Linear Algebra

Ideas and Applications

FOURTH EDITION

RICHARD C. PENNEY

with website

WILEY

LINEAR ALGEBRA

LINEAR ALGEBRA
Ideas and Applications

Fourth Edition

RICHARD C. PENNEY
Purdue University

WILEY

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data:

Penney, Richard C.
Linear algebra : ideas and applications / Richard Penney. – Fourth edition.
pages cm
Includes index.
ISBN 978-1-118-90958-4 (cloth)
1. Algebras, Linear–Textbooks. I. Title.
QA184.2.P46 2015
512′.5–dc23
2015016650

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

CONTENTS

PREFACE

I wrote this book because I have a deep conviction that mathematics is about ideas, not just formulas and algorithms, and not just theorems and proofs. The text covers the material usually found in a one or two semester linear algebra class. It is written, however, from the point of view that knowing *why* is just as important as knowing *how*.

To ensure that the readers see not only why a given fact is true, but also why it is important, I have included a number of the beautiful applications of linear algebra.

Most of my students seem to like this emphasis. For many, mathematics has always been a body of facts to be blindly accepted and used. The notion that they personally can decide mathematical truth or falsehood comes as a revelation. Promoting this level of understanding is the goal of this text.

RICHARD C. PENNEY

West Lafayette, Indiana
Updated with October, 2015

FEATURES OF THE TEXT

Parallel Structure Most linear algebra texts begin with a long, basically computational, unit devoted to solving systems of equations and to matrix algebra and determinants. Students find this fairly easy and even somewhat familiar. But, after a third or more of the class has gone by peacefully, the boom falls. Suddenly, the students are asked to absorb abstract concept after abstract concept, one following on the heels of the other. They see little relationship between these concepts and the first part of the course or, for that matter, anything else they have ever studied. By the time the abstractions can be related to the first part of the course, many students are so lost that they neither see nor appreciate the connection.

This text is different. We have adopted a parallel mode of development in which the abstract concepts are introduced right from the beginning, along with the computational. Each abstraction is used to shed light on the computations. In this way, the students see the abstract part of the text as a natural outgrowth of the computational part. This is not the "mention it early but use it late" approach adopted by some texts. Once a concept such as linear independence or spanning is introduced, it becomes part of the vocabulary to be used frequently and repeatedly throughout the rest of the text.

The advantages of this kind of approach are immense. The parallel development allows us to introduce the abstractions at a slower pace, giving students a whole semester to absorb what was formerly compressed into two-thirds of a semester. Students have time to fully absorb each new concept before taking on another. Since the concepts are utilized as they are introduced, the students see *why* each concept is necessary. The relation between theory and application is clear and immediate.

Gradual Development of Vector Spaces One special feature of this text is its treatment of the concept of vector space. Most modern texts tend to introduce this concept fairly late. We introduce it early because we need it early. Initially, however, we do not develop it in any depth. Rather, we slowly expand the reader's understanding by introducing new ideas as they are needed.

This approach has worked extremely well for us. When we used more traditional texts, we found ourselves spending endless amounts of time trying to explain what a vector space is. Students felt bewildered and confused, not seeing any point to what they were learning. With the gradual approach, on the other hand, the question of what a vector space is hardly arises. *With this approach, the vector space concept seems to cause little difficulty for the students.*

Treatment of Proofs It is essential that students learn to read and produce proofs. Proofs serve both to validate the results and to explain why they are true. For many students, however, linear algebra is their first proof-based course. They come to the subject with neither the ability to read proofs nor an appreciation for their importance.

Many introductory linear algebra texts adopt a formal "definition-theorem-proof" format. In such a treatment, a student who has not yet developed the ability to read abstract mathematics can perceive both the statements of the theorems and their proofs (not to mention the definitions) as meaningless abstractions. They wind up reading only the examples in the hope of finding "patterns" that they can imitate to complete the assignments. In the end, such students wind up only mastering the computational techniques, since this is the only part of the course that has any meaning for them. In essence, we have taught them to be nothing more than slow, inaccurate computers.

Our point of view is different. *This text is meant to be read by the student – all of it!* We always work from the concrete to the abstract, never the opposite. We also make full use of geometric reasoning, where appropriate. We try to explain "analytically, algebraically, and geometrically." We use carefully chosen examples to motivate both the definitions and theorems. Often, the essence of the proof is already contained in the example. Despite this, we give complete and rigorous *student-readable* proofs of most results.

Conceptual Exercises Most texts at this level have exercises of two types: proofs and computations. We certainly do have a number of proofs and we definitely have lots of computations. The vast majority of the exercises are, however, "conceptual, but not theoretical." That is, *each exercise asks an explicit, concrete question which requires the student to think conceptually in order to provide an answer.* Such questions are both more concrete and more manageable than proofs and thus are much better at demonstrating the concepts. They do not require that the student already have facility with abstractions. Rather, they act as a bridge between the abstract proofs and the explicit computations.

Applications Sections Doable as Self-Study Applications can add depth and meaning to the study of linear algebra. Unfortunately, just covering the "essential" topics in the typical first course in linear algebra leaves little time for additional material, such as applications.

Many of our sections are followed by one or more application sections that use the material just studied. This material is designed to be read unaided by the student and thus may be assigned as outside reading. As an aid to this, we have provided two levels of exercises: self-study questions and exercises. The self-study questions are designed to be answerable with a minimal investment of time by anyone who has carefully read and digested the relevant material. The exercises require more thought and a greater depth of understanding. They would typically be used in parallel with classroom discussions.

We feel that, in general, there is great value in providing material that the students are responsible for learning on their own. Learning to read mathematics is the first step in learning to do mathematics. Furthermore, there is no way that we can ever teach everything the students need to know; we cannot even predict what they need to know. Ultimately, the most valuable skill we teach is the ability to teach oneself. The applications form a perfect vehicle for this in that an imperfect mastery of any given application will not impede the student's understanding of linear algebra.

Early Orthogonality Option We have designed the text so that the chapter on orthogonality, with the exception of the last three sections, may be done immediately following Chapter 3 rather than after the section on eigenvalues.

True-False Questions We have included true-false questions for most sections.

Chapter Summaries At the end of each chapter there is a chapter summary that brings together major points from the chapter so students can get an overview of what they just learned.

Student Tested This text has been used over a period of years by numerous instructors at both Purdue University and other universities nationwide. We have incorporated comments from instructors, reviewers, and (most important) students.

Technology Most sections of the text include a selection of computer exercises under the heading Computer Projects. Each exercise is specific to its section and is designed to support and extend the concepts discussed in that section.

These exercises have a special feature: they are designed to be "freestanding." In principle, the instructor should not need to spend any class time at all discussing computing. Everything most students need to know is right there. In the text, the discussion is based on MATLAB®.

Meets LACSG Recommendations The Linear Algebra Curriculum Study Group (LACSG) recommended that the first class in linear algebra be a "student-oriented" class that considers the "client disciplines" and that makes use of technology. The above comments make it clear that this text meets these recommendations. The LACSG also recommended that the first class be "matrix-oriented." We emphasize matrices throughout.

ACKNOWLEDGMENTS

The author would like to thank all of students, teachers, and reviewers of the text who have made comments, many of which have had significant impact on the final form of the fourth edition. In particular, he would like to thank Bill Dunbar and his students, Ken Rubenstein, and Steven Givant for their particularly useful input.

ABOUT THE COMPANION WEBSITE

This book is accompanied by a companion website:

http://www.wiley.com/go/penney/linearalgebra

The website includes:

- Instructors' Solutions Manual
- Figures

CHAPTER 1

SYSTEMS OF LINEAR EQUATIONS

1.1 THE VECTOR SPACE OF $m \times n$ MATRICES

It is difficult to go through life without seeing matrices. For example, the 2014 annual report of Acme Squidget might contain the Table 1.1, which shows how much profit (in millions of dollars) each branch made from the sale of each of the company's three varieties of squidgets in 2014.

TABLE 1.1 Profits: 2014

	Red	*Blue*	*Green*	*Total*
Kokomo	11.4	5.7	6.3	23.4
Philly	9.1	6.7	5.5	21.3
Oakland	14.3	6.2	5.0	25.5
Atlanta	10.0	7.1	5.7	22.8
Total	44.8	25.7	22.5	93.0

If we were to enter this data into a computer, we might enter it as a rectangular array without labels. Such an array is called a **matrix**. The Acme profits for 2014 would be described by the following matrix. This matrix is a 5×4 matrix (read "five by four") in that it has five rows and four columns. We would also say that its "size" is 5×4. In general, a matrix has **size** $m \times n$ if it has m rows and n columns.

Linear Algebra: Ideas and Applications, Fourth Edition. Richard C. Penney.

