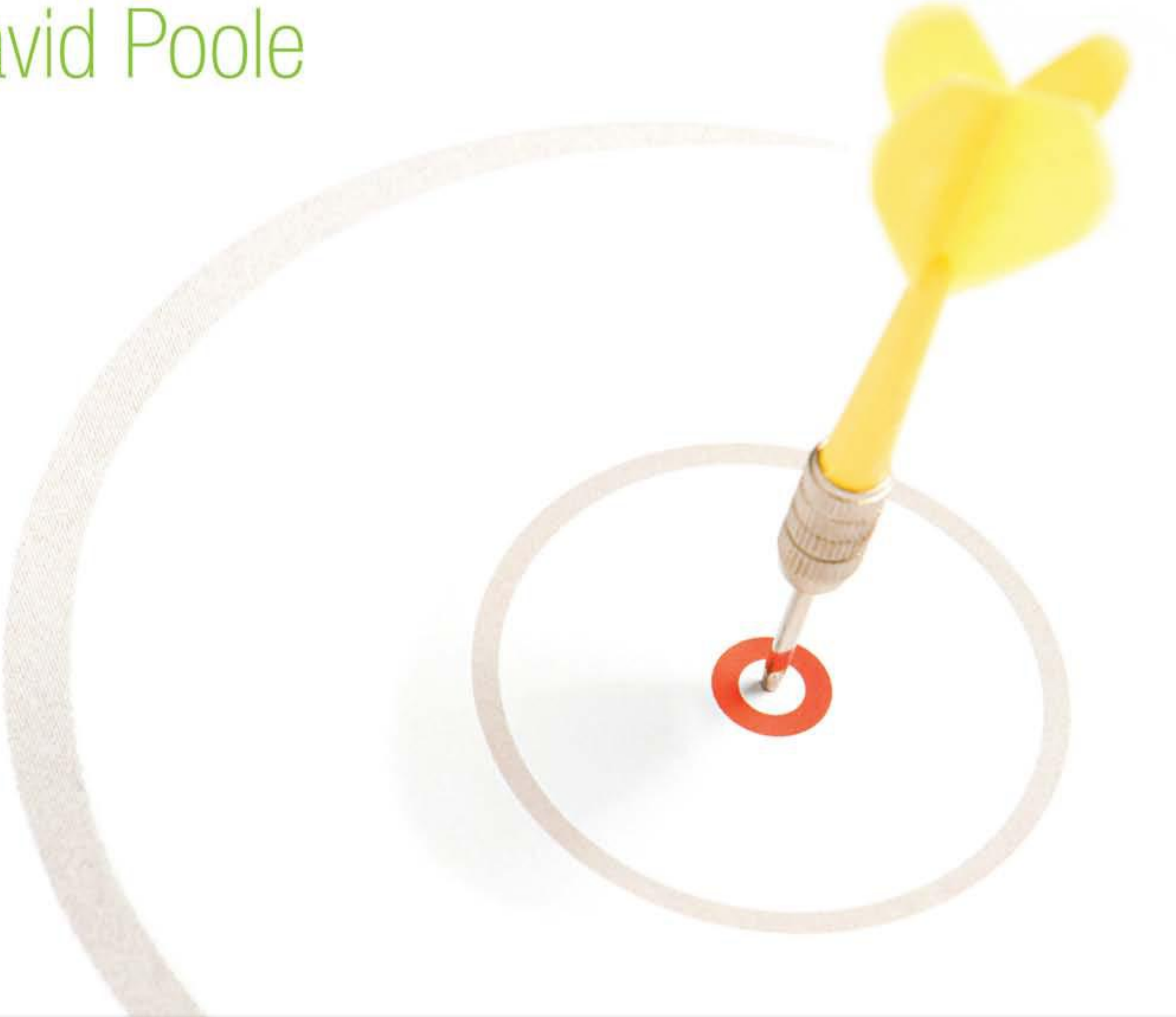
David Poole



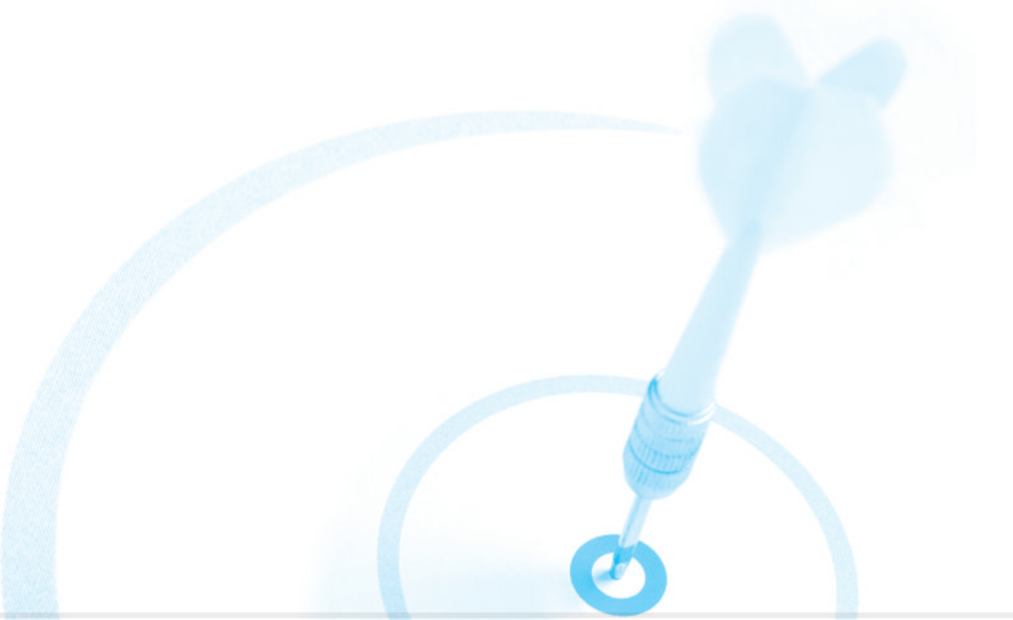
Linear Algebra

A MODERN INTRODUCTION

4th edition

EXPLORATIONS AND VIGNETTES

APPLICATIONS



Linear Algebra

A MODERN INTRODUCTION

Fourth edition

David Poole

Trent University

CENGAGE Learning®

Australia • Brazil • Mexico • Singapore • United Kingdom • United States

Linear Algebra
A Modern Introduction, 4th Edition
David Poole

Product Director: Liz Covello

Product Team Manager: Richard Stratton

Content Developer: Laura Wheel

Product Assistant: Danielle Hallock

Media Developer: Andrew Coppola

Content Project Manager: Alison Eigel Zade

Senior Art Director: Linda May

Manufacturing Planner: Doug Bertke

Rights Acquisition Specialist: Shalice Shah-Caldwell

Production Service & Compositor: MPS Limited

Text Designer: Leonard Massiglia

Cover Designer: Chris Miller

Cover & Interior design Image: Image Source/Getty Images

Library of Congress Control Number: 2013944173

ISBN-13: 978-1-285-46324-7

ISBN-10: 1-285-46324-2

Cengage Learning
200 First Stamford Place, 4th Floor
Stamford, CT 06902
USA

Printed in the United States of America
1 2 3 4 5 6 7 17 16 15 14 13

Dedicated to the memory of Peter Hilton, who was an exemplary mathematician, educator, and citizen—a unit vector in every sense.

Contents

Preface

The last thing one knows when writing a book is what to put first.

—Blaise Pascal
Pensées, 1670

The fourth edition of *Linear Algebra: A Modern Introduction* preserves the approach and features that users found to be strengths of the previous editions. However, I have streamlined the text somewhat, added numerous clarifications, and freshened up the exercises.

I want students to see linear algebra as an exciting subject and to appreciate its tremendous usefulness. At the same time, I want to help them master the basic concepts and techniques of linear algebra that they will need in other courses, both in mathematics and in other disciplines. I also want students to appreciate the interplay of theoretical, applied, and numerical mathematics that pervades the subject.

This book is designed for use in an introductory one- or two-semester course sequence in linear algebra. First and foremost, it is intended for students, and I have tried my best to write the book so that students not only will find it readable but also will *want* to read it. As in the first three editions, I have taken into account the reality that students taking introductory linear algebra are likely to come from a variety of disciplines. In addition to mathematics majors, there are apt to be majors from engineering, physics, chemistry, computer science, biology, environmental science, geography, economics, psychology, business, and education, as well as other students taking the course as an elective or to fulfill degree requirements. Accordingly, the book balances theory and applications, is written in a conversational style yet is fully rigorous, and combines a traditional presentation with concern for student-centered learning.

There is no such thing as a universally best learning style. In any class, there will be some students who work well independently and others who work best in groups; some who prefer lecture-based learning and others who thrive in a workshop setting, doing explorations; some who enjoy algebraic manipulations, some who are adept at numerical calculations (with and without a computer), and some who exhibit strong geometric intuition. In this edition, I continue to present material in a variety of ways—*algebraically, geometrically, numerically*, and *verbally*—so that all types of learners can find a path to follow. I have also attempted to present the theoretical, computational, and applied topics in a flexible yet integrated way. In doing so, it is my hope that all students will be exposed to the many sides of linear algebra.

For more on the recommendations of the Linear Algebra Curriculum Study Group, see *The College Mathematics Journal* **24** (1993), 41–46.

This book is compatible with the recommendations of the Linear Algebra Curriculum Study Group. From a pedagogical point of view, there is no doubt that for most students concrete examples should precede abstraction. I have taken this approach here. I also believe strongly that linear algebra is essentially about vectors and that students need to see vectors first (in a concrete setting) in order to gain some geometric insight. Moreover, introducing vectors early allows students to see how systems of linear equations arise naturally from geometric problems. Matrices then arise equally naturally as coefficient matrices of linear systems and as agents of change (linear transformations). This sets the stage for eigenvectors and orthogonal projections, both of which are best understood geometrically. The dart that appears on the cover of this book symbolizes a vector and reflects my conviction that geometric understanding should precede computational techniques.

I have tried to limit the number of theorems in the text. For the most part, results labeled as theorems either will be used later in the text or summarize preceding work. Interesting results that are not central to the book have been included as exercises or explorations. For example, the cross product of vectors is discussed only in explorations (in Chapters 1 and 4). Unlike most linear algebra textbooks, this book has no chapter on determinants. The essential results are all in Section 4.2, with other interesting material contained in an exploration. The book is, however, comprehensive for an introductory text. Wherever possible, I have included elementary and accessible proofs of theorems in order to avoid having to say, "The proof of this result is beyond the scope of this text." The result is, I hope, a work that is self-contained.

I have not been stingy with the applications: There are many more in the book than can be covered in a single course. However, it is important that students see the impressive range of problems to which linear algebra can be applied. I have included some modern material on finite linear algebra and coding theory that is not normally found in an introductory linear algebra text. There are also several impressive real-world applications of linear algebra and one item of historical, if not practical, interest; these applications are presented as self-contained "vignettes."

I hope that instructors will enjoy teaching from this book. More important, I hope that students using the book will come away with an appreciation of the beauty, power, and tremendous utility of linear algebra and that they will have fun along the way.

What's New in the Fourth Edition

The overall structure and style of *Linear Algebra: A Modern Introduction* remain the same in the fourth edition.

Here is a summary of what is new:

- The applications to coding theory have been moved to the new online Chapter 8.
- To further engage students, five writing projects have been added to the exercise sets. These projects give students a chance to research and write about aspects of the history and development of linear algebra. The explorations, vignettes, and many of the applications provide additional material for student projects.

See pages 49, 82, 283, 301, 443

- There are over 200 new or revised exercises. In response to reviewers' comments, there is now a full proof of the Cauchy-Schwarz Inequality in Chapter 1 in the form of a guided exercise.
- I have made numerous small changes in wording to improve the clarity or accuracy of the exposition. Also, several definitions have been made more explicit by giving them their own definition boxes and a few results have been highlighted by labeling them as theorems.
- All existing ancillaries have been updated.

Features

Clear Writing Style

The text is written is a simple, direct, conversational style. As much as possible, I have used "mathematical English" rather than relying excessively on mathematical notation. However, all proofs that are given are fully rigorous, and Appendix A contains an introduction to mathematical notation for those who wish to streamline their own writing. Concrete examples almost always precede theorems, which are then followed by further examples and applications. This flow—from specific to general and back again—is consistent throughout the book.

Key Concepts Introduced Early

Many students encounter difficulty in linear algebra when the course moves from the computational (solving systems of linear equations, manipulating vectors and matrices) to the theoretical (spanning sets, linear independence, subspaces, basis, and dimension). This book introduces all of the key concepts of linear algebra early, in a concrete setting, before revisiting them in full generality. Vector concepts such as dot product, length, orthogonality, and projection are first discussed in Chapter 1 in the concrete setting of $\mathbb{R}^2$ and $\mathbb{R}^3$ before the more general notions of inner product, norm, and orthogonal projection appear in Chapters 5 and 7. Similarly, spanning sets and linear independence are given a concrete treatment in Chapter 2 prior to their generalization to vector spaces in Chapter 6. The fundamental concepts of subspace, basis, and dimension appear first in Chapter 3 when the row, column, and null spaces of a matrix are introduced; it is not until Chapter 6 that these ideas are given a general treatment. In Chapter 4, eigenvalues and eigenvectors are introduced and explored for 2×2 matrices before their $n \times n$ counterparts appear. By the beginning of Chapter 4, all of the key concepts of linear algebra have been introduced, with concrete, computational examples to support them. When these ideas appear in full generality later in the book, students have had time to get used to them and, hence, are not so intimidated by them.

Emphasis on Vectors and Geometry

In keeping with the philosophy that linear algebra is primarily about vectors, this book stresses geometric intuition. Accordingly, the first chapter is about vectors, and it develops many concepts that will appear repeatedly throughout the text. Concepts such as orthogonality, projection, and linear combination are all found in Chapter 1, as is a comprehensive treatment of lines and planes in $\mathbb{R}^3$ that provides essential insight into the solution of systems of linear equations. This emphasis on vectors, geometry, and visualization is found throughout the text. Linear transformations are introduced as matrix transformations in Chapter 3, with many geometric examples, before general linear transformations are covered in Chapter 6. In Chapter 4, eigenvalues are introduced with "eigenpictures" as a visual aid. The proof of Perron's Theorem is given first heuristically and then formally, in both cases using a geometric argument. The geometry of linear dynamical systems reinforces and summarizes the material on eigenvalues and eigenvectors. In Chapter 5, orthogonal projections, orthogonal complements of subspaces, and the Gram-Schmidt Process are all presented in the concrete setting of $\mathbb{R}^3$ before being generalized to $\mathbb{R}^n$ and, in Chapter 7, to inner product spaces. The nature of the singular value decomposition is also explained informally in Chapter 7 via a geometric argument. Of the more than 300 figures in the text, over 200 are devoted to fostering a geometric understanding of linear algebra.

Explorations

See pages 1, 136, 427, 529

See pages 32, 286, 460, 515, 543, 547

See pages 83, 84, 85, 396, 398

The introduction to each chapter is a guided exploration (Section 0) in which students are invited to discover, individually or in groups, some aspect of the upcoming chapter. For example, "The Racetrack Game" introduces vectors, "Matrices in Action" introduces matrix multiplication and linear transformations, "Fibonacci in (Vector) Space" touches on vector space concepts, and "Taxicab Geometry" sets up generalized norms and distance functions. Additional explorations found throughout the book include applications of vectors and determinants to geometry, an investigation of 3×3 magic squares, a study of symmetry via the tilings of M. C. Escher, an introduction to complex linear algebra, and optimization problems using geometric inequalities. There are also explorations that introduce important numerical considerations and the analysis of algorithms. Having students do some of these explorations is one way of encouraging them to become active learners and to give them "ownership" over a small part of the course.

