MATRICES and LINEAR ALGEBRA

second edition

Hans Schneider

George Phillip Barker

MATRICES and LINEAR ALGEBRA

second edition

Hans Schneider

University of Wisconsin

George Phillip Barker

University of Missouri

Holt, Rinehart and Winston, Inc.
New York • Chicago • San Francisco • Atlanta
Dallas • Montreal • Toronto • London • Sydney

Library of Congress Catalog Card Number: 72-85179
ISBN: 0-03-091402-7
AMS 1970 Subject Classification Number: 1501
Printed in the United States of America
9 8 7 6 5 4 3 2 1 038 6 5 4 3

MATRICES and LINEAR ALGEBRA

preface

to the *second edition*

The primary difference between this new edition and the first one is the addition of several exercises in each chapter and a brand new section in Chapter 7. The exercises, which are both true-false and multiple choice, will enable the student to test his grasp of the definitions and theorems in the chapter. The new section in Chapter 7 illustrates the geometric content of Sylvester's Theorem by means of conic sections and quadric surfaces.

We would also like to thank the correspondents and students who have brought to our attention various misprints in the first edition that we have corrected in this edition .

MADISON, WISCONSIN — *H. S.*
KANSAS CITY, MISSOURI — *G. P. B.*
OCTOBER 1972

preface

to the first edition

Linear algebra is now one of the central disciplines in mathematics. A student of pure mathematics must know linear algebra if he is to continue with modern algebra or functional analysis. Much of the mathematics now taught to engineers and physicists requires it. It is for this reason that the Committee on Undergraduate Programs in Mathematics recommends that linear algebra be taught early in the undergraduate curriculum. In this book, written mainly for students in physics, engineering, economics, and other fields outside mathematics, we attempt to make the subject accessible to a sophomore or even a freshman student with little mathematical experience. After a short introduction to matrices in Chapter 1, we deal with the solving of linear equations in Chapter 2. We then use the insight gained there to motivate the study of abstract vector spaces in Chapter 3. Chapter 4 deals with determinants. Here we give an axiomatic definition, but quickly develop the determinant as a signed sum of products.

For the last thirty years there has been a vigorous and sometimes acrimonious discussion between the proponents of matrices and those of linear transformation. The controversy now appears somewhat absurd, since the level of abstraction that is appropriate is surely determined by the mathematical goal. Thus, if one is aiming to generalize toward ring theory, one should evidently stress linear transformations. On the other hand, if one is looking for the linear algebra analogue of the classical inequalities, then clearly matrices form the natural setting. From a pedagogical point of view, it seems appropriate to us, in the case of sophomore students, first to deal with matrices. We turn to linear transformations in Chapter 5. In Chapter 6, which deals with eigenvalues and similarity, we do some rapid switching between the matrix and the linear transformation points of view. We use whichever approach seems better at any given time. We feel that a student of linear algebra must acquire the skill of switching from one point of view to another to become proficient in this field.

Chapter 7 deals with inner product spaces. In Chapter 8 we deal with systems of linear differential equations. Obviously, for this chapter (and this chapter only) calculus is a prerequisite. There are at least two good reasons for including some linear differential equations in this linear algebra book. First, a student whose only model for a linear transformation is a matrix does not see why the abstract approach is desirable at all. If he is shown that certain differential operators are linear transformations also, then the point of abstraction becomes much more meaningful. Second, the kind of student we have in mind must become familiar with linear differential equations at some stage in his career, and quite often he is aware of this. We have found in teaching this course at the University of Wisconsin that the promise that the subject we are teaching can be applied to differential equations will motivate some students strongly.

We gratefully acknowledge support from the National Science Foundation under the auspices of the Committee on Undergraduate Programs in Mathematics for producing some preliminary notes in linear algebra. These notes were produced by Ken Casey and Ken Kapp, to whom thanks are also due. Some problems were supplied by Leroy Dickey and Peter Smith. Steve Bauman has taught from a preliminary version of this book, and we thank him for suggesting some improvements. We should also like to thank our publishers, Holt, Rinehart and Winston, and their mathematics editor, Robert M. Thrall. His remarks and criticisms have helped us to improve this book.

MADISON, WISCONSIN *H.S.*
JANUARY 1968 *G.P.B.*

MATRICES and LINEAR ALGEBRA

contents

3 VECTOR SPACES

4 DETERMINANTS

5 LINEAR TRANSFORMATIONS

6 EIGENVALUES AND EIGENVECTORS

chapter 1

The Algebra of Matrices

1. MATRICES: DEFINITIONS

This book is entitled *Matrices and Linear Algebra,* and "linear" will be the most common mathematical term used here. This word has many related meanings, and now we shall explain what a linear equation is. An example of a linear equation is $3x_1 + 2x_2 = 5$, where x_1 and x_2 are unknowns. In general an equation is called linear if it is of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{1.1.1}$$

where $x_1, \cdots, x_n$ are unknowns, and $a_1, \cdots, a_n$ and b are numbers. Observe that in a linear equation, products such as x_1x_2 or $x_3{}^4$ and more general functions such as $\sin x_1$ do not occur.

