LINEAR ALGEBRA:

A PATHWAY TO ABSTRACT MATHEMATICS

3rd Ed.

G. Viglino

Ramapo College of New Jersey

September 2023

CONTENTS

CHAPTER 1 MATRICES AND SYSTEMS OF LINEAR EQUATION

CHAPTER 2 VECTOR SPACES

CHAPTER 3 BASES AND DIMENSION

CHAPTER4 LINEARITY

CHAPTER 5 MATRICES AND LINEAR MAPS

PREFACE

There is no mathematical ramp that will enable you to continuously inch your way higher and higher in mathematics. The climb calls for a ladder consisting of discrete steps designed to take you from one mathematical level to another. You are about to take an important step on that ladder, one that will take you to a plateau where mathematical abstraction abounds. Linear algebra rests on a small number of axioms (accepted rules, or “laws”), upon which a beautiful and practical theory emerges.

Technology can be used to reduce the time needed to perform essential but routine tasks. We feature the TI-84+ calculator, but any graphing utility or Computer Algebraic System will do. The real value of whatever technological tool you use is that it will free you to spend more time on the development and comprehension of the theory and its applications. In any event, if you haven’t already discovered in other courses:

MATHEMATICS DOES NOT RUN ON BATTERIES

Systems of linear equations are introduced and analyzed in Chapter 1. Graphing utilities can be used to solve such systems, but understanding what those solutions represent plays a dominant role throughout the text.

We begin Chapter 2 by sowing the seeds for vector spaces in the fertile real number field, where they soon blossom into the concept of an abstract vector. The remainder of Chapter 2 and all of Chapter 3 are dedicated to a study of vector spaces in isolation. Functions from one vector space to another which, in a sense, respect the algebraic structure of those spaces are investigated in Chapters 4 and 5. The sixth chapter focuses on Eigenvalues and Eigenvectors, along with some of their important applications.

The first six chapters may provide a full plate for most one-semester courses. If not, then Chapter 7 (on inner product spaces) is offered for dessert.

We have made every effort to provide a leg-up for the step you are about to take. Our primary goal was to write a readable book, without compromising mathematical integrity. Along the way, you will encounter numerous Check Your Understanding boxes designed to challenge your understanding of each newly introduced concept. Complete solutions to the problems in those boxes appear in Appendix B, but please don’t be in too much of a hurry to look at our solutions. You should **TRY** to solve the problems on your own, for it is only through **ATTEMPTING** to solve a problem that one grows mathematically. In the words of Descartes:

> ***We never understand a thing so well, and make it our own, when we learn it from another, as when we have discovered it for ourselves.***

CHAPTER 1

MATRICES AND SYSTEMS OF LINEAR EQUATIONS

Much of the development of linear algebra calls for the solution and interpretation of systems of linear equations. While the “solution part” can be relegated to a calculator, the “interpretation part” cannot. We focus on the solution-part of the process in Section 1, and on the more important interpretation-part in Sections 2.

§1. SYSTEMS OF LINEAR EQUATIONS

To solve the system of equations:

$$\left.\begin{aligned} 2x + 4y - 4z &= 6 \\ 2x + 6y + 4z &= 0 \\ x + y + 2z &= -2 \end{aligned}\right\}$$

is to determine values of the **variables** (or **unknowns**) x, y, and z for which each of the three equations is satisfied. You certainly solved such systems in earlier courses, and if you take the time to solve the above system, you will find that it has but one solution: $x = -1, y = 1, z = -1$. We can also say that the **three-tuple** $(-1, 1, -1)$ is a solution of the given system of equation, and that $\{(-1, 1, -1)\}$ is its **solution set**. In general:

Brackets are used to denote sets. In particular, $\{(-1, 1, -1)\}$ denotes the set containing but one element—the element $(-1, 1, -1)$.

An (ordered) **n-tuple** is an expression of the form $(c_1, c_2, \ldots, c_n)$, where each c_i is a real number (written $c_i \in \Re$), for $1 \le i \le n$.

We say that the n-tuple $(c_1, c_2, \ldots, c_n)$ is a **solution** of the system of ***m*** **equations in** ***n*** **unknowns**:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n &= b_m \end{aligned}\right\}$$

The x_is denote **variables** (or **unknowns**), while the a_{ij}’s and b_i’s are **constants** (or **scalars**).

if each equation in the system is satisfied when c_i is substituted for x_i, for $1 \le i \le n$.

The set of all solutions of a system of equations is said to be the **solution set** of that system.

EQUIVALENT SYSTEMS OF EQUATIONS

Consider the system of equations:

$$\left.\begin{array}{rcl} -3x+y &=& 2 \\ \dfrac{x}{2}+2y &=& \dfrac{1}{2} \end{array}\right\}$$

As you know, you can perform certain operations on that system which will not alter its solution set. For example, you can:

(1) Interchange the order of the equations:

$$\left.\begin{array}{rcl} -3x+y &=& 2 \\ \dfrac{x}{2}+2y &=& \dfrac{1}{2} \end{array}\right\} \rightarrow \left.\begin{array}{rcl} \dfrac{x}{2}+2y &=& \dfrac{1}{2} \\ -3x+y &=& 2 \end{array}\right\}$$

(2) Multiply both sides of the resulting top equation by 2:

$$\left.\begin{array}{rcl} \dfrac{x}{2}+2y &=& \dfrac{1}{2} \\ -3x+y &=& 2 \end{array}\right\} \rightarrow \left.\begin{array}{rcl} x+4y &=& 1 \\ -3x+y &=& 2 \end{array}\right\}$$

(3) Multiply the resulting top equation by 3 and add it to the bottom equation:

$$\left.\begin{array}{rcl} x+4y &=& 1 \\ -3x+y &=& 2 \end{array}\right\} \rightarrow \left.\begin{array}{rcl} x+4y &=& 1 \\ 13y &=& 5 \end{array}\right\}$$

multiply by 3

$$\begin{array}{rrcl} & 3x+12y &=& 3 \\ & -3x+y &=& 2 \\ \hline \text{add:} & 13y &=& 5 \end{array}$$

You used this third maneuver a lot when eliminating a variable from a given system of equations For example:

$$\left.\begin{array}{lrcl} \text{(i)} & x+3y-z &=& 1 \\ (2) & 2x-5y+3z &=& 3 \\ (3) & -3x+y+2z &=& 2 \end{array}\right\}$$

multiply (1) by -2 and add it to (2)

$$\left.\begin{array}{rcl} -11y+5z &=& 1 \\ 10y-z &=& 5 \end{array}\right\}$$

multiply (1) by 3 and add it to (3)

The above three operations, are said to be **elementary equation operations**:

ELEMENTARY OPERATIONS ON SYSTEMS OF LINEAR EQUATIONS
Interchange the order of any two equations in the system.
Multiply both sides of an equation by a nonzero number.
Add a multiple of one equation to another equation.

EQUIVALENT SYSTEM OF EQUATIONS

Two systems of equations sharing a common solution set are said to be **equivalent.** As you may recall:

THEOREM 1.1 Performing any sequence of elementary operations on a system of linear equations will result in an equivalent system of equations.

AUGMENTED MATRICES

Matrices are arrays of numbers (or expressions representing numbers) arranged in rows and columns:

$$\begin{bmatrix} 2 & 13 & 4 \\ -9 & 7 & 3 \end{bmatrix} \qquad \begin{bmatrix} 4 & 7 \\ 10 & 6 \\ -8 & \sqrt{3} \end{bmatrix} \qquad \begin{bmatrix} 1 & 7 & 0 \\ 3 & 6 & 5 \\ 11 & 2 & -12 \end{bmatrix} \qquad \begin{bmatrix} 12 & 7 & 4 \end{bmatrix} \qquad \begin{bmatrix} 10 \\ 1 \\ 9 \end{bmatrix}$$

(i) (ii) (iii) (iv) (v)

Matrix (i) contains 2 rows and 3 columns and it is said to be a 2×3 (two-by-three) matrix. Similarly, (ii) is a 3×2 matrix, and (iii) is a 3×3 matrix (a **square matrix**). In general, an $m \times n$ matrix is a matrix consisting of m rows and n columns. In particular, (iv) is a 1×3 matrix (a **row matrix**), and (v) is a 3×1 matrix (a **column matrix**).

It is often convenient to represent a system of equations in a more compact matrix form. The rows of the matrix in Figure 1.1(b), for example, concisely represents the equations in Figure 1.1(a). Note that the variables x, y, and z are suppressed in the matrix form, and that the vertical line recalls the equal sign in the equations. Such a matrix is said to be the **augmented matrix** associated with the given system of equations.

AUGMENTED MATRIX

$$\left.\begin{aligned} 2x + 4y - 4z &= 6 \\ 2x + 6y + 4z &= 0 \\ x + y + 2z &= -2 \end{aligned}\right\} \longleftrightarrow \left[\begin{array}{ccc|c} 2 & 4 & -4 & 6 \\ 2 & 6 & 4 & 0 \\ 1 & 1 & 2 & -2 \end{array}\right]$$

System of Equations (a) **Augmented Matrix** (b)

Figure 1.1

Switching two equations in a system of equations results in the switching of the corresponding rows in the associated augmented matrix. Indeed, each of the three previously introduced elementary equation operations corresponds with one of the following elementary matrix row operations:

ELEMENTARY MATRIX ROW OPERATIONS
Interchange the order of any two rows in the matrix.
Multiply each element in a row of the matrix by a nonzero number.
Add a multiple of one row of the matrix to another row of the matrix.

The following terminology is motivated by Theorem 1.1:

DEFINITION 1.1
EQUIVALENT MATRICES

Two matrices are **equivalent** if one can be derived from the other by performing elementary row operations.

Here is where we are at this point:

A system of linear equations can be represented by an augmented matrix, and every augmented matrix represents a system of linear equations. Moreover:

SYSTEMS OF EQUATIONS ASSOCIATED WITH EQUIVALENT AUGMENTED MATRICES ARE THEMSELVES EQUIVALENT (SAME SOLUTION SET).

And here is where we are going:

Suppose you want to solve the system of equations **[1]** in Figure 1.2. Assume, for the time being, that you can go from its augmented matrix (**[2]**) to matrix **[3]**, via elementary row operations. System **[4]**, which is associated with the augmented matrix **[3]**, is easily seen to have the solution: $(x = -1, y = 1, z = -1)$. But this must also be the solution of system **[1]**, since the two systems of equations are also equivalent!

$$\begin{tikzcd}
\textbf{[1]} \quad \left.\begin{aligned} 2x+4y-4z &= 6 \\ 2x+6y+4z &= 0 \\ x+y+2z &= -2 \end{aligned}\right\} \arrow[r, "\textbf{augmented matrix}"] \arrow[d, leftrightarrow, "\textbf{Same Solution Set}"'] & \left[\begin{array}{ccc|c} 2 & 4 & -4 & 6 \\ 2 & 6 & 4 & 0 \\ 1 & 1 & 2 & -2 \end{array}\right] \quad \textbf{[2]} \arrow[d, "\textbf{Elementary Row Operations}"] \\
\textbf{[4]} \quad \left.\begin{aligned} x+0y+0z &= -1 \\ 0x+y+0z &= 1 \\ 0x+0y+z &= -1 \end{aligned}\right\} & \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right] \quad \textbf{[3]} \arrow[l, "\textbf{system of equations}"']
\end{tikzcd}$$

Figure 1.2

The remainder of this section is designed to illustrate a method which can be used to go from matrix **[2]** of Figure 1.2 to matrix **[3]**, via elementary row operations.

Pivoting About a Pivot Point

Capital letters are used to represent matrices, and double subscripted lower case letters for their entries; as in:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$$

.

Note that the first subscript of the element a_{ij} denotes its row: i, and the second subscript, its column: j. For example, if:

$$A = [a_{ij}] = \begin{bmatrix} 2 & 7 & 6 & 2 \\ 3 & 6 & 3 & 1 \\ -1 & 2 & 2 & 3 \end{bmatrix}$$

then $a_{12} = 7, a_{21} = 3, a_{34} = 3$, and so on.

PIVOT POINT
PIVOTING

In the next example, we specify a location in a given matrix (called the **pivot-point**), which contains a non-zero entry. We then illustrate a process (called **pivoting**) designed to turn the given matrix into an equivalent matrix with a 1 in the pivot-point, and with each entry above or below the pivot-point equal to 0. It is a routine process that plays a dominant role in a number of matrix applications, so please make sure that you understand it fully.

The following notation will be used to represent elementary row operations:

ELEMENTARY ROW OPERATION	**NOTATION**
Switch row i with row j:	$R_i \leftrightarrow R_j$
Multiply each entry in row i by a nonzero number c:	$cR_i \rightarrow R_i$
Multiply each entry in row i by a number c, and add the resulting row to row j:	$cR_i + R_j \rightarrow R_j$

EXAMPLE 1.1

Pivot the matrix:

$$A = [a_{ij}] = \begin{bmatrix} \mathbf{2} & 6 & -4 & 2 \\ 3 & 6 & 3 & 15 \\ -1 & 2 & 2 & 3 \end{bmatrix}$$

about the pivot point a_{11} with pivot entry **2**, and then again about the pivot point a_{22} of the resulting equivalent matrix.

SOLUTION:

Step 1. Get a **1** in the pivot-point position by multiplying each entry in row 1 by $\frac{1}{2}$:

$$\begin{bmatrix} \boxed{\mathbf{2}} & 6 & -4 & 2 \\ 3 & 6 & 3 & 15 \\ -1 & 2 & 2 & 3 \end{bmatrix} \xrightarrow{\frac{1}{2}R_1 \rightarrow R_1} \begin{bmatrix} \boxed{\mathbf{1}} & 3 & -2 & 1 \\ 3 & 6 & 3 & 15 \\ -1 & 2 & 2 & 3 \end{bmatrix}$$

$$\begin{array}{rl} -3\times(1\ \ 3\ \ -2\ \ 1) & \\ \xrightarrow{-3R_1} & -3\ \ -9\ \ 6\ \ -3 \\ R_2 \longrightarrow & 3\ \ 6\ \ 3\ \ 15 \\ \hline -3R_1+R_2: & \boxed{\mathbf{0}\ \ -3\ \ 9\ \ 12} \end{array}$$

Step 2. Get **0**'s below (there is no above) the pivot point position. Multiply row 1 by -3 and add it to row 2 (see margin), and then multiply row 1 by 1 and add it to row 3:

$$\begin{bmatrix} 1 & 3 & -2 & 1 \\ 3 & 6 & 3 & 15 \\ -1 & 2 & 2 & 3 \end{bmatrix} \xrightarrow{-3R_1+R_2\to R_2} \begin{bmatrix} 1 & 3 & -2 & 1 \\ \mathbf{0} & -3 & 9 & 12 \\ -1 & 2 & 2 & 3 \end{bmatrix} \xrightarrow{1R_1+R_3\to R_3} \begin{bmatrix} 1 & 3 & -2 & 1 \\ \mathbf{0} & -3 & 9 & 12 \\ \mathbf{0} & 5 & 0 & 4 \end{bmatrix}$$

$$\boxed{\begin{array}{rrrr} 1 & 3 & -2 & 1 \\ -1 & 2 & 2 & 3 \\ \hline \mathbf{0} & 5 & 0 & 4 \end{array}}$$

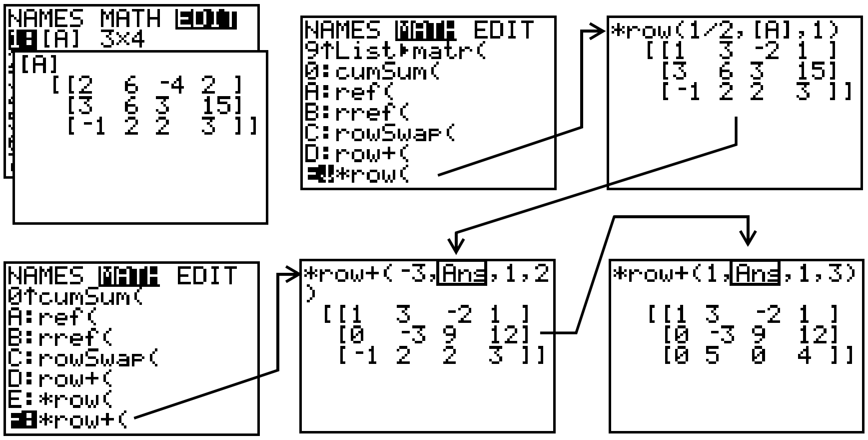
Repeating the above two-step procedure, we now pivot about $a_{22} = -3$:

$$\begin{bmatrix} 1 & 3 & -2 & 1 \\ \mathbf{0} & -3 & 9 & 12 \\ \mathbf{0} & 5 & 0 & 4 \end{bmatrix} \xrightarrow{-\frac{1}{3}R_2\to R_2} \begin{bmatrix} 1 & 3 & -2 & 1 \\ \mathbf{0} & 1 & -3 & -4 \\ \mathbf{0} & 5 & 0 & 4 \end{bmatrix} \xrightarrow{-3R_2+R_1\to R_1} \begin{bmatrix} 1 & \mathbf{0} & 7 & 13 \\ \mathbf{0} & 1 & -3 & -4 \\ \mathbf{0} & 5 & 0 & 4 \end{bmatrix} \xrightarrow{-5R_2+R_3\to R_3} \begin{bmatrix} 1 & \mathbf{0} & 7 & 13 \\ \mathbf{0} & 1 & -3 & -4 \\ \mathbf{0} & \mathbf{0} & 15 & 24 \end{bmatrix}$$

$$\boxed{\begin{array}{rrrr} 0 & -3 & 9 & 12 \\ 1 & 3 & -2 & 1 \\ \hline 1 & \mathbf{0} & 7 & 13 \end{array}} \qquad \boxed{\begin{array}{rrrr} 0 & -5 & 15 & 20 \\ 0 & 5 & 0 & 4 \\ \hline \mathbf{0} & \mathbf{0} & 15 & 24 \end{array}}$$

GRAPHING CALCULATOR GLIMPSE 1.1

We utilize a graphing calculator to perform the first of the two pivoting processes in the above example, and invite you to use your calculator to address the other pivoting process.



```
[A]
  [[2  6 -4  2 ]
   [3  6  3  15]
   [-1 2  2  3 ]]
```

```
*row(1/2,[A],1)
  [[1  3 -2  1 ]
   [3  6  3  15]
   [-1 2  2  3 ]]
```

```
*row+(-3,Ans,1,2
)
  [[1  3  -2  1 ]
   [0  -3  9  12]
   [-1  2  2  3 ]]
```

```
*row+(1,Ans,1,3)
  [[1  3  -2  1 ]
   [0  -3  9  12]
   [0  5   0  4 ]]
```

The TI-84+ calculator is featured throughout the text.

CHECK YOUR UNDERSTANDING 1.1

Pivot about $a_{33} = \mathbf{15}$ to go from: $\begin{bmatrix} 1 & 0 & 7 & 13 \\ 0 & 1 & -3 & -4 \\ 0 & 0 & \mathbf{15} & 24 \end{bmatrix}$ to $\begin{bmatrix} ? & ? & \mathbf{0} & ? \\ ? & ? & \mathbf{0} & ? \\ ? & ? & \mathbf{1} & ? \end{bmatrix}$

Answer: See page B-1.

ROW-REDUCED-ECHELON FORM

A matrix may have many different equivalent forms. Here is the nicest of them all:

DEFINITION 1.2
ROW-REDUCED ECHELON FORM

A matrix is in **row-reduced-echelon form** when it satisfies the following three conditions:

(i) The first non-zero entry in any row is 1 (called its **leading-one),** and all of the entries above or below that leading-one are 0.

(ii) In any two successive rows, not consisting entirely of zeros, the leading-one in the lower row appears further to the right than the leading-one in the row above it.

(iii) All of the rows that consist entirely of zeros are at the bottom of the matrix.

A matrix satisfying (i), (ii) and a slightly weaker form of (i):

> The first non-zero entry in any row is 1**,** and the entries **below** (only) that leading-one are 0

is said to be in **row-echelon form**.

These three matrices are in row-reduced-echelon form:

$$A = \begin{bmatrix} 1 & 0 & 0 & 12 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 0 \end{bmatrix}$$

The matrices

$$\begin{bmatrix} 1 & 0 & 0 & 12 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 1 & 9 & 0 & 12 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

are in row-echelon form

CHECK YOUR UNDERSTANDING 1.2

Determine if the given matrix is in row-reduced-echelon form. If not, list the condition(s) of Definition 1.2 which are not satisfied.

$$(a)\ \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (b)\ \begin{bmatrix} 1 & 0 & -3 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (c)\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 4 \end{bmatrix} \quad (d)\ \begin{bmatrix} 0 & 1 & -3 & 0 & 5 \\ 0 & 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Answer: Yes: (a), (c), and (d).
No: (b) [fails (ii)]

Though a bit tedious, reducing a matrix to its row-reduced-echelon form is a routine task. Just focus on getting those all-important **leading-ones** (which are to be positioned further to the right as you move down), with **zeros above and below** them. Consider the following example:

EXAMPLE 1.2 Perform elementary row operations to obtain the row-reduced-echelon form for the matrix:

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

SOLUTION: Leading-one Step. We could divide the first row by 2 to get a leading-one in that row, but choose to switch the first row and third row instead:

$$\begin{bmatrix} 2 & 4 & -4 & 6 \\ 2 & 6 & 4 & 0 \\ 1 & 1 & 2 & -2 \end{bmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ 2 & 6 & 4 & 0 \\ 2 & 4 & -4 & 6 \end{bmatrix}$$

Zeros-above-and-below Step:

$$\begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ 2 & 6 & 4 & 0 \\ 2 & 4 & -4 & 6 \end{bmatrix} \xrightarrow{\begin{array}{c} \boxed{\begin{array}{rrrr} -2 & -2 & -4 & 4 \\ 2 & 6 & 4 & 0 \\ \hline \mathbf{0} & 4 & 0 & 4 \end{array}} \\ \updownarrow \\ -2R_1+R_2 \to R_2 \end{array}} \begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ \mathbf{0} & 4 & 0 & 4 \\ 2 & 4 & -4 & 6 \end{bmatrix} \xrightarrow{\begin{array}{c} \boxed{\begin{array}{rrrr} -2 & -2 & -4 & 4 \\ 2 & 4 & -4 & 6 \\ \hline \mathbf{0} & 2 & -8 & 10 \end{array}} \\ \updownarrow \\ -2R_1+R_3 \to R_3 \end{array}} \begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ \mathbf{0} & 4 & 0 & 4 \\ \mathbf{0} & 2 & -8 & 10 \end{bmatrix}$$

Next leading-one Step:

$$\begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ \mathbf{0} & 4 & 0 & 4 \\ \mathbf{0} & 2 & -8 & 10 \end{bmatrix} \xrightarrow{\frac{1}{4}R_2 \to R_2} \begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & 2 & -8 & 10 \end{bmatrix}$$

Zeros-above-and-below Step:

$$\begin{bmatrix} \mathbf{1} & 1 & 2 & -2 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & 2 & -8 & 10 \end{bmatrix} \xrightarrow{\begin{array}{c} \boxed{\begin{array}{rrrr} 0 & -1 & 0 & -1 \\ 1 & 1 & 2 & -2 \\ \hline \mathbf{1} & \mathbf{0} & 2 & -3 \end{array}} \\ \updownarrow \\ -1R_2+R_1 \to R_1 \end{array}} \begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & -3 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & 2 & -8 & 10 \end{bmatrix} \xrightarrow{\begin{array}{c} \boxed{\begin{array}{rrrr} 0 & -2 & 0 & -2 \\ 0 & 2 & -8 & 10 \\ \hline \mathbf{0} & \mathbf{0} & -8 & 8 \end{array}} \\ \updownarrow \\ -2R_2+R_3 \to R_3 \end{array}} \begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & -3 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & \mathbf{0} & -8 & 8 \end{bmatrix}$$

Next leading-one Step:

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & -3 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & \mathbf{0} & -8 & 8 \end{bmatrix} \xrightarrow{-\frac{1}{8}R_3 \to R_3} \begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & -3 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 \end{bmatrix}$$

Zeros-above-and-below Step:

$$\boxed{\begin{array}{cccc} 0 & 0 & -2 & 2 \\ 1 & 0 & 2 & -3 \\ \hline \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 \end{array}}$$

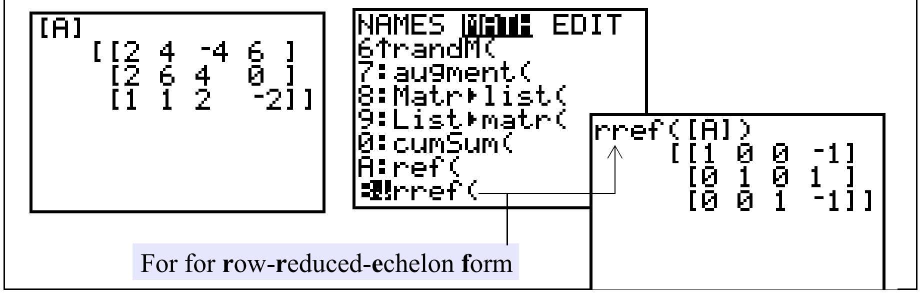
$$\updownarrow$$

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & -3 \\ \mathbf{0} & \mathbf{1} & 0 & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 \end{bmatrix} \xrightarrow{-2R_3 + R_1 \to R_1} \begin{bmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & -1 \\ \mathbf{0} & \mathbf{1} & \mathbf{0} & 1 \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & -1 \end{bmatrix}$$

We are now at a row-reduced-echelon form, and so we stop.

While not difficult, the above example illustrates that obtaining the row-reduced-echelon form of a matrix can be a bit tedious. It's a dirty job, but someone has to do it:

GRAPHING CALCULATOR GLIMPSE 1.2



```
[A]
   [[2 4 -4 6 ]
    [2 6 4  0 ]
    [1 1 2  -2]]
```

```
rref([A])
     [[1 0 0 -1]
      [0 1 0 1 ]
      [0 0 1 -1]]
```

For for **r**ow-**r**educed-**e**chelon **f**orm

In harmony with graphing calculators, we will adopt the notation rref(A) to denote the row-reduced-echelon form of a matrix A.

EXAMPLE 1.3

Solve the system:

$$\left.\begin{aligned} 2x + 4y - 4z &= 6 \\ 2x + 6y + 4z &= 0 \\ x + y + 2z &= -2 \end{aligned}\right\}$$

SOLUTION: All the work has been done:

$$\left.\begin{aligned}2x+4y-4z&=6\\2x+6y+4z&=0\\x+y+2z&=-2\end{aligned}\right\}\xleftrightarrow{\text{augmented matrix}}\begin{array}{c}\begin{array}{ccc}x&y&z\end{array}\\\left[\begin{array}{ccc|c}2&4&-4&6\\2&6&4&0\\1&1&2&-2\end{array}\right]\end{array}\xrightarrow{\text{Example 1.2}}\begin{array}{c}\begin{array}{ccc}x&y&z\end{array}\\\left[\begin{array}{ccc|c}1&0&0&-1\\0&1&0&1\\0&0&1&-1\end{array}\right]\end{array}\leftrightarrow\left.\begin{aligned}x&=-1\\y&=1\\z&=-1\end{aligned}\right\}$$

These two systems of equations are equivalent (same solution sets)

From the above we can easily spot the solution of the given system:

$$(x=-1, y=1, z=-1)$$

CHECK YOUR UNDERSTANDING 1.3

Proceed as in Example 1.3 to solve the given system of equations.

$$\left.\begin{aligned}x+y+z&=6\\3x+2y-z&=4\\3x+y+2z&=11\end{aligned}\right\}$$

Answer:
$x=1, y=2, z=3$

A bit of human-intervention was used in the pivoting process of Example 1.2. If it is "freedom from choice" that you want, then you can use the following algorithm to reduce a given matrix to its row-reduced-echelon form:

Gauss, Karl Friedrich (1777 -1855), the great German mathematician and astronomer.
Wilhelm Jordan (1842-1899) German professor of geodesy.

Gauss-Jordan Elimination Method

Step 1. Locate the left-most column that does not consist entirely of zeros, and pick a nonzero element in that column. Let the position of that chosen element be the pivot-point.

Step 2. Pivot about the pivot-point of Step 1.

Step 3. If necessary, switch the pivot-row with the furthest row above it (nearest the top) that does not already contain a leading-one, to the left of the pivot column.

Step 4. If the matrix is in row-reduced-echelon form, then you are done. If not, return to Step 1, but ignore all rows with established leading-ones for that step of the process.

EXERCISES

Exercises 1-2. Write down the augmented matrix associated with the given system of equations.

1.
$$\left.\begin{aligned} 3x-3y+z &= 2 \\ 5x+5y-9z &= -1 \\ -3x-4y+z &= 0 \end{aligned}\right\}$$

2.
$$\left.\begin{aligned} 2x+3y-4w &= 5 \\ x-4z+w &= -1 \\ x-4y &= 0 \\ -x-y+z+4w &= 9 \end{aligned}\right\}$$

Exercises 3-4. Write down the system of equations associated with the given augmented matrix.

3.
$$\left[\begin{array}{ccc|c} 5 & 1 & 4 & 3 \\ -2 & -3 & 1 & 4 \\ \frac{1}{2} & -1 & 0 & 0 \end{array}\right]$$

4.
$$\left[\begin{array}{cccc|c} 2 & 4 & 1 & 0 & 9 \\ 0 & 5 & 5 & 2 & 2 \\ 2 & 1 & -3 & 8 & 11 \end{array}\right]$$

Exercises 5-8. Perform elementary row operations to obtain the row-reduced echelon form for the given matrix.

5.
$$\begin{bmatrix} 0 & 2 & 4 \\ 1 & 0 & 2 \\ 2 & 4 & 5 \end{bmatrix}$$

6.
$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 2 & 1 & 2 & 1 \\ 0 & 2 & 2 & 4 \end{bmatrix}$$

7.
$$\begin{bmatrix} 0 & 2 & 5 \\ 4 & 4 & 2 \\ 1 & 0 & 3 \\ 2 & 3 & 5 \end{bmatrix}$$

8.
$$\begin{bmatrix} 2 & 3 & 0 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 1 \end{bmatrix}$$

Exercises 9-11. Solve the system of equations corresponding to the given row-reduced-echelon form matrix.

9.
$$\begin{array}{c} \begin{array}{cccc} x & y & z & \end{array} \\ \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 \end{array}\right] \end{array}$$

10.
$$\begin{array}{c} \begin{array}{ccccc} x & y & z & w & \end{array} \\ \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & -3 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}\right] \end{array}$$

11.
$$\begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & \end{array} \\ \left[\begin{array}{ccccc|c} 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{array}\right] \end{array}$$

Exercises 12-15. Proceed as in Example 1.2 to solve the given system of equations.

12.
$$\left.\begin{aligned} x-2y+z &= 1 \\ -3x+5y-2z &= -5 \\ 4x-8y+3z &= 6 \end{aligned}\right\}$$

13.
$$\left.\begin{aligned} x-y-z &= 2 \\ 4x-2y-5z &= -2 \\ -x+3y+6z &= 0 \end{aligned}\right\}$$

14. $$\left.\begin{aligned} 2x+5y-2z &= -1 \\ -4x-y+z &= -4 \\ x+2y-z &= 4 \end{aligned}\right\}$$

15. $$\left.\begin{aligned} 2x-y &= 2z+w+1 \\ w-x &= y \\ 4y+3z &= -\frac{1}{2} \\ x+3y &= w-z \end{aligned}\right\}$$

16. Construct a system of three equations in three unknowns, x, y, and z such that $x = 1, y = 2, z = 3$ is a solution of the system.

17. Construct a system of four equations in four unknowns, x, y, z, and w with solution set $\{(x = 1, y = 2, z = 3, w = 4)\}$.

Exercises 18-20. (Row-Echelon Form) A matrix is said to be in **row-echelon form** if it satisfies all of the conditions of Definition 1.2, except that elements above a leading 1 need not be zero (the entries below leading ones must still be zero). Determine if the matrix is in row-echelon form. If not, indicate why not.

18. $\begin{bmatrix} 1 & 2 & 3 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ 19. $\begin{bmatrix} 1 & 0 & 3 & 2 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ 20. $\begin{bmatrix} 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

Exercises 21-22. Perform elementary row operations to transform the given matrix to the given row-echelon form (see Exercise 18-20).

21. $\begin{bmatrix} 2 & 4 & -4 & 6 \\ 2 & 6 & 4 & 0 \\ 1 & 1 & 2 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & -2 & 3 \\ 0 & 1 & 4 & -3 \\ 0 & 0 & 1 & -1 \end{bmatrix}$ 22. $\begin{bmatrix} 2 & 6 & -4 & 2 \\ 3 & 6 & 3 & 15 \\ -1 & 2 & 2 & 3 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & 1 & -3 & -4 \\ 0 & 0 & 1 & \frac{8}{5} \end{bmatrix}$

Exercises 23-25. Determine the solution set of the system of equations corresponding to the given row-echelon form matrix (see Exercise 18-20).

Note: If you are using a graphing calculator, then you might as well use the row-reduced-echelon command, for that is the most revealing form. If you are doing things by hand, however, you may be able to save some time by going with the row-echelon form.

23. $\begin{array}{ccc|c} x & y & z & \\ 1 & 2 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 2 \end{array}$ 24. $\begin{array}{ccc|c} x & y & z & \\ 1 & 1 & 0 & 2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 0 \end{array}$ 25. $\begin{array}{cccc|c} x & y & z & w & \\ 1 & 2 & 0 & 1 & 2 \\ 0 & 1 & 2 & 3 & -3 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{array}$

26. Offer an argument to justify the following claim:

If the j^{th} column of a matrix A consists entirely of zeros, and if the matrix B is equivalent to A, then the j^{th} column of B also consists entirely of zeros.

In the remaining exercises you are to decide whether the given statement is True or False. If True, then you are to present a general argument to establish the validity of the statement in its most general setting. If False, then you are to exhibit a concrete specific example, called a **counterexample**, showing that the given statement does not hold in general. To illustrate:

Prove or Give a Counterexample:

(a) The sum of any two even integers is again an even integer.

(b) Every odd number is prime.

(a) Yes, each time you add two even integers, out pops another even integer, **suggesting** that statement (a) is true. But you certainly can't check to see if (a) holds for all even integers—case by case—as there are infinitely many such cases. A general argument is needed:

> If a and b are even integers, then $a = 2n$ and $b = 2m$ for some integers n and m. We then have: $a + b = 2n + 2m = 2(n + m)$. Since $a + b$ is itself a multiple of 2, it is even.

(b) Surely (b) is false. Why? Because 9 is odd, but 9 is not prime, that's why. To be sure, we could offer a different counterexample, say 15, or 55, but we did have to come up with a **specific concrete counterexample** to shoot down the claim.

PROVE OR GIVE A COUNTEREXAMPLE

27. The system of equations $\left.\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}\right\}$ has a solution for all $a, b, c, d \in \Re$.

28. The system of equations $\left.\begin{aligned} ax + by &= 1 \\ cx + dy &= 1 \end{aligned}\right\}$ always has a solution for all $a, b, c, d \in \Re$.

29. The system of equations $\left.\begin{aligned} ax + by &= 0 \\ cx + dy &= 0 \end{aligned}\right\}$ can never have more than one solution.

30. The systems of equations associated with the two augmented matrices:

$$\left[\begin{array}{cc|c} a & b & 0 \\ c & d & 0 \end{array}\right] \quad \text{and} \quad \left[\begin{array}{cc|c} a' & b' & 0 \\ c' & d' & 0 \end{array}\right]$$

will have the same solution set only if $a = a'$, $b = b'$, $c = c'$, and $d = d'$.

31. If the matrix A has n rows, and if $\text{rref}(A)$ contains less than n leading ones, then the last row of $\text{rref}(A)$ must consist entirely of zeros.

§2. CONSISTENT AND INCONSISTENT SYSTEMS OF EQUATIONS

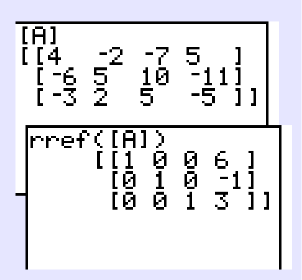
CONSISTENT
INCONSISTENT

A system of equations may have a unique solution, infinitely many solutions, or no solution whatsoever. If it has no solution, then the system is said to be **inconsistent**, otherwise it is said to be **consistent**. As is illustrated in the following examples, the solution set of any system of equations can be spotted from the row-reduced-echelon form of its augmented matrix.

EXAMPLE 1.4 Determine if the following system of equations is consistent. If so, find its solution set.

$$\left.\begin{aligned} 4x-2y-7z &= 5 \\ -6x+5y+10z &= -11 \\ -3x+2y+5z &= -5 \end{aligned}\right\}$$

SOLUTION: Proceeding as in the previous section, we have:



```
[A]
[[4  -2 -7 5  ]
 [-6 5  10 -11]
 [-3 2  5  -5 ]]
rref([A])
     [[1 0 0 6 ]
      [0 1 0 -1]
      [0 0 1 3 ]]
```

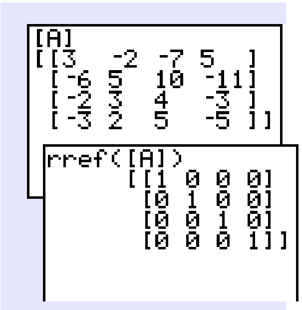
$$\left.\begin{aligned} 4x-2y-7z &= 5 \\ -6x+5y+10z &= -11 \\ -3x+2y+5z &= -5 \end{aligned}\right\} \leftrightarrow \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left[\begin{array}{ccc|c} 4 & -2 & -7 & 5 \\ -6 & 5 & 10 & -11 \\ -3 & 2 & 5 & -5 \end{array}\right] \end{array} \xrightarrow{\text{rref}} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left[\begin{array}{ccc|c} \mathbf{1} & 0 & 0 & 6 \\ 0 & \mathbf{1} & 0 & -1 \\ 0 & 0 & \mathbf{1} & 3 \end{array}\right] \end{array} \leftrightarrow \begin{aligned} x &= 6 \\ y &= -1 \\ z &= 3 \end{aligned}$$

We see that the given system is consistent, and that it has but one solution: $(x=6, y=-1, z=3)$.

EXAMPLE 1.5 Determine if the following system of equations is consistent. If so, find its solution set.

$$\left.\begin{aligned} 3x-2y-7z &= 5 \\ -6x+5y+10z &= -11 \\ -2x+3y+4z &= -3 \\ -3x+2y+5z &= -5 \end{aligned}\right\}$$

SOLUTION: Proceeding as in the previous section, we have:



```
[A]
[[3  -2 -7 5  ]
 [-6 5  10 -11]
 [-2 3  4  -3 ]
 [-3 2  5  -5 ]]
rref([A])
     [[1 0 0 0]
      [0 1 0 0]
      [0 0 1 0]
      [0 0 0 1]]
```

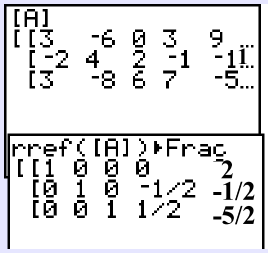
$$\left.\begin{aligned} 3x-2y-7z &= 5 \\ -6x+5y+10z &= -11 \\ -2x+3y+4z &= -3 \\ -3x+2y+5z &= -5 \end{aligned}\right\} \longleftrightarrow \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left[\begin{array}{ccc|c} 3 & -2 & -7 & 5 \\ -6 & 5 & 10 & -11 \\ -2 & 3 & 4 & -3 \\ -3 & 2 & 5 & -5 \end{array}\right] \end{array} \xrightarrow{\text{rref}} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \left[\begin{array}{ccc|c} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{array}\right] \end{array} \to \boxed{\mathbf{0x + 0y + 0z = 1}}$$

Since the equation represented by the last row in the above rref-matrix cannot be satisfied, the given system of equations is inconsistent.

EXAMPLE 1.6 Determine the solution set of the system:

$$\left.\begin{aligned}3x-6y+3w&=9\\-2x+4y+2z-w&=-11\\3x-8y+6z+7w&=-5\end{aligned}\right\}$$

SOLUTION:



```
[A]
[[3   -6 0 3   9 …
 [-2 4   2 -1  -11…
 [3   -8 6 7   -5…
rref([A])►Frac
[[1 0 0 0      2
 [0 1 0 -1/2 -1/2
 [0 0 1 1/2   -5/2
```

$$\left.\begin{aligned}3x-6y+3w&=9\\-2x+4y+2z-w&=-11\\3x-8y+6z+7w&=-5\end{aligned}\right\}\leftrightarrow\begin{array}{c}\begin{matrix}x&y&z&w\end{matrix}\\\left[\begin{array}{cccc|c}3&-6&0&3&9\\-2&4&2&-1&-11\\3&-8&6&7&-5\end{array}\right]\end{array}\xrightarrow{\text{rref}}\begin{array}{c}\begin{matrix}x&y&z&w\end{matrix}\\\left[\begin{array}{cccc|c}\mathbf{1}&0&0&0&2\\0&\mathbf{1}&0&-\frac{1}{2}&-\frac{1}{2}\\0&0&\mathbf{1}&\frac{1}{2}&-\frac{5}{2}\end{array}\right]\end{array}$$

We know that the solution set of the above system of equations coincides with that of the one stemming from the row-reduced-echelon form of its augmented matrix; namely:

$$\left.\begin{aligned}x+0y+0z+0\boldsymbol{w}&=2\\0x+y+0z-\frac{1}{2}\boldsymbol{w}&=-\frac{1}{2}\\0x+0y+z+\frac{1}{2}\boldsymbol{w}&=-\frac{5}{2}\end{aligned}\right\}\text{ or: }\left.\begin{aligned}x&=2\\y&=-\frac{1}{2}+\frac{1}{2}\boldsymbol{w}\\z&=-\frac{5}{2}-\frac{1}{2}\boldsymbol{w}\end{aligned}\right\}$$

Any variable that is not associated with a leading one in the row-reduced echelon form of an augmented matrix is said to be a **free variable**. In the current setting, the variable ***w*** is a free variable (see rref in Figure 1.3).

As you can see, that variable ***w***, which we moved to the right of the equations, can be assigned any value whatsoever, after which the values of x, y, and z (the variables associated with leading-ones in Figure 1.3) are determined. For example, setting $\boldsymbol{w}=\mathbf{0}$ leads to the particular solution:

$$\left(x=2,y=-\frac{1}{2},z=-\frac{5}{2},\boldsymbol{w}=\mathbf{0}\right)$$

We can generate another solution by letting $\boldsymbol{w}=\mathbf{1}$:

$$(x=2,y=0,z=-3,\boldsymbol{w}=\mathbf{1})$$

Indeed, the solutions set of the system of equation is obtained by letting $w=c$, where c can be **any** real number whatsoever:

$$\left\{\left(x=2,y=-\frac{1}{2}+\frac{c}{2},z=-\frac{5}{2}-\frac{c}{2},w=c\right)\middle|\,c\in\Re\right\}$$

Read: ***such that***

We can arrive at a nicer representation of the solution set by replacing each c with $2c$:

$$\left\{\left(x=2,y=-\frac{1}{2}+c,z=-\frac{5}{2}-c,w=2c\right)\middle|\,2c\in\Re\right\}$$

and then observing that as "$2c$ runs through all of the numbers," so does c:

$$\left\{\left(x=2,y=-\frac{1}{2}+c,z=-\frac{5}{2}-c,w=2c\right)\middle|\,c\in\Re\right\}$$

The system of equations of Example 1.4 has a unique solution, that of Example 1.5 has no solutions, and the one in Example 1.6 has infinitely many solutions. These examples cover all of the bases, for if a system of equations has more than one solution then it must have infinitely many solutions (Exercise 22).

CHECK YOUR UNDERSTANDING 1.4

Determine if the system associated with the given row-reduced-echelon augmented matrix is consistent. If it is, find its solution set.

(a) $\left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 5 & 4 \\ 0 & 0 & 0 & 2 \end{array}\right]$ (b) $\left[\begin{array}{ccc|c} 1 & 0 & -2 & 1 \\ 0 & 1 & 5 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$ (c) $\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & 1 & 4 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

Answer: (a) Inconsistent
(b) $\{(1+2r, 4-5r, r) | r \in \Re\}$
(c):
$1-2r-s, r, -2-4s, s | r, s \in R$

Our concern thus far has been with systems of equations with fixed constants on the right side of the equations. We now turn to the question of whether or not a system of equations has a solution for all such constants:

EXAMPLE 1.7 Determine if the following system of equations is consistent for **all** a, b, and c.

$$\left.\begin{aligned} 2x+z &= a \\ 3x+y &= b \\ -x-5y-z &= c \end{aligned}\right\}$$

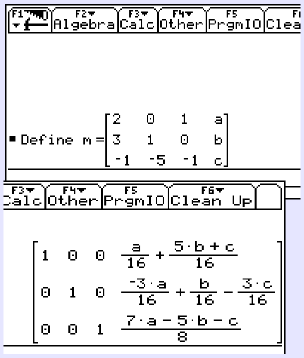
SOLUTION:

$$\left.\begin{aligned} 2x+z &= a \\ 3x+y &= b \\ -x-5y-z &= c \end{aligned}\right\} \longrightarrow \left[\begin{array}{ccc|c} x & y & z & \\ 2 & 0 & 1 & a \\ 3 & 1 & 0 & b \\ -1 & -5 & -1 & c \end{array}\right] \xrightarrow{\frac{1}{2}R_1 \to R_1} \left[\begin{array}{ccc|c} \mathbf{1} & 0 & \frac{1}{2} & \frac{a}{2} \\ 3 & 1 & 0 & b \\ -1 & -5 & -1 & c \end{array}\right]$$

$$\xrightarrow[1R_1+R_3 \to R_3]{-3R_1+R_2 \to R_2} \left[\begin{array}{ccc|c} \mathbf{1} & 0 & \frac{1}{2} & \frac{a}{2} \\ 0 & \mathbf{1} & -\frac{3}{2} & -\frac{3a}{2}+b \\ 0 & -5 & -\frac{1}{2} & \frac{a}{2}+c \end{array}\right] \xrightarrow{5R_2+R_3 \to R_3} \left[\begin{array}{ccc|c} \mathbf{1} & 0 & \frac{1}{2} & \frac{a}{2} \\ 0 & \mathbf{1} & -\frac{3}{2} & -\frac{3a}{2}+b \\ 0 & 0 & -8 & -7a+5b+c \end{array}\right]$$

$$\xrightarrow{-\frac{1}{8}R_3 \to R_3} \left[\begin{array}{ccc|c} \mathbf{1} & 0 & \frac{1}{2} & \frac{a}{2} \\ 0 & \mathbf{1} & -\frac{3}{2} & -\frac{3a}{2}+b \\ 0 & 0 & \mathbf{1} & \frac{7a-5b-c}{8} \end{array}\right] \xrightarrow[\frac{3}{2}R_3+R_2 \to R_2]{-\frac{1}{2}R_3+R_1 \to R_1} \left[\begin{array}{ccc|c} x & y & z & \\ \mathbf{1} & 0 & 0 & \frac{a+5b+c}{16} \\ 0 & \mathbf{1} & 0 & \frac{-3a+b-3c}{16} \\ 0 & 0 & \mathbf{1} & \frac{7a-5b-c}{8} \end{array}\right]$$

Unlike the TI-84+, the TI-89 and above have symbolic capabilities. In particular:



■ Define $m = \begin{bmatrix} 2 & 0 & 1 & a \\ 3 & 1 & 0 & b \\ -1 & -5 & -1 & c \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 & 0 & \frac{a}{16}+\frac{5\cdot b+c}{16} \\ 0 & 1 & 0 & \frac{-3\cdot a}{16}+\frac{b}{16}-\frac{3\cdot c}{16} \\ 0 & 0 & 1 & \frac{7\cdot a-5\cdot b-c}{8} \end{bmatrix}$$

We see that the given system of equations has a solution for **all** a, b, and c; namely: $x = \frac{a+5b+c}{16}$, $y = \frac{-3a+b-3c}{16}$, $z = \frac{7a-5b-c}{8}$

EXAMPLE 1.8 Determine if the following system of equations is consistent for **all possible** values of a, b, c, and d.

$$\left.\begin{aligned} 2x+y+2z+w &= a \\ x-2y+3z+2w &= b \\ 2y+z+4w &= c \\ x-z-4w &= d \end{aligned}\right\}$$

SOLUTION: If you go through the Gauss-Jordon elimination method without making a mistakes you will find that:

$$\begin{tikzcd}
\left.\begin{aligned} 2x+y+2z+w &= a \\ x-2y+3z+2w &= b \\ 0x+2y+z+4w &= c \\ x+0y-z-4w &= d \end{aligned}\right\} \arrow[r, leftrightarrow] & \begin{array}{cccc|c} x & y & z & w & \\ 2 & 1 & 2 & 1 & a \\ 1 & -2 & 3 & 2 & b \\ 0 & 2 & 1 & 4 & c \\ 1 & 0 & -1 & -4 & d \end{array} \arrow[d, "\text{rref}"] \\
\begin{array}{cccc|c} x & y & z & w & \\ \mathbf{1} & 0 & 0 & -2 & \frac{8a-3b-7c}{13} \\ 0 & \mathbf{1} & 0 & 1 & \frac{a-2b+4c}{13} \\ 0 & 0 & \mathbf{1} & 2 & \frac{-2a+4b+5c}{13} \\ 0 & 0 & 0 & 0 & \frac{-10a+7b+12c+13d}{13} \end{array} & \arrow[l]
\end{tikzcd}$$

Figure 1.3

The last row of the above rref-matrix represents the equation:

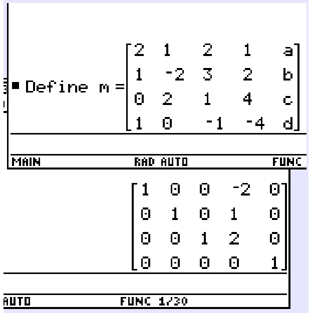
$$0x+0y+0z+0w = \frac{-10a+7b+12c+13d}{13}$$

As such, that matrix tells us that the given system of equation will have a solution if and only if:

$$-10a+7b+12c+13d = 0 \quad (*)$$

In particular, if you choose random numbers for a, b, c, and d, then it is very unlikely that the system will have a solution, for what are the odds that those four numbers will satisfy (*)?

Here, unlike with the smaller system of equations in Example 1.8, the TI-89 (or higher) is of little help:



```
■ Define m = [2  1  2  1  a]
             [1 -2  3  2  b]
             [0  2  1  4  c]
             [1  0 -1 -4  d]
MAIN        RAD AUTO        FUNC

             [1  0  0 -2  0]
             [0  1  0  1  0]
             [0  0  1  2  0]
             [0  0  0  0  1]
AUTO        FUNC 1/30
```

The last row of the above rref matrix tells us that there is no solution to the system, but it "lies," for solutions do exist for certain values of a, b, c, and d [see (*)].

CHECK YOUR UNDERSTANDING 1.5

Determine if the given system of equations has a solution for all a, b, and c. If not, find some specific values of a, b, and c for which a solution does not exist.

$$\text{(a)}\ \left.\begin{aligned} 4x-2y+z &= a \\ -2x+4y+2z &= b \\ 5x-y+4z &= c \end{aligned}\right\} \qquad \text{(b)}\ \left.\begin{aligned} x-4y-4z &= a \\ 2x+8y-12z &= b \\ -x+12y+2z &= c \end{aligned}\right\}$$

Answer: (a) It is consistent for all a, b, and c.
(b) Consistent if and only if
$$c+3a+b = 0$$

COEFFICIENT MATRIX

The **coefficient matrix** of an $m \times n$ system of equation is the $m \times n$ matrix obtained by eliminating the last column of the augmented matrix of the system. For example, referring to system of equations of Examples 1.8, we have:

$$\left.\begin{aligned} 2x + y + 2z + w &= a \\ x - 2y + 3z + 2w &= b \\ 2y + z + 4w &= c \\ x - z - 4w &= d \end{aligned}\right\} \quad \xrightarrow{\text{augmented matrix}} \begin{bmatrix} 2 & 1 & 2 & 1 & a \\ 1 & -2 & 3 & 2 & b \\ 0 & 2 & 1 & 4 & c \\ 1 & 0 & -1 & -4 & d \end{bmatrix} \qquad \xrightarrow{\text{coefficient matrix}} \begin{bmatrix} 2 & 1 & 2 & 1 \\ 1 & -2 & 3 & 2 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & -4 \end{bmatrix}$$

At this point, it behooves us to introduce a bit of notation. To begin with, we will use $\boldsymbol{S}$ to represent a general system of linear equations. We will then let $\text{aug}(\mathbf{S})$ and $\text{coef}(\mathbf{S})$ denote the augmented and coefficient matrices of **S**, respectively. To illustrate:

$$\text{For } \mathbf{S}\text{: } \left.\begin{aligned} 2x + 3y - z &= 1 \\ 3x - y + 2z &= 5 \\ -x + 2y - 3z &= 2 \end{aligned}\right\}, \ \text{aug}(\boldsymbol{S}) = \left[\begin{array}{ccc|c} 2 & 3 & -1 & 1 \\ 3 & -1 & 2 & 5 \\ -1 & 2 & -3 & 2 \end{array}\right], \ \text{and} \ \text{coef}(\boldsymbol{S}) = \begin{bmatrix} 2 & 3 & -1 \\ 3 & -1 & 2 \\ -1 & 2 & -3 \end{bmatrix}$$

The following theorem will enable us to invoke a graphing calculator to resolve the issues of Examples 1.7 and 1.8:

THEOREM 1.2
SPANNING THEOREM

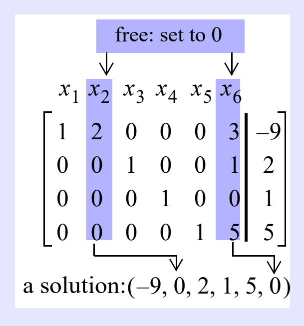
The system of equations:

$$\boldsymbol{S}\text{:} \quad \left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n &= b_m \end{aligned}\right\}$$

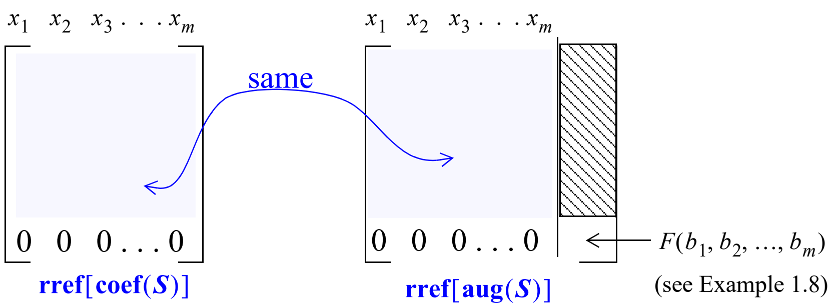
has a solution for all values of $b_1, b_2, \ldots, b_m$ **if and only if** $\text{rref}[\text{coef}(\boldsymbol{S})]$ does not contain a row consisting entirely of zeros.

Let P and Q be two propositions (a proposition is a mathematical statement that is either true or false). To say "P **if and only if** Q," (also written in the form $P \Leftrightarrow Q$) is to say that if P is true then so is Q (also written $P \Rightarrow Q$), and if Q is true then so is P (also written $Q \Rightarrow P$).

PROOF: If $\text{rref}[\text{coef}(\boldsymbol{S})]$ does not contain a row consisting entirely of zeros, then each row of $\text{rref}[\text{coef}(\boldsymbol{S})]$ will have a leading one, as will every row of $\text{rref}[\text{aug}(\boldsymbol{S})]$. For any given values of $b_1, b_2, \ldots, b_m$, a solution for **S** can then be obtained by setting each of the $n - m$ free variables in $\text{rref}[\text{aug}(\boldsymbol{S})]$ to zero, and letting the variable associated with a leading one in the i^{th} row of $\text{rref}[\text{aug}(\boldsymbol{S})]$ equal the last entry in that row (see margin for an illustration).



For the converse, assume that the last row of rref[coef($\boldsymbol{S}$)] consists entirely of zeros. The only difference between rref[coef($\boldsymbol{S}$)] and rref[aug($\boldsymbol{S}$)] is that the latter has an additional column, the last entry of which (as was the case in Figure 1.4) must be a linear expression involving $b_1, b_2, \ldots, b_m$, say $F(b_1, b_2, \ldots, b_m)$:



It follows that for any values of $b_1, b_2, \ldots, b_m$ for which $F(b_1, b_2, \ldots, b_m) \neq 0$, the resulting system of equations will not have a solution, for here is its last equation:

$$0x_1 + 0x_2 + \ldots + 0x_n = F(b_1, b_2, \ldots, b_m)$$

EXAMPLE 1.9 Use the spanning theorem to determine if the given system of equations has a solution for all values of the constants on the right side of the equations.