Applications

See pages 623, 641

See pages 121, 226, 356, 607, 626

The book contains an abundant selection of applications chosen from a broad range of disciplines, including mathematics, computer science, physics, chemistry, engineering, biology, business, economics, psychology, geography, and sociology. Noteworthy among these is a strong treatment of coding theory, from error-detecting codes (such as International Standard Book Numbers) to sophisticated error-correcting codes (such as the Reed-Muller code that was used to transmit satellite photos from space). Additionally, there are five "vignettes" that briefly showcase some very modern applications of linear algebra: the Global Positioning System (GPS), robotics, Internet search engines, digital image compression, and the Codabar System.

Examples and Exercises

There are over 400 examples in this book, most worked in greater detail than is customary in an introductory linear algebra textbook. This level of detail is in keeping with the philosophy that students should want (and be able) to read a textbook. Accordingly, it is not intended that all of these examples be covered in class; many can be assigned for individual or group study, possibly as part of a project. Most examples have at least one counterpart exercise so that students can try out the skills covered in the example before exploring generalizations.

There are over 2000 exercises, more than in most textbooks at a similar level. Answers to most of the computational odd-numbered exercises can be found in the back of the book. Instructors will find an abundance of exercises from which to select homework assignments. The exercises in each section are graduated, progressing from the routine to the challenging. Exercises range from those intended for hand computation to those requiring the use of a calculator or computer algebra system, and from theoretical and numerical exercises to conceptual exercises. Many of the examples and exercises use actual data compiled from real-world situations. For example, there are problems on modeling the growth of caribou and seal populations, radiocarbon dating

See pages 248, 359, 526, 588

of the Stonehenge monument, and predicting major league baseball players' salaries. Working such problems reinforces the fact that linear algebra is a valuable tool for modeling real-life problems.

Additional exercises appear in the form of a review after each chapter. In each set, there are 10 true/false questions designed to test conceptual understanding, followed by 19 computational and theoretical exercises that summarize the main concepts and techniques of that chapter.

Biographical Sketches and Etymological Notes

It is important that students learn something about the history of mathematics and come to see it as a social and cultural endeavor as well as a scientific one. Accordingly, the text contains short biographical sketches about many of the mathematicians who contributed to the development of linear algebra. I hope that these will help to put a human face on the subject and give students another way of relating to the material.

I have found that many students feel alienated from mathematics because the terminology makes no sense to them—it is simply a collection of words to be learned. To help overcome this problem, I have included short etymological notes that give the origins of many of the terms used in linear algebra. (For example, why do we use the word *normal* to refer to a vector that is perpendicular to a plane?)

See page 34

Margin Icons

The margins of the book contain several icons whose purpose is to alert the reader in various ways. Calculus is not a prerequisite for this book, but linear algebra has many interesting and important applications to calculus. The $\frac{dy}{dx}$ icon denotes an example or exercise that requires calculus. (This material can be omitted if not everyone in the class has had at least one semester of calculus. Alternatively, this material can be assigned as projects.) The $a + bi$ icon denotes an example or exercise involving complex numbers. (For students unfamiliar with complex numbers, Appendix C contains all the background material that is needed.) The CAS icon indicates that a computer algebra system (such as Maple, Mathematica, or MATLAB) or a calculator with matrix capabilities (such as almost any graphing calculator) is required—or at least very useful—for solving the example or exercise.

In an effort to help students learn how to read and use this textbook most effectively, I have noted various places where the reader is advised to pause. These may be places where a calculation is needed, part of a proof must be supplied, a claim should be verified, or some extra thought is required. The → icon appears in the margin at such places; the message is "Slow down. Get out your pencil. Think about this."

Technology

This book can be used successfully whether or not students have access to technology. However, calculators with matrix capabilities and computer algebra systems are now commonplace and, properly used, can enrich the learning experience as well as help with tedious calculations. In this text, I take the point of view that students need to master all of the basic techniques of linear algebra by solving by hand examples that are not too computationally difficult. Technology may then be used (in whole or in part) to solve subsequent examples and applications and to apply techniques that rely on earlier ones. For example, when systems of linear equations are first introduced, detailed solutions are provided; later, solutions are simply given, and the reader is expected to verify them. This is a good place to use some form of technology. Likewise, when applications use data that make hand calculation impractical, use technology. All of the numerical methods that are discussed depend on the use of technology.

With the aid of technology, students can explore linear algebra in some exciting ways and discover much for themselves. For example, if one of the coefficients of a linear system is replaced by a parameter, how much variability is there in the solutions? How does changing a single entry of a matrix affect its eigenvalues? This book is not a tutorial on technology, and in places where technology can be used, I have not specified a particular type of technology. The student companion website that accompanies this book offers an online appendix called *Technology Bytes* that gives instructions for solving a selection of examples from each chapter using Maple, Mathematica, and MATLAB. By imitating these examples, students can do further calculations and explorations using whichever CAS they have and exploit the power of these systems to help with the exercises throughout the book, particularly those marked with the CAS icon. The website also contains data sets and computer code in Maple, Mathematica, and MATLAB formats keyed to many exercises and examples in the text. Students and instructors can import these directly into their CAS to save typing and eliminate errors.

Finite and Numerical Linear Algebra

The text covers two aspects of linear algebra that are scarcely ever mentioned together: finite linear algebra and numerical linear algebra. By introducing modular arithmetic early, I have been able to make finite linear algebra (more properly, "linear algebra over finite fields," although I do not use that phrase) a recurring theme throughout the book. This approach provides access to the material on coding theory in Chapter 8 (online). There is also an application to finite linear games in Section 2.4 that students really enjoy. In addition to being exposed to the applications of finite linear algebra, mathematics majors will benefit from seeing the material on finite fields, because they are likely to encounter it in such courses as discrete mathematics, abstract algebra, and number theory.

All students should be aware that in practice, it is impossible to arrive at exact solutions of large-scale problems in linear algebra. Exposure to some of the techniques of numerical linear algebra will provide an indication of how to obtain highly accurate approximate solutions. Some of the numerical topics included in the book are roundoff error and partial pivoting, iterative methods for solving linear systems and computing eigenvalues, the *LU* and *QR* factorizations, matrix norms and condition numbers, least squares approximation, and the singular value decomposition. The inclusion of numerical linear algebra also brings up some interesting and important issues that are completely absent from the *theory* of linear algebra, such as pivoting strategies, the condition of a linear system, and the convergence of iterative methods. This book not only raises these questions but also shows how one might approach them. Gerschgorin disks, matrix norms, and the singular values of a matrix, discussed in Chapters 4 and 7, are useful in this regard.

See pages 83, 84, 124, 180, 311, 392, 555, 561, 568, 590

See pages 319, 563, 600

Appendices

Appendix A contains an overview of mathematical notation and methods of proof, and Appendix B discusses mathematical induction. All students will benefit from these sections, but those with a mathematically oriented major may wish to pay particular attention to them. Some of the examples in these appendices are uncommon (for instance, Example B.6 in Appendix B) and underscore the power of the methods. Appendix C is an introduction to complex numbers. For students familiar with these results, this appendix can serve as a useful reference; for others, this section contains everything they need to know for those parts of the text that use complex numbers. Appendix D is about polynomials. I have found that many students require a refresher about these facts. Most students will be unfamiliar with Descartes's Rule of Signs; it is used in Chapter 4 to explain the behavior of the eigenvalues of Leslie matrices. Exercises to accompany the four appendices can be found on the book's website.

Short answers to most of the odd-numbered computational exercises are given at the end of the book. Exercise sets to accompany Appendixes A, B, C, and D are available on the companion website, along with their odd-numbered answers.

Ancillaries

For Instructors

Enhanced WebAssign®
Printed Access Card: 978-1-285-85829-6
Online Access Code: 978-1-285-85827-2
Exclusively from Cengage Learning, Enhanced WebAssign combines the exceptional mathematics content that you know and love with the most powerful online homework solution, WebAssign. Enhanced WebAssign engages students with immediate feedback, rich tutorial content, and interactive, fully customizable eBooks (YouBook), helping students to develop a deeper conceptual understanding of their subject matter. Flexible assignment options give instructors the ability to release assignments conditionally based on students' prerequisite assignment scores. Visit us at **www.cengage.com/ewa** to learn more.

Cengage Learning Testing Powered by Cognero
Cengage Learning Testing Powered by Cognero is a flexible, online system that allows you to author, edit, and manage test bank content from multiple Cengage Learning solutions; create multiple test versions in an instant; and deliver tests from your LMS, your classroom, or wherever you want.

Complete Solutions Manual
The Complete Solutions Manual provides detailed solutions to all exercises in the text, including *Exploration* and *Chapter Review* exercises. The Complete Solutions Manual is available online.

Instructor's Guide
This online guide enhances the text with valuable teaching resources such as group work projects, teaching tips, interesting exam questions, examples and extra

material for lectures, and other items designed to reduce the instructor's preparation time and make linear algebra class an exciting and interactive experience. For each section of the text, the *Instructor's Guide* includes suggested time and emphasis, points to stress, questions for discussion, lecture materials and examples, technology tips, student projects, group work with solutions, sample assignments, and suggested test questions.

Solution Builder
www.cengage.com/solutionbuilder
Solution Builder provides full instructor solutions to all exercises in the text, including those in the explorations and chapter reviews, in a convenient online format. Solution Builder allows instructors to create customized, secure PDF printouts of solutions matched exactly to the exercises assigned for class.

*Access Cognero and additional instructor resources online at **login.cengage.com**.

For Students

Student Solutions Manual (ISBN-13: 978-1-285-84195-3)
The Student Solutions Manual and Study Guide includes detailed solutions to all odd-numbered exercises and selected even-numbered exercises; section and chapter summaries of symbols, definitions, and theorems; and study tips and hints. Complex exercises are explored through a question-and-answer format designed to deepen understanding. Challenging and entertaining problems that further explore selected exercises are also included.

Enhanced WebAssign® ENHANCED WebAssign
Printed Access Card: 978-1-285-85829-6
Online Access Code: 978-1-285-85827-2
Enhanced WebAssign (assigned by the instructor) provides you with instant feedback on homework assignments. This online homework system is easy to use and includes helpful links to textbook sections, video examples, and problem-specific tutorials.

CengageBrain.com
To access additional course materials and companion resources, please visit www.cengagebrain.com. At the CengageBrain.com home page, search for the ISBN of your title (from the back cover of your book) using the search box at the top of the page. This will take you to the product page where free companion resources can be found.

Acknowledgments

The reviewers of the previous edition of this text contributed valuable and often insightful comments about the book. I am grateful for the time each of them took to do this. Their judgement and helpful suggestions have contributed greatly to the development and success of this book, and I would like to thank them personally:
Jamey Bass, City College of San Francisco; Olga Brezhneva, Miami University; Karen Clark, The College of New Jersey; Marek Elzanowski, Portland State University; Christopher Francisco, Oklahoma State University; Brian Jue, California State University, Stanislaus; Alexander Kheyfits, Bronx Community College/CUNY; Henry Krieger, Harvey Mudd College; Rosanna Pearlstein, Michigan State

University; William Sullivan, Portland State University; Matthias Weber, Indiana University.

I am indebted to a great many people who have, over the years, influenced my views about linear algebra and the teaching of mathematics in general. First, I would like to thank collectively the participants in the education and special linear algebra sessions at meetings of the Mathematical Association of America and the Canadian Mathematical Society. I have also learned much from participation in the Canadian Mathematics Education Study Group and the Canadian Mathematics Education Forum.

I especially want to thank Ed Barbeau, Bill Higginson, Richard Hoshino, John Grant McLoughlin, Eric Muller, Morris Orzech, Bill Ralph, Pat Rogers, Peter Taylor, and Walter Whiteley, whose advice and inspiration contributed greatly to the philosophy and style of this book. My gratitude as well to Robert Rogers, who developed the student and instructor solutions, as well as the excellent study guide content. Special thanks go to Jim Stewart for his ongoing support and advice. Joe Rotman and his lovely book *A First Course in Abstract Algebra* inspired the etymological notes in this book, and I relied heavily on Steven Schwartzman's *The Words of Mathematics* when compiling these notes. I thank Art Benjamin for introducing me to the Codabar system and Joe Grcar for clarifying aspects of the history of Gaussian elimination. My colleagues Marcus Pivato and Reem Yassawi provided useful information about dynamical systems. As always, I am grateful to my students for asking good questions and providing me with the feedback necessary to becoming a better teacher.

I sincerely thank all of the people who have been involved in the production of this book. Jitendra Kumar and the team at MPS Limited did an amazing job producing the fourth edition. I thank Christine Sabooni for doing a thorough copyedit. Most of all, it has been a delight to work with the entire editorial, marketing, and production teams at Cengage Learning: Richard Stratton, Molly Taylor, Laura Wheel, Cynthia Ashton, Danielle Hallock, Andrew Coppola, Alison Eigel Zade, and Janay Pryor. They offered sound advice about changes and additions, provided assistance when I needed it, but let me write the book I wanted to write. I am fortunate to have worked with them, as well as the staffs on the first through third editions.

As always, I thank my family for their love, support, and understanding. Without them, this book would not have been possible.

David Poole
dpoole@trentu.ca

To the Instructor

"Would you tell me, please,
which way I ought to go from here?"
"That depends a good deal on where
you want to get to," said the Cat.

—Lewis Carroll
Alice's Adventures in
Wonderland, 1865

This text was written with flexibility in mind. It is intended for use in a one- or two-semester course with 36 lectures per semester. The range of topics and applications makes it suitable for a variety of audiences and types of courses. However, there is more material in the book than can be covered in class, even in a two-semester course. After the following overview of the text are some brief suggestions for ways to use the book.

An Overview of the Text

Chapter 1: Vectors

See page 1

The racetrack game in Section 1.0 serves to introduce vectors in an informal way. (It's also quite a lot of fun to play!) Vectors are then formally introduced from both algebraic and geometric points of view. The operations of addition and scalar multiplication and their properties are first developed in the concrete settings of $\mathbb{R}^2$ and $\mathbb{R}^3$ before being generalized to $\mathbb{R}^n$. Modular arithmetic and finite linear algebra are also introduced. Section 1.2 defines the dot product of vectors and the related notions of length, angle, and orthogonality. The very important concept of (orthogonal) projection is developed here; it will reappear in Chapters 5 and 7. The exploration "Vectors and Geometry" shows how vector methods can be used to prove certain results in Euclidean geometry. Section 1.3 is a basic but thorough introduction to lines and planes in $\mathbb{R}^2$ and $\mathbb{R}^3$. This section is crucial for understanding the geometric significance of the solution of linear systems in Chapter 2. Note that the cross product of vectors in $\mathbb{R}^3$ is left as an exploration. The chapter concludes with an application to force vectors.

See page 32

See page 48

Chapter 2: Systems of Linear Equations

See page 57

The introduction to this chapter serves to illustrate that there is more than one way to think of the solution to a system of linear equations. Sections 2.1 and 2.2 develop the main computational tool for solving linear systems: row reduction of matrices (Gaussian and Gauss-Jordan elimination). Nearly all subsequent computational methods in the book depend on this. The Rank Theorem appears here for the first time; it shows up again, in more generality, in Chapters 3, 5, and 6. Section 2.3 is very important; it introduces the fundamental notions of spanning sets and linear independence of vectors. Do not rush through this material. Section 2.4 contains six applications from which instructors can choose depending on the time available and the interests of the class. The vignette on the Global Positioning System provides another application that students will enjoy. The iterative methods in Section 2.5 will be optional for many courses but are essential for a course with an applied/numerical focus. The three explorations in this chapter are related in that they all deal with aspects of the use of computers to solve linear systems. All students should at least be made aware of these issues.

