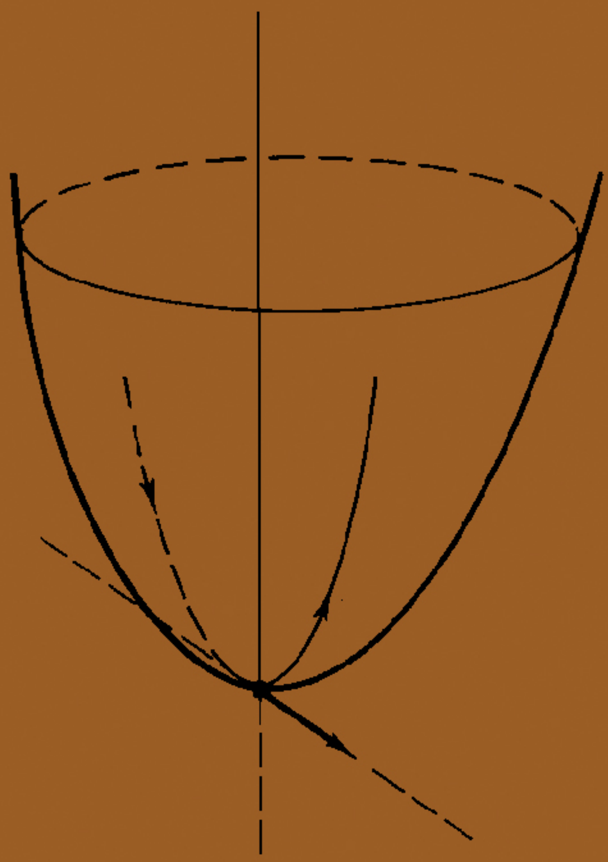

Advanced Calculus of Several Variables

C. H. EDWARDS, JR.

Advanced Calculus
of Several Variables

This page intentionally left blank

Advanced Calculus of Several Variables

C. H. EDWARDS, JR.
THE UNIVERSITY OF GEORGIA

ACADEMIC PRESS New York and London
A Subsidiary of Harcourt Brace Jovanovich, Publishers

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1

Library Library of Congress Cataloging in Publication Data

Edwards, Charles Henry, DATE
Advanced calculus of several variables.

Bibliography: p.
1. Calculus. I. Title.
QA303.E22 515 72-9325
ISBN 0-12-232550-8

AMS (MOS) 1970 Subject Classifications: 26A57, 26A60, 26A63, 26A66

PRINTED IN THE UNITED STATES OF AMERICA

To My Parents

This page intentionally left blank

CONTENTS

IV Multiple Integrals

V Line and Surface Integrals; Differential Forms and Stokes' Theorem

VI The Calculus of Variations

PREFACE

This book has developed from junior–senior level advanced calculus courses that I have taught during the past several years. It was motivated by a desire to provide a modern conceptual treatment of multivariable calculus, emphasizing the interplay of geometry and analysis viä linear algebra and the approximation of nonlinear mappings by linear ones, while at the same time giving equal attention to the classical applications and computational methods that are responsible for much of the interest and importance of this subject.

In addition to a satisfactory treatment of the theory of functions of several variables, the reader will (hopefully) find evidence of a healthy devotion to matters of exposition as such—for example, the extensive inclusion of motivational and illustrative material and applications that is intended to make the subject attractive and accessible to a wide range of "typical" science and mathematics students. The many hundreds of carefully chosen examples, problems, and figures are one result of this expository effort.

This book is intended for students who have completed a standard introductory calculus sequence. A slightly faster pace is possible if the students' first course included some elementary multivariable calculus (partial derivatives and multiple integrals). However this is not essential, since the treatment here of multivariable calculus is fully self-contained. We do not review single-variable calculus, with the exception of Taylor's formula in Section II.6 (Section 6 of Chapter II) and the fundamental theorem of calculus in Section IV.1.

Chapter I deals mainly with the linear algebra and geometry of Euclidean n-space $\mathscr{R}^n$. With students who have taken a typical first course in elementary linear algebra, the first six sections of Chapter I can be omitted; the last two sections of Chapter I deal with limits and continuity for mappings of Euclidean spaces, and with the elementary topology of $\mathscr{R}^n$ that is needed in calculus. The only linear algebra that is actually needed to start Chapter II is a knowledge of the correspondence between linear mappings and matrices. With students having this minimal knowledge of linear algebra, Chapter I might (depending upon the taste of the instructor) best be used as a source for reference as needed.

Chapters II through V are the heart of the book. Chapters II and III treat multivariable differential calculus, while Chapters IV and V treat multivariable integral calculus.

