12). Since the length of S_1' equals 1, we have

$$S_1' = \begin{bmatrix} \cos\alpha \\ \sin\alpha \end{bmatrix}.$$

We know that S_2' is perpendicular to S_1'. Hence, there are two opposite

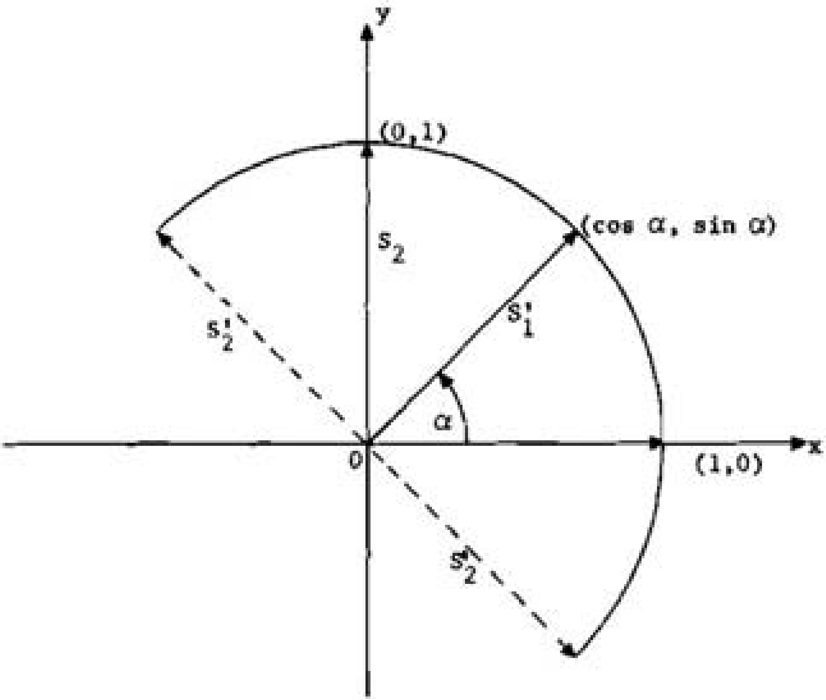


Figure 5-12. A length-preserving transformation.

possibilities for the direction of S'_2, because the angle β (beta) that S'_2 makes with the positive half of the x axis may be either

$$\beta = \alpha + \frac{\pi}{2} \tag{5}$$

or

$$\beta = \alpha - \frac{\pi}{2} . \tag{6}$$

In the first case (5), we have

$$S'_2 = \begin{bmatrix} \cos \beta \\ \sin \beta \end{bmatrix} = \begin{bmatrix} \cos \left(\alpha + \frac{\pi}{2}\right) \\ \sin \left(\alpha + \frac{\pi}{2}\right) \end{bmatrix} = \begin{bmatrix} -\sin \alpha \\ \cos \alpha \end{bmatrix} .$$

In the second case (6), we have

$$S'_2 = \begin{bmatrix} \cos \left(\alpha - \frac{\pi}{2}\right) \\ \sin \left(\alpha - \frac{\pi}{2}\right) \end{bmatrix} = \begin{bmatrix} \sin \alpha \\ -\cos \alpha \end{bmatrix} .$$

Accordingly, any linear transformation f that leaves the length of each vector unchanged must be represented by a matrix having either the form

$$A = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \qquad (7)$$

or the form

$$B = \begin{bmatrix} \cos\alpha & \sin\alpha \\ \sin\alpha & -\cos\alpha \end{bmatrix}. \qquad (8)$$

