Basic Matrix Theory

LEONARD E. FULLER

Basic Matrix Theory

LEONARD E. FULLER

Professor of Mathematics
Kansas State University

PRENTICE-HALL, INC.

Englewood Cliffs, N. J., 1962

Library of Congress Catalog No. 62-12104

Printed in the United States of America

06518—C

Basic Matrix Theory

PRENTICE-HALL MATHEMATICS SERIES

Albert A. Bennett, EDITOR

Preface

This book has been written for the person who needs to use matrices as a tool; therefore, the mathematics involved has been kept at as simple a level as possible. Any person who has had elementary algebra and has a desire to learn something about matrices can understand the contents. To help the reader, many numerical examples are included to illustrate principles and techniques. Since matrices are viewed as a mathematical tool, throughout the book the emphasis is placed on developing skill in their use. No attempt has been made to give specific applications of the theory discussed because these are so varied that one cannot economically do justice to all of them for a given technique. The reader is urged to refer to the literature in his special area of interest for applications of the topics discussed.

The audience for which this text is intended is quite heterogeneous as can be seen by noting the backgrounds of the students in classes using the notes for this book. There were both undergraduate and graduate students enrolled for credit. The undergraduates were chiefly majors in mathematics, some of whom were planning to be high school teachers. The graduate students were in the various fields of engineering, the physical sciences, the social sciences, statistics, and plant and animal genetics, but no graduate students in mathematics were permitted to take the course. In addition to this group, there were several students and faculty members who audited the course. The faculty members were in the fields of engineering, statistics, mathematics, and genetics.

The material covered falls into two categories. The first four chapters contain the basic concepts of matrix theory, whereas the next three chapters are concerned with numerical computation techniques. The first chapters include topics that form the foundation for the remainder of the book. One of the most important of these topics is the concept of elementary operations. For each type, an easy to remember standard notation is given that also facilitates checking for errors in the reduction of matrices. A systematic procedure is outlined for simplifying matrices using elementary row or elementary column operations, or both. This basic procedure is then modified to give a method for evaluating the determinant of a given square matrix. Some of the numerical computation techniques that are considered also utilize the basic elementary operation process.

Chapters 5 through 7 contain a few of the many standard numerical techniques with no pretense at completeness. Again a standard notation is developed and used throughout to make it easier to apply the processes. Both direct and iterative techniques for inverting matrices and solving systems of linear equations are considered. Detailed numerical examples are worked out to show how to apply each of these methods. The emphasis in these chapters is on understanding not only how to use the procedures but also why they are valid.

The author wishes to express his thanks to his wife Frances for her valuable assistance and her patience during the preparation of this book. He is also deeply indebted to Dr. R. G. Sanger for his careful reading of the manuscript and his many valuable suggestions and comments.

L.E.F.

Contents

1

Basic Properties of Matrices

1.1 Introduction

One of the most widely used mathematical concepts is that of a system of linear equations in certain unknowns. Such a system arises in many diverse situations and in a variety of subjects. For such a system, a set of values for the unknowns that will "satisfy" all the equations is desired. In the language of matrices, a system of linear equations can be written in a very simple form. The use of properties of matrices then makes the solution of the system easier to find.

However, this is not the only reason for studying matrix algebra. The sociologist uses matrices whose elements are zeros or ones in talking about dominance within a group. Closely allied to this application are the matrices arising in the study of communication links between pairs of people. In genetics, the relationship between frequencies of mating types in one generation and those in another can be expressed using matrices. In electrical engineering, network analysis is greatly aided by the use of matrix representations.

Today the language of matrices is spreading to more and more fields as its usefulness is becoming recognized. The reader can probably already call to mind instances in his own field where matrices are used. It is hoped that many more applications will occur to him after this study of matrix algebra is completed.

1.2 The Form of a Matrix

A proper way to being a discussion of matrices would be to give a definition. However, before doing this, it should be noted that a simple definition cannot begin to convey the concept that is involved. For this reason, a two part discussion will follow the definition. The first part will be concerned with trying to convey the nature of the form of a matrix. The second part of the discussion will be concerned with the properties of what are called matrix addition, matrix multiplication, and scalar multiplication. This will establish the basic algebra of matrices. Consider the following definition.

Definition 1.1. A matrix is a rectangular array of numbers of some algebraic system.

