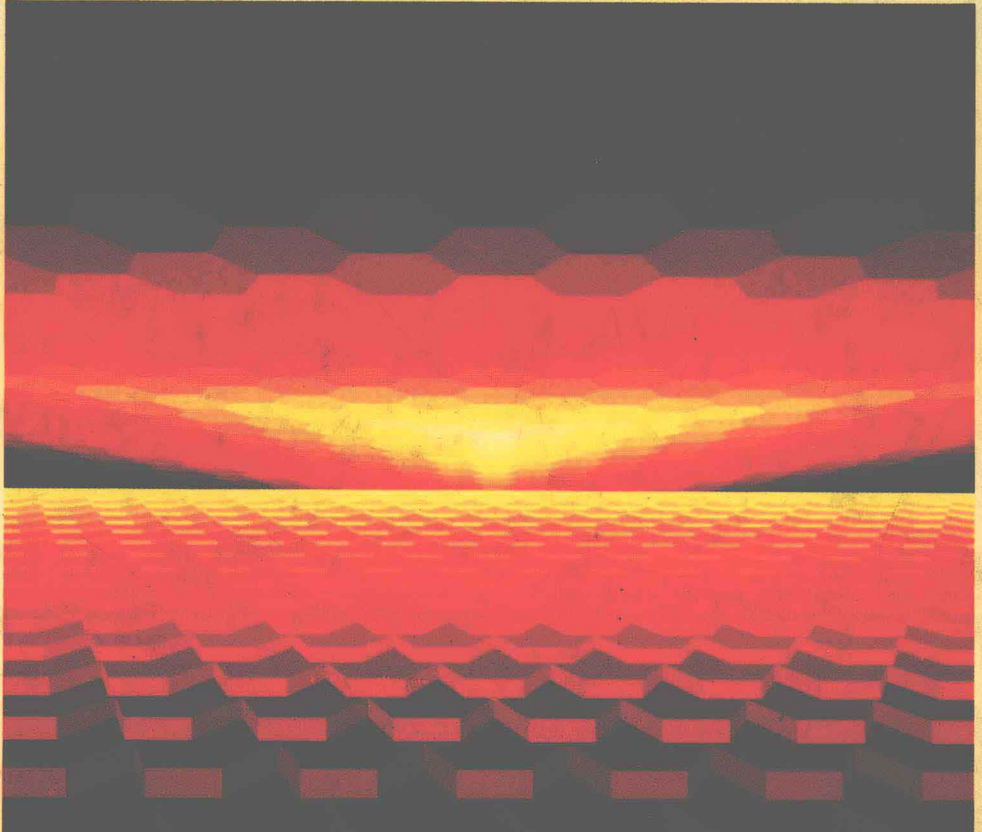
INTRODUCTION TO LINEAR ALGEBRA WITH APPLICATIONS



STEPHEN FRIEDBERG ▪ ARNOLD INSEL

Introduction to Linear Algebra with Applications

Stephen H. Friedberg
Arnold J. Insel

Illinois State University

Prentice-Hall, Englewood Cliffs, N.J.

Library of Congress Cataloging in Publication Data

Friedberg, Stephen H.
Introduction to linear algebra with applications.

Includes index.
1. Algebras, Linear. I. Insel, Arnold J.
II. Title.
QA184.F74 1986 512'.5 85-19126
ISBN 0-13-485988-X

Editorial/production supervision and
interior design: Fay Ahuja
Cover design: Lundgren Graphics, Ltd.
Cover painting: "Horizon Structuré" by Yvaral
Manufacturing buyer: John B. Hall

Printed in the United States of America

10 9 8 7 6 5 4 3 2

ISBN 0-13-485988-X 01

Prentice-Hall International (UK) Limited, *London*
Prentice-Hall of Australia Pty. Limited, *Sydney*
Prentice-Hall Canada Inc., *Toronto*
Prentice-Hall Hispanoamericana, S.A., *Mexico*
Prentice-Hall of India Private Limited, *New Delhi*
Prentice-Hall of Japan, Inc., *Tokyo*
Prentice-Hall of Southeast Asia Pte. Ltd., *Singapore*
Editora Prentice-Hall do Brasil, Ltda., *Rio de Janeiro*
Whitehall Books Limited, *Wellington, New Zealand*

To my wife, Ruth Ann,
and our children, Rachel, Jessica, and Jeremy

S.H.F.