In Chapter II the basic ingredients of single-variable differential calculus are generalized to higher dimensions. We place a slightly greater emphasis than usual on maximum–minimum problems and Lagrange multipliers—experience has shown that this is pedagogically sound from the standpoint of student motivation. In Chapter III we treat the fundamental existence theorems of multivariable calculus by the method of successive approximations. This approach is equally adaptable to theoretical applications and numerical computations.

Chapter IV centers around Sections 4 and 5 which deal with iterated integrals and change of variables, respectively. Section IV.6 is a discussion of improper multiple integrals. Chapter V builds upon the preceding chapters to give a comprehensive treatment, from the viewpoint of differential forms, of the classical material associated with line and surface integrals, Stokes' theorem, and vector analysis. Here, as throughout the book, we are not concerned solely with the development of the theory, but with the development of conceptual understanding and computational facility as well.

Chapter VI presents a modern treatment of some venerable problems of the calculus of variations. The first part of the Chapter generalizes (to normed vector spaces) the differential calculus of Chapter II. The remainder of the Chapter treats variational problems by the basic method of "ordinary calculus"—equate the first derivative to zero, and then solve for the unknown (now a function). The method of Lagrange multipliers is generalized so as to deal in this context with the classical isoperimetric problems.

There is a sense in which the exercise sections may constitute the most important part of this book. Although the mathematician may, in a rapid reading, concentrate mainly on the sequence of definitions, theorems and proofs, this is not the way that a textbook is read by students (nor is it the way a course should be taught). The student's actual course of study may be more nearly defined by the problems than by the textual material. Consequently, those ideas and concepts that are not dealt with by the problems may well remain unlearned by the students. For this reason, a substantial portion of my effort has gone into the approximately 430 problems in the book. These are mainly concrete computational problems, although not all routine ones, and many deal with physical applications. A proper emphasis on these problems, and on the illustrative examples and applications in the text, will give a course taught from this book the appropriate intuitive and conceptual flavor.

I wish to thank the successive classes of students who have responded so enthusiastically to the class notes that have evolved into this book, and who have contributed to it more than they are aware. In addition, I appreciate the excellent typing of Janis Burke, Frances Chung, and Theodora Schultz.

Advanced Calculus of Several Variables

This page intentionally left blank

I

Euclidean Space and Linear Mappings

Introductory calculus deals mainly with real-valued functions of a single variable, that is, with functions from the real line $\mathscr{R}$ to itself. Multivariable calculus deals in general, and in a somewhat similar way, with mappings from one Euclidean space to another. However a number of new and interesting phenomena appear, resulting from the rich geometric structure of n-dimensional Euclidean space $\mathscr{R}^n$.

In this chapter we discuss $\mathscr{R}^n$ in some detail, as preparation for the development in subsequent chapters of the calculus of functions of an arbitrary number of variables. This generality will provide more clear-cut formulations of theoretical results, and is also of practical importance for applications. For example, an economist may wish to study a problem in which the variables are the prices, production costs, and demands for a large number of different commodities; a physicist may study a problem in which the variables are the coordinates of a large number of different particles. Thus a "real-life" problem may lead to a high-dimensional mathematical model. Fortunately, modern techniques of automatic computation render feasible the numerical solution of many high-dimensional problems, whose manual solution would require an inordinate amount of tedious computation.

1 THE VECTOR SPACE $\mathscr{R}^n$

As a set, $\mathscr{R}^n$ is simply the collection of all ordered n-tuples of real numbers. That is,

$$\mathscr{R}^n = \{(x_1, x_2, \ldots, x_n) : \text{each } x_i \in \mathscr{R}\}.$$

Recalling that the Cartesian product $A \times B$ of the sets A and B is by definition the set of all pairs (a, b) such that $a \in A$ and $b \in B$, we see that $\mathscr{R}^n$ can be regarded as the Cartesian product set $\mathscr{R} \times \cdots \times \mathscr{R}$ (n times), and this is of course the reason for the symbol $\mathscr{R}^n$.

The geometric representation of $\mathscr{R}^3$, obtained by identifying the triple (x_1, x_2, x_3) of numbers with that point in space whose coordinates with respect to three fixed, mutually perpendicular "coordinate axes" are x_1, x_2, x_3 respectively, is familiar to the reader (although we frequently write (x, y, z) instead of (x_1, x_2, x_3) in three dimensions). By analogy one can imagine a similar geometric representation of $\mathscr{R}^n$ in terms of n mutually perpendicular coordinate axes in higher dimensions (however there is a valid question as to what "perpendicular" means in this general context; we will deal with this in Section 3).

The elements of $\mathscr{R}^n$ are frequently referred to as *vectors*. Thus a vector is simply an n-tuple of real numbers, and *not* a directed line segment, or equivalence class of them (as sometimes defined in introductory texts).