What does this mean? It simply means that a matrix is first of all a set of numbers arranged in a pattern that suggests the geometric form of a rectangle. Most of the time this will actually be a square. The algebraic system from which the numbers are chosen will be discussed in more detail later. Some simple examples of matrices are as follows:

$$C = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 1 & 3 \end{bmatrix} \qquad D = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} \qquad E = \begin{bmatrix} 3 & 2 \\ 1 & 4 \\ 5 & 2 \end{bmatrix} \qquad F = \begin{bmatrix} 2 & -1 & 3 \\ 5 & 7 & 9 \\ 4 & 13 & 16 \end{bmatrix}$$

The [] that are used to enclose the array bring out the rectangular form. Sometimes large () are used, whereas other authors prefer double vertical lines instead of the []. Regardless of the notation, the numbers of the array are set apart as an entity by the symbolism. These numbers are often referred to as the *elements* of the matrix. The numbers in a horizontal line constitute a *row* of the matrix, those in a vertical line a *column*. The rows are numbered from the top to the bottom, while the columns are numbered from left to right.

It is sometimes necessary in a discussion to refer to a matrix that has been given. To avoid having to write it out completely every time, it is customary to label matrices with capital letters A, B, C, etc. as was done in the example above. In case the matrix being referred to is a general one, then its elements are often denoted with the corresponding small letters with numerical subscripts. The next example will illustrate this symbolism. With this notation one knows at once that capital letters refer to matrices and small letters to the elements.

When it is necessary in a discussion to talk about a general matrix A, it will be assumed that A consists of a set of mn numbers arranged in n rows with m numbers in each row. The individual elements will be denoted as a_{rs} where the r denotes the row in which the element belongs, and the s denotes the column. In other words, a_{23} will be the element in the second row and third column. The double subscript is the *address* of the element; it tells in which row and in which column the element may be found. This is definite since there can be, in each row, only one element that is also in a given column. Consider then the following examples.

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \end{bmatrix}$$

In these general matrices, notice that the first subscript does denote the row in which the element occurs, whereas the second indicates the column of the entry. When the number of rows and the number of columns are known, the shorter notation $A = (a_{rs})$ is often used. The real significance of this idea will appear several times in this chapter.

Another quite useful concept connected with the form of matrices is given by the following definition.

Definition 1.2. The dimension of a matrix A with n rows and m columns is $n \times m$.

In the numerical examples given before, the dimensions are 2×3, 3×1, 3×2, and 3×3, respectively. In the examples of the general matrices, A is of dimension 4×3 and B is of dimension 3×4. The dimension of a matrix is often referred to as the "size" of the matrix.

A special kind of a matrix is one that has only one row or only one column. These are useful enough to have a special name given to them. This designation is indicated in the next definition.

Definition 1.3. A row vector is a $1 \times m$ matrix. A column vector is an $n \times 1$ matrix.

Using these concepts, a matrix can be thought of as being composed of a set of row vectors placed one under the other. These can be numbered in order from top to bottom so that, in the double subscript notation, the first number refers to the row vector to which the element belongs. Similarly, a matrix can be considered as a set of column vectors placed side by side. If these are numbered from left to right, the second subscript of the address of each element would then refer to a column vector in this set.

There are occasions when reference will be made to the row vectors of a matrix or to the column vectors. In this case, the matrix is to be considered as indicated above. Sometimes the term "vectors" of a matrix will be used. In this case, the reference is to either row vectors or column vectors or both.

1.3 The Transpose of a Matrix

Associated with every $1 \times m$ row vector is an $m \times 1$ column vector. This column vector has the same numbers appearing in the same order as in the row vector. The only difference is that they are written vertically for the column vector and horizontally for the row vector. The column vector is referred to as the *transpose* of the row vector. Also, the row vector is called the transpose of the column vector.

This concept is readily extensible to matrices. With a given matrix A, one can associate a matrix A' known as the transpose of A. The column vectors of A' are the transposes of the corresponding row vectors of A; or, viewed another way, the row vectors of A' are the transposes of the corresponding column vectors of A. This concept is expressed in the next definition in terms of the addresses of the elements.

Definition 1.4. If $A = (a_{rs})$, then $A' = (a_{sr})$.

This definition says that the elements in A' are the same as those of A, with reversed interpretation of the subscripts. In other words, the element of A with the address (r, s) in A' has (s, r) as its address in A.

The concept of the transpose can be made clearer by referring to the previous examples. The matrix C is made up of two row vectors. The transposes of these two vectors are

$$\begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}$$

respectively. This means that C' has these two vectors as its column vectors. In other words,

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

Similarly, the transposes of the other matrices are

$$D' = [4 \quad 1 \quad 3] \qquad E' = \begin{bmatrix} 3 & 1 & 5 \\ 2 & 4 & 2 \end{bmatrix} \qquad F' = \begin{bmatrix} 2 & 5 & 4 \\ -1 & 7 & 13 \\ 3 & 9 & 16 \end{bmatrix}$$

For the two general matrices, the transposes are

$$A' = \begin{bmatrix} a_{11} & a_{21} & a_{31} & a_{41} \\ a_{12} & a_{22} & a_{32} & a_{42} \\ a_{13} & a_{23} & a_{33} & a_{43} \end{bmatrix} \qquad B' = \begin{bmatrix} b_{11} & b_{21} & b_{31} \\ b_{12} & b_{22} & b_{32} \\ b_{13} & b_{23} & b_{33} \\ b_{14} & b_{24} & b_{34} \end{bmatrix}$$

In all of these examples notice how the address of the elements of the original matrix are reversed. Of course, the elements whose row and column addresses are the same have their address unchanged in the transpose. Note too, how the column vectors of the transpose are the corresponding row vectors of the original matrix. This also applies to the row vectors of the transpose, for they are the same as the corresponding column vectors of the original matrix.