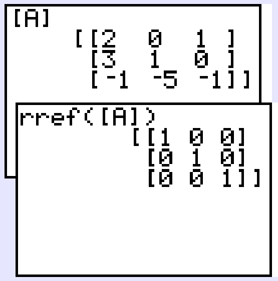
(a) $$\left.\begin{aligned} 2x + z &= a \\ 3x + y &= b \\ -x - 5y - z &= c \end{aligned}\right\}$$

See Example 1.7

(b) $$\left.\begin{aligned} 2x + y + 2z + w &= a \\ x - 2y + 3z + 2w &= b \\ 2y + z + 4w &= c \\ x - z - 4w &= d \end{aligned}\right\}$$

See Example 1.8

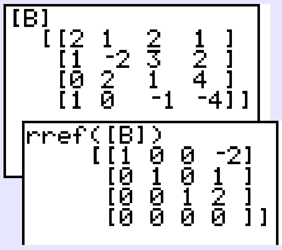
SOLUTION:



```
[A]
      [[2  0  1 ]
       [3  1  0 ]
       [-1 -5 -1]]
rref([A])
             [[1 0 0]
              [0 1 0]
              [0 0 1]]
```

(a) $$\boldsymbol{S}: \left.\begin{aligned} 2x + z &= a \\ 3x + y &= b \\ -x - 5y - z &= c \end{aligned}\right\} \xrightarrow{\text{coef(S)}} \begin{bmatrix} 2 & 0 & 1 \\ 3 & 1 & 0 \\ -1 & -5 & -1 \end{bmatrix} \xrightarrow[\text{(see margin)}]{\text{rref[coef}(\boldsymbol{S})]} \begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & \mathbf{1} & 0 \\ 0 & 0 & \mathbf{1} \end{bmatrix}$$

does not contain a row of zeros:
system has a solution for all values of a, b, and c



```
[B]
    [[2 1  2  1 ]
     [1 -2 3  2 ]
     [0 2  1  4 ]
     [1 0  -1 -4]]
rref([B])
        [[1 0 0 -2]
         [0 1 0 1 ]
         [0 0 1 2 ]
         [0 0 0 0 ]]
```

(b) $$\boldsymbol{S}: \left.\begin{aligned} 2x + y + 2z + w &= a \\ x - 2y + 3z + 2w &= b \\ 2y + z + 4w &= c \\ x - z - 4w &= d \end{aligned}\right\} \xrightarrow{\text{coef(S)}} \begin{bmatrix} 2 & 1 & 2 & 1 \\ 1 & -2 & 3 & 2 \\ 0 & 2 & 1 & 4 \\ 1 & 0 & -1 & -4 \end{bmatrix} \xrightarrow[\text{(see margin)}]{\text{rref[coef}(\boldsymbol{S})]} \begin{bmatrix} \mathbf{1} & 0 & 0 & -2 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & \mathbf{1} & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

contain a row of zeros:
system does not have a solution for all values of a, b, c, and d

Note that while the matrix $\text{rref}[\text{coef}(\boldsymbol{S})]$ in (b) shows that the system S is not consistent for all values of a, b, c, and d, it does not reveal the specific values of a, b, c, and d for which a solution does exist. That information can be derived from the matrix $\text{rref}[\text{aug}(\boldsymbol{S})]$ (see Example 1.8).

CHECK YOUR UNDERSTANDING 1.6

Use the spanning theorem to determine if the given system of equations has a solution for all values of the constants on the right side of the equations.

$$\text{(a)}\ \left.\begin{aligned} 3x + 7y - z &= a \\ 13x - 4y + 2z &= b \\ 2x - 4y + 2z &= c \end{aligned}\right\} \qquad \text{(b)}\ \left.\begin{aligned} x - 3y + w &= a \\ 3x - y + 2z - 3w &= b \\ x + z - 5w &= c \\ 2x - y + z - 2w &= d \end{aligned}\right\}$$

Answer: (a) Yes (b) No

HOMOGENEOUS SYSTEMS OF EQUATIONS

A system with fewer equations than unknowns (“wide”) is said to be **underdetermined.**
A system with more equations than unknowns (“tall”) is said to be **overdetermined.**
A **square** system is a system which contains as many equations as unknowns.

A system of linear equations is said to be **homogeneous** if all of the constants on the right side of the equations are zero:

$$\left.\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & 0 \\ a_{21}x_1 & + & a_{22}x_2 & + & \ldots & + & a_{2n}x_n & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \ldots & + & a_{mn}x_n & = & 0 \end{array}\right\}$$

(A homogeneous system of m equations in n unknowns)

It is easy to see that **every** homogeneous system is consistent, with **trivial solution:** $x_1 = 0, x_2 = 0, \ldots, x_n = 0$. In the event that the homogeneous system is “wide”, then it has more than one solution:

THEOREM 1.3
FUNDAMENTAL THEOREM OF HOMOGENEOUS SYSTEMS OF EQUATIONS

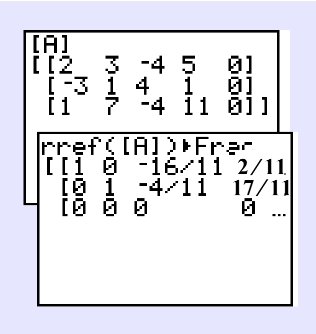
Any homogeneous system $\boldsymbol{S}$ of m linear equations in n unknowns with $n > m$ has nontrivial solutions.

PROOF: Having more columns than rows, $\text{rref}[\text{aug}(\boldsymbol{S})]$ must have free variables, and therefore the system has infinitely many solutions.

EXAMPLE 1.10 Determine the solution set of:

$$\left.\begin{aligned} 2x + 3y - 4z + 5w &= 0 \\ -3x + y + 4z + w &= 0 \\ x + 7y - 4z + 11w &= 0 \end{aligned}\right\}$$

SOLUTION: Theorem 1.3 tells us that the system has nontrivial solutions. Let's find them:



```
[A]
[[2  3  -4 5  0]
 [-3 1  4  1  0]
 [1  7  -4 11 0]]
rref([A])▸Frac
[[1 0 -16/11 2/11
 [0 1 -4/11  17/11
 [0 0 0         0 …
```

$$S:\ \left.\begin{aligned}2x+3y-4z+5w&=0\\-3x+y+4z+w&=0\\x+7y-4z+11w&=0\end{aligned}\right\}\xrightarrow{\text{aug(S)}}\begin{array}{cccc|c}x&y&z&w&\\ \hline 2&3&-4&5&0\\-3&1&4&1&0\\1&7&-4&11&0\end{array}$$

$$\xrightarrow{\text{rref}}\begin{array}{cccc|c}x&y&\overset{\text{free variables}}{z}&w&\\ \hline \mathbf{1}&0&-\frac{16}{11}&\frac{2}{11}&0\\0&\mathbf{1}&-\frac{4}{11}&\frac{17}{11}&0\\0&0&0&0&0\end{array}\longrightarrow\begin{aligned}x&=\frac{16}{11}z-\frac{2}{11}w\\y&=\frac{4}{11}z-\frac{17}{11}w\end{aligned}$$

Figure 1.4

Assigning arbitrary values to the two free variables z and w we arrive at the solution set of the system:

$$\left\{\left(\overbrace{\frac{16}{11}a-\frac{2}{11}b}^{x},\ \overbrace{\frac{4}{11}a-\frac{17}{11}b}^{y},\ \overset{z}{a},\ \overset{w}{b}\right)\middle|\,a,b\in R\right\}=\{(16a-2b,4a-17b,11a,11b)|a,b\in R\}$$

If $\boldsymbol{S}$ is a homogeneous system of equations, then the last column of $\text{rref}[\text{aug}(\boldsymbol{S})]$ will always consist entirely of zeros (Exercise 20). Consequently, when solving a homogeneous system of equations, one might as well start with $\text{coef}(\boldsymbol{S})$ rather than with $\text{aug}(\boldsymbol{S})$ (one less column to carry along in the rref-process, that's all). In particular, the solution set of the homogeneous system of the last example can easily be read from $\text{rref}[\text{coef}(\boldsymbol{S})]$ (just mentally add a **column of zeros** to the right of the matrix):

$$S:\ \left.\begin{aligned}2x+3y-4z+5w&=0\\-3x+y+4z+w&=0\\x+7y-4z+11w&=0\end{aligned}\right\}\xrightarrow{\text{coef(S)}}\begin{array}{cccc}x&y&z&w\\ \hline 2&3&-4&5\\-3&1&4&1\\1&7&-4&11\end{array}\xrightarrow{\text{rref}}\begin{array}{cccc|c}x&y&z&w&\\ \hline \mathbf{1}&0&-\frac{16}{11}&\frac{2}{11}&\mathbf{0}\\0&\mathbf{1}&-\frac{4}{11}&\frac{17}{11}&\mathbf{0}\\0&0&0&0&\mathbf{0}\end{array}$$

CHECK YOUR UNDERSTANDING 1.7

Determine the solution set of:

$$\left.\begin{aligned}2x+3y+4z+5w&=0\\3x+y+4z+w&=0\\x+7y+4z+11w&=0\end{aligned}\right\}$$

Answer:
$\{(4r,-2r,-3r,2r)|r\in R\}$

While underdetermined ("wide") homogeneous systems of equations are guaranteed to always have non-trivial solutions, this is not the case with overdetermined ("tall") systems of equations [see Exercises 27-28], or with square systems of equations [see Exercises 29-30].

We end this section with a rather obvious result, but one that will play an important role in future developments; so much so, that we label it accordingly:

THEOREM 1.4

LINEAR INDEPENDENCE THEOREM

A homogeneous system $\boldsymbol{S}$ of m linear equations in n unknowns has only the trivial solution if and only if $\text{rref}[\text{coef}(\boldsymbol{S})]$ has n leading ones.

PROOF: Since there are n unknowns, to say that $\text{rref}[\text{coef}(\boldsymbol{S})]$ has n leading ones, is to say that it has no free variables.

Exercises

Exercises 1-6. Determine if the system $\boldsymbol{S}$ with given $\operatorname{rref}[\operatorname{aug}(\boldsymbol{S})]$ is consistent. If it is, find its solution set.

1. $\left[\begin{array}{cc|c} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right]$

2. $\left[\begin{array}{cccc|c} 0 & 1 & 0 & 1 & 3 \\ 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$

3. $\left[\begin{array}{ccc|c} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 1 & -2 \end{array}\right]$

4. $\left[\begin{array}{ccc|c} 1 & 0 & 3 & 5 \\ 0 & 1 & -4 & 2 \\ 0 & 0 & 0 & 1 \end{array}\right]$

5. $\left[\begin{array}{cccccc|c} 1 & 0 & 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 1 & 2 & 0 & 1 \end{array}\right]$

6. $\left[\begin{array}{ccccccc|c} 1 & 0 & 0 & 0 & 0 & 5 & -2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 & -1 & 3 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$

Exercises 7-12. Determine if the system of equations is consistent. If it is, find its solution set.

7. $\left.\begin{aligned} 2x+3y+z &= 4 \\ x+y+2z &= 5 \end{aligned}\right\}$

8. $\left.\begin{aligned} x-4y-4z &= 1 \\ 2x+y-2z &= 8 \end{aligned}\right\}$

9. $\left.\begin{aligned} 4x-2y+z &= 4 \\ -2x+4y+2z &= 10 \\ 5x-y+4z &= 2 \end{aligned}\right\}$

10. $\left.\begin{aligned} x-4y-4z &= 1 \\ 2x+8y-12z &= 8 \\ -x+12y+2z &= 1 \end{aligned}\right\}$

11. $\left.\begin{aligned} 2x+3y-z+2w &= 4 \\ x-y+2z-w &= 3 \\ 2y+z-2w &= 1 \\ 6x-3y+6z &= 15 \end{aligned}\right\}$

12. $\left.\begin{aligned} -x+w+y+z &= 3 \\ 6x+4z-2y+3w &= 4 \\ 5y-3x-6w-z &= -1 \\ -20x-7w-10z+8y &= -18 \end{aligned}\right\}$

Exercises 13-14. Does the system of equations have a solution for all a, and b? If not, find some specific values of a and b for which a solution does not exist, and some specific values of a and b, not both zero, for which a solution does exist.

13. $\left.\begin{aligned} 2x+3y+z &= a \\ x+y+2z &= b \end{aligned}\right\}$

14. $\left.\begin{aligned} x-4y-4z &= a \\ 2x+y-2z &= b \end{aligned}\right\}$

Exercises 15-16. Does the system have a solution for all a, b, and c? If not, find some specific values of a, b, and c for which a solution does not exist, and some specific values of a, b, and c, not all zero, for which a solution does exist.

15. $\left.\begin{aligned} x-4y-4z &= a \\ 2x+8y-12z &= b \\ -x+12y+2z &= c \end{aligned}\right\}$

16. $\left.\begin{aligned} 4x-2y+z &= a \\ -2x+4y+2z &= b \\ 5x-y+4z &= c \end{aligned}\right\}$

Exercises 17-19. Use the Spanning Theorem to determine if the system of equations has a solution for all values of a, b, c, and d.

17. $$\left.\begin{aligned} 2x-y &= a \\ z-3w &= b \\ 2x+2z &= c \\ y+2z &= d \end{aligned}\right\}$$

18. $$\left.\begin{aligned} 4x-2y+z &= a \\ -2x+4y+2z &= b \\ 5x-y+4z &= c \\ 2x+y+z &= d \end{aligned}\right\}$$

19. $$\left.\begin{aligned} 4x-2y+z &= a \\ -2x+4y+2z &= b \end{aligned}\right\}$$

20. Let $\boldsymbol{S}$ is a homogeneous system of equations. Prove that the last column of $\text{rref}[\text{aug}(\boldsymbol{S})]$ contains only zeros.

21. Prove that if $m > n$, then the system of equations:

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\quad \vdots \quad &\quad\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n &= b_m \end{aligned}\right\}$$

cannot have a solution for all values of $b_1, b_2, \ldots, b_m$.

22. Show that if $x = x_0, y = y_0$ and $x = x_1, y = y_1$ are solutions of the system: $\left.\begin{aligned} ax+by &= c \\ dx+ey &= f \end{aligned}\right\}$,

then, $x = x_0 + k(x_1 - x_0), y = y_0 + k(y_1 - y_0)$ is also a solution for any given $k \in \Re$.

Suggestion: Substitute the above expressions for x and y into the given system.

Exercises 23-26. Determine the solution set of the given underdetermined ("wide") homogeneous system of equations.

23. $$\left.\begin{aligned} 2x+3y-z &= 0 \\ 4x+6y+2z &= 0 \end{aligned}\right\}$$

24. $$\left.\begin{aligned} 2x+3y-z &= 0 \\ 4x+3y-2z &= 0 \end{aligned}\right\}$$

25. $$\left.\begin{aligned} 2x+3y-z+4w &= 0 \\ -3x-5y+2z-3w &= 0 \\ -x-3y+2z+7w &= 0 \end{aligned}\right\}$$

26. $$\left.\begin{aligned} 2x+3y-2z+4w &= 0 \\ -3x-5y+2z-3w &= 0 \\ -x-3y+2z+7w &= 0 \end{aligned}\right\}$$

Exercises 27-28. Determine if the given overdetermined ("tall") homogeneous system of equations has a unique solution.

27. $$\left.\begin{aligned} 2x+3y-4z &= 0 \\ 3x+2y+z &= 0 \\ x+4y-9z &= 0 \\ -4x-y-6z &= 0 \end{aligned}\right\}$$

28. $$\left.\begin{aligned} 2x+3y+4z+6w &= 0 \\ x+3y+5z+2w &= 0 \\ 2x+y+6z+7w &= 0 \\ 5x+3y+2z+w &= 0 \\ 2x+4y+6z+2w &= 0 \\ 3x+y+4z+w &= 0 \end{aligned}\right\}$$

Exercises 29-30. Determine if the given square homogeneous system of equations has a unique solution.

29. $\left.\begin{aligned} 5x+3y-4z+5w &= 0\\ -x-y+2z-9w &= 0\\ 3x-3y+2z-w &= 0\\ 11x+y-2z-9w &= 0 \end{aligned}\right\}$

30. $\left.\begin{aligned} 2x+5y+z+4w &= 0\\ -3x-2y+4z+6w &= 0\\ 4x+y-2z+6w &= 0\\ 9x-3y-2z+0w &= 0 \end{aligned}\right\}$

Exercises 31-33. For what values of a will the given homogeneous system of equations have a unique solution?

31. $\left.\begin{aligned} x+ay &= 0\\ ax+y &= 0 \end{aligned}\right\}$

32. $\left.\begin{aligned} x+y+z &= 0\\ x+2y-z &= 0\\ -x+y+az &= 0 \end{aligned}\right\}$

33. $\left.\begin{aligned} x+y+z &= 0\\ x+ay-z &= 0\\ -x+y+az &= 0 \end{aligned}\right\}$

Exercises 34-36. For what values of a and b will the given homogeneous system of equations have a unique solution?

34. $\left.\begin{aligned} x+ay &= 0\\ 2x+by &= 0 \end{aligned}\right\}$

35. $\left.\begin{aligned} x+ay &= 0\\ bx+y &= 0 \end{aligned}\right\}$

36. $\left.\begin{aligned} x+y+z &= 0\\ x+ay-z &= 0\\ -x+y+bz &= 0 \end{aligned}\right\}$

37. For what values of a, b, c, and d will the homogeneous system of equations $\left.\begin{aligned} ax+by &= 0\\ cx+dy &= 0 \end{aligned}\right\}$ have a unique solution:

38. Show that if (x_0, y_0) is a solution of a given two by two homogeneous system of equations, then (kx_0, ky_0) is also a solution for any $k \in \Re$.

39. Show that if (x_0, y_0) and (x_1, y_1) are solutions of a given two by two homogeneous system of equations, then (x_0+x_1, y_0+y_1) is also a solution.

40. Let M be the solution set of **S:** $\left.\begin{aligned} a_{11}x+a_{12}y &= b_1\\ a_{21}x+a_{22}y &= b_2 \end{aligned}\right\}$ and let T be the solution set of the corresponding homogeneous system **H:** $\left.\begin{aligned} a_{11}x+a_{12}y &= 0\\ a_{21}x+a_{22}y &= 0 \end{aligned}\right\}$. Show that:

(a) If $(x_0, y_0) \in M$ and $(x_1, y_1) \in M$, then $(x_0-x_1, y_0-y_1) \in T$.

(b) If $(x_0, y_0) \in M$ and $(x_1, y_1) \in T$, then $(x_0+x_1, y_0+y_1) \in M$

41. Let M be the solution set of **S:** $\left.\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}\right\}$ and let T be the solution set of the corresponding homogeneous system **H:** $\left.\begin{aligned} a_{11}x + a_{12}y &= 0 \\ a_{21}x + a_{22}y &= 0 \end{aligned}\right\}$. Show that for any $(x_0, y_0) \in M$, $M = \{(x_0 + x_1, y_0 + y_1) | (x_1, y_1) \in T\}$.

PROVE OR GIVE A COUNTEREXAMPLE

42. The system of equations associated with the augmented matrix $\left[\begin{array}{ccc|c} a & b & c & 0 \\ d & e & f & a \end{array}\right]$ is consistent, independent of the values of the entries a through f.

43. The system of equations associated with the augmented matrix $\left[\begin{array}{ccc|c} a & b & c & 1 \\ d & e & f & 0 \end{array}\right]$ is consistent, independent of the values of the entries a through f.

44. The system of equations associated with the augmented matrix $\left[\begin{array}{cc|c} 1 & 6 & 3 \\ 0 & d & 0 \end{array}\right]$ is consistent if and only if $d = 0$.

45. If a homogeneous system of equations has a nontrivial solution, then it has infinitely many solutions.

46. If the homogeneous system $\left.\begin{aligned} a_{11}x + a_{12}y &= 0 \\ a_{21}x + a_{22}y &= 0 \end{aligned}\right\}$ has only the trivial solution, then the system $\left.\begin{aligned} a_{11}x + a_{12}y &= b_1 \\ a_{21}x + a_{22}y &= b_2 \end{aligned}\right\}$ has a unique solution for all b_1, b_2.

47. Any system $\boldsymbol{S}$ of m linear equations in n unknowns with $n > m$ has nontrivial solutions.

48. A system of n linear equations in m unknowns $\boldsymbol{S}$ is consistent if and only if $\text{rref}[\text{coef}(\boldsymbol{S})]$ has m leading ones.

CHAPTER SUMMARY

***N*-TUPLE**	An (ordered) ***n*-tuple** is an expression of the form $(c_1, c_2, \ldots, c_n)$, where each c_i is a real number, for $1 \le i \le n$.
SOLUTION SET OF A SYSTEM OF EQUATIONS	An n-tuple $(c_1, c_2, \ldots, c_n)$ is a **solution** of the system, S, of m equations in n unknowns $\mathbf{S}: \left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \ldots a_{1n}x_n &=& b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots a_{2n}x_n &=& b_2 \\ \vdots \quad\;\; \vdots \quad\;\; \vdots && \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots a_{mn}x_n &=& b_m \end{array}\right\}$ if each of the m equations is satisfied when c_i is substituted for x_i, for $1 \le i \le n$. The **solution set** of $\mathbf{S}$ is the set of all solutions of $\mathbf{S}$.
CONSISTENT AND INCONSISTENT SYSTEMS OF EQUATIONS	A system of equations is said to be **consistent** if it has non-empty solution set. A system of equations that has no solution is said to be **inconsistent**.
EQUIVALENT SYSTEMS OF EQUATIONS	Two systems of equations are said to be **equivalent** if they have equal solution sets.
OVERDETERMINED, UNDERDETERMINED, AND SQUARE SYSTEMS OF EQUATIONS	A system of m equations in n unknowns is said to be: **Overdetermined** if $n < m$ (more equations than unknowns). **Underdetermined** if $n > m$ (fewer equations than unknowns). **Square** if $n = m$.
ELEMENTARY EQUATION OPERATIONS	The following three operations on a system of linear equations are said to be **elementary equation operations**: Interchange the order of any two equations in the system. Multiply both sides of an equation by a nonzero number. Add a multiple of one equation to another equation.
Elementary row operations do not alter the solution sets of systems of equations.	Performing any number of elementary equation operations on a system of linear equations will result in an equivalent system of equations (same solution set).

<table>
<tr><td>MATRICES</td><td>Matrices are arrays of numbers arranged in rows and columns, such as the matrix A below:
$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix}$ also: $A_{3\times 4} = [a_{ij}]$, or $A = [a_{ij}]$
Since A has 3 rows and 4 columns, it is said to be a three-by-four matrix. When the number of rows of a matrix equals the number of columns, the matrix is said to be a square matrix.</td></tr>
<tr><td>ELEMENTARY ROW OPERATIONS</td><td>The following three operations on any given matrix are said to be elementary row operations:
Interchange the order of any two rows in the matrix.
Multiply each element in a row of the matrix by a nonzero number.
Add a multiple of one row of the matrix to another row of the matrix.</td></tr>
<tr><td>EQUIVALENT MATRICES</td><td>Two matrices are equivalent if one can be derived from the other by means of a sequence of elementary row operations.</td></tr>
<tr><td>AUGMENTED MATRIX</td><td>The augmented matrix of a system of equations S is that matrix aug(S) composed of the coefficients of the equations in, along with the constants to the right of the equations in S. For example:
$\mathbf{S}: \left.\begin{aligned} 2x+y-z &= -2 \\ x+3y+2z &= 9 \\ -x-y+2z &= 1 \end{aligned}\right\} \xrightarrow{\mathbf{aug(S)}} \left[\begin{array}{ccc|c} 2 & 1 & -1 & -2 \\ 1 & 3 & 2 & 9 \\ -1 & -1 & 2 & 1 \end{array}\right]$</td></tr>
<tr><td><i>Equivalent systems of equations corresponding to equivalent augmented matrices.</i></td><td>Two systems of equations, $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$, are equivalent if and only if their corresponding augmented matrices, aug $(\boldsymbol{S_1})$ and aug $(\boldsymbol{S_2})$, are equivalent.</td></tr>
<tr><td>ROW-REDUCED-ECHELON FORM OF A MATRIX</td><td>All rows consisting entirely of zeros are at the bottom of the matrix. All of the other rows contain a leading-1 (with zeros in all entries above or below it), and those leading-ones “move” to the right, as you “move” down the matrix.</td></tr>
<tr><td><i>Gauss-Jordan Elimination Method.</i></td><td>The Gauss-Jordan Elimination Method of page 10 can be used to obtain the row-reduced-echelon form rref(A) of a given matrix A.</td></tr>
</table>

COEFFICIENT MATRIX	The **coefficient matrix** of a system of equations **S** is that matrix coef (**S**) composed of the coefficients of **S**. For example: $\mathbf{S}: \left.\begin{array}{rcl} 2x+y-z &=& -2 \\ x+3y+2z &=& 9 \\ -x-y+2z &=& 1 \end{array}\right\} \xrightarrow{\textbf{coef (S)}} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 3 & 2 \\ -1 & -1 & 2 \end{bmatrix}$ forget about the constants on the right of the equal signs
Spanning Theorem	The system of equations: $\mathbf{S}: \left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \dots a_{1n}x_n &=& b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots a_{2n}x_n &=& b_2 \\ \vdots \quad \vdots \quad \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots a_{mn}x_n &=& b_m \end{array}\right\}$ has a solution for all values of $b_1, b_2, \dots, b_m$ **if and only if** $\text{rref}[\text{coef}(\boldsymbol{S})]$ does not contain a row consisting entirely of zeros.
HOMOGENEOUS SYSTEM OF EQUATIONS	A system of equations of the form: $\mathbf{S}: \left.\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &=& 0 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &=& 0 \\ \vdots \quad \vdots \quad \vdots & & \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &=& 0 \end{array}\right\}$ with zeros to the right of the equal sign, is said to be **homogeneous**.
TRIVIAL SOLUTION	$x_1 = 0, x_2 = 0, \dots, x_n = 0$ is a solution of the above homogeneous system of equation. It is said to be the **trivial solution** of the system.
Fundamental Theorem of Homogeneous Systems	Any homogeneous system **S** of m linear equations in n unknowns with $n > m$ has nontrivial solutions.
You can use $\text{rref}[\text{coef}(\boldsymbol{S})]$ *to solve a homogeneous system of equations* **S**	Let $\boldsymbol{S}$ be a homogeneous system of equations. The only difference between $\text{rref}[\text{aug}(\boldsymbol{S})]$ and $\text{rref}[\text{coef}(\boldsymbol{S})]$ is that the former contains an additional column of zeros. Being aware of this, you might as well focus on $\text{rref}[\text{coef}(\boldsymbol{S})]$ to derive the solution set of $\boldsymbol{S}$.
Linear Independence Theorem	A homogeneous system $\boldsymbol{S}$ of m linear equations in n unknowns has only the trivial solution if and only if $\text{rref}[\text{coef}(\boldsymbol{S})]$ has n leading ones.

CHAPTER 2
VECTOR SPACES

We begin this chapter with a geometrical consideration of vectors as directed line segments in the plane and in three dimensional space, and then extend the vector concept to higher dimensional Euclidean spaces.

Abstraction is the nature of mathematics, and we let the "essence" of Euclidean vector spaces guide us, in Section 2, to the definition of an abstract vector space. In Section 3 we begin to uncover some of the beautiful (and very usefull) theory of abstract vector spaces, an excavation that will keep us well-occupied for the remainder of the text. Subsets of vector spaces which are themselves vector spaces are considered in Section 4. In Section 5, we return to the two and three dimensional Euclidean spaces of the first section and derive a vector representation for lines and planes in those spaces.

§1. VECTORS IN THE PLANE AND BEYOND

We begin by considering vectors in the plane, such as those in Figure 2.1, which are depicted as directed line segments ("arrows"). In that geometrical setting, the arrow is pointing in the direction of the vector, with the length of the arrow representing its magnitude.

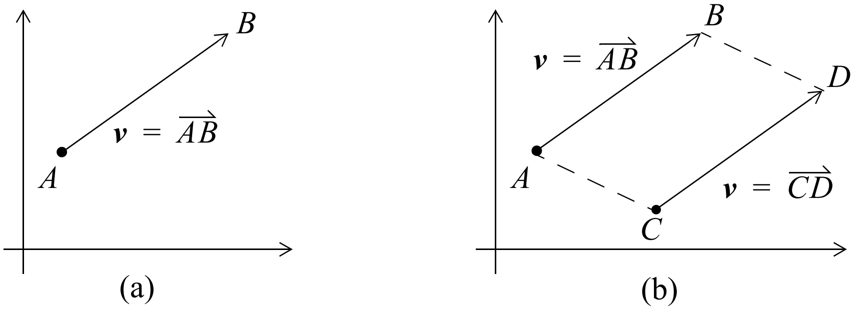
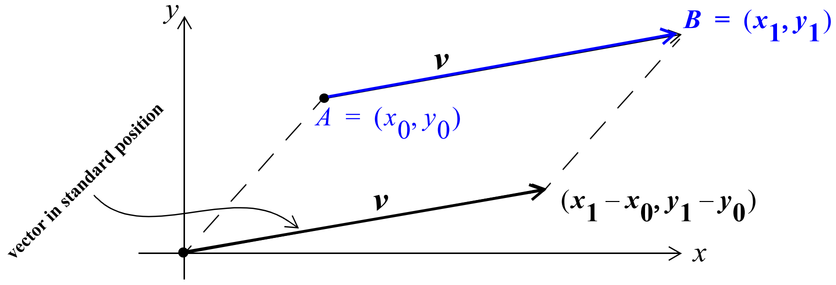


Figure 2.1

Vectors will be denoted by boldface lowercase letters. The vector $\boldsymbol{v} = \overrightarrow{AB}$ in Figure 2.1 is said to have **initial point** A, and **terminal point** B. One defines two vectors to be equal if they have the same magnitude and direction. If you pick up the vector $\boldsymbol{v}$ in Figure 2.1(a) and move it in a parallel fashion as we did in Figure 2.1(b) to the vector with initial point C and terminal point D, then you will still have the same vector:

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

In particular, the vector $\boldsymbol{v}$ in Figure 2.2 with initial point $A = (x_0, y_0)$ and terminal point $B = (x_1, y_1)$ can be moved in a parallel fashion so that its initial point coincides with the origin. When so placed, the vector is said to be in **standard position**.



6

Figure 2.2

3

EXAMPLE 2.1 Sketch the vector with initial point $(-2, 3)$ and terminal point $(4, -1)$. Position that vector in standard position, and identify its terminal point.

SOLUTION: The figure below tells the whole story:

Pick up the top vector and move it 2 units down and 3 units to the right to the right so that its initial point $(-2, 3)$. In the process, the original terminal point $(4, -1)$ is also moved 2 units to the right at 3 units down, coming to rest at $(6, -4)$.

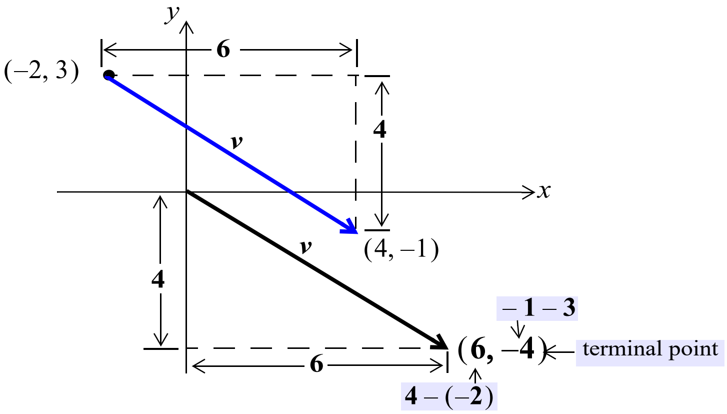


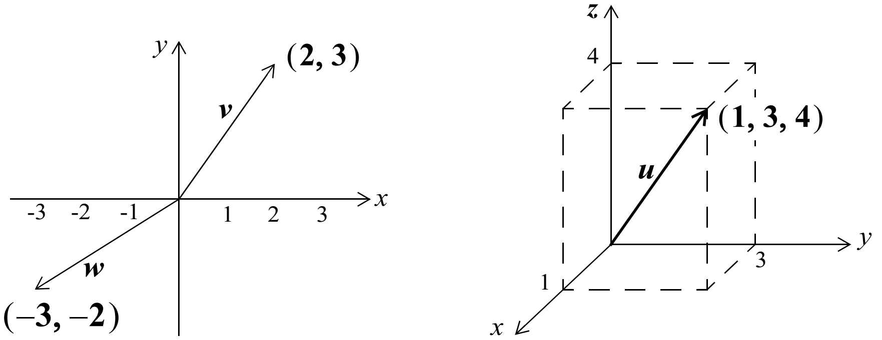
Figure 2.3

A standard position vector in the plane is completely determined by the coordinates of its terminal point. This observation enables us to identify the vector in Figure 2.3 as an **ordered pair** of numbers or **2-tuple**; namely:

$$\boldsymbol{v} = (\mathbf{6}, \mathbf{-4})$$

Note that the two-tuple in the expression $\boldsymbol{v} = (\mathbf{6}, \mathbf{-4})$ appears in bold-face, so as to distinguish it from the form $(6, -4)$ which represents a point in the plane.

In a similar fashion we may refer to the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ in Figure 2.4(a) as $\boldsymbol{v} = (\mathbf{2}, \mathbf{3})$ and $\boldsymbol{w} = (\mathbf{-3}, \mathbf{-2})$.



(a) (b)

Figure 2.4

Likewise, the vector $\boldsymbol{u}$ in the 3 dimensional space of Figure 2.4(b) can be described by the bold-faced 3-tuple $\boldsymbol{u} = (\mathbf{1, 3, 4})$

The beauty of all of this is that while we cannot geometrically represent a vector in 4 dimensional space, we can certainly consider 4-tuples, and beyond. With this in mind, let us agree to denote the set of all (ordered) n-tuples $(a_1, a_2, \ldots, a_n)$ by the symbol $\Re^n$:

$$\Re^n = \{(a_1, a_2, \ldots, a_n) \mid a_i \in \Re\}$$

The real numbers a_i in the n-tuple $(a_1, a_2, \ldots, a_n)$ are said to be the **components** of the n-tuple, and we define two n-tuples to be **equal** if their corresponding components are identical:

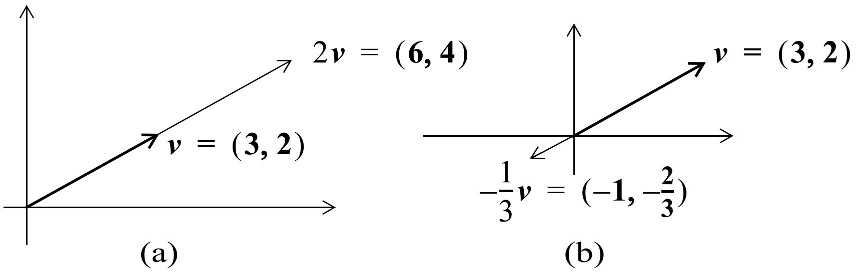
DEFINITION 2.1

***n*-TUPLE EQUALITY**

$(a_1, a_2, \ldots, a_n) = (b_1, b_2, \ldots, b_n)$ if $a_i = b_i$, for $1 \le i \le n$.

SCALAR PRODUCT AND SUMS OF VECTORS

Vectors evolved from the need to adequately represent quantities which are characterized by both magnitude and direction. In a way, these quantities themselves tell us how we should go about defining certain algebraic operations on vectors. Suppose, for example, that the vector $\boldsymbol{v} = (\mathbf{3, 2})$ of Figure 2.5(a) represents a force. Doubling the magnitude of that force without changing its direction would result in the vector force labeled $2\boldsymbol{v}$ in that figure, as that vector is in the same direction as $\boldsymbol{v} = (\mathbf{3, 2})$, with length twice that of $\boldsymbol{v}$.



(a) (b)

Figure 2.5

Similarly, if a force that is one-third that of $\boldsymbol{v} = (\mathbf{3, 2})$ is applied in the opposite direction of $\boldsymbol{v}$, then the vector representing that new force is the vector $-\frac{1}{3}\boldsymbol{v}$ in Figure 2.5(b); for that vector is in the opposite direction of $\boldsymbol{v} = (\mathbf{3, 2})$, with length one-third that of $\boldsymbol{v}$.

Length of $\boldsymbol{v} = (\mathbf{3, 2})$:

$$\sqrt{3^2 + 2^2} = \sqrt{13}$$

Length of $\left(\mathbf{-1, -\frac{2}{3}}\right)$:

$$\sqrt{(-1)^2 + \left(-\frac{2}{3}\right)^2} = \sqrt{1 + \frac{4}{9}} = \sqrt{\frac{13}{9}} = \frac{\sqrt{13}}{3}$$

This stretching or shrinking of a vector, in one direction or the other, is an important operation which we now formalize and extend to the set of n-tuple-vectors for any positive integer n:

DEFINITION 2.2
SCALAR PRODUCT

To any vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ in $\Re^n$, and any $r \in \Re$, we let:

$$r\boldsymbol{v} = (rv_1, rv_2, \ldots, rv_n)$$

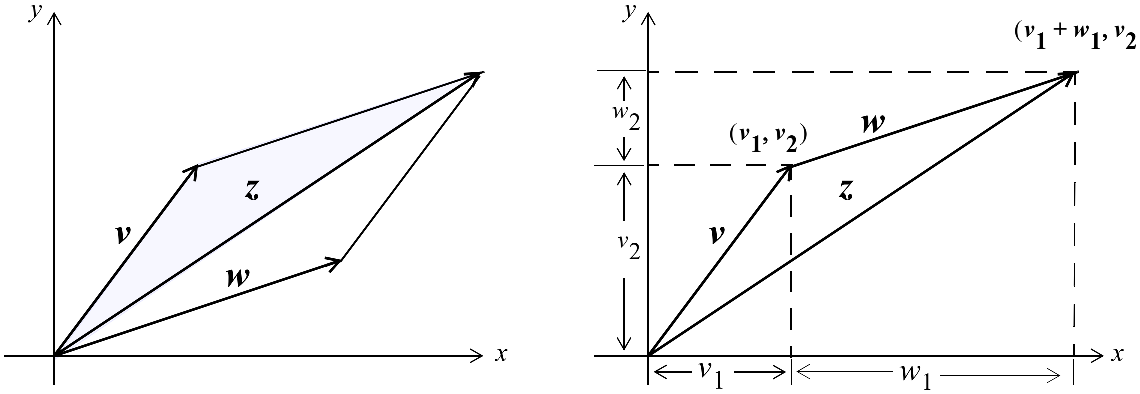
The vector $r\boldsymbol{v}$ is said to be a **scalar multiple of $\boldsymbol{v}$**.

For example:

$$3(\mathbf{1, 5}) = (\mathbf{3, 15})$$
$$-5(\mathbf{1, 0, -4}) = (\mathbf{-5, 0, 20})$$
$$\sqrt{2}(\mathbf{1, 3, -4, 5}) = (\boldsymbol{\sqrt{2}, 3\sqrt{2}, -4\sqrt{2}, 5\sqrt{2}})$$

VECTOR ADDITION

If two people pull on an object positioned at the origin with forces $\boldsymbol{v}$ and $\boldsymbol{w}$, then the observed combined effect is the same as that of one individual pulling with force $\boldsymbol{z}$, where $\boldsymbol{z}$ is the vector coinciding with the diagonal in the parallelogram formed by the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ [Figure 2.6(a)].



(a) (b)

Figure 2.6

The above vector $\boldsymbol{z}$ is said to be the sum of the vectors $\boldsymbol{v}$ and $\boldsymbol{w}$, and is denoted by $\boldsymbol{v} + \boldsymbol{w}$. Figure 2.6(b) reveals that:

$$\boldsymbol{z} = \boldsymbol{v} + \boldsymbol{w} = (v_1, v_2) + (w_1, w_2) = (v_1 + w_1, v_2 + w_2)$$

Generalizing, we have:

DEFINITION 2.3
VECTOR SUM

The **sum** of the vectors $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$ and $\boldsymbol{w} = (w_1, w_2, \ldots, w_n)$ in $\Re^n$, is denoted by $\boldsymbol{v} + \boldsymbol{w}$ and is given by:

$$\boldsymbol{v} + \boldsymbol{w} = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$$

While identical in shape, the "+" in $\boldsymbol{v} + \boldsymbol{w}$ differs in spirit from that in $v_i + w_i$: the latter represents the familiar sum of two numbers, as in $3 + 7$, while the former represents the newly defined sum of two n-tuples, as in:

$$(\mathbf{3, -2}) + (\mathbf{7, 11})$$

EXAMPLE 2.2 For $\boldsymbol{v} = (\mathbf{-2, 3, 1, 5})$, $\boldsymbol{w} = (\mathbf{1, 5, 0, -2})$, and $r = 2$, determine the vector $r\boldsymbol{v} + \boldsymbol{w}$.

SOLUTION:

$$r\boldsymbol{v} + \boldsymbol{w} = 2(\mathbf{-2, 3, 1, 5}) + (\mathbf{1, 5, 0, -2})$$

Definition 2.2:
$$= (\mathbf{-4, 6, 2, 10}) + (\mathbf{1, 5, 0, -2})$$

Definition 2.3:
$$= (\mathbf{-4 + 1, 6 + 5, 2 + 0, 10 - 2})$$
$$= (\mathbf{-3, 11, 2, 8})$$

CHECK YOUR UNDERSTANDING 2.1

For $\boldsymbol{v} = (\mathbf{3, 2, -2})$, $\boldsymbol{w} = (\mathbf{-3, 1, 0})$, $r = 2$, $s = -3$, determine the vector $r\boldsymbol{v} + s\boldsymbol{w}$

Answer: $(\mathbf{15, 1, -4})$

EUCLIDEAN VECTOR SPACES

The set of n-tuples $\Re^n$, together with the above defined operations of vector addition:

$$(\boldsymbol{v_1, v_2, \ldots, v_n}) + (\boldsymbol{w_1, w_2, \ldots, w_n}) = (\boldsymbol{v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n})$$

and scalar multiplication:

$$r(\boldsymbol{v_1, v_2, \ldots, v_n}) = (\boldsymbol{rv_1, rv_2, \ldots, rv_n})$$

is called the **Euclidean *n*-space**.

Every Euclidean space contains a most distinguished vector:

DEFINITION 2.4
ZERO VECTOR

The **zero vector** in $\Re^n$, denoted by **0**, is that vector with each component the number 0:

$$\mathbf{0} = (\mathbf{0, 0, \ldots, 0})$$

For example $\mathbf{0} = (\mathbf{0, 0, 0})$ is the zero vector in $\Re^3$, and $\mathbf{0} = (\mathbf{0, 0, 0, 0})$ is the zero vector in $\Re^4$.

No direction is associated with the zero vector. A zero force, for example, is no force at all, and its "direction" would be a moot point.

Every real number r has an additive inverse $-r$, namely that number which when added to r yields 0. Similarly:

DEFINITION 2.5
ADDITIVE INVERSE OF A VECTOR

For given $\boldsymbol{v} = (\boldsymbol{v_1, v_2, \ldots, v_n}) \in \Re^n$ the **additive inverse** of $\boldsymbol{v}$ is denoted by $-\boldsymbol{v}$ and is given by:

$$-\boldsymbol{v} = (\boldsymbol{-v_1, -v_2, \ldots, -v_n})$$

Note that for every $\boldsymbol{v} = (\boldsymbol{v_1, v_2, \ldots, v_n}) \in \Re^n$:

$$\boldsymbol{v} + (-\boldsymbol{v}) = (\boldsymbol{v_1, v_2, \ldots, v_n}) + (\boldsymbol{-v_1, -v_2, \ldots, -v_n}) = \mathbf{0}$$

We are now in a position to list the properties of Euclidean spaces which will morph into the definition of an abstract vector space in the next section:

THEOREM 2.1 Let $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ be vectors in the Euclidean n-space $\Re^n$, and let r and s be scalars (real numbers). Then:

Addition:

(i) $\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$ (**Commutative Property**)

(ii) $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ (**Vector Associative Property**)

(iii) $\boldsymbol{v}+\boldsymbol{0} = \boldsymbol{v}$ (**Zero Property**)

(iv) $\boldsymbol{v}+(-\boldsymbol{v}) = \boldsymbol{0}$ (**Inverse Property**)

Scalar and Addition:

(v) $r(\boldsymbol{u}+\boldsymbol{v}) = r\boldsymbol{u}+r\boldsymbol{v}$ (**Vector Distributive Property**)

(vi) $(r+s)\boldsymbol{u} = r\boldsymbol{u}+s\boldsymbol{u}$ (**Scalar Distributive Property**)

Scalar:

(vii) $r(s\boldsymbol{u}) = (rs)\boldsymbol{u}$ (**Scalar Associative Property**)

(viii) $1\boldsymbol{u} = \boldsymbol{u}$ (**Identity Property**)

PROOF: We establish (ii) in $\Re^2$, (v) in $\Re^3$, and (vii) in $\Re^3$, and invite you to verify the rest on your own.

To emphasize the important role played by definitions, the symbol $\equiv$ instead of $=$ will temporarily be used to indicate a step in the proof which follows directly from a definition. In addition, the abbreviation "**P**of**R**" will be used to denote that a step follows directly from a **P**roperty of the **R**eal numbers, all of which will be assumed to hold; for example, the additive associative property of the real numbers: $(a+b)+c = a+(b+c)$.

This associative property eliminates the need for including parenthesis when summing more than two vectors. In particular,

$\boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}$

is perfectly well defined.

(ii): $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ (in $\Re^2$):

If $\boldsymbol{u} = (u_1, u_2)$, $\boldsymbol{v} = (v_1, v_2)$, and $\boldsymbol{w} = (w_1, w_2)$, then:

$$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} \equiv [(u_1, u_2)+(v_1, v_2)]+(w_1, w_2)$$

Definition 2.3: $\equiv (u_1+v_1, u_2+v_2)+(w_1, w_2)$

Definition 2.3: $\equiv [(u_1+v_1)+w_1, (u_2+v_2)+w_2]$

POFR: $= [u_1+(v_1+w_1), u_2+(v_2+w_2)]$

Definition 2.3: $\equiv (u_1, u_2)+[(v_1, v_2)+(w_1, w_2)]$

$\equiv \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$

(v): $r(\boldsymbol{u}+\boldsymbol{v}) = r\boldsymbol{u}+r\boldsymbol{v}$ (in $\Re^3$):

If $\boldsymbol{u} = (u_1, u_2, u_3)$ and $\boldsymbol{v} = (v_1, v_2, v_3)$, then:

$$r(\boldsymbol{u}+\boldsymbol{v}) \equiv r[(u_1, u_2, u_3) + (v_1, v_2, v_3)]$$

Definition 2.3: $\equiv r(u_1+v_1, u_2+v_2, u_3+v_3)$

Definition 2.2: $\equiv [r(u_1+v_1), r(u_2+v_2), r(u_3+v_3)]$

Pof**R**: $= (ru_1+rv_1, ru_2+rv_2, ru_3+rv_3)$

Defintion 2.3: $\equiv (ru_1, ru_2, ru_3) + (rv_1, rv_2, rv_3)$

Definition 2.2: $\equiv r(u_1, u_2, u_3) + r(v_1, v_2, v_3)$

$$\equiv r\boldsymbol{u}+r\boldsymbol{v}$$

(vii): $r(s\boldsymbol{u}) = (rs)\boldsymbol{u}$ (in $\boldsymbol{R}^n$):

$$r(s\boldsymbol{u}) \equiv r[s(u_1, u_2, \ldots, u_n)]$$

Definition 2.2: $\equiv r(su_1, su_2, \ldots, su_n)$

Definition 2.2: $\equiv [r(su_1), r(su_2), \ldots, r(su_n)]$

Pof**R**: $= [(rs)u_1, (rs)u_2, \ldots, (rs)u_n]$

Defintion 2.2: $\equiv (rs)(u_1, u_2, \ldots, u_n)$

$$\equiv (rs)\boldsymbol{u}$$

In this, and any other abstract math course:
DEFINITIONS RULE!

Just look at the above proof. It contains but one "logical step," the step labeled **P**of**R**; all other steps hinge on **DEFINITIONS**.

CHECK YOUR UNDERSTANDING 2.2

Answer: See page B-3.

Establish Theorem 2.1(iv), $\boldsymbol{v}+(-\boldsymbol{v}) = \boldsymbol{0}$, in $\Re^3$ and in $\Re^n$.

EXERCISES

Exercises 1-6. Sketch the vector with given initial point A and terminal point B. Sketch the same vector in standard position in the plane, identifying its terminal point.

1. $A = (-2, 1), B = (0, 1)$ 2. $A = (3, 3), B = (0, -1)$ 3. $A = (1, 1), B = (-2, 2)$

4. $A = (1, 0), B = (0, -1)$ 5. $A = (-2, -1), B = (1, -1)$ 6. $A = (2, 2), B = (1, -2)$

Exercises 7-10. Express, as a 3-tuple, the vector with given initial point A and terminal point B.

7. $A = (1, 2, 3), B = (3, 2, 1)$ 8. $A = (-4, 5, 0), B = (2, -5, 1)$

9. $A = (0, 1, -9), B = (-9, 0, 2)$ 10. $A = (-3, 5, -3), B = (3, -5, 3)$

Exercises 11-14. Perform the indicated vector operations.

11. $5(\mathbf{3, -2}) + (\mathbf{0, 1}) + (\mathbf{-2, -4})$ 12. $(\mathbf{2, 5}) + (\mathbf{1, 3}) + [-(\mathbf{-2, 3})]$

13. $-(\mathbf{2, 3, 1, -5}) + [-(\mathbf{1, -2, 0, 0})]$ 14. $-[-(\mathbf{-1, 2, 3, 4})] + (\mathbf{3, 1, 2, -2})$

Exercises 15-18. Find the vector $\boldsymbol{v}$ such that:

15. $\boldsymbol{v} + (\mathbf{2, -4}) = (\mathbf{-4, 2})$ 16. $\boldsymbol{v} + (\mathbf{1, 3, 5}) = (\mathbf{2, 0, -4})$

17. $(\mathbf{4, 2}) + (-\boldsymbol{v}) = (\mathbf{3, 5}) + [-(\mathbf{1, -2})]$ 18. $(\mathbf{4, 3, -1}) + (-\boldsymbol{v}) = \mathbf{2}(\mathbf{1, 3, -2})$

19. For $\boldsymbol{u} = (\mathbf{1, 3})$, $\boldsymbol{v} = (\mathbf{2, 4})$, and $\boldsymbol{w} = (\mathbf{6, -2})$, find scalars r and s such that:

    (a) $r\boldsymbol{u} + s\boldsymbol{v} = \boldsymbol{w}$ (b) $-r\boldsymbol{u} + s\boldsymbol{w} = \boldsymbol{v}$ (c) $r\boldsymbol{v} + (-s\boldsymbol{w}) = \boldsymbol{u}$

20. Find scalars r, s, and t, such that: $r(\mathbf{1, 3, 0}) + s(\mathbf{2, 1, 6}) + t(\mathbf{1, 4, 6}) = (\mathbf{7, 5, 6})$

21. Find scalars r, s, and t, such that: $-r(\mathbf{1, 3, 0}) + s(\mathbf{2, 1, 6}) + [-t(\mathbf{1, 4, 6})] = (\mathbf{7, 5, 6})$

22. Show that there do not exist scalars r, s, and t, such that
$$r(\mathbf{2, 3, 5}) + s(\mathbf{3, 2, 5}) + t(\mathbf{1, 2, 3}) = (\mathbf{1, 2, 4})$$

23. Find the vector $(\boldsymbol{a}, \boldsymbol{b}) \in \Re^2$ of length 5 that has the same direction as the vector with initial point $(1, 3)$ and terminal point $(3, 1)$.

24. Find the vector $(\boldsymbol{a}, \boldsymbol{b}) \in \Re^2$ of length 5 that is in the opposite direction of the vector with initial point $(1, 3)$ and terminal point $(3, 1)$.

25. On page 37, we established Theorem 2.1(ii) for $\Re^2$. Prove that theorem for $\Re^3$ and for $\Re^n$.

26. On page 37, we established Theorem 2.1(v) for $\Re^3$. Prove that theorem for $\Re^2$ and for $\Re^n$.

27. Prove Theorem 2.1(i) for: (a) $\Re^2$ (b) $\Re^3$ (c) $\Re^n$

28. Prove Theorem 2.1(iii) for: (a) $\Re^2$ (b) $\Re^3$ (c) $\Re^n$

29. Prove Theorem 2.1(vi) for: (a) $\Re^2$ (b) $\Re^3$ (c) $\Re^n$

30. Prove Theorem 2.1(viii) for: (a) $\Re^2$ (b) $\Re^3$ (c) $\Re^n$

31. Prove that if $\boldsymbol{v}$, $\boldsymbol{w}$, and $\boldsymbol{z}$, are vectors in $\Re^3$ such that $\boldsymbol{v}+\boldsymbol{w} = \boldsymbol{v}+\boldsymbol{z}$, then $\boldsymbol{w} = \boldsymbol{z}$.

PROVE OR GIVE A COUNTEREXAMPLE

32. For $\boldsymbol{v} \in \Re^n$, if $r\boldsymbol{v} = s\boldsymbol{v}$ then $r = s$.

33. For $\boldsymbol{v} \neq \mathbf{0} \in \Re^n$, if $r\boldsymbol{v} = s\boldsymbol{v}$ then $r = s$.

34. For $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \Re^n$ and $r \neq 0$, if $r\boldsymbol{v}_1 = r\boldsymbol{v}_2$ then $\boldsymbol{v}_1 = \boldsymbol{v}_2$.

35. For $\boldsymbol{v} \in \Re^n$, $r\boldsymbol{v} = \mathbf{0}$ if and only if $r = 0$ or $\boldsymbol{v} = \mathbf{0}$.

§2. ABSTRACT VECTOR SPACES

One of the main objectives of abstract mathematics is to isolate and analyze a particular structure of the real number system, so as to better focus on its "essence." The essence of the vector structure in $\Re^n$ tabulated in Theorem 2.1, page 36, leads us to the definition of an abstract vector space:

The elements of a vector space V are called **vectors**, and will be denoted by bold-faced letters (like $\boldsymbol{v}$). Scalars will continue to be denoted by non-bold-faced letters .

The binary operator need not be represented with a plus-sign—see Example 2.4, page 46.

DEFINITION 2.6

VECTOR SPACE

A **(real) vector space** is a nonempty set V along with two operations, called vector **addition** and **scalar multiplication**. The binary operation of addition assigns to any two element $\boldsymbol{u}$ and $\boldsymbol{v}$ in V, another element $\boldsymbol{u}+\boldsymbol{v}$ in V. The operation of scalar multiplication assigns to any real number r (also called a **scalar**), and any element $\boldsymbol{v}$ in V, another element $r\boldsymbol{v}$ in V. These operations must satisfy the following eight axioms for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, and all $r, s \in \Re$:

Addition:	(i)	$\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$	(**Commutative Axiom**)
	(ii)	$(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$	(**Vector Associative Axiom**)
	(iii)	There is a vector in V, denoted by $\mathbf{0}$ such that $\boldsymbol{v}+\mathbf{0} = \boldsymbol{v}$ for every vector $\boldsymbol{v}$ in V.	(**Zero Axiom**)
	(iv)	For every vector $\boldsymbol{v}$ in V, there is a vector in V, denoted by $-\boldsymbol{v}$ such that $\boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$.	(**Additive Inverse Axiom**)
Scalar and Addition:	(v)	$r(\boldsymbol{u}+\boldsymbol{v}) = (r\boldsymbol{u})+(r\boldsymbol{v})$	(**Vector Distributive Axiom**)
	(vi)	$(r+s)\boldsymbol{v} = (r\boldsymbol{v})+(s\boldsymbol{v})$	(**Scalar Distributive Axiom**)
Scalar:	(vii)	$r(s\boldsymbol{v}) = (rs)\boldsymbol{v}$	(**Scalar Associative Axiom**)
	(viii)	$1\boldsymbol{v} = \boldsymbol{v}$	(**Identity Axiom**)

While eight axioms are specifically listed in the above definition, two more are lurking within the above so-called closure statements:

A set is said to be **closed**, with respect to an operation, if elements of that set subjected to that operation remain in the set.

V is closed under addition:	For every $\boldsymbol{v}$ and $\boldsymbol{w}$ in V, $\boldsymbol{v}+\boldsymbol{w} \in V$
V is closed under scalar multiplication:	For every $\boldsymbol{v} \in V$ and $r \in \Re$, $r\boldsymbol{v} \in V$

It is important to note that while the two plus signs in $(r+s)\boldsymbol{v} = (r\boldsymbol{v})+(s\boldsymbol{v})$ are identical in appearance, they do not represent a common operator:

> The "+" in $(r+s)\boldsymbol{v}$ denotes the sum of two real numbers, as in $2+5 = 7$, while the "+" in $(r\boldsymbol{v})+(s\boldsymbol{v})$ denotes the sum of two vectors, as in $2\boldsymbol{v}+5\boldsymbol{v} = 7\boldsymbol{v}$.

