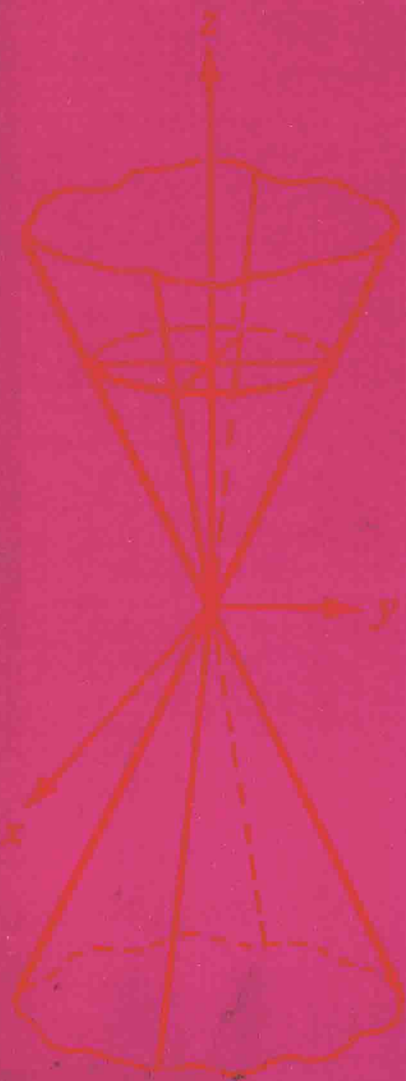

PURE AND APPLIED MATHEMATICS

A Series of Monographs and Textbooks

LINEAR ALGEBRA

With Geometric Applications

Larry E. Mansfield

LINEAR ALGEBRA
With Geometric Applications

LARRY E. MANSFIELD

*Queens College of the City
University of New York
Flushing, New York*

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 75–10345
ISBN: 0–8247–6321–1

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

Preface

Until recently an introduction to linear algebra was devoted primarily to solving systems of linear equations and to the evaluation of determinants. But now a more theoretical approach is usually taken and linear algebra is to a large extent the study of an abstract mathematical object called a vector space. This modern approach encompasses the former, but it has the advantage of a much wider applicability, for it is possible to apply conclusions derived from the study of an abstract system to diverse problems arising in various branches of mathematics and the sciences.

Linear Algebra with Geometric Applications was developed as a text for a sophomore level, introductory course from dittoed material used by several classes. Very little mathematical background is assumed aside from that obtained in the usual high school algebra and geometry courses. Although a few examples are drawn from the calculus, they are not essential and may be skipped if one is unfamiliar with the ideas. This means that very little mathematical sophistication is required. However, a major objective of the text is to develop one's mathematical maturity and convey a sense of what constitutes modern mathematics. This can be accomplished by determining how one goes about solving problems and what constitutes a proof, while mastering computational techniques and the underlying concepts. The study of linear algebra is well suited to this task for it is based on the simple arithmetic properties of addition and multiplication.

Although linear algebra is grounded in arithmetic, so many new concepts must be introduced that the underlying simplicity can be obscured by terminology. Therefore every effort has been made to introduce new terms only when necessary and then only with sufficient motivation. For example, systems of linear equations are not considered until it is clear how they arise, matrix multiplication is not defined until one sees how it will be used, and complex scalars are not introduced until they are actually needed. In addition, examples are presented with each new term. These examples are usually either algebraic or geometric in nature. Heavy reliance is placed on geometric examples because geometric ideas are familiar and they provide good interpretations of linear algebraic concepts. Examples employing polynomials or functions are also easily understood and they supply nongeometric interpretations. Occasionally examples are drawn from other fields to suggest the range of possible application, but this is not done often because it is difficult to clarify a new concept while motivating and solving problems in another field.

The first seven chapters follow a natural development begining with an algebraic approach to geometry and ending with an algebraic analysis of second degree curves and surfaces. Chapter 8 develops canonical forms for matrices under similarity and might be covered at any point after Chapter 5. It is by far the most difficult chapter in the book. The appendix on determinants refers to concepts found in Chapters 4 and 6, but it could be taken up when determinants are introduced in Chapter 3.