In elementary books a pair of equations such as

$$\begin{cases} 3x_1 - 2x_2 + 4x_3 = 1 \\ -x_1 + 5x_2 \qquad\quad = -3 \end{cases} \tag{1.1.2}$$

is called a pair of simultaneous equations. We shall call such a pair a *system* of linear equations. Of course we may have more than three unknowns and more than two equations. Thus the most general system of m equations in n unknowns is

$$\begin{array}{ccccccc} a_{11}x_1 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ a_{m1}x_1 & + & \cdots & + & a_{mn}x_n & = & b_m. \end{array} \tag{1.1.3}$$

The a_{ij} are numbers, and the subscript (i, j) denotes that a_{ij} is the coefficient of x_j in the ith equation.

So far we have not explained what the coefficients of the unknowns are, but we have taken for granted that they are real numbers such as 2, $\sqrt{2}$, or π. The coefficients could just as well be complex numbers. This case would arise if we considered the equations

$$\begin{aligned} ix_1 - (2 + i)x_2 &= 1 \\ 2x_1 + (2 - i)x_2 &= -i \\ x_1 + \qquad 2x_2 &= 3. \end{aligned}$$

Note that a real number is also a complex number (with imaginary part zero), but sometimes it is important to consider either all real numbers or all complex numbers. We shall denote the real numbers by R and the complex numbers by C. The reader who is familiar with abstract algebra will note that R and C are fields. In fact, most of our results could be stated for arbitrary fields. (A reader unfamiliar with abstract algebra should ignore the previous two sentences.) Although we are not concerned with such generality, to avoid stating most theorems twice we shall use the symbol F to stand for either the real numbers R or the complex numbers C. Of course we must be consistent in any particular theorem. Thus in any one theorem if F stands for the real numbers in any place, it must stand for the real numbers in all places. Where convenient we shall call F a number system.

In Chapter 2 we shall study systems of linear equations in greater detail. In this chapter we shall use linear equations only to motivate the concept of a matrix. Matrices will turn out to be extremely useful not only in the study of linear equations but also in much else.

If we consider the system of equations (1.1.2), we see that the arrays of coefficients

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

convey all the necessary information. Conversely, given any arrays like

$$\begin{bmatrix} -2 & 3 & 1 \\ 5 & 0 & 2 \\ \sqrt{2} & 1 & 1 \end{bmatrix} \qquad \begin{bmatrix} 4 \\ 3/2 \\ 0 \end{bmatrix}$$

we can immediately write down a corresponding system of equations

$$\begin{aligned} -2x_1 + 3x_2 + 1x_3 &= 4 \\ 5x_1 + 0x_2 + 2x_3 &= 3/2 \\ \sqrt{2}x_1 + 1x_2 + 1x_3 &= 0. \end{aligned}$$

Let F stand for the real or complex numbers. With this as motivation we adopt the following

■ **(1.1.4) DEFINITION** A *matrix* (over F) is a rectangular array of elements of F. The elements of F that occur in the matrix A are called the *entries* of A.

Examples of matrices are

$$\begin{bmatrix} 3 & -2 & 4 \\ -1 & 5 & 0 \end{bmatrix} \quad \begin{bmatrix} i \\ 1+i \end{bmatrix} \quad [3 \quad 7i \quad 0] \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 2+i & 1 \\ 0 & -1-i \\ 3 & 1 \end{bmatrix}.$$

The general form of a matrix over F is

(1.1.5)
$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix},$$

where each a_{ij} is an element of F, that is, either a real or complex number.

The horizontal array

$$[a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}]$$

is called the ith row of A and we shall denote it by a_{i*}. Similarly, the vertical array

$$\begin{bmatrix} a_{1j} \\ a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ a_{mj} \end{bmatrix}$$

is called the jth column of A, and we shall often denote it by a_{*j}. Observe that a_{ij} is the entry of A occurring in the ith row and the jth column.

If the matrix A has m rows and n columns it is called an $m \times n$ *matrix*. In particular, if $m = n$ the matrix is called a *square matrix*. At times an $n \times n$ square matrix is referred to as a matrix of *order n*. Two other special cases are the $m \times 1$ matrix, referred to as a *column vector*, and the $1 \times n$ matrix, which is called a *row vector*. Examples of each special case are

$$\begin{bmatrix} 2 & 1 \\ 5 & 6i \end{bmatrix} \quad \begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix} \quad [2 \quad \pi i \quad -\tfrac{3}{4} + i/5].$$

Usually we denote matrices by capital letters (A, B, and C), but sometimes the symbols $[a_{ij}]$, $[b_{kl}]$, and $[c_{pq}]$ are used. The entries of the matrix A will be denoted by a_{ij}, those of B by b_{kl}, and so forth.