The set $\mathscr{R}^n$ is endowed with two algebraic operations, called *vector addition* and *scalar multiplication* (numbers are sometimes called scalars for emphasis). Given two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathscr{R}^n$, their *sum* $\mathbf{x} + \mathbf{y}$ is defined by

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n),$$

that is, by coordinatewise addition. Given $a \in \mathscr{R}$, the *scalar multiple* $a\mathbf{x}$ is defined by

$$a\mathbf{x} = (ax_1, \ldots, ax_n).$$

For example, if $\mathbf{x} = (1, 0, -2, 3)$ and $\mathbf{y} = (-2, 1, 4, -5)$ then $\mathbf{x} + \mathbf{y} = (-1, 1, 2, -2)$ and $2\mathbf{x} = (2, 0, -4, 6)$. Finally we write $\mathbf{0} = (0, \ldots, 0)$ and $-\mathbf{x} = (-1)\mathbf{x}$, and use $\mathbf{x} - \mathbf{y}$ as an abbreviation for $\mathbf{x} + (-\mathbf{y})$.

The familiar associative, commutative, and distributive laws for the real numbers imply the following basic properties of vector addition and scalar multiplication:

V1 $\mathbf{x} + (\mathbf{y} + \mathbf{z}) = (\mathbf{x} + \mathbf{y}) + \mathbf{z}$
V2 $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
V3 $\mathbf{x} + \mathbf{0} = \mathbf{x}$
V4 $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
V5 $(ab)\mathbf{x} = a(b\mathbf{x})$
V6 $(a + b)\mathbf{x} = a\mathbf{x} + b\mathbf{x}$
V7 $a(\mathbf{x} + \mathbf{y}) = a\mathbf{x} + a\mathbf{y}$
V8 $1\mathbf{x} = \mathbf{x}$

(Here $\mathbf{x}, \mathbf{y}, \mathbf{z}$ are arbitrary vectors in $\mathscr{R}^n$, and a and b are real numbers.) V1–V8 are all immediate consequences of our definitions and the properties of $\mathscr{R}$. For example, to prove V6, let $\mathbf{x} = (x_1, \ldots, x_n)$. Then

$$\begin{aligned}(a+b)\mathbf{x} &= ((a+b)x_1, \ldots, (a+b)x_n)\\ &= (ax_1 + bx_1, \ldots, ax_n + bx_n)\\ &= (ax_1, \ldots, ax_n) + (bx_1, \ldots, bx_n)\\ &= a\mathbf{x} + b\mathbf{x}.\end{aligned}$$

The remaining verifications are left as exercises for the student.

A *vector space* is a set V together with two mappings $V \times V \to V$ and $\mathscr{R} \times V \to V$, called vector addition and scalar multiplication respectively, such that V1–V8 above hold for all $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z} \in V$ and a, $b \in \mathscr{R}$ (V3 asserts that there exists $\mathbf{0} \in V$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in V$, and V4 that, given $\mathbf{x} \in V$, there exists $-\mathbf{x} \in V$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$). Thus V1–V8 may be summarized by saying that $\mathscr{R}^n$ is a vector space. For the most part, all vector spaces that we consider will be either Euclidean spaces, or subspaces of Euclidean spaces.

By a *subspace* of the vector space V is meant a subset W of V that is itself a vector space (with the same operations). It is clear that the subset W of V is a subspace if and only if it is "closed" under the operations of vector addition and scalar multiplication (that is, the sum of any two vectors in W is again in W, as is any scalar multiple of an element of W)—properties V1–V8 are then inherited by W from V. Equivalently, W is a subspace of V if and only if any linear combination of two vectors in W is also in W (why?). Recall that a *linear combination* of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is a vector of the form $a_1\mathbf{v}_1 + \cdots + a_k\mathbf{v}_k$, where the $a_i \in \mathscr{R}$. The *span* of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathscr{R}^n$ is the set S of all linear combinations of them, and it is said that S is *generated* by the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$.

Example 1 $\mathscr{R}^n$ is a subspace of itself, and is generated by the *standard basis vectors*

$$\begin{aligned}\mathbf{e}_1 &= (1, 0, 0, \ldots, 0),\\ \mathbf{e}_2 &= (0, 1, 0, \ldots, 0),\\ &\vdots\\ \mathbf{e}_n &= (0, 0, 0, \ldots, 0, 1),\end{aligned}$$

since $(x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_2 + x_1\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$. Also the subset of $\mathscr{R}^n$ consisting of the zero vector alone is a subspace, called the *trivial* subspace of $\mathscr{R}^n$.

Example 2 The set of all points in $\mathscr{R}^n$ with last coordinate zero, that is, the set of all $(x_1, \ldots, x_{n-1}, 0) \in \mathscr{R}^n$, is a subspace of $\mathscr{R}^n$ which may be identified with $\mathscr{R}^{n-1}$.