Importance of Problems The role of problems in the study of mathematics cannot be overemphasized. They should not be regarded simply as hurdles to be overcome in assignments and tests. Rather they are the means to understanding the material being presented and to appreciating how ideas can be used. Once the role of problems is understood, it will be seen that the first place to look for problems is not necessarily in problem sets. It is important to be able to find and solve problems while reading the text. For example, when a new concept is introduced, ask yourself what it really means; look for an example in which the property is not present as well as one in which it is, and then note the differences. Numerical examples can be made from almost any abstract expression. Whenever an abstract expression from the text or one of your own seems unclear, replace the symbols with particular numerical expressions. This usually transforms the abstraction into an exercise in arithmetic or a system of linear equations. The next place to look for problems is in worked out examples and proved theorems. In fact, the best way to understand either is by working through the computation or deduction on paper as you read, filling in any steps that may have been omitted. Most of our theorems will have fairly simple proofs which can be constructed with little more than a good understanding of what is being claimed and the knowledge of how each term is defined. This does not mean that you should be able to prove each theorem when you first encounter it, however the attempt to construct a proof will usually aid in understanding the given proof. The problems at the end of each section should be considered next. Solve enough of the computational problems to master the computational techniques, and work on as many of the remaining problems as possible. At the very least, read each problem and determine exactly what is being claimed. Finally you should often try to gain an overview of what you are doing; set yourself the problem of determining how and why a particular concept or technique has come about. In other words, ask yourself what has been achieved, what terms had to be introduced, and what facts were required. This is a good time to see if you can prove some of the essential facts or outline a proof of the main result.

At times you will not find a solution immediately, but simply attempting to set up an example, prove a theorem, or solve a problem can be very useful. Such an attempt can point out that a term or concept is not well understood and thus lead to further examination of some idea. Such an examination will often provide the basis for finding a solution, but even if it does not, it should lead to a fuller understanding of some aspect of linear algebra.

Because problems are so important, an extensive solution section is provided for the problem sets. It contains full answers for all computational problems and some theoretical problems. However, when a problem requires a proof, the actual development of the argument is one objective of the problem. Therefore you will often find a suggestion as to how to begin rather than a complete solution. The way in which such a solution begins is very important; too often an assumption is made at the begining of an argument which amounts to assuming what is to be proved, or a hypothesis is either misused or omitted entirely. One should keep in mind that a proof is viewed in its entirety, so that an argument which begins incorrectly cannot become a proof no matter what is claimed in the last line about having solved the problem. A given suggestion or proof should be used as a last resort, for once you see a completed argument you can no longer create it yourself; creating a proof not only extends your knowledge, but it amounts to participating in the development of linear algebra.

Acknowledgments At this point I would like to acknowledge the invaluable assistance I have received from the many students who worked through my original lecture notes. Their observations when answers did not check or when arguments were not clear have lead to many changes and revisions. I would also like to thank Professor Robert B. Gardner for his many helpful comments and suggestions.

Contents

Appendix: Determinants

1

A Geometric Model

Before begining an abstract study of vector spaces, it is helpful to have a concrete example to use as a guide. Therefore we will begin by defining a particular vector space and after examining a few of its properties, we will see how it may be used in the study of plane geometry.

§1. The Field of Real Numbers

Our study of linear algebra is based on the arithmetic properties of real numbers, and several important terms are derived directly from these properties. Therefore we begin by examining the basic properties of arithmetic. The set of all real numbers will be denoted by R, and the symbol "$\in$" will mean "is a member of." Thus $\sqrt{2} \in R$ can be read as "$\sqrt{2}$ is a member of the real number system" or more simply as "$\sqrt{2}$ is a real number." Now if r, s, and t are any real numbers, then the following properties are satisfied:

Properties of addition:

$r + s \in R$	or R is *closed under addition*
$r + (s + t) = (r + s) + t$	or addition is *associative*
$r + s = s + r$	or addition is *commutative*
$r + 0 = r$	or 0 is an *additive identity*

For any $r \in R$, there is an *additive inverse* $-r \in R$ such that $r + (-r) = 0$.

Properties of multiplication:

$r \cdot s \in R$	or R is *closed under multiplication*
$r \cdot (s \cdot t) = (r \cdot s) \cdot t$	or multiplication is *associative*
$r \cdot s = s \cdot r$	or multiplication is *commutative*
$r \cdot 1 = r$	or 1 is a *multiplicative identity*

For any $r \in R$, $r \neq 0$, there is a *multiplicative inverse* $r^{-1} \in R$ such that $r \cdot (r^{-1}) = 1$.