It is apparent that if the dimension of A is $n \times m$, then the dimension of A' is $m \times n$. This is a consequence of the definitions of transpose and dimension. For instance, the dimensions of the transposes of the numerical examples are 3×2, 1×3, 2×3, and 3×3, respectively. Similarly, the dimension of A' is 3×4, whereas that of B' is 4×3.

Another consequence of the definition of transpose is $(A')' = A$. In other words, the transpose of the transpose of A is A itself. This becomes apparent on considering the row vectors of A. They form the corresponding column vectors of A'. The column vectors of A' in turn determine the corresponding row vectors of its transpose $(A')'$. But this means that the row vectors of A and $(A')'$ are the same so they are the same matrix.

1.4 Submatrices

The last topic concerned with the form of a matrix to be considered is based on the next definition.

Definition 1.5. A submatrix of a matrix A is an array formed by deleting one or more vectors of A.

The definition does not specify whether the vectors deleted are row vectors or column vectors. It also allows the deletion of a combination of row vectors and column vectors. Some examples of what is meant will illustrate the concept.

The deletion of the second column vector of F gives the submatrix

$$\begin{bmatrix} 2 & 3 \\ 5 & 9 \\ 4 & 16 \end{bmatrix}$$

If the third row vector were also deleted, the resulting submatrix would be

$$\begin{bmatrix} 2 & 3 \\ 5 & 9 \end{bmatrix}$$

If the first and third row vectors and the first column vector of F were deleted, there would result the row vector $[7 \quad 9]$. It can be easily seen that there are a variety of submatrices that can be formed from a given matrix.

If the matrix is square, one can draw a diagonal from the upper left corner to the lower right corner. This line would pass through those elements whose row and column subscripts are the same. These elements are known as the *diagonal* elements. For the matrix F, 2, 7 and 16 are the diagonal elements. One can form submatrices by deleting corresponding row and column vectors. If the original matrix is square, the submatrices formed are also square. More important however, their diagonal elements are diagonal elements of the original matrix. These are called principal submatrices. For the matrix F, the principal submatrices are

$$\begin{bmatrix} 2 & -1 \\ 5 & 7 \end{bmatrix} \quad \begin{bmatrix} 2 & 3 \\ 4 & 16 \end{bmatrix} \quad \begin{bmatrix} 7 & 9 \\ 13 & 16 \end{bmatrix} \quad [2] \quad [7] \quad [16]$$

In the first three of these principal submatrices, a single row and column are deleted from the original matrix; in the last three, two rows and columns are deleted. In all six of these matrices, the diagonal elements are diagonal elements of F.

1.5 The Elements of a Matrix

In the definition of a matrix it was stated that the elements belong to an algebraic system. In nearly all of the work that follows, the elements will be real numbers; however, on occasion, they may be complex numbers. In either of these cases, the elements will belong to an algebraic system known as a *field*. These numbers of the algebraic system are often referred to as *scalars*. In the section of this chapter on partitioning, the elements will be matrices of smaller size. This way of considering a matrix can be very useful. Later chapters will have sections depending upon partitioning of matrices.

For matrices whose elements belong to the same algebraic system it is possible to define equality of matrices. Notice that the following definition gives this in terms of equality among the elements.

Definition 1.6. If A and B are both of dimension $n \times m$, and $a_{rs} = b_{rs}$ for all r and s then $A = B$.

This definition says that the matrices must be of the same size and equal element by element. The first requirement is actually implied by the second and is included only for clarity. It should be noted how the definition is made to depend upon the corresponding property of the elements. This will be characteristic of nearly all of the properties of matrices that will be discussed.

1.6 The Algebra of Real Numbers

The arrays of numbers that form a matrix are of little use by themselves. One has to be able to manipulate them according to a set of rules. The set of rules in this case consists of the definitions of three operations. As a consequence of these definitions each of the operations has some important properties. These operations will be known as *addition of matrices*, *multiplication of matrices*, and *scalar multiplication of matrices*. It should be obvious that the first two cannot be the familiar operations upon real numbers that bear these names. However, these new operations will be