In the first instance (7), the transformation f simply rotates the basis vectors S_1 and S_2 through an angle α and we suspect that f is a rotation of the entire plane H through that angle. To verify this observation, we write the vector V in terms of its angle of inclination θ (theta) to the x axis and the length $r = ||V||$; that is, we write

$$V = \begin{bmatrix} r \cos\theta \\ r \sin\theta \end{bmatrix}. \qquad (9)$$

Forming AV from equations (7) and (9), we obtain

$$AV = \begin{bmatrix} r(\cos\theta\cos\alpha - \sin\theta\sin\alpha) \\ r(\sin\theta\cos\alpha + \cos\theta\sin\alpha) \end{bmatrix}.$$

From the formulas of trigonometry,

$$\cos(\theta + \alpha) = \cos\theta\cos\alpha - \sin\theta\sin\alpha,$$
$$\sin(\theta + \alpha) = \sin\theta\cos\alpha + \cos\theta\sin\alpha,$$

we see that

$$AV = \begin{bmatrix} r\cos(\theta + \alpha) \\ r\sin(\theta + \alpha) \end{bmatrix}.$$

Thus, AV is the vector of length r at an angle $\theta + \alpha$ to the horizontal axis. We have proved that the matrix A represents a rotation of H through the angle α.

But suppose f is represented by the matrix B in equation (8) above. This transformation differs from the one represented by A in that the vector S_2' is reflected across the line of the vector S_1'. Consequently, you may suspect that this transformation amounts to a reflection of the plane in the x axis followed by a rotation through the angle α. Since you know that the reflection in the x axis is represented by the matrix

$$J = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

you may therefore expect that

$$B = AJ. \tag{10}$$

We leave this verification as an exercise.

Exercises 5-6

1. Obtain the matrices that rotate H through the following angles:

(a) 180°,	(f) 90°,
(b) 45°,	(g) -120°,
(c) 30°,	(h) 360°,
(d) 60°,	(i) -135°,
(e) 270°,	(j) 150°.

2. Write out the matrices that represent the transformation consisting of a reflection in the x axis followed by the rotations of Exercise 1.

3. Verify Equation (10), above.

4. A linear transformation of H that preserves the length of every vector is called an <u>orthogonal</u> transformation, and the matrix representing the transformation is called an <u>orthogonal</u> matrix. Prove that the transpose of an orthogonal matrix is orthogonal.

5. Show that the inverse of an orthogonal matrix is an orthogonal matrix.

6. Show that the product of two orthogonal matrices is orthogonal.

7. (a) Show that a translation of H in the direction of the vector

$$U = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

and through a distance equal to the length of U is given by the mapping

$$V \longrightarrow V + U.$$

(b) Show that this mapping does not preserve the length of every vector, but that it does preserve the distance between every pair of points in the plane.

(c) Determine whether or not this mapping is linear.

8. Let R_α and R_β denote rotations of H through the angles α and β, respectively. Prove that a rotation through α followed by a rotation through β amounts to a rotation through $\alpha + \beta$; that is, show that

$$R_\beta R_\alpha = R_{\alpha+\beta}.$$

9. Note that the matrix A of Equation (7) is a representation of a complex number. What does the result of Exercise 8 imply for complex numbers?

10. (a) Find a matrix that represents a reflection across the line of the vector

$$\begin{bmatrix} \cos \alpha \\ \sin \alpha \end{bmatrix}.$$

(b) Show that the matrix B of equation (8), above, represents a reflection across the line of some vector.

Appendix

RESEARCH EXERCISES

The exercises in this Appendix are essentially "research-type" problems designed to exhibit aspects of theory and practice in matrix algebra that could not be included in the text. They are especially suited as individual assignments for those students who are prospective majors in the theoretical and practical aspects of the scientific disciplines, and for students who would like to test their mathematical powers. Alternatively, small groups of students might join forces in working them.

1. Quaternions. The algebraic system that is explored in this exercise was invented by the Irish mathematician and physicist, William Rowan Hamilton, who published his first paper on the subject in 1835. It was not until 1858 that Arthur Cayley, an English mathematician and lawyer, published the first research paper on matrices, though the name matrix had previously been applied by James Joseph Sylvester in 1850 to rectangular arrays of numbers. Since Hamilton's system of quaternions is actually an algebra of matrices, it is more easily presented in this guise than in the form in which it was first developed.