It is important to realize that

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

are all distinct matrices. To emphasize this point, we make the following

■ **(1.1.6) DEFINITION** Let A be an $m \times n$ matrix and B a $p \times q$ matrix. Then $A = B$ if and only if $m = p$, $n = q$, and

$$a_{ij} = b_{ij}, \qquad i = 1, \cdots, m, \quad j = 1, \cdots, n.$$

Thus for each pair of integers m, n we may consider two sets of matrices. One is the set of all $m \times n$ matrices with entries from the set of real numbers R, and the other is the set of all $m \times n$ matrices with entries from the set of complex numbers C. We shall denote the first set by $R_{m,n}$ and the second by $C_{m,n}$. To each theorem about $R_{m,n}$ there often corresponds an analogous theorem that can be obtained by consistently replacing $R_{m,n}$ by $C_{m,n}$. In conformity with our use of the symbol F, we shall write $F_{m,n}$ to stand for either $R_{m,n}$ or $C_{m,n}$.

We shall now point out how the use of matrices simplifies the notation for systems of linear equations. Let

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -1 & 5 & 0 \end{bmatrix} \quad b = \begin{bmatrix} 1 \\ -3 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

All the information in (1.1.2) is contained in A, x, and b. (It is convenient here to call x a matrix or column vector even though its entries are unknowns.) Thus we could, purely symbolically at the moment, write

$$Ax = b.$$

Thus we reduce two linear equations in three unknowns to one matrix equation. If we have m linear equations in n unknowns, as in (1.1.3), we can still use the matrix form

$$Ax = b, \tag{1.1.7}$$

where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix}. \tag{1.1.8}$$

For the time being, (1.1.7) is merely a symbolic way of expressing the equations (1.1.3). As another example, let

$$A = \begin{bmatrix} 3 & 1 & 1-i & 0 \\ 2+i & 0 & 2-i & -7 \\ -i & 1 & 1 & -1 \end{bmatrix} \quad b = \begin{bmatrix} i/2 \\ -i/2 \\ 1 \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}.$$

Then $Ax = b$ is shorthand for

$$\begin{aligned} 3x_1 + 1x_2 + (1 - i)x_3 + 0x_4 &= \frac{i}{2} \\ (2 + i)x_1 + 0x_2 + (2 - i)x_3 - 7x_4 &= -\frac{i}{2} \\ -ix_1 + 1x_2 + 1x_3 \qquad - 1x_4 &= 1. \end{aligned}$$

In Section 3 we shall see that the left side of (1.1.7) may be read as the product of two matrices. Note that b and x are column vectors; b is a column vector with m elements and x is a column vector with n elements. This method of writing the linear equations concentrates attention upon the essential item, the coefficient array.

EXERCISES

1. Find the matrices (A, b, x), corresponding to the following systems of equations.

(a) $$\begin{aligned} 2x_1 - 3x_2 &= 4 \\ 4x_1 + 2x_2 &= -6. \end{aligned}$$

(b) $$\begin{aligned} 7x_1 + 3x_2 - x_3 &= 7 \\ x_1 + x_2 \qquad &= 8 \\ 19x_2 - x_3 &= 17. \end{aligned}$$

(c) $$\begin{aligned} - 4w &= 16 \\ 2x + 3y - 5z + 7w &= 11 \\ z + w &= 5. \end{aligned}$$

(d) $$\begin{aligned} 2x + 3y \qquad\qquad &= 6 \\ y + 4z \qquad &= 7 \\ - z + 5w &= 8 \\ 6x \qquad\qquad + 7w &= 9. \end{aligned}$$

(e) $$\begin{aligned} (3 + 2i)z_1 + (-2 + 4i)z_2 &= 2 + i \\ (4 + 4i)z_1 + (-7 + 7i)z_2 &= 4 - i. \end{aligned}$$

(f) $$\begin{aligned} 3z_1 + (4 - 4i)z_2 &= 6 \\ z_1 + (2 + 2i)z_2 &= 7 - i. \end{aligned}$$

2. What systems of equations correspond to the following pairs of matrices?

(a) $$A = \begin{bmatrix} 3 & 2 \\ 4 & 6 \\ -2 & \pi \end{bmatrix}, b = \begin{bmatrix} 17 \\ 9 \\ \sqrt{2} \end{bmatrix}.$$

(b) $$A = \begin{bmatrix} 2 & 4 & 5 & 3 \\ 6 & 1 & 2 & 7 \end{bmatrix}, b = \begin{bmatrix} -1 \\ 2 \end{bmatrix}.$$