Companion website: www.wiley.com/go/penney/linearalgebra

Definition 1.1 *The set of all $m \times n$ matrices is denoted $M(m, n)$.*

$$P = \begin{bmatrix} 11.4 & 5.7 & 6.3 & 23.4 \\ 9.1 & 6.7 & 5.5 & 21.3 \\ 14.3 & 6.2 & 5.0 & 25.5 \\ 10.0 & 7.1 & 5.7 & 22.8 \\ 44.8 & 25.7 & 22.5 & 93.0 \end{bmatrix}$$

Each row of an $m \times n$ matrix may be thought of as a $1 \times n$ matrix. The rows are numbered from top to bottom. Thus, the second row of the Acme profit matrix is the 1×4 matrix

$$[9.1, 6.7, 5.5, 21.3]$$

This matrix would be called the "profit vector" for the Philly branch. (In general, any matrix with only one row is called a **row vector**. For the sake of legibility, we usually separate the entries in row vectors by commas, as above.)

Similarly, a matrix with only one column is called a **column vector**. The columns are numbered from left to right. Thus, the third column of the Acme profit matrix is the column vector

$$\begin{bmatrix} 6.3 \\ 5.5 \\ 5.0 \\ 5.7 \\ 22.5 \end{bmatrix}$$

This matrix is the "green squidget profit vector."

If $A_1, A_2, \dots, A_n$ is a sequence of $m \times 1$ column vectors, then the $m \times n$ matrix A that has the A_i as columns is denoted

$$A = [A_1, A_2, \dots, A_n]$$

Similarly, if $B_1, B_2, \dots, B_m$ is a sequence of $1 \times n$ row vectors, then the $m \times n$ matrix B that has the B_i as rows is denoted

$$B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix}$$

In general, if a matrix is denoted by an uppercase letter, such as A, then the entry in the ith row and jth column may be denoted by either A_{ij} or a_{ij}, using the corresponding lowercase letter. We shall refer to a_{ij} as the "(i,j) entry of A." For example, for the matrix P above, the (2, 3) entry is $p_{23} = 5.5$. Note that the row number comes first. Thus, the most general 2×3 matrix is

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

We will also occasionally write "$A = [a_{ij}]$," meaning that "A is the matrix whose (i,j) entry is a_{ij}."

At times, we want to take data from two tables, manipulate it in some manner, and display it in a third table. For example, suppose that we want to study the performance of each division of Acme Squidget over the two-year period, 2013–2014. We go back to the 2013 annual report, finding the 2013 profit matrix to be

$$Q = \begin{bmatrix} 11.0 & 5.5 & 6.1 & 22.6 \\ 9.0 & 6.3 & 5.3 & 20.6 \\ 14.1 & 5.9 & 4.9 & 24.9 \\ 9.7 & 7.0 & 5.8 & 22.5 \\ 43.8 & 24.7 & 22.1 & 90.6 \end{bmatrix}$$

If we want the totals for the two-year period, we simply add the entries of this matrix to the corresponding entries from the 2014 profit matrix. Thus, for example, over the two-year period, the Kokomo division made $5.5 + 5.7 = 11.2$ million dollars from selling blue squidgets. Totaling each pair of entries, we find the two-year profit matrix to be

$$T = \begin{bmatrix} 22.4 & 11.2 & 12.4 & 46.0 \\ 18.1 & 13.0 & 10.8 & 41.9 \\ 28.4 & 12.1 & 9.9 & 50.4 \\ 19.7 & 14.1 & 11.5 & 45.3 \\ 88.6 & 50.4 & 44.6 & 183.6 \end{bmatrix}$$

In matrix notation, we indicate that T was obtained by summing corresponding entries of Q and P by writing

$$T = Q + P$$

In general, if A and B are $m \times n$ matrices, then $A + B$ is the $m \times n$ matrix defined by the formula

$$A + B = [a_{ij}] + [b_{ij}] = [a_{ij} + b_{ij}]$$

For example

$$\begin{bmatrix} 2 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 2 & 1 \\ 0 & 5 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 5 & 1 \\ 1 & 5 & 0 \end{bmatrix}$$

Addition of matrices of different sizes is not defined.

What if, instead of totals for each division and each product, we wanted two-year *averages*? We would simply multiply each entry of $T = P + Q$ by $\frac{1}{2}$. The notation for this is "$\frac{1}{2}T$." Specifically,

$$\frac{1}{2}T = \begin{bmatrix} 11.20 & 05.60 & 06.20 & 23.00 \\ 09.05 & 06.50 & 05.40 & 20.95 \\ 14.20 & 06.05 & 04.95 & 25.20 \\ 09.85 & 07.05 & 05.75 & 22.65 \\ 44.30 & 25.20 & 22.30 & 91.80 \end{bmatrix}$$

In general, if c is a number and $A = [a_{ij}]$ is an $m \times n$ matrix, we define

$$cA = c[a_{ij}] = [ca_{ij}] = [a_{ij}]c = Ac \tag{1.1}$$

Hence,

$$2\begin{bmatrix} 2 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 4 & 6 & 0 \\ 2 & 0 & -2 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 0 \\ 1 & 0 & -1 \end{bmatrix}2$$

There is also a notion of subtraction of $m \times n$ matrices. In general, if A and B are $m \times n$ matrices, then we define $A - B$ to be the $m \times n$ matrix defined by the formula

$$A - B = [a_{ij}] - [b_{ij}] = [a_{ij} - b_{ij}]$$

Thus,

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

In linear algebra, the terms **scalar** and **number** mean essentially the same thing. Thus, multiplying a matrix by a real number is often called **scalar multiplication.**

The Space $\mathbb{R}^n$

We may think of a 2×1 column vector $X = \begin{bmatrix} x \\ y \end{bmatrix}$ as representing the point in the plane with coordinates (x, y) as in Figure 1.1. We may also think of X as representing

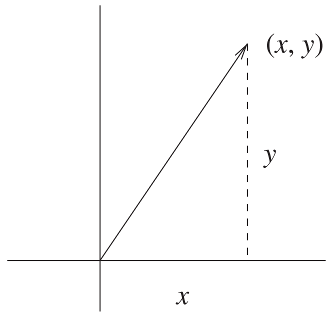

FIGURE 1.1 Coordinates in $\mathbb{R}^2$.

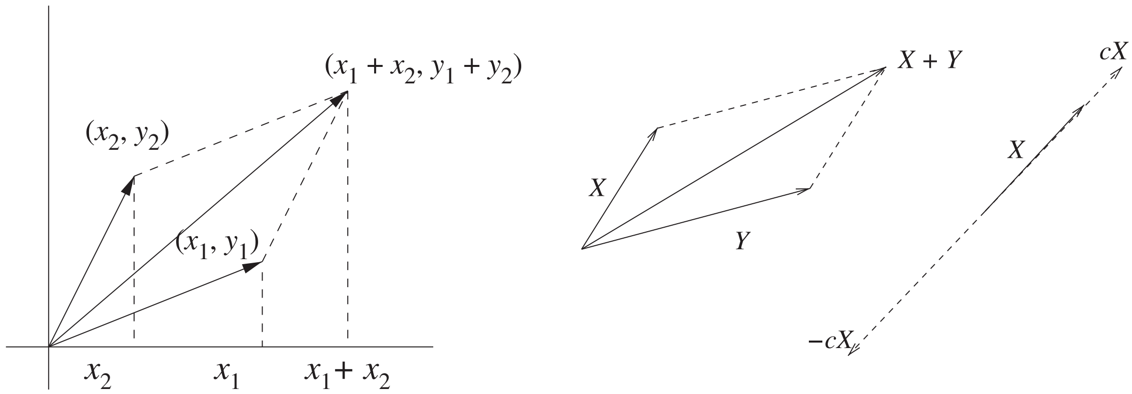

FIGURE 1.2 Vector algebra.

the vector from the point $(0, 0)$ to (x, y)—that is, as an arrow drawn from $(0, 0)$ to (x, y). We will usually denote the set of 2×1 matrices by $\mathbb{R}^2$ when thought of as points in two-dimensional space.

Like matrices, we can add pairs of vectors and multiply vectors by scalars. Specifically, if X and Y are vectors with the same initial point, then $X + Y$ is the diagonal of the parallelogram with sides X and Y beginning at the same initial point (Figure 1.2, right). For a positive scalar c, cX is the vector with the same direction as that of X, but with magnitude expanded (or contracted) by a factor of c.

Figure 1.2 on the left shows that when two elements of $\mathbb{R}^2$ are added, the corresponding vectors add as well. Similarly, multiplication of an element of $\mathbb{R}^2$ by a scalar corresponds to multiplication of the corresponding vector by the same scalar. If $c < 0$, the direction of the vector is reversed and the vector is then expanded or contracted by a factor of $-c$ (Figure 1.2, right).

EXAMPLE 1.1

Compute the sum of the vectors represented by $\begin{bmatrix} -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$ and draw a diagram illustrating your computation.

