INTRODUCTION TO

Linear Algebra and Differential Equations

SAMUEL WOLFENSTEIN

University of Paris

HOLDEN-DAY

San Francisco, Cambridge, London, Amsterdam

Library of Congress Catalog Card Number: 75-75915
Printed in the United States of America

INTRODUCTION TO

Linear Algebra and Differential Equations

For Marlene, Robert and Caroline

PREFACE

"An Introduction to linear algebra with particular application to the theory of linear ordinary differential equations" would be a more accurate but less catchy title. However, although this is essentially a textbook on linear algebra, enough material on analysis has been included to make it self-contained for the reader with no mathematical background beyond the rudiments of calculus. Thus, the title as it stands is not misleading.

This book originates as a practical attempt to implement recommendations of the Committee on the Undergraduate Program in Mathematics, in particular those of its Panel on Physical Sciences and Engineering. This group urged, among other things, that the undergraduate training of all physicists and engineers should include, at the earliest feasible date in their education, a course in linear algebra, and suggested further that "it may be desirable, for mathematical or scheduling purposes, to combine beginning analysis and linear algebra into a single coordinated course to be completed in the sophomore year." Such a course, following the Committee's recommendations,* would include, in addition to the material covered in any standard calculus textbook, the following subjects in analysis:

"(a) Topics in differential equations, including linear differential equations with constant coefficients and first-order systems—linear algebra (including eigenvalue theory, see c below) should be used to treat both homogeneous

* Recommendations on the Undergraduate Mathematics Program for Engineers and Physicists, CUPM, 1962 (rev. ed. 1967).

and nonhomogeneous problems; first-order linear and nonlinear equations, with Picard's method and an introduction to numerical techniques.

(b) Some attempt should be made to fill the gap between the high-school algebra of complex numbers and the use of complex exponentials in the solution of differential equations. In particular some work on the calculus of complex-valued functions of a real variable should be included ...

(c) Vectors in two and three dimensions and the differentiation of vector-valued functions of one variable. [With regard to algebra, the Committee suggests] a course with strong emphasis on the geometrical interpretation of vectors and matrices with applications to mathematics, physics, and engineering. Topics should include the algebra and geometry of vector spaces, linear transformations and matrices, linear equations ... quadratic forms and symmetric matrices, and elementary eigenvalue theory".

Shortly after the publication of this report, when the author was asked to organize the mathematics curriculum of a small experimental college for American students in Europe, he decided to offer a coordinated freshman–sophomore course along the general lines recommended by the CUPM. In the absence of an existing textbook, special material on linear algebra and differential equations was prepared to supplement the traditional-type textbook used for the calculus course. It is this material, modified in the light of four years of classroom testing, that is presented here. A course of approximately a semester based on this text and following any ordinary three-semester calculus course (no particular type of preparation is presupposed) will come close to meeting the CUPM's recommendations.

The book divides naturally into three parts. The first part introduces the notions of differential equations that will provide examples and applications of the linear algebra to be developed in the central chapters. After an introductory chapter on complex-valued and vector-valued functions of a real variable, Chapter 2 introduces ordinary differential equations. The form $Y' = F(x, Y)$, where Y is a vector-valued function of a real variable, is shown to subsume all the cases we are interested in. Existence and uniqueness theorems are stated for this form of equation and their consequences explored. The polygonal method of numerical approximation is described, and a proof of the existence–uniqueness theorems is given based on the convergence of this procedure (a simplified version of Cauchy's original proof). This seems better than Picard's proof, which provides no reason for supposing that the type of approximation procedure which is used in practice will actually work. However, Picard's method is outlined in the exercises.