To my wife, Barbara,
and our children, Tom and Sara

A.J.I.

Preface

In recent years the demand for skills in linear algebra has increased at a rapid pace. In addition to engineering, physics, and economics majors, enrollment in linear algebra classes now includes those majoring in computer science, operations research, psychology, and biology. Of course, the use of linear algebra in multivariate statistics has made the subject a natural requirement for those students wishing to pursue a career in other quantitative areas. The presence of so diverse an audience as well as the importance of applications to mathematics majors has influenced our decision to include a wide variety of significant applications. Rather than postpone these applications to the end of the text, we have made an effort to introduce them as the necessary background is developed.

The material is aimed at the sophomore–junior student. There is no use of calculus until the introduction of function spaces in Chapter 8. The core topics include: the vector space properties and Euclidean n-space, systems of linear equations, matrices, linear transformations, determinants, and eigenvalues and eigenvectors. In addition, orthogonal diagonalization, abstract vector spaces, numerical methods for solving systems and for finding eigenvalues and eigenvectors, and linear programming are covered.

Approach: Through experience, we have discovered that students without a background in abstract mathematics find abstract vector spaces very difficult to comprehend. We have therefore decided to introduce the notion of vector spaces and their properties through the familiar Euclidean n-spaces. What we believe is *unique* about this text is the *spiral* approach that is employed. Instead of introducing all the material about systems, followed by all the material on vector spaces, followed by all the material on linear transformations and matrices, we have made an effort to ease the student gently into the elementary properties of all these topics before the advanced properties are introduced.

Not only are the concepts easier to assimilate, but the student sees the interplay between the various structures early in the development. For example, after linear combinations are introduced, systems are presented as a means of discovering if a given vector is a linear combination of other vectors. Matrices are then needed as a convenient notation for using Gaussian elimination. Once matrices are introduced, their arithmetic is motivated by an example of a Markov process. Matrix multiplication provides the motivation for the definition of a linear transformation and its matrix representation. Now it makes sense to talk about the subspaces associated with a linear transformation, namely, the null space and range. All of this is done *before* the more difficult concepts of linear independence, basis, and dimension are introduced.

We are firm believers in the use of **geometry** to motivate as well as clarify many of the topics of linear algebra. For example, the application of the elementary properties of vector arithmetic are used to show that the diagonals of a rhombus bisect one another. The determination of the null space, range, eigenspaces, and other characteristics of the geometric transformations—rotations, projections, reflections, and shear transformations—permeate the exercises and examples.

The **microcomputer** as a tool is ever present in this text. Throughout the exercise sets, problems preceded by (*) are to be done on a microcomputer. A diskette, which includes a number of useful programs for working these problems (see Appendix B), is available to adopters of the text. In addition, the idea of operation counts (for example, in the solution to systems or in the computation of determinants) is used as a measure of computational efficiency.

Chapter 1 introduces the elementary properties of vector operations; norm and dot product in Euclidean 2-, 3-, and n-space; linear combinations and subspaces; systems of linear equations and Gaussian elimination; and matrices as a tool for manipulating systems. Included are examples illustrating the power of linear algebra to establish results in geometry.

Chapter 2 introduces the elementary properties of matrix arithmetic. Five significant applications of matrices are given to illustrate their power. Left-multiplication by a matrix introduces the concept of a linear transformation. The null space and range of a linear transformation are given as examples of subspaces and as tools for studying additional properties of systems.

Chapter 3 moves the reader into the more sophisticated concepts of linear independence, basis, rank, and dimension. These ideas are united with the earlier concepts of system, null space, and range to provide deeper insights. For example, the dimension theorem is proved and used to establish information about solution spaces. The properties of matrix inverses are established. The construction of the inverse of a matrix is accomplished with the introduction of elementary matrices. The results are then applied to the Leontief closed and open economic models. Finally, additional theoretical results about systems are proved.