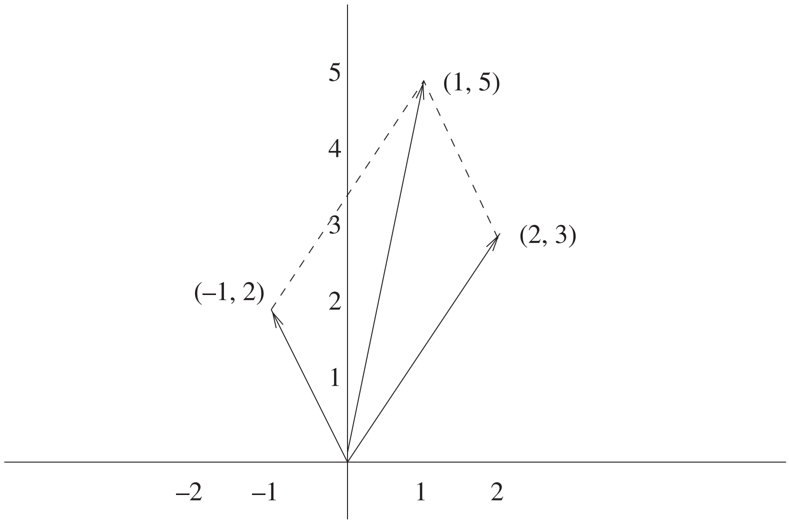

FIGURE 1.3 Example 1.1.

Solution. The sum is computed as follows:

$$\begin{bmatrix} -1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -1+2 \\ 2+3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$$

The vectors (along with their sum) are plotted in Figure 1.3.

Similarly, we may think of the 3×1 matrix

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

as representing either the point (x, y, z) in three-dimensional space or the vector from $(0, 0, 0)$ to (x, y, z) as in Figure 1.4. Matrix addition and scalar multiplication are describable as vector addition just as in two dimensions. We will usually denote the set of 3×1 matrices by $\mathbb{R}^3$ when thought of as points in three-dimensional space.

What about $n \times 1$ matrices? Even though we cannot visualize n dimensions, we still envision $n \times 1$ matrices as somehow representing points in n dimensional space. The set of $n \times 1$ matrices will be denoted as $\mathbb{R}^n$ when thought of in this way.

Definition 1.2 $\mathbb{R}^n$ *is the set of all* $n \times 1$ *matrices.*

Linear Combinations and Linear Dependence

We can use our Acme Squidget profit matrices to demonstrate one of the most important concepts in linear algebra. Consider the last column of the 2014 profit

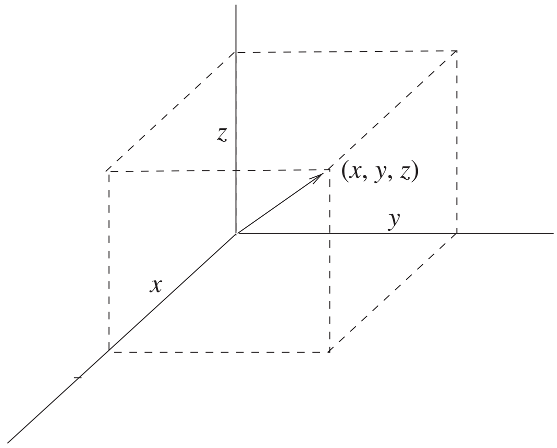

FIGURE 1.4 Coordinates in $\mathbb{R}^3$.

matrix. Since this column represents the total profit for each branch, it is just the sum of the other columns in the profit matrix:

$$\begin{bmatrix} 11.4 \\ 9.1 \\ 14.3 \\ 10.0 \\ 44.8 \end{bmatrix} + \begin{bmatrix} 5.7 \\ 6.7 \\ 6.2 \\ 7.1 \\ 25.7 \end{bmatrix} + \begin{bmatrix} 6.3 \\ 5.5 \\ 5.0 \\ 5.7 \\ 22.5 \end{bmatrix} = \begin{bmatrix} 23.4 \\ 21.3 \\ 25.5 \\ 22.8 \\ 93.0 \end{bmatrix} \tag{1.2}$$

This last column does not tell us anything we did not already know in that we could have computed the sums ourselves. Thus, while it is useful to have the data explicitly displayed, it is not essential. We say that this data is "dependent on" the data in the other columns. Similarly, the last row of the profit matrix is dependent on the other rows in that it is just their sum.

For another example of dependence, consider the two profit matrices Q and P and their average

$$A = \tfrac{1}{2}(Q + P) = \tfrac{1}{2}Q + \tfrac{1}{2}P \tag{1.3}$$

The matrix A depends on P and Q—once we know P and Q, we can compute A.

These examples exhibit an especially simple form of dependence. In each case, the matrix we chose to consider as dependent was produced by multiplying the other matrices by scalars and adding. This leads to the following concept.

Definition 1.3 *Let $S = \{A_1, A_2, \ldots, A_k\}$ be a set of elements of $M(m, n)$. An element C of $M(m, n)$ is* ***linearly dependent on*** *S if there are scalars b_i such that*

$$C = b_1A_1 + b_2A_2 + \cdots + b_kA_k \tag{1.4}$$

We also say that "C is a ***linear combination*** *of the A_i."*

Remark. In set theory, an object that belongs to a certain set is called an **element** of that set. The student must be careful not to confuse the terms "element" and "entry." The matrix below is *one element* of the set of 2×2 matrices. Every element of the set of 2×2 matrices has four *entries*.

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

The expression "$a \in B$" means that a is an element of the set B.

One particular element of $M(m, n)$ is linearly dependent on every other element of $M(m, n)$. This is the $m \times n$ matrix, which has all its entries equal to 0. We denote this matrix by 0. It is referred to as "the zero element of $M(m, n)$." Thus, the zero element of $M(2, 3)$ is

$$0 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The $m \times n$ zero matrix depends on every other $m \times n$ matrix because, for any $m \times n$ matrix A,

$$0A = 0$$

We can also discuss linearly dependent sets of matrices.

Definition 1.4 *Let $S = \{A_1, A_2, \dots, A_k\}$ be a set of elements of $M(m, n)$. Then S is **linearly dependent** if at least one of the A_j is a linear combination of the other elements of S—that is, A_j is a linear combination of the set of elements A_i with $i \neq j$. We also define the set $\{0\}$, where 0 is the zero element of $M(m, n)$, to be linearly dependent. S is said to be **linearly independent** if it is not linearly dependent. Hence, S is linearly independent if none of the A_i are linear combinations of other elements of S.*

Thus, from formula (1.3), the set $S = \{P, Q, A\}$ is linearly dependent. In addition, from formula (1.2), the set of columns of P is linearly dependent. The set of rows of P is also linearly dependent since the last row is the sum of the other rows.

EXAMPLE 1.2

Is $S = \{A_1, A_2, A_3, A_4\}$ a linearly independent set where the A_i are the following 2×2 matrices?

$$A_1 = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 5 \\ 0 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \quad A_4 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Solution. By inspection

$$A_2 = 2A_3 + A_4 = 0A_1 + 2A_3 + A_4$$

showing that S is linearly dependent.

Remark. Note that A_1 is *not* a combination of the other A_i since the (2, 1) entry of A_1 is nonzero, while all the other A_i are zero in this position. This demonstrates that linear dependence does not require that each of the A_i be a combination of the others.

EXAMPLE 1.3

Let B_1, B_2, and B_3 be as shown. Is $S = \{B_1, B_2, B_3\}$ a linearly dependent set?

$$B_1 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -7 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad B_3 = \begin{bmatrix} 1.3 \\ 0 \\ 2.2 \\ 0 \end{bmatrix}$$

Solution. We begin by asking ourselves whether B_1 is linearly dependent on B_2 and B_3—that is, are there scalars a and b such that

$$B_1 = aB_2 + bB_3$$

The answer is no since the last entries of both B_2 and B_3 are 0, while the last entry of B_1 is 1.

Similarly, we see that B_2 is not a linear combination of B_1 and B_3 (from the second entries) and B_3 is not a linear combination of B_1 and B_2 (from the third entries). Thus, the given three matrices form a linearly independent set.

Example 1.3 is an example of the following general principle that we use several times later in the text.

Proposition 1.1 *Suppose that $S = \{A_1, A_2, \ldots, A_p\}$ is a set of $m \times n$ matrices such that each A_k has a nonzero entry in a position where all the other A_q are zero—that is, for each k there is a pair of indices (i, j) such that $(A_k)_{ij} \neq 0$ while $(A_q)_{ij} = 0$ for all $q \neq k$. Then S is linearly independent.*

Proof. Suppose that S is linearly dependent. Then there is a k such that

$$A_k = c_1A_1 + c_2A_2 + \cdots + c_{k-1}A_{k-1} + c_{k+1}A_{k+1} + \cdots + c_qA_q$$

Let (i, j) be as described in the statement of the proposition. Equating the (i, j) entries on both sides of the above equation shows that

$$(A_k)_{ij} = c_1 0 + c_2 0 + \cdots + c_{k-1} 0 + c_{k+1} 0 + \cdots + c_q 0 = 0$$

contradicting the hypothesis that $(A_k)_{ij} \neq 0$. □

Linear independence also has geometric significance. Two vectors X and Y in $\mathbb{R}^2$ will be linearly independent if and only if neither is a scalar multiple of the other—that is, they are noncollinear (Figure 1.5, left). We will prove in Section 2.2 that any three vectors X, Y, and Z in $\mathbb{R}^2$ are linearly dependent (Figure 1.5, right).

In $\mathbb{R}^3$, the set of linear combinations of a pair of linearly independent vectors lies in the plane they determine. Thus, three noncoplanar vectors will be linearly independent (Figure 1.6).