In the present exercise, we shall consider the algebra of 2×2 matrices with complex numbers as entries. The definitions of addition, multiplication, and inversion remain the same. We use C for the set of all complex numbers and we denote by K the set of all matrices

$$\begin{bmatrix} z & w \\ z_1 & w_1 \end{bmatrix},$$

where z, w, z_1, and w_1 are elements of C. As is the case with matrices having real entries, the element

$$\begin{bmatrix} z & w \\ z_1 & w_1 \end{bmatrix}$$

of K has an inverse if and only if

$$zw_1 - wz_1 \neq 0,$$

and then we have

$$\begin{bmatrix} z & w \\ z_1 & w_1 \end{bmatrix}^{-1} = \frac{1}{zw_1 - wz_1} \begin{bmatrix} w_1 & -w \\ -z_1 & z \end{bmatrix}, \quad zw_1 - wz_1 \neq 0.$$

Since 1 is a complex number, the unit matrix is still

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

If

$$z = x + iy,$$

then we write

$$\bar{z} = x - iy$$

and call this number the <u>complex conjugate</u> of z, or simply the <u>conjugate</u> of z.

A <u>quaternion</u> is an element q of K of the particular form

$$\begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix}, \quad z \in C \text{ and } w \in C.$$

We denote by Q the set of all quaternions.

(a) Show that $\delta(q) = x^2 + y^2 + u^2 + v^2$ if $z = x + iy$ and $w = u + iv$. Hence conclude that, since x, y, u, and v are real numbers, $\delta(q) = 0$ if and only if $q = \underline{0}$.

(b) Show that if $q \in Q$, then q has an inverse if and only if $q \neq \underline{0}$, and exhibit the form of q^{-1} if it exists.

Four elements of Q are of particular importance and we give them special names:

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$U = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix},$$

$$V = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

$$W = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}.$$

(c) Show that if

$$q = \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix},$$

where $z = x + iy$ and $w = u + iv$, then

$$q = xI + yU + uV + vW.$$

(d) Prove the following identities involving I, U, V and W:

$$U^2 = V^2 = W^2 = -I$$

and

$$UV = W = -VU, \quad VW = U = -WV, \quad \text{and} \quad WU = V = -UW.$$

(e) Use the preceding two exercises to show that if $q \in Q$ and $r \in Q$, then $q + r$, $q - r$, and qr are all elements of Q.

The <u>conjugate</u> of the element

$$q = \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix}, \quad \text{where } z = x + iy, \quad w = u + iv,$$

is

$$\bar{q} = \begin{bmatrix} \bar{z} & -w \\ \bar{w} & z \end{bmatrix},$$

and the <u>norm</u> and <u>trace</u> are given respectively by

$$|q| = \left[\delta(q)\right]^{1/2}$$

and

$$t(q) = 2x.$$

(f) Show that if $q \in Q$, and if q is invertible, then

$$q^{-1} = \frac{1}{|q|^2}\,\bar{q}.$$

From this conclude that if $q \in Q$, and if q^{-1} exists, then $q^{-1} \in Q$.

(g) Show that each $q \in Q$ satisfies the quadratic equation

$$q^2 - t(q)\,q + |q|^2\,I = \underline{O}.$$

(h) Show that if $q \in Q$, then

$$q\bar{q} = |q|^2\,I.$$

Note that this may be proved by using the result that if

$$q = aI + bU + cV + dW,$$

then

$$\bar{q} = aI - bU - cV - dW,$$

and then using the results given in (d).