Chapter 4 introduces change of coordinate vectors and various matrix representations of a linear transformation. This material forms the necessary background for diagonalization.

Chapter 5 begins with the definition of the determinant of a 2×2 matrix and its properties. These properties are then extended to determinants of $n \times n$ matrices. The basic approach is to establish the properties for elementary matrices and then use the fact, established earlier, that every invertible matrix is a product of elementary matrices. The chapter is concluded with Cramer's rule, the classical adjoint, and an application to cryptography.

Chapter 6 introduces perhaps the most important concept in linear algebra—*diagonalization*. The basic results concerning eigenvalues, eigenvectors, and necessary and sufficient conditions for the diagonalization of a matrix or linear transformation are established. The results are then applied to solving difference equations, examining the long-term behavior of Markov chains, and solving systems of differential equations.

Chapter 7 is concerned with the properties of orthogonal sets. The Gram–Schmidt process is used to prove that every subspace of R^n has an orthonormal basis. Diagonalization of a symmetric matrix by an orthogonal matrix is used to transform a quadratic expression into standard form. Orthogonal projections are carefully developed and applied to derive the least-squares formula. Finally, rotation matrices are applied to computer graphics.

Chapter 8 utilizes most of the previous material to develop abstract vector spaces. The emphasis, however, is on function spaces. The differential operator is given as a special case of a linear transformation on an infinite-dimensional vector space. The chapter concludes with the elementary properties of inner product spaces.

Chapter 9 provides a somewhat more extensive treatment of numerical methods than most texts at this level. For solving systems we describe the direct methods—pivoting, the *LU*- and Cholesky decompositions, and the iterative methods—the Jacobi and the Gauss–Seidel methods. For estimating eigenvalues, Gerschgorin's theorem is proved. Finally, for estimating eigenvalues and eigenvectors, the power and inverse power methods as well as the deflation method are developed and illustrated. Exercises are given which use the programs listed in Appendix B and are included on the diskette. The sections of this chapter are independent and thus may be covered in any order or omitted.

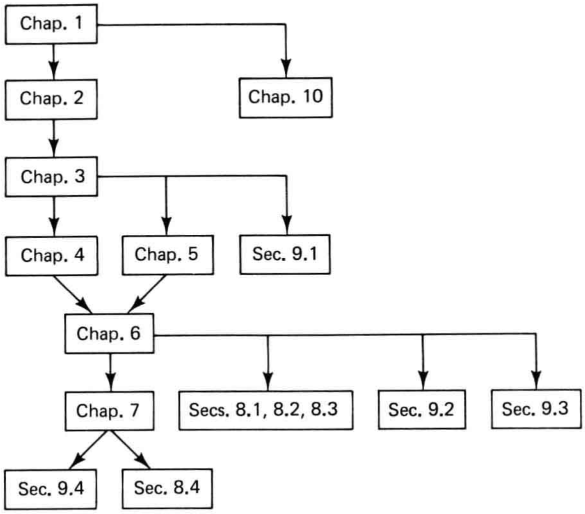
Chapter 10 introduces linear programming. The development is divided into two parts. The first part introduces some terminology and the graphical method. The second part describes the simplex method.

Complex numbers: On the advice of several reviewers, we decided to include examples where complex numbers play an important role, especially when eigenvalues are considered. For example, harmonic motion is illustrated as a case when complex eigenvalues are particularly important. Also, Gerschgorin's theorem takes on a different geometrical interpretation if complex eigenvalues are considered. Appendix A is included to establish the elementary properties of complex numbers. No interruption of the flow of the material will occur if complex numbers are entirely deleted from the presentation.

Notation: The results of problems preceded by a dagger (†) are used in subsequent sections.

Dependencies within the material: All the applications are independent of one another and of the rest of the material. In Section 5.3, the classical adjoint is used only in the application to cryptography. The following flowchart indicates the other dependencies within the material.