In general, the set of all matrices that depends on a given set of matrices is called the span of the set:

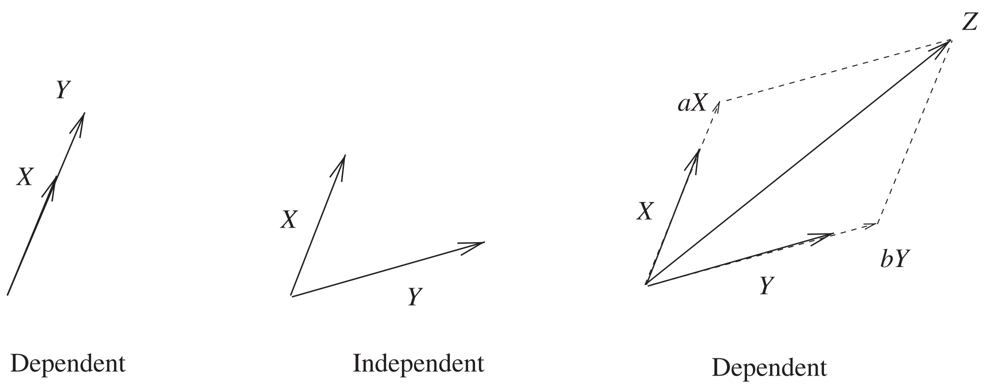

FIGURE 1.5 Dependence in $\mathbb{R}^2$.

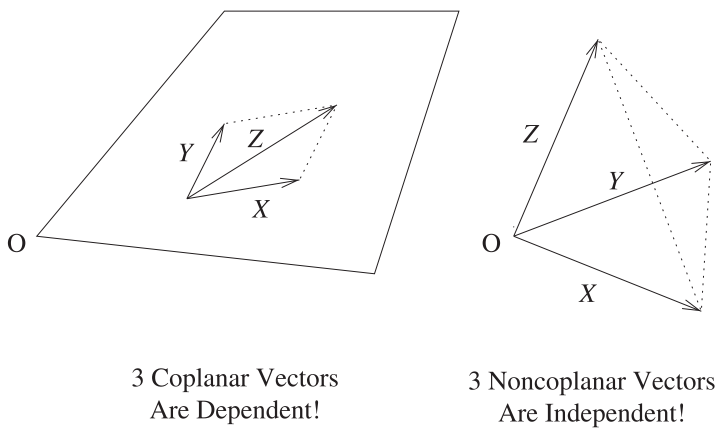

FIGURE 1.6 Three vectors in $\mathbb{R}^3$.

Definition 1.5 *Let $S = \{A_1, A_2, \dots, A_k\}$ be a set of elements of $M(m, n)$. Then* span S (*the* ***span*** *of S*) *is the set of all elements of the form*

$$B = c_1A_1 + c_2A_2 + \cdots + c_kA_k$$

where the c_i are scalars.

The span of S, then, is just the set of all linear combinations of the elements of S. Thus, for example, if B_1, B_2, and B_3 are as in Example 1.3 on page 9, then

$$B_1 + B_2 + 10B_3 = \begin{bmatrix} 8 \\ 2 \\ 22 \\ 1 \end{bmatrix}$$

is one element of span $\{B_1, B_2, B_3\}$.

In $\mathbb{R}^2$ and $\mathbb{R}^3$, the span of a single vector is the line through the origin determined by it. From Figure 1.6, the span of a set of two linearly independent vectors will be the plane they determine.

What Is a Vector Space?

One of the advantages of matrix notation is that it allows us to treat a matrix as if it were one single number. For example, we may solve for Q in formula (1.3) on page 7:

$$\begin{aligned} A &= \tfrac{1}{2}(P + Q) \\ 2A &= Q + P \\ 2A - P &= Q \end{aligned}$$

The preceding calculations used a large number of properties of matrix arithmetic that we have not discussed. In greater detail, our argument was as follows:

$$\begin{aligned} A &= \tfrac{1}{2}(P + Q) \\ 2A &= 2[\tfrac{1}{2}(P + Q)] = \tfrac{2}{2}(Q + P) = Q + P \\ 2A + (-P) &= (Q + P) + (-P) \\ &= Q + (P - P) = Q + 0 = Q \end{aligned}$$

We certainly used the associative law $(A + B) + C = A + (B + C)$, the laws $A + (-A) = 0$ and $A + 0 = A$, as well as several other laws. In Theorem 1.1, we list the most important algebraic properties of matrix addition and scalar multiplication.

These properties are called the **vector space properties**. Experience has proved that these properties are all that one needs to effectively deal with any computations such as those just done with A, P, and Q. For the sake of this list, we let $\mathcal{V} = M(m, n)$ for some fixed m and n.[1] Thus, for example, $\mathcal{V}$ might be the set of all 2×3 matrices.

Theorem 1.1 (The Vector Space Properties). *Let $X, Y,$ and Z be elements of $\mathcal{V}$. Then:*

(a) $X + Y$ *is a well-defined element of* $\mathcal{V}$.
(b) $X + Y = Y + X$ *(commutativity).*
(c) $X + (Y + Z) = (X + Y) + Z$ *(associativity).*
(d) *There is an element denoted* 0 *in* $\mathcal{V}$ *such that* $X + 0 = X$ *for all* $X \in \mathcal{V}$. *This element is referred to as the "zero element."*
(e) *For each* $X \in \mathcal{V}$, *there is an element* $-X \in \mathcal{V}$ *such that* $X + (-X) = 0$.

Additionally, for all scalars k and l:

(f) kX *is a well-defined element of* $\mathcal{V}$.
(g) $k(lX) = (kl)X$.
(h) $k(X + Y) = kX + kY$.
(i) $(k + l)X = kX + lX$.
(j) $1X = X$.

The proofs that the properties from this list hold for $\mathcal{V} = M(m, n)$ are left as exercises for the reader. However, let us prove property (c) as an example of how such a proof should be written.

EXAMPLE 1.4

Prove property (c) for $M(m, n)$.

Solution. Let $X = [x_{ij}]$, $Y = [y_{ij}]$, and $Z = [z_{ij}]$ be elements of $M(m, n)$. Then

$$\begin{aligned} X + (Y + Z) &= [x_{ij}] + ([y_{ij} + z_{ij}]) \\ &= [x_{ij} + (y_{ij} + z_{ij})] \\ &= [(x_{ij} + y_{ij}) + z_{ij}] \quad \text{(from the associative law for numbers)} \\ &= ([x_{ij} + y_{ij}]) + [z_{ij}] = (X + Y) + Z \end{aligned}$$

[1]We use $\mathcal{V}$ in order to avoid the necessity of re-listing these properties when we define the general notion of "vector space."

When we introduced linear independence, we mentioned that for any $m \times n$ matrix A

$$0A = 0$$

This is very simple to prove:

$$0A = 0[a_{ij}] = [0a_{ij}] = 0$$

This proof explicitly uses the fact that we are dealing with matrices. It is possible to give another proof that uses only the vector space properties. We first note from property (i) that

$$0A + 0A = (0 + 0)A = 0A$$

Next, we cancel $0A$ from both sides using the vector space properties:

$$\begin{aligned} -0A + (0A + 0A) &= -0A + 0A && \text{Property (e)} \\ (-0A + 0A) + 0A &= 0 && \text{Property (c)} \\ 0 + 0A &= 0 && \text{Properties (b) and (e)} \\ 0A &= 0 && \text{Properties (b) and (d)} \end{aligned} \tag{1.5}$$

Both proofs are valid for matrices. We, however, prefer the second. Since it used only the vector space properties, it will be valid in any context in which these properties hold. For example, let $\mathcal{F}(\mathbb{R})$ denote the set of all real-valued functions which are defined for all real numbers. Thus, the functions $y = e^x$ and $y = x^2$ are two elements of $\mathcal{F}(\mathbb{R})$. We define addition and scalar multiplication for functions by the formulas

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ (cf)(x) &= cf(x) \end{aligned} \tag{1.6}$$

Thus, for example,

$$y = 3e^x - 7x^2$$

defines an element of $\mathcal{F}(\mathbb{R})$. Since addition and scalar multiplication of functions are defined using the corresponding operations on numbers, it is easily proved that the vector space properties (a)–(j) hold if we interpret A, B, and C as functions rather than matrices. (See Example 1.5.)

Thus, we can automatically state that $0f(x) = 0$, where $f(x)$ represents any function and 0 is the zero function. Admittedly, this is not an exciting result. (Neither, for that matter, is $0A = 0$ for matrices.) However, it demonstrates an extremely important principle: *Anything we prove about matrices using only the vector space properties will be true in any context for which these properties hold.*

As we progress in our study of linear algebra, it will be important to keep track of exactly which facts can be proved directly from the vector space properties and which require additional structure. We do this with the concept of "vector space."

Definition 1.6 *A set $\mathcal{V}$ is a* ***vector space*** *if it has a rule of addition and a rule of scalar multiplication defined on it so that all the vector space properties* (a)–(j) *from Theorem 1.1 hold. By a* ***rule of addition*** *we mean a well-defined process for taking an arbitrary pair of elements X and Y from $\mathcal{V}$ and producing a third element $X + Y$ in $\mathcal{V}$. (Note that the sum must lie in $\mathcal{V}$.) By a* ***rule of scalar multiplication*** *we mean a well-defined process for taking an arbitrary scalar c and an arbitrary element X of $\mathcal{V}$ and producing a second element cX of $\mathcal{V}$.*

The following theorem summarizes our discussion of functions. We leave most of the proof as an exercise.