Example 3 Given $(a_1, a_2, \ldots, a_n) \in \mathscr{R}^n$, the set of all $(x_1, x_2, \ldots, x_n) \in \mathscr{R}^n$ such that $a_1x_1 + \cdots + a_nx_n = 0$ is a subspace of $\mathscr{R}^n$ (see Exercise 1.1).

Example 4 The span S of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathscr{R}^n$ is a subspace of $\mathscr{R}^n$ because, given elements $\mathbf{a} = \sum_1^k a_i \mathbf{v}_i$ and $\mathbf{b} = \sum_1^k b_i \mathbf{v}_i$ of S, and real numbers r and s, we have $r\mathbf{a} + s\mathbf{b} = \sum_1^k (ra_i + sb_i)\mathbf{v}_i \in S$.

Lines through the origin in $\mathscr{R}^3$ are (essentially by definition) those subspaces of $\mathscr{R}^3$ that are generated by a single nonzero vector, while planes through the origin in $\mathscr{R}^3$ are those subspaces of $\mathscr{R}^3$ that are generated by a pair of non-collinear vectors. We will see in the next section that every subspace V of $\mathscr{R}^n$ is generated by some finite number, at most n, of vectors; the dimension of the subspace V will be defined to be the minimal number of vectors required to generate V. Subspaces of $\mathscr{R}^n$ of all dimensions between 0 and n will then generalize lines and planes through the origin in $\mathscr{R}^3$.

Example 5 If V and W are subspaces of $\mathscr{R}^n$, then so is their intersection $V \cap W$ (the set of all vectors that lie in both V and W). See Exercise 1.2.

Although most of our attention will be confined to subspaces of Euclidean spaces, it is instructive to consider some vector spaces that are not subspaces of Euclidean spaces.

Example 6 Let $\mathscr{F}$ denote the set of all real-valued functions on $\mathscr{R}$. If $f + g$ and af are defined by $(f+g)(x) = f(x) + g(x)$ and $(af)(x) = af(x)$, then $\mathscr{F}$ is a vector space (why?), with the zero vector being the function which is zero for all $x \in \mathscr{R}$. If $\mathscr{C}$ is the set of all continuous functions and $\mathscr{P}$ is the set of all polynomials, then $\mathscr{P}$ is a subspace of $\mathscr{C}$, and $\mathscr{C}$ in turn is a subspace of $\mathscr{F}$. If $\mathscr{P}_n$ is the set of all polynomials of degree at most n, then $\mathscr{P}_n$ is a subspace of $\mathscr{P}$ which is generated by the polynomials $1, x, x^2, \ldots, x^n$.

Exercises

1.1 Verify Example 3.

1.2 Prove that the intersection of two subspaces of $\mathscr{R}^n$ is also a subspace.

1.3 Given subspaces V and W of $\mathscr{R}^n$, denote by $V + W$ the set of all vectors $v + w$ with $v \in V$ and $w \in W$. Show that $V + W$ is a subspace of $\mathscr{R}^n$.

1.4 If V is the set of all $(x, y, z) \in \mathscr{R}^3$ such that $x + 2y = 0$ and $x + y = 3z$, show that V is a subspace of $\mathscr{R}^3$.

1.5 Let $\mathscr{D}_0$ denote the set of all differentiable real-valued functions on $[0, 1]$ such that $f(0) = f(1) = 0$. Show that $\mathscr{D}_0$ is a vector space, with addition and multiplication defined as in Example 6. Would this be true if the condition $f(0) = f(1) = 0$ were replaced by $f(0) = 0, f(1) = 1$?

1.6 Given a set S, denote by $\mathscr{F}(S, \mathscr{R})$ the set of all real-valued functions on S, that is, all maps $S \to R$. Show that $\mathscr{F}(S, \mathscr{R})$ is a vector space with the operations defined in Example 6. Note that $\mathscr{F}(\{1, \ldots, n\}, \mathscr{R})$ can be interpreted as $\mathscr{R}^n$ since the function $\varphi \in \mathscr{F}(\{1, \ldots, n\}, \mathscr{R})$ may be regarded as the n-tuple $(\varphi(1), \varphi(2), \ldots, \varphi(n))$.

2 SUBSPACES OF $\mathscr{R}^n$

In this section we will define the dimension of a vector space, and then show that $\mathscr{R}^n$ has precisely $n - 1$ types of *proper* subspaces (that is, subspaces other than $\mathbf{0}$ and $\mathscr{R}^n$ itself)—namely, one of each dimension 1 through $n - 1$.

In order to define dimension, we need the concept of linear independence. The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are said to be *linearly independent* provided that no one of them is a linear combination of the others; otherwise they are *linearly dependent.* The following proposition asserts that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly independent if and only if $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$ implies that $x_1 = x_2 = \cdots = x_k = 0$. For example, the fact that $x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n = (x_1, x_2, \ldots, x_n)$ then implies immediately that the standard basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in $\mathscr{R}^n$ are linearly independent.