The final property states that multiplication distributes over addition and ties the two operations together:

$$r \cdot (s + t) = r \cdot s + r \cdot t \qquad \text{a } \textit{distributive law}.$$

This is a rather special list of properties. On one hand, none of the properties can be derived from the others, while on the other, many properties of real numbers are omitted. For example, it does not contain properties of order or the fact that every real number can be expressed as a decimal. Only certain properties of the real number system have been included, and many other mathematical systems share them. Thus if r, s, and t are thought of as complex numbers and R is replaced by C, representing the set of all complex numbers, then all the above properties are still valid. In general, an algebraic system satisfying all the preceding properties is called a *field*. The real number system and the complex number system are two different fields, and there are many others. However, we will consider only the field of real numbers in the first five chapters.

Addition and multiplication are *binary operations*, that is they are only defined for two elements. This explains the need for associative laws. For if addition were not associative, then $r + (s + t)$ need not equal $(r + s) + t$ and $r + s + t$ would be undefined. The field properties listed above may seem obvious, but it is not to difficult to find binary operations that violate any or all of them.

One phrase in the preceding list which will appear repeatedly is the statement that a set is closed under an operation. The statement is defined for a set of numbers and the operations of addition and multiplication as follows:

Definition Let S be a set of real numbers. S is *closed under addition* if $r + t \in S$ for all $r, t \in S$. S is *closed under multiplication* if $r \cdot t \in S$ for all $r, t \in S$.

For example, if S is the set containing only the numbers 1, 3, 4, then S is not closed under either addition or multiplication. For $3 + 4 = 7 \notin S$ and $3 \cdot 4 = 12 \notin S$, yet both 3 and 4 are in S. As another example, the set of all odd integers is closed under multiplication but is not closed under addition.

Some notation is useful when working with sets. When the elements are easily listed, the set will be denoted by writing the elements within brackets. Therefore $\{1, 3, 4\}$ denotes the set containing only the numbers 1, 3, and 4. For larger sets, a set-building notation is used which denotes an arbitrary member of the set and states the conditions which must be satisfied by any member of the set. This notation is $\{\cdots \mid \cdots\}$ and may be read as "the set of all $\cdots$ such that $\cdots$." Thus the set of odd integers could be written as:

$\{x \mid x$ is an odd integer$\}$, "the set of all x such that x is an odd integer." Or it could be written as $\{2n + 1 \mid n$ is an integer$\}$.

Problems

1. Write out the following notations in words:

   a. $7 \in R$. b. $\sqrt{-6} \notin R$. c. $\{0, 5\}$. d. $\{x \mid x \in R, x < 0\}$.
   e. $\{x \in R \mid x^2 = -1\}$.

2. a. Show by example that the set of odd integers is not closed under addition.
   b. Prove that the set of odd integers is closed under multiplication.

3. Determine if the following sets are closed under addition or multiplication:

   a. $\{1, -1\}$. b. $\{5\}$. c. $\{x \in R \mid x < 0\}$. d. $\{2n \mid n$ is an integer$\}$.
   e. $\{x \in R \mid x \geq 0\}$.

4. Using the property of addition as a guide, give a formal definition of what it means to say that "addition of real numbers is commutative."

5. A distributive law is included in the properties of the real number system. State another distributive law which holds and explain why it was not included.

§2. The Vector Space $\mathscr{V}_2$

It would be possible to begin a study of linear algebra with a formal definition of an abstract vector space. However, it is more fruitful to consider an example of a particular vector space first. The formal definition will essentially be a selection of properties possessed by the example. The idea is the same as that used in defining a field by selecting certain properties of real numbers. The mathematical problem is to select enough properties to give the essential character of the example while at the same time not taking so many that there are few examples that share them. This procedure obviously cannot be carried out with only one example in hand, but even with several examples the resulting definition might appear arbitrary. Therefore one should not expect the example to point directly to the definition of an abstract vector space, rather it should provide a first place to interpret abstract concepts.

As with the real number system, a vector space will be more than just a collection of elements; it will also include the algebraic structure imposed by operations on the elements. Therefore to define the vector space $\mathscr{V}_2$, both its elements and its operations must be given.

The elements of $\mathscr{V}_2$ are defined to be all ordered pairs of real numbers and are called *vectors*. The operations of $\mathscr{V}_2$ are addition and scalar multiplication as defined below:

Vector addition: The sum of two vectors (a_1, b_1) and (a_2, b_2) is defined by: $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$.

For example, $(2, -5) + (4, 7) = (2 + 4, -5 + 7) = (6, 2)$.

Scalar multiplication: For any real number r, called a *scalar*, and any vector (a, b) in $\mathscr{V}_2$, $r(a, b)$ is a scalar multiple and is defined by $r(a, b) = (ra, rb)$.