By the same token, the two "products" in the expression $(rs)\boldsymbol{v}$ also denote distinct operators:

> The rs in $(rs)\boldsymbol{v}$ denotes the product of two real numbers, resulting in another number, as in $2\cdot 5 = 10$, while the scalar product $10\boldsymbol{v}$ represents a vector.

We also point out that, by convention, no meaning is attributed to an expression of the form $\boldsymbol{v}r$, wherein a vector $\boldsymbol{v}$ appears to the left of a scalar r.

We already have infinitely many vector spaces at our disposal, namely the Euclidean *n*-spaces. We now turn our attention to several others.

MATRIX SPACES

EXAMPLE 2.3 The set of two-by-two matrices:

$$M_{2\times 2} = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a, b, c, d \in R \right\}$$

with addition and scalar multiplication given by:

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

$$r\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$$

is a vector space.

SOLUTION: We content ourselves with verifying Axiom (iii) (the zero axiom), and Axiom (iv) (the additive inverse axiom).

Axiom (iii): Let $\mathbf{0} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$. Then, for any $\boldsymbol{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ we have:

$$\boldsymbol{v}+\mathbf{0} \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \equiv \begin{bmatrix} a+0 & b+0 \\ c+0 & d+0 \end{bmatrix} \overset{\textbf{PofR}}{=} \begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \boldsymbol{v}$$

We are again using ≡ to indicate that equality follows from a definition, and "**P**of**R**" for "**P**roperty of the **R**eal numbers."

Axiom (iv): For any given $\boldsymbol{v} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we show there exists a vector $-\boldsymbol{v}$,

namely $-\boldsymbol{v} = \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix}$, such that $\boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$:

$$\boldsymbol{v} + (-\boldsymbol{v}) \equiv \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} -a & -b \\ -c & -d \end{bmatrix} \equiv \begin{bmatrix} (a-a) & (b-b) \\ (c-c) & (d-d) \end{bmatrix} \overset{\textbf{PofR}}{=} \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \equiv \mathbf{0}$$

Generalizing Example 2.3 to accommodate matrices of all dimensions, we have:

THEOREM 2.2

MATRIX SPACE

Let $M_{m \times n}$ denote the set of all $m \times n$ matrices.

For $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$ in $M_{m \times n}$, let:

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

(The ij^{th} entry of the sum matrix is the sum of the ij^{th} entry in matrix $\boldsymbol{A}$ with the ij^{th} entry in the matrix $\boldsymbol{B}$.)

For $r \in \Re$ and $\boldsymbol{A} = [a_{ij}] \in M_{m \times n}$, let:

$$r\boldsymbol{A} = r[a_{ij}] = [ra_{ij}]$$

(The ij^{th} entry in the matrix $r\boldsymbol{A}$ is r times the ij^{th} entry in the matrix $\boldsymbol{A}$.)

The set $M_{m \times n}$ with the above operations is a vector space.

PROOF: We content ourselves with verifying Axioms (i) and (vi).

Axiom (i): For every $\boldsymbol{A} = [a_{ij}]$ and $\boldsymbol{B} = [b_{ij}]$:

$$\boldsymbol{A} + \boldsymbol{B} \equiv [a_{ij}] + [b_{ij}] \equiv [a_{ij} + b_{ij}] \overset{\textbf{PofR}}{=} [b_{ij} + a_{ij}] \equiv [b_{ij}] + [a_{ij}] \equiv \boldsymbol{B} + \boldsymbol{A}$$

Axiom (vi): For scalars r and s, and $\boldsymbol{A} = [a_{ij}]$:

$$\begin{aligned}(r+s)\boldsymbol{A} \equiv (r+s)[a_{ij}] \equiv [(r+s)a_{ij}] &\overset{\textbf{PofR}}{=} [ra_{ij} + sa_{ij}] \\ &\equiv [ra_{ij}] + [sa_{ij}] \\ &\equiv r[a_{ij}] + s[a_{ij}] \equiv r\boldsymbol{A} + s\boldsymbol{A}\end{aligned}$$

CHECK YOUR UNDERSTANDING 2.3

Verify the associative axiom $r(s\boldsymbol{v}) = (rs)\boldsymbol{v}$ for the vector space $M_{m \times n}$ of Theorem 2.2.

Answer: See page B-3.

POLYNOMIAL SPACES

A function of the form $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$, with $a_n \neq 0$ is said to be a **polynomial function** of degree *n*. For any given integer $n \geq 0$, P_n will represent the set of polynomials of degree less than or equal to *n*

In particular:
$P_0(x) = \{a_0 | a_0 \in \Re\} = \Re$

We note that the polynomial:

$$a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0$$

can be written the other way around:

$$a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

and can also be expressed in Sigma-notation form:

$$\sum_{i=0}^{n} a_i x^i = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n$$

The Greek letter $\sum$ (Sigma) is used to denote a sum.

THEOREM 2.3

POLYNOMIAL SPACES

The set of polynomials P_n of degree less than or equal to *n*, with operations:

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} (a_i + b_i) x^i$$

$$r\left(\sum_{i=0}^{n} a_i x^i\right) = \sum_{i=0}^{n} r a_i x^i$$

is a vector space.

PROOF: We establish the two distributive axioms, and relegate the remaining six to the exercises.

Axiom (v) $r(\boldsymbol{u} + \boldsymbol{v}) = (r\boldsymbol{u}) + (r\boldsymbol{v})$:

$$r\left(\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i\right) \equiv r\sum_{i=0}^{n} (a_i + b_i) x^i \equiv \sum_{i=0}^{n} r(a_i + b_i) x^i$$

$$\underset{PofR}{\overset{\uparrow}{=}} \sum_{i=0}^{n} (ra_i + rb_i) x^i \equiv \sum_{i=0}^{n} ra_i x^i + \sum_{i=0}^{n} rb_i x^i$$

$$\equiv \left(r\sum_{i=0}^{n} a_i x^i\right) + \left(r\sum_{i=0}^{n} b_i x^i\right)$$

Axiom (vi) $(r+s)\boldsymbol{v} = (r\boldsymbol{v}) + (s\boldsymbol{v})$:

$$(r+s)\sum_{i=0}^{n} a_i x^i \equiv \sum_{i=0}^{n}(r+s)a_i x^i \overset{\textbf{PofR}}{=} \sum_{i=0}^{n}(ra_i x^i + sa_i x^i)$$

$$\equiv \sum_{i=0}^{n} ra_i x^i + \sum_{i=0}^{n} sa_i x^i \equiv \left(r\sum_{i=0}^{n} a_i x^i\right) + \left(s\sum_{i=0}^{n} a_i x^i\right)$$

CHECK YOUR UNDERSTANDING 2.4

Referring to Theorem 2.3, verify the commutative axiom:

$$\sum_{i=0}^{n} a_i x^i + \sum_{i=0}^{n} b_i x^i = \sum_{i=0}^{n} b_i x^i + \sum_{i=0}^{n} a_i x^i$$

Answer: See page B-3.

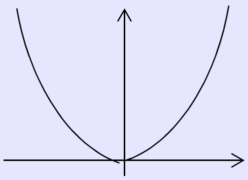
FUNCTION SPACES

You are probably accustomed of thinking that a function is some sort of dynamic creature that "takes numbers to numbers." At this point, however, you want to think of a function as being an object, in the same way that you see the number 5 as an object. Indeed, the set of all such functions $f: X \to \Re$ from a set X (the **domain** of the function) to the set $\Re$ of real numbers, can be turned into a vector space:

All "objects" in mathematics are sets, and functions are no exceptions. The function $\boldsymbol{f}$ given by $f(x) = x^2$, for example, is that subset of the plane, typically called the graph of f:

$$\boldsymbol{f} = \{(x, x^2) | (x \in R)\}$$

Pictorially:



THEOREM 2.4
FUNCTION SPACE

Let $F(X)$ denote the set of all **real-valued** functions defined on a non-empty set X:

$$F(X) = \{\boldsymbol{f} | \boldsymbol{f}: X \to \boldsymbol{R}\}$$

For $\boldsymbol{f}$ and $\boldsymbol{g}$ in $F(X)$, and $r \in \boldsymbol{R}$, let $\boldsymbol{f}+\boldsymbol{g}$, and $r\boldsymbol{f}$ be given by:

$$(\boldsymbol{f}+\boldsymbol{g})(x) = f(x) + g(x) \quad (*)$$
$$\text{and} \quad (r\boldsymbol{f})(x) = rf(x)$$

With respect to these operations, $F(X)$ is a vector space.

A function $\boldsymbol{f}: X \to \Re$ is defined to be equal to a function $\boldsymbol{g}: X \to \Re$, if $f(x) = g(x)$ for every $x \in X$.

PROOF: We verifying Axioms (i), (iii), (iv) and (v):

The fact that $F(X)$ is closed under addition and scalar multiplication is self-evident.

Axiom (i) (Commutative Axiom). For $\boldsymbol{f}, \boldsymbol{g} \in F(X)$ and $x \in X$:

$$(\boldsymbol{f}+\boldsymbol{g})(x) \equiv f(x) + g(x) \overset{\text{PofR}}{=} g(x) + f(x) \equiv (\boldsymbol{g}+\boldsymbol{f})(x)$$

Since $(\boldsymbol{f}+\boldsymbol{g})(x) = (\boldsymbol{g}+\boldsymbol{f})(x)$ for every $x \in X$, $\boldsymbol{f}+\boldsymbol{g} = \boldsymbol{g}+\boldsymbol{f}$.

As you can see, we elected to use the letter $\boldsymbol{Z}$, rather than the symbol $\mathbf{0}$, for our zero vector. It's just that an expression like $\mathbf{0}(x)$ would strongly suggest that a multiplication by zero is being performed, which is not the case.

Axiom (iii) (Zero Axiom). Let $\boldsymbol{Z} : X \rightarrow \Re$ be the function given by $\boldsymbol{Z}(x) = 0$ for every $x \in X$. For any $f \in F$ and $x \in X$:

$$(\boldsymbol{f}+\boldsymbol{Z})(x) \equiv f(x) + Z(x) \equiv f(x) + 0 \overset{\text{PofR}}{=} f(x) \equiv \boldsymbol{f}(x)$$

Since $(\boldsymbol{f}+\boldsymbol{Z})(x) = \boldsymbol{f}(x)$ for every $x \in X$, $\boldsymbol{f}+\boldsymbol{Z} = \boldsymbol{f}$.

Axiom (iv) (Additive Inverse Axiom). For given $\boldsymbol{f} \in F(X)$ let $-\boldsymbol{f}$ be the function given by $(-\boldsymbol{f})(x) = -f(x)$. Then, for any $x \in X$:

$$[\boldsymbol{f}+(-\boldsymbol{f})](x) \equiv f(x) + (-f(x)) \overset{\text{PofR}}{=} 0 \equiv \boldsymbol{Z}(x)$$

Since $[\boldsymbol{f}+(-\boldsymbol{f})](x) = \boldsymbol{Z}(x)$ for every $x \in X$, $\boldsymbol{f}+(-\boldsymbol{f}) = \boldsymbol{Z}$.

Axiom (v) [Distributive Axiom (vector)]: For any $\boldsymbol{f}, \boldsymbol{g} \in F(X)$, $x \in X$ and $r \in \boldsymbol{R}$:

$$\begin{aligned}[r(\boldsymbol{f}+\boldsymbol{g})](x) \equiv r[(\boldsymbol{f}+\boldsymbol{g})(x)] \equiv r[f(x)+g(x)] &\overset{\text{PofR}}{=} rf(x) + rg(x) \\ &\equiv (r\boldsymbol{f})(x) + (r\boldsymbol{g})(x)\end{aligned}$$

Since $[r(\boldsymbol{f}+\boldsymbol{g})](x) = (r\boldsymbol{f})(x) + (r\boldsymbol{g})(x)$ for every $x \in X$,
$r(\boldsymbol{f}+\boldsymbol{g}) = r\boldsymbol{f} + r\boldsymbol{g}$.

CHECK YOUR UNDERSTANDING 2.5

Verify the distributive axiom $(r+s)\boldsymbol{v} = (r\boldsymbol{v}) + (s\boldsymbol{v})$ for the function space of Theorem 2.4.

Answer: See page B-3.

ADDITIONAL EXAMPLES

As you will see in the next two examples, addition and scalar multiplication in a vector space can be somewhat "counter-intuitive." Moreover, both the zero vector $\mathbf{0}$ and the additive inverse vector $-\boldsymbol{v}$, may appear somewhat strange in a vector space.

EXAMPLE 2.4 Show that the set $\Re^+ = \{x|x > 0\}$ of positive real numbers with a binary operator of multiplication: ab and scalar operation: a^r is a vector space.

SOLUTION: $\Re^+$ is certainly closed under both of the above operations. Moreover:

Axiom (i). For every $a, b \in \Re^+$: $ab = ba$

Axiom (ii). For every $a, b, c \in \Re^+$: $a(bc) = (ab)c$

Axiom (iii). For every $a \in \Re^+$: $a \cdot 1 = a$, so 1 is the zero vector.

Axiom (iv). For every $a \in \Re^+$: $a \cdot \frac{1}{a} = 1$, so $\frac{1}{a}$ is the inverse of a.

Axiom (v). For every $a, b \in \Re^+$**,** and every $r \in \Re$, $(\boldsymbol{ab})^r = a^r b^r$.

You are invited to establish the remaining three axioms of Definition 2.6 in the exercises, thereby establishing the fact that $\Re^+$, with given operations, is a vector space.

EXAMPLE 2.5 Show that the set $V = \{(x, y)|x, y \in \Re\}$ under "addition":

$$(\boldsymbol{x}, \boldsymbol{y}) + (\boldsymbol{x}', \boldsymbol{y}') = (x + x' - 1, y + y' + 1)$$

and scalar multiplication

$$\boldsymbol{r}(\boldsymbol{x}, \boldsymbol{y}) = (rx - r + 1, ry + r - 1)$$

is a vector space:

SOLUTION: V is certainly closed under both of the above operations. We content ourselves by establishing the zero and inverse axioms, and leave it for you to verify the remaining six axioms in the exercises.

Zero Axiom: Does there exist a vector $\mathbf{0}$ such that $\boldsymbol{v} + \mathbf{0} = \boldsymbol{v}$ for every $\boldsymbol{v} \in V$? Don't be to quick to say "no," basing you answer on the observation that

$$(x, y) + (0, 0) = (x + 0 - 1, y + 0 + 1) = (x - 1, y + 1) \neq (x, y)$$

But you have no right to assume that if a zero vector exists, then it must be the one you would like it to be! Putting partiality aside, let's see if we can find a vector $\mathbf{0} = (a, b)$ such that $(x, y) + \mathbf{0} = (x, y)$, for every $(x, y) \in V$:

$$\begin{aligned}(x,y)+\mathbf{0} &= (x,y)\\ (x,y)+(\boldsymbol{a},\boldsymbol{b}) &= (x,y)\\ (x+a-1,y+b+1) &= (x,y) \Rightarrow \begin{cases} x+a-1 = x \Rightarrow \boldsymbol{a} = \mathbf{1} \\ y+b+1 = y \Rightarrow \boldsymbol{b} = \mathbf{-1} \end{cases}\end{aligned}$$

That's right, in this vector space, $\mathbf{0} = (1,-1)$, for:

$$(x,y)+(\mathbf{1},-1) = (x+\mathbf{1}-1, y+(-1)+1) = (x,y)$$

Additive Inverse Axiom: Now that we have a zero vector, $\mathbf{0} = (1,-1)$, we can ask if, for any given $\boldsymbol{v} = (x,y)$, there exists a vector $-\boldsymbol{v} = (a,b)$ such that $(x,y)+(a,b) = (1,-1)$. There does:

$$\begin{aligned}(x,y)+(\boldsymbol{a},\boldsymbol{b}) &= (1,-1)\\ (x+a-1,y+b+1) &= (1,-1) \Rightarrow \begin{cases} x+a-1 = 1 \Rightarrow \boldsymbol{a} = -\boldsymbol{x}+\mathbf{2} \\ y+b+1 = -1 \Rightarrow \boldsymbol{b} = -\boldsymbol{y}-\mathbf{2} \end{cases}\end{aligned}$$

So, the additive inverse of (x,y) turns out to be $(-x+2,-y-2)$, for:

$$(x,y)+(-\boldsymbol{x}+\mathbf{2},-\boldsymbol{y}-\mathbf{2}) = (x-\boldsymbol{x}+\mathbf{2}-1, y-\boldsymbol{y}-\mathbf{2}+1) = \underset{\substack{\uparrow\\ \text{the zero vector}}}{(1,-1)}$$

CHECK YOUR UNDERSTANDING 2.6

Verify that Axioms (iii) and (iv) of Definition 2.6 are satisfied for the set $V = \{(x,y,z) \mid x,y,z \in \Re\}$ with imposed addition:

$$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})+(\boldsymbol{x'},\boldsymbol{y'},\boldsymbol{z'}) = (\boldsymbol{x}+\boldsymbol{x'}-1, \boldsymbol{y}+\boldsymbol{y'}+\mathbf{2}, \boldsymbol{z}+\boldsymbol{z'}-\mathbf{3})$$

and scalar multiplication:

$$r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = (\boldsymbol{rx}-\boldsymbol{r}+\mathbf{1}, \boldsymbol{ry}-\mathbf{2}\boldsymbol{r}-\mathbf{2}, \boldsymbol{rz}-\mathbf{3}\boldsymbol{z}+\mathbf{3})$$

Answer: Zero vector: $(\mathbf{1},-\mathbf{2},\mathbf{3})$
Inverse of $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})$: $(-\boldsymbol{x}+\mathbf{2},-\boldsymbol{y}+\mathbf{4},-\boldsymbol{z}+\mathbf{6})$

EXAMPLE 2.6

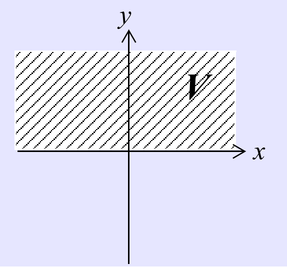
Let V denote the upper-half plane:

$$V = \{(x,y) \mid x,y \in \Re, y \geq 0\}$$

with standard addition and scalar multiplication:

$$\begin{aligned}(x,y)+(x',y') &= (x+x',y+y')\\ r(x,y) &= (rx,ry)\end{aligned}$$

Is V a vector space?



SOLUTION: No, since V is not closed under scalar multiplication. But we can't just say this, for our claim has to be established. We have to demonstrate that the scalar product involving some **specific** element of V and some **specific** real number ends up being outside of V, and so we shall:

$$(5,2) \in V \quad \text{but} \quad -7(5,2) = (-35,-14) \notin V$$

EXAMPLE 2.7 Let X be the set of two-tuples with operations:

$$\begin{aligned}(x, y) + (x', y') &= (x + x', y + y') \\ r(x, y) &= (rx, y)\end{aligned}$$

Is X a vector space?

SOLUTION: X is certainly closed under both operations. We need not challenge the first four axioms of Definition 2.6, as they involve only the addition operator which coincides with that of Euclidean two-space. The scalar operator, however, is a bit different from that of $\Re^2$, and we must therefore determine whether or not Axioms (v) through (viii) are satisfied. What you may want to do is to quickly check to see if they hold for some specific vectors and scalars:

Let's Challenge Axiom (vi), $(r + s)\boldsymbol{v} = (r\boldsymbol{u}) + (s\boldsymbol{v})$, with $r = 7$, $s = 3$, and $\boldsymbol{u} = (4, -2)$:

$$(7 + 3)(4, -2) = 10(4, -2) = (40, -2)$$

and $7(4, -2) + 3(4, -2) = (28, -2) + (12, -2) = (40, -4)$ Oops!

We need go no further, for the above shows that Axiom (vi) does not hold, and can therefore conclude that under the given operations, X is not a vector space.

CHECK YOUR UNDERSTANDING 2.7

THE TRIVIAL VECTOR SPACE:

Define addition and scalar multiplication on the set $V = \{\mathbf{0}\}$ and verify that V is a vector space with respect to those operations.

Answer: See page B-4.

EXERCISES

Exercises 1-13. Verify that the set S, with given operations, fails to be a vector space.

1. $S = \Re$, $x + y = x - y$, and $r\boldsymbol{x} = rx$.

2. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (x + x', 0)$, and $r(x, y) = (rx, ry)$.

3. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (xx', yy')$, and $r(x, y) = (rx, ry)$.

4. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (0, 0)$, and $r(x, y) = (rx, ry)$.

5. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (2x + 2x', y + y')$, and $r(x, y) = (rx, ry)$.

6. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (x + x', y + y')$, and $r(x, y) = (0, 0)$.

7. $S = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (x + y', y + x')$, and $r(x, y) = (rx, ry)$.

8. $S = \{(x, y, 0) | x, y \in \Re\}$, $(x, y, 0) + (x', y', 0) = (x + x', y + y', 0)$, and
$r(x, y, 0) = (rx, ry, 0)$.

9. $S = M_{2\times 2}$, $\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a + a' & 0 \\ 0 & d + d' \end{bmatrix}$, and $r\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$.

10. $S = M_{2\times 2}$, $\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a + c' & b + d' \\ c + a' & d + b' \end{bmatrix}$, and $r\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$.

11. $S = \{ax^2 + bx + c\}$, $(ax^2 + bx + c) + (a'x^2 + b'x + c') = (a + a')x^2 + (b + b')x$ and
$r(ax^2 + bx + c) = (ra)x^2 + (rb)x + (rc)$.

12. $S = \{0, 1\}$; $0 + 0 = 0$, $1 + 1 = 0$, $0 + 1 = 1 + 0 = 1$; and $r0 = 0$, $r1 = 1$.

13. $S = \{\boldsymbol{a} \in \Re | \boldsymbol{a} > 0\}$, $\boldsymbol{a} + \boldsymbol{b} = \ln a + \ln b$, and $r\boldsymbol{a} = \ln a^r$.

Exercises 14-16. Verify that the set V, with given operations, is a vector space.

14. $V = \{1\}$, $1 + 1 = 1$, and $r1 = 1$.

15. $V = \{(x, y, 0) | x, y \in \Re\}$, $(x, y, 0) + (x', y', 0) = (x + x', y + y', 0)$, and
$r(x, y, 0) = (rx, ry, 0)$.

16. $V = \{(x, y) | x, y \in \Re\}$, $(x, y) + (x', y') = (x + x' + 2, y + y')$,
$r(x, y) = (rx + 2r - 2, ry)$.

17. Complete the proof of Theorem 2.2.

18. Complete the proof of Theorem 2.3.

19. Complete the proof of Theorem 2.4.

20. Establish the remaining three axioms for the space of Example 2.4.

21. Establish the remaining six axioms for the space of Example 2.5.

22. A polynomial is an expression of the form $p(x) = \sum_{i=0}^{\infty} a_i x^i$ for which there exists an m such that $a_i = 0$ for $i > m$. Show that, with respect to the following operations, the set P of all polynomials is a vector space:

$$\sum_{i=0}^{\infty} a_i x^i + \sum_{i=0}^{\infty} b_i x^i = \sum_{i=0}^{\infty} (a_i + b_i) x^i \text{ and } \mathrm{r} \sum_{i=0}^{\infty} a_i x^i = \sum_{i=0}^{\infty} (r a_i) x^i$$

PROVE OR GIVE A COUNTEREXAMPLE

23. Let V be a vector space, and let $\boldsymbol{v} \in V$. If $r\boldsymbol{v} = \mathbf{0}$, then $r = 0$.

24. Let V be a vector space, and let $\boldsymbol{v} \in V$. If $r\boldsymbol{v} = \mathbf{0}$ and $\boldsymbol{v} \neq \mathbf{0}$, then $r = 0$.

25. Let V be a vector space, and let $\boldsymbol{v} \in V$. If $r\boldsymbol{v} = \mathbf{0}$ and $r \neq 0$, then $\boldsymbol{v} = \mathbf{0}$.

26. Let V be a vector space, and let $\boldsymbol{v} \in V$. If $r\boldsymbol{v} = s\boldsymbol{v}$, then $r = s$.

27. Let V be a vector space, and let $\boldsymbol{v} \in V$. If $r\boldsymbol{v} = s\boldsymbol{v}$ and $\boldsymbol{v} \neq \mathbf{0}$, then $r = s$.

§3. PROPERTIES OF VECTOR SPACES

For the sake of convenience, we again list the vector space axioms:

Addition:	(i) $\mathbf{u}+\mathbf{v} = \mathbf{v}+\mathbf{u}$	(**Commutative Axiom**)
	(ii) $(\mathbf{u}+\mathbf{v})+\mathbf{w} = \mathbf{u}+(\mathbf{v}+\mathbf{w})$	(**Vector Associative Axiom**)
	(iii) There is a vector in V, denoted by $\mathbf{0}$ such that $\mathbf{v}+\mathbf{0} = \mathbf{v}$ for every vector $\mathbf{v}$ in V.	(**Zero Axiom**)
	(iv) For every vector $\mathbf{v}$ in V, there is a vector in V, denoted by $-\mathbf{v}$ such that $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$.	(**Additive Inverse Axiom**)
Scalar and Addition:	(v) $r(\mathbf{u}+\mathbf{v}) = (r\mathbf{u})+(r\mathbf{v})$	(**Vector Distributive Axiom**)
	(vi) $(r+s)\mathbf{v} = (r\mathbf{v})+(s\mathbf{v})$	(**Scalar Distributive Axiom**)
Scalar:	(vii) $r(s\mathbf{v}) = (rs)\mathbf{v}$	(**Scalar Associative Axiom**)
	(viii) $1\mathbf{v} = \mathbf{v}$	(**Identity Axiom**)

For aesthetic reasons, a set of axioms should be independent, in that no part of an axiom is a consequence of the rest. One should not, for example, replace Axiom (iii) with:

> There is a vector in V, denoted by $\mathbf{0}$ such that $\mathbf{v}+\mathbf{0} = \mathbf{v}$ and $\mathbf{0}+\mathbf{v} = \mathbf{v}$ for every vector $\mathbf{v}$ in V.

Reason: Axiom (i) already implies that the $\mathbf{0}$ of Axiom (iii) can be on either side of the $\mathbf{v}$. The same can be said about the vector $-v$ in Axiom (iv). Bringing us to:

THEOREM 2.5 Let V be a vector space.

(a) For every vector $\mathbf{v}$ in V:

$$\mathbf{v}+\mathbf{0} = \mathbf{0}+\mathbf{v} = \mathbf{v}$$

(b) For every vector $\mathbf{v}$ in V, there exists a vector $-\mathbf{v}$ such that:

$$\mathbf{v}+(-\mathbf{v}) = (-\mathbf{v})+\mathbf{v} = -\mathbf{v}+\mathbf{v} = 0$$

Our next theorem tells us that there is **but one 0** vector in any vector space, and that every vector $\mathbf{v}$ has a **unique** additive inverse $-\mathbf{v}$. While you might have taken these two facts for granted, neither is given to you free of charge:

THEOREM 2.6 Let V be any vector space, then:

(a) There is but one vector $\mathbf{0}$ which satisfies the property that $\boldsymbol{v}+\mathbf{0} = \boldsymbol{v}$ for every $\boldsymbol{v}$ in V.

(b) For any given vector $\boldsymbol{v}$ in V, there is but one vector $-\boldsymbol{v}$ in V such that $\boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$.

PROOF:

Strategy for (a): Assume that $\mathbf{0}$ and $\bar{\mathbf{0}}$ are any two zeros, and then go on to show $\mathbf{0} = \bar{\mathbf{0}}$.

$\boldsymbol{v}+\mathbf{0} = \boldsymbol{v}$ and $\boldsymbol{v}+\bar{\mathbf{0}} = \boldsymbol{v}$

$\downarrow$

$\mathbf{0} = \bar{\mathbf{0}}$

Let $\mathbf{0}$ and $\bar{\mathbf{0}}$ be such that, for every vector $\boldsymbol{v}$:

$$(*)\ \ \boldsymbol{v}+\mathbf{0} = \boldsymbol{v} \quad \text{and} \quad (**)\ \boldsymbol{v}+\bar{\mathbf{0}} = \boldsymbol{v}$$

Substituting $\bar{\mathbf{0}}$ for $\boldsymbol{v}$ in (*), we have (i): $\bar{\mathbf{0}}+\mathbf{0} = \bar{\mathbf{0}}$.

Substituting $\mathbf{0}$ for $\boldsymbol{v}$ in (**), we also have (ii): $\mathbf{0}+\bar{\mathbf{0}} = \mathbf{0}$.

Then:

$$\bar{\mathbf{0}} \overset{\overset{\text{(i)}}{\downarrow}}{=} \bar{\mathbf{0}}+\mathbf{0} \underset{\underset{\text{commutativity}}{\uparrow}}{=} \mathbf{0}+\bar{\mathbf{0}} \overset{\overset{\text{(ii)}}{\downarrow}}{=} \mathbf{0}$$

Strategy for (b): Assume that a vector $\boldsymbol{v}$ has two additive inverses, $-\boldsymbol{v}$ and $\neg\boldsymbol{v}$, and then go on to show that $-\boldsymbol{v} = \neg\boldsymbol{v}$.

$\boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$ and $\boldsymbol{v}+(\neg\,\boldsymbol{v}) = \boldsymbol{v}$

$\downarrow$

$-\boldsymbol{v} = \neg\boldsymbol{v}$

Let $-\boldsymbol{v}$ and $\neg\boldsymbol{v}$ be such that:

$$(*)\ \ \boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0} \text{ and } (**)\ \ \boldsymbol{v}+(\neg\,\boldsymbol{v}) = \mathbf{0}$$

Adding $\neg\boldsymbol{v}$ to both sides of (*) we have:

$$\begin{aligned} & \neg\boldsymbol{v}+[\boldsymbol{v}+(-\boldsymbol{v})] = \neg\boldsymbol{v}+\mathbf{0} \\ \text{Axioms (ii) and (iii):}\quad & (\neg\boldsymbol{v}+\boldsymbol{v})+(-\boldsymbol{v}) = \neg\boldsymbol{v} \\ \text{Axiom (i) and (**):}\quad & \mathbf{0}+(-\boldsymbol{v}) = \neg\boldsymbol{v} \\ \text{Theorem 2.5(a):}\quad & -\boldsymbol{v} = \neg\boldsymbol{v} \end{aligned}$$

The above proof illustrates the important fact that a mathematical theory is based on a set of rules or axioms, on which sit logically derived results, or theorems. Once established, a theorem can be used to prove other theorems. At some point, the axioms and theorems kind of blend into each other—they are just facts, with some of them being dictated (the axioms), while others are established (the theorems).

Answer: See page B-4.

CHECK YOUR UNDERSTANDING 2.8

Show that in any vector space:

$$\text{If:}\quad \mathbf{v}+\mathbf{z} = \mathbf{w}+\mathbf{z}$$
$$\text{Then:}\quad \mathbf{v} = \mathbf{w}$$

Two different zeros come into play in the following theorem:

> The real number 0 that is involved in the scalar product at the left of the equality, and the vector **0** that appears to the right of the equality.

THEOREM 2.7 For any vector $\mathbf{v}$ in a vector space V:

$$0\mathbf{v} = \mathbf{0}$$

PROOF: At times, as is the case here, a proof almost writes itself, once an appropriate initial step is taken (in this case, to write 0 as $0+0$):

$$0\mathbf{v} \overset{\overset{\textbf{PofR}}{\downarrow}}{=} (0+0)\mathbf{v} \overset{\overset{\text{Axiom (ii) (Distributive Axiom)}}{\downarrow}}{=} 0\mathbf{v}+0\mathbf{v}$$

The end is now in sight: just add the additive inverse of the vector $0\mathbf{v}$ to both sides of the equation:

$$\begin{aligned}
0\mathbf{v} &= 0\mathbf{v}+0\mathbf{v}\\
-0\mathbf{v}+0\mathbf{v} &= -0\mathbf{v}+(0v+0v)\\
\text{Axiom (iv) and (ii):}\qquad \mathbf{0} &= (-0\mathbf{v}+0\mathbf{v})+0\mathbf{v}\\
\text{Axiom (iv):}\qquad \mathbf{0} &= \mathbf{0}+0\mathbf{v}\\
\text{Theorem 2.5:}\qquad \mathbf{0} &= 0\mathbf{v}
\end{aligned}$$

In words, the above theorem tells us that:

Multiplying any vector by the scalar 0 results in the vector **0**.

Answer: See page B-4.

CHECK YOUR UNDERSTANDING 2.9

Prove that in every vector space V, $r\mathbf{0} = \mathbf{0}$ for every $r \in \Re$.

The above CYU together with Theorem 2.7 tells us that if either $r = 0$ or $\mathbf{v} = \mathbf{0}$, then the scalar product $r\mathbf{v} = \mathbf{0}$. The converse also hods:

THEOREM 2.8 In any vector space V:

$$\text{If } r\mathbf{v} = \mathbf{0} \text{ then } r = 0 \text{ or } \mathbf{v} = \mathbf{0}$$

PROOF: If $r = 0$, then surely the statement $r = 0$ or $\boldsymbol{v} = \boldsymbol{0}$ holds, and we are done. If $r \neq 0$, then:

$$\begin{aligned} r\boldsymbol{v} &= \boldsymbol{0} \\ \frac{1}{r}(r\boldsymbol{v}) &= \frac{1}{r}\boldsymbol{0} \\ \text{Axiom (vii) and CYU 2.8:} \quad \left(\frac{1}{r}\cdot r\right)\boldsymbol{v} &= \boldsymbol{0} \\ 1\boldsymbol{v} &= \boldsymbol{0} \\ \text{Axiom (viii):} \quad \boldsymbol{v} &= \boldsymbol{0} \end{aligned}$$

CHECK YOUR UNDERSTANDING 2.10

Establish the following Cancellation Properties:

(a) If $r \neq 0$ and $r\boldsymbol{v} = r\boldsymbol{w}$, then $\boldsymbol{v} = \boldsymbol{w}$.

(b) If $\boldsymbol{v} \neq \boldsymbol{0}$ and $r\boldsymbol{v} = s\boldsymbol{v}$, then $r = s$.

Answer: See page B-4.

Here is what the next theorem is saying:

Multiplying any vector by the scalar -1 results in the additive inverse of that vector.

THEOREM 2.9 For any vector $\boldsymbol{v}$ in a vector space V:

$$-1\boldsymbol{v} = -\boldsymbol{v}$$

PROOF:

Strategy: Show that if you add $-1\boldsymbol{v}$ to $\boldsymbol{v}$ you end up with the vector **0.**

$$-1\boldsymbol{v}+\boldsymbol{v} \overset{\text{Axiom (viii)}}{=} -1\boldsymbol{v}+1\boldsymbol{v} \underset{\text{Axiom (vi)}}{=} (-1+1)\boldsymbol{v} \underset{\text{PofR}}{=} 0\boldsymbol{v} \underset{\text{Theorem 2.7}}{=} \boldsymbol{0}$$

CHECK YOUR UNDERSTANDING 2.11

Establish the following results, for any $\boldsymbol{v}$ in a vector space V, and any $r \in \Re$.

(a) $-(-\boldsymbol{v}) = \boldsymbol{v}$ (b) $(-r)\boldsymbol{v} = -(r\boldsymbol{v})$ (c) $r(-\boldsymbol{v}) = -(r\boldsymbol{v})$

Answer: See page B-5.

SUBTRACTION

We all want to replace the expression $\boldsymbol{v}+(-\boldsymbol{w})$ with $\boldsymbol{v}-\boldsymbol{w}$. Let's do it, but officially:

DEFINITION 2.7
"SUBTRACTION"

For vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ in a vector space V, we define $\boldsymbol{v}$ **minus** $\boldsymbol{w}$, denoted by $\boldsymbol{v}-\boldsymbol{w}$, to be the vector given by:

$$\boldsymbol{v}-\boldsymbol{w} = \boldsymbol{v}+(-\boldsymbol{w})$$

A definition is the introduction of a new word in the language of mathematics. As such, one must understand all of the words used in its description. This is so in Definition 2.7, where the new word "$\boldsymbol{v}-\boldsymbol{w}$" on the left of the equal sign is described by previously defined words "$\boldsymbol{v}+(-\boldsymbol{w})$" on the right of the equal sign.

Here are a few results featuring the operation of subtraction. They are very reminiscent of familiar subtraction operations of real numbers. This should come as no surprise since the real number system is itself a vector space.

THEOREM 2.10

For any vectors $\boldsymbol{v}$, $\boldsymbol{w}$, and $\boldsymbol{z}$ in a vector space V, and scalars r, and s in $\Re$:

(a) $\boldsymbol{v}-\boldsymbol{v} = \mathbf{0}$

(b) $(\boldsymbol{v}+\boldsymbol{w})-\boldsymbol{z} = \boldsymbol{v}+(\boldsymbol{w}-\boldsymbol{z})$

(c) $(\boldsymbol{v}+\boldsymbol{w})-\boldsymbol{w} = \boldsymbol{v}$

PROOF:

(a) $\boldsymbol{v}-\boldsymbol{v} \underset{\substack{\uparrow\\ \text{Definition 2.7}}}{=} \boldsymbol{v}+(-\boldsymbol{v}) = \mathbf{0}$

(b) $(\boldsymbol{v}+\boldsymbol{w})-\boldsymbol{z} \underset{\uparrow}{=} (\boldsymbol{v}+\boldsymbol{w})+(-\boldsymbol{z}) = \boldsymbol{v}+[\boldsymbol{w}+(-\boldsymbol{z})] \underset{\uparrow}{=} \boldsymbol{v}+(\boldsymbol{w}-\boldsymbol{z})$

Definition 2.7

(c) $(\boldsymbol{v}+\boldsymbol{w})-\boldsymbol{w} \underset{\substack{\uparrow\\ \text{(b)}}}{=} \boldsymbol{v}+(\boldsymbol{w}-\boldsymbol{w}) \underset{\substack{\uparrow\\ \text{(a)}}}{=} \boldsymbol{v}+\mathbf{0} = \boldsymbol{v}$

CHECK YOUR UNDERSTANDING 2.12

(a) Show that for any two vectors $\boldsymbol{v}$ and $\boldsymbol{w}$:

$$-(\boldsymbol{v}+\boldsymbol{w}) = -\boldsymbol{v}-\boldsymbol{w}$$

(b) Use the Principle of Mathematical Induction (see Appendix A) to show that for any n vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$:

$$-(\boldsymbol{v}_1+\boldsymbol{v}_2+\cdots+\boldsymbol{v}_n) = -\boldsymbol{v}_1-\boldsymbol{v}_2-\cdots-\boldsymbol{v}_n$$

Answer: See page B-5.

We complete this section with a list of results; some of which we proved, some of which appeared in Check Your Understanding boxes, and others which you are invited to establish in the exercises.

THEOREM 2.11 For every $\boldsymbol{v}$, $\boldsymbol{w}$, and $\mathbf{z}$ in a vector space V, and every $r, s, t \in \Re$:

(i) There exists a unique vector $\mathbf{0} \in V$ $\boldsymbol{v} + \mathbf{0} = \mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$.

(ii) There exists a unique vector $-\boldsymbol{v} \in V$ such that $\boldsymbol{v} + (-\boldsymbol{v}) = -\boldsymbol{v} + \boldsymbol{v} = \mathbf{0}$.

Cancellation Properties

(iii) If $\boldsymbol{v} + \boldsymbol{z} = \boldsymbol{w} + \boldsymbol{z}$, then $\boldsymbol{v} = \boldsymbol{w}$

(iv) If $r \neq 0$ and $r\boldsymbol{v} = r\boldsymbol{w}$, then $\boldsymbol{v} = \boldsymbol{w}$.

(v) If $\boldsymbol{v} \neq \mathbf{0}$ and $r\boldsymbol{v} = s\boldsymbol{v}$, then $r = s$.

(vi) $0\boldsymbol{v} = \mathbf{0}$

(vii) $r\mathbf{0} = \mathbf{0}$

(viii) $r\boldsymbol{v} = \mathbf{0}$ if and only if $r = 0$ or $\boldsymbol{v} = \mathbf{0}$

(ix) $r(s\boldsymbol{v} + t\boldsymbol{w}) = (rs)\boldsymbol{v} + (rt)\boldsymbol{w}$

(x) $-1\boldsymbol{v} = -\boldsymbol{v}$

(xi) $-(-\boldsymbol{v}) = \boldsymbol{v}$

(xii) $r(-\boldsymbol{v}) = (-r)\boldsymbol{v} = -(r\boldsymbol{v})$

(xiii) $r(\boldsymbol{v} - \boldsymbol{w}) = r\boldsymbol{v} - (r\boldsymbol{w})$

(xiv) $(r - s)\boldsymbol{v} = r\boldsymbol{v} - (s\boldsymbol{v})$

(xv) $-(\boldsymbol{v} + \boldsymbol{w}) = -\boldsymbol{v} - \boldsymbol{w}$

(xvi) $\boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{w} + \boldsymbol{v}$

(xvii) $\boldsymbol{v} - (\boldsymbol{w} + \boldsymbol{z}) = (\boldsymbol{v} - \boldsymbol{w}) - \boldsymbol{z}$

(xviii) $\boldsymbol{v} - (\boldsymbol{w} - \boldsymbol{z}) = (\boldsymbol{v} - \boldsymbol{w}) + \boldsymbol{z}$

EXERCISES

Exercises 1-8. Prove:

1. Theorem 2.11 (iv): If $r \neq 0$ and $r\boldsymbol{v} = r\boldsymbol{w}$, then $\boldsymbol{v} = \boldsymbol{w}$.
2. Theorem 2.11 (v): If $\boldsymbol{v} \neq \mathbf{0}$ and $r\boldsymbol{v} = s\boldsymbol{v}$, then $r = s$.
3. Theorem 2.11 (ix): $r(s\boldsymbol{v} + t\boldsymbol{w}) = (rs)\boldsymbol{v} + (rt)\boldsymbol{w}$.
4. Theorem 2.11 (xiii): $r(\boldsymbol{v} - \boldsymbol{w}) = r\boldsymbol{v} - (r\boldsymbol{w})$.
5. Theorem 2.11 (xiv): $(r - s)\boldsymbol{v} = r\boldsymbol{v} - (s\boldsymbol{v})$.
6. Theorem 2.11 (xvi): $\boldsymbol{v} - \boldsymbol{w} = -\boldsymbol{w} + \boldsymbol{v}$.
7. Theorem 2.11 (xvii): $\boldsymbol{v} - (\boldsymbol{w} + \boldsymbol{z}) = (\boldsymbol{v} - \boldsymbol{w}) - \boldsymbol{z}$.
8. Theorem 2.11 (xviii): $\boldsymbol{v} - (\boldsymbol{w} - \boldsymbol{z}) = (\boldsymbol{v} - \boldsymbol{w}) + \boldsymbol{z}$.

9. Show that for any vector $\boldsymbol{v}$ in a vector space V, and any $r \in \Re$: $r\boldsymbol{v} = -(-r\boldsymbol{v})$.

10. Show that for any vector $\boldsymbol{v}$ in a vector space V and any integer $n \geq 1$: $n\boldsymbol{v} = (n-1)\boldsymbol{v} + \boldsymbol{v}$.

11. Let $\boldsymbol{v}$, $\boldsymbol{w}$, and $\boldsymbol{z}$ be any vectors in a vector space V, and let $a, b, c \in \Re$, with $a \neq 0$. Show that if $a\boldsymbol{v} + b\boldsymbol{w} = c\boldsymbol{z}$, then $\boldsymbol{v} = \frac{c}{a}\boldsymbol{z} - \frac{b}{a}\boldsymbol{w}$.

12. Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be vectors in a vector space V, with $\boldsymbol{v} \neq \mathbf{0}$. Show that if $r\boldsymbol{v} + \boldsymbol{w} = s\boldsymbol{v} + \boldsymbol{w}$, then $r = s$.

13. Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be vectors in a vector space V. Show that if $r \neq 1$ and $r\boldsymbol{v} + \boldsymbol{w} = \boldsymbol{v} + r\boldsymbol{w}$, then $\boldsymbol{v} = \boldsymbol{w}$.

14. Show that for any $\boldsymbol{v}$ and $\boldsymbol{w}$ in a vector space V, and for any $a, b \in \Re$:

$$(a + b)(\boldsymbol{v} + \boldsymbol{w}) = a\boldsymbol{v} + b\boldsymbol{v} + a\boldsymbol{w} + b\boldsymbol{w}$$

15. Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be non-zero vectors in a vector space V. Show that if $r\boldsymbol{v} + s\boldsymbol{w} = \mathbf{0}$, with not both r and s equal to 0, then there exist unique numbers a and b such that $\boldsymbol{v} = a\boldsymbol{w}$ and $\boldsymbol{w} = b\boldsymbol{v}$. Hint: Show that the condition that not both r and s equal 0 implies that **neither** is 0.

PROVE OR GIVE A COUNTEREXAMPLE

16. All vector spaces contain infinitely many vectors.

17. Any vector space that contains more than one vector must contain an infinite number of vectors.

18. For any vector $\boldsymbol{v}$ in a vector space V and any $r \in \Re$: $r\boldsymbol{v} = (r-1)\boldsymbol{v} + \boldsymbol{v}$

19. Let V and W be vector spaces. Let $V \times W = \{(\boldsymbol{v}, \boldsymbol{w}) | \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$ with operations given by:
$$(v_1, w_1) + (v_2, w_2) = (v_1 + v_2, w_1 + w_2) \text{ and } r(v, w) = (rv, rw)$$
Then $V \times W$ is a vector space.

20. Let V and W be vector spaces. Let $V \times W = \{(\boldsymbol{v}, \boldsymbol{w}) | \boldsymbol{v} \in V, \boldsymbol{w} \in W\}$ with operations given by:
$$(v_1, w_1) + (v_2, w_2) = (v_1 - v_2, w_1 - w_2) \text{ and } r(v, w) = (rv, rw)$$
Then $V \times W$ is a vector space.

§4. SUBSPACES

DEFINITION 2.8
SUBSPACE

A **subspace** of a vector space V is a nonempty subset S of V which, together with the imposed operations of addition and scalar multiplication of V, is itself a vector space.

If S is to be a subspace of V, then it is itself a vector space and must therefore be closed under addition and scalar multiplication:

If $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$ are in S, then $\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$.

If $\boldsymbol{s} \in S$ and $r \in \Re$, then $r\boldsymbol{s} \in S$.

In addition, the eight axioms listed in Definition 2.6, page 40, must also hold for S. Actually, the eight axioms come "free of charge," once closure is established; for:

THEOREM 2.12 A subset S of a vector space V is a subspace of V **if and only if**:

1. S is nonempty.
2. S is closed under addition.
3. S is closed under scalar multiplication.

PROOF: If S is a subspace of the vector space V, then it is itself a vector space and must therefore satisfy the three stated conditions. We now show that if those three conditions hold, then S is a subspace of V.

Of the eight axioms of Definition 2.6, we need not worry about Axioms (i), (ii), (v), (vi), (vii), and (viii):

(i) $\boldsymbol{v} + \boldsymbol{u} = \boldsymbol{u} + \boldsymbol{v}$	(ii) $(\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w} = \boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w})$
(v) $r(\boldsymbol{u} + \boldsymbol{v}) = r\boldsymbol{u} + r\boldsymbol{v}$	(vi) $(r + s)\boldsymbol{u} = r\boldsymbol{u} + s\boldsymbol{u}$
(vii) $r(s\boldsymbol{u}) = (rs)\boldsymbol{u}$	(viii) $1\boldsymbol{u} = \boldsymbol{u}$

Why not? Because, since they hold for all $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ in the given vector space V, then they will surely hold for all $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$ in the subset S of V.

Why do we have to worry about the zero axiom? Because though we know that there is a vector $\boldsymbol{0}$ in V such that $\boldsymbol{v} + \boldsymbol{0} = \boldsymbol{v}$ for every $\boldsymbol{v} \in V$ (and therefore for every $\boldsymbol{s} \in S$), we have no assurance that $\boldsymbol{0}$ sits in S. To see that it does, take any vector $\boldsymbol{v}$ in S (we are given that S is nonempty), and then scalar multiply that vector by $\boldsymbol{0}$:

$$\overset{\text{Since } \boldsymbol{v} \in S \text{ and } S \text{ is closed under scalar mutliplication}}{0\boldsymbol{v} \underset{\substack{\uparrow \\ \text{Theorem 2.7, page 53}}}{=} \mathbf{0} \in S}$$

We now complete the proof by showing that if $\boldsymbol{v} \in S$, then its additive inverse $-\boldsymbol{v}$ is also in S:

$$\overset{\text{Since } \boldsymbol{v} \in S \text{ and } S \text{ is closed under scalar mutliplication}}{-1\boldsymbol{v} \underset{\substack{\uparrow \\ \text{Theorem 2.9, page 54}}}{=} -\boldsymbol{v} \in S}$$

When challenging the "S is nonempty" condition of Theorem 2.12, one typically looks to see if the zero vector is contained in S. For:

If $\mathbf{0} \in S$, then S is certainly nonempty.

If $\mathbf{0} \notin S$, then S is not a subspace, period.

EXAMPLE 2.8 Verify that

$$S = \{(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) | c = a + b\}$$

is a subspace of $\Re^3$.

SOLUTION: We show that S satisfies the three conditions of Theorem 2.12.

1. Since the sum of the first two components of $\mathbf{0} = (\mathbf{0}, \mathbf{0}, \mathbf{0})$ is equal to its third component, $\mathbf{0} \in S$ (and therefore S is not empty).

2. To show that S is closed under addition, we take two arbitrary elements of S:

The "ticket" to be in S is that the third component is equal to the sum of its first two components.

$$\boldsymbol{s}_1 = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} + \boldsymbol{b}), s_2 = (\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{c} + \boldsymbol{d})$$

and consider their sum:

$$\boldsymbol{s}_1 + \boldsymbol{s}_2 = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} + \boldsymbol{b}) + (\boldsymbol{c}, \boldsymbol{d}, \boldsymbol{c} + \boldsymbol{d}) = (\boldsymbol{a} + \boldsymbol{c}, \boldsymbol{b} + \boldsymbol{d}, \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} + \boldsymbol{d})$$

Since $s_1 + s_2$ has the "ticket," it is in S.

Since the third component of $\boldsymbol{s}_1 + \boldsymbol{s}_2$, $\boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c} + \boldsymbol{d}$, equals the sum of its first two components, $\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$.

3. We now show S is closed under scalar multiplication.

For $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} + \boldsymbol{b}) \in S$ and $r \in \Re$:

$$r(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} + \boldsymbol{b}) = (\boldsymbol{ra}, \boldsymbol{rb}, \boldsymbol{r}(\boldsymbol{a} + \boldsymbol{b})) = (\boldsymbol{ra}, \boldsymbol{rb}, \boldsymbol{ra} + \boldsymbol{rb}) \underset{\substack{\uparrow \\ \text{third component is sum of first two}}}{\in} S$$

CHECK YOUR UNDERSTANDING 2.13

Show that:

$$S = \left\{ \begin{bmatrix} a & 2a \\ -a & 0 \end{bmatrix} \middle| a \in \Re \right\}$$

is a subspace of the vector space $M_{2 \times 2}$ of Example 2.3, page 41.

Answer: See page B-5.

The following theorem merges two of the properties of Theorem 2.12 into one:

THEOREM 2.13 A nonempty subset S of a vector space V is a subspace of V **if and only if**:
For every $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S$ and $r \in \Re$, $r\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$.

PROOF: For S a nonempty subspace of V, let $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S$ and $r \in \Re$. Since S is closed under scalar multiplication: $r\boldsymbol{s}_1 \in S$. Since S is closed under addition: $r\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$.

Conversely, assume that for every $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S$ and $r \in \Re$:

$$r\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S \ (*)$$

We show that S is closed under addition and scalar multiplication:

For any given $s_1, s_2 \in S$, simply choose $r = 1$, and apply (*):

$$1\boldsymbol{s}_1 + \boldsymbol{s}_2 = \boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$$

To show that S is also closed under scalar multiplication, we first observe that (*) implies that $\mathbf{0} \in S$:

Since S is nonempty, we can choose an $\boldsymbol{s} \in S$. Letting $\boldsymbol{s}_1 = \boldsymbol{s}_2 = \boldsymbol{s}$, and $r = -1$ in (*) brings us to:

$$-1\boldsymbol{s} + \boldsymbol{s} \in S$$
$$\mathbf{0} \in S$$

Now consider any $\boldsymbol{s} \in S$ and $r \in \Re$. Appealing to (*) with $\boldsymbol{s}_1 = \boldsymbol{s}$ and $\boldsymbol{s}_2 = \mathbf{0} \in S$, we find that:

$$r\boldsymbol{s} + \mathbf{0} = r\boldsymbol{s} \in S$$

EXAMPLE 2.9 Let $\boldsymbol{u}$ and $\boldsymbol{v}$ be any two vectors in a vector space V. Show that the set:

$$S = \{a\boldsymbol{u} + b\boldsymbol{v} | a, b \in \Re\}$$

is a subspace of V.

SOLUTION: Since $\mathbf{0} = 0\boldsymbol{u} + 0\boldsymbol{v} \in S$, $\mathbf{0} \in S$: S is not empty.

$$\begin{gathered} s_1, s_2 \in S \text{ and } r \in R \\ \Big\downarrow \\ rs_1 + s_2 \in S \end{gathered}$$

For $a\boldsymbol{u} + b\boldsymbol{v} \in S$ and $c\boldsymbol{u} + d\boldsymbol{v} \in S$, and for $r \in \Re$:

$$\begin{aligned} r(a\boldsymbol{u} + b\boldsymbol{v}) + (c\boldsymbol{u} + d\boldsymbol{v}) &= (ra\boldsymbol{u} + c\boldsymbol{u}) + (rb\boldsymbol{v} + d\boldsymbol{v}) \\ &= (ra + c)\boldsymbol{u} + (rb + d)\boldsymbol{v} \in S \end{aligned}$$

It is of the form $A\boldsymbol{u} + B\boldsymbol{v}$
(has the “ticket”)

CHECK YOUR UNDERSTANDING 2.14

Show that $S = \{(x, y, z) | x + y + z = 0\}$ is a subspace of $\Re^3$.

Answer: See Page B-5.

EXAMPLE 2.10

Let $F(\Re)$ denote the function space $F(X)$ of Theorem 2.4, page 44, with domain $X = \Re$. Show that: $S = \{\boldsymbol{f} \in F(\Re) | f(9) = 0\}$ is a subspace of $F(\Re)$.

The “ticket” needed for a function $\boldsymbol{f}$ to get into S is that it maps 9 to 0.

Solution: As you recall, the zero in $F(\Re)$ turned out to be the function $\boldsymbol{Z}$ given by: $Z(x) = 0$ for every number x. In particular, since $Z(9) = 0$, $\boldsymbol{Z} \in S$. Hence, $S \neq \varnothing$.

If $\boldsymbol{f}, \boldsymbol{g} \in S$ and $r \in \Re$, then:

$$(r\boldsymbol{f} + \boldsymbol{g})(9) = rf(9) + g(9) \overset{\text{since } \boldsymbol{f} \text{ and } \boldsymbol{g} \text{ are in } S}{=} r0 + 0 = 0$$

$r\boldsymbol{f} + \boldsymbol{g} \in S$
(has the “ticket”)

$$\begin{gathered} \boldsymbol{f}, \boldsymbol{g} \in S \text{ and } r \in R \\ \Big\downarrow \\ r\boldsymbol{f} + \boldsymbol{g} \in S \end{gathered}$$

CHECK YOUR UNDERSTANDING 2.15

Let $F(\Re)$ denote the function space of Theorem 2.4, page 44 (with domain $X = \Re$). Determine whether or not

$$S = \{\boldsymbol{f} \in F(\Re) | f(0) = 9\}$$

is a subspace of $F(R)$.

Answer: Not a subspace.

EXAMPLE 2.11

Show that for any $a, b \in \Re$:

$$S = \{(x, y) | ax + by = c\}$$

is a subspace of $\Re^2$ **if and only if** $c = 0$.

Solution: We have to establish two results:

The “**if** -condition:” If $c = 0$, then S is a subspace of $\Re^2$.

The “**only if**-condition:” If $c \neq 0$, then S is not a subspace of $\Re^2$.

The "only if-condition" is easily dispensed with:

If $c \neq 0$, then $\mathbf{0} = (\mathbf{0}, \mathbf{0}) \notin S$, for: $a \cdot 0 + b \cdot 0 = 0 \neq c$. Since S does not contain **0**, S is not a subspace of $\Re^2$.