(i) Show that if $q \in Q$ and $r \in Q$, then

$$|qr| = |q|\,|r|$$

and

$$|q + r| \leq |q| + |r|.$$

The geometry of quaternions constitutes a very interesting subject. It requires the representation of a quaternion

$$q = aI + bU + cV + dW$$

as a point with coordinates (a, b, c, d) in four-dimensional spaces. The subset of elements,

$$Q_1 = \{q\colon q \in Q \text{ and } |q| = 1\},$$

is a group and is represented geometrically as the hypersphere with equation

$$a^2 + b^2 + c^2 + d^2 = 1.$$

2. Nonassociative Algebras

The algebra of matrices (we restrict our attention in this exercise to the set M of 2×2 matrices) has an associative but not a commutative multiplication. "Algebras" with nonassociative multiplication have become increasingly important in recent years—for example, in mathematical genetics. Genetics is a subdiscipline of biology and is concerned with transmission of hereditary traits. Nonassociative "algebras" are important also in the study of quantum mechanics, a subdiscipline of physics. We give first a simple example of a Lie algebra (named after the Norwegian geometer Sophus Lie).

If $A \in M$ and $B \in M$, we write

$$A \circ B = AB - BA$$

and read this "A op B," "op" being an abbreviation for operation.

(a) Prove the following properties of $\circ$:

(i) $A \circ B = - B \circ A$,

(ii) $A \circ A = \underline{0}$,

(iii) $A \circ (B \circ C) + B \circ (C \circ A) + C \circ (A \circ B) = \underline{0}$,

(iv) $A \circ I = \underline{0} = I \circ A$.

(b) Give an example to show that $A \circ (B \circ C)$ and $(A \circ B) \circ C$ are different and hence that $\circ$ is not an associative operation.

Despite these strange properties, $\circ$ behaves nicely relative to ordinary matrix addition.

(c) Show that $\circ$ distributes over addition:

$$A \circ (B + C) = (A \circ B) + (A \circ C)$$

and

$$(A + B) \circ C = (A \circ C) + (B \circ C).$$

(d) Show that $\circ$ behaves nicely relative to multiplication by a number.

It will be recalled that A^{-1} is termed the multiplicative inverse of A and is defined as the element B satisfying the relationships

$$AB = I = BA.$$

But it must also be recalled that this definition was motivated by the fact that

$$AI = A = IA,$$

that is, by the fact that I is a multiplicative unit.

(e) Show that there is no o unit.

We know, from the foregoing work, that o is neither commutative nor associative. Here is another kind of operation, called Jordan multiplication: If $A \in M$ and $B \in M$, we define

$$AjB = \frac{(AB + BA)}{2}.$$

We see at once that

$$AjB = BjA,$$

so that Jordan multiplication is a commutative operation; but it is not associative.

(f) Determine all of the properties of the operation j that you can. For example, does j distribute over addition?

3. The Algebra of Subsets

We have seen that there are interesting algebraically defined subsets of M, the set of all 2×2 matrices. One such subset, for example, is the set Z, which is isomorphic with the set of complex numbers. Much of higher mathematics is concerned with the "global structure" of "algebras," and generally this involves the consideration of subsets of the "algebras" being studied. In this exercise, we shall generally underscore letters to denote subsets of M.

If $\underline{A}$ and $\underline{B}$ are subsets of M, then

$$\underline{A} + \underline{B}$$

is the set of all elements of the form

$$A + B, \text{ where } A \in \underline{A} \text{ and } B \in \underline{B}.$$

In set-builder notation this may be written

$$\underline{A} + \underline{B} = \{A + B : A \in \underline{A} \text{ and } B \in \underline{B}\}.$$

* By an <u>additive subset</u> of M is meant a subset $\underline{A} \subset M$ such that

$$\underline{A} + \underline{A} \subset \underline{A}.$$

(a) Determine which of the following are additive subsets of M:

(i) $\{\underline{O}\}$,

(ii) $\{I\}$,

(iii) M,

(iv) Z,

(v) M_1, the set of all A in M with $\delta(A) = 1$,

(vi) the set of all elements of M whose entries are nonnegative.