Numbering: We have numbered our theorems, lemmas, and corollaries by chapter



and section. For example, Theorem 3.2.1 is the first theorem in the second section of Chapter 3. Tables and figures are each numbered similarly. Examples are numbered sequentially within each section.

We would like to express our thanks to the following people who carefully reviewed our manuscript and made many helpful suggestions: Carl C. Cowen, Mathematics Department, Purdue University; Vincent Giambalvo, Mathematics Department, University of Connecticut; Terry L. Herdman, Mathematics Department, Virginia Polytechnic Institute and State University; Kenneth Kalmanson, Computer Science Department, Montclair State College; Robert H. Lohman, Mathematical Sciences Department, Clemson University.

In addition, special thanks go to the staff of Prentice-Hall, especially Fay Ahuja, production editor, and Robert Sickles, mathematics editor. We are also grateful to Ken Hirschel for proofreading the galleys.

Finally, we are indebted to our colleagues and students for their insightful comments and their encouragement.

SHF
AJI

Contents

1

Introduction to the Vector Space Properties of R^n and Systems of Linear Equations

In this chapter we introduce the *vector space* properties of Euclidean n-space. Although the notion of *vector* was introduced in the nineteenth century primarily by the Irish mathematician W. R. Hamilton, its usefulness in real-world applications, particularly in physics, was not recognized until the twentieth century. More recently, the important properties of vectors have been exploited in such areas as the social and biological sciences, as well as in statistics.

We begin this chapter with the most important Euclidean space, namely, the plane. It is in this space that we can take advantage of our ability to visualize many of the geometric properties of vectors. With this introduction in place, it will be easier to understand the properties of vectors in Euclidean n-space.

1.1 VECTORS IN R^2

Certain physical quantities, such as length, area, speed, and mass, can be described by a single number or magnitude. However, others, such as velocity and force, require both a magnitude and a direction.

Example 1 The *velocity* of an object is described by giving its speed and direction. The *speed*, a nonnegative number, is the magnitude of the velocity. Suppose that $\mathbf{x}$ denotes the velocity of an object traveling 20 miles per hour in a northeast direction. Geometrically, $\mathbf{x}$ can be represented as an arrow, that is, a directed line segment of length 20 which points northeast (see Figure 1.1.1).

Figure 1.1.1

Example 2 Let **f** be the force exerted by the weight of a 200-pound man. The magnitude of this force is 200 and the direction is toward the center of the earth. Geometrically, this force can be represented as an arrow of length 200 which points from the man to the earth's center (see Figure 1.1.2).

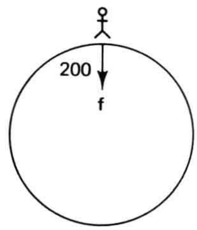


Figure 1.1.2

Any quantity determined by both a magnitude and a direction is called a ***vector***. The velocity and force described in Examples 1 and 2, respectively, are each examples of vectors. Acceleration is another example of a vector. Two vectors are considered to be ***equal*** if they have the same direction and magnitude.

A vector may be represented geometrically as an arrow or directed line segment. The line segment is pointing in the direction of the vector, and its length is the magnitude of the vector. The tail and head of this arrow are called the ***initial*** and ***terminal points***, respectively. The arrow whose initial point is A and whose terminal point is B will be denoted by $\overline{AB}$.

Referring to Figure 1.1.3, we see that $\overline{AB}$ and $\overline{CD}$ have the same magnitude and direction. Therefore, they represent the same vector. The fact that different directed line segments may represent the same vector is very useful. For example, the vector that represents the velocity of a car traveling northeast at 30 miles per hour on a particular road also represents the velocity of that car if it were traveling at the same speed and direction on another road. The arrow $\overline{PQ}$ is pointing in the same direction

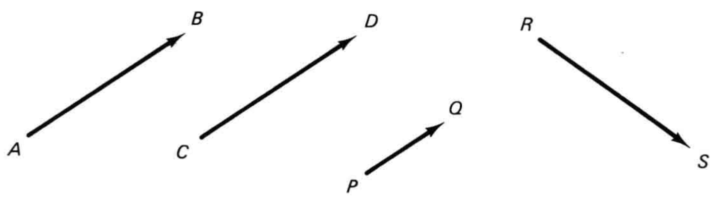


Figure 1.1.3

as $\overline{AB}$ but it is shorter and so it does not represent the same vector as $\overline{AB}$. The arrow $\overline{RS}$ has the same length as $\overline{AB}$ but it is pointing in a different direction. Therefore, it does not represent the same vector as $\overline{AB}$.