The remainder of this first part, Chapters 3 and 4, provide the specific examples that will be needed further along; that is, the solution in closed form of first-order linear equations and of homogeneous linear equations with constant coefficients. Some additional material on " solution " of equations is expounded (special equations that can be solved in closed form, formal technique of series solution, some practical methods for nonhomogeneous linear equations with constant coefficients). There are also a very few simple, physical applications, including damped harmonic motion and electric circuits, with use of the complex exponential function. It will be found that these introductory chapters contain virtually all the material traditionally taught in a course on "elementary" differential equations. Hence use of this textbook would be in line with the CUPM recommendation (see "A General Curriculum in Mathematics for Colleges"*) that this be eliminated as a separate course.

Chapters 5 through 8, the central chapters of the book, present vector spaces, inner-product spaces, linear transformations and matrices. The presentation is axiomatic and (it is hoped) rigorous, but with systematic use of geometrical illustration. It is found that even students having no previous acquaintance with abstract algebra can follow the material when it is presented in this manner. Each point is illustrated by many examples, taken from spaces of n-tuples, of polynomials and of functions. Thus, from the outset, the student sees the manifold applications of a single theory and the advantage of the abstract approach. In Chapter 7, the general theory of linear differential equations is worked out as illustration of a theorem on linear transformations, while the method of " variation of parameters " is presented in the following chapter as an application of matrix algebra. The chapter on inner-product spaces also contains an introduction to Fourier expansions.

The last part, Chapters 9 and 10, contains an introduction to eigenvalue theory, with the obvious application to systems of linear differential equations with constant coefficients. Determinants are introduced for the first time in Chapter 9 and (following the CUPM recommendations) "treated with all possible brevity." " Reduction " of square matrices is approached through the theory of invariant subspaces. A brief, final section deals with symmetric matrices and quadratic forms.

Several appendices contain material for reference or for supplementing the text. Appendix 3, in particular, on the Gauss–Jordan reduction scheme, should be familiar to students from the modern high-school programs. Some teachers might wish to present it at the beginning of the linear algebra before going on to abstract vector spaces.

Some of the exercises provide examples to test whether the text has been assimilated. Others provide practice in computation or develop supplementary results. Some of the latter (Peano's existence theorem, duality of vector

* Report to the Mathematical Association of America, *CUPM*, 1965, p. 13.

spaces and the Jordan canonical form, for example) are fairly ambitious projects and could be assigned to gifted students, perhaps for class presentation.

The author's debts to previous writers on the subjects presented in this book are too many and too obvious to be acknowledged here. It is a particular pleasure to thank N. Bourbaki for permission to reproduce the passage quoted on p. 94. Finally a special word of thanks is in order to the American students in Europe who not only served as guinea pigs for the testing of this material, but also offered many pertinent and useful suggestions.

S. W. Paris-Arcueil, July 1968

INTRODUCTION TO

Linear Algebra and Differential Equations

CONTENTS

1

VECTOR FUNCTIONS AND COMPLEX FUNCTIONS

Since it is desirable in mathematics (and perhaps even in other realms of discourse) to know precisely what one is talking about, this preliminary chapter will define concepts and fix notation to be employed consistently hereafter. It is assumed that the reader is acquainted with the simpler properties of real numbers encountered in courses in arithmetic, algebra and beginning calculus. Most of the propositions stated in this chapter follow easily from the definitions and these simple properties. For a few more advanced results, requiring properties of real numbers perhaps less familiar to the reader, proofs are sketched in the exercises at the end of the chapter.

1. THE SPACES $\mathbf{R}^m$

Let $S_1, S_2, \ldots, S_m$ be any m sets, where m is a positive integer. We define their Cartesian product (in that order) and let $S_1 \times S_2 \times \cdots \times S_m$ denote the set of all ordered m-tuples $(x_1, x_2, \ldots, x_m)$ such that $x_1 \in S_1, x_2 \in S_2, \ldots, x_m \in S_m$. We recall that two ordered m-tuples are equal if, and only if, all of their corresponding components are equal; in other words, $(x_1, x_2, \ldots, x_m) = (y_1, y_2, \ldots, y_m)$ says the same thing as: $x_1 = y_1, x_2 = y_2, \ldots, x_m = y_m$.