Proposition 2.1 The vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent if and only if there exist numbers $x_1, x_2, \ldots, x_k$, not all zero, such that $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$.

PROOF If there exist such numbers, suppose, for example, that $x_1 \neq 0$. Then

$$\mathbf{v}_1 = -\frac{x_2}{x_1}\mathbf{v}_2 - \cdots - \frac{x_k}{x_1}\mathbf{v}_k,$$

so $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are linearly dependent. If, conversely, $\mathbf{v}_1 = a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k$, then we have $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k = \mathbf{0}$ with $x_1 = -1 \neq 0$ and $x_i = a_i$ for $i > 1$. ∎

Example 1 To show that the vectors $\mathbf{x} = (1, 1, 0)$, $\mathbf{y} = (1, 1, 1)$, $\mathbf{z} = (0, 1, 1)$ are linearly independent, suppose that $a\mathbf{x} + b\mathbf{y} + c\mathbf{z} = \mathbf{0}$. By taking components of this vector equation we obtain the three scalar equations

$$\begin{aligned} a + b \phantom{{}+c} &= 0, \\ a + b + c &= 0, \\ b + c &= 0. \end{aligned}$$

Subtracting the first from the second, we obtain $c = 0$. The last equation then gives $b = 0$, and finally the first one gives $a = 0$.

Example 2 The vectors $\mathbf{x} = (1, 1, 0)$, $\mathbf{y} = (1, 2, 1)$, $\mathbf{z} = (0, 1, 1)$ are linearly dependent, because $\mathbf{x} - \mathbf{y} + \mathbf{z} = \mathbf{0}$.

It is easily verified (Exercise 2.7) that any two collinear vectors, and any three coplanar vectors, are linearly dependent. This motivates the following definition of the dimension of a vector space. The vector space V has *dimension* n, $\dim V = n$, provided that V contains a set of n linearly independent vectors, while any $n+1$ vectors in V are linearly dependent; if there is no integer n for which this is true, then V is said to be *infinite-dimensional.* Thus the dimension of a finite-dimensional vector space is the largest number of linearly independent vectors which it contains; an infinite-dimensional vector space is one that contains n linearly independent vectors for every positive integer n.

Example 3 Consider the vector space $\mathscr{F}$ of all real-valued functions on $\mathscr{R}$. The functions $1, x, x^2, \ldots, x^n$ are linearly independent because a polynomial $a_0 + a_1 x + \cdots + a_n x^n$ can vanish identically only if all of its coefficients are zero. Therefore $\mathscr{F}$ is infinite-dimensional.

One certainly expects the above definition of dimension to imply that Euclidean n-space $\mathscr{R}^n$ does indeed have dimension n. We see immediately that its dimension is at least n, since it contains the n linearly independent vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$. To show that the dimension of $\mathscr{R}^n$ is precisely n, we must prove that any $n+1$ vectors in $\mathscr{R}^n$ are linearly dependent.

Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are $k > n$ vectors in $\mathscr{R}^n$, and write

$$\mathbf{v}_j = (a_{1j}, a_{2j}, \ldots, a_{nj}), \qquad j = 1, \ldots, k.$$

We want to find real numbers $x_1, \ldots, x_k$, not all zero, such that

$$\begin{aligned} \mathbf{0} &= x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_k\mathbf{v}_k \\ &= \sum_{j=1}^{k} x_j(a_{1j}, a_{2j}, \ldots, a_{nj}). \end{aligned}$$

This will be the case if $\sum_{j=1}^{k} a_{ij}x_j = 0$, $i = 1, \ldots, n$. Thus we need to find a nontrivial solution of the homogeneous linear equations

$$\begin{aligned} a_{11}x_2 + a_{12}x_2 + \cdots + a_{1k}x_k &= 0, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2k}x_k &= 0, \\ \vdots \qquad\qquad\qquad & \ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nk}x_k &= 0. \end{aligned} \tag{1}$$

By a *nontrivial* solution $(x_1, x_2, \ldots, x_k)$ of the system (1) is meant one for which not all of the x_i are zero. But $k > n$, and (1) is a system of n homogeneous linear equations in the k unknowns $x_1, \ldots, x_k$. (Homogeneous meaning that the right-hand side constants are all zero.)

It is a basic fact of linear algebra that any system of homogeneous linear equations, with more unknowns than equations, has a nontrivial solution. The proof of this fact is an application of the elementary algebraic technique of elimination of variables. Before stating and proving the general theorem, we consider a special case.