Two directed line segments such as $\overline{AB}$ and $\overline{CD}$ in Figure 1.1.3 which determine the same vector, that is, which have the same direction and are of equal length, are called ***equivalent***. The vector they both represent is denoted by either $\overrightarrow{AB}$ or $\overrightarrow{CD}$. We may write

$$\overrightarrow{AB} = \overrightarrow{CD}$$

Vectors in the Plane

It is useful to consider vectors in the context of a rectangular coordinate system. Consider the familiar coordinate system of the xy-plane. We identify a point in the plane with its coordinates. $O = (0, 0)$ is called the ***origin*** of the system. Consider a vector $\overrightarrow{PQ}$ in the xy-plane. Suppose that $P = (a, b)$ and $Q = (c, d)$ (see Figure 1.1.4). Setting $C = (c - a, d - b)$, $A = (c - a, 0)$, and $R = (c, b)$, we see that the (right) triangles OCA and PQR have legs of equal length and so are congruent. Therefore, the vectors $\overrightarrow{OC}$ and $\overrightarrow{PQ}$ have the same magnitude. Since $\overrightarrow{OA}$ and $\overrightarrow{PR}$ have the same direction, it follows that $\overrightarrow{OC}$ and $\overrightarrow{PQ}$ also have the same direction and hence are equal. Summarizing, we have:

For any points $P = (a, b)$ and $Q = (c, d)$ in the plane, the vector $\overrightarrow{PQ}$ can be identified with the vector $\overrightarrow{OC}$, where $C = (c - a, d - b)$.

We call $c - a$ the ***first component*** (or ***x-component***) and $d - b$ the ***second component*** (or ***y-component***) of the vector $\overrightarrow{PQ}$. In this manner, any vector $\mathbf{x}$ in the xy-plane can be associated with a unique ordered pair of real numbers, namely, its components. In our notation above, if $\mathbf{x} = \overrightarrow{PQ}$, we may also write $\mathbf{x} = (c - a, d - b)$. Conversely, given any point $D = (p, q)$ in the xy-plane, the vector $\mathbf{y} = \overrightarrow{OD}$ has components p and q and we may write $\mathbf{y} = (p, q)$. Thus, there is a one-to-one correspondence between vectors in the plane and ordered pairs. We also denote the set of all vectors in the plane by R^2.

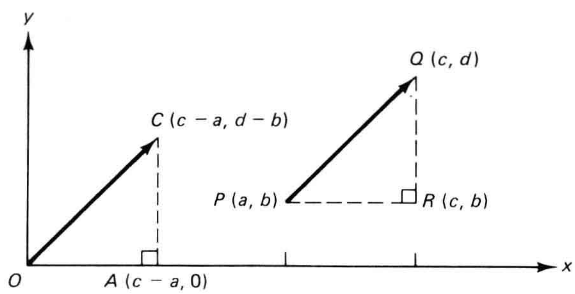


Figure 1.1.4

Some authors use a different notation for points than for vectors. Usually, the context in which an ordered pair is used will make it clear whether we mean a point or a vector.

From this definition of components, we see that two vectors are equal if and only if their corresponding components are equal.

Example 3 Let $A = (2, 3)$, $B = (5, 2)$, $C = (3, 8)$, and $D = (6, 7)$. We wish to show that $\overrightarrow{AB} = \overrightarrow{CD}$. Clearly, the first component of $\overrightarrow{AB}$ is $5 - 2 = 3$, and the second component is $2 - 3 = -1$. Similarly, the first and second components of $\overrightarrow{CD}$ are 3 and -1, respectively. Therefore, $\overrightarrow{AB} = \overrightarrow{CD}$.