For example, $5(3, -4) = (15, -20)$.

Now the set of all ordered pairs of real numbers together with the operations of vector addition and scalar multiplication forms the vector space $\mathscr{V}_2$. The numbers a and b in the vector (a, b) are called the *components* of the vector. Since vectors are ordered pairs, two vectors (a, b) and (c, d) are *equal* if their corresponding components are equal, that is if $a = c$ and $b = d$.

One point that should be made immediately is that the term "vector" may be applied to many different objects, so in other situations the term may apply to something quite different from an ordered pair of real numbers. In this regard it is commonly said that a vector has magnitude and direction, but this is not true for vectors in $\mathscr{V}_2$.

The strong similarity between the vectors of $\mathscr{V}_2$ and the names for points in the Cartesian plane will be utilized in time. However, these are quite different mathematical objects, for $\mathscr{V}_2$ has no geometric properties and the Cartesian plane does not have algebraic properties. Before relating the vector space $\mathscr{V}_2$ with geometry, much can be said of its algebraic structure.

Theorem 1.1 (Basic properties of addition in $\mathscr{V}_2$) If U, V, and W are any vectors in $\mathscr{V}_2$, then

1. $U + V \in \mathscr{V}_2$
2. $U + (V + W) = (U + V) + W$
3. $U + V = V + U$
4. $U + (0, 0) = U$
5. For any vector $U \in \mathscr{V}_2$, there exists a vector $-U \in \mathscr{V}_2$ such that $U + (-U) = (0, 0)$.

Proof Each of these is easily proved using the definition of addition in $\mathscr{V}_2$ and the properties of addition for real numbers. For example, to prove part 3, let $U = (a, b)$ and $V = (c, d)$ where $a, b, c, d \in R$, then

$$\begin{aligned} U + V &= (a, b) + (c, d) \\ &= (a + c, b + d) && \text{Definition of vector addition} \\ &= (c + a, d + b) && \text{Addition in } R \text{ is commutative} \\ &= (c, d) + (a, b) && \text{Definition of addition in } \mathscr{V}_2 \\ &= V + U. \end{aligned}$$

The proof of 2 is similiar, using the fact that addition in R is associative, and 4 follows from the fact that zero is an additive identity in R. Using the above notation, $U + V = (a + c, b + d)$ and $a + c, b + d \in R$ since R is closed under addition. Therefore $U + V \in \mathscr{V}_2$ if $U, V \in \mathscr{V}_2$ and 1 holds. Part 5 follows from the fact that every real number has an additive inverse, thus if $U = (a, b)$

$$U + (-a, -b) = (a, b) + (-a, -b) = (a - a, b - b) = (0, 0)$$

and $(-a, -b)$ can be called $-U$.

Each property in Theorem 1.1 arises from a property of addition in R and gives rise to similar terminology. (1) states that $\mathscr{V}_2$ is *closed under addition*. From (2) and (3) we say that addition in $\mathscr{V}_2$ is *associative* and *commutative*, respectively. The fourth property shows that the vector $(0, 0)$ is an *identity for addition* in $\mathscr{V}_2$. Therefore $(0, 0)$ is called the *zero vector* of the vector space $\mathscr{V}_2$ and it will be denoted by $\mathbf{0}$. Finally the fifth property states that every vector U has an *additive inverse* denoted by $-U$.

Other properties of addition and a list of basic properties for scalar multiplication can be found in the problems below.

Problems

1. Find the following vector sums:

a. $(2, -5) + (3, 2)$.
b. $(-6, 1) + (5, -1)$.
c. $(2, -3) + (-2, 3)$.
d. $(1/2, 1/3) + (-1/4, 2)$.

2. Find the following scalar multiples:

a. $\frac{1}{2}(4, -5)$. b. $0(2, 6)$. c. $3(2, -1/3)$. d. $(-1)(3, -6)$.

3. Solve the following equations for the vector U:

   a. $U + (2, -3) = (4, 7)$. c. $(-5, 1) + U = (0, 0)$.
   b. $3U + (2, 1) = (1, 0)$. d. $2U + (-4)(3, -1) = (1, 6)$.

4. Show that for all vectors $U \in \mathscr{V}_2$, $U + \mathbf{0} = U$.

5. Prove that vector addition in $\mathscr{V}_2$ is associative; give a reason for each step.

6. Suppose an operation were defined on pairs of vectors from $\mathscr{V}_2$ by the formula $(a, b)\circ(c, d) = ac + bd$. Would $\mathscr{V}_2$ be closed under this operation?