Theorem 1.2 *The set $\mathcal{F}(\mathbb{R})$ of real valued functions on $\mathbb{R}$ is a vector space under the operations defined by formula* (1.6).

EXAMPLE 1.5

Prove vector space property (h) for $\mathcal{F}(\mathbb{R})$.

Solution. Let $f(x)$ and $g(x)$ be real-valued functions and let $k \in \mathbb{R}$. Then

$$\begin{aligned}(k(f+g))(x) &= k((f+g)(x))\\ &= k(f(x)+g(x))\\ &= kf(x)+kg(x)\\ &= (kf+kg)(x)\end{aligned}$$

showing that $k(f + g) = kf + kg$, as desired.

Any concept defined for $M(m, n)$ solely in terms of addition and scalar multiplication will be meaningful in any vector space $\mathcal{V}$. One simply replaces $M(m, n)$ by $\mathcal{V}$ where $\mathcal{V}$ is a general vector space. Specifically:

(a) *The concept an element C in $\mathcal{V}$ depending on a set $S = \{A_1, A_2, \ldots, A_k\}$ of elements of $\mathcal{V}$ is defined as in Definition* 1.3.
(b) *The concepts of linear independence/dependence for a set $S = \{A_1, A_2, \ldots, A_k\}$ of elements of $\mathcal{V}$ are defined as in Definition* 1.4.
(c) *The concept of the span of a set $S = \{A_1, A_2, \ldots, A_k\}$ of elements of $\mathcal{V}$ is defined as in Definition* 1.5.

EXAMPLE 1.6

Show that the set of functions $\{\sin^2 x, \cos^2 x, 1\}$ is linearly dependent in $\mathcal{F}(\mathbb{R})$.

Solution. This is clear from the formula

$$\sin^2 x + \cos^2 x = 1$$

Theorem 1.1 states that for each m and n, $M(m, n)$ is a vector space. The set of all possible matrices is not a vector space, at least under the usual rules of addition and scalar multiplication. This is because we cannot add matrices unless they are the same size: for example, we cannot add a 2×2 matrix to a 2×3 matrix. Thus, our "rule of addition" is not valid for all matrices.

At the moment, the $M(m, n)$ spaces, along with $\mathcal{F}(\mathbb{R})$, are the only vector spaces we know. This will change in Section 1.5 where we describe the concept of "subspace of a vector space." However, if we say that something is "true for all vector spaces," we are implicitly stating that it can be proved solely on the basis of the vector space properties. Thus, the property that $0A = 0$ is true for all vector spaces. Another important vector space property is the following. The proof (which *must* use only the vector space properties or their consequences) is left as an exercise.

Proposition 1.2 *Let X be an element of a vector space $\mathcal{V}$. Then* $(-1)X = -X$.

Before ending this section, we need to make a comment concerning notation. Writing column vectors takes considerable text space. There is a handy space-saving notation that we shall use often. Let A be an $m \times n$ matrix. The "main diagonal" of A refers to the entries of the form a_{ii}. (Note that all these entries lie on a diagonal line starting at the upper left-hand corner of A.) If we flip A along its main diagonal, we obtain an $n \times m$ matrix, which is denoted A^t and called the **transpose** of A. Mathematically, A^t is the $n \times m$ matrix $[b_{ij}]$ defined by the formula

$$b_{ij} = a_{ji}$$

Thus if

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}, \text{ then } A^t = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$

Notice that the columns of A become rows in A^t. Thus, $[2, 3, -4, 10]^t$ is a space efficient way of writing the column vector

$$\begin{bmatrix} 2 \\ 3 \\ -4 \\ 10 \end{bmatrix}$$

Remark. The reader will discover in later sections that the transpose of a matrix has importance far beyond typographical convenience.

Why Prove Anything?

There is a fundamental difference between mathematics and science. Science is founded on experimentation. If certain principles (such as Newton's laws of motion) seem to be valid every time experiments are done to verify them, they are accepted as a "law."

They will remain a law only as long as they agree with experimental evidence. Thus, Newton's laws were eventually replaced by the theory of relativity when they were found to conflict with the experiments of Michelson and Morley. Mathematics, on the other hand, is based on *proof*. No matter how many times some mathematical principle is observed to hold, we will not accept it as a "theorem" until we can produce a logical argument that shows the principle can *never* be violated.

One reason for this insistence on proof is the wide applicability of mathematics. Linear algebra, for example, is essential to a staggering array of disciplines including (to mention just a few) engineering (all types), biology, physics, chemistry, economics, social sciences, forestry, and environmental science. We must be certain that our "laws" hold, regardless of the context in which they are applied. Beyond this, however, proofs also serve as explanations of *why* our laws are true. We cannot say that we truly understand some mathematical principle until we can prove it.

Mastery of linear algebra, of course, requires that the student learn a body of computational techniques. Beyond this, however, the student should read and, most important, *understand* the proofs. The student will also be asked to create his or her own proofs. This is because it cannot be truly said that we understand something until we can explain it to someone else.

In writing a proof, the student should always bear in mind that *proofs are communication.* One should envision the "audience" as another student who wants to be convinced of the validity of what is being proved. This other student will question anything that is not a simple consequence of something that he or she already understands.

True-False Questions: Justify your answers.

1.1 A subset of a linearly independent set is linearly independent.

1.2 A subset of a linearly dependent set is linearly dependent.

1.3 A set that contains a linearly independent set is linearly independent.

1.4 A set that contains a linearly dependent set is linearly dependent.

1.5 If a set of elements of a vector space is linearly dependent, then each element of the set is a linear combination of the other elements of the set.

1.6 A set of vectors that contains the zero vector is linearly dependent.

1.7 If X is in the span of A_1, A_2, and A_3, then the set $\{X, A_1, A_2, A_3\}$ is linearly independent as long as the A_i are independent.

1.8 If $\{X, A_1, A_2, A_3\}$ is linearly dependent then X is in the span of A_1, A_2, and A_3.

1.9 The following set of vectors is linearly independent:

$$[1,0,1,1,0]^t, \quad [0,1,0,2,0]^t, \quad [2,0,0,3,4]^t$$

1.10 The following matrices form a linearly independent set:

$$\begin{bmatrix} 2 & 1 \\ 3 & 1 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 \\ 3 & -1 \\ 2 & 2 \end{bmatrix}, \quad \begin{bmatrix} 4 & 5 \\ 9 & -1 \\ 5 & 5 \end{bmatrix}$$

1.11 If $\{A_1, A_2, A_3\}$ is a linearly dependent set of matrices, then $\{A_1^t, A_2^t, A_3^t\}$ is also a linearly dependent set.

1.12 The set of functions $\{\tan^2 x, \sec^2 x, 1\}$ is a linearly independent set of elements of the vector space of all continuous functions on the interval $(-\pi/2, \pi/2)$.

EXERCISES

A check mark ✓ next to an exercise indicates that there is an answer/solution provided in the Student Resource Manual. A double check mark ✓✓ indicates that there is also an answer/hint provided in the Answers and Hints section at the back of the text.

1.1 In each case, explicitly write out the matrix A, where $A = [a_{ij}]$. Also, give the third row (written as a row vector) and the second column (written as a column vector).

(a) ✓✓ $a_{ij} = 2i - 3j$, where $1 \le i \le 3$ and $1 \le j \le 4$

(b) ✓ $a_{ij} = i^2 j^3$, where $1 \le i \le 3$ and $1 \le j \le 2$

(c) ✓✓ $a_{ij} = \cos(ij\pi/3)$, where $1 \le i \le 3$ and $1 \le j \le 2$

1.2 For the matrices A, B, and C below, compute (in the order indicated by the parentheses) $(A+B)+C$ and $A+(B+C)$ to illustrate that $(A+B)+C = A + (B+C)$. Also illustrate the distributive law by computing $3(A+B)$ and $3A + 3B$.

$$A = \begin{bmatrix} 1 & 1 & 2 \\ 0 & 1 & -2 \\ 2 & 0 & 1 \\ 3 & 2 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 0 & 1 \\ 4 & 1 & 0 \\ 2 & 3 & 2 \\ -1 & 2 & -2 \end{bmatrix}, \quad C = \begin{bmatrix} 3 & 1 & 3 \\ 4 & 2 & -2 \\ 4 & 3 & 3 \\ 2 & 4 & -1 \end{bmatrix}$$

1.3 ✓✓The set of matrices $\{A, B, C\}$ from Exercise 1.2 is linearly dependent. Express one element of this set as a linear combination of the others. You should be able to solve this by inspection (guessing).

1.4 Let A, B, and C be as in Exercise 1.2. Give a fourth matrix D (reader's choice) that belongs to the span of these matrices.

1.5 Each of the following sets of matrices is linearly dependent. Demonstrate this by explicitly exhibiting one of the elements of the set as a linear combination of the others. You should be able to find the constants by inspection (guessing).