See pages 72, 205, 386, 486

See page 121

See pages 83, 84, 85

Chapter 3: Matrices

This chapter contains some of the most important ideas in the book. It is a long chapter, but the early material can be covered fairly quickly, with extra time allowed for the crucial material in Section 3.5. Section 3.0 is an exploration that introduces the notion of a linear transformation: the idea that matrices are not just static objects but rather a type of function, transforming vectors into other vectors. All of the basic facts about matrices, matrix operations, and their properties are found in the first two sections. The material on partitioned matrices and the multiple representations of the matrix product is worth stressing, because it is used repeatedly in subsequent sections. The Fundamental Theorem of Invertible Matrices in Section 3.3 is very important and will appear several more times as new characterizations of invertibility are presented. Section 3.4 discusses the very important *LU* factorization of a matrix. If this topic is not covered in class, it is worth assigning as a project or discussing in a workshop. The point of Section 3.5 is to present many of the key concepts of linear algebra (subspace, basis, dimension, and rank) in the concrete setting of matrices before students see them in full generality. Although the examples in this section are all familiar, it is important that students get used to the new terminology and, in particular, understand what the notion of a basis means. The geometric treatment of linear transformations in Section 3.6 is intended to smooth the transition to general linear transformations in Chapter 6. The example of a projection is particularly important because it will reappear in Chapter 5. The vignette on robotic arms is a concrete demonstration of composition of linear (and affine) transformations. There are four applications from which to choose in Section 3.7. Either Markov chains or the Leslie model of population growth should be covered so that they can be used again in Chapter 4, where their behavior will be explained.

See page 136

See pages 172, 206, 296, 512, 605

See page 226

See pages 230, 239

Chapter 4: Eigenvalues and Eigenvectors

The introduction Section 4.0 presents an interesting dynamical system involving graphs. This exploration introduces the notion of an eigenvector and foreshadows the power method in Section 4.5. In keeping with the geometric emphasis of the book, Section 4.1 contains the novel feature of "eigenpictures" as a way of visualizing the eigenvectors of 2×2 matrices. Determinants appear in Section 4.2, motivated by their use in finding the characteristic polynomials of small matrices. This "crash

See page 253

See page 284

See page 286

course" in determinants contains all the essential material students need, including an optional but elementary proof of the Laplace Expansion Theorem. The vignette "Lewis Carroll's Condensation Method" presents a historically interesting, alternative method of calculating determinants that students may find appealing. The exploration "Geometric Applications of Determinants" makes a nice project that contains several interesting and useful results. (Alternatively, instructors who wish to give more detailed coverage to determinants may choose to cover some of this exploration in class.) The basic theory of eigenvalues and eigenvectors is found in Section 4.3, and Section 4.4 deals with the important topic of diagonalization. Example 4.29 on powers of matrices is worth covering in class. The power method and its variants, discussed in Section 4.5, are optional, but all students should be aware of the method, and an applied course should cover it in detail. Gerschgorin's Disk Theorem can be covered independently of the rest of Section 4.5. Markov chains and the Leslie model of population growth reappear in Section 4.6. Although the proof of Perron's Theorem is optional, the theorem itself (like the stronger Perron-Frobenius Theorem) should at least be mentioned because it explains *why* we should expect a unique positive eigenvalue with a corresponding positive eigenvector in these applications. The applications on recurrence relations and differential equations connect linear algebra to discrete mathematics and calculus, respectively. The matrix exponential can be covered if your class has a good calculus background. The final topic of discrete linear dynamical systems revisits and summarizes many of the ideas in Chapter 4, looking at them in a new, geometric light. Students will enjoy reading how eigenvectors can be used to help rank sports teams and websites. This vignette can easily be extended to a project or enrichment activity.

See pages 325, 330

See page 348

Chapter 5: Orthogonality

See page 366

The introductory exploration, "Shadows on a Wall," is mathematics at its best: it takes a known concept (projection of a vector onto another vector) and generalizes it in a useful way (projection of a vector onto a subspace—a plane), while uncovering some previously unobserved properties. Section 5.1 contains the basic results about orthogonal and orthonormal sets of vectors that will be used repeatedly from here on. In particular, orthogonal matrices should be stressed. In Section 5.2, two concepts from Chapter 1 are generalized: the orthogonal complement of a subspace and the orthogonal projection of a vector onto a subspace. The Orthogonal Decomposition Theorem is important here and helps to set up the Gram-Schmidt Process. Also note the quick proof of the Rank Theorem. The Gram-Schmidt Process is detailed in Section 5.3, along with the extremely important *QR* factorization. The two explorations that follow outline how the *QR* factorization is computed in practice and how it can be used to approximate eigenvalues. Section 5.4 on orthogonal diagonalization of (real) symmetric matrices is needed for the applications that follow. It also contains the Spectral Theorem, one of the highlights of the theory of linear algebra. The applications in Section 5.5 are quadratic forms and graphing quadratic equations. I always include at least the second of these in my course because it extends what students already know about conic sections.

See pages 396, 398

See pages 408, 415

Chapter 6: Vector Spaces

See page 427

The Fibonacci sequence reappears in Section 6.0, although it is not important that students have seen it before (Section 4.6). The purpose of this exploration is to show that familiar vector space concepts (Section 3.5) can be used fruitfully in a new setting. Because all of the main ideas of vector spaces have already been introduced in Chapters 1–3, students should find Sections 6.1 and 6.2 fairly familiar. The emphasis here should be on using the vector space axioms to prove properties rather than relying on computational techniques. When discussing change of basis in Section 6.3, it is helpful to show students how to use the notation to remember how the construction works. Ultimately, the Gauss-Jordan method is the most efficient here. Sections 6.4 and 6.5 on linear transformations are important. The examples are related to previous results on matrices (and matrix transformations). In particular, it is important to stress that the kernel and range of a linear transformation generalize the null space and column space of a matrix. Section 6.6 puts forth the notion that (almost) all linear transformations are essentially matrix transformations. This builds on the information in Section 3.6, so students should not find it terribly surprising. However, the examples should be worked carefully. The connection between change of basis and similarity of matrices is noteworthy. The exploration "Tilings, Lattices, and the Crystallographic Restriction" is an impressive application of change of basis. The connection with the artwork of M. C. Escher makes it all the more interesting. The applications in Section 6.7 build on previous ones and can be included as time and interest permit.

See page 515

Chapter 7: Distance and Approximation

Section 7.0 opens with the entertaining "Taxicab Geometry" exploration. Its purpose is to set up the material on generalized norms and distance functions (metrics) that follows. Inner product spaces are discussed in Section 7.1; the emphasis here should be on the examples and using the axioms. The exploration "Vectors and Matrices with Complex Entries" shows how the concepts of dot product, symmetric matrix, orthogonal matrix, and orthogonal diagonalization can be extended from real to complex vector spaces. The following exploration, "Geometric Inequalities and Optimization Problems," is one that students typically enjoy. (They will have fun seeing how many "calculus" problems can be solved without using calculus at all!) Section 7.2 covers generalized vector and matrix norms and shows how the condition number of a matrix is related to the notion of ill-conditioned linear systems explored in Chapter 2. Least squares approximation (Section 7.3) is an important application of linear algebra in many other disciplines. The Best Approximation Theorem and the Least Squares Theorem are important, but their proofs are intuitively clear. Spend time here on the examples—a few should suffice. Section 7.4 presents the singular value decomposition, one of the most impressive applications of linear algebra. If your course gets this far, you will be amply rewarded. Not only does the SVD tie together many notions discussed previously; it also affords some new (and quite powerful) applications. If a CAS is available, the vignette on digital image compression is worth presenting; it is a visually impressive display of the power of linear algebra and a fitting culmination to the course. The further applications in Section 7.5 can be chosen according to the time available and the interests of the class.

See page 529

See page 543

See page 547

See page 607

Chapter 8: Codes

This online chapter contains applications of linear algebra to the theory of codes. Section 8.1 begins with a discussion of how vectors can be used to design

See page 626

error-detecting codes such as the familiar Universal Product Code (UPC) and International Standard Book Number (ISBN). This topic only requires knowledge of Chapter 1. The vignette on the Codabar system used in credit and bank cards is an excellent classroom presentation that can even be used to introduce Section 8.1. Once students are familiar with matrix operations, Section 8.2 describes how codes can be designed to correct as well as detect errors. The Hamming codes introduced here are perhaps the most famous examples of such error-correcting codes. Dual codes, discussed in Section 8.3, are an important way of constructing new codes from old ones. The notion of orthogonal complement, introduced in Chapter 5, is the prerequisite concept here. The most important, and most widely used, class of codes is the class of linear codes that is defined in Section 8.4. The notions of subspace, basis, and dimension are key here. The powerful Reed-Muller codes used by NASA spacecraft are important examples of linear codes. Our discussion of codes concludes in Section 8.5 with the definition of the minimum distance of a code and the role it plays in determining the error-correcting capability of the code.

How to Use the Book

Students find the book easy to read, so I usually have them read a section before I cover the material in class. That way, I can spend class time highlighting the most important concepts, dealing with topics students find difficult, working examples, and discussing applications. I do not attempt to cover all of the material from the assigned reading in class. This approach enables me to keep the pace of the course fairly brisk, slowing down for those sections that students typically find challenging.

In a two-semester course, it is possible to cover the entire book, including a reasonable selection of applications. For extra flexibility, you might omit some of the topics (for example, give only a brief treatment of numerical linear algebra), thereby freeing up time for more in-depth coverage of the remaining topics, more applications, or some of the explorations. In an honors mathematics course that emphasizes proofs, much of the material in Chapters 1–3 can be covered quickly. Chapter 6 can then be covered in conjunction with Sections 3.5 and 3.6, and Chapter 7 can be integrated into Chapter 5. I would be sure to assign the explorations in Chapters 1, 4, 6, and 7 for such a class.

For a one-semester course, the nature of the course and the audience will determine which topics to include. Three possible courses are described below and on the following page. The basic course, described first, has fewer than 36 hours suggested, allowing time for extra topics, in-class review, and tests. The other two courses build on the basic course but are still quite flexible.

A Basic Course

A course designed for mathematics majors and students from other disciplines is outlined on the next page. This course does not mention general vector spaces at all (all concepts are treated in a concrete setting) and is very light on proofs. Still, it is a thorough introduction to linear algebra.

Section	Number of Lectures	Section	Number of Lectures
1.1	1	**3.6**	1–2
1.2	1–1.5	**4.1**	1
1.3	1–1.5	**4.2**	2
2.1	0.5–1	**4.3**	1
2.2	1–2	**4.4**	1–2
2.3	1–2	**5.1**	1–1.5
3.1	1–2	**5.2**	1–1.5
3.2	1	**5.3**	0.5
3.3	2	**5.4**	1
3.5	2	**7.3**	2

Total: 23–30 lectures

Because the students in a course such as this one represent a wide variety of disciplines, I would suggest using much of the remaining lecture time for applications. In my course, I do code vectors in Section 8.1, which students really seem to like, and at least one application from each of Chapters 2–5. Other applications can be assigned as projects, along with as many of the explorations as desired. There is also sufficient lecture time available to cover some of the theory in detail.

A Course with a Computational Emphasis

For a course with a computational emphasis, the basic course outlined on the previous page can be supplemented with the sections of the text dealing with numerical linear algebra. In such a course, I would cover part or all of Sections 2.5, 3.4, 4.5, 5.3, 7.2, and 7.4, ending with the singular value decomposition. The explorations in Chapters 2 and 5 are particularly well suited to such a course, as are almost any of the applications.

A Course for Students Who Have Already Studied Some Linear Algebra

Some courses will be aimed at students who have already encountered the basic principles of linear algebra in other courses. For example, a college algebra course will often include an introduction to systems of linear equations, matrices, and determinants; a multivariable calculus course will almost certainly contain material on vectors, lines, and planes. For students who have seen such topics already, much early material can be omitted and replaced with a quick review. Depending on the background of the class, it may be possible to skim over the material in the basic course up to Section 3.3 in about six lectures. If the class has a significant number of mathematics majors (and especially if this is the only linear algebra course they will take), I would be sure to cover Sections 6.1–6.5, 7.1, and 7.4 and as many applications as time permits. If the course has science majors (but not mathematics majors), I would cover Sections 6.1 and 7.1 and a broader selection of applications, being sure to include the material on differential equations and approximation of functions. If computer science students or engineers are prominently represented, I would try to do as much of the material on codes and numerical linear algebra as I could.

There are many other types of courses that can successfully use this text. I hope that you find it useful for your course and that you enjoy using it.

To the Student

"Where shall I begin, please your Majesty?" he asked.
"Begin at the beginning," the King said, gravely, "and go on till you come to the end: then stop."

—Lewis Carroll
Alice's Adventures in Wonderland, 1865

Linear algebra is an exciting subject. It is full of interesting results, applications to other disciplines, and connections to other areas of mathematics. The *Student Solutions Manual and Study Guide* contains detailed advice on how best to use this book; following are some general suggestions.

Linear algebra has several sides: There are *computational techniques, concepts,* and *applications*. One of the goals of this book is to help you master all of these facets of the subject and to see the interplay among them. Consequently, it is important that you read and understand each section of the text before you attempt the exercises in that section. If you read only examples that are related to exercises that have been assigned as homework, you will miss much. Make sure you understand the definitions of terms and the meaning of theorems. Don't worry if you have to read something more than once before you understand it. Have a pencil and calculator with you as you read. Stop to work out examples for yourself or to fill in missing calculations. The icon in the margin indicates a place where you should pause and think over what you have read so far.

Answers to most odd-numbered computational exercises are in the back of the book. Resist the temptation to look up an answer before you have completed a question. And remember that even if your answer differs from the one in the back, you may still be right; there is more than one correct way to express some of the solutions. For example, a value of $1/\sqrt{2}$ can also be expressed as $\sqrt{2}/2$ and the set of all scalar multiples of the vector $\begin{bmatrix} 3 \\ 1/2 \end{bmatrix}$ is the same as the set of all scalar multiples of $\begin{bmatrix} 6 \\ 1 \end{bmatrix}$.

As you encounter new concepts, try to relate them to examples that you know. Write out proofs and solutions to exercises in a logical, connected way, using complete sentences. Read back what you have written to see whether it makes sense. Better yet, if you can, have a friend in the class read what you have written. If it doesn't make sense to another person, chances are that it doesn't make sense, period.

You will find that a calculator with matrix capabilities or a computer algebra system is useful. These tools can help you to check your own hand calculations and are indispensable for some problems involving tedious computations. Technology also enables you to explore aspects of linear algebra on your own. You can play "what if?" games: What if I change one of the entries in this vector? What if this matrix is of a different size? Can I force the solution to be what I would like it to be by changing something? To signal places in the text or exercises where the use of technology is recommended, I have placed the icon CAS in the margin. The companion website that accompanies this book contains computer code working out selected exercises from the book using Maple, Mathematica, and MATLAB, as well as *Technology Bytes*, an appendix providing much additional advice about the use of technology in linear algebra.

You are about to embark on a journey through linear algebra. Think of this book as your travel guide. Are you ready? Let's go!

1 Vectors

Here they come pouring out of the blue.
Little arrows for me and for you.
—Albert Hammond and
Mike Hazelwood
Little Arrows
Dutchess Music/BMI, 1968

1.0 Introduction: The Racetrack Game

Many measurable quantities, such as length, area, volume, mass, and temperature, can be completely described by specifying their magnitude. Other quantities, such as velocity, force, and acceleration, require both a magnitude and a direction for their description. These quantities are *vectors*. For example, wind velocity is a vector consisting of wind speed and direction, such as 10 km/h southwest. Geometrically, vectors are often represented as arrows or directed line segments.

Although the idea of a vector was introduced in the 19th century, its usefulness in applications, particularly those in the physical sciences, was not realized until the 20th century. More recently, vectors have found applications in computer science, statistics, economics, and the life and social sciences. We will consider some of these many applications throughout this book.

This chapter introduces vectors and begins to consider some of their geometric and algebraic properties. We begin, though, with a simple game that introduces some of the key ideas. [You may even wish to play it with a friend during those (very rare!) dull moments in linear algebra class.]