Consider again the velocity vector $\mathbf{x}$ of Example 1. If the positive direction of the y-axis points north and the initial point of the directed line segment representing $\mathbf{x}$ is placed at the origin, it can be shown using trigonometry that the terminal point of $\mathbf{x}$ is the ordered pair $(10\sqrt{2}, 10\sqrt{2})$. Thus, these coordinates are the components of the vector $\mathbf{x}$ [see Figure 1.1.5(a)], that is, $\mathbf{x} = (10\sqrt{2}, 10\sqrt{2})$.

The components of the force $\mathbf{f}$ of Example 2 depend on the orientation of the coordinate system. If the coordinate system is oriented so that the man is at the origin and the negative direction of the y-axis points toward the earth's center, the first component of $\mathbf{f}$ is 0 and the second is -200 [see Figure 1.1.5(b)]. So $\mathbf{f} = (0, -200)$.

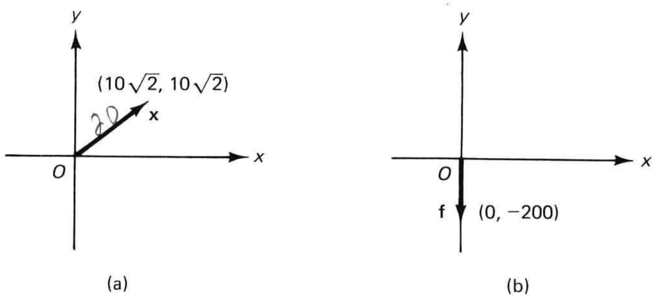


Figure 1.1.5

Vectors have uses other than to represent physical quantities such as velocity and force. For example, a vector might be constructed for each person in a group in which the first component gives the height (in feet) and the second component gives the weight (in pounds). In this case, we have as many vectors as there are people in the group. In the study of Markov chains in Chapter 6, the components represent particular probabilities.

Example 4 Consider a system of two linear equations in two unknowns:

$$
\begin{aligned}
2x + y &= 0 \\
3x - 2y &= 7
\end{aligned}
$$

We shall learn how to solve such systems later in this chapter. It is easy to verify that $x = 1$ and $y = -2$ is a solution to this system. Rather than expressing this solution by two equations, we can make use of one (vector) equation using the vector $\mathbf{y} = (x, y) = (1, -2)$. This notation will be of considerable use in Chapter 2.

Vectors would be of little interest if they were used merely to describe properties, as in the examples above. They are useful because there is an arithmetic defined on them. This arithmetic allows us to make certain calculations in order to draw new conclusions about vector properties.

Suppose that a train is traveling at a speed of 60 miles per hour while a passenger is walking toward the front of the train at 2 miles per hour. Then the velocity of the passenger relative to the ground is $60 + 2 = 62$. Similarly, if the passenger is walking toward the back of the train at 2 miles per hour, her velocity relative to the train is -2 miles per hour, and her velocity relative to the ground is $60 + (-2) = 58$ miles per hour. It is clear that the operation of addition can be used to combine velocities.

Now we consider a comparable situation in the xy-plane. Suppose that a pilot aims an airplane in a particular direction and that the airplane moves at a certain "airspeed," that is, the speed relative to the surrounding air. Then the airplane has a velocity $\mathbf{x}$ relative to the air around it. Now suppose that at the same time the air or wind is moving relative to the ground with a given speed and direction. Suppose that the motion of the wind has a velocity $\mathbf{y}$. We combine the velocity of the airplane relative to the air and the velocity of the wind relative to the ground to determine the velocity of the airplane relative to the ground. Call this combined velocity $\mathbf{z}$. Let us see how to compute $\mathbf{z}$. Imposing a coordinate system whose positive x-axis points east and whose positive y-axis points north, suppose that $\mathbf{x} = (a, b)$ and $\mathbf{y} = (c, d)$. Then a and c represent the horizontal (east-west) components of $\mathbf{x}$ and $\mathbf{y}$, respectively. As with the example of the passenger on the train, it follows from physics that these components can be added to yield the horizontal component of the combined velocity, namely, $a + c$. Similarly, $b + d$ is the vertical (north-south) component of the combined velocity. Therefore, $\mathbf{z} = (a + c, b + d)$.