(a) ✓✓ $\{[1, 1, 2], [0, 0, 1], [1, 1, 4]\}$

(b) $\{[1, 0, 0], [0, 1, 0], [0, 0, 1], [1, 2, 3]\}$

(c) ✓✓ $\left\{ \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \right\}$

(d) ✓✓ $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 9 \\ 12 \\ 15 \end{bmatrix} \right\}$

(e) $\left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & -4 \end{bmatrix} \right\}$

(f) ✓ $\left\{ \begin{bmatrix} 3 & -1 & 2 \\ 0 & 1 & 4 \end{bmatrix}, \begin{bmatrix} -9 & 3 & -6 \\ 0 & -3 & -12 \end{bmatrix} \right\}$

(g) $\left\{ \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix} \right\}$

1.6 ✓✓Write the second row of the Acme profit matrix P (Table (1.1) on page 2) as a linear combination of the other rows.

1.7 Write the first column of the Acme profit matrix P (Table (1.1) on page 2) as a linear combination of the other columns.

1.8 Verify the Remark following Example 1.2 on page 8, that is, show that A_1 is not a linear combination of A_2, A_3, and A_4.

1.9 ✓✓What general feature of the following matrices makes it clear that they are independent?

$$\begin{bmatrix} 1 \\ 0 \\ 3 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 0 \\ -5 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ 1 \\ 13 \\ 0 \end{bmatrix}$$

1.10 ✓Prove that the rows of the following matrix are linearly independent. [*Hint:* Assume $A_3 = xA_1 + yA_2$, where A_i is the ith row of A. Prove first that $x = 0$ and then show $y = 0$, which is impossible. Repeat for the other rows. [In Section 2.1 we discuss a more efficient way of solving such problems.]

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 5 & 6 \\ 0 & 0 & 8 \end{bmatrix}$$

1.11 Prove that the columns of the matrix in Exercise 1.10 are linearly independent. [*Hint:* See the hint for Exercise 1.10.]

1.12 Let $X = [1, -1, 0]$ and $Y = [1, 0, 0]$. Give four row vectors (reader's choice) that belong to the span of X and Y. Give an element of $M(1, 3)$ that does not belong to the span of X and Y.

1.13 ✓✓Let $X = [-1, 1, -1]$ and $Y = [-1, 3, 2]$.

(a) Find an element in the span of X and Y such that each of its entries is positive.

(b) Show that every element $[x, y, z]$ of the span of X and Y satisfies $5x + 3y - 2z = 0$.

(c) Give an element of $M(1, 3)$ that does not belong to the span of X and Y.

1.14 Find two elements of $\mathbb{R}^4$ which belong to the span of the following vectors. Find an element of $\mathbb{R}^4$ which does not belong to their span. [*Hint:* Compute the sum of the entries of each of the given vectors.]

$$X_1 = [1, 1, -1, -1]^t, \quad X_2 = [2, -1, -3, 2]^t, \quad X_3 = [1, 3, -2, -2]^t$$

1.15 Find an element in the span of the vectors $X = [-1, 2, 1]^t$ and $Y = [2, 5, 1]^t$ which has its third entry equal to 0 and its other two entries positive.

1.16 ✓Let $X = [1, -1, 0]^t$ and $Y = [1, 0, -1]^t$. Are there any elements in their span with all entries positive? Explain.

1.17 Let $X = [1, -2, 4]^t$ and $Y = [-1, 2, 3]^t$. Are there any elements in their span with all entries positive? Explain.

1.18 Let $X = \begin{bmatrix} 2 & 3 \\ 1 & 4 \end{bmatrix}$ and $Y = \begin{bmatrix} -2 & -3 \\ 4 & -1 \end{bmatrix}$. Are there any elements in their span in the vector space $M(2, 2)$ with all entries positive? Explain.

1.19 For each of the following sets of functions either find a function $f(x)$ in their span such that $f(x) > 0$ for all x or prove that no such function exists.

(a) $\{\sin x, 1\}$ **(b)** $\{\cos x, 1\}$

(c) ✓ $\{\sin x, \cos x\}$

1.20 The xy plane in $\mathbb{R}^3$ is the set of elements of the form $[x, y, 0]^t$. Find a nonzero element of the xy plane that belongs to the span of the vectors X and Y from (a) Exercise 1.16 and (b) Exercise 1.17. (c)✓Find two nonzero vectors X and Y in $\mathbb{R}^3$, $X \neq Y$, for which there are NO *nonzero* vectors Z in the xy plane that also belong to the span of X and Y.

1.21 Let $X = [x_1, y_1, z_1]$ and $Y = [x_2, y_2, z_2]$ be elements of $M(1, 3)$. Suppose that a, b, and c are such that $ax_i + by_i + cz_i = 0$ for $i = 1, 2$. Show that every element $[x, y, z]^t$ of the span of X and Y satisfies $ax + by + cz = 0$.

1.22 In Exercise 1.16, find constants a, b, c, not all zero, such that every element $[x, y, z]$ of the span of X and Y satisfies the equation $ax + by + cz = 0$. Repeat for Exercise 1.17. Explain geometrically why such constants should exist. [*Hint:* The equation $ax + by + cz = 0$ describes a plane through 0, as long as at least one of a, b, and c is nonzero.]

1.23 ✓✓Let X, Y, and Z be as shown. Give four matrices (reader's choice) that belong to their span. Give a matrix that does not belong to their span.

$$X = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \quad Y = \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}, \quad Z = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

1.24 In the following questions we investigate geometric significance of spanning and linear independence.

(a) Sketch the span of $[1, 2]^t$ in $\mathbb{R}^2$.

(b) What do you guess that the span of $\{[1, 2]^t, [1, 1]^t\}$ in $\mathbb{R}^2$ is? Draw a diagram to support your guess.

(c) Do you think that it is possible to find three linearly independent matrices in $M(2, 1)$?

(d) On the basis of your answer to (c), do you think that it is possible to construct a 2×3 matrix with linearly independent columns? How about a 3×2 matrix with linearly independent rows?

(e) Sketch (as best you can) the span of $\{[1, 1, 0]^t, [0, 0, 1]^t\}$ in $\mathbb{R}^3$.

(f) How does the span of $\{[1, 1, 1]^t, [1, 1, 0]^t\}$ in $\mathbb{R}^3$ compare with that in part (e)?

(g) Sketch the span of $\{[1, 1, 1]^t, [2, 2, 2]^t\}$ in $\mathbb{R}^3$. Why is this picture so different from that in part (f)? Bring the phrase " linearly dependent" into your discussion.

1.25 ✓✓Suppose that V and W both belong to the span of X and Y in some vector space. Show that all linear combinations of V and W also belong to this span.

1.26 ✓✓The columns of the following matrix A are linearly dependent. Exhibit one column as a linear combination of the other columns.

$$\begin{bmatrix} 6 & 6 & 4 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix}$$

1.27 Let A be as in Exercise 1.26. Exhibit one row of A as a linear combination of the other rows.

1.28 Is it possible to find a 2×2 matrix whose rows are linearly dependent but whose columns are linearly independent? Prove your answer.

1.29 ✓✓Construct an example of your own choice of a 4×4 matrix with linearly dependent rows having all of its entries nonzero.

1.30 Construct an example of your own choice of a 4×4 matrix with linearly dependent columns having all of its entries nonzero.

1.31 ✓Let $S = \{A, B, C, D\}$ be some set of four elements of some vector space. Suppose that $D = 2A + B + 3C$ and $C = A - B$. (a) Is $\{A, B, D\}$ linearly dependent? Explain. (b) Is $\{A, C, D\}$ linearly dependent? (c) What can you conclude (if anything) about the linear dependence of $\{A, B\}$?

1.32 Let $S = \{A, B, C, D\}$ be some set of four elements of some vector space. Suppose that $A = B - 3C + D$ and $C = A - B$. (a) Is $\{A, B, D\}$ linearly dependent? Explain. (b) Is $\{A, C, D\}$ linearly dependent? (c) What can you conclude (if anything) about the linear dependence of $\{A, C\}$?

1.33 The following sets of functions are linearly dependent in $\mathcal{F}(\mathbb{R})$. Show this by expressing one of them as a linear combination of the others. (You may need to look up the definitions of the sinh and cosh functions as well as some trigonometric identities in a calculus book.)

(a) $\{3\sin^2 x, -5\cos^2 x, 119\}$

(b) ✓✓ $\{2e^x, 3e^{-x}, \sinh x\}$

(c) $\{\sinh x, \cosh x, e^{-x}\}$

(d) ✓✓ $\{\cos(2x), \sin^2 x, \cos^2 x\}$

(e) $\{\cos(2x), 1, \cos^2 x\}$

(f) ✓✓ $\{(x+3)^2, 1, x, x^2\}$

(g) $\{x^2 + 3x + 3, x + 1, 2x^2\}$

(h) ✓✓ $\{\sin x, \sin(x + \frac{\pi}{4}), \cos(x + \frac{\pi}{4})\}$

(i) ✓✓ $\{\ln[(x^2+1)^3/(x^4+7)], \ln\sqrt{x^2+1}, \ln(x^4+7)\}$

1.34 ✓Give two examples of functions in the span of the functions $\{1, x, x^2\}$. Describe in words what the span of these three functions is. [Some useful terminology: the function $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial of degree less than or equal to n. Its degree equals n if $a_n \neq 0$.]