To establish the "**if**-condition," we assume the $c = 0$, and first observe that S is not empty:

$$\text{Since } a \cdot 0 + b \cdot 0 = 0 = c,\ \mathbf{0} = (\mathbf{0}, \mathbf{0}) \in S.$$

We complete the proof by showing that if $(\boldsymbol{x_1}, \boldsymbol{y_1}), (\boldsymbol{x_2}, \boldsymbol{y_2}) \in S$ and $r \in \Re$, then:

$$r(\boldsymbol{x_1}, \boldsymbol{y_1}) + (\boldsymbol{x_2}, \boldsymbol{y_2}) = (\boldsymbol{rx_1} + \boldsymbol{x_2}, \boldsymbol{ry_1} + \boldsymbol{y_2}) \in S$$

which is to say, that:

$$\boldsymbol{a(rx_1 + x_2) + b(ry_1 + y_2)} = \mathbf{0}$$

Here goes:

$$\begin{aligned} \boldsymbol{a(rx_1 + x_2) + b(ry_1 + y_2)} &= arx_1 + ax_2 + bry_1 + by_2 \\ &= r(ax_1 + by_1) + (ax_2 + by_2) \\ \text{Since } (x_1, y_1), (x_2, y_2) \in S: \quad &= r0 + 0 = \mathbf{0} \end{aligned}$$

CHECK YOUR UNDERSTANDING 2.16

In Example 1.10, page 20, we showed that:

$$S = \{(16a - 2b, 4a - 17b, 11a, 11b) | a, b \in \Re\}$$

is the solution set of the homogeneous system of equations:

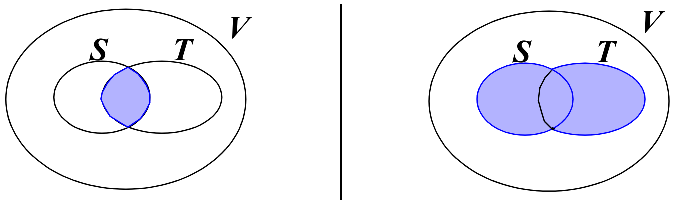
$$\left.\begin{aligned} 2x + 3y - 4z + 5w &= 0 \\ -3x + y + 4z + w &= 0 \\ x + 7y - 4z + 11w &= 0 \end{aligned}\right\}$$

Show that S is a subspace of $\Re^4$.

Answer: See page B-5.

INTERSECTION AND UNION OF SUBSPACES

Let S and T be subspaces of a vector space V. Is their intersection $S \cap T = (\boldsymbol{v} | \boldsymbol{v} \in S \text{ and } \boldsymbol{v} \in T)$ [Figure 2.7(a)] necessarily a subspace of V? Is their union $S \cup T = (\boldsymbol{v} | \boldsymbol{v} \in S \text{ or } \boldsymbol{v} \in T)$ [Figure 2.7(b)] necessarily a subspace of V? The answer to the first question is "Yes" (Theorem 2.14 below), while the answer to the second question is "No" (Example 2.12 below).



$S \cap T = (\boldsymbol{v} | \boldsymbol{v} \in S \text{ and } \boldsymbol{v} \in T)$

S intersect ***T***

(a)

$S \cup T = (\boldsymbol{v} | \boldsymbol{v} \in S \text{ or } \boldsymbol{v} \in T)$

S union ***T***

(b)

Figure 2.7

In the exercises you are asked to show that the intersection of any number of subspaces of V is again a subspace of V.

THEOREM 2.14 If S and T are subspaces of a space V, then so is their intersection:

$$S \cap T = \{\boldsymbol{v} | \boldsymbol{v} \in S \text{ and } \boldsymbol{v} \in T\}$$

PROOF: $S \cap T$ is not empty:

Since $\mathbf{0} \in S$ and $\mathbf{0} \in T$ (why?), $\mathbf{0} \in S \cap T$.

$\boldsymbol{u}, \boldsymbol{v} \in S \cap T$ and $r \in R$

$\Downarrow$

$r\boldsymbol{u} + \boldsymbol{v} \in S \cap T$

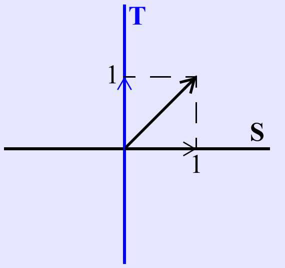
Let $\boldsymbol{u}, \boldsymbol{v} \in S \cap T$ and $r \in R$. Being in the intersection of S and T, $\boldsymbol{u}$ and $\boldsymbol{v}$ are both in S and in T. Since S and T are subspaces, $r\boldsymbol{u} + \boldsymbol{v} \in S$ and $r\boldsymbol{u} + \boldsymbol{v} \in T$. Being in both S and T, $r\boldsymbol{u} + \boldsymbol{v} \in S \cap T$.

EXAMPLE 2.12 Show that the union of two subspaces S and T of a vector space V:

$$S \cup T = (\boldsymbol{v} | \boldsymbol{v} \in S \text{ or } \boldsymbol{v} \in T)$$

need not be a subspace of V.

SOLUTION: While one can easily show that the set $S \cup T$ is nonempty, and that it is closed under scalar multiplication, one cannot show that it is closed under addition, for it need not be! And how do we show that this is the case? By exhibiting a **specific** vector space V, along with two **specific** subspaces S and T, such that their union fails to be closed under addition. Let's do it:



Let $V = \Re^2$, $S = \{(\boldsymbol{x}, \mathbf{0}) | x \in \Re\}$, and $T = \{(\mathbf{0}, \boldsymbol{y}) | y \in \Re\}$. We leave it for you to verify that both S and T are subspaces of $\Re^2$. To see that $S \cup T$ is not closed under addition, simply note that while $(\mathbf{1}, \mathbf{0}) \in S \cup T$ and $(\mathbf{0}, \mathbf{1}) \in S \cup T$, $(\mathbf{1}, \mathbf{0}) + (\mathbf{0}, \mathbf{1}) = (\mathbf{1}, \mathbf{1}) \notin S \cup T$ (for $(1, 1)$ is neither in S nor in T).

CHECK YOUR UNDERSTANDING 2.17

PROVE OR GIVE A COUNTEREXAMPLE:

If S and T are subsets of a vector space V, and if $S \cap T$ is a subspace of V, then either S is a subspace of V or T is a subspace of V.

Answer: See page B-6.

EXERCISES

Exercises 1-6. Determine if the given subset S of the vector space $\Re^2$ is a subspace of $\Re^2$. Justify your answer.

1. $S = \{(x, y) | y = 2x\}$
2. $S = \{(x, y) | y = 2x + 1\}$
3. $S = \{(x, y) | x + y = 0\}$
4. $S = \{(x, y) | x \le y\}$
5. $S = \{(x, y) | y = x^2\}$
6. $S = \{(x, y) | xy = 0\}$

Exercises 7-12. Determine if the given subset S of the vector space $\Re^3$ is a subspace of $\Re^3$. Justify your answer.

7. $S = \{(x, y, z) | z = 2x - y\}$
8. $S = \{(x, y, z) | z = 2x + y\}$
9. $S = \{(x, y, z) | z = 2x - y + 1\}$
10. $S = \{(x, y, z) | z = 2x + y + 1\}$
11. $S = \{(x, y, z) | |x| + |y| + z = 0\}$
12. $S = \{(x, y, z) | y = 0\}$

Exercises 13-18. Determine if the given subset S of the matrix space $M_{2 \times 2}$ of Example 2.3, page 41 is a subspace of $M_{2 \times 2}$. Justify your answer.

13. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| d = 0 \right\}$
14. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| d = a + b \right\}$
15. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a = d = 0, c = 2b \right\}$
16. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a + b = 2c - 3d \right\}$
17. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a + b + c + d = 1 \right\}$
18. $S = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \middle| a + b + c + d = 0 \right\}$

Exercises 19-22. Determine if the given subset S of the polynomial space $P_2(x)$ of Theorem 2.3, page 43, is a subspace of $P_2(x)$. Justify your answer.

19. $S = \{ax^2 + bx + c | a = 2\}$
20. $S = \{ax^2 + bx + c | b = 0\}$
21. $S = \{ax^2 + bx + c | a + b + c = 0\}$
22. $S = \{ax^2 + bx + c | b = 2a\}$

Exercises 23-26. Determine if the given subset S of the polynomial space $P_3(x)$ of Theorem 2.3, page 43, is a subspace of $P_3(x)$. Justify your answer.

23. $S = \{ax^2 + bx + c \mid b = 0\}$
24. $S = \{ax^2 + bx + c \mid b = -a + 2c\}$
25. $S = \{ax^2 + bx + c \mid a + b = 0\}$
26. $S = \{ax^2 + bx + c \mid b = 2a, c = a + 1\}$

Exercises 27-38. Determine if the given subset S of the function space $F(\Re)$ of Theorem 2.4 (with $X = \Re$), page 44, is a subspace of $F(\Re)$. Justify your answer.

27. $S = \{f \mid f(0) = 0\}$
28. $S = \{f \mid f(1) = 0\}$
29. $S = \{f \mid f(1) = 1\}$
30. $S = \{f \mid f(2x) = 2f(x)\}$
31. The subset of even functions: $S = \{f \mid f(-x) = f(x)\}$
32. The subset of odd functions: $S = \{f \mid f(-x) = -f(x)\}$
33. The subset of increasing functions: $S = \{f \mid \mathrm{a} < b \Rightarrow f(a) < f(b)\}$
34. The subset of decreasing functions: $S = \{f \mid \mathrm{a} < b \Rightarrow f(a) > f(b)\}$
35. The subset of bounded functions: $S = \{f \mid |f(x)| < M$ for every $x \in \Re$, for some $M \geq 0\}$
36. **(Calculus dependent)** $S = \{f \mid f$ is continuous$\}$
37. **(Calculus dependent)** $S = \{f \mid f$ is differentiable$\}$
38. **(Calculus dependent)** $S = \{f \mid f$ is integrable$\}$
39. Let V be a vector space. Show that:
    (a) $\{\mathbf{0}\}$ is a subspace of V.
    (b) V is a subspace of V.

40. (PMI) Establish the following generalization of Theorem 2.14.

(a) If $S_1, S_2, \ldots, S_n$ are subspaces of a vector space V, then so is their intersection:

$$\bigcap_{i=1}^{n} S_i = S_1 \cap S_2 \cap \ldots \cap S_n$$

(b) If $S_1, S_2, \ldots, S_n, \ldots$ are subspaces of a vector space V, then so is their intersection:

$$\bigcap_{i=1}^{\infty} S_i = S_1 \cap S_2 \cap \ldots \cap S_n \cap \ldots$$

(c) Let A be a nonempty set. If S_α is a subspace of a vector space V for every $\alpha \in A$, then the set $\bigcap_{\alpha \in A} S_\alpha$ is also a subspace of V.

41. (PMI) Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$ be vectors in a vector space V. Show that $S = \{a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \dots + a_n\boldsymbol{v}_n | a_i \in \Re, 1 \le i \le n\}$ is a subspace of V.

42. Let S and T be subspaces of a vector space V. Show that $S + T = \{\boldsymbol{s} + \boldsymbol{t} | \boldsymbol{s} \in S \text{ and } \boldsymbol{t} \in T\}$ is a subspace of V.

43. Let S and T be subspaces of a vector space V, with $S \cap T = \{\boldsymbol{0}\}$. Show that every vector in the subspace $S + T$ of the previous exercise can be **uniquely** expressed as a sum of a vector in S with a vector in T.

    Suggestion: Show that if $s + t = s_1 + t_1$, then $s = s_1$ and $t = t_1$.

PROVE OR GIVE A COUNTEREXAMPLE

44. If S and T are both subsets of a vector space V, and if neither S nor W is a subspace of V, then $S \cap T$ cannot be a subspace of V.

45. If S and T are both subsets of a vector space V, and if neither S nor W is a subspace of V, then $S \cup T$ cannot be a subspace of V.

46. If S and T are subspaces of a vector space V, then $S + T = S \cup T$ (see Exercise 43).

47. If S and T are subspaces of a vector space V, then $S \cup T \subseteq S + T$ (see Exercise 43).

48. If S, T, and W are subspaces of a vector space V, then $(S \cap T) + (T \cap W)$ is also a subspace of V (see Exercise 43).

49. If S, T, and W are subspaces of a vector space V, then $S \cap (T + W) = (S \cap T) + W$ (see Exercise 42).

50. If S and T are subspaces of a vector space V with $S \cap T = \{\boldsymbol{0}\}$, then $S \cup T$ is a subspace of V.

51. If S is a subspace of a vector space V, and if T is a subspace of S, then T is a subspace of V.

52. If a vector space has two distinct subspaces, then it has infinitely many distinct subspaces.

§5. LINES AND PLANES

This chapter began with a consideration of vectors in the plane and in three dimensional space, both from a geometrical point of view (directed line segments, or “arrows”), and from an analytical perspective (2-tuples and 3-tuples). The main focus of this section is to determine and classify all of the subspaces of those Euclidean spaces. Additional insight for the material of this section will surface in the following chapter on Dimension Theory. This section, in turn, is a nice lead-in to the following chapter, for Euclidean spaces have a dimension component built right into their terminology. It should come as no surprise, for example, to find that the Euclidean spaces $\Re^2$ and $\Re^3$ will turn out to have dimensions 2 and 3, respectively.

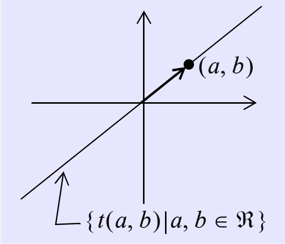
SUBSPACES OF $\Re^2$

A subset of a vector space V that is neither V or $\{\mathbf{0}\}$ is said to be a **proper subset** of V. The following theorem serves to characterize the proper subspaces of $\Re^2$:

THEOREM 2.15 S is a **proper** subspace of $\Re^2$ if and only if
$$S = \{t(\boldsymbol{a}, \boldsymbol{b}) | t \in \Re\} \text{ for } (\boldsymbol{a}, \boldsymbol{b}) \neq (\mathbf{0}, \mathbf{0})$$

PROOF: Let $S = \{t(\boldsymbol{a}, \boldsymbol{b}) | t \in \Re\}$ for $(\boldsymbol{a}, \boldsymbol{b}) \neq (\mathbf{0}, \mathbf{0})$. Appealing to Theorem 2.13, page 61, we first observe that S is not empty [it clearly contains $(\boldsymbol{a}, \boldsymbol{b})$]. Moreover, for $t_1(\boldsymbol{a}, \boldsymbol{b}), t_2(\boldsymbol{a}, \boldsymbol{b}) \in S$ and $r \in \Re$:

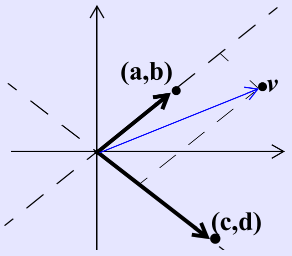
$$r[t_1(\boldsymbol{a}, \boldsymbol{b})] + t_2(\boldsymbol{a}, \boldsymbol{b}) = (rt_1 + t_2)(\boldsymbol{a}, \boldsymbol{b}) \in S$$



At this point we know that S is a non-empty subspace of $\Re^2$. Can it be all of $\Re^2$? No, for the set $\{(ta, tb) | (t \in \Re)\}$ is the line in the plane passing through the origin and the point (a, b) (see margin).

Conversely, assume that T is a **proper** subspace of $\Re^2$. Since it is not empty, it contains a nonzero vector $(\boldsymbol{a}, \boldsymbol{b})$. Since T is a subspace, every scalar multiple of $(\boldsymbol{a}, \boldsymbol{b})$ must be in T, which is to say: $S = \{t(\boldsymbol{a}, \boldsymbol{b}) | t \in \Re\} \subseteq T$. We now show that, in fact, $S = T$, by demonstrating that if T were to contain any vector $(\boldsymbol{c}, \boldsymbol{d}) \notin S$ then T must be all of $\Re^2$ (see margin)”:

For any vector $v \in R^2$:



Can we find scalars A, B, such that:

$$A(\boldsymbol{a}, \boldsymbol{b}) + B(\boldsymbol{c}, \boldsymbol{d}) = (\boldsymbol{x}, \boldsymbol{y}) \Rightarrow \begin{matrix} Aa + Bc = x \\ Ab + Bd = y \end{matrix} ?$$

Yes, providing the last row of rref$\begin{bmatrix} a & c \\ b & d \end{bmatrix}$ does not consist entirely of zeros (see the Spanning Theorem, page 18) — and it doesn’t (Exercise 61).

While no line in the plane that does not pass through the origin can represent a subspace of $\Re^2$ (why not?), every line in the plane is parallel to one that does pass through the origin—a translation of a subspace of $\Re^2$:

The vector $\boldsymbol{v}$ is said to be a **direction vector** for the line, and the vector $\boldsymbol{u}$ is said to be a **translation vector**.

THEOREM 2.16 Let L be the line passing through two distinct points $P = (x_1, y_1)$ and $Q = (x_2, y_2)$ in the plane [Figure 2.9(a)]. Then, **in terms of vectors**:

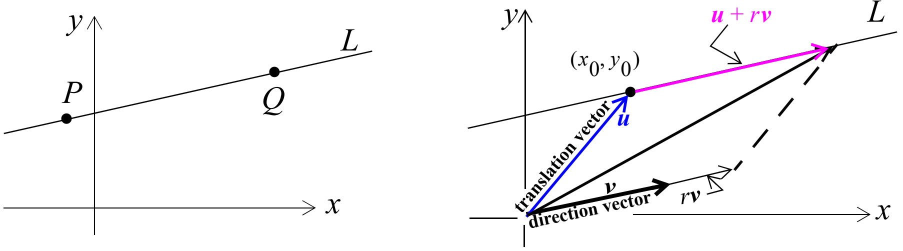
$$\boldsymbol{L} = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\} \qquad (*)$$

where $\boldsymbol{v} = \overrightarrow{PQ} = (\boldsymbol{x_2 - x_1, y_2 - y_1})$, and $\boldsymbol{u} = (\boldsymbol{x_0, y_0})$, for (x_0, y_0) any chosen point on L [see Figure 2.8(b)].

[(*) is said to be a **vector form** representation for the line L.]

PROOF: Figure 2.8(b) may serve as a geometrical "proof" of the theorem; for if you take the vector $\boldsymbol{v}$ which is in the same direction as the line L, and stretch it every which way, then you will get a line that is parallel to L passing through the origin. Adding the vector $\boldsymbol{u}$ to every $r\boldsymbol{v}$ "moves" that line up to L, in a parallel fashion.

Those not satisfied with this geometrical proof are invited to consider Exercise 62.



(a) (b)

Figure 2.8

EXAMPLE 2.13 Find a vector form representation for the line L passing through the points $(1, 5)$ and $(2, 3)$

Note that the set:

$L = \{(1, 5) + r(1, -2) \mid r \in \Re\}$

$= \{(1 + r, 5 - 2r) \mid r \in \Re\}$

This brings us to the so-called **parametric** representation of L:

$x = 1 + r, y = 5 - 2r$

SOLUTION: We take $\boldsymbol{v} = (\boldsymbol{2 - 1, 3 - 5}) = (\boldsymbol{1, -2})$ to be our direction vector, and $\boldsymbol{u} = (\boldsymbol{1, 5})$ as our translation vector, leading us to the vector form: $L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\} = \{(\boldsymbol{1, 5}) + r(\boldsymbol{1, -2}) \mid r \in \Re\}$.

CHECK YOUR UNDERSTANDING 2.18

(a) Referring to Example 2.13, find the vector form representation for the line L, when $\boldsymbol{v}$ is the vector from $(2, 3)$ to $(1, 5)$, and $\boldsymbol{u} = (2, 3)$.

(b) Your vector form representation in (a) will look different from that of Example 2.12: $L = \{(1, 5) + r(1, -2) \mid r \in \Re\}$. Appearances aside, show that your set of two-tuples in (a) and ours of Example 2.13 are one and the same.

Answer: See page B6.

SUBSPACES OF $\Re^3$

We now move up a notch and focus our attention on the Euclidean space $\Re^3$. The first order of business is to arrive at vector form representations for lines and planes in $\Re^3$. As it was in $\Re^2$, a line in three-space is determined by any two points, P and Q, in $\Re^3$. In the exercises, you are invited to establish the following result (compare with Theorem 2.16):

THEOREM 2.17 Let L be the line passing through two distinct points $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$ in three-space. Then, in **vector form**:

$$L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\}$$

where $\boldsymbol{v}$ is the **direction** vector:

$$\boldsymbol{v} = \overrightarrow{PQ} = (\boldsymbol{x_2 - x_1, y_2 - y_1, z_2 - z_1})$$

and $\boldsymbol{u} = (\boldsymbol{x_0, y_0, z_0})$ is a **translation vector**, with (x_0, y_0, z_0) any chosen point on L.

One cannot envision a line in $\Re^n$ for $n > 3$. We can, however define, in vector form, the line passing through:

$$P = (a_1, a_2, \ldots, a_n)$$

and $Q = (b_1, b_2, \ldots, b_n)$

in $\Re^n$ to be the set:

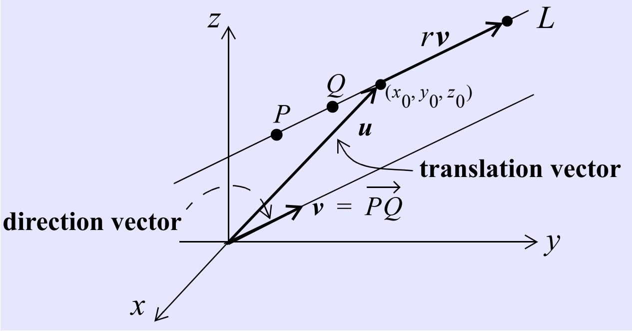
$$L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\}$$

where:

$$\boldsymbol{v} = (\boldsymbol{b_1 - a_1, \ldots, b_n - a_n})$$

and:

$$\boldsymbol{u} = (\boldsymbol{a_1, a_2, \ldots, a_n})$$



EXAMPLE 2.14 Find a vector form representation for the line L passing through the points $(2, 0, -3)$ and $(1, 4, 2)$.

SOLUTION: We take the vector from $(2, 0, -3)$ to $(1, 4, 2)$ $\boldsymbol{v} = (\mathbf{1 - 2, 4 - 0, 2 - (-3)}) = (\mathbf{-1, 4, 5})$ as our direction vector. Selecting the translation vector $\boldsymbol{u} = (\mathbf{2, 0, -3})$, we have:

$$L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\} = \{(2, 0, -3) + r(-1, 4, 5) \mid r \in \Re\}$$

The line can also be expressed in parametric form (see margin note of Example 2.13): $x = 2 - r, y = 4r, z = 3 + 5r$

CHECK YOUR UNDERSTANDING 2.19

Consider the line $L = \{(1, 3, 5) + r(2, 1, -1) \mid r \in \Re\}$. Find two points on L, both different from $(1, 3, 5)$, and proceed as in Example 2.14 to obtain another representation for the set L. Verify that "your set" is equal to the set $\{(1, 3, 5) + r(2, 1, -1) \mid r \in \Re\}$.

Answer: See page B6.

In Theorem 2.16, we noted that the line L in $\Re^2$ which passes through the origin and the point (a, b) (distinct from the origin) can be expressed in vector form: $L = \{r\boldsymbol{v} | r \in \Re\}$, where $\boldsymbol{v} = (\boldsymbol{a}, \boldsymbol{b})$. A similar result, which you are invited to establish in the exercises, holds for a plane in $\Re^3$:

THEOREM 2.18 Let P be a plane in $\Re^3$ passing through the origin [Figure 2.9(a)]. Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be any two points on P which do not both lie on a common line passing through the origin. Then, in **vector form**:

$$P = \{r\boldsymbol{u} + s\boldsymbol{v} | r, s \in \Re\}$$

where $\boldsymbol{u} = (\boldsymbol{x_1}, \boldsymbol{y_1}, \boldsymbol{z_1})$ and $\boldsymbol{v} = (\boldsymbol{x_2}, \boldsymbol{y_2}, \boldsymbol{z_2})$.

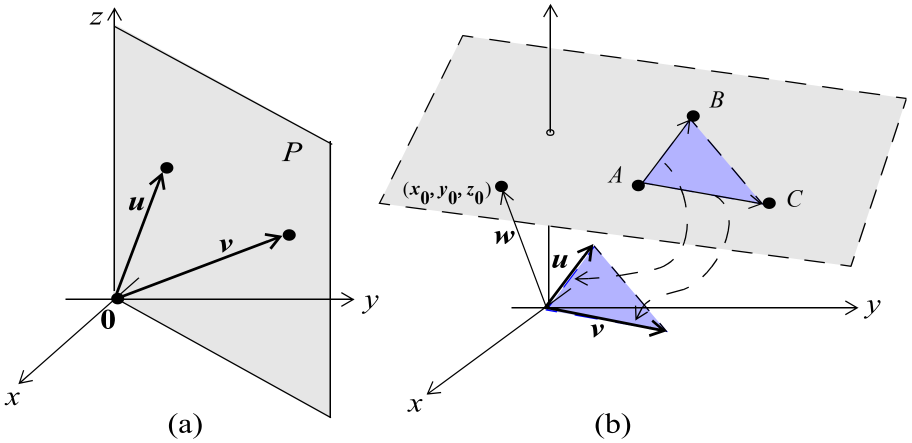


Figure 2.9

In this general setting, we have (compare with Theorem 2.17):

THEOREM 2.19 Let P be the plane passing through three non-colinear (not lying on a common line) points $A = (x_1, y_1, z_1)$, $B = (x_2, y_2, z_2)$ and $C = (x_3, y_3, z_3)$ [Figure 2.9(b)]. Then, in **vector form**:

$$P = \{\boldsymbol{w} + r\boldsymbol{u} + s\boldsymbol{v} | r \in R, s \in \Re\}$$

where $\boldsymbol{u} = \overrightarrow{AB} = (\boldsymbol{x_2} - \boldsymbol{x_1}, \boldsymbol{y_2} - \boldsymbol{y_1}, \boldsymbol{z_2} - \boldsymbol{z_1})$

$$\boldsymbol{v} = \overrightarrow{AC} = (\boldsymbol{x_3} - \boldsymbol{x_1}, \boldsymbol{y_3} - \boldsymbol{y_1}, \boldsymbol{z_3} - \boldsymbol{z_1}),$$

and $\boldsymbol{w} = (\boldsymbol{x_0}, \boldsymbol{y_0}, \boldsymbol{z_0})$ for (x_0, y_0, z_0) any chosen point on P.

u and ***v*** are said to be **direction vectors**, *and* ***w*** *is said to be a* ***translation vector****.*

EXAMPLE 2.15 Find a vector form representation for the Plane P passing through the points $(3, -2, 2)$, $(2, 5, -3)$ and $(4, 1, -3)$.

SOLUTION: We elect $(3, -2, 2)$ to play the role of both ***w*** and of (x_1, y_1, z_1) in Theorem 2.19; with:

$$\boldsymbol{u} = (\mathbf{2, 5, -3}) - (\mathbf{3, -2, 2}) = (\mathbf{-1, 7, -5})$$

and

$$\boldsymbol{v} = (\mathbf{4, 1, -3}) - (\mathbf{3, -2, 2}) = (\mathbf{1, 3, -5})$$

Then:

$$\begin{aligned} P &= \{\boldsymbol{w} + r\boldsymbol{u} + s\boldsymbol{v} \mid r, s \in \Re\} \\ &= \{(3, -2, 2) + r(-1, 7, -5) + s(1, 3, -5) \mid r, s \in \Re\} \\ &= \{(3 - r + s, -2 + 7r + 3s, 2 - 5r - 5s) \mid r, s \in \Re\} \end{aligned}$$

P consists of all points (x, y, z) such that:

$$x = 3 - r + s$$
$$y = -2 + 7r + 3s$$
$$z = 2 - 5r - 5s$$

The above is said to be a parametric representation of the plane (with parameters r and s).

CHECK YOUR UNDERSTANDING 2.20

(a) Repeat our solution to Example 2.15, but this time letting $(4, 1, -3)$ play the role of ***w***, instead of $(3, -2, 2)$; all of the rest remaining as before.

(b) Your set representation for P in (a) will look different from that of Example 2.15. Appearances aside, show that your set in (a) and ours of Example 2.15 are one and the same.

Answer: See page B6.

We previously showed that the only proper subspaces of $\Re^2$ are the lines passing through the origin (as sets). The following theorem, a proof of which is relegated to the exercises, settles the subspace issue in $\Re^3$:

THEOREM 2.20 S is a proper subspace of $\Re^3$ **if and only if** S is the set of points on a line that passes through the origin, or the set of points on a plane that passes through the origin.

EXERCISES

Exercises 1-4. Determine a vector form representation for the line in $\Re^2$ passing through the origin and the given point.

1. $(1, 5)$
2. $(-2, 4)$
3. $(5, -1)$
4. $(-2, -2)$

Exercises 5-8. Determine a vector form representation for the line in $\Re^2$ passing through the two given points.

5. $(1, 3), (2, -4)$
6. $(3, 5), (5, 5)$
7. $(3, 5), (3, 7)$
8. $(2, 4), (-5, -2)$

Exercises 9-12. Determine two different vector form representations for the line in $\Re^2$ passing through the two given points, and then proceed to show that the set of points associated with the two vector forms are one and the same.

9. $(1, 3), (2, -4)$
10. $(3, 5), (5, 5)$
11. $(3, 5), (3, 7)$
12. $(2, 4), (-5, -2)$

Exercises 13-20. Determine a vector form representation for the line in $\Re^2$ that passes through the point $(3, 7)$ and is parallel[1] to the line of:

13. Exercise 1
14. Exercise 2
15. Exercise 3
16. Exercise 4
17. Exercise 5
18. Exercise 6
19. Exercise 7
20. Exercise 8

Exercises 21-28. Determine a vector form representation for the line in $\Re^2$ that passes through the point $(3, 7)$ and is perpendicular[2] to the line of:

21. Exercise 1
22. Exercise 2
23. Exercise 3
24. Exercise 4
25. Exercise 5
26. Exercise 6
27. Exercise 7
28. Exercise 8

Exercises 29-32. Determine a vector form representation for the line in $\Re^3$ passing through the origin and the given point.

29. $(2, 4, 5)$
30. $(1, 0, 0)$
31. $(-2, 4, 0)$
32. $(-4, -4, -1)$

Exercises 33-36. Determine a vector form representation for the line in $\Re^3$ passing through the two given points.

33. $(2, 4, 5), (3, 1, 1)$
34. $(0, 1, 2), (1, 0, 2)$
35. $(3, 4, -1), (2, 1, 0)$
36. $(5, -1, -1), (2, 2, 2)$

---

[1] Two lines are parallel they have equal slopes, or if both lines are vertical.

[2] Two lines are perpendicular if and only if the slope of one is the negative reciprocal of the slope of the other, or if one line is horizontal and the other is vertical.

Exercises 37-40. Determine two different vector form representations for the line in $\Re^3$ passing through the two given points, and then proceed to show that the set of points associated with the two vector forms are one and the same.

37. $(2, 4, 5), (3, 1, 1)$
38. $(0, 1, 2), (1, 0, 2)$
39. $(3, 4, -1), (2, 1, 0)$
40. $(5, -1, -1), (2, 2, 2)$

Exercises 41-48. Determine a vector form representation for the line in $\Re^3$ that passes through the point $(1, 2, -1)$ and is parallel to the line of:

41. Exercise 29
42. Exercise 30
43. Exercise 31
44. Exercise 32
45. Exercise 33
46. Exercise 34
47. Exercise 35
48. Exercise 36

Exercises 49-52. Determine a vector form representation for the plane passing through the origin and the two given points.

49. $(1, 3, 2), (2, 1, 4)$
50. $(3, 1, 4), (2, 0, 0)$
51. $(2, 0, 0), (0, 2, 0)$
52. $(-3, -2, -1), (2, 4, 1)$

Exercises 53-56. Determine a vector form representation for the plane passing through the three given points.

53. $(3, 4, 1), (2, 1, 5), (1, 1, -1)$
54. $(2, -1, 0), (2, -1, 1), (1, 2, 3)$
55. $(2, 4, -3), (5, 1, 5), (4, 1, -1)$
56. $(-2, -1, 1), (2, -3, 1), (1, 2, 1)$

Exercises 57-60. Determine two different vector form representations for the plane passing through the two given points, and then proceed to show that the set of points associated with the two vector forms are one and the same.

57. $(3, 4, 1), (2, 1, 5), (1, 1, -1)$
58. $(2, -1, 0), (2, -1, 1), (1, 2, 3)$
59. $(2, 4, -3), (5, 1, 5), (4, 1, -1)$
60. $(-2, -1, 1), (2, -3, 1), (1, 2, 1)$

61. Complete the proof of Theorem 2.15. Incidentally:
    You can also show directly that if $\boldsymbol{v_1} = (\boldsymbol{a}, \boldsymbol{b})$ and $\boldsymbol{v_2} = (\boldsymbol{c}, \boldsymbol{d})$ are such that $\boldsymbol{v_2} \neq r\boldsymbol{v_1}$ for any $r \in \Re$, then for any $\boldsymbol{v} \in \Re^2$ there exist $r_1, r_2 \in \Re$ such that $\boldsymbol{v} = r_1\boldsymbol{v_1} + r_2\boldsymbol{v_2}$.
62. Prove Theorem 2.16.
63. Prove Theorem 2.17.
64. Prove Theorem 2.18.
65. Prove Theorem 2.19.
66. Prove Theorem 2.20.

CHAPTER SUMMARY

EUCLIDEAN VECTOR SPACE $\mathfrak{R}^n$	The set of n-tuples, with addition and scalar multiplication given by: $(v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n)$ $r(v_1, v_2, \ldots, v_n) = (rv_1, rv_2, \ldots, rv_n)$
ABSTRACT VECTOR SPACE	A nonempty set V, closed under addition and scalar multiplication, satisfying the following eight properties: (i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (iii) There is a vector in V, denoted by $\mathbf{0}$ such that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for every vector $\mathbf{v}$ in V. (iv) For every vector $\mathbf{v}$ in V, there is a vector $-\mathbf{v}$ in V, such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$. (v) $r(\mathbf{u} + \mathbf{v}) = r\mathbf{u} + r\mathbf{v}$ (vi) $(r + s)\mathbf{u} = r\mathbf{u} + s\mathbf{u}$ (vii) $r(s\mathbf{u}) = (rs)\mathbf{u}$ (viii) $1\mathbf{u} = \mathbf{u}$
SUBTRACTION	$\mathbf{v} - \mathbf{w} = \mathbf{v} + (-\mathbf{w})$
Uniqueness of $\mathbf{0}$ and $-\mathbf{v}$	There is but one vector $\mathbf{0}$ which satisfies the property that $\mathbf{v} + \mathbf{0} = \mathbf{v}$ for every $\mathbf{v}$ in V. For any given vector $\mathbf{v}$ in V, there is but one vector $-\mathbf{v}$ in V such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.
Cancellation Properties	If $\mathbf{v} + \mathbf{z} = \mathbf{w} + \mathbf{z}$, then $\mathbf{v} = \mathbf{w}$ If $r \neq 0$ and $r\mathbf{v} = r\mathbf{w}$, then $\mathbf{v} = \mathbf{w}$. If $\mathbf{v} \neq \mathbf{0}$ and $r\mathbf{v} = s\mathbf{v}$, then $r = s$.
Zero Properties	$0\mathbf{v} = \mathbf{0}$ $r\mathbf{0} = \mathbf{0}$ $r\mathbf{v} = \mathbf{0}$ if and only if $r = 0$ or $\mathbf{v} = \mathbf{0}$
Inverse Properties	$-1\mathbf{v} = -\mathbf{v}$ $-(-\mathbf{v}) = \mathbf{v}$ $r(-\mathbf{v}) = (-r)\mathbf{v} = -(r\mathbf{v})$

SUBSPACE	A nonempty subset S of V which is itself a vector space under the vector addition and scalar multiplication operations of the space V.
Closure says it all	A nonempty subset S of a vector space V is a subspace of V if and only if it is closed under addition and under scalar multiplication.
A one liner	A nonempty subset S of a vector space V is a subspace of V if and only if for every $\boldsymbol{s}_1, \boldsymbol{s}_2 \in S$ and $r \in \Re$, $r\boldsymbol{s}_1 + \boldsymbol{s}_2 \in S$.
Intersection of subspaces	The intersection of any collection of subspaces in a vector space V is again a subspace of V.
PROPER SUBSPACES	A subspace of a vector space V distinct from the trivial subspace $\{\mathbf{0}\}$ and V itself is said to be a proper subspace of V.
Vector form of lines	Let L be the line in $\Re^2$ passing through the origin, and let (a, b) be any point on L distinct from the origin, then, in vector form: $L = \{r\boldsymbol{v} \mid r \in \Re\}$, where $\boldsymbol{v} = (a, b)$ Let L be the line in the plane passing through $P = (x_1, y_1)$ and $Q = (x_2, y_2)$. Then, in vector form, $L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in \Re\}$ where $\boldsymbol{v} = \overrightarrow{PQ}$, and $\boldsymbol{u} = (\boldsymbol{x_0}, \boldsymbol{y_0})$, with (x_0, y_0) any chosen point on L. Let L be the line in $\Re^3$ passing through two distinct points $P = (x_1, y_1, z_1)$ and $Q = (x_2, y_2, z_2)$. Then $L = \{\boldsymbol{u} + r\boldsymbol{v} \mid r \in R\}$ where $\boldsymbol{v}$ is the **direction** vector $\boldsymbol{v} = \overrightarrow{PQ} = (x_2 - x_1, y_2 - y_1, z_2 - z_1)$ and $\boldsymbol{u} = (x_0, y_0, z_0)$ is a **translation vector**, with (x_0, y_0, z_0) any chosen point on L.
Vector form of planes	Let P be a plane in $\Re^3$ passing through the origin. Let (x_1, y_1, z_1) and (x_2, y_2, z_2) be any two points on P which do not both lie on a common line passing through the origin. Then, in **vector form**: $P = \{r\boldsymbol{u} + s\boldsymbol{v} \mid r, s \in \Re\}$ where $\boldsymbol{u} = (\boldsymbol{x_1}, \boldsymbol{y_1}, \boldsymbol{z_1})$ and $\boldsymbol{v} = (\boldsymbol{x_2}, \boldsymbol{y_2}, \boldsymbol{z_2})$. Let P be the plane passing through three non-colinear points $P_1 = (x_1, y_1, z_1)$, $P_2 = (x_2, y_2, z_2)$ and $P_3 = (x_3, y_3, z_3)$. Then $P = \{\boldsymbol{w} + r\boldsymbol{u} + s\boldsymbol{v} \mid r \in R, s \in \Re\}$ where $\boldsymbol{u} = \overrightarrow{P_1P_2}$, $\boldsymbol{v} = \overrightarrow{P_1P_3}$, and $\boldsymbol{w} = (x_0, y_0, z_0)$ is any chosen point on P

CHAPTER 3

BASES AND DIMENSION

It is easy to see that every vector $(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z})$ in $\Re^3$ can uniquely be expressed in terms of the three vectors $(\mathbf{1}, \mathbf{0}, \mathbf{0})$, $(\mathbf{0}, \mathbf{1}, \mathbf{0})$, and $(\mathbf{0}, \mathbf{0}, \mathbf{1})$. For example:

$$(\mathbf{5}, \mathbf{2}, \mathbf{-9}) = 5(\mathbf{1}, \mathbf{0}, \mathbf{0}) + 2(\mathbf{0}, \mathbf{1}, \mathbf{0}) + (-9)(\mathbf{0}, \mathbf{0}, \mathbf{1})$$

In this chapter, we consider an arbitrary vector space V to see if we can find a set of vectors $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ in V, called a basis for V, such that every vector $\boldsymbol{v}$ in V can be **uniquely** expressed in the form:

$$\boldsymbol{v} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_n\boldsymbol{v_n}$$

for some scalars $c_1, c_2, \ldots, c_n$. As you will see, while many such sets of vectors $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ may exist, the number of vectors in those sets will always be the same. For example, if you find a basis for a vector space V that contains 5 vectors, $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}, \boldsymbol{v_4}, \boldsymbol{v_5}\}$, and someone else finds another basis, that other basis must also consist of 5 vectors. That being the case, we will then be in a position to say that the vector space V has dimension 5.

§1. SPANNING SETS

Using vector addition and scalar multiplication in $\Re^2$, one can build the vector $\boldsymbol{v} = (\mathbf{9}, \mathbf{14})$ from the vectors $\boldsymbol{v_1} = (\mathbf{3}, \mathbf{4})$ and $\boldsymbol{v_2} = (\mathbf{1}, \mathbf{2})$:

$$(\mathbf{9}, \mathbf{14}) = 2(\mathbf{3}, \mathbf{4}) + 3(\mathbf{1}, \mathbf{2})$$

and we say that $\boldsymbol{v} = (\mathbf{9}, \mathbf{14})$ is a linear combination of $\boldsymbol{v_1} = (\mathbf{3}, \mathbf{4})$ and $\boldsymbol{v_2} = (\mathbf{1}, \mathbf{2})$. In general:

DEFINITION 3.1

LINEAR COMBINATION

A vector $\boldsymbol{v}$ in a vector space V is said to be a **linear combination** of vectors $\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}$ in V, if there exist scalars $c_1, c_2, \ldots, c_n$ such that:

$$\boldsymbol{v} = c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_n\boldsymbol{v_n}$$

EXAMPLE 3.1

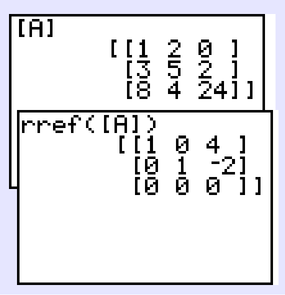
Determine whether or not the vector $(\mathbf{0}, \mathbf{2}, \mathbf{24})$ in $\Re^3$ is a linear combination of the vectors $(\mathbf{1}, \mathbf{3}, \mathbf{8})$ and $(\mathbf{2}, \mathbf{5}, \mathbf{4})$.

SOLUTION: We are to see if we can find scalars a and b such that:

$$(\mathbf{0, 2, 24}) = a(\mathbf{1, 3, 8}) + b(\mathbf{2, 5, 4})$$

$$\text{or:}\quad (\boldsymbol{a + 2b, 3a + 5b, 8a + 4b}) = (\mathbf{0, 2, 24})$$

Equating coefficients, we come to the following system of three equations in two unknowns:



```
[A]
      [[1 2 0 ]
       [3 5 2 ]
       [8 4 24]]
rref([A])
      [[1 0 4 ]
       [0 1 -2]
       [0 0 0 ]]
```

$$\mathbf{S}:\ \left.\begin{aligned} a+2b &= 0\\ 3a+5b &= 2\\ 8a+4b &= 24\end{aligned}\right\}\xrightarrow{\textbf{aug}(S)}\left[\begin{array}{cc|c} a & b & \\ 1 & 2 & 0\\ 3 & 5 & 2\\ 8 & 4 & 24\end{array}\right]\xrightarrow{\textbf{rref}}\left[\begin{array}{cc|c} a & b & \\ 1 & 0 & 4\\ 0 & 1 & -2\\ 0 & 0 & 0\end{array}\right]$$

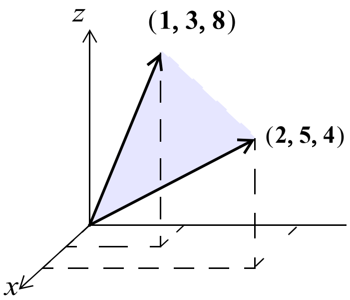
Solution: $a = 4, b = -2$

Conclusion: $(\mathbf{0, 2, 24})$ is a linear combination of the vectors $(\mathbf{1, 3, 8})$ and $(\mathbf{2, 5, 4})$:

$$(\mathbf{0, 2, 24}) = 4(\mathbf{1, 3, 8}) + (-2)(\mathbf{2, 5, 4})$$

Some Added Insight on Example 3.1

The set of all linear combinations of $(\mathbf{1, 3, 8})$ and $(\mathbf{2, 5, 4})$ is the plane in $\Re^3$ containing those vectors. As such, were we to pick an arbitrary point in $\Re^3$, say $(\mathbf{2, 4, 3})$, then there would be little chance that it would lie in that plane, and would therefore not be a linear combination of the two given vectors:



Note that, except for the last column, this augmented matrix is the same as that of Example 3.1.

$$\mathbf{S}:\ \left.\begin{aligned} a+2b &= 2\\ 3a+5b &= 4\\ 8a+4b &= 3\end{aligned}\right\}\xrightarrow{\textbf{aug}(S)}\left[\begin{array}{cc|c} a & b & \\ 1 & 2 & 2\\ 3 & 5 & 4\\ 8 & 4 & 3\end{array}\right]\xrightarrow{\textbf{rref}}\left[\begin{array}{cc|c} a & b & \\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{array}\right]$$

$$\left.\begin{aligned} a+0b &= 0\\ 0a+b &= 0\\ 0a+0b &= 1\end{aligned}\right\}$$

No solution!

CHECK YOUR UNDERSTANDING 3.1

Determine if the given vector is a linear combination of the vectors $(\mathbf{1, 3, 8})$ and $(\mathbf{2, 5, 4})$.

(a) $(\mathbf{-2, -3, 8})$ (b) $(\mathbf{-2, -4, 8})$

Answers: (a) No (b) Yes

THEOREM 3.1 The set of linear combinations of the set of vectors $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ in a vector space V is a subspace of V.

PROOF: Let $S = \{c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \ldots + c_n\boldsymbol{v_n}\}$. We first observe that since $\boldsymbol{0} = 0\boldsymbol{v_1} + 0\boldsymbol{v_2} + \ldots + 0\boldsymbol{v_n} \in S$, S is not empty. Moreover, for $\boldsymbol{u} = a_1\boldsymbol{v_1} + a_2\boldsymbol{v_2} + \ldots + a_n\boldsymbol{v_n}$ and $\boldsymbol{w} = b_1\boldsymbol{v_1} + b_2\boldsymbol{v_2} + \ldots + b_n\boldsymbol{v_n}$ in S, and $r \in \Re$:

$$\begin{aligned} r\boldsymbol{u} + \boldsymbol{w} &= r(a_1\boldsymbol{v_1} + a_2\boldsymbol{v_2} + \ldots + a_n\boldsymbol{v_n}) + (b_1\boldsymbol{v_1} + b_2\boldsymbol{v_2} + \ldots + b_n\boldsymbol{v_n}) \\ &= (ra_1 + b_1)\boldsymbol{v_1} + (ra_2 + b_2)\boldsymbol{v_2} + \ldots + (ra_n + b_n)\boldsymbol{v_n} \underset{\underset{\text{a linear combination of the } v_i\text{'s}}{\uparrow}}{\in} S \end{aligned}$$

$$\begin{tikzcd} \boldsymbol{u}, \boldsymbol{w} \in S \text{ and } r \in \Re \arrow[d] \\ r\boldsymbol{u} + \boldsymbol{w} \in S \end{tikzcd}$$

DEFINITION 3.2
SPANNING

The set of linear combinations of a set of vectors $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ in a vector space V is called the **subspace of V spanned** by $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ and will be denoted by $\mathrm{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$.

In the event that $V = \mathrm{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$, $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ is said to **span** V.

EXAMPLE 3.2 Determine if the vectors $2x^2 + 3x - 1$, $x - 5$, and $x^2 - 1$ span the space $P_2(x)$ of polynomials of degree less than or equal to two.

SOLUTION: We consider an arbitrary vector $ax^2 + bx + c$ in $P_2(x)$ to see whether or not we can find scalars r, s, and t such that:

$$r(2x^2 + 3x - 1) + s(x - 5) + t(x^2 - 1) = ax^2 + bx + c$$

Expanding and combining the left side of the above equation brings us to:

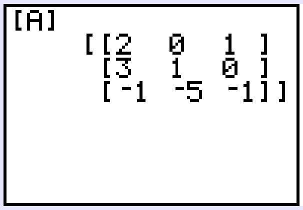
$$(2r + t)x^2 + (3r + s)x + (-r - 5s - t) = ax^2 + bx + c$$

Equating coefficients we come to the following system of three equations, in the **unknowns** r, s, and t (the a, b, and c are **not** variables, they are the fixed coefficients of the polynomial $ax^2 + bx + c$):

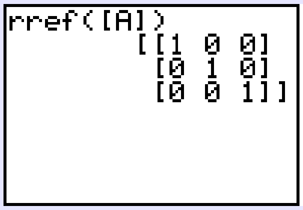
$$\mathbf{S}: \quad \left.\begin{aligned} 2r + t &= a \\ 3r + s &= b \\ -r - 5s - t &= c \end{aligned}\right\}$$

System **S** was solved directly in Example 1.7, page 16. In that example, we labeled the variables x, y, and z, instead of r, s, and t.

Can we find values for r, s, and t such that the above three equations hold? The Spanning Theorem (page 18), tells us that the answer is "yes" if and only if $\text{rref}[\text{coef}(S)]$ does not contain a row consisting entirely of zeros. Let's see:



```
[A]
     [[2   0   1 ]
      [3   1   0 ]
      [-1  -5  -1]]
```



```
rref([A])
          [[1 0 0]
           [0 1 0]
           [0 0 1]]
```

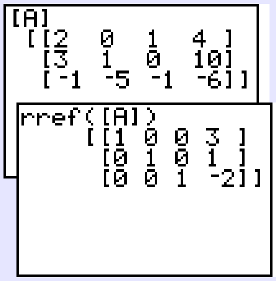
Since the rref-matrix does not contain a row of zeros, system **S** has a solution for all values of a, b, and c, and we conclude that the vectors $2x^2+3x-1$, $x-5$, and x^2-1 span the space $P_2(x)$.

While the above does not tell you how to build ax^2+bx+c from $2x^2+3x-1$, $x-5$, and x^2-1, it does tell you that it can be done, for each and every polynomial in $P_2(x)$. If you want to see how to build any particular polynomial, say the polynomial $4x^2+10x-6$, then that's not a problem, for the task reduces to finding scalars r, s, and t, such that:

$$4x^2+10x-6 = r(2x^2+3x-1)+s(x-5)+t(x^2-1)$$

Or: $$4x^2+10x-6 = (2r+t)x^2+(3r+s)x+(-r-5s-t)$$

Equating coefficients, we have:



```
[A]
 [[2   0   1   4 ]
  [3   1   0   10]
  [-1  -5  -1  -6]]
rref([A])
      [[1 0 0 3 ]
       [0 1 0 1 ]
       [0 0 1 -2]]
```

$$\textbf{S:}\quad \left.\begin{aligned} 2r+t &= 4\\ 3r+s &= 10\\ -r-5s-t &= -6 \end{aligned}\right\} \xrightarrow{\textbf{aug}(S)} \begin{array}{c} \begin{array}{cccc} r & s & t & \end{array}\\ \left[\begin{array}{ccc|c} 2 & 0 & 1 & 4\\ 3 & 1 & 0 & 10\\ -1 & -5 & -1 & -6 \end{array}\right]\end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{cccc} r & s & t & \end{array}\\ \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3\\ 0 & 1 & 0 & 1\\ 0 & 0 & 1 & -2 \end{array}\right]\end{array}$$

Conclusion:

$$4x^2+10x-6 = 3(2x^2+3x-1)+1(x-5)+(-2)(x^2-1)$$

CHECK YOUR UNDERSTANDING 3.2

(a) Use the Spanning Theorem (page 18) to show that the vectors $\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}, \begin{bmatrix}1 & 0\\ 1 & 0\end{bmatrix}, \begin{bmatrix}0 & 1\\ 0 & 1\end{bmatrix}, \begin{bmatrix}0 & 4\\ 2 & 0\end{bmatrix}$ span $M_{2\times 2}$.

(b) Express $\begin{bmatrix}-1 & 5\\ 1 & 13\end{bmatrix}$ as a linear combination of the above four vectors.

Answer: See page B7.

EXAMPLE 3.3 Do the vectors $(\mathbf{2, 1, 0, 1})$, $(\mathbf{1, -2, 2, 0})$, $(\mathbf{2, 3, 1, -1})$ and $(\mathbf{1, 2, 4, -4})$ span $\Re^4$?

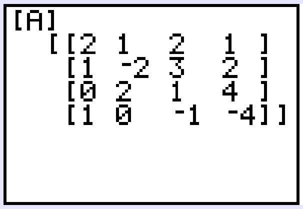
Solution: Let $(\boldsymbol{a, b, c, d})$ be an arbitrary vector in $\Re^4$. Are there scalars x, y, z, and w such that:

$$(\boldsymbol{a, b, c, d}) = x(\mathbf{2, 1, 0, 1}) + y(\mathbf{1, -2, 2, 0}) + z(\mathbf{2, 3, 1, -1}) + w(\mathbf{1, 2, 4, -4})$$

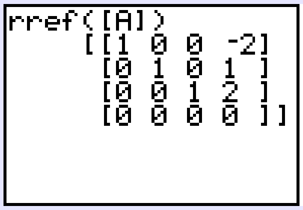
which is to say: does the following system of equations have a solution for all values of a, b, c, and d?

$$\mathbf{S:}\quad \left.\begin{aligned} 2x + y + 2z + w &= a \\ x - 2y + 3z + 2w &= b \\ 0x + 2y + z + 4w &= c \\ x + 0y - z - 4w &= d \end{aligned}\right\}$$

The Spanning Theorem tells us that it does not, as rref[coef($\boldsymbol{S}$)] contains a row consisting entirely of zeros:



```
[A]
  [[2 1  2  1 ]
   [1 -2 3  2 ]
   [0 2  1  4 ]
   [1 0  -1 -4]]
```



```
rref([A])
    [[1 0 0 -2]
     [0 1 0 1 ]
     [0 0 1 2 ]
     [0 0 0 0 ]]
```

While the above argument establishes the fact that the vectors $(\mathbf{2, 1, 0, 1})$, $(\mathbf{1, -2, 2, 0})$, $(\mathbf{2, 3, 1, -1})$ and $(\mathbf{1, 2, 4, -4})$ do not span $\Re^4$, it does not define for us the subspace of $\Re^4$ spanned by those four vectors; bringing us to:

EXAMPLE 3.4 Determine the subspace of $\Re^4$ spanned by $(\mathbf{2, 1, 0, 1})$, $(\mathbf{1, -2, 2, 0})$, $(\mathbf{2, 3, 1, -1})$ and $(\mathbf{1, 2, 4, -4})$.

Solution: We are to find the set of all vectors $(\boldsymbol{a, b, c, d})$ for which there exist scalars x, y, z, and w such that:

$$(\boldsymbol{a, b, c, d}) = x(\mathbf{2, 1, 0, 1}) + y(\mathbf{1, -2, 2, 0}) + z(\mathbf{2, 3, 1, -1}) + w(\mathbf{1, 2, 4, -4})$$

which again boils down to a consideration of a system of equations:

$$\mathbf{S:}\quad \left.\begin{aligned} 2x + y + 2z + w &= a \\ x - 2y + 3z + 2w &= b \\ 0x + 2y + z + 4w &= c \\ x + 0y - z - 4w &= d \end{aligned}\right\}$$

for to say that $(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})$ is in the space spanned by the four given vectors is to say that system **S** has a solution for those numbers a, b, c, and d. That system appeared earlier in Example 1.8, page 17 where it was noted that the given system of equation has a solution if and only if:

$$-10\mathbf{a} + 7\mathbf{b} + 12\mathbf{c} + 13\mathbf{d} = 0$$

Bringing us to a representation for the space spanned by the four given vectors:

$$\text{Span}\{(\mathbf{2}, \mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{-2}, \mathbf{2}, \mathbf{0}), (\mathbf{2}, \mathbf{3}, \mathbf{1}, \mathbf{-1}), (\mathbf{1}, \mathbf{2}, \mathbf{4}, \mathbf{-4})\}$$
$$= \{(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) | -10a + 7b + 12c + 13d = 0\}$$

CHECK YOUR UNDERSTANDING 3.3

Determine the space spanned by the vectors $(\mathbf{2}, \mathbf{1}, \mathbf{5})$, $(\mathbf{1}, \mathbf{-2}, \mathbf{2})$, $(\mathbf{0}, \mathbf{5}, \mathbf{1})$. It is not all of $\Re^3$, exhibit a vector in $\Re^3$ that is not in $\text{Span}\{(\mathbf{2}, \mathbf{1}, \mathbf{5}), (\mathbf{1}, \mathbf{-2}, \mathbf{2}), (\mathbf{0}, \mathbf{5}, \mathbf{1})\}$.

Answer: See page B8.

The following example differs from the previous two in that it is not confined to a specific concrete vector space, like $\Re^4$.