The operation of combining two vectors by adding the corresponding components is called ***vector addition***.

Definition Let $\mathbf{x} = (a, b)$ and $\mathbf{y} = (c, d)$ be two vectors in R^2. We define the ***sum*** of $\mathbf{x}$ and $\mathbf{y}$, denoted by $\mathbf{x} + \mathbf{y}$, to be the vector in R^2 defined by

$$\mathbf{x} + \mathbf{y} = (a + c, b + d)$$

For example, if $\mathbf{x} = (1, -2)$ and $\mathbf{y} = (2, 4)$, then $\mathbf{x} + \mathbf{y} = (3, 2)$.

There is also a geometric interpretation of vector addition which is quite useful. Using facts about congruent (right) triangles (see Figure 1.1.6), it follows that if the sides of a parallelogram are determined by the vectors $\mathbf{x}$ and $\mathbf{y}$, the sum $\mathbf{z} = \mathbf{x} + \mathbf{y}$ is given by a diagonal of the parallelogram. This result is called the ***parallelogram law of vector addition***.

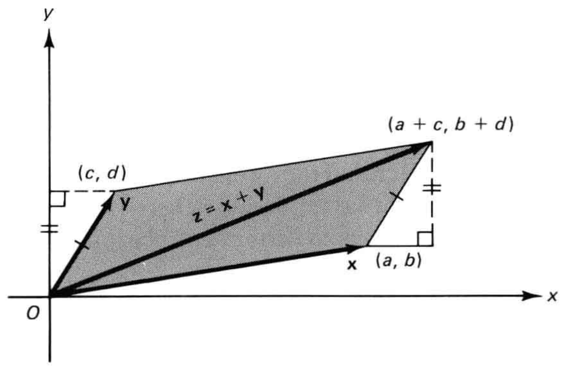


Figure 1.1.6

It is interesting to consider the parallelogram law in the context of our airplane example. Recall that the velocity (relative to the air) of the airplane is given by **x**, the wind velocity is given by **y**, and the resulting velocity (relative to the ground) of the airplane is given by **z**. If **x** and **y** are in approximately the same direction, then the airplane's speed (relative to the ground), which is given by the magnitude of **z**, is increased [see Figure 1.1.7(a)]. On the other hand, if **x** and **y** are approximately in opposite directions, then the magnitude of **z** is decreased [see Figure 1.1.7(b)]. Of course, this agrees with our intuition about the effect of wind on the velocity of an object.

The same physical interpretation that we have given to the sum of velocity vectors may also be applied to the sum of vectors that represent forces [see Exercise 2(b) and (d)].

Vector addition may be easily extended to sums of more than two vectors. For example, if **x**, **y**, and **u** are three vectors, we may define their sum **z** as $\mathbf{z} = (\mathbf{x} + \mathbf{y}) + \mathbf{u}$. It is easy to see that **z** may also be determined by $\mathbf{x} + (\mathbf{y} + \mathbf{u})$. For example, if $\mathbf{x} = (a, b)$, $\mathbf{y} = (c, d)$, and $\mathbf{u} = (e, f)$, then

$$\begin{aligned}
\mathbf{z} &= (\mathbf{x} + \mathbf{y}) + \mathbf{u} \\
&= (a + c, b + d) + (e, f) \\
&= ((a + c) + e, (b + d) + f) \\
&= (a + (c + e), b + (d + f)) \\
&= (a, b) + (c + e, d + f) \\
&= \mathbf{x} + (\mathbf{y} + \mathbf{u})
\end{aligned}$$

The result above is the ***associative law for vector addition***. It allows us to omit parentheses and write $\mathbf{z} = \mathbf{x} + \mathbf{y} + \mathbf{u}$. The short proof of this fact illustrates a technique which will be employed again. Namely, the associative law of addition of real numbers has been used to establish the associative law of vector addition. The fact