The game is played on graph paper. A track, with a starting line and a finish line, is drawn on the paper. The track can be of any length and shape, so long as it is wide enough to accommodate all of the players. For this example, we will have two players (let's call them Ann and Bert) who use different colored pens to represent their cars or bicycles or whatever they are going to race around the track. (Let's think of Ann and Bert as cyclists.)

Ann and Bert each begin by drawing a dot on the starting line at a grid point on the graph paper. They take turns moving to a new grid point, subject to the following rules:

1. Each new grid point and the line segment connecting it to the previous grid point must lie entirely within the track.
2. No two players may occupy the same grid point on the same turn. (This is the "no collisions" rule.)
3. Each new move is related to the previous move as follows: If a player moves a units horizontally and b units vertically on one move, then on the next move he or she must move between $a - 1$ and $a + 1$ units horizontally and between

The Irish mathematician William Rowan Hamilton (1805–1865) used vector concepts in his study of complex numbers and their generalization, the quaternions.

$b - 1$ and $b + 1$ units vertically. In other words, if the second move is c units horizontally and d units vertically, then $|a - c| \leq 1$ and $|b - d| \leq 1$. (This is the "acceleration/deceleration" rule.) Note that this rule forces the first move to be 1 unit vertically and/or 1 unit horizontally.

A player who collides with another player or leaves the track is eliminated. The winner is the first player to cross the finish line. If more than one player crosses the finish line on the same turn, the one who goes farthest past the finish line is the winner.

In the sample game shown in Figure 1.1, Ann was the winner. Bert accelerated too quickly and had difficulty negotiating the turn at the top of the track.

To understand rule 3, consider Ann's third and fourth moves. On her third move, she went 1 unit horizontally and 3 units vertically. On her fourth move, her options were to move 0 to 2 units horizontally and 2 to 4 units vertically. (Notice that some of these combinations would have placed her outside the track.) She chose to move 2 units in each direction.

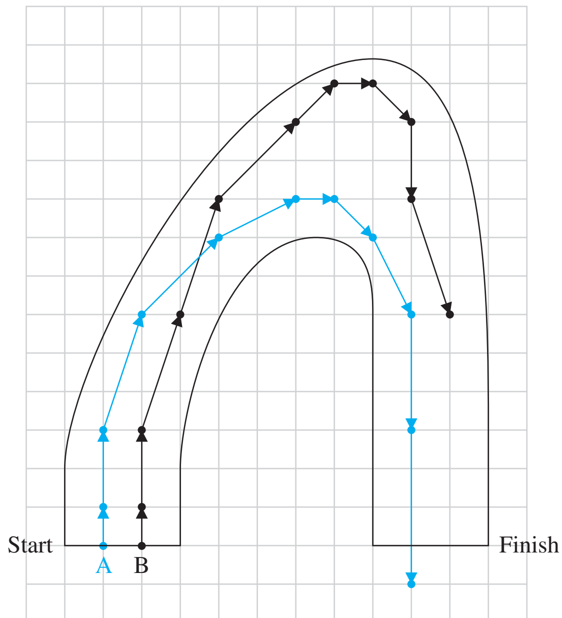


Figure 1.1
A sample game of racetrack

Problem 1 Play a few games of racetrack.

Problem 2 Is it possible for Bert to win this race by choosing a different sequence of moves?

Problem 3 Use the notation $[a, b]$ to denote a move that is a units horizontally and b units vertically. (Either a or b or both may be negative.) If move $[3, 4]$ has just been made, draw on graph paper all the grid points that could possibly be reached on the next move.

Problem 4 What is the net effect of two successive moves? In other words, if you move $[a, b]$ and then $[c, d]$, how far horizontally and vertically will you have moved altogether?

Problem 5 Write out Ann's sequence of moves using the $[a, b]$ notation. Suppose she begins at the origin $(0, 0)$ on the coordinate axes. Explain how you can find the coordinates of the grid point corresponding to each of her moves *without looking at the graph paper*. If the axes were drawn differently, so that Ann's starting point was not the origin but the point $(2, 3)$, what would the coordinates of her final point be?

Although simple, this game introduces several ideas that will be useful in our study of vectors. The next three sections consider vectors from geometric and algebraic viewpoints, beginning, as in the racetrack game, in the plane.

1.1 The Geometry and Algebra of Vectors

Vectors in the Plane

The Cartesian plane is named after the French philosopher and mathematician René Descartes (1596–1650), whose introduction of coordinates allowed *geometric* problems to be handled using *algebraic* techniques.

The word *vector* comes from the Latin root meaning "to carry." A vector is formed when a point is displaced—or "carried off"—a given distance in a given direction. Viewed another way, vectors "carry" two pieces of information: their length and their direction.

When writing vectors by hand, it is difficult to indicate boldface. Some people prefer to write $\vec{v}$ for the vector denoted in print by $\mathbf{v}$, but in most cases it is fine to use an ordinary lowercase v. It will usually be clear from the context when the letter denotes a vector.

The word *component* is derived from the Latin words *co*, meaning "together with," and *ponere*, meaning "to put." Thus, a vector is "put together" out of its components.

We begin by considering the Cartesian plane with the familiar x- and y-axes. A ***vector*** is a *directed line segment* that corresponds to a *displacement* from one point A to another point B; see Figure 1.2.

The vector from A to B is denoted by $\overrightarrow{AB}$; the point A is called its ***initial point,*** or ***tail,*** and the point B is called its ***terminal point,*** or ***head***. Often, a vector is simply denoted by a single boldface, lowercase letter such as $\mathbf{v}$.

The set of all points in the plane corresponds to the set of all vectors whose tails are at the origin O. To each point A, there corresponds the vector $\mathbf{a} = \overrightarrow{OA}$; to each vector $\mathbf{a}$ with tail at O, there corresponds its head A. (Vectors of this form are sometimes called *position vectors*.)

It is natural to represent such vectors using coordinates. For example, in Figure 1.3, $A = (3, 2)$ and we write the vector $\mathbf{a} = \overrightarrow{OA} = [3, 2]$ using square brackets. Similarly, the other vectors in Figure 1.3 are

$$\mathbf{b} = [-1, 3] \quad \text{and} \quad \mathbf{c} = [2, -1]$$

The individual coordinates (3 and 2 in the case of $\mathbf{a}$) are called the ***components*** of the vector. A vector is sometimes said to be an *ordered pair* of real numbers. The order is important since, for example, $[3, 2] \neq [2, 3]$. In general, two vectors are equal if and only if their corresponding components are equal. Thus, $[x, y] = [1, 5]$ implies that $x = 1$ and $y = 5$.

It is frequently convenient to use ***column vectors*** instead of (or in addition to) ***row vectors.*** Another representation of $[3, 2]$ is $\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. (The important point is that the

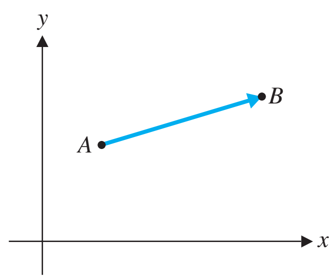


Figure 1.2

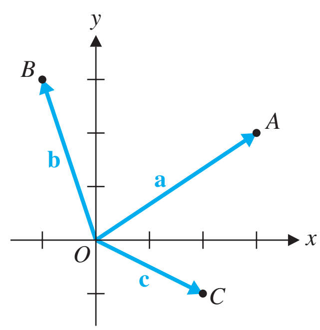


Figure 1.3

components are *ordered*.) In later chapters, you will see that column vectors are somewhat better from a computational point of view; for now, try to get used to both representations.

It may occur to you that we cannot really draw the vector $[0, 0] = \overrightarrow{OO}$ from the origin to itself. Nevertheless, it is a perfectly good vector and has a special name: the ***zero vector***. The zero vector is denoted by $\mathbf{0}$.

$\mathbb{R}^2$ is pronounced "r two."

The set of all vectors with two components is denoted by $\mathbb{R}^2$ (where $\mathbb{R}$ denotes the set of real numbers from which the components of vectors in $\mathbb{R}^2$ are chosen). Thus, $[-1, 3.5]$, $[\sqrt{2}, \pi]$, and $[\frac{5}{3}, 4]$ are all in $\mathbb{R}^2$.

Thinking back to the racetrack game, let's try to connect all of these ideas to vectors whose tails are not at the origin. The etymological origin of the word *vector* in the verb "to carry" provides a clue. The vector $[3, 2]$ may be interpreted as follows: Starting at the origin O, travel 3 units to the right, then 2 units up, finishing at P. The same displacement may be applied with other initial points. Figure 1.4 shows two equivalent displacements, represented by the vectors $\overrightarrow{AB}$ and $\overrightarrow{CD}$.

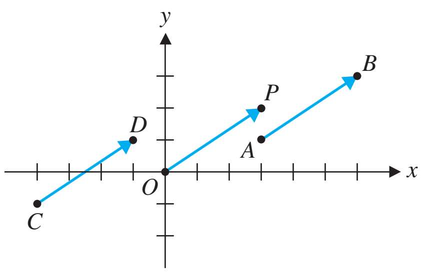


Figure 1.4

When vectors are referred to by their coordinates, they are being considered *analytically*.

We define two vectors as *equal* if they have the same length and the same direction. Thus, $\overrightarrow{AB} = \overrightarrow{CD}$ in Figure 1.4. (Even though they have different initial and terminal points, they represent the same displacement.) Geometrically, two vectors are equal if one can be obtained by sliding (or *translating*) the other parallel to itself until the two vectors coincide. In terms of components, in Figure 1.4 we have $A = (3, 1)$ and $B = (6, 3)$. Notice that the vector $[3, 2]$ that records the displacement is just the difference of the respective components:

$$\overrightarrow{AB} = [3, 2] = [6 - 3, 3 - 1]$$

Similarly,

$$\overrightarrow{CD} = [-1 - (-4), 1 - (-1)] = [3, 2]$$

and thus $\overrightarrow{AB} = \overrightarrow{CD}$, as expected.

A vector such as $\overrightarrow{OP}$ with its tail at the origin is said to be in ***standard position***. The foregoing discussion shows that every vector can be drawn as a vector in standard position. Conversely, a vector in standard position can be redrawn (by translation) so that its tail is at any point in the plane.

Example 1.1

If $A = (-1, 2)$ and $B = (3, 4)$, find $\overrightarrow{AB}$ and redraw it (a) in standard position and (b) with its tail at the point $C = (2, -1)$.

Solution We compute $\overrightarrow{AB} \approx [3 - (-1), 4 - 2] = [4, 2]$. If $\overrightarrow{AB}$ is then translated to $\overrightarrow{CD}$, where $C = (2, -1)$, then we must have $D = (2 + 4, -1 + 2) = (6, 1)$. (See Figure 1.5.)

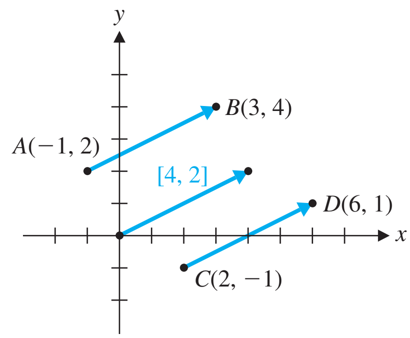


Figure 1.5

New Vectors from Old

As in the racetrack game, we often want to "follow" one vector by another. This leads to the notion of ***vector addition,*** the first basic vector operation.

If we follow **u** by **v**, we can visualize the total displacement as a third vector, denoted by **u** + **v**. In Figure 1.6, **u** = [1, 2] and **v** = [2, 2], so the net effect of following **u** by **v** is

$$[1 + 2, 2 + 2] = [3, 4]$$

which gives **u** + **v**. In general, if **u** = $[u_1, u_2]$ and **v** = $[v_1, v_2]$, then their ***sum*** **u** + **v** is the vector

$$\mathbf{u} + \mathbf{v} = [u_1 + v_1, u_2 + v_2]$$

It is helpful to visualize **u** + **v** geometrically. The following rule is the geometric version of the foregoing discussion.

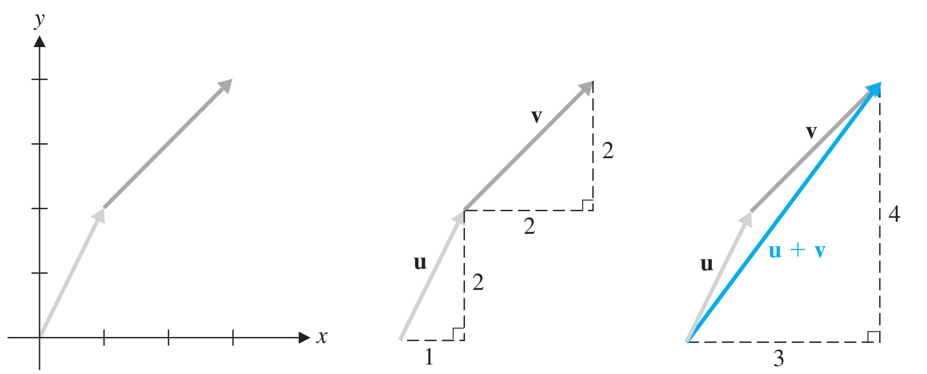


Figure 1.6
Vector addition

The Head-to-Tail Rule

Given vectors **u** and **v** in $\mathbb{R}^2$, translate **v** so that its tail coincides with the head of **u**. The ***sum*** **u** + **v** of **u** and **v** is the vector from the tail of **u** to the head of **v**. (See Figure 1.7.)

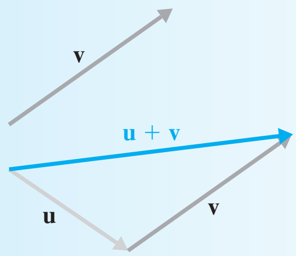


Figure 1.7
The head-to-tail rule

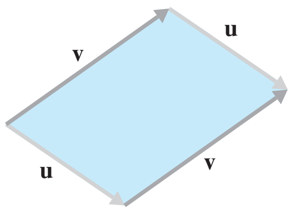


Figure 1.8
The parallelogram determined by **u** and **v**

By translating **u** and **v** parallel to themselves, we obtain a parallelogram, as shown in Figure 1.8. This parallelogram is called the *parallelogram determined by* **u** *and* **v**. It leads to an equivalent version of the head-to-tail rule for vectors in standard position.

The Parallelogram Rule

Given vectors **u** and **v** in $\mathbb{R}^2$ (in standard position), their ***sum*** **u** + **v** is the vector in standard position along the diagonal of the parallelogram determined by **u** and **v**. (See Figure 1.9.)

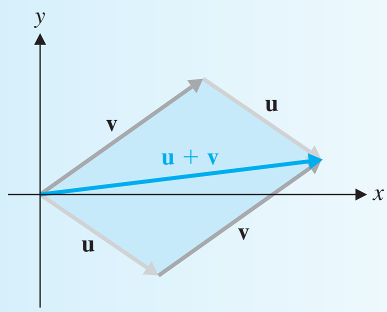


Figure 1.9
The parallelogram rule

Example 1.2

If **u** = [3, −1] and **v** = [1, 4], compute and draw **u** + **v**.

Solution We compute **u** + **v** = [3 + 1, −1 + 4] = [4, 3]. This vector is drawn using the head-to-tail rule in Figure 1.10(a) and using the parallelogram rule in Figure 1.10(b).

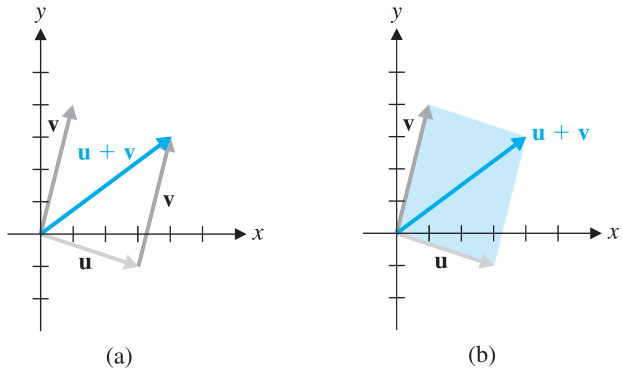


Figure 1.10

The second basic vector operation is ***scalar multiplication.*** Given a vector **v** and a real number c, the ***scalar multiple*** $c\mathbf{v}$ is the vector obtained by multiplying each component of **v** by c. For example, $3[-2, 4] = [-6, 12]$. In general,

$$c\mathbf{v} = c\,[v_1, v_2] = [cv_1, cv_2]$$

Geometrically, $c\mathbf{v}$ is a "scaled" version of **v**.