(b) Prove that if $\underline{A}$, $\underline{B}$, and $\underline{C}$ are subsets of M, then

(i) $\underline{A} + \underline{B} = \underline{B} + \underline{A}$,

(ii) $\underline{A} + (\underline{B} + \underline{C}) = (\underline{A} + \underline{B}) + \underline{C}$,

(iii) and if $\underline{A} \subset \underline{B}$ then $\underline{A} + \underline{C} \subset \underline{B} + \underline{C}$.

(c) Prove that if $\underline{A}$ and $\underline{B}$ are additive subsets of M, then

$$\underline{A} + \underline{B}$$

is also an additive subset of M.

Let V denote the set of all column vectors

$$\begin{bmatrix} x \\ y \end{bmatrix}$$

with $x \in R$ and $y \in R$.

(d) Show that if v is a fixed element of V, then

$$\left\{A\colon A \in M \text{ and } Av = \begin{bmatrix} 0 \\ 0 \end{bmatrix}\right\}$$

is an additive subset of M. Notice also that if $Av = \underline{0}$ then $(-A)v = \underline{0}$.

If $\underline{A}$ and $\underline{B}$ are subsets of M, then

$$\underline{AB}$$

is the set of all

$$AB,\ A \in \underline{A} \text{ and } B \in \underline{B}.$$

Using set-builder notation, we can write this in the form

$$\underline{AB} = \{AB\colon A \in \underline{A} \text{ and } B \in \underline{B}\}.$$

A subset $\underline{A}$ of M is multiplicative if

$$\underline{AA} \subset \underline{A}.$$

(e) Which of the subsets in part (a) are multiplicative?

(f) Show that if $\underline{A}$, $\underline{B}$, and $\underline{C}$ are subsets of M, then

(i) $\underline{A}(\underline{BC}) = (\underline{AB})\underline{C}$,

(ii) and if $\underline{A} \subset \underline{B}$, then $\underline{AC} \subset \underline{BC}$.

(g) Give an example of two subsets $\underline{A}$ and $\underline{B}$ of M such that

$$\underline{AB} \neq \underline{BA}.$$

(h) Determine which of the following subsets are multiplicative:

(i) $\{\underline{0}, I\}$,

(ii) $\{I, -I\}$,

(iii) the set of all elements of M with negative entries,

(iv) the set of all elements of M for which the upper left-hand entry is less than 1,

(v) the set of all elements of M of the form

$$\begin{bmatrix} x & y \\ 0 & 1 \end{bmatrix},$$

with $0 \leq x$, $0 \leq y$, and $x + y \leq 1$.

The exercises stated above are suggestions as to how this "algebra of subsets" works. There are many other results that come to mind, but we shall leave them to you to find. Here are some clues: How would you define $t\underline{A}$ if $t \in R$ and $\underline{A} \subset M$? Is $(-1)\underline{A} = -\underline{A}$? Wait a minute! What does $-\underline{A}$ mean? What does $\underline{A}^7$ mean? Does set multiplication distribute over addition, over union, over intersection? Do not expect that even your teacher knows the answer to all of these possible questions. Few people know all of them and fewer still, of those who know them, remember them. If you conjecture that something is true but the proof of it escapes you, then try to construct an example to show that it is false. If this does not work, try proving it again, and so on.

4. Analysis and Synthesis of Proofs

This is an exercise in analysis and synthesis, taking an old proof to pieces and using the pattern to make a new proof. In describing his activities, a mathematician is likely to put at the very top that of creating new results. But "result" in mathematics usually means "theorem and proof." The mathematician does not by any means limit his methods in conjecturing a new theorem: He guesses, uses analogies, draws diagrams and figures, sets up physical models, experiements, computes; no holds are barred. Once he has his conjecture firmly in mind, he is only half through, for he still must construct a proof. One way of doing this is to analyze proofs of known theorems that are somewhat like the theorem he is trying to prove and then synthesize a proof of the new theorem. Here we ask you to apply this process of analysis and synthesis of proofs to the algebra of matrices. To accomplish this, we shall introduce some new operations among matrices by analogy with the old operations.