EXAMPLE 3.5 Let the set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ and $\{\mathbf{w_1}, \mathbf{w_2}\}$ be such that $\mathbf{w_i} \in \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$ for $1 \leq i \leq 2$. Show that $\text{Span}\{\mathbf{w_1}, \mathbf{w_2}\} \subseteq \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}$.

SOLUTION: In any non-routine problem, it is important that you chart out, in one way or another, what is given and that which is to be established:

> We are **given** that the vectors $\mathbf{w_1}$ and $\mathbf{w_2}$ are in the space spanned by the three vectors $\mathbf{v_1}, \mathbf{v_2},$ and $\mathbf{v_3}$, and are **to show** that for any given scalars a and b, the vector $a\mathbf{w_1} + b\mathbf{w_2}$ can be written as a linear combination of the vectors $\mathbf{v_1}, \mathbf{v_2},$ and $\mathbf{v_3}$; which is to say that we **can find** scalars c_1, c_2, c_3 such that:

$$a\mathbf{w_1} + b\mathbf{w_2} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + c_3\mathbf{v_3}$$

From the **given** information, we know that there exist scalars a_1, a_2, a_3 and b_1, b_2, b_3 such that:

$$\mathbf{w_1} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + a_3\mathbf{v_3} \text{ and } \mathbf{w_2} = b_1\mathbf{v_1} + b_2\mathbf{v_2} + b_3\mathbf{v_3}$$

Consequently:

$$\begin{aligned} a\mathbf{w}_1 + b\mathbf{w}_2 &= a(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + a_3\mathbf{v}_3) + b(b_1\mathbf{v}_1 + b_2\mathbf{v}_2 + b_3\mathbf{v}_3) \\ &= aa_1\mathbf{v}_1 + aa_2\mathbf{v}_2 + aa_3\mathbf{v}_3 + bb_1\mathbf{v}_1 + bb_2\mathbf{v}_2 + bb_3\mathbf{v}_3 \\ &= (\underbrace{aa_1 + bb_1}_{\swarrow})\mathbf{v}_1 + (aa_2 + bb_2)\mathbf{v}_2 + (aa_3 + bb_3)\mathbf{v}_3 \\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 \end{aligned}$$

The following theorem generalizes the situation of Example 3.5.

THEOREM 3.2 Let the set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ and $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ be such that

$\mathbf{w}_i \in \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for $1 \le i \le m$ .Then:

$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$.

PROOF: If $\mathbf{w} \in \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ then, for some scalars c_i:

$$\mathbf{w} = c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \ldots + c_m\mathbf{w}_m$$

We also know that for each $1 \le i \le m$, there exist scalars $a_{i1}, a_{i2}, \ldots, a_{in}$, such that:

$$\mathbf{w}_i = a_{i1}\mathbf{v}_1 + a_{i2}\mathbf{v}_2 + \ldots + a_{in}\mathbf{v}_n$$

Consequently:

$$\begin{aligned} \mathbf{w} &= c_1\mathbf{w}_1 + \ldots + c_m\mathbf{w}_m \\ &= c_1(a_{11}\mathbf{v}_1 + \ldots + a_{1n}\mathbf{v}_n) + \ldots + c_m(a_{m1}\mathbf{v}_1 + \ldots + a_{mn}\mathbf{v}_n) \\ &= (c_1a_{11} + \ldots + c_ma_{m1})\mathbf{v}_1 + \ldots + (c_1a_{1n} + \ldots + c_ma_{mn})\mathbf{v}_n \end{aligned}$$

Since $\mathbf{w}$ can be expressed as a linear combination of $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, $\mathbf{w} \in \text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$.

Consequently:

$$\text{Span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\} \subseteq \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$$

CHECK YOUR UNDERSTANDING 3.4

Prove that for any three vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ in a vector space V:

$$\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\} = \text{Span}\{\mathbf{v}_1, \mathbf{v}_1 + \mathbf{v}_2, \mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3\}$$

Answer: See page B8.

EXERCISES

Exercises 1-4. Determine whether or not the given vector in $\Re^3$ is a linear combination of the vectors $(\mathbf{-1, 2, 1})$ and $(\mathbf{2, 0, 3})$.

1. $(\mathbf{2, 3, 4})$
2. $(\mathbf{1, 4, -2})$
3. $(\mathbf{-5, 6, 0})$
4. $(\mathbf{-1, 10, 11})$

Exercises 5-8. Determine whether or not the given vector in $M_{2\times 2}$ is a linear combination of the vectors $\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}$.

5. $\begin{bmatrix} 5 & 9 \\ 7 & 3 \end{bmatrix}$
6. $\begin{bmatrix} 4 & 9 \\ 7 & 3 \end{bmatrix}$
7. $\begin{bmatrix} 11 & 6 \\ 6 & 7 \end{bmatrix}$
8. $\begin{bmatrix} 6 & 11 \\ 7 & 6 \end{bmatrix}$

Exercises 9-12. Determine whether or not the given vector in P_3 is a linear combination of the vectors $x^3 + 1, 2x^2 - 3,$ and $x + 1$.

9. $2x^3 + x + 1$
10. $3x^2 - x - 2$
11. $2x^3 + 4x^2 - x - 6$
12. $2x^3 + 6x^2 + x - 6$

Exercises 13-16. Show that the given vector in the function space $F(x)$ of Theorem 2.4, page 44, is in the space spanned by the vectors $\sin x, \cos x, \sin^2 x, \cos^2 x, \tan^2 x, \cot^2 x$.

13. $\cos 2x$
14. $\sin\left(\frac{\pi}{2} - x\right)$
15. $\sin\left(\frac{\pi}{7} - x\right)$
16. $\sin\left(\frac{\pi}{7} - x\right)$

Exercises 17-21. Determine if the given vectors span $\Re^4$. If not, find a specific vector in $\Re^4$ which is not contained in that subspace.

17. $(\mathbf{1, 0, 0, 0}), (\mathbf{0, 2, 0, 0}), (\mathbf{0, 0, 3, 0}), (\mathbf{0, 0, 0, 4})$
18. $(\mathbf{1, 0, 0, 0}), (\mathbf{1, 1, 0, 0}), (\mathbf{1, 1, 1, 0}), (\mathbf{1, 1, 1, 1})$
19. $(\mathbf{1, 1, 1, 1}), (\mathbf{1, 0, 0, 1}), (\mathbf{0, 1, 1, 0}), (\mathbf{6, 4, 4, 4})$
20. $(\mathbf{2, 1, 3, 1}), (\mathbf{1, 2, 1, 3}), (\mathbf{3, 1, 1, 2}), (\mathbf{1, 1, 2, 3})$
21. $(\mathbf{-1, -2, -3, -4}), (\mathbf{-3, 1, 1, 2}), (\mathbf{1, 2, 1, 3}), (\mathbf{7, 0, -3, -2})$

Exercises 22-25. Determine if the given vectors span P_3. If not, find a specific vector in P_3 which is not contained in that subspace.

22. $1, x, 2x^2, 3x^3$
23. $1 + x^2, 1 + x, -x + x^2, 1 + x + x^2 + x^3$
24. $x, 1 + x, 1 + x + x^2, 1 + x + x^2 + x^3$
25. $1, x, x^2 + 1, x^3 - x^2$
26. For what values of c do the vectors $(2, 1, 3), (4, 3, 5), (0, 0, c)$ span $\Re^3$?
27. For what values of c do the vectors $c, 2x, 2x^2, 2x^2 + 2x + c$ span P_2?

28. For what values of a and b do the vectors $(\boldsymbol{a}, \boldsymbol{b})$ and $(-\boldsymbol{b}, \boldsymbol{a})$ span $\Re^2$?

29. Show that for any given set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, $\boldsymbol{v}_i \in S\text{pan}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ for every $1 \le i \le n$.

30. Let the set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ be such that $\boldsymbol{w}_i \in \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ for $1 \le i \le m$ and $\boldsymbol{v}_i \in \text{Span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ for $1 \le i \le n$. Prove that $\text{Span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\} = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$.

31. Show that if $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ span a vector space V, then for any vector $\boldsymbol{v}_4$ the vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4$ also span V.

32. Show that a nonempty subset S of as vector space V is a subspace of V if and only if $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subseteq S$ for every $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subseteq S$.

33. Let P denote the vector space of all polynomials of Exercise 23, page 50. Show that no finite set of vectors in P spans P.

34. Let S be a subset of a vector space V. Prove that $\text{Span}(S)$ is the intersections of all subspaces of V which contain the set S.

PROVE OR GIVE A COUNTEREXAMPLE

35. If the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ span V, then so do the vectors $\boldsymbol{u}$ and $\boldsymbol{u} + \boldsymbol{v}$.

36. If the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ span V, then so do the vectors $\boldsymbol{u}$ and $\boldsymbol{u} - \boldsymbol{v}$.

37. If the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ are contained in the space spanned by the vectors $\boldsymbol{w}$ and $\boldsymbol{z}$, then $\text{Span}\{\boldsymbol{u}, \boldsymbol{v}\} = \text{Span}\{\boldsymbol{w}, \boldsymbol{z}\}$.

38. If $\text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\} = V$, and if $\boldsymbol{v}_i \in \text{Span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ for $1 \le i \le 3$, then $\text{Span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\} = V$.

39. If S_1 and S_2 are finite sets of vectors in a vector space V, then:
$$\text{Span}(S_1 \cap S_2) = \text{Span}(S_1) \cap \text{Span}(S_2).$$

40. If S_1 and S_2 are finite sets of vectors in a vector space V, then:
$$\text{Span}(S_1 \cap S_2) \subseteq \text{Span}(S_1) \cap \text{Span}(S_2)$$

41. If S_1 and S_2 are finite sets of vectors in a vector space V, then:
$$\text{Span}(S_1 \cup S_2) = \text{Span}(S_1) \cup \text{Span}(S_2).$$

42. If S_1 and S_2 are subspaces of a vector space V, then:
$$\text{Span}(S_1 \cup S_2) = \text{Span}(S_1) \cup \text{Span}(S_2).$$

§2. LINEAR INDEPENDENCE

The subspace of $\Re^3$ spanned by the vectors $(\mathbf{1, 0, 0})$, $(\mathbf{0, 1, 0})$, $(\mathbf{2, 5, 0})$ is built from those vectors. But there is a kind of inefficiency with the three building blocks $\{(\mathbf{1, 0, 0}), (\mathbf{0, 1, 0}), (\mathbf{2, 5, 0})\}$, in that whatever can be built from those three vectors can be built with just two of them; as one of them, say the vector $(\mathbf{2, 5, 0})$, can itself be constructed from the other two:

$$(\mathbf{2, 5, 0}) = 2(\mathbf{1, 0, 0}) + 5(\mathbf{0, 1, 0})$$

The following concept, as you will soon see, addresses the above "inefficiency issue:"

Note that if each $c_i = 0$, then surely
$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$
will equal zero.
To say that $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ is linearly independent, is to say that **no other** linear combination of the vectors equals $\mathbf{0}$.

DEFINITION 3.3
Linearly Independent

A set of vectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ are **linearly independent** if:

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n = \mathbf{0} \Rightarrow \text{each } c_i = \mathbf{0}$$

A collection of vectors that are not linearly independent is said to be **linearly dependent**.

EXAMPLE 3.6 Is $\{x^2, x^2 + x, x^2 + x + 1\}$ a linearly independent set in the vector space $P_2(x)$?

SOLUTION: To resolve the issue we consider the equation:

$$ax^2 + b(x^2 + x) + c(x^2 + x + 1) = 0x^2 + 0x + 0$$

Equating coefficients, brings us to the following system of equations:

$$\left.\begin{aligned} a + b + c &= 0 \\ b + c &= 0 \\ c &= 0 \end{aligned}\right\}$$

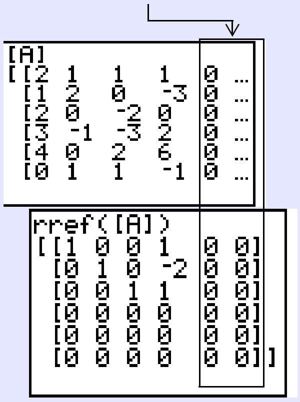
Working from the bottom up, we see that $c = 0, b = 0, a = 0$ is the only solution for the above system of equations, and therefore conclude that the given set of vectors is linearly independent.

EXAMPLE 3.7 Is $\left\{\begin{bmatrix} 2 & 1 & 2 \\ 3 & 4 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & -2 \\ -3 & 2 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -3 & 0 \\ 2 & 6 & -1 \end{bmatrix}\right\}$ a linearly independent set in $M_{2 \times 3}$?

SOLUTION: If:

$$\boldsymbol{a}\begin{bmatrix} 2 & 1 & 2 \\ 3 & 4 & 0 \end{bmatrix} + \boldsymbol{b}\begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 1 \end{bmatrix} + \boldsymbol{c}\begin{bmatrix} 1 & 0 & -2 \\ -3 & 2 & 1 \end{bmatrix} + \boldsymbol{d}\begin{bmatrix} 1 & -3 & 0 \\ 2 & 6 & -1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Most graphing calculators do not have the capability of “rref-ing” a “tall matrix.” But you can always add enough zero columns to arrive at a square matrix:



```
[A]
[[2 1   1   1  0 …
 [1 2   0   -3 0 …
 [2 0   -2  0  0 …
 [3 -1  -3  2  0 …
 [4 0   2   6  0 …
 [0 1   1   -1 0 …
```

```
rref([A])
[[1 0 0 1   0 0]
 [0 1 0 -2  0 0]
 [0 0 1 1   0 0]
 [0 0 0 0   0 0]
 [0 0 0 0   0 0]
 [0 0 0 0   0 0]]
```

Then:

$$\begin{bmatrix} 2a+b+c+d & a+2b-3d & 2a-2c \\ 3a-b-3c+2d & 4a+2c+6d & b+c-d \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Leading us to the homogeneous system:

$$S: \left.\begin{array}{rcl} 2a+b+c+d &=& 0 \\ a+2b-3d &=& 0 \\ 2a-2c &=& 0 \\ 3a-b-3c+2d &=& 0 \\ 4a+2c+6d &=& 0 \\ b+c-d &=& 0 \end{array}\right\} \xrightarrow{\text{coef(S)}} \begin{array}{c} \begin{matrix} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} \end{matrix} \\ \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 0 & -3 \\ 2 & 0 & -2 & 0 \\ 3 & -1 & -3 & 2 \\ 4 & 0 & 2 & 6 \\ 0 & 1 & 1 & -1 \end{bmatrix} \end{array} \xrightarrow{\text{rref}} \begin{array}{c} \begin{matrix} \mathbf{a} & \mathbf{b} & \mathbf{c} & \mathbf{d} \end{matrix} \\ \begin{bmatrix} \mathbf{1} & 0 & 0 & 1 \\ 0 & \mathbf{1} & 0 & -2 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Since rref[coef(S)] has a free variable, the system has nontrivial solutions, and we therefore conclude that the given set of vectors is linearly dependent.

CHECK YOUR UNDERSTANDING 3.5

Is $\{x^2, 2x^2+x, x-3\}$ a linearly independent set in P_2?

Answer: Yes.

EXAMPLE 3.8

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ be a linearly independent set of vectors in a vector space V. Show that $\{\boldsymbol{v}_1, \boldsymbol{v}_2+\boldsymbol{v}_1, \boldsymbol{v}_3-\boldsymbol{v}_2\}$ is also linearly independent.

SOLUTION: We start with: $a\boldsymbol{v}_1 + b(\boldsymbol{v}_2+\boldsymbol{v}_1) + c(\boldsymbol{v}_3-\boldsymbol{v}_2) = 0$ and go on to show that $a = b = c = 0$:

$$a\boldsymbol{v}_1 + b(\boldsymbol{v}_2+\boldsymbol{v}_1) + c(\boldsymbol{v}_3-\boldsymbol{v}_2) = \mathbf{0}$$

$$\text{regroup:} \quad (a+b)\boldsymbol{v}_1 + (b-c)\boldsymbol{v}_2 + c\boldsymbol{v}_3 = \mathbf{0}$$

Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is linearly independent:

$$\left.\begin{array}{rcl} a+b &=& 0 \\ b-c &=& 0 \\ c &=& 0 \end{array}\right\} \Rightarrow a = 0, b = 0, c = 0$$

CHECK YOUR UNDERSTANDING 3.6

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ be a linearly independent set of vectors in a vector space V. Show that $\{\boldsymbol{v}_1, \boldsymbol{v}_1+\boldsymbol{v}_2, \boldsymbol{v}_1+\boldsymbol{v}_2+\boldsymbol{v}_3, \boldsymbol{v}_1+\boldsymbol{v}_2+\boldsymbol{v}_3+\boldsymbol{v}_4\}$ is also a linearly independent set.

Answer: See page B-8.

Here is a useful consequence of the Linear Independence Theorem of page 22:

THEOREM 3.3
LINEAR INDEPENDENCE THEOREM FOR $\Re^n$

A set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$ in $\Re^n$ is linearly independent if and only if the row-reduced-echelon form of the $n \times m$ matrix with i^{th} column the (vertical) n-tuple $\boldsymbol{v}_i$ has m leading ones.

PROOF: For $1 \le i \le m$, let $\boldsymbol{v}_i = (\boldsymbol{a_{1i}}, \boldsymbol{a_{2i}}, \dots, \boldsymbol{a_{ni}})$. To challenge linear independence of those m vectors, we consider the vector equation:

$$c_1\boldsymbol{v}_1 + \dots + c_m\boldsymbol{v}_m = \mathbf{0}:$$

$$c_1(\boldsymbol{a_{11}}, \boldsymbol{a_{21}}, \dots, \boldsymbol{a_{n1}}) + c_2(\boldsymbol{a_{12}}, \boldsymbol{a_{22}}, \dots, \boldsymbol{a_{n2}}) + \dots + c_m(\boldsymbol{a_{1m}}, \boldsymbol{a_{2m}}, \dots, \boldsymbol{a_{nm}}) = (\mathbf{0}, \dots, \mathbf{0})$$

$$(\boldsymbol{c_1a_{11}} + \boldsymbol{c_2a_{12}} + \cdots + \boldsymbol{c_ma_{1m}}, \dots, \boldsymbol{c_1a_{n1}} + \boldsymbol{c_2a_{n2}} + \cdots + \boldsymbol{c_ma_{nm}}) = (\mathbf{0}, \dots, \mathbf{0})$$

Equating coefficients brings us to the following homogeneous system of n equations in m unknowns:

$$\mathbf{S:}\quad \left.\begin{array}{c} a_{11}\boldsymbol{c_1} + a_{12}\boldsymbol{c_2} + \dots + a_{1m}\boldsymbol{c_m} = 0 \\ a_{21}\boldsymbol{c_1} + a_{22}\boldsymbol{c_2} + \dots + a_{2m}\boldsymbol{c}_m = 0 \\ \vdots \quad \vdots \quad \vdots \quad \vdots \qquad \vdots \quad \vdots \\ a_{n1}\boldsymbol{c_1} + a_{n2}\boldsymbol{c_2} + \dots + a_{nm}\boldsymbol{c_m} = 0 \end{array}\right\} \xrightarrow{\mathbf{coef}(\boldsymbol{S})} \left[\begin{array}{cccc} \boldsymbol{c_1} & \boldsymbol{c_2} & & \boldsymbol{c_m} \\ a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{array}\right]$$

Applying the Linearly Independent Theorem of page 22, we conclude that the above homogeneous system of equations has only the trivial solution $c_1 = c_2 = \dots = c_m = 0$ if and only if $\text{rref}[\text{coef}(\boldsymbol{S})]$ has m leading ones.

EXAMPLE 3.9 Determine if:

$$\{(\mathbf{1}, \mathbf{3}, \mathbf{-2}, \mathbf{1}), (\mathbf{2}, \mathbf{1}, \mathbf{3}, \mathbf{2}), (\mathbf{1}, \mathbf{3}, \mathbf{1}, \mathbf{4}), (\mathbf{4}, \mathbf{7}, \mathbf{2}, \mathbf{7})\}$$

is a linearly independent set of vectors in $\Re^4$.

SOLUTION:

$$\begin{bmatrix} 1 & 2 & 1 & 4 \\ 3 & 1 & 3 & 7 \\ -2 & 3 & 1 & 2 \\ 1 & 2 & 4 & 7 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} \mathbf{1} & 0 & 0 & 1 \\ 0 & \mathbf{1} & 0 & 1 \\ 0 & 0 & \mathbf{1} & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Since the above matrix does not have 4 leading ones, the given set of vectors is **not** linearly independent.

CHECK YOUR UNDERSTANDING 3.7

Use Theorem 3.3 to show that there cannot exist a set of five linearly independent vectors in $\Re^4$.

Answer: See page B-9.

THEOREM 3.4 Any set of vectors containing the zero vector **0** is linearly dependent.

PROOF: Let $\{\mathbf{0}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be a subset of a vector space V. Since:

$$\underset{\uparrow \neq 0}{\mathbf{1}}(\mathbf{0}) + 0\boldsymbol{v}_2 + 0\boldsymbol{v}_3 + \ldots + 0\boldsymbol{v}_m = \mathbf{0}$$

$\{\mathbf{0}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is not linearly independent.

The above theorem tells us that $\{\mathbf{0}\}$ is a linearly dependent set in any vector space V. As for the rest:

THEOREM 3.5 A finite set of vectors, distinct from $\{\mathbf{0}\}$, is **linearly independent** if and only if no vector in the set can be expressed as a linear combination of the rest.

CONTRAPOSITIVE PROOF

Let P and Q be two propositions.

You can prove that:

$$P \Rightarrow Q$$

by showing that:

$$\text{Not-}Q \Rightarrow \text{Not-}P$$

(After all if *Not-Q* implies *Not*-P, then you certainly cannot have P without having Q: think about it)

PROOF: To establish the fact that linear independence implies that no vector in the set can be expressed as a linear combination of the rest, we show that if some vector in the set can be expressed as a linear combination of the rest, then the set is not linearly independent (see margin). Here goes:

Assume that one of the vectors in $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ can be expressed as a linear combination of the rest. Since one can always reorder the given vectors we may assume, without loss of generality, that $\boldsymbol{v}_1$ is a linear combination of the rest:

$$\boldsymbol{v}_1 = a_2\boldsymbol{v}_2 + a_3\boldsymbol{v}_3 + \cdots + a_n\boldsymbol{v}_n$$

Leading us to:

$$\mathbf{1}\boldsymbol{v}_1 - a_2\boldsymbol{v}_2 - a_3\boldsymbol{v}_3 - \cdots - a_n\boldsymbol{v}_n = 0$$

Since the coefficient of $\boldsymbol{v}_1$ is not zero, we conclude that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly dependent.

$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ NOT linearly independent

↓

some v_i can be expressed as a linear combination of the rest

To establish the converse we again turn to a contrapositive proof:

If $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ is not linearly independent, then we can find scalars $c_1, c_2, \ldots, c_n$, not all zero, such that:

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \mathbf{0}$$

Assuming, without loss of generality, that $c_1 \neq 0$ we find that we can express $\boldsymbol{v}_1$ as a linear combination of the rest:

$$\boldsymbol{v}_1 = -\frac{c_2}{c_1}\boldsymbol{v}_2 - \frac{c_3}{c_1}\boldsymbol{v}_3 - \cdots - \frac{c_n}{c_1}\boldsymbol{v}_n$$

The next theorem tells us that whatever can be built from a collection of linearly independent vectors, can only be built in one way:

In the exercises you are invited to establish the converse of this theorem.

THEOREM 3.6 Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a linearly independent set. If $a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n = b_1\boldsymbol{v}_1 + \ldots + b_n\boldsymbol{v}_n$, then:

$$a_i = b_i, \text{ for } 1 \leq i \leq n.$$

$$\sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n} b_i v_i$$
$$\Downarrow$$
$$a_i = b_i$$

PROOF:

$$a_1\mathbf{v_1} + a_2\mathbf{v_2} + \cdots + a_n\mathbf{v_n} = b_1\mathbf{v_1} + b_2\mathbf{v_2} + \cdots + b_n\mathbf{v_n}$$
$$a_1\mathbf{v_1} + a_2\mathbf{v_2} + \cdots + a_n\mathbf{v}_n - (b_1\mathbf{v_1} + b_2\mathbf{v_2} + \cdots + b_n\mathbf{v_n}) = \mathbf{0}$$
$$(a_1 - b_1)\mathbf{v_1} + (a_2 - b_2)\mathbf{v_2} + \cdots + (a_n - b_n)\mathbf{v_n} = \mathbf{0}$$

by linear Ind.: $(a_1 - b_1) = 0, (a_2 - b_2) = 0, \ldots, (a_n - b_n) = 0$

$$a_1 = b_1, a_2 = b_2, \ldots, a_n = b_n$$

CHECK YOUR UNDERSTANDING 3.8

Show that the vectors $(\mathbf{2, 1, 3}), (\mathbf{5, 0, 2}), (\mathbf{11, 3, 11})$ are linearly dependent in $\Re^3$, and that $(\mathbf{8, 4, 12})$ is in the space spanned by those vectors. Express $(\mathbf{8, 4, 12})$ as a linear combination of those vectors in two distinct ways.

Answer: See page B-9.

A linearly independent set of vectors S in a vector space V may not be able to accommodate additional vectors without losing its independence. This is not the case if $V \neq \text{Span}(S)$:

THEOREM 3.7
EXPANSION THEOREM

If $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a linearly independent set of vectors, and if $\mathbf{v} \notin \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$, then $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}, \mathbf{v}\}$ is also linearly independent.

PROOF: Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be linearly independent, and let $\mathbf{v} \notin \text{Span}(S)$. We show that $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}, \mathbf{v}\}$ is linearly independent by showing that no vector in $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}, \mathbf{v}\}$ can be expressed as a linear combination of the rest.

To begin with, we observe that $\mathbf{v}$ cannot be expressed as a linear combination of the rest, as $\mathbf{v} \notin \text{Span}(S)$. Suppose then that some other vector in $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}, \mathbf{v}\}$, say for definiteness the vector $\mathbf{v_1}$, can be expressed as a linear combination of the rest:

$$\mathbf{v_1} = c_2\mathbf{v_2} + c_3\mathbf{v_3} + \ldots + c_n\mathbf{v_n} + \mathbf{c}\mathbf{v}$$

Since $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is linearly independent, $\mathbf{c} \neq 0$ (why?). But then:

$$\mathbf{v} = \frac{1}{\mathbf{c}}\mathbf{v_1} - \frac{c_2}{\mathbf{c}}\mathbf{v_2} - \frac{c_3}{\mathbf{c}}\mathbf{v_3} - \cdots - \frac{c_n}{\mathbf{c}}\mathbf{v_n}$$

again contradicting the given condition that $\mathbf{v} \notin \text{Span}(S)$.

CHECK YOUR UNDERSTANDING 3.9

Find a linearly independent set of four vectors in P_3 which includes the two vectors $x^3 + x$ and -7.

Answer: See page B-9.

EXERCISES

Exercises 1-6. Determine if the given set of vectors is a linearly independent set in $\Re^3$.

1. $\{(\mathbf{2}, \mathbf{1}, \mathbf{5}), (\mathbf{4}, \mathbf{1}, \mathbf{10})\}$

2. $\{(\mathbf{0}, \mathbf{0}, \mathbf{0}), (\mathbf{-6}, \mathbf{4}, \mathbf{-5})\}$

3. $\{(\mathbf{2}, \mathbf{1}, \mathbf{5}), (\mathbf{4}, \mathbf{1}, \mathbf{10}), (\mathbf{4}, \mathbf{-1}, \mathbf{10})\}$

4. $\left\{(\mathbf{3}, \mathbf{-2}, \tfrac{\mathbf{5}}{\mathbf{2}}), (\mathbf{-6}, \mathbf{-4}, \mathbf{-5}), (\mathbf{1}, \mathbf{2}, \mathbf{0})\right\}$

5. $\{(\mathbf{1}, \mathbf{3}, \mathbf{4}), (\mathbf{2}, \mathbf{5}, \mathbf{-1}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{2}, \mathbf{2}, \mathbf{0})\}$

6. $\{(\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{0}, \mathbf{2}), (\mathbf{1}, \mathbf{2}, \mathbf{3})\}$

Exercises 7-12. Determine if the given set of vectors is a linearly independent set in $M_{2 \times 2}$.

7. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}\right\}$

8. $\left\{\begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ -5 & -6 \end{bmatrix}\right\}$

9. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix}\right\}$

10. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 2 & 5 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$

12. $\left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} -2 & 2 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

Exercises 13-17. Determine whether the given set of vectors is a linearly independent set in P_3.

13. $\{x+1, x^2+x+1, x^3\}$

14. $\{x+1, x^2+1, 3x-5, 17\}$

15. $\{3x^3, 3x^2+3x+3, 3x^3+3x^2+6x+6, 3x+3\}$

16. $\{2x^3, 3x^2+3x+3, 3x^3+3x^2+6x+6, 3x+3\}$

17. $\{2x^3, 3x^2+3x+3, 3x^3+3x^2+6x+6, 3x+5\}$

Exercises 18-27. Determine if the given set of vectors in the function space $F(X)$ of Theorem 2.4, page 44, is linearly independent.

18. $\{5, \sin x\} \subset F(\Re)$

19. $\{\sin^2 x, \cos^2 x\} \subset F(\Re)$

20. $\{5, \sin^2 x, \cos^2 x\} \subset F(\Re)$

21. $\{\sin^2 x, \cos^2 x, \cos 2x\} \subset F(\Re)$

22. $\{x^2, \sin x\} \subset F(\Re)$

23. $\{\sin^2 x, \cos^2 x, 2\} \subset F(\Re)$

24. $\{e^x, e^{-x}\} \subset F(\Re)$

25. $\{e^x, e^{2x}\} \subset F(\Re)$

26. $\{1, e^x+e^{-x}, e^x-e^{-x}\} \subset F(\Re)$

27. $\{\ln x, \ln(x^2)\} \subset F(\Re^+)$, where $\Re^+$ denotes the set of positive numbers.

28. For what real numbers a is $\{(\mathbf{1}, \mathbf{1}, \boldsymbol{a}), (\mathbf{1}, \boldsymbol{a}, \mathbf{1}), (\boldsymbol{a}, \mathbf{1}, \mathbf{1})\}$ a linearly dependent set in $\Re^3$?

29. For what real numbers a is $\{2x + a + 1, x + 2\}$ a linearly dependent set in $P_1(x)$?

30. For what real numbers a is $\left\{ax^2 - \frac{x}{2} - \frac{1}{2}, -\frac{x^2}{2} + ax - \frac{1}{2}, -\frac{x^2}{2} - \frac{x}{2} + a\right\}$ a linearly dependent set in $P_2(x)$?

31. Find a value of a for which $\{\cos(x + a), \sin x\}$ is a linearly dependent set in the function space $F(\Re)$?

32. Find a value of a for which $\{\sin(x + a), \sin x, \cos x\}$ is a linearly dependent set in the function space $F(\Re)$?

33. Let $\boldsymbol{v}$ be any nonzero vector in a vector space V. Prove that $\{\boldsymbol{v}\}$ is a linearly independent set.

34. Prove that every nonempty subset of a linearly independent set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ is again linearly independent.

35. Prove that if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ is a linearly dependent set in a vector space V, then so is the set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\} \cup \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}$ for any set of vectors $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}$ in V.

36. Establish the converse of Theorem 3.6.

37. Let $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$ be a set of vectors in a space V. Show that if there exists **any vector** $\boldsymbol{v} \in \text{Span}(S)$ which can be **uniquely** expressed as a linear combination of the vectors in S then S is linearly independent.

38. Show that $\{(1, 0), (0, 1)\}$ is a linearly independent set in the vector space of Example 2.5, page 47.

39. Let $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$ be linearly independent sets of vectors in a vector space V with $\text{Span}(S) \cap \text{Span}(T) = \{\mathbf{0}\}$. Prove that $S \cup T$ is also a linearly independent set.

PROVE OR GIVE A COUNTEREXAMPLE

40. If $\{\boldsymbol{u}, \boldsymbol{v}\}$ is a linearly dependent set, then $\boldsymbol{u} = r\boldsymbol{v}$ for some scalar r.

41. If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a linearly dependent set, then $\boldsymbol{u} = r\boldsymbol{v} + s\boldsymbol{w}$ for some scalars r and s.

42. If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a linearly independent set of vectors in a vector space V, then $\{\boldsymbol{u}, \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}+\boldsymbol{w}\}$ is also linearly independent.

43. If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a linearly independent set of vectors in a vector space V, then $\{\boldsymbol{u}-\boldsymbol{v}, \boldsymbol{v}-\boldsymbol{w}, \boldsymbol{w}-\boldsymbol{u}\}$ is also linearly independent.

44. For any three nonzero distinct vectors $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ in a vector space V, $\{\boldsymbol{u}-\boldsymbol{v}, \boldsymbol{v}-\boldsymbol{w}, \boldsymbol{w}-\boldsymbol{u}\}$ is linearly dependent.

45. If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a linearly independent set of vectors in a vector space V, and if $a \in \Re$ then $\{a\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is also linearly independent.

46. If $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is a linearly independent set of vectors in a vector space V, and if a is any nonzero number, then $\{a\boldsymbol{u}, \boldsymbol{u}+a\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}+a\boldsymbol{w}\}$ is also linearly independent.

47. If $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ and $T = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ are linearly independent sets of vectors in a vector space V, then $S \cup T$ is also a linearly independent set.

48. If $\{\boldsymbol{u}, \boldsymbol{v}\}$ and $\{\boldsymbol{v}, \boldsymbol{w}\}$ are linearly independent sets of vectors in a vector space V, then $\{\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}\}$ is also a linearly independent set.

§3. BASES

So far we have considered sets of vectors $S = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ in a space V that satisfy one of two properties:

(1) ***S* spans *V***: Every vector in V can be built from those in S.

(2) ***S* is linearly independent**: No vector in S can be built from the rest.

In a way, (1) and (2) are tugging in numerically opposite directions:

From the spanning point of view:

The more vectors in S the better.

From the linear independence point of view:

The fewer vectors in S the better.

Sometimes, the set of vectors S is not too big, nor too small — it's just right:

DEFINITION 3.4
BASIS

A set of vectors $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ in a vector space V is said to be a **basis for *V*** if:

(1) β spans V

and: (2) β is linearly independent.

In the exercises you are asked to verify the fact that:

$S_2 = \{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$ is a basis for $\Re^2$,

$S_3 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \{(\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{0}, \mathbf{1})\}$ is a basis for $\Re^3$, and that, in general:

STANDARD BASES IN $\Re^n$

$S_n = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$, where each entry in the n-tuple e_i is 0 with the exception of the i^{th} entry which equals 1, is a basis for $\Re^n$, called the **standard basis** of $\Re^n$.

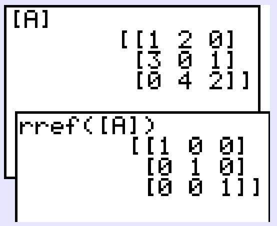
EXAMPLE 3.10

Show that $\{(\mathbf{1}, \mathbf{3}, \mathbf{0}), (\mathbf{2}, \mathbf{0}, \mathbf{4}), (\mathbf{0}, \mathbf{1}, \mathbf{2})\}$ is a basis for $\Re^3$.

SOLUTION: Appealing to Definition 3.4, we challenge the given set of vectors on two fronts: spanning, and linear independence.

Spanning. For any given $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) \in \Re^3$, we are to determine if there exist scalars x, y, and z such that:

$$x(\mathbf{1}, \mathbf{3}, \mathbf{0}) + y(\mathbf{2}, \mathbf{0}, \mathbf{4}) + z(\mathbf{0}, \mathbf{1}, \mathbf{2}) = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$$



```
[A]
        [[1 2 0]
         [3 0 1]
         [0 4 2]]
rref([A])
        [[1 0 0]
         [0 1 0]
         [0 0 1]]
```

If you take the time to solve the system directly, you will find that:

$$x = \frac{2a+2b-c}{8}$$

$$y = \frac{6a-2b+c}{16}$$

$$z = \frac{-6a+2b+3c}{8}$$

Expanding the left side, and equating coefficients, we come to the following 3×3 system of equations:

$$\mathbf{S:}\quad \left.\begin{aligned} x+2y &= a\\ 3x+z &= b\\ 4y+2z &= c\end{aligned}\right\} \xrightarrow{\text{coef}(S)} \begin{bmatrix}1&2&0\\3&0&1\\0&4&2\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}\mathbf{1}&0&0\\0&\mathbf{1}&0\\0&0&\mathbf{1}\end{bmatrix}$$

the three given vectors

Figure 3.1

Applying the Spanning Theorem (page 18) we conclude that $\{(\mathbf{1},\mathbf{3},\mathbf{0}),(\mathbf{2},\mathbf{0},\mathbf{4}),(\mathbf{0},\mathbf{1},\mathbf{2})\}$ spans $\Re^3$.

Linear independence. A consequence of:

$$\begin{bmatrix}1&2&0\\3&0&1\\0&4&2\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}\mathbf{1}&0&0\\0&\mathbf{1}&0\\0&0&\mathbf{1}\end{bmatrix}$$

the three given vectors

and the Linear Independence Theorem for $\Re^n$ (page 88).

> Since $\{(\mathbf{1},\mathbf{3},\mathbf{0}),(\mathbf{2},\mathbf{0},\mathbf{4}),(\mathbf{0},\mathbf{1},\mathbf{2})\}$ is a linearly independent set which spans $\Re^3$, it is a basis for $\Re^3$.

EXAMPLE 3.11 Determine if the following set of matrices constitute a basis for the vector space $M_{2\times 2}$:

$$\left\{\begin{bmatrix}2&1\\3&0\end{bmatrix},\begin{bmatrix}1&1\\2&2\end{bmatrix},\begin{bmatrix}-3&0\\-6&-5\end{bmatrix},\begin{bmatrix}0&4\\1&5\end{bmatrix}\right\}$$

SOLUTION: The problem boils down to a consideration of the coefficient matrix of a system of equations. What system? Well, if you take the spanning approach, then you will be looking at the vector equation:

$$x\begin{bmatrix}2&1\\3&0\end{bmatrix}+y\begin{bmatrix}1&1\\2&2\end{bmatrix}+z\begin{bmatrix}-3&0\\-6&-5\end{bmatrix}+w\begin{bmatrix}0&4\\1&5\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix} \qquad (*)$$

to see if it can be solved for any given matrix $\begin{bmatrix}a&b\\c&d\end{bmatrix}$.

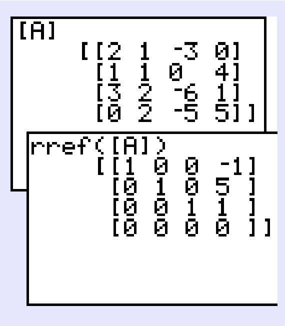
On the other hand, if you take the linear independent approach, then you will consider the vector equation:

$$x\begin{bmatrix}2&1\\3&0\end{bmatrix}+y\begin{bmatrix}1&1\\2&2\end{bmatrix}+z\begin{bmatrix}-3&0\\-6&-5\end{bmatrix}+w\begin{bmatrix}0&4\\1&5\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}\qquad(**)$$

to see if it has only the trivial solution.

In either case, by equating entries on both sides of the vector equations, you arrive at systems of equations:

$$\begin{array}{cc}\text{Form }(*): & \text{From }(**):\\ \begin{aligned}2x+y-3z+0w&=a\\x+y+0z+4w&=b\\3x+2y-6z+w&=c\\0x+2y-5z+5w&=d\end{aligned} & \begin{aligned}2x+y-3z+0w&=0\\x+y+0z+4w&=0\\3x+2y-6z+w&=0\\0x+2y-5z+5w&=0\end{aligned}\end{array}$$



```
[A]
     [[2 1 -3 0]
      [1 1 0  4]
      [3 2 -6 1]
      [0 2 -5 5]]
rref([A])
     [[1 0 0 -1]
      [0 1 0 5 ]
      [0 0 1 1 ]
      [0 0 0 0 ]]
```

Yes, the two systems differ to the right of the equal signs, but both share a common coefficient matrix, which "twice" reveals the fact that the four given vectors do **not** constitute a basis for $M_{2\times 2}$. Once, by the Spanning Theorem of page 18, and then again by the Linear Independence Theorem of page 22:

$$\begin{bmatrix}2&1&-3&0\\1&1&0&4\\3&2&-6&1\\0&2&-5&5\end{bmatrix}\xrightarrow{\text{rref}}\begin{bmatrix}\mathbf{1}&0&0&-1\\0&\mathbf{1}&0&5\\0&0&\mathbf{1}&1\\0&0&0&0\end{bmatrix}$$

only 3 leading ones and 4 variables
so the four vectors are not linearly independent

contains a zero row
so the four vectors do not span

CHECK YOUR UNDERSTANDING 3.10

Determine if the following set of matrices constitute a basis for the matrix space $M_{2\times 2}$:

$$\left\{\begin{bmatrix}2&1\\3&0\end{bmatrix},\begin{bmatrix}1&1\\2&2\end{bmatrix},\begin{bmatrix}-3&0\\-5&-5\end{bmatrix},\begin{bmatrix}0&4\\1&5\end{bmatrix}\right\}$$

Answer: See page B-10.

In Example 3.10 we found that the vector space $\Re^3$ has a basis consisting of three vectors, and every fiber in your body is probably suggesting that each and every basis for $\Re^3$ must also consist of 3 vectors. Those fibers are correct. Indeed we will prove that if a vector space V has a basis consisting of n vectors, then every basis for V must again contain n vectors. Our proof will make use of the following fundamental result:

In words: There cannot be more lineally independent vectors than the number of vectors in any spanning set.

THEOREM 3.8
SPAN VERSES INDEPENDENT

If $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ spans V, and if $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a linearly independent subset of V, then $n \le m$.

PROOF: Assume that $n > m$.

(We will show that this assumption contradicts the given condition that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a linearly independent set).

We begin by expressing each $\mathbf{v}_i$ as a linear combination of the vectors in the spanning set $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$:

$$\begin{aligned} \mathbf{v}_1 &= c_{11}\mathbf{w}_1 + c_{12}\mathbf{w}_2 + \ldots + c_{1n}\mathbf{w}_m \\ \mathbf{v}_2 &= c_{21}\mathbf{w}_1 + c_{22}\mathbf{w}_2 + \ldots + c_{2n}\mathbf{w}_m \\ \vdots\ & \quad\ \ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \\ \mathbf{v}_n &= c_{n1}\mathbf{w}_1 + c_{n2}\mathbf{w}_2 + \ldots + c_{nm}\mathbf{w}_m \end{aligned} \qquad (*)$$

Now Consider the following homogeneous system of n linear equations in m unknowns, with $n > m$:

Note that the coefficients of the i^{th} equation coincide with those of the i^{th} column of (*).

$$\begin{aligned} c_{11}x_1 + c_{21}x_2 + \ldots + c_{n1}x_m &= 0 \\ c_{12}x_1 + c_{22}x_2 + \ldots + c_{n2}x_m &= 0 \\ \vdots \quad\ \ \vdots \qquad\qquad\quad \vdots \quad &\ \ \ \vdots \\ c_{1n}x_1 + c_{2n}x_2 + \ldots + c_{nm}x_m &= 0 \end{aligned} \qquad (**)$$

The Fundamental Theorem of Homogeneous Systems (page 20), tells us that (**) has a **nontrivial** solution: $(r_1, r_2, \ldots, r_n)$ (not all of the r_is are zero). We now show that $r_1\mathbf{v_1} + r_2\mathbf{v_2} + \ldots + r_n\mathbf{v}_n$ equals $\mathbf{0}$, contradicting the assumption that $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v}_n\}$ is a linearly independent set (for some $r_i \neq 0$):

$$\begin{aligned}
r_1\mathbf{v_1} + r_2\mathbf{v_2} + \ldots + r_n\mathbf{v}_n &\overset{\text{From } (*)}{=} r_1(c_{11}\mathbf{w_1} + c_{12}\mathbf{w_2} + \ldots + c_{1n}\mathbf{w}_m) \\
&\quad + r_2(c_{21}\mathbf{w_1} + c_{22}\mathbf{w_2} + \ldots + c_{2n}\mathbf{w}_m) \\
&\quad \vdots \\
&\quad + r_n(c_{n1}\mathbf{w_1} + c_{n2}\mathbf{w_2} + \ldots + c_{mn}\mathbf{w}_m) \\
\text{Regrouping:} \quad &= (c_{11}r_1 + c_{21}r_2 + \ldots + c_{n1}r_n)\mathbf{w_1} \\
&\quad + (c_{12}r_1 + c_{22}r_2 + \ldots + c_{n2}r_n)\mathbf{w_2} \\
&\quad \vdots \\
&\quad + (c_{1n}r_1 + c_{2n}r_2 + \ldots + c_{mn}r_n)\mathbf{w}_m \\
\text{see margin:} \quad &= 0\mathbf{w_1} + 0\mathbf{w_2} + \ldots + 0\mathbf{w_n} = \mathbf{0}
\end{aligned}$$

Since $(r_1, r_2, \ldots, r_n)$ is a solution of (**):

$$\begin{aligned}
c_{11}r_1 + c_{21}r_2 + \ldots + c_{n1}r_n &= 0 \\
&\vdots \\
c_{1n}r_1 + c_{2n}r_2 + \ldots + c_{mn}r_n &= 0
\end{aligned}$$

We are now in a position to show that all bases of a vector space must contain the same number of elements:

THEOREM 3.9 If $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_n}\}$ and $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ are bases for a vector space V, then $n = m$.

PROOF: Since $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_n}\}$ is a basis, it spans V, and since $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ is a basis, it is a linearly independent set. Applying Theorem 3.8, we have $n \geq m$.

One more time: Since $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_m}\}$ is a basis, it spans V, and since $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_n}\}$ is a basis, it is linearly independent. Applying Theorem 3.8, we also have $m \geq n$.

Since $n \geq m$ and $m \geq n$, $n = m$.

DEFINITION 3.5
DIMENSION

A vector space with basis $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is said to be finite dimensional of **dimension** n. The symbol $\dim(V)$ will be used to denote the dimension of the vector space V.

In the exercises you are asked to show that the polynomial space of Exercise 23 of page 50 is an infinite dimensional space.

The trivial vector space $V = \{\mathbf{0}\}$ of CYU 2.6, page 48, has no basis (see Theorem 3.4, page 89). Nonetheless, it is said to be of **dimension zero**. We also point out that a vector space that is not finite dimensional is said to be **infinite dimensional**.

CHECK YOUR UNDERSTANDING 3.11

Prove that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$ is a basis for a vector space V if and only if every vector in V can uniquely be expressed as a linear combination of the vectors in S.

Answer: See page B-10.

Let V be a space of dimension $\boldsymbol{n}$. The following theorem says that any spanning set of $\boldsymbol{n}$ vectors in V must also be linearly independent, and that any linearly independent set of $\boldsymbol{n}$ vectors must also span V:

So, if the number of vectors **equals** the dimension of the space, then to show that those vectors is a basis you do not have to establish both linear independence and spanning, for **either implies the other**.

THEOREM 3.10 Let V be a vector space of dimension $\boldsymbol{n} > 0$, and let S be a set of $\boldsymbol{n}$ vectors in V. Then:

S spans V **if and only if** S is linearly independent.

PROOF: We first show that the assumption that $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ spans V and is not linearly independent leads to a contradiction:

If S is **not** linearly independent, then some vector in S is a linear combination of the remaining elements in S (Theorem 3.5, page 91). Assume, without loss of generality, that it is the vector $\mathbf{v}_n$:

$$\mathbf{v}_n = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_{n-1}\mathbf{v}_{n-1}$$

Let $v \in V$. Since S spans V:

$$\begin{aligned}\mathbf{v} &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_n\mathbf{v}_n\\ &= c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \ldots + c_{n-1}\mathbf{v}_{n-1} + c_n(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_{n-1}\mathbf{v}_{n-1})\\ &= (c_1 + c_na_1)\mathbf{v}_1 + (c_2 + c_na_2)\mathbf{v}_2 + \ldots + (c_{n-1} + c_na_{n-1})\mathbf{v}_{n-1}\end{aligned}$$

The above argument shows that the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{n-1}\}$ spans V — a contradiction, since a space of dimension n, which necessarily contains a basis of n elements, and therefore a linearly independent set of n elements, cannot contain a spanning set of $n - 1$ vectors (Theorem 3.8).

We now show that the assumption that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a linearly independent set which does not span V will also lead to a contradiction:

Let $\mathbf{v}_{n+1} \notin \text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. The Expansion Theorem of page 90 tells us that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \mathbf{v}_{n+1}\}$ is still a linearly independent set. This leads to a contradiction since a space of dimension n, which necessarily contains a spanning set of n elements, cannot contain a linearly independent set of $n+1$ vectors (Theorem 3.8).

The following result is essentially a restatement of Theorem 3.10. It underlines the fact that you can show that a set of ***n*** vectors in an ***n***-dimensional vector space is a basis by **EITHER** showing that they span the space, **OR** by showing that it is a linearly independent set—**you DON'T have to do both**:

THEOREM 3.11 Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a set of $\boldsymbol{n}$ vectors in a vector space V of dimension $\boldsymbol{n}$. The following are equivalent:

(i) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis.

(ii) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ spans V.

(iii) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly independent.

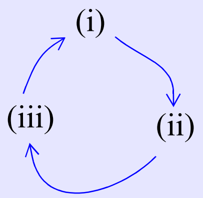
PROOF: We can easily show that $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$:

$(i) \Rightarrow (ii)$: By Definition 3.4.

$(ii) \Rightarrow (iii)$: By Theorem 3.10.

$(iii) \Rightarrow (i)$: By Theorem 3.10 and Definition 3.4.

The cycle:



insures that the validity of any of the three propositions implies that of the other two.

CHECK YOUR UNDERSTANDING 3.12

Prove that the vector space of Example 2.5, page 47: $V = \{(x, y) | x, y \in \Re\}$, under the operations:

$$(x, y) + (x', y') = (x + x' - 1, y + y' + 1)$$
$$r(x, y) = (rx - r + 1, ry + r - 1)$$

has dimension 2.

Answer: See page B-10.

STRETCHING OR SHRINKING TO A BASIS

A basis has to be both a spanning and a linearly independent set of vectors. Help is on the way for any set of vectors that falls short on either of those two counts:

THEOREM 3.12 Let V be a nonzero space of dimension n.

Expansion Theorem: (a) Any linearly independent set of vectors in V can be extended to a basis for V.

Reduction Theorem: (b) Any spanning set of vectors in V can be reduced to a basis for V.

Procedure: Keep adding vectors, while maintaining linear independence, till you end up with n linearly independent vectors.

PROOF: (a) Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be a linearly independent set of vectors in V. If it spans V, then it is already a basis for V and we are done. If not, then take any vector $v_{m+1} \notin \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ and add it to the set: $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m, \boldsymbol{v}_{m+1}\}$. This brings us to a larger set of vectors which, by the Expansion Theorem of page 90, is again linearly independent. Continue this process until the set contains n linearly independent vectors: $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m, \ldots, \boldsymbol{v}_n\}$, and then apply Theorem 3.11 to conclude that it is a basis for V.

Procedure: Keep throwing vectors away, while maintaining spanning, till you end up with n spanning vectors.

(b) Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be a spanning set of vectors in V. Since V contains a linearly independent set of n elements (any basis will do), $m \geq n$ (Theorem 3.8). If $m = n$ then we are done (Theorem 3.11). If $m > n$, then the spanning set $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ cannot be linearly independent; for if it were, then it would be a basis, and all bases have n elements. Find a vector in $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ that is a linear combination of the rest and remove it from that set to arrive at a smaller spanning set. Continue this "tossing out" process until you arrive at a spanning set consisting of n elements—a basis for V (Theorem 3.11).

EXAMPLE 3.12 (a) Expand the linearly independent set:

$$L = \{(\mathbf{1, 2, 3, 4}), (\mathbf{-4, 1, 0, 1}), (\mathbf{3, 3, 1, 2})\}$$

to a basis for $\Re^4$.

(b) Reduce the spanning set:

$$S = \{x+1, 2x^2-3, 2x-3, x^2+4, x^2-x-3\}$$

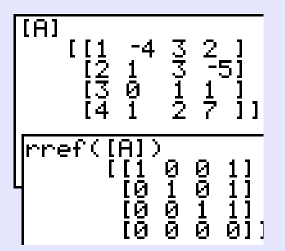
to a basis for P_2.

SOLUTION:

(a) We are given that L is linearly independent, and need to complement it with an additional 4-tuple, while maintaining linear independence. From earlier discussions, you know that if you randomly pick a 4-tuple, then the probability that it will be in Span(L) is nil. Let's go with $(\mathbf{2, -5, 1, 7})$:

$$\beta = \{(\mathbf{1, 2, 3, 4}), (\mathbf{-4, 1, 0, 1}), (\mathbf{3, 3, 1, 2}), (\mathbf{2, -5, 1, 7})\}$$

We, of course, have to demonstrate that the "gods were not against us," and do so via the Linearly Independence Theorem for $\Re^n$ of page 88:



```
[A]
    [[1  -4 3  2 ]
     [2  1  3  -5]
     [3  0  1  1 ]
     [4  1  2  7 ]]
rref([A])
         [[1  0  0  1]
          [0  1  0  1]
          [0  0  1  1]
          [0  0  0  0]]
```

$$\begin{bmatrix} 1 & -4 & 3 & 2 \\ 2 & 1 & 3 & -5 \\ 3 & 0 & 1 & 1 \\ 4 & 1 & 2 & 7 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} \mathbf{1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \end{bmatrix}$$

The above shows that β is a linearly independent set. Since it consists of 4 vectors, and since $\Re^4$ is of dimension 4 [Exercise 1(c)], β is a basis for $\Re^4$.

(b) It is easy to see that $\{1, x, x^2\}$ is a basis for the vector space $P_2(x)$, and that consequently P_2 is of dimension 3.

Since $S = \{x+1, 2x^2-3, 2x-3, x^2+4, x^2-x-3\}$ contains five vectors, we have to throw two of them away in such a manner so as to end up with three vectors that still span P_2; or, equivalently, with three linearly independent vectors. We leave it for you to verify that $\{x+1, 2x^2-3, 2x-3\}$ is a linearly independent set. As such, it must be a basis for the three dimensional space P_2.

CHECK YOUR UNDERSTANDING 3.13

(a) Expand the linearly independent set $L = \left\{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\right\}$ to a basis for $M_{2 \times 2}$

(b) Reduce the set:
$S = \{(3, -1, 2), (-9, 3, -6), (1, 2, -2), (-5, 4, -6), (6, -2, 4)\}$
to a basis for $\text{Span}(S)$. Dose S span $\Re^3$?

Answer: See page B-11.

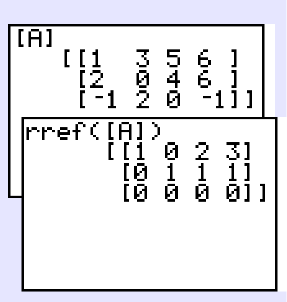
The next example reveals a systematic approach that can be used to reduce a set of vectors S in $\Re^n$ to a basis for $\text{Span}(S)$:

EXAMPLE 3.13 Find a subset of:
$S = \{(\mathbf{1}, \mathbf{2}, -\mathbf{1}), (\mathbf{3}, \mathbf{0}, \mathbf{2}), (\mathbf{5}, \mathbf{4}, \mathbf{0}), (\mathbf{6}, \mathbf{6}, -\mathbf{1})\}$
which is a basis for $\text{Span}(S)$.

SOLUTION: To determine which of the four vectors can be discarded, we challenge their linear independence, and turn to the equation:

$$a(\mathbf{1}, \mathbf{2}, -\mathbf{1}) + b(\mathbf{3}, \mathbf{0}, \mathbf{2}) + c(\mathbf{5}, \mathbf{4}, \mathbf{0}) + d(\mathbf{6}, \mathbf{6}, -\mathbf{1}) = (\mathbf{0}, \mathbf{0}, \mathbf{0}) \quad (*)$$

Equating coefficients leads us to the following homogeneous system of equations:



```
[A]
   [[1  3 5 6 ]
    [2  0 4 6 ]
    [-1 2 0 -1]]
rref([A])
      [[1 0 2 3]
       [0 1 1 1]
       [0 0 0 0]]
```

$$\begin{tikzcd}
\mathbf{S}: \left.\begin{aligned} a+3b+5c+6d &= 0\\ 2a+0b+4c+6d &= 0\\ -a+2b+0c-d &= 0\end{aligned}\right\} \arrow[r, "\mathbf{coef}(S)"] \arrow[d, leftrightarrow, "\text{equivalent systems}"] & \begin{array}{c} \begin{matrix} a & b & c & d \end{matrix} \\ \begin{bmatrix} 1 & 3 & 5 & 6 \\ 2 & 0 & 4 & 6 \\ -1 & 2 & 0 & -1 \end{bmatrix} \end{array} \arrow[r, "\mathbf{rref}"] & \begin{array}{c} \begin{matrix} a & b & c & d \end{matrix} \\ \begin{bmatrix} \mathbf{1} & 0 & 2 & 3 \\ 0 & \mathbf{1} & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \arrow[dll] \\
\left.\begin{aligned} a+2c+3d &= 0\\ b+c+d &= 0\end{aligned}\right\} & &
\end{tikzcd}$$

Note that these are the four given vectors, but written in **column** form

Figure 3.2

Let's agree to call a vector in

$$S = \{(\mathbf{1}, \mathbf{2}, -\mathbf{1}), (\mathbf{3}, \mathbf{0}, \mathbf{2}), (\mathbf{5}, \mathbf{4}, \mathbf{0}), (\mathbf{6}, \mathbf{6}, -\mathbf{1})\}$$

that occupies the same column-location in coef(S) as that of a leading-one-column in the rref-matrix, a **leading-one** vector [$(\mathbf{1}, \mathbf{2}, -\mathbf{1})$ and $(\mathbf{3}, \mathbf{0}, \mathbf{2})$ are leading-one vectors]. We now proceed to show that those leading-one vectors constitute a basis for $\text{Span}(S)$.