All the sets S_i above need not be distinct. If they are all identical, for example, if $S_i = S$, $i = 1, \ldots, m$, the Cartesian product is written S^m. We agree that $S^1 = S$ (and we do not define S^0).

In what follows, we shall be particularly interested in the products $\mathbf{R}^m$, where $\mathbf{R}$ is the set of all real numbers. $\mathbf{R}^m$ then is the set of all ordered m-tuples of real numbers. When $m = 1$, 2 or 3, $\mathbf{R}^m$ can be represented pictorially or geometrically in a familiar manner (*see* Fig. 1-1). For larger values of m, this is no longer possible, but geometrical language may still be used.

Let us speak of $\mathbf{R}^m$ as a "real m-dimensional space" or "real m space," and call the elements of $\mathbf{R}^m$ "points" or alternatively, "vectors." These terms are,

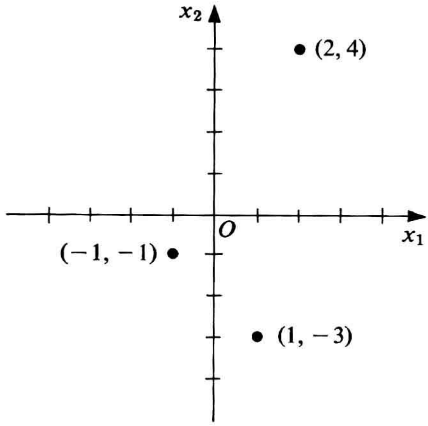

Figure 1-1

for the present, strictly synonymous. Our refraining from such incongruous terminology as "length of a point" is for aesthetic not mathematical reasons.

Let us denote vectors by capital letters; thus, $X = (x_1, \ldots, x_m)$. The ith component x_i of X (also called ith *coordinate* or ith *projection*) will be denoted by $\mathrm{pr}_i X$. For each i, $i = 1, \ldots, m$, pr_i is a function whose domain is $\mathbf{R}^m$ and whose range is $\mathbf{R}$.

Vectors are added componentwise. That is to say, if $X = (x_1, \ldots, x_m)$ and $Y = (y_1, \ldots, y_m)$, we *define*:

$$X + Y = (x_1 + y_1, \ldots, x_m + y_m) \tag{1}$$

Note that addition is defined only when the two vectors have the same number of components. Subtraction of vectors is defined in the same way.

In any space $\mathbf{R}^m$, the vector $(0, 0, \ldots, 0)$ has special properties which are easily verified. Thus, for any vector X, $X - X = (0, \ldots, 0)$ and $X + (0, \ldots, 0) = X$. To provide a slightly less cumbersome name for this interesting object, we define 0_m to be the vector of $\mathbf{R}^m$ each of whose components is 0. We shall *always* omit the subscript "m" when no ambiguity results from doing so. That is, we write $X - X = 0$, $X + 0 = X$.

Other vectors for which we want a special symbol are those of the form $(0, \ldots, 0, 1, 0, \ldots, 0)$. We shall define $E_{j,m}$ to be the vector of $\mathbf{R}^m$ whose jth component is 1 and each of whose other components is 0, and, again, we shall drop the second subscript whenever possible. Thus, in $\mathbf{R}^4$ for example, $E_3 = (0, 0, 1, 0)$.

If c is any real number, and X is defined as above, we *define*:

$$cX = (cx_1, \ldots, cx_m) \tag{2}$$

We do not define the symbol "Xc". It is obvious that, for any vector X,

$0X = 0$ (where the use of the symbol "0" to denote two different objects in the same equation cannot give rise to any ambiguity, since, in each case, there is only one interpretation that makes sense), and

$$0 - X = (-1)X \tag{3}$$

We shall denote the common value of the two members of (3) by the symbol $-X$. Note also the following identity, for any X of $\mathbf{R}^m$:

$$X = \sum_{i=1}^{m} (\mathrm{pr}_i X)E_i \tag{4}$$

Observe that an equation between vectors, with indicated additions and multiplications, is strictly equivalent to a finite number of equations between real numbers. Thus, for example,

$$(2, 1, -4) - 2(1, 2, -2) + 3E_2 = 0$$

says nothing more or less than that $2 - 2(1) + 3(0) = 0$, $1 - 2(2) + 3(1) = 0$ and $-4 - 2(-2) + 3(0) = 0$. Performing operations of this kind requires only a knowledge of elementary arithmetic, so it hardly seems worthwhile to propose further numerical examples. If the reader can work one such problem, he can work them all.