1.35 Repeat Exercise 1.34 for the polynomials $\{1, x, x^2, x^3\}$.

1.36 ✓✓Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

(a) Use the definition of matrix addition to prove that the only 2×2 matrix B such that $A + B = A$ is the zero matrix. The point of this problem is that one should think of $A + 0 = A$ as the defining property of the zero matrix.

(b) Use the definition of matrix addition to prove that the only 2×2 matrix B such that $A + B = 0$ is

$$B = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$$

The point of this problem is that one should think of $A + (-A) = 0$ as the defining property of $-A$.

1.37 Prove vector space properties (b), (e), (g), (h), and (i)✓for $M(m, n)$.

1.38 Let B, C, and X be elements of some vector space. In the following discussion, we solved the equation $3X + B = C$ for X. At each step we used one of the vector space properties. Which property was used? [*Note:* We define $C - B = C + (-B)$.]

$$\begin{aligned}
3X + B &= C && \\
(3X + B) + (-B) &= C + (-B) && \text{Properties (a) and (e)} \\
3X + (B + (-B)) &= C - B && \text{Definition of } C - B \text{ and property (?)} \\
3X + 0 &= C - B && \text{Property (?)} \\
3X &= C - B && \text{Property (?)} \\
\frac{1}{3}(3X) &= \frac{1}{3}(C - B) && \text{Property (?)} \\
\left(\frac{1}{3}3\right) X &= \frac{1}{3}(C - B) && \text{Property (?)} \\
1X &= \frac{1}{3}(C - B) && \\
X &= \frac{1}{3}(C - B) && \text{Property (?)}
\end{aligned}$$

1.39 Let X and Y be elements of some vector space. Prove, putting in every step, that $-(2X + 3Y) = (-2)X + (-3)Y$. You may find Proposition 1.2 useful.

1.40 ✓✓Let X, Y, and Z be elements of some vector space. Suppose that there are scalars a, b, and c such that $aX + bY + cZ = 0$. Show that if $a \neq 0$, then

$$X = \left(-\frac{b}{a}\right) Y + \left(-\frac{c}{a}\right) Z$$

Do your proof in a step-by-step manner to demonstrate the use of each vector space property needed. [*Note*: In a vector space, $X + Y + Z$ denotes $X + (Y + Z)$.]

1.41 Prove that in any vector space, if $X + Y = 0$, then $Y = -X$. (Begin by adding $-X$ to both sides of the given equality.)

1.42 Prove Proposition 1.2. [*Hint:* From Exercise 1.41, it suffices to prove that $X + (-1)X = 0$.]

1.1.1 Computer Projects

Our goal in this discussion is to plot some elements of the span of the vectors $A = [1, 1]$ and $B = [2, 3]$ using MATLAB. Before we begin, however, let us make a few general comments. When you start up MATLAB, you will see something like >> followed by a blank line. If the instructions ask you to enter $2 + 2$, then you should type $2 + 2$ on the screen behind the >> prompt and then press the enter key. Try it!

```
>> 2+2
ans =
     4
```

Entering matrices into MATLAB is not much more complicated. Matrices begin with "[" and end with "]". Entries in rows are separated with either commas or spaces. Thus, after starting MATLAB, our matrices A and B would be entered as shown. Note that MATLAB repeats our matrix, indicating that it has understood us.

```
>> A = [1  1]
A =
    1     1
>> B = [1  3]
B =
    1     3
```

Next we construct a few elements of the span of A and B. If we enter "2∗A+B", MATLAB responds

```
ans =
     3    5
```

(Note that ∗ is the symbol for "times." MATLAB will complain if you simply write 2A+B.)

If we enter (−5) ∗A +7 ∗B, MATLAB responds

```
ans =
     2    16
```

Thus, the vectors $[3, 5]$ and $[2, 16]$ both belong to the span.

We can get MATLAB to automatically generate elements of the span. Try entering the word "rand". This should cause MATLAB to produce a random number between 0 and 1. Enter "rand" again. You should get a different random number. It follows that entering the command C=rand∗A+rand∗B should produce random linear combinations of A and B. Try it!

To see more random linear combinations of these vectors, push the up-arrow key. This should return you to the previous line. Now you can simply hit "enter" to produce a new random linear combination. By repeating this process, you can produce as many random elements of the span as you wish.

Next, we will plot our linear combinations. Begin by entering the following lines. Here "figure" creates a figure window, "hold on" tells MATLAB to plot all points on the same graph, and "axis([-5,5,-5,5])" tells MATLAB to show the range $-5 \leq x \leq 5$ and $-5 \leq y \leq 5$ on the axes. The command "hold on" will remain in effect until we enter "hold off":

```
>> figure
>> hold on
>> axis([-5,5,-5,5])
```

A window (the Figure window) showing a blank graph should pop up.

Points are plotted in MATLAB using the command "plot." For example, entering plot(3,4) will plot the point $(3, 4)$. Return to the MATLAB Command window and try plotting a few points of your own choosing. (To see your points, you will either need to return to the Figure window or move and resize the Command and Figure windows so that you can see them both at the same time. Moving between windows is accomplished by pulling down the Window menu.) When you are finished, clear the figure window by entering "cla" and then enter the following line:

```
C=rand*A+rand*B plot(C(1),C(2))
```

This will plot one point in the span. [C(1) is the first entry of C and C(2) is the second.] You can plot as many points as you wish by using the up-arrow key as before.

EXERCISES

1. Plot the points $[1, 1]$, $[1, -1]$, $[-1, 1]$, and $[-1, -1]$ all on the same figure. When finished, clear the figure window by entering the "cla" command.

2. Enter the vectors A and B from the discussion above.

- **(a)** Get MATLAB to compute several different linear combinations of them. (Reader's choice.)
- **(b)** Use C=rand*A+rand*B to create several "random" linear combinations of A and B.
- **(c)** Plot enough points in the span of A and B to obtain a discernible geometric figure. Be patient. This may require plotting over 100 points. What kind of geometric figure do they seem to form? What are the coordinates of the vertices?

  Note: If your patience runs thin, you might try entering the following three lines. The ";" keeps MATLAB from echoing the command every time it is being executed.

```
for i=1:200
  C=rand*A+rand*B; plot(C(1),C(2));
end
```

This causes MATLAB to execute any commands between the "for" and "end" statements 200 times.

(d) The plot in part (c) is only part of the span. To see more of the span, enter the commands

```
for i=1:200
  C=2*rand*A+rand*B; plot(C(1),C(2),'r');
end
```

The "r" in the plot command tells MATLAB to plot in red.

3. Describe in words the set of points $s{*}A + t{*}B$ for $-2 \leq s \leq 2$ and $-2 \leq t \leq 2$. Create a MATLAB plot that shows this set reasonably well. Use yet another color. (Enter "help plot" to see the choice of colors.) [*Hint:* "rand" produces random numbers between 0 and 1. What would "rand-0.5" produce?]

4. In Exercise 1.13 on page 19, it was stated that each element of the span of X and Y satisfies $5x + 3y - 2z = 0$.

(a) Check this by generating a random matrix C in the span of X and Y and computing 5∗C(1)+3*C(2)-2*C(3). Repeat with another random element of the span.

(b) Plot a few hundred elements of this span in $\mathbb{R}^3$. Before doing so, close the Figure window by selecting Close from the File menu. Next, enter "figure", then "axis([-4,4,-4,4,-4,4])", and "hold on". A three-dimensional graph should pop up. The command plot3(C(1),C(2),C(3)) plots the three-dimensional vector C.

Describe the geometric figure so obtained. What are the coordinates of the vertices? Why is this to be expected?

1.1.2 Applications to Graph Theory I

Figure 1.7 represents the route map of an airline that serves four cities, A, B, C, and D. Each arrow represents a daily flight between the two cities.

The information from this diagram can be represented in tabular form, where the numbers represent the number of daily flights between the cities:

from/to	A	B	C	D
A	0	1	0	1
B	1	0	0	1
C	0	1	0	1
D	1	0	2	0

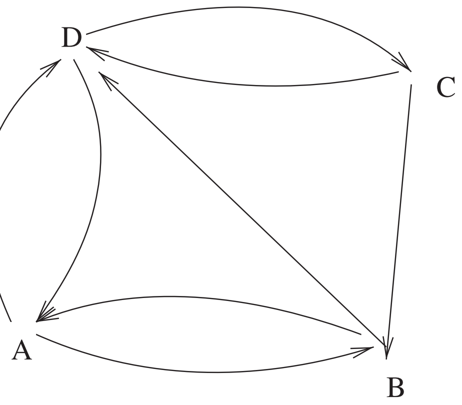

FIGURE 1.7 Route map.

The 4×4 matrix obtained by deleting the labels from the preceding table is what we refer to as the route matrix. The route matrix could be stored in a computer. One could then, for example, use this information as the basis for a computer program to find the shortest connection for a customer.

Route maps are examples of what are called **directed graphs**. In general, a directed graph is a finite set of points (called **vertices**), together with arrows connecting some of the vertices. A directed graph may be described using a matrix just as was done for route maps. Specifically, if the vertices are $V_1, V_2, \ldots, V_n$, then the graph will be represented by the matrix A, where a_{ij} is the number of arrows from V_i to V_j.