Example 1.3

If $\mathbf{v} = [-2, 4]$, compute and draw $2\mathbf{v}$, $\frac{1}{2}\mathbf{v}$, and $-2\mathbf{v}$.

Solution We calculate as follows:

$$\begin{aligned} 2\mathbf{v} &= [2(-2), 2(4)] = [-4, 8] \\ \tfrac{1}{2}\mathbf{v} &= [\tfrac{1}{2}(-2), \tfrac{1}{2}(4)] = [-1, 2] \\ -2\mathbf{v} &= [-2(-2), -2(4)] = [4, -8] \end{aligned}$$

These vectors are shown in Figure 1.11.

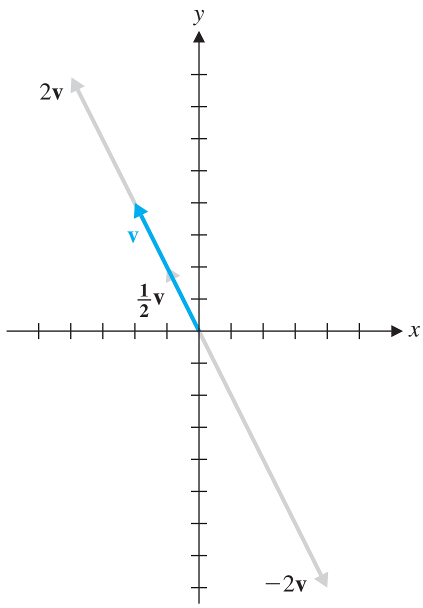


Figure 1.11

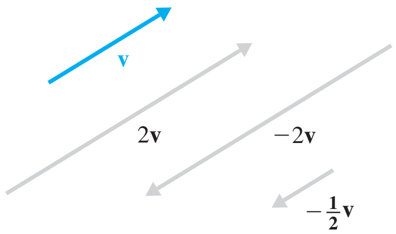


Figure 1.12

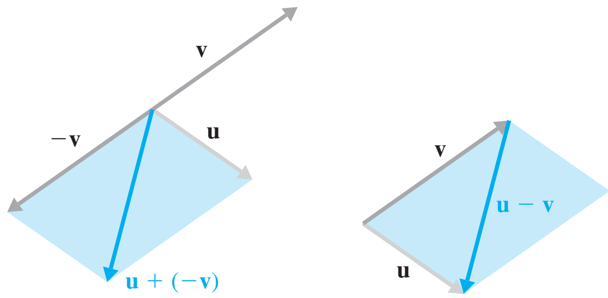


Figure 1.13
Vector subtraction

> The term *scalar* comes from the Latin word *scala*, meaning "ladder." The equally spaced rungs on a ladder suggest a scale, and in vector arithmetic, multiplication by a constant changes only the scale (or length) of a vector. Thus, constants became known as scalars.

Observe that $c\mathbf{v}$ has the same direction as $\mathbf{v}$ if $c > 0$ and the opposite direction if $c < 0$. We also see that $c\mathbf{v}$ is $|c|$ times as long as $\mathbf{v}$. For this reason, in the context of vectors, constants (i.e., real numbers) are referred to as ***scalars.*** As Figure 1.12 shows, when translation of vectors is taken into account, two vectors are scalar multiples of each other if and only if they are ***parallel.***

A special case of a scalar multiple is $(-1)\mathbf{v}$, which is written as $-\mathbf{v}$ and is called the ***negative of* v.** We can use it to define ***vector subtraction:*** The ***difference*** of $\mathbf{u}$ and $\mathbf{v}$ is the vector $\mathbf{u} - \mathbf{v}$ defined by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-\mathbf{v})$$

Figure 1.13 shows that $\mathbf{u} - \mathbf{v}$ corresponds to the "other" diagonal of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$.

Example 1.4

If $\mathbf{u} = [1, 2]$ and $\mathbf{v} = [-3, 1]$, then $\mathbf{u} - \mathbf{v} = [1 - (-3), 2 - 1] = [4, 1]$.

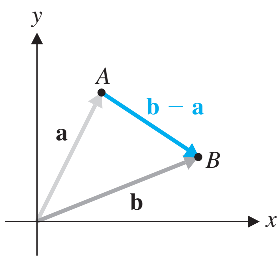


Figure 1.14

The definition of subtraction in Example 1.4 also agrees with the way we calculate a vector such as $\overrightarrow{AB}$. If the points A and B correspond to the vectors $\mathbf{a}$ and $\mathbf{b}$ in standard position, then $\overrightarrow{AB} = \mathbf{b} - \mathbf{a}$, as shown in Figure 1.14. [Observe that the head-to-tail rule applied to this diagram gives the equation $\mathbf{a} + (\mathbf{b} - \mathbf{a}) = \mathbf{b}$. If we had accidentally drawn $\mathbf{b} - \mathbf{a}$ with its head at A instead of at B, the diagram would have read $\mathbf{b} + (\mathbf{b} - \mathbf{a}) = \mathbf{a}$, which is clearly wrong! More will be said about algebraic expressions involving vectors later in this section.]

Vectors in $\mathbb{R}^3$

Everything we have just done extends easily to three dimensions. The set of all *ordered triples* of real numbers is denoted by $\mathbb{R}^3$. Points and vectors are located using three mutually perpendicular coordinate axes that meet at the origin O. A point such as $A = (1, 2, 3)$ can be located as follows: First travel 1 unit along the x-axis, then move 2 units parallel to the y-axis, and finally move 3 units parallel to the z-axis. The corresponding vector $\mathbf{a} = [1, 2, 3]$ is then $\overrightarrow{OA}$, as shown in Figure 1.15.

Another way to visualize vector $\mathbf{a}$ in $\mathbb{R}^3$ is to construct a box whose six sides are determined by the three coordinate planes (the xy-, xz-, and yz-planes) and by three planes through the point $(1, 2, 3)$ parallel to the coordinate planes. The vector $[1, 2, 3]$ then corresponds to the diagonal from the origin to the opposite corner of the box (see Figure 1.16).

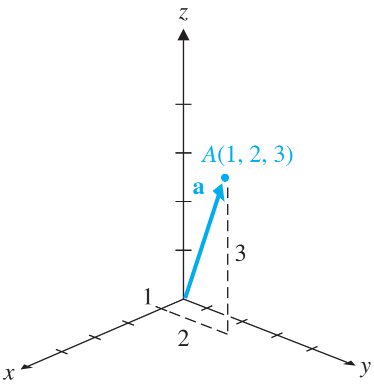


Figure 1.15

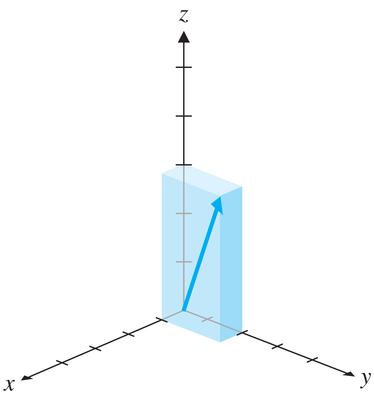


Figure 1.16

The "componentwise" definitions of vector addition and scalar multiplication are extended to $\mathbb{R}^3$ in an obvious way.

Vectors in $\mathbb{R}^n$

In general, we define $\mathbb{R}^n$ as the set of all *ordered n-tuples* of real numbers written as row or column vectors. Thus, a vector $\mathbf{v}$ in $\mathbb{R}^n$ is of the form

$$[v_1, v_2, \ldots, v_n] \quad \text{or} \quad \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

The individual entries of $\mathbf{v}$ are its components; v_i is called the ith component.

We extend the definitions of vector addition and scalar multiplication to $\mathbb{R}^n$ in the obvious way: If $\mathbf{u} = [u_1, u_2, \ldots, u_n]$ and $\mathbf{v} = [v_1, v_2, \ldots, v_n]$, the ith component of $\mathbf{u} + \mathbf{v}$ is $u_i + v_i$ and the ith component of $c\mathbf{v}$ is just cv_i.

Since in $\mathbb{R}^n$ we can no longer draw pictures of vectors, it is important to be able to calculate with vectors. We must be careful not to assume that vector arithmetic will be similar to the arithmetic of real numbers. Often it is, and the algebraic calculations we do with vectors are similar to those we would do with scalars. But, in later sections, we will encounter situations where vector algebra is quite *unlike* our previous experience with real numbers. So it is important to verify any algebraic properties before attempting to use them.

One such property is *commutativity* of addition: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ for vectors $\mathbf{u}$ and $\mathbf{v}$. This is certainly true in $\mathbb{R}^2$. Geometrically, the head-to-tail rule shows that both $\mathbf{u} + \mathbf{v}$ and $\mathbf{v} + \mathbf{u}$ are the main diagonals of the parallelogram determined by $\mathbf{u}$ and $\mathbf{v}$. (The parallelogram rule also reflects this symmetry; see Figure 1.17.)

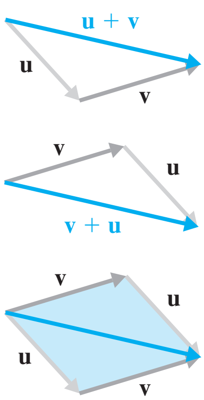


Figure 1.17
$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$

Note that Figure 1.17 is simply an illustration of the property $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$. It is not a proof, since it does not cover every possible case. For example, we must also include the cases where $\mathbf{u} = \mathbf{v}$, $\mathbf{u} = -\mathbf{v}$, and $\mathbf{u} = \mathbf{0}$. (What would diagrams for these cases look like?) For this reason, an algebraic proof is needed. However, it is just as easy to give a proof that is valid in $\mathbb{R}^n$ as to give one that is valid in $\mathbb{R}^2$.

The following theorem summarizes the algebraic properties of vector addition and scalar multiplication in $\mathbb{R}^n$. The proofs follow from the corresponding properties of real numbers.

Theorem 1.1 Algebraic Properties of Vectors in $\mathbb{R}^n$

Let **u, v,** and **w** be vectors in $\mathbb{R}^n$ and let c and d be scalars. Then

a. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ — Commutativity
b. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ — Associativity
c. $\mathbf{u} + \mathbf{0} = \mathbf{u}$
d. $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$
e. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ — Distributivity
f. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ — Distributivity
g. $c(d\mathbf{u}) = (cd)\mathbf{u}$
h. $1\mathbf{u} = \mathbf{u}$

The word *theorem* is derived from the Greek word *theorema,* which in turn comes from a word meaning "to look at." Thus, a theorem is based on the insights we have when we look at examples and extract from them properties that we try to prove hold in general. Similarly, when we understand something in mathematics—the proof of a theorem, for example—we often say, "I see."

Remarks

- Properties (c) and (d) together with the commutativity property (a) imply that $\mathbf{0} + \mathbf{u} = \mathbf{u}$ and $-\mathbf{u} + \mathbf{u} = \mathbf{0}$ as well.
- If we read the distributivity properties (e) and (f) from right to left, they say that we can *factor* a common scalar or a common vector from a sum.

Proof We prove properties (a) and (b) and leave the proofs of the remaining properties as exercises. Let $\mathbf{u} = [u_1, u_2, \ldots, u_n]$, $\mathbf{v} = [v_1, v_2, \ldots, v_n]$, and $\mathbf{w} = [w_1, w_2, \ldots, w_n]$.

(a)
$$\begin{aligned}\mathbf{u} + \mathbf{v} &= [u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n]\\ &= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n]\\ &= [v_1 + u_1, v_2 + u_2, \ldots, v_n + u_n]\\ &= [v_1, v_2, \ldots, v_n] + [u_1, u_2, \ldots, u_n]\\ &= \mathbf{v} + \mathbf{u}\end{aligned}$$

The second and fourth equalities are by the definition of vector addition, and the third equality is by the commutativity of addition of real numbers.

(b) Figure 1.18 illustrates associativity in $\mathbb{R}^2$. Algebraically, we have

$$\begin{aligned}(\mathbf{u} + \mathbf{v}) + \mathbf{w} &= ([u_1, u_2, \ldots, u_n] + [v_1, v_2, \ldots, v_n]) + [w_1, w_2, \ldots, w_n]\\ &= [u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n] + [w_1, w_2, \ldots, w_n]\\ &= [(u_1 + v_1) + w_1, (u_2 + v_2) + w_2, \ldots, (u_n + v_n) + w_n]\\ &= [u_1 + (v_1 + w_1), u_2 + (v_2 + w_2), \ldots, u_n + (v_n + w_n)]\\ &= [u_1, u_2, \ldots, u_n] + [v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n]\\ &= [u_1, u_2, \ldots, u_n] + ([v_1, v_2, \ldots, v_n] + [w_1, w_2, \ldots, w_n])\\ &= \mathbf{u} + (\mathbf{v} + \mathbf{w})\end{aligned}$$

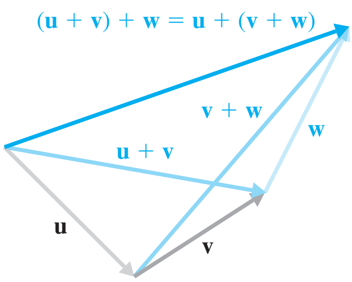


Figure 1.18

The fourth equality is by the associativity of addition of real numbers. Note the careful use of parentheses.

By property (b) of Theorem 1.1, we may unambiguously write $\mathbf{u} + \mathbf{v} + \mathbf{w}$ without parentheses, since we may group the summands in whichever way we please. By (a), we may also rearrange the summands—for example, as $\mathbf{w} + \mathbf{u} + \mathbf{v}$—if we choose. Likewise, sums of four or more vectors can be calculated without regard to order or grouping. In general, if $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ are vectors in $\mathbb{R}^n$, we will write such sums without parentheses:

$$\mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_k$$

The next example illustrates the use of Theorem 1.1 in performing algebraic calculations with vectors.

Example 1.5

Let $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{x}$ denote vectors in $\mathbb{R}^n$.
(a) Simplify $3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a})$.
(b) If $5\mathbf{x} - \mathbf{a} = 2(\mathbf{a} + 2\mathbf{x})$, solve for $\mathbf{x}$ in terms of $\mathbf{a}$.

Solution We will give both solutions in detail, with reference to all of the properties in Theorem 1.1 that we use. It is good practice to justify all steps the first few times you do this type of calculation. Once you are comfortable with the vector properties, though, it is acceptable to leave out some of the intermediate steps to save time and space.

(a) We begin by inserting parentheses.