Figure 3.2 tells us that system $\boldsymbol{S}$ will be satisfied for **any** a, b, c, and d for which: $\left.\begin{aligned} a+2c+3d &= 0\\ b+c+d &= 0\end{aligned}\right\}$ (**)

Note that c and d are the free variables in rref[coef (s)]

Setting $c = \mathbf{1}$ and $d = \mathbf{0}$ in (**) leads us to $\left.\begin{aligned} a+2 &= 0\\ b+1 &= 0\end{aligned}\right\}$, with solution $a = -2$ and $b = -1$. Substituting these values in (*) we have:

$$-2(\mathbf{1}, \mathbf{2}, -\mathbf{1}) - 1(\mathbf{3}, \mathbf{0}, \mathbf{2}) + \mathbf{1}(\mathbf{5}, \mathbf{4}, \mathbf{0}) + \mathbf{0}(\mathbf{6}, \mathbf{6}, -\mathbf{1}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$$

$$\text{or:}\quad (\mathbf{5}, \mathbf{4}, \mathbf{0}) = 2(\mathbf{1}, \mathbf{2}, -\mathbf{1}) + 1(\mathbf{3}, \mathbf{0}, \mathbf{2})$$

Conclusion: the non-leading-one vector **(5, 4, 0)** can be expressed as a linear combination of the two leading-one vectors.

Setting $d = 1$ and $c = 0$ in (**) we find that $a = -3$ and $b = -1$, bringing us to:

$$-3(\mathbf{1}, \mathbf{2}, -\mathbf{1}) - 1(\mathbf{3}, \mathbf{0}, \mathbf{2}) + 0(\mathbf{5}, \mathbf{4}, \mathbf{0}) + 1(\mathbf{6}, \mathbf{6}, -\mathbf{1}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$$

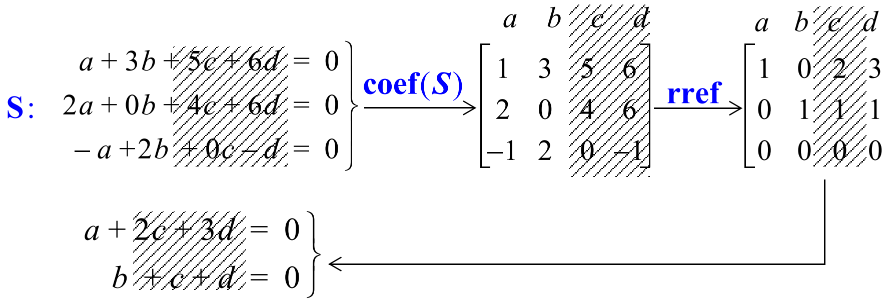
$$\text{or:}\quad (\mathbf{6}, \mathbf{6}, -\mathbf{1}) = 3(\mathbf{1}, \mathbf{2}, -\mathbf{1}) + 1(\mathbf{3}, \mathbf{0}, \mathbf{2})$$

Conclusion: the non-leading-one vector **(6, 6, -1)** can be expressed as a linear combination of the two leading-one vectors.

Since $(\mathbf{5}, \mathbf{4}, \mathbf{0})$ and $(\mathbf{6}, \mathbf{6}, -\mathbf{1})$ are linear combinations of $(\mathbf{1}, \mathbf{2}, -\mathbf{1}), (\mathbf{3}, \mathbf{0}, \mathbf{2})$:

$$\text{Span}\{(\mathbf{1}, \mathbf{2}, -\mathbf{1}), (\mathbf{3}, \mathbf{0}, \mathbf{2}), (\mathbf{5}, \mathbf{4}, \mathbf{0}), (\mathbf{6}, \mathbf{6}, -\mathbf{1})\} = \text{Span}\{(\mathbf{1}, \mathbf{2}, -\mathbf{1}), (\mathbf{3}, \mathbf{0}, \mathbf{2})\}$$

Covering up the two non-leading-one columns in the development of Figure 3.2:



we see that the only solution of: $\boldsymbol{a}(\mathbf{1, 2, -1}) + b(\mathbf{3, 0, 2}) = (\mathbf{0, 0, 0})$ is the trivial solution, and that the vectors $(\mathbf{1, 2, -1}), (\mathbf{3, 0, 2})$ are therefore linearly independent.

SUMMARIZING

To find a basis for the space spanned by the vectors $\boldsymbol{v}_1 = (\mathbf{1, 2, -1})$, $\boldsymbol{v}_2 = (\mathbf{3, 0, 2})$, $\boldsymbol{v}_3 = (\mathbf{5, 4, 0})$, and $\boldsymbol{v}_4 = (\mathbf{6, 6, -1})$, we constructed the 3×4 matrix $A = \begin{bmatrix} 1 & 3 & 5 & 6 \\ 2 & 0 & 4 & 6 \\ -1 & 2 & 0 & -1 \end{bmatrix}$ with columns the given four-tuples. We then showed that the leading-one vectors, namely $\boldsymbol{v}_1 = (\mathbf{1, 2, -1})$ and $\boldsymbol{v}_2 = (\mathbf{3, 0, 2})$, constituted a basis for the space spanned by the four given vectors.

We state, without proof, a generalization of the above observation:

THEOREM 3.13 Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$ be vectors in $\Re^n$. If A is the $n \times m$ matrix whose i^{th} column is the n-tuple $\boldsymbol{v}_i$, then the set consisting of those vectors $\boldsymbol{v}_i$, where the i^{th} column of $\text{rref}(A)$ contains a leading one, is a basis for the space spanned by $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$.

Answer: See page B-11.

CHECK YOUR UNDERSTANDING 3.14

Use the above theorem to address the problem of CYU 3.13(b).

‘

EXERCISES

1. (a) Prove that $S_2 = \{\boldsymbol{e}_1, \boldsymbol{e}_2\} = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$ is a basis for $\Re^2$. Express $(\mathbf{-3}, \mathbf{\tfrac{5}{2}})$ as a linear combination of the vectors in S_2.

   (b) Prove that $S_3 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\} = \{(\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{0}, \mathbf{1})\}$ is a basis for $\Re^3$. Express $(\mathbf{3}, \sqrt{\mathbf{2}}, \mathbf{0})$ as a linear combination of the vectors in S_3.

   (c) Prove that $S_n = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ is a basis for $\Re^n$.

2. (a) Prove that $\beta = \{(\mathbf{2}, \mathbf{0}, \mathbf{5}), (\mathbf{0}, \mathbf{1}, \mathbf{10}), (\mathbf{1}, \mathbf{2}, \mathbf{0})\}$ is a basis for $\Re^3$, and express $(\mathbf{7}, \mathbf{0}, \mathbf{-5})$ as a linear combinations of the vectors in β.

   (b) Show that $S = \{(\mathbf{2}, \mathbf{0}, \mathbf{5}), (\mathbf{0}, \mathbf{1}, \mathbf{10}), (\mathbf{1}, \mathbf{2}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$ is not a basis for $\Re^3$, and find two different representations for the vector $(\mathbf{7}, \mathbf{0}, \mathbf{-5})$ as a linear combination of the vectors in S.

3. (a) Prove that $\beta = \{2x^2 + 3, x^2 - x, x - 5\}$ is a basis for P_2, and express $x^2 + 3x - 1$ as a linear combinations of the vectors in β.

   (b) Show that $S = \{2x^2, 3x, 5, x - 4\}$ is not a basis for P_2, and find two different representations for the vector $x^2 + 3x - 1$ as a linear combination of the vectors in S.

Exercises 4-7. Determine if the given set of vectors is a basis for $\Re^3$.

4. $\{(\mathbf{2}, \mathbf{1}, \mathbf{5}), (\mathbf{4}, \mathbf{1}, \mathbf{10}), (\mathbf{1}, \mathbf{2}, \mathbf{3})\}$

5. $\{(\mathbf{0}, \mathbf{0}, \mathbf{0}), (\mathbf{-1}, \mathbf{2}, \mathbf{-5}), (\mathbf{2}, \mathbf{-1}, \mathbf{-5})\}$

6. $\{(\mathbf{2}, \mathbf{1}, \mathbf{5}), (\mathbf{4}, \mathbf{1}, \mathbf{10}), (\mathbf{4}, \mathbf{3}, \mathbf{10})\}$

7. $\{(\sqrt{\mathbf{2}}, \boldsymbol{\pi}, \boldsymbol{e}), (\boldsymbol{\pi}, \boldsymbol{e}, \sqrt{\mathbf{2}}), (\boldsymbol{e}, \sqrt{\mathbf{2}}, \boldsymbol{\pi})\}$

8. (a) Prove that the matrix space $M_{2 \times 2}$ has dimension 4.

   (b) Prove that the matrix space $M_{m \times n}$ has dimension $m \times n$.

Exercises 9-12. Determine if the given set of vectors is a basis for $M_{2 \times 2}$.

9. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & 3 \\ 4 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 4 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 4 & 1 \\ 2 & 3 \end{bmatrix}\right\}$

10. $\left\{\begin{bmatrix} -1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}\right\}$

11. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 0 & 4 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 3 & 4 \end{bmatrix}\right\}$

12. $\left\{\begin{bmatrix} 2 & \frac{1}{2} \\ 3 & -1 \end{bmatrix}, \begin{bmatrix} 13 & -11 \\ \frac{13}{2} & 4 \end{bmatrix}, \begin{bmatrix} \frac{1}{3} & -1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ \frac{1}{6} & 0 \end{bmatrix}\right\}$

Exercises 13-15. Determine if the given set of vectors is a basis for $M_{3 \times 2}$.

13. $\begin{bmatrix} 0 & 1 \\ 3 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 3 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 3 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 3 & 0 \\ 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

14. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$

15. $\begin{bmatrix} 1 & 2 \\ 3 & -1 \\ -2 & -3 \end{bmatrix}, \begin{bmatrix} 2 & -2 \\ -2 & -1 \\ 3 & -4 \end{bmatrix}, \begin{bmatrix} 3 & 2 \\ 1 & -2 \\ -3 & -1 \end{bmatrix}, \begin{bmatrix} -1 & 2 \\ 3 & 1 \\ -2 & 3 \end{bmatrix}, \begin{bmatrix} -1 & -2 \\ 3 & 1 \\ 2 & -3 \end{bmatrix}$

16. (a) Prove that the polynomial space $P_3 = \{a_3x^3 + a_2x^2 + a_1x + a_0 | a_i \in R, 0 \le i \le 3\}$ is of dimension 4.

(b) Prove that the polynomial space:

$$P_n = \{a_nx^n + a_{n-1}x^{n-1} + \ldots + a_1x + a_0 | a_i \in R, 0 \le i \le n\}$$

is of dimension $n + 1$.

Exercises 17-20. Determine if the given set of vectors is a basis for P_3.

17. $\{x^3 + 1, 2x^3 - x^2 + 2, 3x + 1, 4x^3 - x^2 + 9x + 7\}$

18. $\{x^3 + 1, 2x^3 - x^2 + 2, 3x + 1, 4x^3 - x^2 + 9x + 8\}$

19. $\{2x^3 + 3x^2 + x - 1, -x^3 - 9x^2 + 2x + 2, x^3 - x^2 + 2x + 2, -3x^3 + 2x^2 - x + 1\}$

20. $\{2x^3 + 3x^2 + x - 1, -x^3 - x^2 + 2x + 2, x^3 - x^2 + 2x + 2, -3x^3 + 2x^2 - x + 1\}$

Exercises 21-24. Extend the given linearly independent set of vectors to a basis for $\Re^4$.

21. $\{(\mathbf{1, 3, 4, 1}), (\mathbf{-1, 2, 0, 1}), (\mathbf{1, 1, 2, 2})\}$

22. $\{(\mathbf{1, 2, 1, -2}), (\mathbf{1, 2, 1, 2}), (\mathbf{1, 2, 3, 4})\}$

23. $\{(\mathbf{2, 1, 3, -1}), (\mathbf{1, 3, 0, 2})\}$

24. $\{(\mathbf{2, 2, 1, 1})\}$

Exercises 25-28. Extend the given linearly independent set of vectors to a basis for P_4.

25. $\{x^4 + 1, x^3 + 2x^2, 3x^3 + x - 5, 9\}$

26. $\{3x - 2, 3x^2 - 2, 3x^3 - 2, 3x^4 - 2\}$

27. $\{x^4 + x^3 + x^2, x^3 + x^2 + x, x^2 + x + 1\}$

28. $\{2x^4 + x^3 + x^2 + x + 1, x^4 - 5\}$

Exercises 29-30. Does the given set of vectors span $M_{2 \times 2}$? If so, reduce the set to obtain a basis for $M_{2 \times 2}$.

29. $\left\{ \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 3 & 3 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right\}$

30. $\left\{ \begin{bmatrix} -1 & 0 \\ 2 & 0 \end{bmatrix}, \begin{bmatrix} 5 & 0 \\ -10 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$

Exercises 31-32. Does the given set of vectors span P_3? If so, reduce the set to obtain a basis for P_3.

31. $\{2x, 3x^2, 6x^2 + 4x, x^3 + 1, 3x^2 + x, x^3 + 3x^2 + x + 1\}$

32. $\{2x^3 + x - 1, 4x^3 - 2x^2 - 6, 4x^2 + 2, 5x^3 + 6x^2 - 2, x^3 + x - 1, x^3 - 2x^2, -x^3 - 2x^2\}$

Exercises 33-37. Use Theorem 3.13 to find a subset of the given set of vectors S in $\Re^n$ which is a basis for $\text{Span}(S)$. If necessary expand that basis for $\text{Span}(S)$ to a basis for the corresponding Euclidean space.

33. $S = \{(\mathbf{2, 1, 4}), (\mathbf{-1, 3, 2}), (\mathbf{5, -1, 6}), (\mathbf{4, 2, 8})\}$

34. $S = \{(\mathbf{1, 1, 3}), (\mathbf{-1, 3, 2}), (\mathbf{1, 5, 8}), (\mathbf{3, 2, -1})\}$

35. $S = \{(\mathbf{1, 1, 3}), (\mathbf{-1, 3, 2}), (\mathbf{1, 5, 8}), (\mathbf{3, 2, -1})\}$

36. $S = \{(\mathbf{2, 1, 3, 0}), (\mathbf{0, 4, 4, 2}), (\mathbf{-2, 3, 1, 2}), (\mathbf{1, 1, 5, 8}), (\mathbf{3, 2, 8, 8})\}$

37. $S = \{(\mathbf{1, 3, 1, 3, 2}), (\mathbf{2, 4, 1, 4, 2}), (\mathbf{1, 1, 0, 1, 0}), (\mathbf{1, 1, 2, 0, 2}), (\mathbf{2, 2, 1, 1, 1}), (\mathbf{1, 2, 3, 4, 5})\}$

Exercises 38-39. Find a subset of the given set of vectors S which is a basis for $\text{Span}(S)$. If necessary, expand that basis for $\text{Span}(S)$ to a basis for P_4.

38. $S = \{2x^4 + x^3 + 3x + 1, -x^4 - x^3 + x^2, x^3 - 2x^3 + 2x^2 + 3x, -2x^4 - 2x^3 + 2x^2 - 1\}$

39. $S = \{5, -x^3 - x, x^4 + x^3 + x^2 + x + 1, 2x^4 - 2x^2, 2x^4 + 2x^2\}$

Exercises 40-41. Find a basis for the space spanned by the given set of vectors in the function space vectors $F(\Re)$ of Theorem 2.4, page 44.

40. $\{5, \sin x, \cos x, \sin^2 x, \cos^2 x, \sin(5 + x)\}$

41. $\{\sin x, \cos x, \sin^2 x, \cos^2 x, \cos 2x, \sin 2x\}$

42. Show that $\{(\boldsymbol{a, b, c}) | a + 2b + c = 0\}$ is a subspace of $\Re^3$, and then find a basis for that subspace.

43. Show that $\{(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) | a + b = c - d\}$ is a subspace of $\Re^4$, and then find a basis for that subspace.

44. Show that $\{ax^2 + bx + c | a = b - c\}$ is a subspace of P_2, and then find a basis for that subspace.

45. Show that $\{ax^3 + bx^2 + cx + d | a = b, c = a + b + d\}$ is a subspace of P_3, and then find a basis for that subspace.

46. Find all values of c for which $\{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (\boldsymbol{c}^2, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \boldsymbol{c}, \mathbf{1})\}$ is a basis for $\Re^3$.

47. Find all values of c for which $\left\{ \begin{bmatrix} c & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 2c \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -c & 0 \end{bmatrix}, \begin{bmatrix} c & 0 \\ 0 & c \end{bmatrix} \right\}$ is a basis for $M_{2 \times 2}$.

48. Find a basis for the vector space of Example 2.5, page 47.

49. Suppose $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for a vector space V. For what values of a and b is $\{a\boldsymbol{v}_1, b\boldsymbol{v}_2, (a - b)\boldsymbol{v}_3\}$ a basis for V?

50. Let S is a subspace of V with $\dim(S) = \dim(V) = n$. Prove that $S = V$.

51. Suppose that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s\}$ is a linearly independent set of vectors in a space V of dimension n, and that $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots \boldsymbol{w}_t\}$ spans V. Prove that $s \le n \le t$.

52. A set of vectors S in a finite dimensional vector space V is said to be a **maximal linearly independent set** if it is not a proper subset of any linearly independent set. Prove that a set of vectors is a basis for V if and only if it is a maximal linearly independent set.

53. A set of vectors S in a finite dimensional vector space V is said to be a **minimal spanning set** if no proper subset of S spans V. Prove that a set of vectors is a basis for V if and only if it is a minimal spanning set.

54. Let H and K be finite dimensional subspace of a vector space V with $H \cap K = \{\mathbf{0}\}$, and let $H + K = \{\boldsymbol{h} + \boldsymbol{k} | \boldsymbol{h} \in H \text{ and } \boldsymbol{k} \in K\}$. Prove that $\dim(H + K) = \dim(H) + \dim(K)$. (Note: you were asked to show that $H + K$ is a subspace of V in Exercise 42, page 67.)

55. Let H and K be finite dimensional subspace of a vector space V, and let $H + K = \{\boldsymbol{h} + \boldsymbol{k} | \boldsymbol{h} \in H \text{ and } \boldsymbol{k} \in K\}$. Prove that:

$$\dim(H + K) = \dim(H) + \dim(K) - \dim(H \cap K)$$

Suggestion: Start with a basis for $H \cap K$ and extend it to a basis for $H + K$.

56. Prove that the polynomial space P of Exercise 22, page 50, is not finite dimensional by showing that it does not have a finite base.

57. **(Calculus dependent)** Show that $S = \{p \in P | p'(0) = 0\}$ is a subspace of the polynomial space P of Exercise 22, page 50. Find a basis for S.

58. Prove that a vector space V is infinite dimensional (not finite dimensional) if and only if for any positive integer n, there exists a set of n linearly independent vectors in V.

PROVE OR GIVE A COUNTEREXAMPLE

59. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for a vector space V, and if c_1, c_2, and c_3 are nonzero scalars, then $\{c_1\boldsymbol{v}_1, c_2\boldsymbol{v}_2, c_3\boldsymbol{v}_3\}$ is also a basis for V.

60. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}\}$ is a linearly independent set of vectors in a space V of dimension n, and if $\boldsymbol{v}_n \notin \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}\}$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n\}$ is a basis for V.

61. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}\}$ is a linearly independent set of vectors in a space V of dimension n, and if $\boldsymbol{v}_n \notin \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}\}$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_{n-1}, \boldsymbol{v}_n\}$ is a basis for V.

62. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}\}$ is a spanning set of vectors in a space V of dimension n, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for V.

63. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1}\}$ is a spanning set of vectors in a space V of dimension n, and if $\boldsymbol{v}_{n+1} \notin \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for V.

64. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ is a basis for a vector space V, then $\{\boldsymbol{v}_1, \boldsymbol{v}_1 + \boldsymbol{v}_2, \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3\}$ is also a basis for V.

65. It is possible to have a basis for the polynomial space $P_2(x)$ which consists entirely of polynomials of degree 2.

66. Let $\{\boldsymbol{v}_\mathbf{1}, \boldsymbol{v}_\mathbf{2}, \ldots, \boldsymbol{v}_{\boldsymbol{n}+\mathbf{1}}\}$ be a spanning set for a space V of dimension n satisfying the property that $\boldsymbol{v}_\mathbf{1} + \boldsymbol{v}_\mathbf{2} + \ldots + \boldsymbol{v}_{\boldsymbol{n}+\mathbf{1}} = \mathbf{0}$. If you delete any vector from the set $\{\boldsymbol{v}_\mathbf{1}, \boldsymbol{v}_\mathbf{2}, \ldots, \boldsymbol{v}_{\boldsymbol{n}+\mathbf{1}}\}$, then the resulting set of n vectors will be a basis for V.
    Note: In set notation, an element cannot be repeated. In particular, no two of the $\boldsymbol{v}$s in $\{\boldsymbol{v}_\mathbf{1}, \boldsymbol{v}_\mathbf{2}, \ldots, \boldsymbol{v}_{\boldsymbol{n}+\mathbf{1}}\}$ are the same.

67. If V is a space of dimension n, then V contains a subspace of dimension m for every integer $0 \le m \le n$.

CHAPTER SUMMARY

LINEAR COMBINATION	A vector $\mathbf{v}$ in a vector space V is said to be a **linear combination** of vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ in V, if there exists scalars $c_1, c_2, \ldots, c_n$ such that: $\mathbf{v} = c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}$
SPANNING	The set of linear combinations of $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a subspace of V. It is said to be the **subspace of V spanned** by those vectors, and is denoted by $\text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$: $\text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\} = \{c_1\mathbf{v_1} + c_2\mathbf{v_2} + \ldots + c_n\mathbf{v_n}\}$ If $V = \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$, then $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is said to **span** the vector space V.
If every vector in a set S_1 is contained in the space spanned by another set S_2, then $\text{Span}(S_1)$ *is a subset of* $\text{Span}(S_2)$.	Let the set of vectors $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ and $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ be such that $\mathbf{w_i} \in \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ for $1 \le i \le m$. Then: $\text{Span}\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\} \subseteq \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$
	If $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\} \subseteq \text{Span}\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$ and $\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\} \subseteq \text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ then: $\text{Span}\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\} = \text{Span}\{\mathbf{w_1}, \mathbf{w_2}, \ldots, \mathbf{w_m}\}$
LINEARLY INDEPENDENT SET	The vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are **linearly independent** if: $a_1\mathbf{v_1} + \cdots + a_n\mathbf{v_n} = \mathbf{0} \Rightarrow a_i = 0, 1 \le i \le n$ If the vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are not linearly independent then they are said to be **linearly dependent**.
Unique representation.	The vectors $\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}$ are linearly independent **if and only if** $a_1\mathbf{v_1} + \ldots + a_n\mathbf{v_n} = b_1\mathbf{v_1} + \ldots + b_n\mathbf{v_n}$ implies that $a_i = b_i$, for $1 \le i \le n$.
No vector can be built from the rest.	A collection of **two or more** vectors is linearly independent **if and only if** none of those vectors can be expressed as a linear combination of the rest.
Expanding a linearly independent set.	Let $S = \{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ be a linearly independent set. If $\mathbf{v} \notin \text{Span}(S)$, then $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}, \mathbf{v}\}$ is again a linearly independent set.

Linear Independence Theorem.	A homogeneous system of m linear equations in n unknowns **S** has only the trivial solution **if and only if** $\text{rref}[\text{coef}(S)]$ has n leading ones.
Linear independence in $\boldsymbol{R}^n$.	A set of vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ in $\Re^n$ is linearly independent **if and only if** the row-reduced-echelon form of the $n \times m$ matrix whose i^{th} columns is the (vertical) n-tuple $\boldsymbol{v}_i$ has m leading ones.
BASIS	A set of vectors $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ in a vector space V is said to be a **basis for** $\boldsymbol{V}$ if: (1) β spans V and: (2) β is linearly independent.
A spanning set can't have fewer elements than the number of elements in ***any*** *linearly independent set.*	If $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ spans V, and if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a linearly independent subset of V, then $n \le m$.
All bases for a vector space contain the same number of vectors.	If $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ are bases for a vector space V, then $n = m$.
You can show that a set of ***n*** *vectors in an* ***n****-dimensional vector space is a basis by* ***EITHER*** *showing that they span the space,* ***OR*** *by showing that it is a linearly independent set—****you DON'T have to do both****:*	Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a set of n vectors in a vector space V of dimension n. The following are equivalent: (i) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis. (ii) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ spans V. (iii) $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is linearly independent.
Expansion Theorem	Any linearly independent set of vectors in V can be extended to a basis for V.
Reduction Theorem	Any spanning set of vectors in V can be reduced to a basis for V.
Reducing a set of vectors S in $\Re^n$ *to a basis for* Span(S)	Let $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$ be vectors in $\Re^n$. If A is the $n \times m$ matrix whose i^{th} column is the n-tuple $\boldsymbol{v}_i$, then the set consisting of those vectors $\boldsymbol{v}_i$, where the i^{th} column of $\text{rref}(A)$ contains a leading one, is a basis for the space spanned by $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m$.

CHAPTER 4
LINEARITY

In this chapter we turn our attention to functions from one vector space to another which, in a sense, preserve the algebraic structure of those spaces. Such functions, called linear transformations, are introduced in Section 1. As you will see, each linear transformation $T: V \to W$ gives rise to two important subspace; one, called the kernel of *T*, resides in the vector space *V*, while the other, called the image of *T*, lives in *W*. Those two subspaces are investigated in Section 2. Linear transformations which are also one-to-one and onto are called isomorphisms, and they are considered in Section 3.

§1. LINEAR TRANSFORMATIONS

Up to now we have focused our attention exclusively on the internal nature of a given vector space. The time has come to consider functions from one vector space and another:

A linear transformation is also called a **linear function,** or a **linear map**. A linear map $T: V \to V$ from a vector space to itself is said to be a **linear operator.**

DEFINITION 4.1
LINEAR TRANSFORMATION

A function $T: V \to W$ from a vector space *V* to a vector space *W* is said to be a **linear transformation** if for every $\boldsymbol{v}, \boldsymbol{v}' \in V$ and $r \in \Re$:

$$(1): \quad T(\boldsymbol{v} + \boldsymbol{v}') = T(\boldsymbol{v}) + T(\boldsymbol{v}')$$
$$\text{and } (2): \quad T(r\boldsymbol{v}) = rT(\boldsymbol{v})$$

Let's focus a bit on the statement:

$$T(\boldsymbol{v} + \boldsymbol{v}') = T(\boldsymbol{v}) + T(\boldsymbol{v}')$$

for $T: V \to W$. There are two plus signs in the equation, but the one on the left is happening in the space *V*, while the one on the right takes place in the space *W*. What the statement is saying is that you can perform the sum in the vector space *V* and then carry the result over to the vector space *W* via the transformation *T*, or you can first carry those two vectors $\boldsymbol{v}$ and $\boldsymbol{v}'$ over to *W* (via *T*) and then perform their sum in the space *W*. In the same fashion, $T(r\boldsymbol{v}) = rT(\boldsymbol{v})$ is saying that you can perform scalar multiplication in *V* and then carry the result over to *W*, or you can first carry the vector $\boldsymbol{v}$ to *W* and then perform the scalar multiplication in that vector space.

EXAMPLE 4.1 Show that the function $T: \Re^3 \to \Re^2$ given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{2a} + \boldsymbol{b}, -\boldsymbol{c})$ is linear.

SOLUTION: Let $(a, b, c), (a', b', c') \in \Re^3$ and $r \in \Re$.

A smoother approach:

$$\begin{aligned} &T[(a,b,c)+(a',b',c')] \\ &= T(a+a', b+b', c+c') \\ &= [2(a+a')+(b+b'), -(c+c')] \\ &= (2a+2a'+b+b', -c-c') \\ &= (2a+b,-c)+(2a'+b',-c') \\ &= T(a,b,c)+T(a',b',c') \end{aligned}$$

T preserves sums:

$$\begin{aligned} T[(a,b,c)+(a',b',c')] &= T(a+a', b+b', c+c') \\ T(a,b,c) = (2a+b,-c){:} \quad &= [2(a+a')+(b+b'), -(c+c')] \\ &= (2a+2a'+b+b', -c-c') \leftarrow \\ \text{and:} \quad T(a,b,c)+T(a',b',c') &= (2a+b,-c)+(2a'+b',-c') \\ &= (2a+b+2a'+b', -c-c') \leftarrow \quad \text{same} \end{aligned}$$

T preserves scalar products:

$$\begin{aligned} T[r(a,b,c)] = T(ra, rb, rc) &= (2ra+rb, -rc) \\ &= r(2a+b,-c) = rT(a,b,c) \end{aligned}$$

EXAMPLE 4.2

Determine if the function $f\colon \Re^3 \to \Re^3$ given by $f(a,b,c) = (2c, b^2, -a)$ is linear.

SOLUTION: If you are undecided on whether or not f is linear, you may want to challenge its linearity with a couple of specific voters

$$\begin{aligned} f[(2,1,3)+(4,0,1)] &\overset{?}{=} f(2,1,3)+f(4,0,1) \\ f(6,1,4) &\overset{?}{=} f(2,1,3)+f(4,0,1) \\ f(a,b,c) = (2c, b^2, -a){:} \quad (8,1,-6) &\overset{?}{=} (6,1,-2)+(2,0,-4) \\ (8,1,-6) &= (8,1,-6) \quad \text{Yes!} \end{aligned}$$

The above establishes nothing. It certainly does not show that the function f is not linear, nor does it establish its linearity since we have but demonstrated that it "works" for the two chosen vectors $(2,1,3)$ and $(4,0,1)$. Let's try another pair:

$$\begin{aligned} f[(3,-2,5)+(4,9,-3)] &\overset{?}{=} f(3,-2,5)+f(4,9,-3) \\ f(7,7,2) &\overset{?}{=} f(3,-2,5)+f(4,9,-3) \\ (4,49,-7) &\overset{?}{=} (10,4,-3)+(-6,81,-4) \\ (4,49,-7) &\overset{?}{=} (4,85,-7) \quad \textbf{No!} \end{aligned}$$

Since $f[(3,-2,5)+(4,9,-3)] \neq f(3,-2,5)+f(4,9,-3)$, the function is not linear.

You can also show that the above function is not linear by demonstrating, for example, that $f[2(1,2,3)] \neq 2f(1,2,3)$. To show that a function is not linear you need only come up with a specific counterexample which "shoots down" **EITHER** (1) or (2) of Definition 4.1.

CHECK YOUR UNDERSTANDING 4.1

Is the function $f\colon \Re^2 \to \Re^3$ given by $f(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a}+\boldsymbol{b}, 2\boldsymbol{b}, \boldsymbol{a}-\boldsymbol{b})$ linear? Justify your answer.

Answer: See page B-12.

Given that a linear transformation preserves vector space structures, it may come as no surprise to find that it maps zeros to zeros, and inverses to inverses:

In order to distinguish where the different zeros preside, we are using $\mathbf{0}_V$ and $\mathbf{0}_W$ to indicate the zero is in the vector space V and W, respectively.

THEOREM 4.1 If $T\colon V \to W$ is linear, then:

(a) $$T(\underset{\uparrow\ \text{zero in the space } V}{\mathbf{0}_V}) = \overset{\text{zero in the space } W\ \downarrow}{\mathbf{0}_W}$$

(b) $$T(\underset{\uparrow\ \text{inverse of } \boldsymbol{v} \text{ in the space } V}{-\boldsymbol{v}}) = \overset{\text{inverse of the vector } T(\boldsymbol{v}) \text{ in the space W}\ \downarrow}{\boxed{-T(\boldsymbol{v})}}$$

PROOF: (a) Note that three different zeros are featured below:

$$T(\mathbf{0}_V) \underset{\uparrow\ \text{Theorem 2.7, page 53}}{=} T(0\cdot\mathbf{0}_V) \overset{\text{by linearity}\ \downarrow}{=} 0T(\mathbf{0}_V) \underset{\uparrow}{=} \mathbf{0}_W$$

(b) $$T(-\boldsymbol{v}) \underset{\uparrow\ \text{Theorem 2.8, page 54}}{=} T[(-1)\boldsymbol{v}] \overset{\text{by linearity}\ \downarrow}{=} (-1)T(\boldsymbol{v}) \underset{\uparrow}{=} -T(\boldsymbol{v})$$

CHECK YOUR UNDERSTANDING 4.2

Use Theorem 4.1(a) to show that the function $f\colon \Re^2 \to P_2(x)$ given by $f(a, b) = ax^2 + bx + 1$ is not linear.

Answer: See page B-12.

The following Theorem compresses the two conditions of Definition 4.1 into one:

You can perform the vector operations in V and then apply T to that result: $T(r\boldsymbol{v}+\boldsymbol{w})$, or you can first apply T and then perform the vector operations in W: $rT(\boldsymbol{v})+T(\boldsymbol{w})$. Either way, you will end up at the same vector in W.

THEOREM 4.2 $T\colon V \to W$ is linear **if and only if**:

$$T(r\boldsymbol{v}+\boldsymbol{v}') = rT(\boldsymbol{v}) + T(\boldsymbol{v}')$$

for all $\boldsymbol{v}, \boldsymbol{v}' \in V$ and $r \in \Re$

PROOF: If T is linear, then for all $\boldsymbol{v}, \boldsymbol{v}' \in V$ and $r \in \Re$:

$$T(r\boldsymbol{v}+\boldsymbol{v}') \underset{\text{(1) of Definition 4.1}\ \uparrow}{=} T(r\boldsymbol{v}) + T(\boldsymbol{v}') \underset{\uparrow\ \text{(1) of Definition 4.1}}{=} rT(\boldsymbol{v}) + T(\boldsymbol{v}')$$

Conversely, if $T(r\mathbf{v}+\mathbf{v'}) = rT(\mathbf{v})+T(\mathbf{v'})$ for all $\mathbf{v}, \mathbf{v'} \in V$ and $r \in \Re$, then:

$$(1):\ T(\mathbf{v}+\mathbf{v'}) \underset{\uparrow}{=} T(1\mathbf{v}+\mathbf{v'}) = 1T(\mathbf{v})+T(\mathbf{v'}) \underset{\uparrow}{=} T(\mathbf{v})+T(\mathbf{v'})$$

Axiom (viii), page 40

Theorem 4.1(a)

$$(2):\ T(r\mathbf{v}) = T(r\mathbf{v}+0) \underset{\uparrow}{=} rT(\mathbf{v})+T(0) \overset{\downarrow}{=} rT(\mathbf{v})+0 \underset{\uparrow}{=} rT(\mathbf{v})$$

Axiom (iii), page 40

You are invited to establish the following generalization of the above result in the exercises:

THEOREM 4.3 Let $T: V \to W$ be linear. For any vectors $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ in V, and any scalars $a_1, a_2, \dots, a_n$:

$$T(a_1\mathbf{v_1}+\dots+a_n\mathbf{v_n}) = a_1T(\mathbf{v_1})+\dots+a_nT(\mathbf{v_n})$$

Answer: See page B-12.

CHECK YOUR UNDERSTANDING 4.3

Use Theorem 4.2 to establish the linearity of the function of CYU 4.1.

The following theorem tells us that linear transformations map subspaces to subspaces:

THEOREM 4.4 Let $T: V \to W$ be linear. If S is a subspace of V, then $T(S) = \{T(s) | s \in S\}$ is a subspace of W.

See Theorem 2.13, page 61

PROOF: For $T(s_1), T(s_2) \in T(S)$, and $r \in \Re$:

$$rT(s_1)+T(s_2) \underset{\uparrow}{=} T(rs_1+s_2) \underset{\uparrow}{\in} T(S)$$

since T is linear

since S is a subspace, $rs_1+s_2 \in S$

CHECK YOUR UNDERSTANDING 4.4

PROVE OR GIVE A COUNTEREXAMPLE: Let $T: V \to W$ be linear, and $S \subseteq V$. If $T(S)$ is a subspace of W, then S must be a subspace of V.

A LINEAR MAP IS COMPLETELY DETERMINED BY ITS ACTION ON A BASIS

Suppose you have a run-of-the-mill function $f: \Re^2 \to \Re^3$, and you know that $f(\mathbf{1, 0}) = (\mathbf{1, 2, 3})$ and that $f(\mathbf{0, 1}) = (\mathbf{2, 3, -5})$. What can you say about $f(3, 5)$? Nothing. But if $T: \Re^2 \to \Re^3$ is linear, with $T(\mathbf{1, 0}) = (\mathbf{1, 2, 3})$ and $T(\mathbf{0, 1}) = (\mathbf{2, 3, -5})$, then:

$$\begin{aligned} T(\mathbf{3}, \mathbf{5}) &= T[\mathbf{3}(\mathbf{1}, \mathbf{0}) + \mathbf{5}(\mathbf{0}, \mathbf{1})] \\ \text{by linearity:} \quad &= 3T(\mathbf{1}, \mathbf{0}) + 5T(\mathbf{0}, \mathbf{1}) \\ &= 3(\mathbf{1}, \mathbf{2}, \mathbf{3}) + 5(\mathbf{2}, \mathbf{3}, -\mathbf{5}) = (\mathbf{13}, \mathbf{21}, -\mathbf{16}) \end{aligned}$$

Yes: A linear transformation $T: V \to W$ is **COMPLETELY DETERMINED** by its action on a basis of V

In particular, if you have two linear transformations that act identically on a basis, then those two transformations must, in fact, be one and the same:

THEOREM 4.5 Let V be a finite dimensional space with basis $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$. If $T: V \to W$ and $L: V \to W$ are linear maps such that $T(\mathbf{v_i}) = L(\mathbf{v_i})$ for $1 \le i \le n$, then $T(\mathbf{v}) = L(\mathbf{v})$ for every $\mathbf{v} \in V$.

PROOF: Let $\mathbf{v} \in V$. Since $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a basis for V, there exist scalars $a_1, a_2, \ldots, a_n$ such that:

$$\mathbf{v} = a_1\mathbf{v_1} + a_2\mathbf{v_2} + \ldots + a\mathbf{v_n}$$

We then have:

$$\begin{aligned} T(\mathbf{v}) = T(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n) &\overset{\text{linearity of } T}{=} a_1T(\mathbf{v}_1) + a_2T(\mathbf{v}_2) + \ldots + a_nT(\mathbf{v}_n) \\ \text{Since } T(v_i) = L(v_i): \quad &= a_1L(\mathbf{v}_1) + a_2L(\mathbf{v}_2) + \ldots + a_nL(\mathbf{v}_n) \\ \text{linearity of } L: \quad &= L(a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \ldots + a_n\mathbf{v}_n) = L(\mathbf{v}) \end{aligned}$$

CHECK YOUR UNDERSTANDING 4.5

PROVE OR GIVE A COUNTEREXAMPLE: Theorem 4.5 still holds if $\{\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_n}\}$ is a spanning set for V (not necessarily a basis).

Answer: See page B-12.

The next theorem gives a method for constructing **all** linear maps from one vector space to another:

THEOREM 4.6

LINEAR CONSTRUCTION

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be a basis for a vector space V, and let $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w_n}$ be n arbitrary vectors (not necessarily distinct) in a vector space W. There is a unique linear transformation $T: V \to W$ which maps $\mathbf{v}_i$ to $\mathbf{w}_i$ for $1 \le i \le n$; and it is given by:

$$T(a_1\mathbf{v}_1 + \ldots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \ldots + a_n\mathbf{w}_n$$

$$\begin{tikzcd} \boldsymbol{T(rv+v')} \arrow[d, squiggly] \\ \boldsymbol{rT(v)+T(v')} \end{tikzcd}$$

PROOF: From Theorem 4.4, we know that there can be at most one linear transformation $T\colon V \to W$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for $1 \le i \le n$. We complete the proof by establishing the fact that the above function T is indeed linear:

For $\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n$ and $\boldsymbol{v}' = b_1\boldsymbol{v}_1 + \ldots + b_n\boldsymbol{v}_n$ in V, and $r \in \Re$:

$$\begin{aligned}
\boldsymbol{T(rv+v')} &= T\left(r\sum_{i=1}^{n} a_i\boldsymbol{v}_i + \sum_{i=1}^{n} b_i\boldsymbol{v}_i\right) \\
&= T\left(\sum_{i=1}^{n}(ra_i+b_i)\boldsymbol{v}_i\right) = \sum_{i=1}^{n}(ra_i+b_i)w_i \qquad \text{definition of } T \\
&= r\sum_{i=1}^{n} a_iw_i + \sum_{i=1}^{n} b_iw_i = \boldsymbol{rT(v)+T(v')}
\end{aligned}$$

EXAMPLE 4.3 Let $T\colon \Re^2 \to \Re^3$ be the linear transformation which maps $(\mathbf{1, 1})$ to $(\mathbf{3, 2, 4})$ and $(\mathbf{4, 1})$ to $(\mathbf{1, 0, -5})$. Determine:

(a) $T(\mathbf{9, 6})$ (b) $\boldsymbol{T(a, b)}$

SOLUTION: (a) It is easy to see that $\{(1, 1), (4, 1)\}$ is a basis for $\Re^2$. We express $(9, 6)$ as a linear combination of the vectors in that basis:

$$(\mathbf{9, 6}) = r(\mathbf{1, 1}) + s(\mathbf{4, 1}) \longrightarrow \quad \begin{aligned} 9 &= r+4s \\ 6 &= r+s \\ \hline -\colon\ 3 &= \quad 3s \Rightarrow s = 1 \\ 6 &= r+1 \Rightarrow r = 5 \end{aligned}$$

Applying the linear transformation to $(\mathbf{9, 6}) = 5(\mathbf{1, 1}) + 1(\mathbf{4, 1})$ we have:

$$\begin{aligned}
T(\mathbf{9, 6}) = T[\mathbf{5(1, 1) + 1(4, 1)}] &= 5T(\mathbf{1, 1}) + 1T(\mathbf{4, 1}) \\
&= 5(\mathbf{3, 2, 4}) + (\mathbf{1, 0, -5}) \\
&= (\mathbf{15, 10, 20}) + (\mathbf{1, 0, -5}) \\
&= (\mathbf{16, 10, 15})
\end{aligned}$$

(b) Repeating the above process for (a, b), we have:

$$(\boldsymbol{a, b}) = r(\mathbf{1, 1}) + s(\mathbf{4, 1}) \longrightarrow \quad \begin{aligned} a &= r+4s \\ b &= r+s \\ \hline -\colon\ a-b &= \quad 3s \Rightarrow s = \frac{a-b}{3} \\ b &= r+\frac{a-b}{3} \Rightarrow r = \frac{4b-a}{3} \end{aligned}$$

Consequently:

$$T(\boldsymbol{a},\boldsymbol{b}) = rT(\mathbf{1},\mathbf{1})+sT(\mathbf{4},\mathbf{1}) = \frac{4b-a}{3}(\mathbf{3},\mathbf{2},\mathbf{4})+\frac{a-b}{3}(\mathbf{1},\mathbf{0},\mathbf{-5})$$

$$\text{details omitted:} \quad = \left(\frac{-2\boldsymbol{a}+11\boldsymbol{b}}{\mathbf{3}},\frac{-2\boldsymbol{a}+8\boldsymbol{b}}{\mathbf{3}},\frac{-9\boldsymbol{a}+21\boldsymbol{b}}{\mathbf{3}}\right)$$

CHECK YOUR UNDERSTANDING 4.6

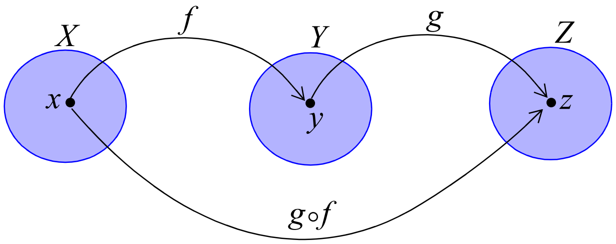
Let $T: \boldsymbol{R}^3 \to P_2(x)$ be the linear transformation which maps $(\mathbf{1},\mathbf{0},\mathbf{0})$ and $(\mathbf{0},\mathbf{2},\mathbf{0})$ to $2x^2+x$, and maps $(1, 1, 1)$ to $x-5$. Determine:

(a) $T(\mathbf{3},\mathbf{4},\mathbf{2})$ (b) $T(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})$.

Answer: See page B-12.

Composition of Linear Functions

As you may recall from earlier courses, for given functions $f: X \to Y$ and $g: Y \to Z$, the **composition** of f followed by g is that function, denoted by $g\circ f$, that is given by $(g\circ f)(x) = g[f(x)]$ (first apply f, and then apply g):



The following theorem tells us that the composition of linear functions is again a linear function:

THEOREM 4.7 If $T: V \to W$ and $L: W \to Z$ are linear, then the composition $L\circ T: V \to Z$ is also linear.

Proof: For $\mathbf{v}, \mathbf{v}' \in V$ and $r \in \Re$ we show that **the** function $L\circ T$ satisfies the condition of Theorem 4.1:

$$\begin{tikzcd}[row sep=large] (\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{r}\boldsymbol{v}+\boldsymbol{v}') \arrow[d, squiggly] \\ \boldsymbol{r}(\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{v})+(\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{v}') \end{tikzcd}$$

$$\begin{aligned} (\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{r}\boldsymbol{v}+\boldsymbol{v}') &= L[T(r\mathbf{v}+\mathbf{v}')] \overset{\text{linearity of } T}{=} L[rT(\mathbf{v})+T(\mathbf{v}')] \\ \text{linearity of L:} \quad &= rL[T(\mathbf{v})]+L[T(\mathbf{v}')] \\ \text{Definition of composition:} \quad &= \boldsymbol{r}(\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{v})+(\boldsymbol{L}\circ\boldsymbol{T})(\boldsymbol{v}') \end{aligned}$$

EXAMPLE 4.4 Let $T: \Re^3 \to M_{2\times 2}$ and $L: M_{2\times 2} \to P_2$ be given by:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \begin{bmatrix} a+b & 0 \\ c & -b \end{bmatrix}$$

and:

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = bx^2 + (a-c)x + c$$

(a) Show that T and L are linear.

(b) Show directly that the composite function $L \circ T: \Re^3 \to P_2$ is again linear.

SOLUTION: Linearity of T:

$$\begin{aligned} T[r(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) + (\boldsymbol{a}', \boldsymbol{b}', \boldsymbol{c}')] &= T[(\boldsymbol{ra + a}', \boldsymbol{rb + b}', \boldsymbol{rc + c}')] \\ &= \begin{bmatrix} (ra+a') + (rb+b') & 0 \\ rc + c' & -(rb+b') \end{bmatrix} \\ &= r\begin{bmatrix} a+b & 0 \\ c & -b \end{bmatrix} + \begin{bmatrix} a'+b' & 0 \\ c' & -b' \end{bmatrix} \\ &= rT(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) + T(\boldsymbol{a}', \boldsymbol{b}', \boldsymbol{c}') \end{aligned}$$

Linearity of L:

$$\begin{aligned} L\left(r\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) &= T\left(\begin{bmatrix} ra+a' & rb+b' \\ rc+c' & rd+d' \end{bmatrix}\right) \\ &= (rb+b')x^2 + [(ra+a') - (rc+c')]x + (rc+c') \\ &= r[bx^2 + (a-c)x + c] + [b'x^2 + (a'-c')x + c'] \\ &= rT\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) + T\left(\begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}\right) \end{aligned}$$

The composite function $L \circ T: \Re^3 \to P_2$:

$$\begin{aligned} (L \circ T)(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = L[T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})] &= L\left(\begin{bmatrix} a+b & 0 \\ c & -b \end{bmatrix}\right) \\ &= 0x^2 + [(a+b) - c]x + c \\ &= (a+b-c)x + c \end{aligned}$$

Linearity of $L\circ T\colon \Re^3 \to P_2$:

$$\begin{aligned}(L\circ T)[r(\mathbf{a}, \mathbf{b}, \mathbf{c}) + (\mathbf{a}', \mathbf{b}', \mathbf{c}')] &= (L\circ T)[(\mathbf{ra + a}', \mathbf{rb + b}', \mathbf{rc + c}')]\\ &= [(ra + a') + (rb + b') - (rc + c')]x + (rc + c')\\ &= r[(a + b - c)x + c] + [(a' + b' - c')x + c']\\ &= r(L\circ T)(\mathbf{a}, \mathbf{b}, \mathbf{c}) + (L\circ T)(\mathbf{a}', \mathbf{b}', \mathbf{c}')\end{aligned}$$

CHECK YOUR UNDERSTANDING 4.7

(a) Show that $\{(\mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{1})\}$ and $\{(\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{1})\}$ are bases for $\Re^2$ and $\Re^3$, respectively.

(b) Let $T\colon \Re^2 \to \Re^3$ be the linear transformation which maps $(\mathbf{1}, \mathbf{0})$ to $(\mathbf{0}, \mathbf{2}, \mathbf{0})$ and $(\mathbf{1}, \mathbf{1})$ to $(\mathbf{1}, \mathbf{0}, \mathbf{1})$. Let $L\colon \Re^3 \to \Re^2$ be the linear transformation which maps $(\mathbf{0}, \mathbf{1}, \mathbf{0})$ to $(\mathbf{0}, \mathbf{1})$ and both $(\mathbf{1}, \mathbf{1}, \mathbf{0})$ and $(\mathbf{0}, \mathbf{1}, \mathbf{1})$ to $(\mathbf{1}, \mathbf{0})$. Determine the map $L\circ T\colon \Re^2 \to \Re^2$.

Answer: See page B-13.

EXERCISES

Exercises 1-18. Determine whether or not the given function f is a linear transformation.

1. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = -\mathbf{5}\boldsymbol{x}$.
2. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \mathbf{0}$.
3. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \mathbf{1}$.
4. $f\colon \Re^2 \to \Re$, where $f(\boldsymbol{x}, \boldsymbol{y}) = \mathbf{2}\boldsymbol{x} - \mathbf{3}\boldsymbol{y}$.
5. $f\colon \Re^2 \to \Re$, where $f(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{xy}$.
6. $f\colon \Re^2 \to \Re^2$, where $f(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y})$.
7. $f\colon \Re^2 \to \Re^3$, where $f(\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \boldsymbol{xy}, \boldsymbol{y})$.
8. $f\colon \Re^2 \to \Re^3$, where $f(\boldsymbol{x}, \boldsymbol{y}) = (\frac{\boldsymbol{x}}{\mathbf{2}}, \boldsymbol{y}, \frac{\boldsymbol{y}}{\mathbf{2}})$.
9. $f\colon \Re^3 \to P_2(x)$, where $f(a, b, c) = ax^2 + bx + c$.
10. $f\colon P_2 \to \Re$, where $f(ax^2 + bx + c) = \boldsymbol{a} + \boldsymbol{b} + \boldsymbol{c}$.
11. $f\colon P_2 \to \Re$, where $f(ax^2 + bx + c) = \boldsymbol{abc}$.
12. $f\colon P_2 \to P_2$, where $f(ax^2 + bx + c) = a(x+1)^2 + b(x+1) + c$.
13. $f\colon P_2 \to P_3$, where $f(ax^2 + bx + c) = ax^3 + bx^2 + cx$.
14. $f\colon \Re^3 \to M_{2 \times 3}$, where $f(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \begin{bmatrix} 0 & a & b \\ c & b & a \end{bmatrix}$.
15. $f\colon M_{2 \times 2} \to \Re$, where $f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \boldsymbol{ad} - \boldsymbol{bc}$.
16. $f\colon M_{2 \times 2} \to M_{2 \times 2}$, where $f\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a-b & 0 \\ 0 & cd \end{bmatrix}$.
17. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \mathbf{sin}\boldsymbol{x}$.
18. $f\colon \Re \to V$, where $f(\boldsymbol{x}) = 2x$, and V is the vector space of Example 2.5, page 47.

19. Let the linear map $T: \Re^2 \to \Re^3$ be such that:

$$T(\mathbf{1, 1}) = (\mathbf{1, 2, 0}),\ T(\mathbf{0, 2}) = (\mathbf{1, 0, 2})$$

Find: (a) $T(\mathbf{5, 3})$ (b) $T(\boldsymbol{a, b})$

20. Let the linear map $T: \Re^3 \to \Re^2$ be such that:

$$T(\mathbf{1, 0, 0}) = (\mathbf{3, 2}),\ T(\mathbf{0, 1, 0}) = (\mathbf{2, 3}),\ \text{and}\ T(\mathbf{0, 0, 1}) = (\mathbf{1, 1})$$

Find: (a) $T(\mathbf{5, 3, -2})$ (b) $T(\boldsymbol{a, b, c})$

21. Let the linear map $T: \Re^2 \to M_{2 \times 2}$ be such that:

$$T(\mathbf{1, 0}) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix},\ T(\mathbf{1, 1}) = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$$

Find: (a) $T(\mathbf{5, 3})$ (b) $T(\boldsymbol{a, b})$

22. Let the linear map $T: \Re^2 \to M_{2 \times 2}$be such that:

$$T(\mathbf{1, 1}) = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix},\ T(\mathbf{3, 1}) = \begin{bmatrix} 2 & 1 \\ 3 & 2 \end{bmatrix}$$

Find: (a) $T(\mathbf{5, 3})$ (b) $T(\boldsymbol{a, b})$

23. Show that there cannot exist a linear transformation $T: R^2 \to R^2$ such that:

$$T(\mathbf{1, 2}) = (\mathbf{5, 3}),\ T(\mathbf{5, 3}) = (\mathbf{1, 2}),\ \text{and}\ T(\mathbf{1, 1}) = (\mathbf{2, 2})$$

24. Show that there cannot exist a linear transformation $T: \boldsymbol{R}^2 \to P_2$ such that:

$$T(\mathbf{1, 2}) = x^2 + 2,\ T(\mathbf{5, 3}) = 5x + 3,\ \text{and}\ T(\mathbf{1, 1}) = x^2 + x + 1$$

25. Show that the identity function I: V $\to V$, given by $I(\boldsymbol{v}) = \boldsymbol{v}$ for every $\boldsymbol{v}$ in V, is linear.

26. Show that the zero function $Z: V \to W$, given by $Z(\boldsymbol{v}) = \mathbf{0}$ for every $\boldsymbol{v}$ in V, is linear. (Referring to the equation $Z(\boldsymbol{v}) = \mathbf{0}$, where does $\mathbf{0}$ live?)

27. In precalculus and calculus, functions of the form $y = ax + b$ are typically called linear functions. Give necessary and sufficient conditions for a function of that form to be a linear operator on the vector space $\Re$.

28. Show that for any $a \in \Re$ the function $T_a: F(\Re) \to \Re$, where $T_a(f) = \boldsymbol{f}(\boldsymbol{a})$ is linear. (See Theorem 2.4, page 44.)

29. **(Calculus Dependent)** Let $D(\Re)$ be the subspace of the function space $F(\Re)$ consisting of all differentiable functions. Let $T: D(\Re) \to F(\Re)$ be given by $T(f) = f'$, where f' denotes the derivative of f. Show that T is linear.

30. **(Calculus Dependent)** Show that the function $T\colon P_2 \to P_4$, given by $T(p(x)) = 5x^3\frac{d}{dx}(px+5)$ is linear.

31. **(Calculus Dependent)** Show that if the linear function $T\colon P_2 \to P_1$ is such that $T(x^2) = 2x$, $T(x) = 1$, and $T(1) = 0$, then T is the derivative function.

32. **(Calculus Dependent)** Let $I[a, b]$ denote the vector space of all real-valued functions that are integrable over the interval $[a, b]$. Let $T\colon I[a, b] \to \Re$ be given by $T(f) = \int_a^b f(x)dx$. Show that T is linear.

33. Let $T\colon V \to W$ be linear and let S be a subspace of V. Show that $\{T(\boldsymbol{v}) | \boldsymbol{v} \in S\}$ is a subspace of W.