Example 4 Consider the following three equations in four unknowns:

$$\begin{aligned} x_1 + 2x_2 - x_3 + 2x_4 &= 0, \\ x_1 - x_2 + 2x_3 + x_4 &= 0, \\ 2x_1 + x_2 - x_3 - x_4 &= 0. \end{aligned} \tag{2}$$

We can eliminate x_1 from the last two equations of (2) by subtracting the first equation from the second one, and twice the first equation from the third one. This gives two equations in three unknowns:

$$\begin{aligned} -3x_2 + 3x_3 - x_4 &= 0, \\ -3x_2 + x_3 - 5x_4 &= 0. \end{aligned} \tag{3}$$

Subtraction of the first equation of (3) from the second one gives the single equation

$$-2x_3 - 4x_4 = 0 \tag{4}$$

in two unknowns. We can now choose x_4 arbitrarily. For instance, if $x_4 = 1$, then $x_3 = -2$. The first equation of (3) then gives $x_2 = -\frac{3}{7}$, and finally the first equation of (2) gives $x_1 = -\frac{22}{7}$. So we have found the nontrivial solution $(-\frac{22}{7}, -\frac{3}{7}, -2, 1)$ of the system (2).

The procedure illustrated in this example can be applied to the general case of n equations in the unknowns $x_1, \ldots, x_k$, $k > n$. First we order the n equations so that the first equation contains x_1, and then eliminate x_1 from the remaining equations by subtracting the appropriate multiple of the first equation from each of them. This gives a system of $n - 1$ homogeneous linear equations in the $k - 1$ variables $x_2, \ldots, x_k$. Similarly we eliminate x_2 from the last $n - 2$ of these $n - 1$ equations by subtracting multiples of the first one, obtaining $n - 2$ equations in the $k - 2$ variables $x_3, x_4, \ldots, x_k$. After $n - 2$ steps of this sort, we end up with a single homogeneous linear equation in the $k - n + 1$ unknowns x_n, $x_{n+1}, \ldots, x_k$. We can then choose arbitrary nontrivial values for the "extra" variables $x_{n+1}, x_{n+2}, \ldots, x_k$ (such as $x_{n+1} = 1$, $x_{n+2} = \cdots = x_k = 0$), solve the final equation for x_n, and finally proceed backward to solve successively for each of the eliminated variables $x_{n-1}, x_{n-2}, \ldots, x_1$. The reader may (if he likes) formalize this procedure to give a proof, by induction on the number n of equations, of the following result.

Theorem 2.2 If $k > n$, then any system of n homogeneous linear equations in k unknowns has a nontrivial solution.

By the discussion preceding Eqs. (1) we now have the desired result that $\dim \mathscr{R}^n = n$.

Corollary 2.3 Any $n + 1$ vectors in $\mathscr{R}^n$ are linearly dependent.

We have seen that the linearly independent vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ generate $\mathscr{R}^n$. A set of linearly independent vectors that generates the vector space V is called a *basis* for V. Since $\mathbf{x} = (x_1, x_2, \ldots, x_n) = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n$, it is clear that the basis vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$ generate V *uniquely*; that is, if $\mathbf{x} = y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \cdots + y_n\mathbf{e}_n$ also, then $x_i = y_i$ for each i. Thus each vector in $\mathscr{R}^n$ can be expressed in one and only one way as a linear combination of $\mathbf{e}_1, \ldots, \mathbf{e}_n$. Any set of n linearly independent vectors in an n-dimensional vector space has this property.

Theorem 2.4 If the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in the n-dimensional vector space V are linearly independent, then they constitute a basis for V, and furthermore generate V uniquely.

PROOF Given $\mathbf{v} \in V$, the vectors $\mathbf{v}, \mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly dependent, so by Proposition 2.1 there exist numbers $x, x_1, \ldots, x_n$, not all zero, such that

$$x\mathbf{v} + x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = \mathbf{0}.$$

If $x = 0$, then the fact that $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent implies that $x_1 = \cdots = x_n = 0$. Therefore $x \neq 0$, so we solve for $\mathbf{v}$:

$$\mathbf{v} = -\frac{x_1}{x}\mathbf{v}_1 - \frac{x_2}{x}\mathbf{v}_2 + \cdots - \frac{x_n}{x}\mathbf{v}_n.$$

Thus the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ generate V, and therefore constitute a basis for V. To show that they generate V uniquely, suppose that

$$a_1\mathbf{v}_1 + \cdots + a_n\mathbf{v}_n = a_1'\mathbf{v}_1 + \cdots + a_n'\mathbf{v}_n.$$

Then

$$(a_1 - a_1')\mathbf{v}_1 + \cdots + (a_n - a_n')\mathbf{v}_n = \mathbf{0}.$$

So, since $\mathbf{v}_1, \ldots, \mathbf{v}_n$ are linearly independent, it follows that $a_i - a_i' = 0$, or $a_i = a_i'$, for each i. ∎

There remains the possibility that $\mathscr{R}^n$ has a basis which contains fewer than n elements. But the following theorem shows that this cannot happen.