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= (3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a})) + 2(\mathbf{b} - \mathbf{a}) \\
&= (3\mathbf{a} + (-2\mathbf{a} + 5\mathbf{b})) + (2\mathbf{b} - 2\mathbf{a}) && \text{(a), (e)} \\
&= ((3\mathbf{a} + (-2\mathbf{a})) + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(b)} \\
&= ((3 + (-2))\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) && \text{(f)} \\
&= (1\mathbf{a} + 5\mathbf{b}) + (2\mathbf{b} - 2\mathbf{a}) \\
&= ((\mathbf{a} + 5\mathbf{b}) + 2\mathbf{b}) - 2\mathbf{a} && \text{(b), (h)} \\
&= (\mathbf{a} + (5\mathbf{b} + 2\mathbf{b})) - 2\mathbf{a} && \text{(b)} \\
&= (\mathbf{a} + (5 + 2)\mathbf{b}) - 2\mathbf{a} && \text{(f)} \\
&= (7\mathbf{b} + \mathbf{a}) - 2\mathbf{a} && \text{(a)} \\
&= 7\mathbf{b} + (\mathbf{a} - 2\mathbf{a}) && \text{(b)} \\
&= 7\mathbf{b} + (1 - 2)\mathbf{a} && \text{(f), (h)} \\
&= 7\mathbf{b} + (-1)\mathbf{a} \\
&= 7\mathbf{b} - \mathbf{a}
\end{aligned}$$

You can see why we will agree to omit some of these steps! In practice, it is acceptable to simplify this sequence of steps as

$$\begin{aligned}
3\mathbf{a} + (5\mathbf{b} - 2\mathbf{a}) + 2(\mathbf{b} - \mathbf{a}) &= 3\mathbf{a} + 5\mathbf{b} - 2\mathbf{a} + 2\mathbf{b} - 2\mathbf{a} \\
&= (3\mathbf{a} - 2\mathbf{a} - 2\mathbf{a}) + (5\mathbf{b} + 2\mathbf{b}) \\
&= -\mathbf{a} + 7\mathbf{b}
\end{aligned}$$

or even to do most of the calculation mentally.

(b) In detail, we have

$$\begin{aligned}
5\mathbf{x} - \mathbf{a} &= 2(\mathbf{a} + 2\mathbf{x}) & \\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 2(2\mathbf{x}) & \text{(e)} \\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + (2 \cdot 2)\mathbf{x} & \text{(g)} \\
5\mathbf{x} - \mathbf{a} &= 2\mathbf{a} + 4\mathbf{x} & \\
(5\mathbf{x} - \mathbf{a}) - 4\mathbf{x} &= (2\mathbf{a} + 4\mathbf{x}) - 4\mathbf{x} & \\
(-\mathbf{a} + 5\mathbf{x}) - 4\mathbf{x} &= 2\mathbf{a} + (4\mathbf{x} - 4\mathbf{x}) & \text{(a), (b)} \\
-\mathbf{a} + (5\mathbf{x} - 4\mathbf{x}) &= 2\mathbf{a} + \mathbf{0} & \text{(b), (d)} \\
-\mathbf{a} + (5 - 4)\mathbf{x} &= 2\mathbf{a} & \text{(f), (c)} \\
-\mathbf{a} + (1)\mathbf{x} &= 2\mathbf{a} & \\
\mathbf{a} + (-\mathbf{a} + \mathbf{x}) &= \mathbf{a} + 2\mathbf{a} & \text{(h)} \\
(\mathbf{a} + (-\mathbf{a})) + \mathbf{x} &= (1 + 2)\mathbf{a} & \text{(b), (f)} \\
\mathbf{0} + \mathbf{x} &= 3\mathbf{a} & \text{(d)} \\
\mathbf{x} &= 3\mathbf{a} & \text{(c)}
\end{aligned}$$

Again, we will usually omit most of these steps.

Linear Combinations and Coordinates

A vector that is a sum of scalar multiples of other vectors is said to be a *linear combination* of those vectors. The formal definition follows.

Definition A vector $\mathbf{v}$ is a ***linear combination*** of vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ if there are scalars $c_1, c_2, \ldots, c_k$ such that $\mathbf{v} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$. The scalars $c_1, c_2, \ldots, c_k$ are called the ***coefficients*** of the linear combination.

Example 1.6

The vector $\begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$ is a linear combination of $\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}$, $\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}$, and $\begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix}$, since

$$3\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} + 2\begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} - \begin{bmatrix} 5 \\ -4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -2 \\ -1 \end{bmatrix}$$

Remark Determining whether a given vector is a linear combination of other vectors is a problem we will address in Chapter 2.

In $\mathbb{R}^2$, it is possible to depict linear combinations of two (nonparallel) vectors quite conveniently.

Example 1.7

Let $\mathbf{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$. We can use $\mathbf{u}$ and $\mathbf{v}$ to locate a new set of axes (in the same way that $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ locate the standard coordinate axes). We can use

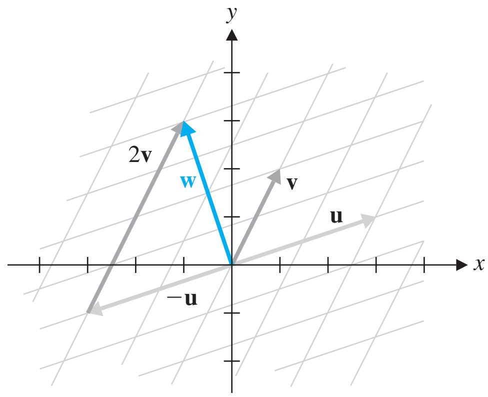


Figure 1.19

these new axes to determine a ***coordinate grid*** that will let us easily locate linear combinations of **u** and **v**.

As Figure 1.19 shows, **w** can be located by starting at the origin and traveling $-\mathbf{u}$ followed by $2\mathbf{v}$. That is,

$$\mathbf{w} = -\mathbf{u} + 2\mathbf{v}$$

We say that the coordinates of **w** with respect to **u** and **v** are -1 and 2. (Note that this is just another way of thinking of the coefficients of the linear combination.) It follows that

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

(Observe that -1 and 3 are the coordinates of **w** with respect to $\mathbf{e}_1$ and $\mathbf{e}_2$.)

Switching from the standard coordinate axes to alternative ones is a useful idea. It has applications in chemistry and geology, since molecular and crystalline structures often do not fall onto a rectangular grid. It is an idea that we will encounter repeatedly in this book.

Binary Vectors and Modular Arithmetic

We will also encounter a type of vector that has no geometric interpretation—at least not using Euclidean geometry. Computers represent data in terms of 0s and 1s (which can be interpreted as off/on, closed/open, false/true, or no/yes). ***Binary vectors*** are vectors each of whose components is a 0 or a 1. As we will see in Chapter 8, such vectors arise naturally in the study of many types of codes.

In this setting, the usual rules of arithmetic must be modified, since the result of each calculation involving scalars must be a 0 or a 1. The modified rules for addition and multiplication are given below.

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

The only curiosity here is the rule that $1 + 1 = 0$. This is not as strange as it appears; if we replace 0 with the word "even" and 1 with the word "odd," these tables simply summarize the familiar *parity rules* for the addition and multiplication of even and odd integers. For example, $1 + 1 = 0$ expresses the fact that the sum of two odd integers is an even integer. With these rules, our set of scalars {0, 1} is denoted by $\mathbb{Z}_2$ and is called the set of ***integers modulo 2.***

Example 1.8

In $\mathbb{Z}_2$, $1 + 1 + 0 + 1 = 1$ and $1 + 1 + 1 + 1 = 0$. (These calculations illustrate the parity rules again: The sum of three odds and an even is odd; the sum of four odds is even.)

We are using the term *length* differently from the way we used it in $\mathbb{R}^n$. This should not be confusing, since there is no *geometric* notion of length for binary vectors.

With $\mathbb{Z}_2$ as our set of scalars, we now extend the above rules to vectors. The set of all n-tuples of 0s and 1s (with all arithmetic performed modulo 2) is denoted by $\mathbb{Z}_2^n$. The vectors in $\mathbb{Z}_2^n$ are called ***binary vectors of length n.***

Example 1.9

The vectors in $\mathbb{Z}_2^2$ are [0, 0], [0, 1], [1, 0], and [1, 1]. (How many vectors does $\mathbb{Z}_2^n$ contain, in general?)

Example 1.10

Let $\mathbf{u} = [1, 1, 0, 1, 0]$ and $\mathbf{v} = [0, 1, 1, 1, 0]$ be two binary vectors of length 5. Find $\mathbf{u} + \mathbf{v}$.

Solution The calculation of $\mathbf{u} + \mathbf{v}$ takes place over $\mathbb{Z}_2$, so we have

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= [1, 1, 0, 1, 0] + [0, 1, 1, 1, 0] \\ &= [1 + 0, 1 + 1, 0 + 1, 1 + 1, 0 + 0] \\ &= [1, 0, 1, 0, 0] \end{aligned}$$

It is possible to generalize what we have just done for binary vectors to vectors whose components are taken from a finite set $\{0, 1, 2, \ldots, k\}$ for $k \geq 2$. To do so, we must first extend the idea of binary arithmetic.

Example 1.11

The ***integers modulo 3*** is the set $\mathbb{Z}_3 = \{0, 1, 2\}$ with addition and multiplication given by the following tables:

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

Observe that the result of each addition and multiplication belongs to the set {0, 1, 2}; we say that $\mathbb{Z}_3$ is ***closed*** with respect to the operations of addition and multiplication. It is perhaps easiest to think of this set in terms of a 3-hour clock with 0, 1, and 2 on its face, as shown in Figure 1.20.

The calculation $1 + 2 = 0$ translates as follows: 2 hours after 1 o'clock, it is 0 o'clock. Just as 24:00 and 12:00 are the same on a 12-hour clock, so 3 and 0 are equivalent on this 3-hour clock. Likewise, all multiples of 3—positive and negative—are equivalent to 0 here; 1 is equivalent to any number that is 1 more than a multiple of 3 (such as -2, 4, and 7); and 2 is equivalent to any number that is 2 more than a multiple of 3 (such as −1, 5, and 8). We can visualize the number line as wrapping around a circle, as shown in Figure 1.21.

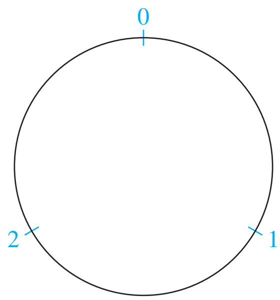


Figure 1.20
Arithmetic modulo 3

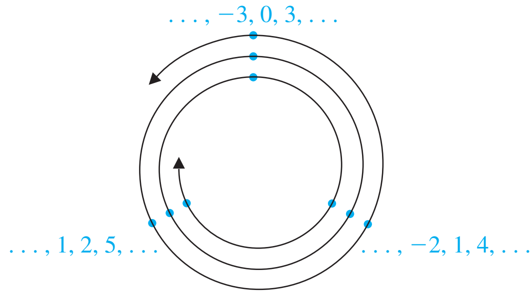


Figure 1.21

Example 1.12

To what is 3548 equivalent in Z_3?

Solution This is the same as asking where 3548 lies on our 3-hour clock. The key is to calculate how far this number is from the nearest (smaller) multiple of 3; that is, we need to know the *remainder* when 3548 is divided by 3. By long division, we find that $3548 = 3 \cdot 1182 + 2$, so the remainder is 2. Therefore, 3548 is equivalent to 2 in $\mathbb{Z}_3$.

In courses in abstract algebra and number theory, which explore this concept in greater detail, the above equivalence is often written as $3548 = 2 \pmod 3$ or $3548 \equiv 2 \pmod 3$, where $\equiv$ is read "is congruent to." We will not use this notation or terminology here.

Example 1.13

In $\mathbb{Z}_3$, calculate $2 + 2 + 1 + 2$.

Solution 1 We use the same ideas as in Example 1.12. The ordinary sum is $2 + 2 + 1 + 2 = 7$, which is 1 more than 6, so division by 3 leaves a remainder of 1. Thus, $2 + 2 + 1 + 2 = 1$ in $\mathbb{Z}_3$.

Solution 2 A better way to perform this calculation is to do it step by step entirely in $\mathbb{Z}_3$.

$$\begin{aligned} 2 + 2 + 1 + 2 &= (2 + 2) + 1 + 2 \\ &= 1 + 1 + 2 \\ &= (1 + 1) + 2 \\ &= 2 + 2 \\ &= 1 \end{aligned}$$

Here we have used parentheses to group the terms we have chosen to combine. We could speed things up by simultaneously combining the first two and the last two terms:

$$\begin{aligned} (2 + 2) + (1 + 2) &= 1 + 0 \\ &= 1 \end{aligned}$$

Repeated multiplication can be handled similarly. The idea is to use the addition and multiplication tables to reduce the result of each calculation to 0, 1, or 2.

Extending these ideas to vectors is straightforward.

Example 1.14

In $\mathbb{Z}_3^5$, let $\mathbf{u} = [2, 2, 0, 1, 2]$ and $\mathbf{v} = [1, 2, 2, 2, 1]$. Then

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= [2, 2, 0, 1, 2] + [1, 2, 2, 2, 1] \\ &= [2 + 1, 2 + 2, 0 + 2, 1 + 2, 2 + 1] \\ &= [0, 1, 2, 0, 0] \end{aligned}$$

Vectors in $\mathbb{Z}_3^5$ are referred to as ***ternary vectors of length*** **5.**

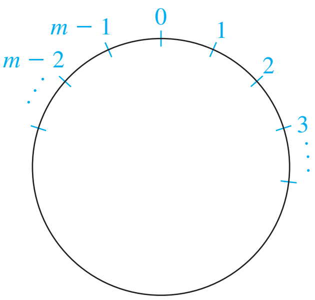


Figure 1.22
Arithmetic modulo m

In general, we have the set $\mathbb{Z}_m = \{0, 1, 2, \ldots, m - 1\}$ of ***integers modulo m*** (corresponding to an m-hour clock, as shown in Figure 1.22). A vector of length n whose entries are in $\mathbb{Z}_m$ is called an ***m-ary vector of length n***. The set of all m-ary vectors of length n is denoted by $\mathbb{Z}_m^n$.

Exercises 1.1

1. Draw the following vectors in standard position in $\mathbb{R}^2$:

(a) $\mathbf{a} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$ **(b)** $\mathbf{b} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

(c) $\mathbf{c} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$ **(d)** $\mathbf{d} = \begin{bmatrix} 3 \\ -2 \end{bmatrix}$

2. Draw the vectors in Exercise 1 with their tails at the point $(2, -3)$.

3. Draw the following vectors in standard position in $\mathbb{R}^3$:

(a) $\mathbf{a} = [0, 2, 0]$ **(b)** $\mathbf{b} = [3, 2, 1]$
(c) $\mathbf{c} = [1, -2, 1]$ **(d)** $\mathbf{d} = [-1, -1, -2]$

4. If the vectors in Exercise 3 are translated so that their heads are at the point $(3, 2, 1)$, find the points that correspond to their tails.

5. For each of the following pairs of points, draw the vector $\overrightarrow{AB}$. Then compute and redraw $\overrightarrow{AB}$ as a vector in standard position.

(a) $A = (1, -1), B = (4, 2)$
(b) $A = (0, -2), B = (2, -1)$
(c) $A = (2, \frac{3}{2}), B = (\frac{1}{2}, 3)$
(d) $A = (\frac{1}{3}, \frac{1}{3}), B = (\frac{1}{6}, \frac{1}{2})$

6. A hiker walks 4 km north and then 5 km northeast. Draw displacement vectors representing the hiker's trip and draw a vector that represents the hiker's net displacement from the starting point.

Exercises 7–10 refer to the vectors in Exercise 1. Compute the indicated vectors and also show how the results can be obtained geometrically.

7. $\mathbf{a} + \mathbf{b}$ **8.** $\mathbf{b} - \mathbf{c}$
9. $\mathbf{d} - \mathbf{c}$ **10.** $\mathbf{a} + \mathbf{d}$

Exercises 11 and 12 refer to the vectors in Exercise 3. Compute the indicated vectors.

11. $2\mathbf{a} + 3\mathbf{c}$ **12.** $3\mathbf{b} - 2\mathbf{c} + \mathbf{d}$

13. Find the components of the vectors $\mathbf{u}$, $\mathbf{v}$, $\mathbf{u} + \mathbf{v}$, and $\mathbf{u} - \mathbf{v}$, where $\mathbf{u}$ and $\mathbf{v}$ are as shown in Figure 1.23.