34. **(PMI)** Use the Principle of Mathematical Induction to prove Theorem 4.3.

35. Let $L(V, W) = \{\boldsymbol{T}\colon V \to W | \boldsymbol{T} \text{ is linear}\}$, with addition and scalar multiplication given by:

$$(\boldsymbol{T}_1 + \boldsymbol{T}_2)(\boldsymbol{v}) = \boldsymbol{T}_1(\boldsymbol{v}) + \boldsymbol{T}_2(\boldsymbol{v}) \text{ and } (r\boldsymbol{T})(\boldsymbol{v}) = r[\boldsymbol{T}(\boldsymbol{v})]$$

Show that $L(V, W)$ is a vector space.

36. (a) Show that if a function $f\colon \Re \to \Re$ satisfies the property that $f(r\boldsymbol{x}) = rf(\boldsymbol{x})$ for every $r \in \Re$ and $\boldsymbol{x} \in \Re$, then f is a linear function: which is ti say, that it must also satisfy the property that $f(\boldsymbol{x}_1 + \boldsymbol{x}_2) = f(\boldsymbol{x}_1) + f(\boldsymbol{x}_2)$ for every $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \Re$.

(b) Give an example of a function $f\colon \Re^2 \to \Re^2$ satisfying the property that $f[r(\boldsymbol{a}, \boldsymbol{b})] = rf(\boldsymbol{a}, \boldsymbol{b})$ for every $r \in \Re$ and every $(\boldsymbol{a}, \boldsymbol{b}) \in \Re^2$ but which does not satisfy the property that $f\colon [(\boldsymbol{a}, \boldsymbol{b}) + (\boldsymbol{c}, \boldsymbol{d})] = f(\boldsymbol{a}, \boldsymbol{b}) + f(\boldsymbol{c}, \boldsymbol{d})$.

(Note: It is by no means a trivial task to establish the existence of a non-linear function $f\colon V \to W$ which satisfies the property that $f(v_1 + v_2) = f(v_1) + f(v_2)$ for every $v_1, v_2 \in V$.)

37. Let $f\colon V \to W$ satisfy the condition that $f(\boldsymbol{v}_1 + \boldsymbol{v}_2) = f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2)$ for every $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Show that:

(a) $f(3\boldsymbol{v}) = 3f(\boldsymbol{v})$ for every $\boldsymbol{v} \in V$. (b) $f(\frac{2}{3}\boldsymbol{v}) = \frac{2}{3}f(\boldsymbol{v})$ for every $\boldsymbol{v} \in V$.

38. Let $f\colon V \to W$ satisfy the condition that $f(\boldsymbol{v}_1 + \boldsymbol{v}_2) = f(\boldsymbol{v}_1) + f(\boldsymbol{v}_2)$ for every $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$. Show that:

(a) **(PMI)** $f(n\boldsymbol{v}) = nf(\boldsymbol{v})$ for every $\boldsymbol{v} \in V$, and for every integer n.

(b) $f(\frac{a}{b}\boldsymbol{v}) = \frac{a}{b}f(\boldsymbol{v})$ for every $\boldsymbol{v} \in V$, and for every rational number $\frac{a}{b}$.

Exercises 39-44. Determine the indicated composition of the given linear functions, and then directly verify its linearity.

39. $L \circ T\colon \Re^2 \to \Re^2$, where $T\colon \Re^2 \to \Re$ is given by $T(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a} + \boldsymbol{b}$ and $L\colon \Re \to \Re^2$ by $L(\boldsymbol{a}) = (\boldsymbol{2a}, -\boldsymbol{a})$.

40. $L \circ T\colon \Re^2 \to \Re$, where $T\colon \Re^2 \to \Re^2$ is given by $T(\boldsymbol{a}, \boldsymbol{b}) = (-\boldsymbol{a}, \boldsymbol{a} + \boldsymbol{b})$ and $L\colon \Re^2 \to \Re$ by $L(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a} + \boldsymbol{2b}$

41. $L \circ T\colon \Re^2 \to P_2$, where $T\colon \Re^2 \to \Re$ is given by $T(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{3a} + \boldsymbol{b}$ and $L\colon \Re \to P_2$ by $L(\boldsymbol{a}) = ax^2 + 2a$.

42. $L \circ T\colon P_2 \to M_{2 \times 2}$, where $T\colon P_2 \to \Re^3$ is given by $T(ax^2 + bx + c) = (\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ and $L\colon \Re^3 \to M_{2 \times 2}$ by $L(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \begin{bmatrix} 2a & b \\ 0 & -c \end{bmatrix}$.

43. $K \circ L \circ T\colon \Re \to \Re^3$, where $T\colon \Re \to \Re^2$ is given by $T(\boldsymbol{a}) = (\boldsymbol{a}, -\boldsymbol{a})$, $L\colon \Re^2 \to \Re^2$ by $L(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{2a}, \boldsymbol{a} + \boldsymbol{b})$, and $K\colon \Re^2 \to \Re^3$ by $K(\boldsymbol{a}, \boldsymbol{b}) = (-\boldsymbol{a}, \boldsymbol{2b}, \boldsymbol{a} + \boldsymbol{b})$.

44. $K \circ L \circ T\colon \Re^3 \to P_2$, where $T\colon \Re^3 \to \Re^2$ is given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{a} - \boldsymbol{c})$, $L\colon \Re^2 \to \Re$ by $L(\boldsymbol{a}, \boldsymbol{b}) = \boldsymbol{a} - \boldsymbol{b}$, and $K\colon \Re \to P_2$ by $K(\boldsymbol{a}) = ax^2 + 2ax + 3a$.

45. **(PMI)** Let $L_i\colon V_i \to V_{i+1}$ be linear, for $1 \le i \le n$. Show that $L_n \circ \cdots \circ L_2 \circ L_1\colon V_1 \to V_{n+1}$ is linear.

PROVE OR GIVE A COUNTEREXAMPLE

46. For any $a \in \Re$ the function $T_a\colon V \to V$ given by $T_a(\boldsymbol{v}) = a\boldsymbol{v}$ is linear.

47. For any $\boldsymbol{v}_0 \in V$ the function $T_{v_0}\colon V \to V$ given by $T_{v_0}(\boldsymbol{v}) = \boldsymbol{v} + \boldsymbol{v}_0$ is linear.

48. Let $T\colon V \to W$ be linear. If $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a linearly independent subset of W then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a linearly independent subset of V.

49. Let $T\colon V \to W$ be linear. If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a linearly independent subset of V then $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a linearly independent subset of W.

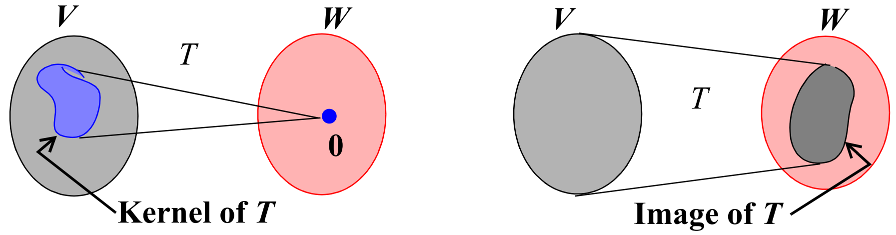
50. If for given functions $f\colon V \to W$ and $g\colon W \to Z$ the composite function $g \circ f\colon V \to Z$ is linear, then both f and g must be linear.

51. If for given functions $f\colon V \to W$ and $g\colon W \to Z$ the composite function $g \circ f\colon V \to Z$ is linear, then f must be linear.

52. If, for given functions $f\colon V \to W$ and $g\colon W \to Z$, the composite function $g \circ f\colon V \to Z$ is linear, then g must be linear.

§2. KERNEL AND IMAGE

For any given transformation $T\colon V \to W$, we define the kernel of T to be the set of vectors in V which map to the zero vector in W [see Figure 4.1(a)], and we define the set of all vectors in W which are "hit" by some $T(\boldsymbol{v})$ to be the image of T [see Figure 4.1(b)].



(a) (b)

Figure 4.1

More formally:

DEFINITION 4.2

KERNEL

Let $T\colon V \to W$ be linear. The **kernel** (or **null space**) of T is denoted by $\operatorname{Ker}(T)$ and is given by:

$$\operatorname{Ker}(T) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \mathbf{0}\}$$

IMAGE

The **image** of T is denoted by $\operatorname{Im}(T)$ and is given by:

$$\operatorname{Im}(T) = \{\boldsymbol{w} \in W \mid T(\boldsymbol{v}) = \boldsymbol{w} \text{ for some } \boldsymbol{v} \in V\}$$

Both the kernel and image of a linear transformation turn out to be subspaces of their respective vector space:

THEOREM 4.8 Let $T\colon V \to W$ be linear. Then:

(a) $\operatorname{Ker}(T)$ is a subspace of V.

(b) $\operatorname{Im}(T)$ is a subspace of W.

PROOF: We employ Theorem 2.13, page 61 to establish both parts of the theorem.

$\boldsymbol{v}_1, \boldsymbol{v}_2 \in \operatorname{Ker}(T)$ and $r \in R$

$$\rightsquigarrow$$

$r\boldsymbol{v}_1 + \boldsymbol{v}_2 \in \operatorname{Ker}(T)$

(a) Since $T(\mathbf{0}) = \mathbf{0}$, $\operatorname{Ker}(T)$ is nonempty.

Let $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \operatorname{Ker}(T)$ and $r \in \Re$. Then:

$$T(r\boldsymbol{v}_1 + \boldsymbol{v}_2) \overset{\text{linearity}}{=} rT(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) \underset{\text{since } \boldsymbol{v}_1, \boldsymbol{v}_2 \in \ker(T)}{=} r \cdot \mathbf{0} + \mathbf{0} = \mathbf{0}$$

Since T maps $r\boldsymbol{v}_1 + \boldsymbol{v}_2$ to $\mathbf{0}$, $r\boldsymbol{v}_1 + \boldsymbol{v}_2 \in \operatorname{Ker}(T)$.

$$\begin{tikzcd} w_1, w_2 \in \mathrm{Im}(T) \text{ and } r \in R \arrow[d, squiggly] \\ rw_1 + w_2 \in \mathrm{Im}(T) \end{tikzcd}$$

(b) Since $T(\mathbf{0}) = \mathbf{0}$, $\mathrm{Im}(T)$ is nonempty.

Let $w_1, w_2 \in \mathrm{Im}(T)$ and $r \in R$. Choose vectors v_1, v_2 such that $T(v_1) = w_1$ and $T(v_2) = w_2$ (how do we know that such vectors exist?). Then:

$$T(rv_1 + v_2) = rT(v_1) + T(v_2) = rw_1 + w_2$$

Since we found a vector in V which maps to $rw_1 + w_2$, $rw_1 + w_2 \in \mathrm{Im}(T)$.

DEFINITION 4.3

Let $T\colon V \to W$ be linear.

NULLITY

The dimension of $\mathrm{Ker}(T)$ is called the **nullity** of T, and is denoted by $\mathrm{nullity}(T)$.

RANK

The dimension of $\mathrm{Im}(T)$ is called the **rank** of T, and is denoted by $\mathrm{rank}(T)$.

The following theorem will be useful in determining the rank of a linear transformation.

In particular, if
$$\beta = \{v_1, v_2, \ldots, v_n\}$$
is a basis for V, then
$$\{T(v_1), T(v_2), \ldots, T(v_n)\}$$
will span $\mathrm{Im}(T)$.

THEOREM 4.9

Let $T\colon V \to W$ be linear. If

$$\mathrm{Span}\{v_1, v_2, \ldots, v_n\} = V$$

Then:

$$\mathrm{Span}\{T(v_1), T(v_2), \ldots, T(v_n)\} = \mathrm{Im}(T)$$

PROOF: We show that any $w \in \mathrm{Im}(T)$ can be written as a linear combinations of the vectors $\{T(v_1), T(v_2), \ldots, T(v_n)\}$:

$$\begin{tikzcd} w \in \mathrm{Im}(T) \arrow[d, squiggly] \\ w = \sum_{i=1}^{n} a_i T(v_i) \end{tikzcd}$$

Let $w \in \mathrm{Im}(T)$, and let $v \in V$ be such that $T(v) = w$. Since $S = \{v_1, v_2, \ldots, v_n\}$ spans V, we can express v as a linear combination of the vectors in S:.

$$v = a_1 v_1 + a_2 v_2 + \ldots + a_n v_n$$

By linearity:

$$T(v) = w = a_1 T(v_1) + a_2 T(v_2) + \ldots + a_n T(v_n)$$

EXAMPLE 4.5

(a) Show that the function$T\colon \Re^3 \to P_2$ given by:

$$T(a, b, c) = (a + b)x^2 + cx + c \text{ is linear.}$$

(b) Determine $\mathrm{Im}(T)$ and $\mathrm{rank}(T)$.

(c) Determine $\mathrm{Ker}(T)$ and $\mathrm{nullity}(T)$.

SOLUTION: (a) For $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}), (\boldsymbol{a'}, \boldsymbol{b'}, \boldsymbol{c'}) \in \Re^3$ and $r \in \Re$:

$$\begin{aligned} T[r(a, b, c) + (a', b', c')] &= T(ra + a', rb + b', rc + c') \\ &= [(ra + a') + (rb + b')]x^2 + (rc + c')x + (rc + c') \\ \text{regrouping:} \quad &= r[(a + b)x^2 + cx + c] + [(a' + b')x^2 + c'x + c'] \\ &= rT(a, b, c) + T(a', b', c') \end{aligned}$$

(b) By Theorem 4.9, we know that the vectors:

$$T(\mathbf{1, 0, 0}) = x^2, T(\mathbf{0, 1, 0}) = x^2, T(\mathbf{0, 0, 1}) = x + 1$$

span the image of T. Since x^2 and $x + 1$ are linearly independent, $\{x^2, x + 1\}$ is a basis for $\mathrm{Im}(T)$. Consequently, $\mathrm{rank}(T) = 2$.

(c) By definition:

$$\begin{aligned} \mathrm{Ker}(T) &= \{(\boldsymbol{a, b, c}) \mid T(\boldsymbol{a, b, c}) = \mathbf{0}\} \\ &= \{(\boldsymbol{a, b, c}) \mid (a + b)x^2 + cx + c = 0x^2 + 0x + 0\} \end{aligned}$$

This leads us to the following system of equations:

$$\left.\begin{aligned} a + b &= 0 \\ c &= 0 \end{aligned}\right\} \Rightarrow c = 0 \text{ and } a = -b$$

Thus: $\mathrm{Ker}(T) = \{(\boldsymbol{a, -a, 0}) \mid a \in R\}$. Since $\{(\mathbf{1, -1, 0})\}$ is a basis for $\mathrm{Ker}(T)$, $\mathrm{nullity}(T) = 1$.

CHECK YOUR UNDERSTANDING 4.8

(a) Show that the function $T: P_2 \to M_{2 \times 2}$ given by: $T(ax^2 + bx + c) = \begin{bmatrix} a & b \\ c & a \end{bmatrix}$ is linear.

(b) Determine $\mathrm{rank}(T)$ and $\mathrm{nullity}(T)$.

Answer: (a) See page B-13.
(b) $\mathrm{rank}(T) = 3$,
$\mathrm{nullity}(T) = 0$

THEOREM 4.10
DIMENSION THEOREM

Let V be a vector space of dimension n, and let $T: V \to W$ be linear. Then:

$$\mathrm{rank}(T) + \mathrm{nullity}(T) = n$$

PROOF: Start with a basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_k\}$ for $\mathrm{Ker}(T)$ and extend it (if necessary) to a basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \ldots, \boldsymbol{v}_{k+t}\}$ for V [see Theorem 3.12(a), page 100], where: $\boldsymbol{k + t = n}$—the dimension of V. We will show that the t vectors $\{T(\boldsymbol{v}_{k+1}), \ldots, T(\boldsymbol{v}_{k+t})\}$ constitute a basis for $\mathrm{Im}(T)$. This will complete the proof, for we will then have:

$$\mathbf{rank}(\boldsymbol{T}) + \mathbf{nullity}(\boldsymbol{T}) = \boldsymbol{k} + \boldsymbol{t} = \boldsymbol{n}$$

As you know, to establish the fact that $\{T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_{k+t})\}$ is a basis for $\mathrm{Im}(T)$ we have to verify that:

(1) The vectors $T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_{k+t})$ span the space $\mathrm{Im}(T)$.

And that:

Why can't you simply show just one of the two?

(2) The vectors $T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_{k+t})$ are linearly independent.

Let's do it:

(1) Let $\mathbf{w} \in \mathrm{Im}(T)$. Choose $\mathbf{v} \in V$ such that $T(\mathbf{v}) = \mathbf{w}$.
Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_{k+t}\}$ is a basis for V, we can express $\mathbf{v}$ as a linear combination of those $k+t$ vectors:

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \cdots + a_k\mathbf{v}_k + a_{k+1}\mathbf{v}_{k+1} + \cdots + a_{k+t}\mathbf{v}_{k+t}$$

By linearity:

$$\mathbf{w} = T(\mathbf{v}) = \underbrace{a_1T(\mathbf{v_1}) + a_2T(\mathbf{v_2}) + \cdots + a_kT(\mathbf{v_k})}_{\text{Since } v_1, v_2, \ldots, v_k \in \mathrm{Ker}(T)} + a_{k+1}T(\mathbf{v}_{k+1}) + \cdots + a_{k+t}T(\mathbf{v_{k+t}})$$

$$= \mathbf{0} + a_{k+1}T(\mathbf{v}_{k+1}) + \ldots + a_{k+t}T(\mathbf{v_{k+t}}) = a_{k+1}T(\mathbf{v}_{k+1}) + \ldots + a_{k+t}T(\mathbf{v_{k+t}})$$

a linear combination of the vectors $T(\mathbf{v}_{k+1}), \ldots, T(\mathbf{v}_{k+t})$

(2): Assume that:

$$b_1T(\mathbf{v}_{k+1}) + b_2T(\mathbf{v}_{k+2}) + \ldots + b_tT(\mathbf{v}_{k+t}) = \mathbf{0}$$

By linearity:

$$T[b_1(\mathbf{v}_{k+1}) + b_2(\mathbf{v}_{k+2}) + \ldots + b_t(\mathbf{v}_{k+t})] = \mathbf{0}$$

And therefore:

$$b_1(\mathbf{v}_{k+1}) + b_2(\mathbf{v}_{k+2}) + \ldots + b_t(\mathbf{v}_{k+t}) \in \mathrm{Ker}(T)$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_k\}$ is a basis for $\mathrm{Ker}(T)$, we can then find scalars $c_1, \ldots, c_k$ such that:

$$b_1(\mathbf{v}_{k+1}) + \ldots + b_t(\mathbf{v}_{k+t}) = c_1(\mathbf{v}_1) + \ldots + c_k(\mathbf{v}_k)$$

Or:

$$(-c_1)(\mathbf{v}_1) + \ldots + (-c_k)(\mathbf{v}_k) + b_1(\mathbf{v}_{k+1}) + \ldots + b_t(\mathbf{v}_{k+t}) = \mathbf{0}$$

Since $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k, \mathbf{v}_{k+1}, \ldots, \mathbf{v}_{k+t}\}$ is a basis, it is a linearly independent set of vectors. Consequently those c's and b's must all be zero; in particular: $b_1 = 0, b_2 = 0, \ldots, b_t = 0$.

$$b_1T(\mathbf{v}_{k+1}) + \ldots + b_tT(\mathbf{v}_{k+t}) = 0$$
$$\Big\downarrow$$
$$b_1 = 0, b_2 = 0, \ldots, b_t = 0$$

EXAMPLE 4.6 Determine bases for the kernel and image of the linear function $T: P_3 \to M_{2\times 2}$ given by:

$$T(ax^3+bx^2+cx+d) = \begin{bmatrix} a-b & c \\ 2c & a+d \end{bmatrix}$$

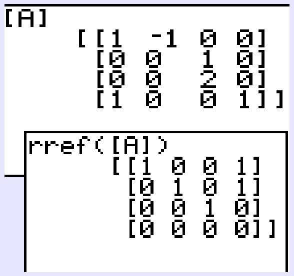
SOLUTION: We go for the kernel, as it is generally easier to find than the image apace:

To say that: $T(ax^3+bx^2+cx+d) = \mathbf{0}$

Is to say that: $\begin{bmatrix} a-b & c \\ 2c & a+d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Equating coefficients leads us to a homogeneous system of equations:

System **S** is certainly easy enough to solve directly. Still:



```
[A]
     [[1  -1  0  0]
      [0  0   1  0]
      [0  0   2  0]
      [1  0   0  1]]
```

```
rref([A])
     [[1 0 0 1]
      [0 1 0 1]
      [0 0 1 0]
      [0 0 0 0]]
```

each of these four equations is equal to 0.

$$\textbf{S:}\ \left.\begin{aligned} a-b&=0\\ c&=0\\ 2c&=0\\ a+d&=0\end{aligned}\right\}\xrightarrow{\textbf{coef[S]}}\begin{array}{cccc} a & b & c & d\\ 1 & -1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 2 & 0\\ 1 & 0 & 0 & 1\end{array}\xrightarrow{\textbf{rref}}\begin{array}{cccc} a & b & c & d\\ \mathbf{1} & 0 & 0 & 1\\ 0 & \mathbf{1} & 0 & 1\\ 0 & 0 & \mathbf{1} & 0\\ 0 & 0 & 0 & 0\end{array}\longrightarrow\left.\begin{aligned} a&=-d\\ b&=-d\\ c&=0\\ d&\text{ is free}\end{aligned}\right\}$$

From the above, we see that:

$$\text{Ker}(T) = \{-dx^3-dx^2+d \mid d\in R\} = \{d(-x^3-x^2+1)\mid d\in R\}$$

Conclusion: $\{-x^3-x^2+1\}$ is a basis for $\text{Ker}(T)$.

Knowing that $\text{nullity}(T) = 1$, we turn to Theorem 4.10 and conclude that

$$\text{rank}(T) = \dim(P_3)-\text{nullity}(T) = \underset{\uparrow}{4}-1 = 3$$

See Exercise 16, page 105

At this point, the easiest way to find a basis for $\text{Im}(T)$ is to apply T to 3 vectors in P_3 making sure that you end up with 3 **linearly independent** vectors in $\text{Im}(T)$—a basis for $\text{Im}(T)$. Which 3 vectors should we start with? Basically, you can take any 3 **randomly** chosen vectors, and the chances are that they will do fine (think about it); we will go with the 3 vectors x^3, x^2, and $x+1$:

$$T\Bigl(\overset{\mathbf{1x^3+0x^2+0x+0}}{\overset{\uparrow}{\mathbf{x^3}}}\Bigr) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ T\Bigl(\overset{\mathbf{0x^3+1x^2+0x+0}}{\overset{\uparrow}{\mathbf{x^2}}}\Bigr) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix},\ T\Bigl(\overset{\mathbf{0x^3+0x^2+1x+1}}{\overset{\uparrow}{\mathbf{x+1}}}\Bigr) = \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$

Recall that:

$$T(ax^3+bx^2+cx+d) = \begin{bmatrix} a-b & c \\ 2c & a+d \end{bmatrix}$$

We leave it to you to verify that above three vectors, which are certainly in Im(T) (why?), are linearly independent, and therefore constitute a basis for the 3-dimensional space Im(T).

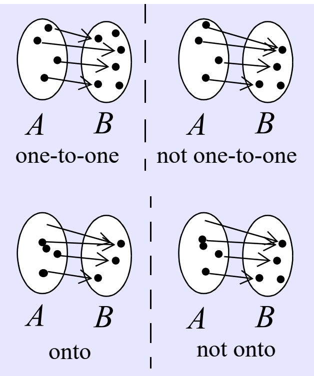
CHECK YOUR UNDERSTANDING 4.9

Determine bases for the kernel and image of the linear function $T\colon \Re^3 \to \Re^4$ given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (2\boldsymbol{a}, \boldsymbol{b}+\boldsymbol{c}, \boldsymbol{c}, \boldsymbol{b})$.

Answer: See page B-14.

ONE-TO-ONE AND ONTO FUNCTIONS

As you may recall:



DEFINITION 4.4

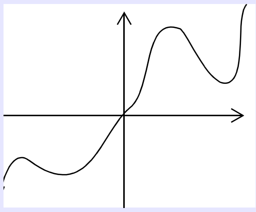
ONE-TO-ONE A function f from a set A to a set B is said to be **one-to-one** if:

$$f(a) = f(a') \Rightarrow a = a'$$

ONTO A function f from a set A to a set B is said to be **onto** if for every $b \in B$ there exist $a \in A$ such that $f(a) = b$.

When dealing with a linear transformation, we have:

The first part of this theorem is telling is that if a linear map is "one-to-one at **0**," then it is one-to-one everywhere. Certainly not true for other functions:



THEOREM 4.11 (a) A linear transformation $T\colon V \to W$ is one-to-one if and only if $\text{Ker}(T) = \{\mathbf{0}\}$.

(b) A linear transformation $T\colon V \to W$ is onto if and only if $\text{Im}(T) = W$.

PROOF: (a) Let T be one-to-one, and suppose that $T(\boldsymbol{v}) = \mathbf{0}$. Since $T(\mathbf{0}) = \mathbf{0}$ [Theorem 4.1(a), page 113], and since there can be but one $\boldsymbol{v} \in V$ that is mapped to $\mathbf{0} \in W$ (T is one-to-one), $\boldsymbol{v} = \mathbf{0}$.

Conversely, assume that $T(\boldsymbol{v}) = \mathbf{0} \Rightarrow \boldsymbol{v} = \mathbf{0}$. Then:

$$\begin{aligned} T(\boldsymbol{v}_1) &= T(\boldsymbol{v}_2) \\ T(\boldsymbol{v}_1) - T(\boldsymbol{v}_2) &= \mathbf{0} \\ T\text{ is linear:}\quad T(\boldsymbol{v}_1 - \boldsymbol{v}_2) &= \mathbf{0} \\ T(\boldsymbol{v}) = \mathbf{0} \Rightarrow \boldsymbol{v} = \mathbf{0}\text{:}\quad \boldsymbol{v}_1 - \boldsymbol{v}_2 &= \mathbf{0} \\ \boldsymbol{v}_1 &= \boldsymbol{v}_2 \end{aligned}$$

(b) Follows directly from the definition of onto (Definition 4.4) and the definition of $\text{Im}(T)$ (Definition 4.2).

EXAMPLE 4.7 Show that the linear function $T: \Re^3 \to \Re^3$ given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}+\boldsymbol{c}, 3\boldsymbol{b}, \boldsymbol{c}-\boldsymbol{b})$ is one-to-one.

$$\begin{tikzcd}T(a, b, c) = \mathbf{0} \arrow[d] \\ (a, b, c) = \mathbf{0}\end{tikzcd}$$

SOLUTION: We show that $T(\boldsymbol{v}) = \mathbf{0} \Rightarrow \boldsymbol{v} = \mathbf{0}$:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a}+\boldsymbol{c}, 3\boldsymbol{b}, \boldsymbol{c}-\boldsymbol{b}) = (\mathbf{0}, \mathbf{0}, \mathbf{0})$$

equating coefficients:

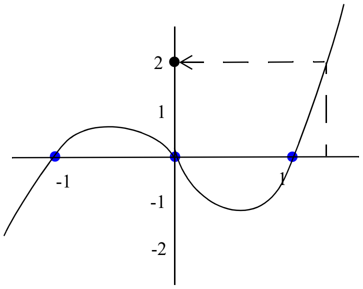
$$3b = 0 \Rightarrow \boldsymbol{b} = \mathbf{0}$$

$$c - b = 0 \Rightarrow c = b \Rightarrow \boldsymbol{c} = \mathbf{0}$$

$$a + c = 0 \Rightarrow a = -c \Rightarrow \boldsymbol{a} = \mathbf{0}$$

Applying Theorem 4.11, we conclude that T is one-to-one.

Consider the adjacent graph of the function $f(x) = x^3 - x$. As you can see, there are some y's which are only "hit by one x" (like $y = \mathbf{2}$), and there are some y's which are "hit by more than one x" (like $y = \mathbf{0}$) — the function is kind of one-to-one in some places, and not one-to-one in other places. Linear transformations are not so fickle; if a linear transformation is "one-to-one anywhere" (not just at $\mathbf{0}$) then it is one-to-one everywhere:



CHECK YOUR UNDERSTANDING 4.10

Let $T: V \to W$ be linear. Show that if there exists $\bar{\boldsymbol{v}} \in V$ such that $T(\boldsymbol{v}) = T(\bar{\boldsymbol{v}}) \Rightarrow \boldsymbol{v} = \bar{\boldsymbol{v}}$ then T is one-to-one.

(So, one-to-one anywhere implies one-to-one everywhere.)

Answer: See page B-14.

EXERCISES

Exercises 1-18. Determine if the given function is linear. If it is, find a basis for its kernel and its image space.

1. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = -\boldsymbol{5x}$

2. $f\colon \Re \to \Re^2$, where $f(\boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{2x})$

3. $f\colon \Re \to \Re^2$, where $f(\boldsymbol{x}) = (\boldsymbol{x}, -\boldsymbol{x})$

4. $f\colon \Re \to \Re^2$, where $f(\boldsymbol{a}) = (\boldsymbol{a}, \boldsymbol{a}^2)$

5. $f\colon \Re^2 \to \Re^2$, where $f(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{ab}, \boldsymbol{a})$

6. $f\colon \Re^2 \to \Re^2$, where $f(\boldsymbol{a}, \boldsymbol{b}) = (-\boldsymbol{2b}, \boldsymbol{a})$

7. $f\colon \Re^2 \to P_1$, where $f(\boldsymbol{a}, \boldsymbol{b}) = ax + b$

8. $f\colon \Re^3 \to \Re^3$, where $f(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{0}, \boldsymbol{b}, \boldsymbol{c})$

9. $f\colon M_{2\times 2} \to P_2$, where $f\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (a-b) + (b-c)x + (c-a)x^2$

10. $f\colon \Re^2 \to M_{2\times 2}$, where $f(\boldsymbol{a}, \boldsymbol{b}) = \begin{bmatrix} a & b \\ a+b & a-b \end{bmatrix}$

11. $f\colon \Re^4 \to M_{2\times 2}$, where $f(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = \begin{bmatrix} a & 2b \\ cd & c+d \end{bmatrix}$

12. $f\colon P_2 \to \Re^3$, where $f[p(x)] = [\boldsymbol{p}(\boldsymbol{1}), \boldsymbol{p}(\boldsymbol{2}), \boldsymbol{p}(\boldsymbol{3})]$

13. $f\colon P_2 \to \Re^2$, where $f[p(x)] = [\boldsymbol{p}(\boldsymbol{0}), \boldsymbol{p}(\boldsymbol{1})]$

14. $f\colon P_3 \to P_4$, where $f[p(x)] = xp(x)$

15. $f\colon P_3 \to P_3$, where $f[p(x)] = p(x) + p(1)$

16. $f\colon P_3 \to M_{2\times 2}$, where $f[p(x)] = \begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix}$

17. $f\colon \Re^3 \to M_{2\times 2}$, where $f(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \begin{bmatrix} a-b & b-c \\ a+b & b+c \end{bmatrix}$

18. $f\colon \Re^4 \to M_{2\times 2}$, where $f(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = \begin{bmatrix} 2a & c+d \\ c-d & b \end{bmatrix}$

Exercises 19-27. (Calculus Dependent) Show that the given function is linear. Find a basis for its kernel and its image space.

19. f: $P_3 \to P_3$, where $f[p(x)] = p'(x)$

20. f: $P_3 \to P_3$, where $f[p(x)] = p'(x)$

21. f: $P_3 \to P_3$, where $f[p(x)] = p'(x) + p(x)$

22. f: $P_2 \to P$, where $f[p(x)] = p''(x)$

23. f: $P_2 \to P_3$, where $f[p(x)] = 2p(x) - 3p'(x)$

24. f: $P_1 \to \Re$ where $f[p(x)] = \int_0^1 p(x)dx$

25. f: $P_2 \to \Re$ where $f[p(x)] = \int_0^1 p(x)dx$

26. f: $P_1 \to P_2$ where $f[p(x)] = \int_0^x p(t)dt$

27. f: $P_2 \to P_3$ where $f[p(x)] = \int_0^x p(t)dt$

28. Let T: $\Re^3 \to \Re^3$ be given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a} - \boldsymbol{b}, \boldsymbol{a} + \boldsymbol{c}, \boldsymbol{c})$.
    (a) Which, if any of the following 3-tuples are in the kernel of T?
        (i) $(\mathbf{-3, 3, 0})$ (ii) $(\mathbf{0, -3, -3})$ (iii) $(\mathbf{3, 0, 0})$
    (b) Which, if any of the following 3-tuples are in the image T?
        (i) $(\mathbf{-3, 3, 0})$ (ii) $(\mathbf{0, -3, -3})$ (iii) $(\mathbf{3, 0, 0})$

29. Let T: $\Re^3 \to P_3$ be given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = ax^3 + ax + (b + c)$.
    (a) Which, if any of the following 3-tuples are in the kernel of T?
        (i) $(\mathbf{1, 0, 0})$ (ii) $(\mathbf{0, -3, 3})$ (iii) $(\mathbf{0, 0, 1})$
    (b) Which, if any of the following 3-tuples are in the image of T?
        (i) $x^3 + x$ (ii) 5 (iii) $x^3 + 5$

30. Determine a basis for the kernel and image of the linear transformation T: $\Re^3 \to \Re^3$ which maps $(\mathbf{1, 0, 0})$ to $(\mathbf{1, 1, 1})$, $(\mathbf{0, 1, 0})$ to $(\mathbf{3, -2, 5})$, and $(\mathbf{0, 0, 1})$ to $(\mathbf{-2, 3, -4})$.

31. Determine a basis for the kernel and image of the linear transformation T: $\Re^3 \to M_{2 \times 2}$ which maps $(\mathbf{1, 0, 0})$ to $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$, $(\mathbf{0, 1, 0})$ to $\begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}$, and $(\mathbf{0, 0, 1})$ to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

32. Determine a basis for kernel and image of the linear transformation $T\colon M_{2\times 2} \to P_4$ which maps $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ to $2x^4+1$, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$ to x^4-x, $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ to 5, and $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ to x^2.

33. Find, if one exists, a linear transformation $T\colon \Re^2 \to \Re^3$ such that:
    (a) $\{(\mathbf{9, 1, 3})\}$ is a basis for $\operatorname{Im}(T)$.
    (b) $\{(\mathbf{9, 0, 0}), (\mathbf{0, 1, 0})\}$ is a basis for $\operatorname{Im}(T)$.
    (c) $\{(\mathbf{9, 1, 3}), (\mathbf{3, 1, 9}), (\mathbf{9, 3, 1})\}$ is a basis for $\operatorname{Im}(T)$.

34. Find, if one exists, a linear transformation $T\colon \Re^3 \to \Re^3$ such that:
    (a) $\{(\mathbf{9, 1, 3})\}$ is a basis for $\operatorname{Ker}(T)$.
    (b) $\{(\mathbf{9, 0, 0}), (\mathbf{0, 1, 0})\}$ is a basis for $\operatorname{Ker}(T)$.
    (c) $\{(\mathbf{9, 1, 3}), (\mathbf{3, 1, 9}), (\mathbf{9, 3, 1})\}$ is a basis for $\operatorname{Ker}(T)$.

35. Let $\dim(V) = n$, and let $T\colon V \to V$ be the linear operator $T(\boldsymbol{v}) = r\boldsymbol{v}$, for $r \in \Re$. Enumerate the possible values of $\operatorname{nullity}(T)$ and $\operatorname{rank}(T)$.

36. Let $T\colon V \to W$ be linear with $\dim(V) = 2$ and $\dim(W) = 3$. Enumerate the possible values of $\operatorname{nullity}(T)$ and $\operatorname{rank}(T)$.

37. Let $T\colon V \to W$ be linear with $\dim(V) = 3$ and $\dim(W) = 2$. Enumerate the possible values of $\operatorname{nullity}(T)$ and $\operatorname{rank}(T)$.

38. Let $T\colon P_3 \to \Re^5$ be a one-to-one linear map. Determine the rank and nullity of T.

39. Let $T\colon \Re^5 \to P_3$ be an onto linear map. Determine the rank and nullity of T.

40. Let $T\colon V \to V$ be a linear operator, with $\dim(V) = n$. Prove that $\operatorname{Im}(T) = V$ if and only if T is one-to-one.

41. Let $T\colon V \to W$ be linear, with $\dim(V) = \dim(W)$. Prove that $\{\boldsymbol{v}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ is a basis for V if and only if $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a basis for W.

42. Give an example of a linear transformation $T\colon V \to W$ such that $\operatorname{Im}(T) = \operatorname{Ker}(T)$.

43. Let $T\colon V \to W$ be a linear transformation, with $\dim(V) = n$. Prove that T is one-to-one if and only if $\operatorname{rank}(T) = n$.

44. Let $T\colon V \to W$ be linear, with $\dim(V) = \dim(W)$. Prove that T is one-to-one if and only if T is onto.

45. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear.
    (a) Prove that $\operatorname{Ker}(T) \subseteq \operatorname{Ker}(L \circ T)$.
    (b) Give an example for which $\operatorname{Ker}(T) = \operatorname{Ker}(L \circ T)$.
    (c) Give an example for which $\operatorname{Ker}(T) \subset \operatorname{Ker}(L \circ T)$.

PROVE OR GIVE A COUNTEREXAMPLE

46. If $\text{span}[T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)] = \text{Im}(T)$ then $\text{span}(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n) = V$.

47. There exists a one-to-one linear map $T\colon M_{2 \times 3} \to P_4$.

48. There exists a one-to-one linear map $T\colon P_4 \to M_{2 \times 3}$.

49. There exists an onto linear map $T\colon M_{2 \times 3} \to P_4$.

50. There exists an onto linear map $T\colon P_4 \to M_{2 \times 3}$.

51. If $T\colon V \to W$ is linear and $\dim(V) < \dim(W)$, then T cannot be onto.

52. If $T\colon V \to W$ is linear and $\dim(V) < \dim(W)$, then T cannot be one-to-one.

53. If $T\colon V \to W$ is linear and $\dim(V) > \dim(W)$, then T cannot be onto.

54. If $T\colon V \to W$ is linear and $\dim(V) > \dim(W)$, then T cannot be one-to-one.

55. There exists a linear transformation $T\colon \Re^2 \to \Re^4$ such that $\text{rank}(T) = \text{nullity}(T)$.

56. There exists a linear transformation $T\colon \Re^3 \to \Re^4$ such that $\text{rank}(T) = \text{nullity}(T)$.

57. There exists a linear transformation $T\colon \Re^2 \to \Re^4$ such that $\text{rank}(T) < \text{nullity}(T)$.

58. There exists a linear transformation $T\colon \Re^2 \to \Re^4$ such that $\text{rank}(T) > \text{nullity}(T)$.

59. If $T\colon V \to W$ is linear and if W is finite dimensional, then V is finite dimensional.

60. If $T\colon V \to W$ is linear and if V is finite dimensional, then W is finite dimensional.

61. If $T\colon V \to W$ is linear and if V is finite dimensional, then $\text{Im}(T)$ is finite dimensional.

62. If $T\colon V \to W$ is linear and if $\text{Im}(T)$ is finite dimensional, then V is finite dimensional.If $T\colon V \to W$ is linear and if $\text{Ker}(T)$ is finite dimensional, then either V or W is finite dimensional.

63. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear. If $\dim(V) = 3, \dim(W) = 2$, and $\text{nullity}(T) = 1$, then $\text{rank}(L \circ T) \le 1$.

64. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear. If $\dim(V) = 3, \dim(W) = 3$, $\text{rank}(T) = 1$, and $\text{nullity}(L) = 2$, then $\text{rank}(L \circ T) = 1$.

65. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear. If $\dim(V) = 3, \dim(W) = 3$, $\text{rank}(T) = 1$, and $\text{nullity}(L) = 2$, then $\text{rank}(L \circ T) \ge 1$.

66. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear, with $\dim(W) = n$ and $\dim(Z) = m$. If T is one-to-one and L is onto, then $\dim(V) = n - m$.

§3. ISOMORPHISMS

We can all agree that there is little difference between the vector space $\Re^n$ with its horizontal n-tuples, and the space $M_{n\times 1}$ of "vertical n-tuples". In this section, we show that there is, in fact, little difference between the vector space $\Re^n$ and any n dimensional vector space whatsoever.

BIJECTIONS AND INVERSE FUNCTIONS

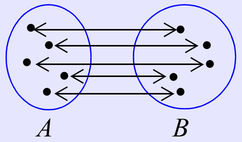
One-to-one and onto functions were previously defined on page 129. Of particular interest are functions that satisfy both properties:

DEFINITION 4.5
BIJECTION

A function $f: A \to B$ that is both one-to-one and onto is said to be a **bijection**.

Roughly speaking:

> A bijection $f: A \to B$ serves to pair of each elements of A with those of B (see margin).



a bijection

THEOREM 4.12 If $f: A \to B$ and $g: B \to C$ are bijections, then the composite function $g\circ f: A \to C$ is also a bijection.

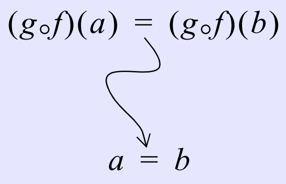
PROOF: $g\circ f$ is one-to-one:

$$(g\circ f)(a) = (g\circ f)(b)$$

Definition of composition: $g[f(a)] = g[f(b)]$

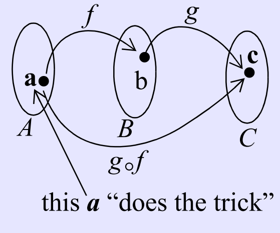
Since g is one-to-one: $f(a) = f(b)$

Since f is one-to-one: $a = b$



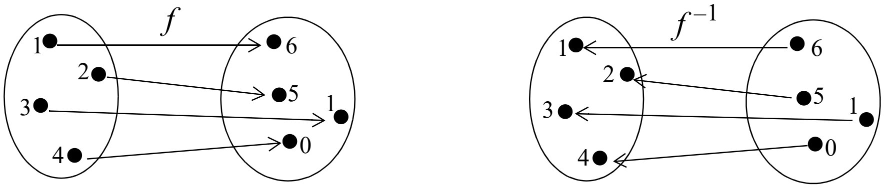
$g\circ f$ is onto: Let $c \in C$ be given (see margin). Since g is onto there exists $b \in B$ such that $g(b) = c$. Since f is onto, there exists $a \in A$ such that $f(a) = b$. We then have:

$$(g\circ f)(a) = g[f(a)] = g(b) = c$$



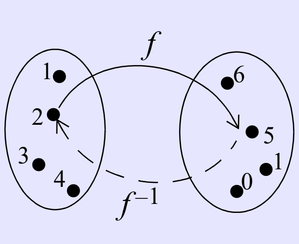
Figuratively speaking, if you reverse the arrows of the bijection $f: \{1, 2, 3, 4\} \to \{6, 5, 1, 0\}$ of Figure 4.2(a), you end up with the bijection $f^{-1}: \{6, 5, 1, 0\} \to \{1, 2, 3, 4\}$ of Figure 4.2(b), which is called the **inverse** of f.

Only bijections
$f: X \to Y$
have inverses
$f^{-1}: Y \to X$



(a) (b)

Figure 4.2



The relationship between the functions f and f^{-1} depicted in the margin reveals the fact that each function "undoes" the work of the other. For example:

$$(f^{-1} \circ f)(\mathbf{2}) = f^{-1}[f(2)] = f^{-1}(5) = \mathbf{2}$$

and

$$(f \circ f^{-1})(\mathbf{5}) = f[f^{-1}(5)] = f(2) = \mathbf{5}$$

In general:

DEFINITION 4.6

INVERSE FUNCTIONS

The inverse of a bijection $f: X \to Y$ is the function $f^{-1}: Y \to X$ such that:

$$(f^{-1} \circ f)(x) = x \text{ for every } x \text{ in } X$$

and

$$(f \circ f^{-1})(y) = y \text{ for every } y \text{ in } Y.$$

CHECK YOUR UNDERSTANDING 4.11

Prove that if $f: X \to Y$ is a bijection, then so is $f^{-1}: Y \to X$.

Answer: See page B-14.

BACK TO LINEAR ALGEBRA

THEOREM 4.13 If the bijection $T: V \to W$ is linear, then its inverse, $T^{-1}: W \to V$ is also linear.

PROOF: Let $\mathbf{w}, \mathbf{w}' \in W$ and $r \in \Re$ be given. Let $\mathbf{v}, \mathbf{v}' \in V$ be such that $T(\mathbf{v}) = \mathbf{w}$ and $T(\mathbf{v}') = \mathbf{w}'$. Then:

$$T^{-1}(r\mathbf{w} + \mathbf{w}') = T^{-1}(rT(\mathbf{v}) + T(\mathbf{v}')) \overset{\text{linearity of } T}{=} T^{-1}[T(r\mathbf{v} + \mathbf{v}')]$$

$$\underset{\text{Definition of composition}}{=} (T^{-1} \circ T)(r\mathbf{v} + \mathbf{v}') \underset{\text{Definition 4.6}}{=} r\mathbf{v} + \mathbf{v}' \underset{\text{since } T(\mathbf{v}) = \mathbf{w} \text{ and } T(\mathbf{v}') = \mathbf{w}'}{=} rT^{-1}(\mathbf{w}) + T^{-1}(\mathbf{w}')$$

EXAMPLE 4.8 Show that the linear map $T\colon \Re^3 \to \Re^3$ given by:

$$T(\boldsymbol{a, b, c}) = (\boldsymbol{a + c, 3b, c - b})$$

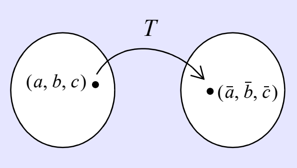
is a bijection. Find its inverse and show, directly, that T^{-1} is linear.

SOLUTION:

***T* IS ONE-TO-ONE.** We start with $T(\boldsymbol{a, b, c}) = T(\boldsymbol{a', b', c'})$ and go on to show that $(\boldsymbol{a, b, c}) = (\boldsymbol{a', b', c'})$ (see Definition 4.4, page 129):

$$T(\boldsymbol{a, b, c}) = T(\boldsymbol{a', b', c'}) \Rightarrow (\boldsymbol{a + c, 3b, c - b}) = (\boldsymbol{a' + c', 3b', c' - b'})$$

$$\Rightarrow \left.\begin{array}{r} a + c = a' + c' \\ 3b = 3b' \\ c - b = c' - b' \end{array}\right\} \Rightarrow \begin{array}{l} a = a' \\ b = b' \\ c = c' \end{array} \quad \begin{array}{l} \text{then} \\ \text{then} \end{array}$$



***T* IS ONTO.** We start with $(\boldsymbol{a', b', c'})$ and find an $(\boldsymbol{a, b, c})$ such that $T(\boldsymbol{a, b, c}) = (\boldsymbol{a', b', c'})$ (see Definition 4.4, page 129):

$$T(\boldsymbol{a, b, c}) = (\boldsymbol{a + c, 3b, c - b}) = (\boldsymbol{a', b', c'})$$

Equating coefficients brings us to the system of equations:

$$\left.\begin{array}{r} a + c = a' \\ 3b = b' \\ c - b = c' \end{array}\right\} \Rightarrow \begin{array}{l} \boldsymbol{a} = a' - \bar{c} - \dfrac{b'}{3} \\ \boldsymbol{b} = \dfrac{b'}{3} \\ \boldsymbol{c} = c' + \dfrac{b'}{3} \end{array} \quad \begin{array}{l} \text{then} \\ \text{then} \end{array}$$

Let's make sure that $(\boldsymbol{a, b, c}) = \left(\boldsymbol{a' - c' - \frac{b'}{3}, \frac{b'}{3}, c' + \frac{b'}{3}}\right)$ works:

$$T(\boldsymbol{a, b, c}) = T\left(\boldsymbol{a' - c' - \frac{b'}{3}, \frac{b'}{3}, c' + \frac{b'}{3}}\right)$$

$$= \left[\left(\boldsymbol{a' - \bar{c} - \frac{b'}{3}}\right) + \left(\boldsymbol{c' + \frac{b'}{3}}\right), 3\left(\boldsymbol{\frac{b'}{3}}\right), \left(\boldsymbol{c' + \frac{b'}{3}}\right) - \left(\boldsymbol{\frac{b'}{3}}\right)\right] = (\boldsymbol{a', b', c'})$$

FINDING THE INVERSE OF *T*. Since

$$T\left(\boldsymbol{a' - c' - \frac{b'}{3}, \frac{b'}{3}, c' + \frac{b'}{3}}\right) = (\boldsymbol{a', b', c'}):$$

$$T^{-1}(\boldsymbol{a', b', c'}) = \left(\boldsymbol{a' - c' - \frac{b'}{3}, \frac{b'}{3}, c' + \frac{b'}{3}}\right)$$

Dropping the primes, we show, directly, that $T^{-1}\colon \Re^3 \to \Re^3$ given by

$T^{-1}(\boldsymbol{a, b, c}) = \left(\boldsymbol{a - c - \frac{b}{3}, \frac{b}{3}, c + \frac{b}{3}}\right)$ **IS LINEAR**:

$$\begin{tikzcd}T^{-1}[r(a,b,c)+(a',b',c')] \arrow[d, "="] \\ rT^{-1}(a,b,c)+T^{-1}(a',b',c')\end{tikzcd}$$

$$\begin{aligned}
T^{-1}[\boldsymbol{r(a,b,c)+(a',b',c')}] &= T^{-1}(\boldsymbol{ra+a', rb+b', rc+c'}) \\
&= \left(\boldsymbol{ra+a'-rc-c'-\frac{rb}{3}-\frac{b'}{3}, \frac{rb}{3}+\frac{b'}{3}, rc+c'+\frac{rb}{3}+\frac{b'}{3}}\right) \\
&= r\left(\boldsymbol{a-c-\frac{b}{3}, \frac{b}{3}, c+\frac{b}{3}}\right)+\left(\boldsymbol{a'-c'-\frac{b'}{3}, \frac{b'}{3}, c'+\frac{b'}{3}}\right) \\
&= rT^{-1}(\boldsymbol{a,b,c})+T^{-1}(\boldsymbol{a',b',c'})
\end{aligned}$$

CHECK YOUR UNDERSTANDING 4.12

Show that the linear map $T: \Re^2 \rightarrow P_1$ given by:

$$T(\boldsymbol{a,b}) = (a+b)x-a$$

is a bijection. Find its inverse, and show directly that it is linear.

Answer: See page B-14.

Roughly speaking, two vector spaces will be considered to be the "same" if the vectors of one space can be pared off with those of the other, while preserving the vector space structures of those spaces. More formally:

DEFINITION 4.7 ISOMORPHISM

A linear map $T: V \rightarrow W$ which is one-to-one and onto is said to be an **isomorphism** from the vector space V to the vector space W.

EXAMPLE 4.9

Show that the function $T: \Re^3 \rightarrow P_2$ given by:

$$T(\boldsymbol{a,b,c}) = -ax^2+(b+c)x+3c$$

is an isomorphism.

SOLUTION: We start off by establishing linearity, for we can then take advantage of previously established theory to show that the function is one-to-one and onto:

Linearity:
$$\begin{aligned}
T[\boldsymbol{r(a,b,c)+(a',b',c')}] &= T(\boldsymbol{ra+a', rb+b', rc+c'}) \\
&= -(ra+a')x^2+[(rb+b')+(rc+c')]x+3(rc+c') \\
&= r[-ax^2+(b+c)x+3c]+[-a'x^2+(b'+\bar{c})x+3c'] \\
&= rT(\boldsymbol{a,b,c})+T(\boldsymbol{a',b',c'})
\end{aligned}$$

One-to-one: We show that $\text{Ker}(T) = \mathbf{0}$ [see Theorem 4.11(a), page 129]:

$$T(\boldsymbol{a,b,c}) = \mathbf{0} \Rightarrow -ax^2+(b+c)x+3c = 0x^2+0x+0$$

Equation coefficients:

$$-a = 0, b+c = 0, 3c = 0$$

$$\text{or:} \qquad a = b = c = 0$$

Onto: We show that $\text{Im}(T) = P_2$ [see Theorem 4.11(b)]:
The Dimension Theorem of page 126, tells us that:

$$\text{rank}(T) + \text{nullity}(T) = 3$$

Knowing that the nullity is 0, we conclude that $\text{rank}(T) = 3$. Since P_2 is of dimension 3: $\text{Im}(T) = P_2$ (Exercise 50, page 107).

This theorem asserts that "*isomorphic*" is an **equivalence relation** on any set of vector spaces. See Exercises 37-39.

THEOREM 4.14

(a) Every vector space is isomorphic to itself.

(b) If V is isomorphic to W, then W is isomorphic to V.

(c) If V is isomorphic to W, and W is isomorphic to Z, then V is isomorphic to Z.

PROOF: (a) The identity map $I_V: V \to V$ is easily seen to be an isomorphism.

(b) To say that V is isomorphic to W is to say that there exists an isomorphism $T: V \to W$. Since $T^{-1}: W \to V$ is also a linear bijection (CYU 4.12 and Theorems 4.13), W is isomorphic to V.

(c) Let $T: V \to W$ and $L: W \to Z$ be isomorphisms. Since $L \circ T: V \to Z$ is both a bijection (Theorem 4.12) and linear (Theorem 4.7, page 117), V is isomorphic to Z.

Theorem 4.14(b) enables us to formulate the following definition:

DEFINITION 4.8
ISOMORPHIC SPACES

Two vector spaces V and W are **isomorphic**, written $V \cong W$, if there exists an isomorphism from one of the vector spaces to the other.

CHECK YOUR UNDERSTANDING 4.13

Prove that $\Re^4 \cong M_{2 \times 2}$. (You have to exhibit an isomorphism from one of the spaces to the other, whichever you prefer).

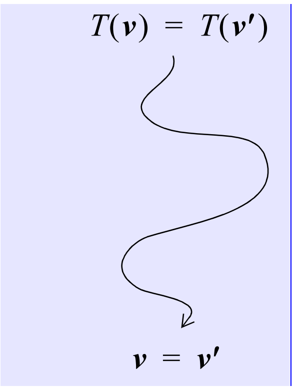
Answer: See page B-15.

The following lovely result says that all n dimensional vector spaces are isomorphic to the Euclidean n-space.

THEOREM 4.15 If V is a vector space of dimension n, then:

$$V \cong \Re^n$$

PROOF: Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for V, and let $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ be the standard basis for $\Re^n$ (see page 94). Consider the function T: V $\to \Re^n$, given by: $T(a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n) = a_1\boldsymbol{e}_1 + \ldots + a_n\boldsymbol{e}_n$. Theorem 4.6, page 115 assures us that T is linear. We complete the proof by showing that T is both one-to-one and onto (and therefore an isomorphism).



***T* is one-to-one:** Let $T(\boldsymbol{v}) = T(\boldsymbol{v}')$, for:

$$\boldsymbol{v} = a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n \text{ and } \boldsymbol{v}' = b_1\boldsymbol{v}_1 + \ldots + b_n\boldsymbol{v}_n$$

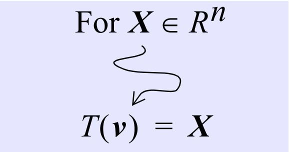
Then:

$$T(a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n) = T(b_1\boldsymbol{v}_1 + \ldots + b_n\boldsymbol{v}_n)$$

$$a_1T(\boldsymbol{v}_1) + \ldots + a_nT(\boldsymbol{v}_n) = b_1T(\boldsymbol{v}_1) + \ldots + b_nT(\boldsymbol{v}_n)$$

$$a_1\boldsymbol{e}_1 + \ldots + a_n\boldsymbol{e}_n = b_1\boldsymbol{e}_1 + \ldots + b_n\boldsymbol{e}_{1n}$$

Since $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ is a linearly independent set of vectors we have, by Theorem 3.6, page 89, that $a_i = b_i$, for $1 \le i \le n$; in other words: $\boldsymbol{v} = \boldsymbol{v}'$.



***T* is onto:** For $\boldsymbol{X} = a_1\boldsymbol{e}_1 + \ldots + a_n\boldsymbol{e}_n \in \Re^n$:

$$T(a_1\boldsymbol{v}_1 + \ldots + a_n\boldsymbol{v}_n) = a_1T(\boldsymbol{v}_1) + \ldots + a_nT(\boldsymbol{v}_n)$$

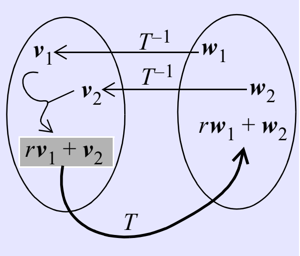
$$= a_1\boldsymbol{e}_1 + \ldots + a_n\boldsymbol{e}_n = \boldsymbol{X}$$

CHECK YOUR UNDERSTANDING 4.14

Let V and W be finite-dimensional vector spaces. Show that $V \cong W$ if and only if $\dim(V) = \dim(W)$.