Theorem 2.5 If $\dim V = n$, then each basis for V consists of exactly n vectors.

PROOF Let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n$ be n linearly independent vectors in V. If there were a basis $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m$ for V with $m < n$, then there would exist numbers $\{a_{ij}\}$ such that

$$\begin{aligned}\mathbf{w}_1 &= a_{11}\mathbf{v}_1 + \cdots + a_{m1}\mathbf{v}_m,\\ &\vdots\\ \mathbf{w}_n &= a_{1n}\mathbf{v}_1 + \cdots + a_{mn}\mathbf{v}_m.\end{aligned}$$

Since $m < n$, Theorem 2.2 supplies numbers $x_1, \ldots, x_n$ *not all zero*, such that

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0, \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0. \end{aligned}$$

But this implies that

$$\begin{aligned} x_1\mathbf{w}_1 + \cdots + x_n\mathbf{w}_n &= \sum_{j=1}^{n} x_j(a_{1j}\mathbf{v}_1 + \cdots + a_{mj}\mathbf{v}_m) \\ &= \sum_{i=1}^{m} (a_{i1}x_1 + \cdots + a_{in}x_n)\mathbf{v}_i \\ &= \mathbf{0}, \end{aligned}$$

which contradicts the fact that $\mathbf{w}_1, \ldots, \mathbf{w}_n$ are linearly independent. Consequently no basis for V can have $m < n$ elements. ∎

We can now completely describe the general situation as regards subspaces of $\mathscr{R}^n$. If V is a subspace of $\mathscr{R}^n$, then $k = \dim V \leqq n$ by Corollary 2.3, and if $k = n$, then $V = \mathscr{R}^n$ by Theorem 2.4. If $k > 0$, then any k linearly independent vectors in V generate V, and no basis for V contains fewer than k vectors (Theorem 2.5).

Exercises

2.1 Why is it true that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ are linearly dependent if any one of them is zero? If any subset of them is linearly dependent?

2.2 Which of the following sets of vectors are bases for the appropriate space $\mathscr{R}^n$?
(a) (1, 0) and (1, 1).
(b) (1, 0, 0), (1, 1, 0), and (0, 0, 1).
(c) (1, 1, 1), (1, 1, 0), and (1, 0, 0).
(d) (1, 1, 1, 0), (1, 0, 0, 0), (0, 1, 0, 0), and (0, 0, 1, 0).
(e) (1, 1, 1, 1), (1, 1, 1, 0), (1, 1, 0, 0), and (1, 0, 0, 0).

2.3 Find the dimension of the subspace V of $\mathscr{R}^4$ that is generated by the vectors (0, 1, 0, 1), (1, 0, 1, 0), and (1, 1, 1, 1).

2.4 Show that the vectors (1, 0, 0, 1), (0, 1, 0, 1), (0, 0, 1, 1) form a basis for the subspace V of $\mathscr{R}^4$ which is defined by the equation $x_1 + x_2 + x_3 - x_4 = 0$.

2.5 Show that any set $\mathbf{v}_1, \ldots, \mathbf{v}_k$, of linearly independent vectors in a vector space V can be extended to a basis for V. That is, if $k < n = \dim V$, then there exist vectors $\mathbf{v}_{k+1}, \ldots, \mathbf{v}_n$ in V such that $\mathbf{v}_1, \ldots \mathbf{v}_n$ is a basis for V.

2.6 Show that Theorem 2.5 is equivalent to the following theorem: Suppose that the equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0, \\ &\vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n &= 0 \end{aligned}$$

have only the trivial solution $x_1 = \cdots = x_n = 0$. Then, for each $\mathbf{b} = (b_1, \ldots, b_n)$, the equations

$$a_{11}x_1 + \cdots + a_{1n}x_n = b_1,$$
$$\vdots$$
$$a_{n1}x_1 + \cdots + a_{nn}x_n = b_n$$

have a *unique* solution. *Hint:* Consider the vectors $\mathbf{a}_j = (a_{1j}, a_{2j}, \ldots, a_{nj}), j = 1, \ldots, n$.

2.7 Verify that any two collinear vectors, and any three coplanar vectors, are linearly dependent.

3 INNER PRODUCTS AND ORTHOGONALITY

In order to obtain the full geometric structure of $\mathscr{R}^n$ (including the concepts of distance, angles, and orthogonality), we must supply $\mathscr{R}^n$ with an inner product. An *inner* (scalar) *product* on the vector space V is a function $V \times V \to \mathscr{R}$, which associates with each pair $(\mathbf{x}, \mathbf{y})$ of vectors in V a real number $\langle \mathbf{x}, \mathbf{y} \rangle$, and satisfies the following three conditions:

SP1 $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ if $\mathbf{x} \neq \mathbf{0}$ (positivity).
SP2 $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$ (symmetry).
SP3 $\langle a\mathbf{x} + b\mathbf{y}, \mathbf{z} \rangle + a\langle \mathbf{x}, \mathbf{z} \rangle + b\langle \mathbf{y}, \mathbf{z} \rangle$.