14. In Figure 1.24, $A, B, C, D, E,$ and F are the vertices of a regular hexagon centered at the origin.

Express each of the following vectors in terms of $\mathbf{a} = \overrightarrow{OA}$ and $\mathbf{b} = \overrightarrow{OB}$:

(a) $\overrightarrow{AB}$ **(b)** $\overrightarrow{BC}$
(c) $\overrightarrow{AD}$ **(d)** $\overrightarrow{CF}$
(e) $\overrightarrow{AC}$ **(f)** $\overrightarrow{BC} + \overrightarrow{DE} + \overrightarrow{FA}$

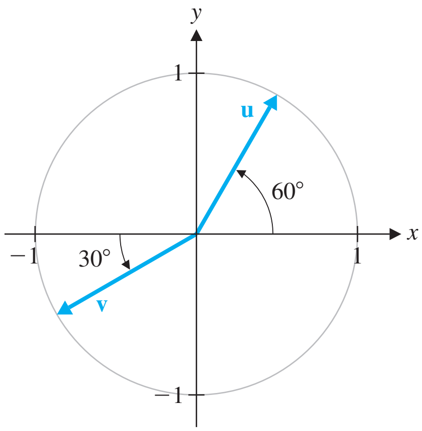


Figure 1.23

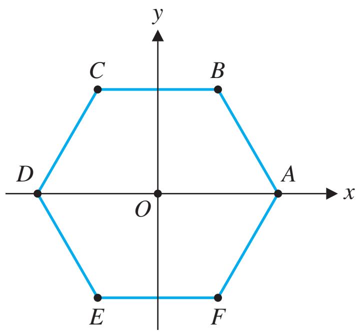


Figure 1.24

In Exercises 15 and 16, simplify the given vector expression. Indicate which properties in Theorem 1.1 you use.

15. $2(\mathbf{a} - 3\mathbf{b}) + 3(2\mathbf{b} + \mathbf{a})$

16. $-3(\mathbf{a} - \mathbf{c}) + 2(\mathbf{a} + 2\mathbf{b}) + 3(\mathbf{c} - \mathbf{b})$

In Exercises 17 and 18, solve for the vector **x** *in terms of the vectors* **a** *and* **b**.

17. $\mathbf{x} - \mathbf{a} = 2(\mathbf{x} - 2\mathbf{a})$

18. $\mathbf{x} + 2\mathbf{a} - \mathbf{b} = 3(\mathbf{x} + \mathbf{a}) - 2(2\mathbf{a} - \mathbf{b})$

In Exercises 19 and 20, draw the coordinate axes relative to **u** *and* **v** *and locate* **w**.

19. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = 2\mathbf{u} + 3\mathbf{v}$

20. $\mathbf{u} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \mathbf{w} = -\mathbf{u} - 2\mathbf{v}$

In Exercises 21 and 22, draw the standard coordinate axes on the same diagram as the axes relative to **u** *and* **v**. *Use these to find* **w** *as a linear combination of* **u** *and* **v**.

21. $\mathbf{u} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$

22. $\mathbf{u} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \mathbf{w} = \begin{bmatrix} 2 \\ 9 \end{bmatrix}$

23. Draw diagrams to illustrate properties (d) and (e) of Theorem 1.1.

24. Give algebraic proofs of properties (d) through (g) of Theorem 1.1.

In Exercises 25–28, **u** *and* **v** *are binary vectors. Find* **u** + **v** *in each case.*

25. $\mathbf{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

26. $\mathbf{u} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

27. $\mathbf{u} = [1, 0, 1, 1], \mathbf{v} = [1, 1, 1, 1]$

28. $\mathbf{u} = [1, 1, 0, 1, 0], \mathbf{v} = [0, 1, 1, 1, 0]$

29. Write out the addition and multiplication tables for $\mathbb{Z}_4$.

30. Write out the addition and multiplication tables for $\mathbb{Z}_5$.

In Exercises 31–43, perform the indicated calculations.

31. $2 + 2 + 2$ in $\mathbb{Z}_3$

32. $2 \cdot 2 \cdot 2$ in $\mathbb{Z}_3$

33. $2(2 + 1 + 2)$ in $\mathbb{Z}_3$

34. $3 + 1 + 2 + 3$ in $\mathbb{Z}_4$

35. $2 \cdot 3 \cdot 2$ in $\mathbb{Z}_4$

36. $3(3 + 3 + 2)$ in $\mathbb{Z}_4$

37. $2 + 1 + 2 + 2 + 1$ in $\mathbb{Z}_3$, Z_4, and $\mathbb{Z}_5$

38. $(3 + 4)(3 + 2 + 4 + 2)$ in $\mathbb{Z}_5$

39. $8(6 + 4 + 3)$ in $\mathbb{Z}_9$

40. 2^{100} in $\mathbb{Z}_{11}$

41. $[2, 1, 2] + [2, 0, 1]$ in $\mathbb{Z}_3^3$

42. $2[2, 2, 1]$ in $\mathbb{Z}_3^3$

43. $2([3, 1, 1, 2] + [3, 3, 2, 1])$ in $\mathbb{Z}_4^4$ and $\mathbb{Z}_5^4$

In Exercises 44–55, solve the given equation or indicate that there is no solution.

44. $x + 3 = 2$ in $\mathbb{Z}_5$

45. $x + 5 = 1$ in $\mathbb{Z}_6$

46. $2x = 1$ in $\mathbb{Z}_3$

47. $2x = 1$ in $\mathbb{Z}_4$

48. $2x = 1$ in $\mathbb{Z}_5$

49. $3x = 4$ in $\mathbb{Z}_5$

50. $3x = 4$ in $\mathbb{Z}_6$

51. $6x = 5$ in $\mathbb{Z}_8$

52. $8x = 9$ in $\mathbb{Z}_{11}$

53. $2x + 3 = 2$ in $\mathbb{Z}_5$

54. $4x + 5 = 2$ in $\mathbb{Z}_6$

55. $6x + 3 = 1$ in $\mathbb{Z}_8$

56. **(a)** For which values of a does $x + a = 0$ have a solution in $\mathbb{Z}_5$?
(b) For which values of a and b does $x + a = b$ have a solution in $\mathbb{Z}_6$?
(c) For which values of a, b, and m does $x + a = b$ have a solution in $\mathbb{Z}_m$?

57. **(a)** For which values of a does $ax = 1$ have a solution in $\mathbb{Z}_5$?
(b) For which values of a does $ax = 1$ have a solution in $\mathbb{Z}_6$?
(c) For which values of a and m does $ax = 1$ have a solution in $\mathbb{Z}_m$?

1.2 Length and Angle: The Dot Product

It is quite easy to reformulate the familiar geometric concepts of length, distance, and angle in terms of vectors. Doing so will allow us to use these important and powerful ideas in settings more general than $\mathbb{R}^2$ and $\mathbb{R}^3$. In subsequent chapters, these simple geometric tools will be used to solve a wide variety of problems arising in applications—even when there is no geometry apparent at all!

The Dot Product

The vector versions of length, distance, and angle can all be described using the notion of the dot product of two vectors.

Definition If

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

then the ***dot product*** $\mathbf{u}\cdot\mathbf{v}$ of $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$\mathbf{u}\cdot\mathbf{v} = u_1v_1 + u_2v_2 + \cdots + u_nv_n$$

In words, $\mathbf{u}\cdot\mathbf{v}$ is the sum of the products of the corresponding components of $\mathbf{u}$ and $\mathbf{v}$. It is important to note a couple of things about this "product" that we have just defined: First, $\mathbf{u}$ and $\mathbf{v}$ must have the same number of components. Second, the dot product $\mathbf{u}\cdot\mathbf{v}$ is a *number,* not another vector. (This is why $\mathbf{u}\cdot\mathbf{v}$ is sometimes called the ***scalar product*** of $\mathbf{u}$ and $\mathbf{v}$.) The dot product of vectors in $\mathbb{R}^n$ is a special and important case of the more general notion of ***inner product,*** which we will explore in Chapter 7.

Example 1.15

Compute $\mathbf{u}\cdot\mathbf{v}$ when $\mathbf{u} = \begin{bmatrix} 1 \\ 2 \\ -3 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} -3 \\ 5 \\ 2 \end{bmatrix}$.

Solution $\mathbf{u}\cdot\mathbf{v} = 1\cdot(-3) + 2\cdot 5 + (-3)\cdot 2 = 1$

Notice that if we had calculated $\mathbf{v}\cdot\mathbf{u}$ in Example 1.15, we would have computed

$$\mathbf{v}\cdot\mathbf{u} = (-3)\cdot 1 + 5\cdot 2 + 2\cdot(-3) = 1$$

That $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$ in general is clear, since the individual products of the components commute. This commutativity property is one of the properties of the dot product that we will use repeatedly. The main properties of the dot product are summarized in Theorem 1.2.

Theorem 1.2 Let $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{w}$ be vectors in $\mathbb{R}^n$ and let c be a scalar. Then

a. $\mathbf{u}\cdot\mathbf{v} = \mathbf{v}\cdot\mathbf{u}$ — Commutativity
b. $\mathbf{u}\cdot(\mathbf{v}+\mathbf{w}) = \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{w}$ — Distributivity
c. $(c\mathbf{u})\cdot\mathbf{v} = c(\mathbf{u}\cdot\mathbf{v})$
d. $\mathbf{u}\cdot\mathbf{u} \geq 0$ and $\mathbf{u}\cdot\mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$

Proof We prove (a) and (c) and leave proof of the remaining properties for the exercises.

(a) Applying the definition of dot product to $\mathbf{u}\cdot\mathbf{v}$ and $\mathbf{v}\cdot\mathbf{u}$, we obtain

$$\begin{aligned}\mathbf{u}\cdot\mathbf{v} &= u_1v_1 + u_2v_2 + \cdots + u_nv_n \\ &= v_1u_1 + v_2u_2 + \cdots + v_nu_n \\ &= \mathbf{v}\cdot\mathbf{u}\end{aligned}$$

where the middle equality follows from the fact that multiplication of real numbers is commutative.

(c) Using the definitions of scalar multiplication and dot product, we have

$$\begin{aligned}(c\mathbf{u})\cdot\mathbf{v} &= [cu_1, cu_2, \ldots, cu_n]\cdot[v_1, v_2, \ldots, v_n] \\ &= cu_1v_1 + cu_2v_2 + \cdots + cu_nv_n \\ &= c(u_1v_1 + u_2v_2 + \cdots + u_nv_n) \\ &= c(\mathbf{u}\cdot\mathbf{v})\end{aligned}$$

Remarks

- Property (b) can be read from right to left, in which case it says that we can factor out a common vector $\mathbf{u}$ from a sum of dot products. This property also has a "right-handed" analogue that follows from properties (b) and (a) together: $(\mathbf{v}+\mathbf{w})\cdot\mathbf{u} = \mathbf{v}\cdot\mathbf{u} + \mathbf{w}\cdot\mathbf{u}$.
- Property (c) can be extended to give $\mathbf{u}\cdot(c\mathbf{v}) = c(\mathbf{u}\cdot\mathbf{v})$ (Exercise 58). This extended version of (c) essentially says that in taking a scalar multiple of a dot product of vectors, the scalar can first be combined with whichever vector is more convenient. For example,

$$\left(\tfrac{1}{2}[-1,-3,2]\right)\cdot[6,-4,0] = [-1,-3,2]\cdot\left(\tfrac{1}{2}[6,-4,0]\right) = [-1,-3,2]\cdot[3,-2,0] = 3$$

With this approach we avoid introducing fractions into the vectors, as the original grouping would have.

- The second part of (d) uses the logical connective *if and only if*. Appendix A discusses this phrase in more detail, but for the moment let us just note that the wording signals a *double implication*—namely,

$$\text{if } \mathbf{u} = \mathbf{0}, \text{ then } \mathbf{u}\cdot\mathbf{u} = 0$$

and

$$\text{if } \mathbf{u}\cdot\mathbf{u} = 0, \text{ then } \mathbf{u} = \mathbf{0}$$

Theorem 1.2 shows that aspects of the algebra of vectors resemble the algebra of numbers. The next example shows that we can sometimes find vector analogues of familiar identities.

Example 1.16 Prove that $(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) = \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v}$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$.

Solution

$$\begin{aligned}(\mathbf{u}+\mathbf{v})\cdot(\mathbf{u}+\mathbf{v}) &= (\mathbf{u}+\mathbf{v})\cdot\mathbf{u} + (\mathbf{u}+\mathbf{v})\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{v}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{v} + \mathbf{u}\cdot\mathbf{v} + \mathbf{v}\cdot\mathbf{v}\\ &= \mathbf{u}\cdot\mathbf{u} + 2(\mathbf{u}\cdot\mathbf{v}) + \mathbf{v}\cdot\mathbf{v}\end{aligned}$$

(Identify the parts of Theorem 1.2 that were used at each step.)

Length

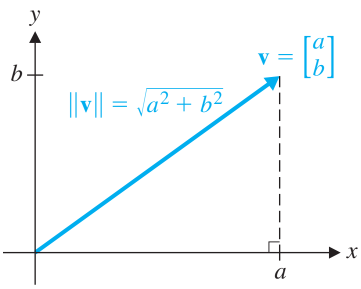


Figure 1.25

To see how the dot product plays a role in the calculation of lengths, recall how lengths are computed in the plane. The Theorem of Pythagoras is all we need.

In $\mathbb{R}^2$, the length of the vector $\mathbf{v} = \begin{bmatrix} a \\ b \end{bmatrix}$ is the distance from the origin to the point (a, b), which, by Pythagoras' Theorem, is given by $\sqrt{a^2 + b^2}$, as in Figure 1.25. Observe that $a^2 + b^2 = \mathbf{v}\cdot\mathbf{v}$. This leads to the following definition.

Definition The ***length*** (or ***norm***) of a vector $\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ in $\mathbb{R}^n$ is the nonnegative scalar $\|\mathbf{v}\|$ defined by

$$\|\mathbf{v}\| = \sqrt{\mathbf{v}\cdot\mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}$$

In words, the length of a vector is the square root of the sum of the squares of its components. Note that the square root of $\mathbf{v}\cdot\mathbf{v}$ is always defined, since $\mathbf{v}\cdot\mathbf{v} \geq 0$ by Theorem 1.2(d). Note also that the definition can be rewritten to give $\|\mathbf{v}\|^2 = \mathbf{v}\cdot\mathbf{v}$, which will be useful in proving further properties of the dot product and lengths of vectors.

Example 1.17

$$\|[2, 3]\| = \sqrt{2^2 + 3^2} = \sqrt{13}$$

Theorem 1.3 lists some of the main properties of vector length.

Theorem 1.3 Let $\mathbf{v}$ be a vector in $\mathbb{R}^n$ and let c be a scalar. Then

a. $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = \mathbf{0}$
b. $\|c\mathbf{v}\| = |c|\,\|\mathbf{v}\|$

Proof Property (a) follows immediately from Theorem 1.2(d). To show (b), we have

$$\|c\mathbf{v}\|^2 = (c\mathbf{v})\cdot(c\mathbf{v}) = c^2(\mathbf{v}\cdot\mathbf{v}) = c^2\|\mathbf{v}\|^2$$

using Theorem 1.2(c). Taking square roots of both sides, using the fact that $\sqrt{c^2} = |c|$ for any real number c, gives the result.

A vector of length 1 is called a ***unit vector***. In $\mathbb{R}^2$, the set of all unit vectors can be identified with the *unit circle*, the circle of radius 1 centered at the origin (see Figure 1.26). Given any nonzero vector $\mathbf{v}$, we can always find a unit vector in the same direction as $\mathbf{v}$ by dividing $\mathbf{v}$ by its own length (or, equivalently, *multiplying* by $1/\|\mathbf{v}\|$). We can show this algebraically by using property (b) of Theorem 1.3 above: If $\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v}$, then

$$\|\mathbf{u}\| = \|(1/\|\mathbf{v}\|)\mathbf{v}\| = |1/\|\mathbf{v}\||\,\|\mathbf{v}\| = (1/\|\mathbf{v}\|)\|\mathbf{v}\| = 1$$

and $\mathbf{u}$ is in the same direction as $\mathbf{v}$, since $1/\|\mathbf{v}\|$ is a positive scalar. Finding a unit vector in the same direction is often referred to as ***normalizing*** a vector (see Figure 1.27).