Authors of "plane analytic geometry" books claim to prove that they can define a one–one correspondence between the "points of a plane" and the elements of $\mathbf{R}^2$ in such a way that, if (x_1, x_2) and (y_1, y_2) correspond to the "points" A and B, the "distance from A to B" is equal to $\sqrt{[(x_1 - y_1)^2 + (x_2 - y_2)^2]}$. Similar claims are made in books on "solid analytic geometry." We shall not examine these claims here, but rather we shall *define*, for any $X = (x_1, \ldots, x_m)$ of $\mathbf{R}^m$, $|X|$ to be the non-negative real number, called the *absolute value* (or *length*, *magnitude*, *Euclidean norm*) of X, such that

$$|X|^2 = {x_1}^2 + {x_2}^2 + \cdots + {x_m}^2 \tag{5}$$

(Note that, when $m = 1$, $|x|$ is the absolute value with which you are already familiar.) Given two points X and Y, $|X - Y|$ will be called the *Euclidean distance* between them. Thus we have

$$|X - Y| = \sqrt{[(x_1 - y_1)^2 + \cdots + (x_m - y_m)^2]} \tag{6}$$

as a matter of definition.

There are other possible ways of defining distance. We might, for example, consider, for the distance between X and Y, the apparently simpler formula:

$$|x_1 - y_1| + \cdots + |x_m - y_m|$$

This formula would be appropriate (within $m = 2$) for measuring the distance from house to house if one lived in a city with square blocks and traveled by

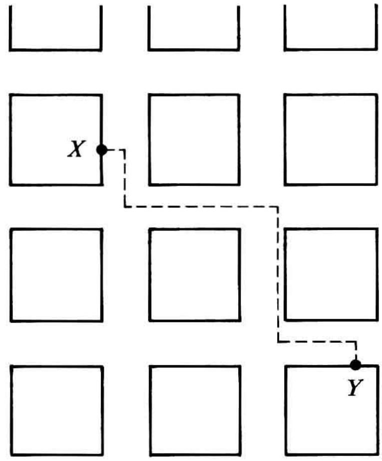

Figure 1-2

car rather than by helicopter (Fig. 1-2). However, we shall not pursue considerations of this type, since our purpose in discussing distance in $\mathbf{R}^m$ is not to plot itineraries in "hyperspace," but rather to extend the notion of *limit* to $\mathbf{R}^m$ for $m > 1$. For real numbers, this concept was defined in terms of the absolute value $|x - y|$, which, intuitively, is the "distance" between x and y on the "number line." What we want to do in $\mathbf{R}^m$ is to use the same definitions over again, only replacing $|x - y|$ by an appropriate function of X and Y. The Euclidean distance could be used for this purpose, as could the "city-block" distance considered above. We shall in fact not use the latter at all; but we shall define a third distance function that has all the desirable properties of the other two and is considerably easier to work with than either.

For any vector $X = (x_1, \ldots, x_m)$, we define the *maximum norm* of X, written $\|X\|$, by

$$\|X\| = \max_{i=1,\ldots,m} |x_i| \tag{7}$$

(where for typographical reasons we shall omit the subscripts beneath the "max" whenever feasible). In words: the maximum norm of a vector is the absolute value of its numerically largest component.

In this book, the word "norm," without any qualifying adjective, will always refer to the maximum norm, and we shall not use the double vertical bars to denote anything else. (For a more general definition of norm, see Exercise 6 at the end of this chapter.) Other writers adopt other conventions.