The third of these conditions is linearity in the first variable; symmetry then gives linearity in the second variable also. Thus an inner product on V is simply a positive, symmetric, bilinear function on $V \times V$. Note that SP3 implies that $\langle \mathbf{0}, \mathbf{0} \rangle = 0$ (see Exercise 3.1).

The *usual inner product* on $\mathscr{R}^n$ is denoted by $\mathbf{x} \cdot \mathbf{y}$ and is defined by

$$\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n, \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_n)$, $\mathbf{y} = (y_1, \ldots, y_n)$. It should be clear that this definition satisfies conditions SP1, SP2, SP3 above. There are many inner products on $\mathscr{R}^n$ (see Example 2 below), but we shall use only the usual one.

Example 1 Denote by $\mathscr{C}[a, b]$ the vector space of all continuous functions on the interval $[a, b]$, and define

$$\langle f, g \rangle = \int_a^b f(t)g(t)\, dt$$

for any pair of functions $f, g \in \mathscr{C}[a, b]$. It is obvious that this definition satisfies conditions SP2 and SP3. It also satisfies SP1, because if $f(t_0) \neq 0$, then by continuity $(f(t))^2 > 0$ for all t in some neighborhood of t_0, so

$$\langle f, f \rangle = \int_a^b f(t)^2\, dt > 0.$$

Therefore we have an inner product on $\mathscr{C}[a, b]$.

Example 2 Let a, b, c be real numbers with $a > 0$, $ac - b^2 > 0$, so that the quadratic form $q(\mathbf{x}) = ax_1^2 + 2bx_1x_2 + cx_2^2$ is positive-definite (see Section II.4). Then $\langle \mathbf{x}, \mathbf{y} \rangle = ax_1y_1 + bx_1y_2 + bx_2y_1 + cx_2y_2$ defines an inner product on $\mathscr{R}^2$ (why?). With $a = c = 1$, $b = 0$ we obtain the usual inner product on $\mathscr{R}^2$.

An inner product on the vector space V yields a notion of the length or "size" of a vector $\mathbf{x} \in V$, called its *norm* $|\mathbf{x}|$. In general, a *norm* on the vector space V is a real-valued function $\mathbf{x} \to |\mathbf{x}|$ on V satisfying the following conditions:

N1	$\lvert\mathbf{x}\rvert > 0$ if $\mathbf{x} \neq 0$	(positivity),
N2	$\lvert a\mathbf{x}\rvert = \lvert a\rvert\,\lvert\mathbf{x}\rvert$	(homogeneity),
N3	$\lvert\mathbf{x} + \mathbf{y}\rvert \leqq \lvert\mathbf{x}\rvert + \lvert\mathbf{y}\rvert$	(triangle inequality),

for all $\mathbf{x}, \mathbf{y} \in V$ and $a \in \mathscr{R}$. Note that N2 implies that $|\mathbf{0}| = 0$.

The norm associated with the inner product $\langle\ ,\ \rangle$ on V is defined by

$$|\mathbf{x}| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} \tag{2}$$

It is clear that SP1–SP3 and this definition imply conditions N1 and N2, but the triangle inequality is not so obvious; it will be verified below.

The most commonly used norm on $\mathscr{R}^n$ is the *Euclidean norm*

$$|\mathbf{x}| = (x_1^2 + \cdots + x_n^2)^{1/2},$$

which comes in the above way from the usual inner product on $\mathscr{R}^n$. Other norms on $\mathscr{R}^n$, not necessarily associated with inner products, are occasionally employed, but henceforth $|\mathbf{x}|$ will denote the Euclidean norm unless otherwise specified.

Example 3 $\|\mathbf{x}\| = \max\{|x_1|, \ldots, |x_n|\}$, the maximum of the absolute values of the coordinates of $\mathbf{x}$, defines a norm on $\mathscr{R}^n$ (see Exercise 3.2).

Example 4 $|\mathbf{x}|_1 = |x_1| + |x_2| + \cdots + |x_n|$ defines still another norm on $\mathscr{R}^n$ (again see Exercise 3.2).

A norm on V provides a definition of the *distance* $d(\mathbf{x}, \mathbf{y})$ between any two points $\mathbf{x}$ and $\mathbf{y}$ of V:

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|.$$

Note that a distance function d defined in this way satisfies the following three conditions:

D1	$d(\mathbf{x}, \mathbf{y}) > 0$ unless $\mathbf{x} = \mathbf{y}$	(positivity),
D2	$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$	(symmetry),
D3	$d(\mathbf{x}, \mathbf{z}) \leqq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$	(triangle inequality),