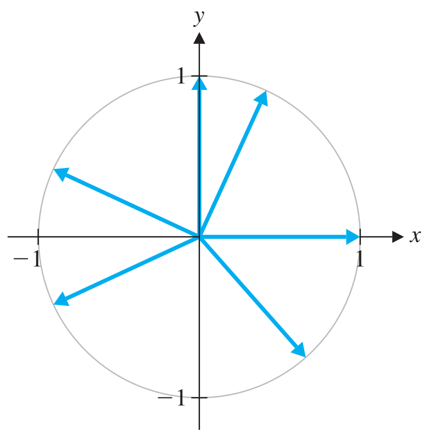


Figure 1.26
Unit vectors in $\mathbb{R}^2$

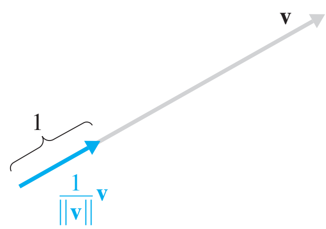


Figure 1.27
Normalizing a vector

Example 1.18

In $\mathbb{R}^2$, let $\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$. Then $\mathbf{e}_1$ and $\mathbf{e}_2$ are unit vectors, since the sum of the squares of their components is 1 in each case. Similarly, in $\mathbb{R}^3$, we can construct unit vectors

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Observe in Figure 1.28 that these vectors serve to locate the positive coordinate axes in $\mathbb{R}^2$ and $\mathbb{R}^3$.

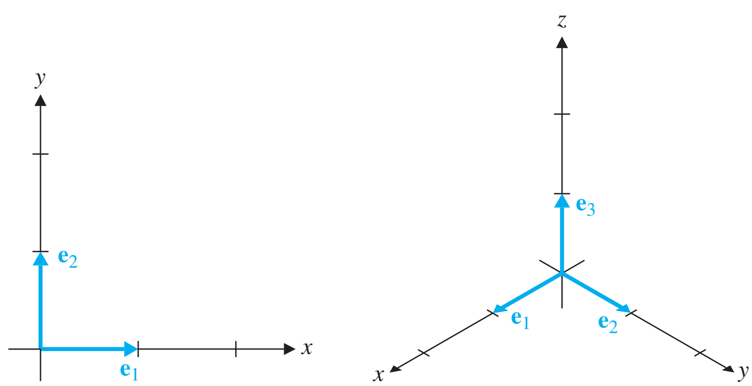


Figure 1.28
Standard unit vectors in $\mathbb{R}^2$ and $\mathbb{R}^3$

In general, in $\mathbb{R}^n$, we define unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$, where $\mathbf{e}_i$ has 1 in its ith component and zeros elsewhere. These vectors arise repeatedly in linear algebra and are called the ***standard unit vectors.***

Example 1.19

Normalize the vector $\mathbf{v} = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}$.

Solution $\|\mathbf{v}\| = \sqrt{2^2 + (-1)^2 + 3^2} = \sqrt{14}$, so a unit vector in the same direction as $\mathbf{v}$ is given by

$$\mathbf{u} = (1/\|\mathbf{v}\|)\mathbf{v} = (1/\sqrt{14})\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{14} \\ -1/\sqrt{14} \\ 3/\sqrt{14} \end{bmatrix}$$

Since property (b) of Theorem 1.3 describes how length behaves with respect to scalar multiplication, natural curiosity suggests that we ask whether length and vector addition are compatible. It would be nice if we had an identity such as $\|\mathbf{u} + \mathbf{v}\| = \|\mathbf{u}\| + \|\mathbf{v}\|$, but for almost any choice of vectors $\mathbf{u}$ and $\mathbf{v}$ this turns out to be false. [See Exercise 52(a).] However, all is not lost, for it turns out that if we replace the = sign by $\leq$, the resulting inequality is true. The proof of this famous and important result—the Triangle Inequality—relies on another important inequality—the Cauchy-Schwarz Inequality—which we will prove and discuss in more detail in Chapter 7.

Theorem 1.4 **The Cauchy-Schwarz Inequality**

For all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$,

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\| \, \|\mathbf{v}\|$$

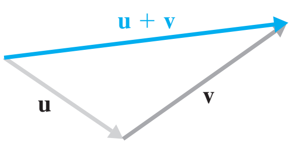


Figure 1.29
The Triangle Inequality

See Exercises 71 and 72 for algebraic and geometric approaches to the proof of this inequality.

In $\mathbb{R}^2$ or $\mathbb{R}^3$, where we can use geometry, it is clear from a diagram such as Figure 1.29 that $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ for all vectors $\mathbf{u}$ and $\mathbf{v}$. We now show that this is true more generally.

Theorem 1.5 **The Triangle Inequality**

For all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$,

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Proof Since both sides of the inequality are nonnegative, showing that the *square* of the left-hand side is less than or equal to the *square* of the right-hand side is equivalent to proving the theorem. (Why?) We compute

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{u} + 2(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v} && \text{By Example 1.9} \\ &\le \|\mathbf{u}\|^2 + 2|\mathbf{u} \cdot \mathbf{v}| + \|\mathbf{v}\|^2 \\ &\le \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\,\|\mathbf{v}\| + \|\mathbf{v}\|^2 && \text{By Cauchy-Schwarz} \\ &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \end{aligned}$$

as required.

Distance

The distance between two vectors is the direct analogue of the distance between two points on the real number line or two points in the Cartesian plane. On the number line (Figure 1.30), the distance between the numbers a and b is given by $|a - b|$. (Taking the absolute value ensures that we do not need to know which of a or b is larger.) This distance is also equal to $\sqrt{(a-b)^2}$, and its two-dimensional generalization is the familiar formula for the distance d between points (a_1, a_2) and (b_1, b_2)—namely, $d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}$.

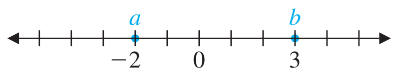


Figure 1.30
$d = |a - b| = |-2 - 3| = 5$

In terms of vectors, if $\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$, then d is just the length of $\mathbf{a} - \mathbf{b}$, as shown in Figure 1.31. This is the basis for the next definition.

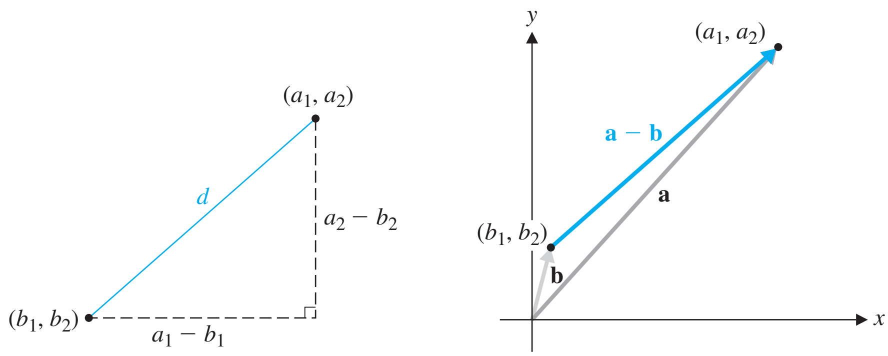


Figure 1.31
$d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2} = \|\mathbf{a} - \mathbf{b}\|$

Definition The ***distance*** $d(\mathbf{u}, \mathbf{v})$ between vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ is defined by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$$

Example 1.20

Find the distance between $\mathbf{u} = \begin{bmatrix} \sqrt{2} \\ 1 \\ -1 \end{bmatrix}$ and $\mathbf{v} = \begin{bmatrix} 0 \\ 2 \\ -2 \end{bmatrix}$.

Solution We compute $\mathbf{u} - \mathbf{v} = \begin{bmatrix} \sqrt{2} \\ -1 \\ 1 \end{bmatrix}$, so

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(\sqrt{2})^2 + (-1)^2 + 1^2} = \sqrt{4} = 2$$

Angles

The dot product can also be used to calculate the angle between a pair of vectors. In $\mathbb{R}^2$ or $\mathbb{R}^3$, the angle between the nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ will refer to the angle θ determined by these vectors that satisfies $0 \le \theta \le 180°$ (see Figure 1.32).

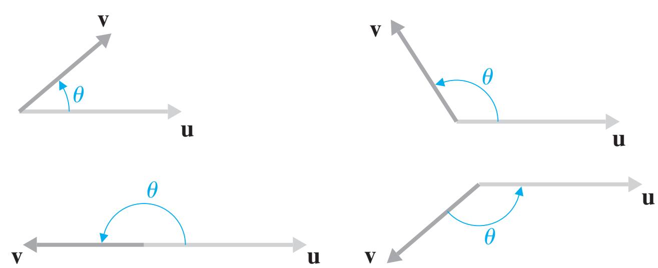


Figure 1.32
The angle between **u** and **v**

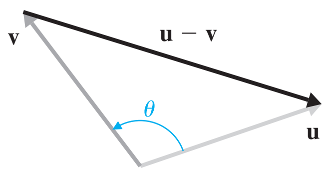


Figure 1.33

In Figure 1.33, consider the triangle with sides $\mathbf{u}$, $\mathbf{v}$, and $\mathbf{u} - \mathbf{v}$, where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$. Applying the law of cosines to this triangle yields

$$\|\mathbf{u} - \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

Expanding the left-hand side and using $\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v}$ several times, we obtain

$$\|\mathbf{u}\|^2 - 2(\mathbf{u} \cdot \mathbf{v}) + \|\mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - 2\|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$$

which, after simplification, leaves us with $\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\|\,\|\mathbf{v}\| \cos\theta$. From this we obtain the following formula for the cosine of the angle θ between nonzero vectors $\mathbf{u}$ and $\mathbf{v}$. We state it as a definition.

Definition For nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$,

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$$

Example 1.21

Compute the angle between the vectors $\mathbf{u} = [2, 1, -2]$ and $\mathbf{v} = [1, 1, 1]$.

Solution We calculate $\mathbf{u} \cdot \mathbf{v} = 2 \cdot 1 + 1 \cdot 1 + (-2) \cdot 1 = 1$, $\|\mathbf{u}\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$, and $\|\mathbf{v}\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$. Therefore, $\cos\theta = 1/3\sqrt{3}$, so $\theta = \cos^{-1}(1/3\sqrt{3}) \approx 1.377$ radians, or $78.9°$.

Example 1.22

Compute the angle between the diagonals on two adjacent faces of a cube.

Solution The dimensions of the cube do not matter, so we will work with a cube with sides of length 1. Orient the cube relative to the coordinate axes in $\mathbb{R}^3$, as shown in Figure 1.34, and take the two side diagonals to be the vectors [1, 0, 1] and [0, 1, 1]. Then angle θ between these vectors satisfies

$$\cos\theta = \frac{1 \cdot 0 + 0 \cdot 1 + 1 \cdot 1}{\sqrt{2}\sqrt{2}} = \frac{1}{2}$$

from which it follows that the required angle is $\pi/3$ radians, or $60°$.

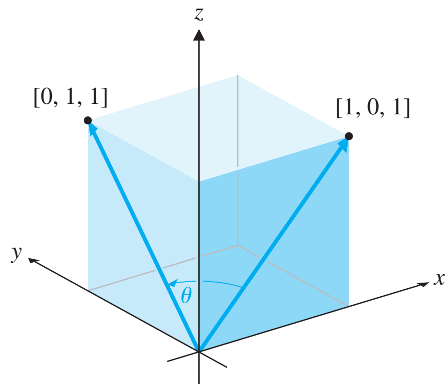


Figure 1.34

(Actually, we don't need to do any calculations at all to get this answer. If we draw a third side diagonal joining the vertices at (1, 0, 1) and (0, 1, 1), we get an equilateral triangle, since all of the side diagonals are of equal length. The angle we want is one of the angles of this triangle and therefore measures 60°. Sometimes, a little insight can save a lot of calculation; in this case, it gives a nice check on our work!)

Remarks

- As this discussion shows, we usually will have to settle for an approximation to the angle between two vectors. However, when the angle is one of the so-called special angles (0°, 30°, 45°, 60°, 90°, or an integer multiple of these), we should be able to recognize its cosine (Table 1.1) and thus give the corresponding angle exactly. In all other cases, we will use a calculator or computer to approximate the desired angle by means of the inverse cosine function.

Table 1.1 Cosines of Special Angles

θ	0°	30°	45°	60°	90°
$\cos\theta$	$\frac{\sqrt{4}}{2} = 1$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2} = \frac{1}{\sqrt{2}}$	$\frac{\sqrt{1}}{2} = \frac{1}{2}$	$\frac{\sqrt{0}}{2} = 0$

- The derivation of the formula for the cosine of the angle between two vectors is valid only in $\mathbb{R}^2$ or $\mathbb{R}^3$, since it depends on a geometric fact: the law of cosines. In $\mathbb{R}^n$, for $n > 3$, the formula can be taken as a *definition* instead. This makes sense, since the Cauchy-Schwarz Inequality implies that $\left|\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}\right| \le 1$, so $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\,\|\mathbf{v}\|}$ ranges from -1 to 1, just as the cosine function does.

Orthogonal Vectors

The word *orthogonal* is derived from the Greek words *orthos,* meaning "upright," and *gonia*, meaning "angle." Hence, orthogonal literally means "right-angled." The Latin equivalent is *rectangular*.

The concept of perpendicularity is fundamental to geometry. Anyone studying geometry quickly realizes the importance and usefulness of right angles. We now generalize the idea of perpendicularity to vectors in $\mathbb{R}^n$, where it is called ***orthogonality.***

In $\mathbb{R}^2$ or $\mathbb{R}^3$, two nonzero vectors $\mathbf{u}$ and $\mathbf{v}$ are perpendicular if the angle θ between them is a right angle—that is, if $\theta = \pi/2$ radians, or 90°. Thus, $\dfrac{\mathbf{u}\cdot\mathbf{v}}{\|\mathbf{u}\|\|\mathbf{v}\|} = \cos 90° = 0$, and it follows that $\mathbf{u}\cdot\mathbf{v} = 0$. This motivates the following definition.

Definition Two vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$ are ***orthogonal*** to each other if $\mathbf{u}\cdot\mathbf{v} = 0$.

Since $\mathbf{0}\cdot\mathbf{v} = 0$ for every vector $\mathbf{v}$ in $\mathbb{R}^n$, the zero vector is orthogonal to every vector.

Example 1.23 In $\mathbb{R}^3$, $\mathbf{u} = [1, 1, -2]$ and $\mathbf{v} = [3, 1, 2]$ are orthogonal, since $\mathbf{u}\cdot\mathbf{v} = 3 + 1 - 4 = 0$.

Using the notion of orthogonality, we get an easy proof of Pythagoras' Theorem, valid in $\mathbb{R}^n$.

Theorem 1.6 **Pythagoras' Theorem**

For all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$, $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if and only if $\mathbf{u}$ and $\mathbf{v}$ are orthogonal.

Proof From Example 1.16, we have $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2$ for all vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^n$. It follows immediately that $\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2$ if and only if $\mathbf{u}\cdot\mathbf{v} = 0$. See Figure 1.35.

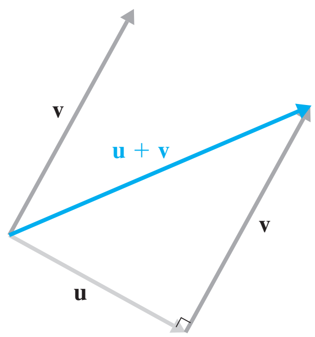


Figure 1.35

The concept of orthogonality is one of the most important and useful in linear algebra, and it often arises in surprising ways. Chapter 5 contains a detailed treatment of the topic, but we will encounter it many times before then. One problem in which it clearly plays a role is finding the distance from a point to a line, where "dropping a perpendicular" is a familiar step.