For simplicity of computation, we shall use only 2×2 matrices.

To start with, we introduce new operations in the set of real numbers, R. If $x \in R$ and $y \in R$, we define

$x \wedge y =$ the smaller of x and y (read: "x cap y")

and

$x \vee y =$ the larger of x and y (read: "x cup y").

(a) Show that if $x \in R$, $y \in R$, and $z \in R$, then

(i) $x \wedge y = y \wedge x$,

(ii) $x \vee y = y \vee x$,

(iii) $x \wedge (y \wedge z) = (x \wedge y) \wedge z$,

(iv) $x \vee (y \vee z) = (x \vee y) \vee z$,

(v) $x \wedge x = x$,

(vi) $x \vee x = x$,

(vii) $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$,

(viii) $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$.

Although the foregoing operations may seem a little unusual, you will have no difficulty in proving the above statements. They are not difficult to remember if you notice the following facts:

The even-numbered results can be obtained from the odd-numbered results by interchanging $\wedge$ and $\vee$, and conversely.

The first states that $\wedge$ is commutative and the third states that $\wedge$ is associative. The fifth is new but the seventh states that $\wedge$ distributes over $\vee$.

To define the matrix operations, let us think of $\wedge$ as the analog of <u>multiplication</u> and $\vee$ as the analog of <u>addition</u> and let us begin with our new matrix "multiplication."

We <u>define</u>

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \wedge \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} (a \wedge x) \vee (b \wedge z) & (a \wedge y) \vee (b \wedge w) \\ (c \wedge x) \vee (d \wedge z) & (c \wedge y) \vee (d \wedge w) \end{bmatrix}.$$

This is simply the row-by-column operations, except that $\wedge$ is used in place of multiplication and $\vee$ is used in place of addition. To see this more clearly, we write

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} ax + bz & ay + bw \\ cx + dz & cy + dw \end{bmatrix}.$$

(b) Write out a proof that if A, B, and C are elements of M, then

$$A(BC) = (AB)C.$$

Be sure not to omit any steps in the proof. Using this as a pattern, write out a proof that

$$A \wedge (B \wedge C) = (A \wedge B) \wedge C,$$

verifying at each step that you have the necessary results from (a) to make the proof sound. List all the properties of the two pairs of operations that you need, such as associativity, commutativity, and distributivity.

(c) Using the analogy between $\vee$ and addition, define $A \vee B$ for elements A and B of M.

(d) State and prove, for the new operations, analogs of all the rules you know for the operations of matrix addition and multiplication.

BIBLIOGRAPHY

C. B. Alledoerfer and C. O. Oakley, "Fundamentals of Freshman Mathematics," McGraw-Hill Book Company, Inc., New York, 1959.

Richard V. Andree, "Selection from Modern Abstract Algebra," Henry Holt and Company, 1958.

R. A. Beaumont and R. W. Ball, "Introduction to Modern Algebra and Matrix Theory," Rinehart and Company, New York, 1954.

Garrett Birkhoff and Saunders MacLane, "A Survey of Modern Algebra," The Macmilian Company, New York, 1959.

R. E. Johnson, "First Course in Abstract Algebra," Prentice-Hall, Inc., New York, 1958.

John G. Kemeny, J. Laurie Snell, and Gerald L. Thompson, "Introduction to Finite Mathematics," Prentice-Hall, Inc., New York, 1957.

Lowell J. Paige and J. Dean Swift, "Elements of Linear Algebra," Ginn and Company, 1961.

M. J. Weiss, "Higher Algebra for the Undergraduate," John Wiley and Sons, Inc., New York, 1949.

INDEX