7. The following is a proof of the fact that the additive identity of $\mathscr{V}_2$ is unique, that is, there is only one vector that is an additive identity. Find the reasons for the six indicated steps.

   Suppose there is a vector $W \in \mathscr{V}_2$ such that $U + W = U$ for all $U \in \mathscr{V}_2$. Then, since $\mathbf{0}$ is an additive identity, it must be shown that $W = \mathbf{0}$.

Let $U = (a, b)$ and $W = (x, y)$	a. ?
then $U + W = (a, b) + (x, y)$	
$= (a + x, b + y)$	b. ?
but $U + W = (a, b)$	c. ?
therefore $a + x = a$ and $b + y = b$	d. ?
so $x = 0$ and $y = 0$	e. ?
and $W = \mathbf{0}$.	f. ?

8. Following the pattern in problem 7, prove that each vector in $\mathscr{V}_2$ has a unique additive inverse.

9. Prove that the following properties of scalar multiplication hold for any vectors $U, V \in \mathscr{V}_2$ and any scalars $r, s \in R$:

   a. $rU \in \mathscr{V}_2$
   b. $r(U + V) = rU + rV$
   c. $(r + s)U = rU + sU$
   d. $(rs)U = r(sU)$
   e. $1U = U$ where 1 is the number 1.

10. Show that, for all $U \in \mathscr{V}_2$: a. $0U = \mathbf{0}$. b. $-U = (-1)U$.

11. Addition in R and therefore in $\mathscr{V}_2$ is both associative and commutative, but not all operations have these properties. Show by examples that if subtraction is viewed as an operation on real numbers, then it is neither commutative nor associative. Do the same for division on the set of all positive real numbers.

§3. Geometric Representation of Vectors of $\mathscr{V}_2$

The definition of an abstract vector space is to be based on the example provided by $\mathscr{V}_2$, and we have now obtained all the properties necessary for the definition. However, aside from the fact that $\mathscr{V}_2$ has a rather simple definition, the preceding discussion gives no indication as to why one would want to study it, let alone something more abstract. Therefore the remainder of this chapter is devoted to examing one of the applications for linear algebra, by drawing the connection between linear algebra and Euclidean geometry. We will first find that the Cartesian plane serves as a good model for algebraic concepts, and then begin to see how algebraic techniques can be used to solve geometric problems.

Let E^2 denote the Cartesian plane, that is the Euclidean plane with a Cartesian coordinate system. Every point of the plane E^2 is named by an ordered pair of numbers, the coordinates of the point, and thus can be used to represent a vector of $\mathscr{V}_2$ pictorially. That is, for every vector $U = (a, b)$, there is a point in E^2 with Cartesian coordinates a and b that can be used as a picture of U. And conversely, for every point with coordinates (x, y) in the plane, there is a vector in $\mathscr{V}_2$ which has x and y as components. Now if the vectors of $\mathscr{V}_2$ are represented as points in E^2, how are the operations of vector addition and scalar multiplication represented?

Suppose $U = (a, b)$ and $V = (c, d)$ are two vectors in $\mathscr{V}_2$, then $U + V = (a + c, b + d)$, and Figure 1 gives a picture of these three vectors in E^2. A little plane geometry shows that when the four vectors $\mathbf{0}$, U, V, and $U + V$ are viewed as points in E^2, they lie at the vertices of a parallelogram, as in Figure 2. Thus the sum of two vectors U and V can be pictured as the fourth vertex of the parallelogram having the two lines from the origin to U and V as two sides.

To see how scalar multiples are represented, let $U = (a, b)$ and $r \in R$, then $rU = (ra, rb)$. If $a \neq 0$, then the components of the scalar multiple rU satisfy the equation $rb = (b/a)ra$. That is, if $rU = (x, y)$, then the components

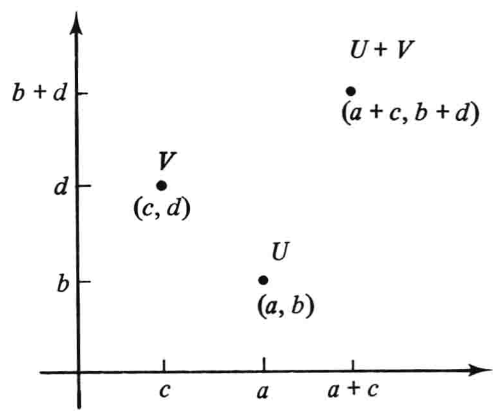

Figure 1