Graph theory may also be applied to anthropology. Suppose an anthropologist is studying generational dominance in an extended family. The family members are M (mother), F (father), S1 (first born son), S2 (second born son), D1 (first born daughter), D2 (second born daughter), MGM (maternal grandmother), MGF (maternal grandfather), PGM (paternal grandmother), and PGF (paternal grandfather). The anthropologist represents the dominance relationships by a directed graph where an arrow is drawn from each individual to any individual he or she directly dominates. In the exercises you will study the dominance relationship given in Figure 1.8.

We will say more about the matrix of a graph in Section 3.2.

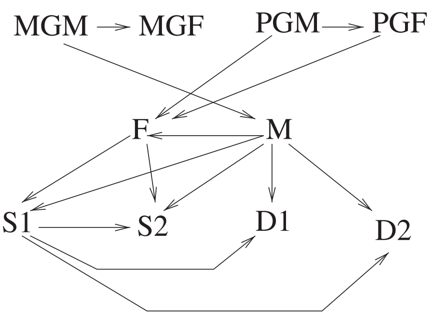

FIGURE 1.8 Dominance in an extended family.

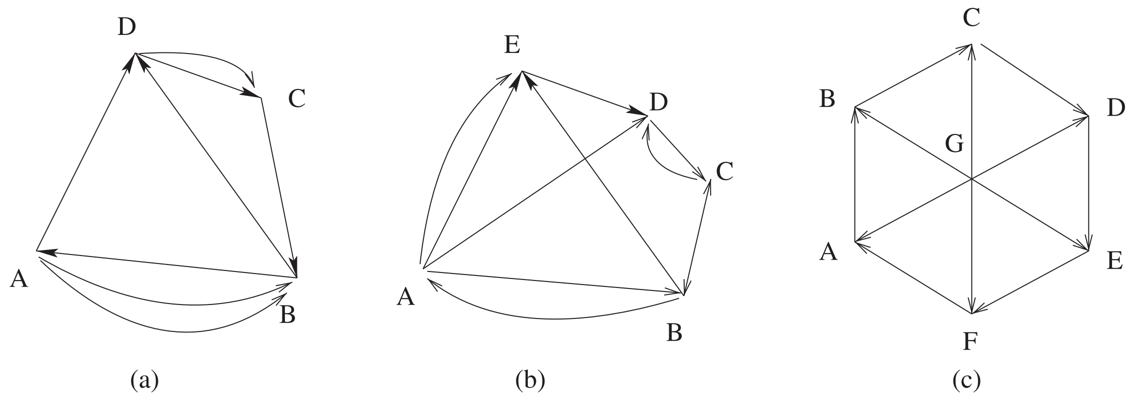

FIGURE 1.9 Exercise 1.

Self-Study Questions

1.1 ✓Give the route matrix for each of the route maps in Figure 1.9 where the nodes are listed in alphabetical order.

1.2 ✓For each of the following matrices, draw a route map that could correspond to the given matrix:

$$\text{(a)}\quad \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \text{(b)}\quad \begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

1.3 ✓Why are entries on the diagonal in a route matrix always zero?

EXERCISES

1.43 Under what circumstances does a route matrix A satisfy $A = A^t$?

1.44 What would be the significance for the route if all the entries in a given column were zero? What if a given row were zero? What if a row and a column were zero?

1.45 ✓Give the dominance matrix (the matrix for the graph) for the dominance relationship described by Figure 1.8.

1.46 Suppose that A is the matrix of a dominance relationship. Explain why $a_{ij}a_{ji} = 0$.

1.47 ✓We say that two points A and B of a directed graph are two-step connected if there is a point C such that $A \to C \to B$. Thus, for example, in the route map in Figure 1.7, A and C are two-step connected, but D and C are not. Also A

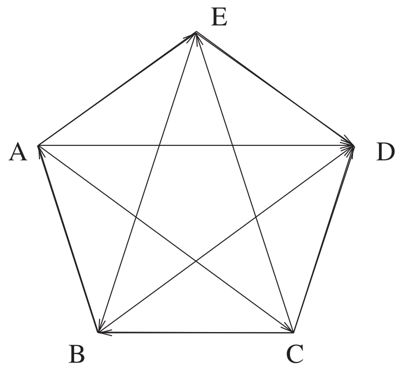

FIGURE 1.10 Exercise 1.48.

is two-step connected with itself. Give the two-step route matrix for the route map in Figure 1.7.

1.48 Figure 1.10 shows the end-of-season results from an athletic conference with teams A–D can be described using a graph. The arrows indicate which team beat which.

1. Find the matrix A for the graph in Figure 1.10.

2. Compute the win–loss record of team C.

1.2 SYSTEMS

An equation in variables $x_1, x_2, \dots, x_n$ is a **linear equation** if and only if it is expressible in the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b \tag{1.7}$$

where a_i and b are all scalars. By a **solution** to equation (1.7) we mean a column vector $[x_1, x_2, \dots, x_n]^t$ of values for the variables that make the equation valid. Thus, $X = [1, 2, -1]^t$ is a solution of the equation

$$2x + 3y + z = 7$$

because

$$2(1) + 3(2) + (-1) = 7$$

More generally, a set of linear equations in a particular collection of variables is called a **linear system** of equations. Thus, the general system of linear equations in the variables $x_1, x_2, \dots, x_n$ may be written as follows:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.8}$$

A solution to the system is a column vector that is a solution to each equation in the system. The set of all column vectors that solve the system is the **solution set** for the system.

In particular,

$$\begin{aligned} x + 2y + z &= 1 \\ 3x + y + 4z &= 0 \\ 2x + 2y + 3z &= 2 \end{aligned} \tag{1.9}$$

is a linear system in the variables x, y, and z.

Finding all solutions to this system is not hard. We begin by subtracting three times the first equation from the second, producing

$$\begin{aligned} x + 2y + \ \ z &= 1 \\ -5y + \ \ z &= -3 \\ 2x + 2y + 3z &= 2 \end{aligned} \tag{1.10}$$

Any x, y, and z that satisfy the original system also satisfy the system above.

Conversely, notice that we can transform the above system back into the original by *adding* three times the first equation onto the second equation. Thus, any variables that satisfy the second system must also satisfy the first. Hence, both systems have the same solution set. We say that these systems are equivalent:

Definition 1.7 *Two systems of linear equations in the same variables are* ***equivalent*** *if they have the same solution set.*

To continue the solution process, we next subtract twice the first equation from the third, producing

$$\begin{aligned} x + 2y + z &= 1 \\ -5y + z &= -3 \\ -2y + z &= 0 \end{aligned}$$

Note that we have eliminated all occurrences of x from the second and third equations. This system is equivalent with our second system for similar reasons that the second

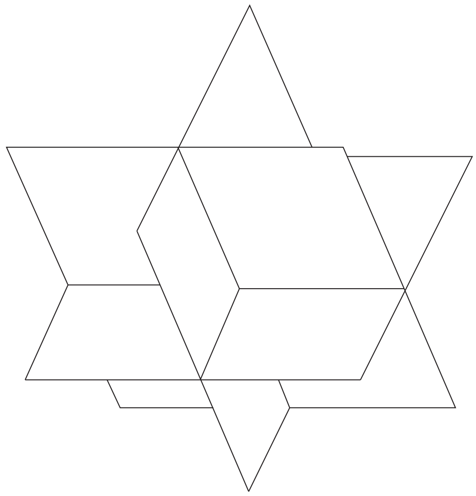

FIGURE 1.11 Only one solution.

system was equivalent with the first. It follows that this system has the same solution set as the original system.

Next, we eliminate y from the third equation by subtracting twice the second from five times the third, again producing an equivalent system:

$$\begin{aligned} x + 2y + \quad z &= 1 \\ -5y + \quad z &= -3 \\ 3z &= 6 \end{aligned} \tag{1.11}$$

Thus, $z = 2$. Then, from the second equation, $y = 1$, and finally, from the first equation, $x = -3$. Thus, our only solution is $[-3, 1, 2]^t$.

The fact that there was only one solution can be understood geometrically. Each of the equations in system (1.9) describes a plane in $\mathbb{R}^3$. A point that satisfies each equation in the system must lie on all three planes. Typically, three planes in $\mathbb{R}^3$ intersect at precisely one point, as shown in Figure 1.11.

Remark. The method we used to compute the solution from system (1.11) is referred to as **back substitution**. In general, in back substitution, we solve the last equation for one variable and then substitute the result into the preceding equations, obtaining a system with one fewer variable and one fewer equation, to which the same process may be repeated. In this way, we obtain all solutions to the system.

The process we used to reduce system (1.9) to system (1.11) is called **Gaussian elimination**. The general idea is to use the first equation to eliminate all occurrences of the first variable from the equations below it. One then attempts to use the second equation to eliminate the next variable from all equations below it, and so on. In the end, the last variable is determined first (z in our example) and then the others are determined by substitution as in the example.

We will describe Gaussian elimination in detail in the next section, after considering several more examples. First, however, we introduce a "shorthand" notation