Answer: See page B-15.

A ROSE BY ANY OTHER NAME



Let $T: V \to W$ be an isomorphism. Being a bijection it links every element in V with a unique element in W (every element in V has its own W-counterpart, and vice versa). Moreover, if you know how to function algebraically in V, then you can also figure out how to function algebraically in W. Suppose, for example, that you forgot how to add or scalar multiply in the space W, but remember how to add and scalar multiply in V. To figure out $r\boldsymbol{w}_1 + \boldsymbol{w}_2$ in W you can take the "T^{-1} bridge" back to V and find **the** vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ such that $T(\boldsymbol{v}_1) = \boldsymbol{w}_1$ and $T(\boldsymbol{v}_2) = \boldsymbol{w}_2$. Do the calculations $r\boldsymbol{v}_1 + \boldsymbol{v}_2$ in V and then take the "T bridge" back to W to find the value of the vector $r\boldsymbol{w}_1 + \boldsymbol{w}_2$, for it coincides with the vector $T(r\boldsymbol{v}_1 + \boldsymbol{v}_2)$:

$$T(r\boldsymbol{v}_1 + \boldsymbol{v}_2) = rT(\boldsymbol{v}_1) + T(\boldsymbol{v}_2) = r\boldsymbol{w}_1 + \boldsymbol{w}_2$$

Indeed, the intimacy between isomorphic vector spaces is so great that isomorphic spaces are said to be equal up to an isomorphism. Basically, if a vector space W is isomorphic to V, then the two spaces can only differ from each other in "appearance." For example, $M_{2\times 2}$ looks different than $\Re^4$, but you can easily link its elements with those of V:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \xleftrightarrow{\text{link}} (a, b, c, d)$$

And that linkage preserving the algebraic structure:

$$r\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \xleftrightarrow{\quad\text{link}\quad} r(a, b, c, d) + (a', b', c', d')$$

We note that all vector space properties are preserved under an isomorphisms. In particular, as you are asked to verify in the exercises, a linear map $T: V \to W$ maps:

Linearly independent sets in V to linearly independent sets in W.
And it maps spanning sets in V to spanning sets in W.

CHECK YOUR UNDERSTANDING 4.15

Let $L: V \to W$ be an isomorphism. Prove that if $\{v_1, v_2, \dots, v_n\}$ is a basis for V, then $\{L(v_1), L(v_2), \dots, L(v_n)\}$ is a basis for W.

Answer: See page B-16.

At times, one can take advantage of established properties of Euclidean spaces to address issues in other vector spaces:

EXAMPLE 4.10

Find a subset of the set:

$$S = \left\{ \begin{bmatrix} 1 & 2 \\ 3 & -2 \end{bmatrix}, \begin{bmatrix} -3 & 2 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} 11 & -2 \\ 3 & -13 \end{bmatrix}, \begin{bmatrix} 3 & 6 \\ 9 & -6 \end{bmatrix}, \begin{bmatrix} -6 & 14 \\ -6 & 4 \end{bmatrix} \right\}$$

which is a basis for $\text{Span}(S)$.

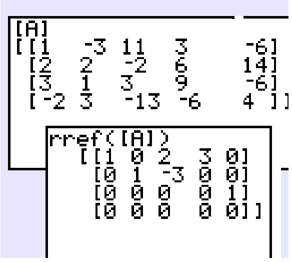
SOLUTION: Lets move the elements of S over to $\Re^4$ via the isomorphism $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a, b, c, d)$:

$$T(S) = \left\{ \begin{matrix} (1, 2, 3, -2), (-3, 2, 1, 3), (11, -2, 3, -13) \\ (3, 6, 9, -6), (-6, 14, -6, 4) \end{matrix} \right\}$$

For $f: X \to Y$, and $S \subseteq X$:
$f(S) = \{f(s) | s \in S\}$

Applying Theorem 3.13, page 103, to the vectors in $T(S)$ we see that the first, second, and fifth vector in $T(S)$ constitute a basis for $\text{Span}[T(S)]$:

$$\begin{bmatrix} 1 & -3 & 11 & 3 & -6 \\ 2 & 2 & -2 & 6 & 14 \\ 3 & 1 & 3 & 9 & -6 \\ -2 & 3 & -13 & -6 & 4 \end{bmatrix} \xrightarrow{\textbf{rref}} \begin{bmatrix} \mathbf{1} & 0 & 2 & 3 & 0 \\ 0 & \mathbf{1} & -3 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$



```
[A]
[[1   -3 11  3     -6]
 [2   2  -2  6     14]
 [3   1  3   9     -6]
 [-2 3   -13 -6    4 ]]
rref([A])
  [[1 0 2  3 0]
   [0 1 -3 0 0]
   [0 0 0  0 1]
   [0 0 0  0 0]]
```

Utilizing the result of the CYU 4.16, we conclude that the corresponding first, second, and fifth matrix in S constitute a basis for $\text{Span}(S)$.

CHECK YOUR UNDERSTANDING 4.16

Proceed as in Example 4.10 to find a subset of the set

$$S = \{2x^3 - 3x^2 + 5x - 1, x^3 - x^2 + 8x - 3, x^2 + 11x - 5, -x^3 + 2x^2 + 3x - 2\}$$

in P_3 which is a basis for Span (S).

Answer: See page B-16.

Let f be a bijection from a vector space V (with addition denoted by $\boldsymbol{v_1} + \boldsymbol{v_2}$ and scalar multiplication by $r\boldsymbol{v}$) to a **set** X. Just as the bijection f can be thought of as simply "renaming" each $\boldsymbol{v} \in V$ with its counterpart in X, so then can f be used to "carry" the vector space structure of V onto the set X; specifically:

THEOREM 4.16 Let f be a bijection from a vector space V to a set X. With addition and scalar multiplication on X defined as follows:

$$x_1 \oplus x_2 \overset{(*)}{=} f[\underset{\nwarrow}{f^{-1}}(x_1) + \underset{\nearrow}{f^{-1}}(x_2)] \quad \text{and} \quad r \otimes x \overset{(**)}{=} f[\underset{\nearrow}{rf^{-1}}(x)]$$

Go back to V and and sum in V.

Then carry the sum back to X.

Similarly

the **set X evolves into a vector space**. Moreover f itself turns out to be an isomorphism from the vector space V to the vector space X.

PROOF: The set X is clearly closed with respect to both (*) and (**) operations. We will content ourselves by verifying the zero and inverse axioms of Definition 2.6, page 40:

$f(\mathbf{0})$ turns out to be the zero in X.

Zero Axiom. Let $\mathbf{0}$ be the zero vector in V. We show that $f(\mathbf{0})$ is the zero vector in X.

For $x \in X$, let $\boldsymbol{v}$ be the vector in V such that $f(\boldsymbol{v}) = x$. Then:

$$\begin{aligned} f(\mathbf{0}) \oplus x = f(\mathbf{0}) \oplus f(\boldsymbol{v}) &= f[f^{-1}(f(\mathbf{0})) + f^{-1}(f(v))] \\ &= f(\mathbf{0} + \boldsymbol{v}) = f(\boldsymbol{v}) = x \end{aligned}$$

$f(-v)$ turns out to be the inverse of $f(\boldsymbol{v})$.

Inverse Axiom: For $x \in X$, let $\boldsymbol{v}$ be the vector in V such that $f(\boldsymbol{v}) = x$. We show that $f(-\boldsymbol{v})$ is the inverse of $x = f(\boldsymbol{v})$:

$$\begin{aligned} x \oplus f(-\boldsymbol{v}) = f(\boldsymbol{v}) \oplus f(-\boldsymbol{v}) &= f[f^{-1}(f(v)) + f^{-1}(f(-v))] \\ &= f[v + (-v)] = f(\mathbf{0}) \leftarrow \text{the zero in } X \end{aligned}$$

In the exercises you are invited to verify that the remaining axioms of Definition 2.6 are also satisfied, thereby establishing the fact that X with the above specified addition and scalar multiplication is indeed a **vector space**. We now show that the given bijection f from the vector space V to the above vector space X is an isomorphism. Actually, since f was given to be a bijection, we need only establish the linearity of f. Let's do it:

For $\boldsymbol{v_1}, \boldsymbol{v_2} \in V$, let $x_1 = f(\boldsymbol{v_1})$ and $x_2 = f(v_2)$. Then, for any $r \in \Re$:

$$f(r\boldsymbol{v_1} + \boldsymbol{v_2}) \underset{\uparrow}{=} r \otimes x_1 \oplus x_2 \underset{\uparrow}{=} rf(\boldsymbol{v_1}) + f(\boldsymbol{v_2})$$

by (*) and (**)

EXAMPLE 4.11

Let $\Re^2$ denote the Euclidean two-space, and let X be the set $X = \{(x, y) | x, y \in R\}$.

(a) Show that the function f: $\Re^2 \to X$ given by $f(x, y) = (x - 1, x + y)$ is a bijection, and find its inverse f^{-1}: $X \to \Re^2$.

(b) Determine the vector space structure on X induced by the function f, as is described in Theorem 4.16.

(c) Identify the zero in the above vector space X, and the inverse of the vector $(x, y) \in X$.

Solution: (a) f is **one-to-one**:

$$f(x_1, y_1) = f(x_2, y_2)$$

$$(x_1 - 1, x_1 + y_1) = (x_2 - 1, x_2 + y_2) \Rightarrow \begin{cases} x_1 - 1 = x_2 - 1 \\ x_1 + y_1 = x_2 + y_2 \end{cases} \Rightarrow \begin{cases} x_1 = x_2 \\ y_1 = y_2 \end{cases}$$

f **is onto:** Let $(x, y) \in X$. Then:

$$f(x + 1, y - x - 1) = (x + 1 - 1, x + 1 + (y - x - 1)) = (x, y)$$

From the above we can easily see that:

$$f^{-1}(x, y) = (x + 1, y - x - 1)$$

(b) Theorem 4.16 assures us that the set X achieves "vector space-hood" when it is augmented with the following operations:

$$(x_1, y_1) \oplus (x_2, y_2) = f[(x_1 + 1, y_1 - x_1 - 1) + (x_2 + 1, y_2 - x_2 - 1)]$$

$$= f(x_1 + x_2 + 2, y_1 + y_2 - x_1 - x_2 - 2)$$

$f(x, y) = (x - 1, x + y)$: $$= (x_1 + x_2 + 1, y_1 + y_2)$$

$$r \otimes (x, y) = f[r(x + 1, -x + y - 1)]$$

$$= f(rx + r, -rx + ry - r) = (rx + r - 1, ry)$$

(c) Theorem 4.16 assures us that $f(x, y) = (x - 1, x + y)$ is a linear map (in fact an isomorphism). That being the case, $f(0, 0) = (-1, 0)$ must be the zero vector in X. Less there be any doubt:

$$(-1, 0) \oplus (x, y) = f[f^{-1}(-1, 0) + f^{-1}(x, y)]$$

$f^{-1}(x, y) = (x + 1, y - x - 1)$: $$= f[(0, 0) + (x + 1, -x + y - 1)]$$

$$= f(x + 1, -x + y - 1) \underset{\substack{\uparrow \\ f(x, y) = (x - 1, x + y)}}{=} (x, y)$$

As for the inverse of $(x, y) \in X$:

$$-(x, y) = f[-f^{-1}(x, y)] = f[-(x+1, y-x-1)]$$

$$f^{-1}(x, y) = (x+1, y-x-1): \quad = f(-x-1, -y+x+1) \underset{\substack{\uparrow \\ f(x, y) = (x-1, x+y)}}{=} (-x-2, -y)$$

Let's challenge the above formula with the vector $(3, 2) \in X$. The formulas tells us that $-(3, 2) = (-3-2, -2) = (-5, -2)$. If that is correct, then $(3, 2) \oplus (-5, -2)$ has to be the zero vector $(-1, 0)$, and it is:

$$\begin{aligned}(3, 2) \oplus (-5, -2) &= f[f^{-1}(3, 2) + f^{-1}(-5, -2)] \\ &= f[(4, -2) + (-4, 2)] = f(0, 0) = (-1, 0)\end{aligned}$$

CHECK YOUR UNDERSTANDING 4.17

Let $\Re^3$ denote the Euclidean three-space, and let X be the set $X = \{(x, y, z) | x, y \in R\}$.

(a) Show that the function $f: \Re^3 \to X$ given by $f(x, y, z) = (2x, x+z, -z)$ is a bijection, and find its inverse.

(b) Determine the vector space structure on X induced by the function f, as is described in Theorem 4.16.

(c) Identify the zero in the above vector space X, and the inverse of the vector $(x, y, z) \in X$.

Answer: See page B-16.

It can be shown that there exists a bijection from $\Re^n$ to R for any positive integer n. Consequently:

THEOREM 4.17 Every Euclidean vector space $\Re^n$ (and therefore every finite dimensional vector space) sits (isomorphically) in the **set** R of real numbers.

PROOF: A direct consequence of Theorem 4.16.

EXERCISES

Exercises 1-6. Determine if the given linear function f is a bijection. If so, find its inverse f^{-1} and show directly that it is also linear.

1. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \boldsymbol{-5x}$.
2. $f\colon \Re \to \Re^2$, where $f(\boldsymbol{x}) = (\boldsymbol{x, -x})$.
3. $f\colon \Re^2 \to \Re^2$, where $f(\boldsymbol{a, b}) = (\boldsymbol{-2b, a})$.
4. $f\colon \Re^2 \to P_1$, where $f(\boldsymbol{a, b}) = ax + b$.
5. $f\colon \Re^3 \to P_2$, where $f(\boldsymbol{a, b, c}) = bx^2 + cx - a$.
6. $f\colon \Re^2 \to M_{2 \times 2}$, where $f(\boldsymbol{a, b}) = \begin{bmatrix} a & b \\ a+b & a-b \end{bmatrix}$.

Exercises 7-17. Determine if the given function is an isomorphism.

7. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \boldsymbol{-5x}$.
8. $f\colon \Re \to \Re$, where $f(\boldsymbol{x}) = \boldsymbol{x + 1}$.
9. $f\colon \Re^2 \to P_1$, where $f(\boldsymbol{a, b}) = (a + b)x$.
10. $f\colon \Re^2 \to P_1$, where $f(\boldsymbol{a, b}) = ax + b$.
11. $f\colon \Re^3 \to P_2$,where $f(\boldsymbol{a, b, c}) = cx^2 + bx - a$.
12. $f\colon P_2 \to P_2$, where $f[p(x)] = p(x + 1)$.
13. $f\colon P_2 \to P_3$, where $f[p(x)] = xp(x)$.
14. $f\colon P_2 \to \Re^3$, where $f[p(x)] = [\boldsymbol{p(1), p(2), p(3)}]$.
15. $f\colon P_3 \to P_3$, where $f[p(x)] = p(x) + p(1)$.
16. $f\colon P_3 \to M_{2 \times 2}$, where $f[p(x)] = \begin{bmatrix} p(1) & p(2) \\ p(3) & p(4) \end{bmatrix}$.
17. $f\colon \Re^4 \to M_{2 \times 2}$, where $f(\boldsymbol{a, b, c, d}) = \begin{bmatrix} 2a & c \\ d & b \end{bmatrix}$.

Exercises 18-22. (Calculus dependent) Determine if the given function is an isomorphism.

18. $f\colon P_2 \to P_2$, where $f[p(x)] = p'(x)$.

19. $f\colon P_2 \to P_2$, where $f[p(x)] = p'(x) + p(x)$.

20. $f\colon P_2 \to P_2$, where $f[p(x)] = p''(x)$.

21. $f\colon P_2 \to P_2$, where $f[p(x)] = 2p(x) - 3$.

22. $f\colon P_2 \to P_2$, where $f[p(x)] = p(x) + \int_0^1 p(x)dx$.

Exercises 23-24. As is done in Example 4.11, show that the given function f is a bijection and find its inverse. Determine the vector space structure on the given set X induced by f and identify the zero and the inverse of the vector $(x, y) \in X$ in the resulting space X.

23. $f\colon \Re^2 \to X = \{(x, y) | x, y \in R\}$ given by $f(x, y) = (x + y + 3, x - 4)$.

24. $f\colon \Re^3 \to X = \{(x, y, z) | x, y, z \in R\}$ given by $f(x, y, z) = (x + 1, 2y - 2, 3z + 4)$.

25. Show that if the functions $f\colon X \to Y$ and $g\colon Y \to Z$ have inverses, then the function $g \circ f\colon X \to Z$ also has an inverse and that $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

26. For $r \in \Re$, let $f_r\colon V \to V$ be given by $f_r(\boldsymbol{v}) = r\boldsymbol{v}$. For what values of r is f_r an isomorphism?

27. For $\boldsymbol{v}_0$ a vector in the space V let $f_{\boldsymbol{v}_0}\colon V \to V$ be given by $f_{\boldsymbol{v}_0}(\boldsymbol{v}) = \boldsymbol{v} + \boldsymbol{v}_0$.

    (a) Show that $f_{\boldsymbol{v}_0}$ is a bijection.

    (b) Give necessary and sufficient conditions for $f_{\boldsymbol{v}_0}$ to be an isomorphism.

28. Find a specific isomorphism from $V = \{ax^3 + (a + b)x + c | a, b, c \in \Re\}$ to $\Re^3$.

29. Show that the vector space $\Re^+$ of Example 2.4, page 46, is isomorphic to the vector space of real numbers, $\Re$.

30. Find an isomorphism between the vector space of Example 2.5, page 47 and $\Re^2$.

31. Suppose that a linear transformation $T\colon V \to W$ is one-to-one, and that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a linearly independent subset of V. Show that $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a linearly independent subset of W. (In particular, the above holds if T is an isomorphism.)

32. Suppose that a linear transformation $T\colon V \to W$ is onto, and that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a spanning set for V. Show that $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a spanning set for W. (In particular, the above holds if T is an isomorphism.)

33. Prove that a linear transformation $T: V \to W$ is an isomorphism if and only if for any given basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ for V, $\{T(\boldsymbol{v}_1), T(\boldsymbol{v}_2), \ldots, T(\boldsymbol{v}_n)\}$ is a basis for W.

34. Let V be a vector space of dimension n, and let $L(V, \Re)$ be the vector space of linear transformations from V to $\Re$ (see Exercise 35, page 122). Prove that $L(V, \Re)$ is also of dimension n and is therefore isomorphic to V. (The space $L(V, \Re)$ is called the **dual space of *V*.**)

    Suggestion: For $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis for V, show that $\{\boldsymbol{T}_1, \boldsymbol{T}_2, \ldots, \boldsymbol{T}_n\}$ is a basis for $L(V, \Re)$, where $\boldsymbol{T}_i$ is the linear transformation given by: $\boldsymbol{T}_i(\boldsymbol{v}_j) = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$.

35. Let V be a vector space of dimension n, and let W be a vector space of dimension m. Let $L(V, W)$ be the vector space of linear transformations from V to W (see Exercise 35, page 122). Prove that $L(V, \Re) \cong M_{m \times n}$.

    Suggestion: For $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ a basis for V, and $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ a basis for W, show that $\{\boldsymbol{T}_{ij}\}$ is a basis for $L(V, W)$, where $\boldsymbol{T}_{ij}$ is the linear transformation given by: $\boldsymbol{T}_{ij}(v_j) = \begin{cases} w_j \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$.

Exercises 37-39. (Equivalence Relation) A **relation** on a set X is a set $S = \{(a, b) | a, b \in X\}$. If $(a, b) \in S$, then we will say that a is related to b, and write $a \sim b$. An **equivalence relation** on a set X is a relation which satisfies the following three properties:

(i) Reflexive property: $a \sim a$.

(ii) Symmetric property: If $a \sim b$, then $b \sim a$.

(iii) Transitive property: If $a \sim b$ and $b \sim c$, then $a \sim c$.

36. A **partition** of a set X is a collection of mutually disjoint (nonempty) subsets of X whose union equals X. (In words: a partition breaks the set X into disjoint pieces.)

    (a) Let $\sim$ be an equivalence relation of a set X, and for any given $x \in X$ define the equivalence class of x to be the set of all elements of X that are related to x: $[x] = \{x' \in X | x \sim x'\}$. Prove that $\{[x] | x \in X\}$ is a partition of X.

    (b) Let S be a partition of a set X. Prove that there exists an equivalence relation $\sim$ on X such that the $S = \{[x] | x \in X\}$, where $[x] = \{x' \in X | x \sim x'\}$.

37. Show that the relation defined by $\frac{a}{b} \sim \frac{c}{d}$ if and only if $ad = bc$ is an equivalence relation on the set Q of rational numbers ("fractions").

38. Show that the relation $V \sim W$ if the vector space V is isomorphic to the vector space W is an equivalence relation on any set of vector spaces.

PROVE OR GIVE A COUNTEREXAMPLE

39. (a) If $f\colon X \to Y$ is an onto function, then so is the function $g \circ f\colon X \to Z$ onto for any function $g\colon Y \to Z$.

    (b) If $g\colon Y \to Z$ is an onto function, then so is the function $g \circ f\colon X \to Z$ onto for any function $f\colon X \to Y$.

40. (a) Let $f\colon X \to Y$ and $g\colon Y \to Z$. If $g \circ f\colon X \to Z$ is onto, then f must also be onto.

    (b) Let $f\colon X \to Y$ and $g\colon Y \to Z$. If $g \circ f\colon X \to Z$ is onto then g must also be onto.

41. (a) If $f\colon X \to Y$ is a one-to-one function, then so is the function $g \circ f\colon X \to Z$ one-to-one for any function $g\colon Y \to Z$.

    (b) If $g\colon Y \to Z$ is a one-to-one function, then so is the function $g \circ f\colon X \to Z$ one-to-one for any function $f\colon X \to Y$.

42. If $T\colon V \to W$ and $L\colon V \to W$ are isomorphisms, then $T = L$.

43. If $T\colon V \to W$ is an isomorphism, and if $a \neq 0$, then $T_a\colon V \to W$ given by $T_a(\boldsymbol{v}) = aT(\boldsymbol{v})$ is also an isomorphism.

44. Let $T\colon V \to W$ and $L\colon W \to Z$ be linear. If $g \circ f\colon V \to Z$ is an isomorphism, then T and W must both be isomorphisms.

45. If $T\colon V \to W$ and $L\colon V \to W$ are isomorphisms, then so is the function $T + L\colon V \to W$ given by $(T + L)\boldsymbol{v} = T(\boldsymbol{v}) + L(\boldsymbol{v})$ an isomorphism.

CHAPTER SUMMARY

LINEAR TRANSFORMATION	A function $T: V \to W$ from a vector space V to a vector space W is said to be a **linear transformation** if for all $\mathbf{v}, \mathbf{w} \in V$ and $r \in \Re$: $T(\mathbf{v}+\mathbf{w}) = T(\mathbf{v}) + T(\mathbf{w})$ and $T(r\mathbf{v}) = rT(\mathbf{v})$
The two conditions for linearity can be incorporated into one statement.	$T: V \to W$ is linear if and only if: $T(r\mathbf{v}+\mathbf{w}) = rT(\mathbf{v}) + T(\mathbf{w})$ for all $\mathbf{v}, \mathbf{w} \in V$ and $r \in \Re$.
The above result can be extended to encompass n-vectors and scalars.	Let $T: V \to W$ be linear. For any vectors $\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}$ in V, and any scalars $\mathbf{a_1}, \mathbf{a_2}, \dots, \mathbf{a_n}$: $T(a_1\mathbf{v_1} + \dots + a_n\mathbf{v_n}) = a_1T(\mathbf{v_1}) + \dots + a_nT(\mathbf{v_n})$
Linear transformations map zeros to zeros and inverses to inverses.	If $T: V \to W$ is linear, then: $T(\mathbf{0}) = \mathbf{0}$ and $T(-\mathbf{v}) = -T(\mathbf{v})$
A linear transformation is completely determined by its action on a basis.	Let V be a finite dimensional space with basis $\{\mathbf{v_1}, \mathbf{v_2}, \dots, \mathbf{v_n}\}$. If $T: V \to W$ and $L: V \to W$ are linear maps such that $T(\mathbf{v_i}) = L(\mathbf{v_i})$ for $1 \le i \le n$, then $T(\mathbf{v}) = L(\mathbf{v})$ for every $\mathbf{v} \in V$.
A method for constructing all linear transformations from a finite dimensional vector space to any other vector space.	Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be a basis for a vector space V, and let $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w_n}$ be n arbitrary vectors (not necessarily distinct) in a vector space W. There is then a unique linear transformation $L: V \to W$ which maps $\mathbf{v}_i$ to $\mathbf{w}_i$ for $1 \le i \le n$; and it is given by: $L(a_1\mathbf{v}_1 + \dots + a_n\mathbf{v}_n) = a_1\mathbf{w}_1 + \dots + a_n\mathbf{w}_n$
The composition of linear maps is linear.	If $T: V \to W$ and $L: W \to Z$ are linear, then the composition $L \circ T: V \to Z$ is also linear.
KERNEL	Let $T: V \to W$ be linear. The **kernel** (or **null space**) of T is denoted by $\mathrm{Ker}(T)$ and is defined by: $\mathrm{Ker}(T) = \{\mathbf{v} \in V \mid T(\mathbf{v}) = \mathbf{0}\}$
IMAGE	The **image** of T is denoted by $\mathrm{Im}(T)$ and is defined by: $\mathrm{Im}(T) = \{\mathbf{w} \in W \mid T(\mathbf{v}) = \mathbf{w} \text{ for some } \mathbf{v} \in V\}$
Both the kernel and image of a linear transformation are subspaces.	Let $T: V \to W$ be linear. Then: $\mathrm{Ker}(T)$ is a subspace of V $\mathrm{Im}(T)$ is a subspace of W

NULLITY	Let $T: V \to W$ be linear. The dimension of $\mathrm{Ker}(T)$ is called the **nullity** of T, and is denoted by $\mathrm{nullity}(T)$.
RANK	The dimension of $\mathrm{Im}(T)$ is called the **rank** of T, and is denoted by $\mathrm{rank}(T)$.
The Dimension Theorem.	Let V be a vector space of dimension n, and let $T: V \to W$ be linear. Then: $\mathrm{rank}(T) + \mathrm{nullity}(T) = n$
ONE-TO-ONE	A function f from a set A to a set B is said to be **one-to-one** if: $f(a) = f(a') \Rightarrow a = a'$
ONTO	A function f from a set A to a set B is said to be **onto** if for every $b \in B$ there exist $a \in A$ such that $f(a) = b$.
	A linear transformation $T: V \to W$ is one-to-one if and only if $\mathrm{Ker}(T) = \mathbf{0}$. A linear transformation $T: V \to W$ is onto if and only if $\mathrm{Im}(T) = W$.
BIJECTION	A function f: A $\to B$ that is both one-to-one and onto is said to be a **bijection**.
The composite of bijections is again a bijection.	If f: A $\to B$ and $g: B \to C$ are bijections, then the composite function $g \circ f: A \to C$ is also a bijection.
INVERSE FUNCTION	The inverse of a bijection $f: X \to Y$ is the function $f^{-1}: Y \to X$ such that $(f^{-1} \circ f)(x) = x$ for every x in X, and $(f \circ f^{-1})(y) = y$ for every y in Y.
The inverse of a linear bijection is again linear.	If the bijection $T: V \to W$ is linear, then its inverse, $T^{-1}: W \to V$ is also linear.
ISOMORPHISM	A bijection $T: V \to W$ that is also a linear transformation is said to be an **isomorphism** from the vector space V to the vector space W. If there exists an isomorphism from V to W we then say that V and W are **isomorphic**, and write: $V \cong W$.
Every vector space is isomorphic to itself. If V is isomorphic to W, then W is isomorphic to V. If V is isomorphic to W, and W is isomorphic to Z, then V is isomorphic to Z.	$V \cong V$ $V \cong W \Rightarrow W \cong V$ $V \cong W$ and $W \cong Z \Rightarrow V \cong Z$
All n-dimensional vector spaces are isomorphic to Euclidean n-space.	If $\dim(V) = n$, then $V \cong \boldsymbol{R}^n$.

CHAPTER 5

MATRICES AND LINEAR MAPS

It turns out that linear transformations can, in a sense, be represented by matrices. Such representations are developed and scrutinized in Sections three and four — a development that rests on the matrix theory presented in the first two sections of the chapter.

§1. MATRIX MULTIPLICATION

The matrix space $M_{m \times n}$ of Theorem 2.2, page 42, comes equipped with a scalar multiplication. We now turn our attention to another form of multiplication, one that involves a pair of matrices (instead of a scalar and a matrix).

Left to one's own devices, one would probably define the product of matrices in the following fashion:

Take two matrices of equal dimension, and simply multiply corresponding entries to obtain their product.	As with: $\begin{bmatrix} 2 & 3 \\ 5 & 4 \end{bmatrix}\begin{bmatrix} 5 & 3 \\ 0 & 6 \end{bmatrix} = \begin{bmatrix} 10 & 9 \\ 0 & 24 \end{bmatrix}$

The above might be "natural," but as it turns out, not very useful. Here, as you will see, is a most useful definition:

In general, we will use: $A_{m \times n} = [a_{ij}]$ or $[a_{ij}]_{m \times n}$ to denote an m by n matrix with entries a_{ij}.

DEFINITION 5.1
MATRIX MULTIPLICATION

If $A_{m \times r} = [a_{ij}]$ and $B_{r \times n} = [b_{ij}]$, then:

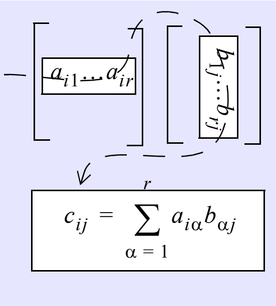
$$\boldsymbol{A_{m \times r} B_{r \times n}} = \boldsymbol{C_{m \times n}} = [c_{ij}]$$

where:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj}$$

Using Sigma notation, we have:

$$c_{ij} = \sum_{\alpha = 1}^{r} a_{i\alpha}b_{\alpha j}$$



IN WORDS: To get $\boldsymbol{c_{ij}}$ of $C = AB$, run across the i^{th} row of A and down the j^{th} column of B, multiplying and adding along the way (see margin).

Note: The above is meaningful only if the number of columns of the matrix on the left equals the number of rows of the matrix on the right.

EXAMPLE 5.1 Find the product $C = AB$, if:

$$A = \begin{bmatrix} 2 & 0 & 1 \\ 5 & 2 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 0 & 4 & 3 \\ 1 & 5 & 3 & 2 \end{bmatrix}$$

SOLUTION: Since the number of columns of A equals the number of rows of B, the product AB is defined:

$$A_{2\times \underbrace{3}} B_{\underbrace{3}\times 4} = C_{2\times 4}$$

same

Below, we illustrate the process leading to the value of c_{11} (run across the first row of A and down the first column of B), and for c_{23} (run across the second row of A and down the third column of B):

$$\begin{bmatrix} \mathbf{2} & \mathbf{0} & \mathbf{1} \\ 5 & 2 & 4 \end{bmatrix}\begin{bmatrix} \mathbf{1} & 2 & 3 & 5 \\ \mathbf{2} & 0 & 4 & 3 \\ \mathbf{1} & 5 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{3} & & & \\ & & & \end{bmatrix} \qquad \mathbf{2\cdot 1+0\cdot 2+1\cdot 1}$$

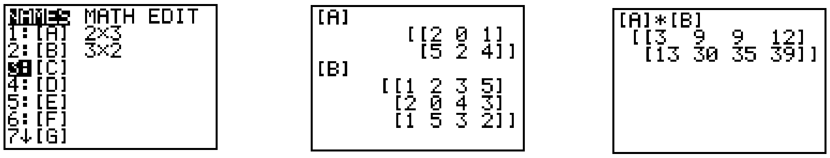
$$\begin{bmatrix} 2 & 0 & 1 \\ \mathbf{5} & \mathbf{2} & \mathbf{4} \end{bmatrix}\begin{bmatrix} 1 & 2 & \mathbf{3} & 5 \\ 2 & 0 & \mathbf{4} & 3 \\ 1 & 5 & \mathbf{3} & 2 \end{bmatrix} = \begin{bmatrix} 3 & & & \\ & & \mathbf{35} & \end{bmatrix} \qquad \mathbf{5\cdot 3+2\cdot 4+4\cdot 3}$$

At this point, we are confident that you can find the remaining entries:

$$\begin{bmatrix} 2 & 0 & 1 \\ 5 & 2 & 4 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 & 5 \\ 2 & 0 & 4 & 3 \\ 1 & 5 & 3 & 2 \end{bmatrix} = \begin{bmatrix} \mathbf{3} & 9 & 9 & 12 \\ 13 & 30 & \mathbf{35} & 39 \end{bmatrix}$$

In any event:

GRAPHING CALCULATOR GLIMPSE 5.1



```
NAMES MATH EDIT
1:[A] 2×3
2:[B] 3×2
3:[C]
4:[D]
5:[E]
6:[F]
7↓[G]
```

```
[A]
        [[2 0 1]
         [5 2 4]]
[B]
      [[1 2 3 5]
       [2 0 4 3]
       [1 5 3 2]]
```

```
[A]*[B]
 [[3  9  9  12]
  [13 30 35 39]]
```

CHECK YOUR UNDERSTANDING 5.1

(a) Perform the product AB for:

$$A = \begin{bmatrix} 3 & 5 \\ 4 & 2 \\ 9 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 6 & 4 \\ 3 & 5 \end{bmatrix}$$

(b) Explain why the product BA is not defined.

Answers:

(a) $\begin{bmatrix} 33 & 37 \\ 30 & 26 \\ 54 & 36 \end{bmatrix}$

(b) See page B-18.

In some ways, matrices behave differently than numbers. For one thing, even when both products AB and BA are defined, as is the case when the matrices A and B are square matrices of the same dimension, those products need not be equal:

Matrix multiplication is **not** commutative.

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix} \quad \text{while} \quad \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 23 & 34 \\ 31 & 46 \end{bmatrix}$$

Some familiar multiplication properties do however carry over to matrices:

THEOREM 5.1 Assuming that the matrix dimensions are such that the given operations are defined, we have:

Distributive Properties
(i) $A(B+C) = AB+AC$
(ii) $(A+B)C = AC+BC$

Associative Properties
(iii) $A(BC) = (AB)C$
(iv) $r(AB) = (rA)B = A(rB)$

PROOF: We establish (i): $A(B+C) = AB+AC$, and invite you to verify the rest in the exercises. Let:

$$A_{m \times r} = [a_{ij}] \qquad B_{r \times n} = [b_{ij}] \quad C_{r \times n} = [c_{ij}]$$

$$D_{m \times n} = A(B+C) = [d_{ij}] \qquad E_{m \times n} = AB+AC = [e_{ij}]$$

The properties of this theorem are not particularly difficult to establish. The trick is to carefully keep track of the entries of the matrices along the way,

We need to show that $d_{ij} = e_{ij}$. Let's do it:

$$d_{ij} = \sum_{t=1}^{r} a_{it}(b_{tj}+c_{tj}) \overset{\text{distribute}}{=} \sum_{t=1}^{r} a_{it}b_{tj} + \sum_{t=1}^{r} a_{it}c_{tj} = e_{ij}$$

run across the i^{th} row of A and down the j^{th} column of $B+C$

across the i^{th} colum of A down the j^{th} column of B

across the i^{th} colum of A down the j^{th} column of C

POWERS OF SQUARE MATRICES

If A is a square matrix, what should A^2 represent? That's right: $A^2 = AA$. In a more general setting, we have:

Why we are restricting this discussion to square matrices? Because:

$$\begin{bmatrix} 2 & 1 \\ 3 & 5 \\ 4 & 2 \end{bmatrix}^2 = ?$$

DEFINITION 5.2
POWERS

For A a square matrix:

$$A^2 = AA, A^3 = AAA, \text{ and } A^n = AA^{n-1}$$

While the definition of A^n mimics that of a^n for $a \in R$, not all of the familiar properties of exponents hold for matrices, even when the matrix expressions are well-defined. The property $(ab)^n = a^n b^n$, for example, does **not** carry over to matrices:

$$\left(\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}\right)^2 = \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix}^2 = \begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 20 & 8 \\ 8 & 4 \end{bmatrix}$$

$$\text{while: } \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}^2 \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}^2 = \left(\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\right)\left(\begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}\right)$$

$$= \begin{bmatrix} 1 & 6 \\ 0 & 4 \end{bmatrix}\begin{bmatrix} 6 & 4 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 18 & 16 \\ 8 & 8 \end{bmatrix}$$

In $\Re = M_{1 \times 1}$, $ab = ba$. Why not for $M_{2 \times 2}$?

$$\text{And we see that: } \left(\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}\right)^2 \neq \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}^2 \begin{bmatrix} 2 & 2 \\ 1 & 0 \end{bmatrix}^2$$

The following familiar properties are, however, a directly consequence of Definition 5.2:

THEOREM 5.2 For A a square matrix, and positive integers n and m:

$$A^n A^m = A^{n+m} \text{ and } (A^n)^m = A^{nm}$$

CHECK YOUR UNDERSTANDING 5.2

PROVE OR GIVE A COUNTEREXAMPLE:

For any two matrices $A, B \in M_{2 \times 2}$:

$$(A + B)^2 = A^2 + 2AB + B^2$$

Answer: See page B-18.

COLUMN AND ROW SPACES

DEFINITION 5.3
COLUMN AND ROW SPACE

The **columns space** of a matrix $A \in M_{m \times n}$, is the subspace of $\Re^m$ spanned by the columns of A.

The **row space** of a matrix $A \in M_{m \times n}$, is the subspace of $\Re^n$ spanned by the rows of A.

The following result may be a bit surprising in that the rows and columns of the matrix $A \in M_{m \times n}$ need not even live in the same Euclidean spaces: The rows of A are in $\Re^n$ while its columns are in $\Re^m$:

THEOREM 5.3 The dimension of the column space of any matrix $A \in M_{m \times n}$ equals the dimension of its row space.

PROOF: We know, from Theorem 3.13, page 103, that the dimension of the column space of $A \in M_{m \times n}$ equals the number of leading ones in $\mathrm{rref}(A)$. To see that the dimension of the row space of A is also equal to the number of leading ones in $\mathrm{rref}(A)$ we reason as follows:

> Let B be a matrix obtained by performing any of the three elementary row operations on A. Since each row in B is a linear combination of the rows in A, and vice version, the space spanned by the rows of A equals that spanned by the rows of B. Since $\mathrm{rref}(A)$ is derived from A through a sequence of elementary row operations, the row space of A equals that of $\mathrm{rref}(A)$, which is easily seen to be the non-zero rows of $\mathrm{rref}(A)$; each of which contains a leading one.

DEFINITION 5.4
RANK OF A MATRIX
For $A \in M_{m \times n}$, the rank of A, denoted by $\mathrm{rank}(A)$ is the common dimension of the row and column space of $A \in M_{m \times n}$.

EXAMPLE 5.2 Find a basis for both the row and the column space of $A = \begin{bmatrix} 3 & -9 & 1 & -5 & 6 \\ -1 & 3 & 2 & 4 & -2 \\ 2 & -6 & -2 & -6 & 4 \end{bmatrix}$.

SOLUTION: $A = \begin{bmatrix} 3 & -9 & 1 & -5 & 6 \\ -1 & 3 & 2 & 4 & -2 \\ 2 & -6 & -2 & -6 & 4 \end{bmatrix} \xrightarrow{\mathrm{rref}} \begin{bmatrix} 1 & -3 & 0 & -2 & 2 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.

A basis for the row space of A: $\{(\mathbf{1, -3, 0. -2, 2}), (\mathbf{0, 0, 1, 1, 0})\}$.

A basis for the column space of A: $\{(\mathbf{3, -1, 2}), (\mathbf{1, 2, -2})\}$
(see Theorem 3.13, page 103)

While Theorem 3.13 assures us that he columns of A associated with the leading one columns of $\mathrm{rref}(A)$ constitute a basis for the column space of A, the rows of A associated with the leading one rows of $\mathrm{rref}(A)$ need not be a basis for the row space of A. A case in point:

CHECK YOUR UNDERSTANDING 5.3

Find a basis for the column and the row space of:

$$A = \begin{bmatrix} 2 & 5 & -3 & 4 \\ -4 & -10 & 6 & -8 \\ 0 & 1 & 2 & -4 \end{bmatrix}.$$

Answer: See page B-18.

SYSTEM OF EQUATIONS REVISITED

As you know, a system of equations can take the form of an augmented matrix [see page 3, and Figure (a) and (b) below]. That system can also be represented as the product of its coefficient matrix with a (column) variable matrix [see Figure 5.1(c) and margin].

matrix multiplication

$$\begin{bmatrix} 2 & 4 & -4 \\ 2 & 6 & 4 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ -2 \end{bmatrix}$$

$$\begin{bmatrix} 2x+4y-4z \\ 2x+6y+4z \\ x+y+2z \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ -2 \end{bmatrix}$$

$$\left.\begin{aligned} 2x+4y-4z &= 6 \\ 2x+6y+4z &= 0 \\ x+y+2z &= -2 \end{aligned}\right\}$$

$$\left.\begin{aligned} 2x+4y-4z &= 6 \\ 2x+6y+4z &= 0 \\ x+y+2z &= -2 \end{aligned}\right\} \longleftrightarrow \left[\begin{array}{ccc|c} 2 & 4 & -4 & 6 \\ 2 & 6 & 4 & 0 \\ 1 & 1 & 2 & -2 \end{array}\right] \longleftrightarrow \begin{bmatrix} 2 & 4 & -4 \\ 2 & 6 & 4 \\ 1 & 1 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ -2 \end{bmatrix}$$

System of Equations (a) **Augmented Matrix** (b) **Matrix Product Form** (c)

Figure 5.1

The next result tells us that the product of $A \in M_{m \times n}$ with a column variable matrix $X_{n \times 1}$ is a linear combination of the columns of A. Specifically:

THEOREM 5.4 Let $A_{m \times n} = [a_{ij}]$. For each $1 \le j \le n$ let $C_j = [a_{ij}]_{m \times 1}$ (the j^{th} "column matrix" of A). Then, for any $X = [x_i]_{n \times 1}$:

$$AX = \sum_{i=1}^{n} x_i C_i \text{ (see margin).}$$

$$\begin{array}{c} \begin{matrix} C_1 & C_2 & & C_n \end{matrix} \\ \begin{bmatrix} a_{11} & a_{12} & & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ & & & \\ a_{m1} & a_{m2} & & a_{mn} \end{bmatrix} \end{array} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} =$$

$$x_1 \underset{C_1}{\begin{bmatrix} a_{11} \\ a_{21} \\ \\ a_{m1} \end{bmatrix}} + x_2 \underset{C_2}{\begin{bmatrix} a_{12} \\ a_{22} \\ \\ a_{m2} \end{bmatrix}} + \cdots x_n \underset{C_n}{\begin{bmatrix} a_{1n} \\ a_{2n} \\ \\ a_{mn} \end{bmatrix}}$$

PROOF: For $B = AX = [b_i]_{m \times 1}$, and $D = \sum\limits_{i=1}^{n} x_i C_i = [d_i]_{m \times 1}$:

$$b_i = \sum_{j=1}^{m} a_{ij} x_j = \sum_{j=1}^{m} x_j a_{ij} = d_i$$

Consider the following system of equations $\boldsymbol{AX} = \boldsymbol{B}$:

$$\underset{\boldsymbol{A}}{\begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 8 & 3 & -5 \\ 1 & -2 & 1 & 6 \end{bmatrix}} \underset{\boldsymbol{X}}{\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix}} = \underset{\boldsymbol{B}}{\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}} \Rightarrow \begin{bmatrix} 2x + 1y + 3z + 4w \\ 1x + 8y + 3z - 5w \\ 1x - 2y + 1z + 6w \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Invoking Theorem 5.2, we find that:

$$x\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + y\begin{bmatrix} 1 \\ 8 \\ -2 \end{bmatrix} + z\begin{bmatrix} 3 \\ 3 \\ 1 \end{bmatrix} + w\begin{bmatrix} 4 \\ -5 \\ 6 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

It follows that the given system $\boldsymbol{AX} = \boldsymbol{B}$ is consistent if and only if $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ is in the column space of A.

In general:

THEOREM 5.5 A system of linear equations $\boldsymbol{AX} = \boldsymbol{B}$ is consistent if and only if $\boldsymbol{B}$ is in the column space of A.

CHECK YOUR UNDERSTANDING 5.4

Prove that the solution set of any homogeneous system of m equations in n unknowns is a subspace of $\Re^n$.

Answer: See page B-18.

FROM MATRICES TO LINEAR TRANSFORMATIONS

Note that:
$A_{m \times n}X_{n \times z} \in M_{m \times z}$

While matrix spaces only come equipped with vector addition and scalar multiplication, matrix multiplication can serve to define a linear map between such spaces — providing their dimensions "match up:"

THEOREM 5.6 For $A \in M_{m \times n}$ and any positive integer z, the map $T_A:(M_{n \times z} \to M_{m \times z})$ given by $T_A(B) = AB$ is linear.

PROOF: For $B_1, B_2 \in M_{n \times z}$ and $r \in \Re$

$$T_A(rB_1 + B_2) = A(rB_1 + B_2) \overset{\text{Theorem 5.1(iv):}}{\overset{\downarrow}{=}} A(rB_1) + AB_2$$

$$\text{Theorem 5.1(iv):} \quad = r(AB_1) + AB_2 = rT_A(B_1) + T_A(B_2)$$

For notational convenience we will let the symbol $\overline{\Re^n}$ denote the space $M_{n \times 1}$ ("vertical n-tuples"). We then have:

Note that X is a $n \times 1$ "vertical n-tuple," and that $T_A(X)$ is a vertical m-tuple.

THEOREM 5.7 For $A \in M_{m \times n}$ the map $T_A\colon \overline{\Re^n} \to \overline{\Re^m}$ given by $T_A(X) = AX$ is linear.

PROOF: Simply set $z = 1$ in Theorem 5.6

DEFINITION 5.5
NULL SPACE

For $\boldsymbol{A} \in M_{m \times n}$, the **null space of A**, denoted by $\text{null}(A)$, is the set:

$$\text{null}(\boldsymbol{A}) = \{X \in \overline{\Re^n} \mid \boldsymbol{A}X = \mathbf{0}\}$$

THEOREM 5.8 The null space of $A \in M_{m \times n}$ is a subspace of $\overline{\Re^n}$.

PROOF: Follows from Theorem 4.8(a), page 124, and the fact that:

$$\text{null}(\boldsymbol{A}) = \text{Ker}(T_A)$$

Note: The dimension of the null space of A is called the **nullity** of A. [Approproate terminology, in that $\text{null}(\boldsymbol{A}) = \text{Ker}(T_A)$]

EXAMPLE 5.3 Find a basis for the null space of the matrix:

$$A = \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 8 & 3 & -10 \\ 1 & -2 & 1 & 6 \end{bmatrix}$$

SOLUTION: By definition, the null space of A is the solution set of the homogeneous system of equations:

$$\begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 8 & 3 & -10 \\ 1 & -2 & 1 & 6 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From:
$$\begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} 2 & 1 & 3 & 4 \\ 1 & 8 & 3 & -10 \\ 1 & -2 & 1 & 6 \end{bmatrix} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} \mathbf{1} & 0 & \frac{7}{5} & \frac{14}{5} \\ 0 & \mathbf{1} & \frac{1}{5} & -\frac{8}{5} \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

vertical 3-tuple

We see that: $\text{null}(A) = \left\{ \overline{\left(-\frac{7}{5}a - \frac{14}{5}b, -\frac{1}{5}a + \frac{8}{5}b, a, b\right)} \middle| a, b \in R \right\}$

$$= \{\overline{(-7a - 14b, -a + 8b, 5a, 5b)} \mid a, b \in R\}$$

Letting $a = 1$ and $b = 0$ we obtain the vector $\overline{(-7, -1, 5, 0)}$.

Letting $a = 0$ and $b = 1$ we arrive at the vector $\overline{(-14, 8, 0, 5)}$.

The above two vectors are clearly independent, and also span $\text{null}(A)$:

$$\overline{(-7a - 14b, -a + 8b, 5a, 5b)} = a\overline{(-7, -1, 5, 0)} + b\overline{(-14, 8, 0, 5)}$$

It follows that $\{\overline{(-7, -1, 5, 0)}, \overline{(-14, 8, 0, 5)}\}$ is a basis for $\text{null}(A)$.

We arrived at the two vectors in the basis for $\text{null}(A)$ in the above example by setting each of the two free variables to 1 and the other free variable to 0. Generalizing:

THEOREM 5.9 The nullity of $A \in M_{m \times n}$ equals the number of free variables in $\text{rref}(A)$.

CHECK YOUR UNDERSTANDING 5.5

Find a basis for:

$$\text{null}\left(\begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 4 & -2 & -7 \\ 3 & 0 & 1 & -2 \end{bmatrix}\right)$$

Answer: See page B-18.

You get to draw the final curtain of this section:

CHECK YOUR UNDERSTANDING 5.6

Let $A \in M_{m \times n}$ and let $T_A: \overline{\Re^n} \to \overline{\Re^m}$ be the linear map given by $T_A(X) = AX$. Show that:

$$\text{nullity}(A) = \text{nullity}(T_A) \text{ and } \text{rank}(A) = \text{rank}(T_A)$$

Answer: See page B-18.

EXERCISES

Exercises 1-5. Perform the given matrix operations.

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

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

3. $\begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 3 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\left(\begin{bmatrix} 0 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix} + \begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 0 \end{bmatrix}\right)\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 0 & 3 & 1 \\ 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} 2 & 1 & 1 \\ 4 & 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 1 & 3 \\ 0 & 1 \end{bmatrix}$

Exercises 5-8. Find a basis for the null space of A. Determine $\text{rank}(A)$ along with a basis for the column and row space of A.

5. $A = \begin{bmatrix} 3 & 2 & 1 & 5 \\ 10 & 6 & 1 & 10 \\ 4 & 2 & -1 & 0 \end{bmatrix}$ 6. $A = \begin{bmatrix} 2 & 0 & 1 \\ 3 & -1 & 0 \\ 6 & 3 & 3 \\ 1 & 4 & 2 \end{bmatrix}$ 7. $A = \begin{bmatrix} 1 & 1 & 0 & 1 \\ -1 & 0 & -1 & 4 \\ 5 & 3 & -2 & 7 \\ 5 & 2 & -1 & 2 \end{bmatrix}$ 8. $A = \begin{bmatrix} 1 & -1 & 5 & 5 \\ 1 & 0 & 3 & 2 \\ 0 & -1 & -2 & -1 \\ 1 & 4 & 7 & 2 \end{bmatrix}$

9. (a) Show that each column of $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 7 & 1 \\ 2 & 0 & 4 \end{bmatrix}$ (as a vertical two-tuple) is a linear combination of the columns of $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}$.

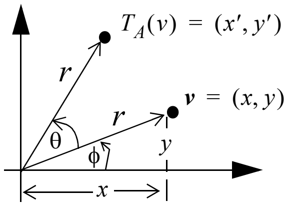
   (b) Show that each row of $\begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 3 & 7 & 1 \\ 2 & 0 & 4 \end{bmatrix}$ (as a horizontal three-tuple) is a linear combination of the rows of $\begin{bmatrix} 3 & 7 & 1 \\ 2 & 0 & 4 \end{bmatrix}$.

10. Let $A \in M_{m \times n}$ and let $X_i \in M_{n \times 1}$ be the column matrix whose i^{th} entry is 1 and all other entries are 0. Show that AX_i is the i^{th} column of A, for $1 \le i \le n$.

11. (a) (Dilation and Contraction) Let $A = \begin{bmatrix} r & 0 \\ 0 & r \end{bmatrix}$ for $r > 0$. Show that $T_A \colon \Re^2 \to \Re^2$ maps every point in the plane to a point r times a far from the origin.

    (b) (Reflection about the x-axis) Let $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. Show that $T_A \colon \Re^2 \to \Re^2$ reflects every point in the plane about the x-axis.

(c) Find the matrix A for which $T_A\colon \Re^2 \to \Re^2$ reflects the point $(\boldsymbol{x}, \boldsymbol{y})$ to the point twice the length of $(\boldsymbol{x}, \boldsymbol{y})$ and reflected about the y-axis.



(d) (Rotation about the Origin) Let $A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. Show that $T_A\colon \Re^2 \to \Re^2$ rotates the vector $(\boldsymbol{x}, \boldsymbol{y})$ by θ in a counterclockwise direction.

Suggestion: Start with $x' = r\cos(\varphi + \theta)$ and $y' = r\sin(\varphi + \theta)$.

12. Show that for any given linear transformation $T\colon \Re^n \to \Re^m$ there exists a unique matrix $A \in M_{m \times n}$ such that $T(X) = AX$.

13. Let $A, B \in M_{m \times n}$. Prove that if $AC = BC$ for every $C \in M_{n \times 1}$, then $A = B$.

14. Determine all $A \in M_{2 \times 2}$ such that $AB = BA$ for every $B \in M_{2 \times 2}$.

15. Prove Theorem 5.1(ii).

16. Prove Theorem 5.1(iii).

17. Prove Theorem 5.1(iv).

18. A square matrix $A = [a_{ij}]_{n \times n}$ for which $a_{ij} = 0$ if $i \neq j$ is said to be a **diagonal matrix**. Show that if $A = [a_{ij}]_{n \times n}$ is a diagonal matrix and if $X = [x_i] \in \Re^n$ is a column matrix, then $AX = [a_{ii}x_i] \in \Re^n$. For example: $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{bmatrix} 9 \\ 6 \\ -1 \end{bmatrix} = \begin{bmatrix} 18 \\ 30 \\ -7 \end{bmatrix}$.

19. The **transpose** of a matrix $A = [a_{ij}] \in M_{m \times n}$ is the matrix $A^{\mathrm{T}} = [\bar{a}_{ij}] \in M_{n \times m}$, where $\bar{a}_{ij} = a_{ji}$. In other words, the transpose of $\boldsymbol{A}$ is that matrix obtained by interchanging the rows and columns of A.

    (a) Show that for any matrix A, $(A^{\mathrm{T}})^{\mathrm{T}} = A$.

    (b) Show that for any matrix A and any scalar r, $(rA)^{\mathrm{T}} = rA^{\mathrm{T}}$.

    (c) Show that for any $A, B \in M_{m \times n}$, $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$.

    (d) Show that for any $A, B \in M_{m \times n}$, $(A - B)^{\mathrm{T}} = A^{\mathrm{T}} - B^{\mathrm{T}}$.

    (e) Show that for $A \in M_{n \times n}$, $(AA)^{\mathrm{T}} = A^{\mathrm{T}}A^{\mathrm{T}}$.

    (f) Show that for $A \in M_{m \times n}$ and $B \in M_{n \times r}$, $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

20. A square matrix A is **symmetric** if the transpose of A equals A: $A^{\mathrm{T}} = A$ (see Exercise 19).

    (a) Show that the set of symmetric matrices in the space $M_{n \times n}$ is a subspace of $M_{n \times n}$.

    (b) Let $A, B \in M_{m \times n}$ be symmetric matrices. Show that $A + B$ is symmetric.

    (c) Let $A \in M_{n \times n}$. Show that $A + A^{\mathrm{T}}$ is symmetric.

    (d) Let $A \in M_{n \times n}$. Show that AA^{T} is symmetric.

    (e) Let $A, B \in M_{n \times n}$ be symmetric. Show that AB is symmetric if and only if $AB = BA$

21. A square matrix A is said to be **skew-symmetric** if $A^{\mathrm{T}} = -A$ (see Exercise 19).
    - (a) Show that if the square matrix $A = [a_{ij}]$ is skew symmetric, then each diagonal entry a_{ii} must be zero.
    - (b) Show that the square matrix $A = [a_{ij}]$ is skew symmetric if and only if $a_{ij} = -a_{ji}$ for all i, j.
    - (c) Let $A \in M_{n \times n}$. Show that $A - A^{\mathrm{T}}$ is skew symmetric.
    - (d) Show that every square matrix A can be uniquely written as $A = S + K$, where S is a symmetric matrix and K is a skew symmetric matrix.

22. A matrix $A \in M_{n \times n}$ is said to be **idempotent** if $A^2 = A$.
    - (a) Find an idempotent 2×2 matrix A containing no zero entry.
    - (b) Let A and B be idempotent square matrices of the same dimension. Prove that if $AB = BA$, then AB is idempotent.

23. A matrix $A \in M_{n \times n}$ is said to be **nilpotent** if $A^k = 0$ for some integer k.
    - (a) Find a nilpotent 2×2 matrix A distinct from the zero matrix.
    - (b) Let A and B be two nilpotent matrices of the same dimension. Prove that if $AB = BA$, then AB is again nilpotent.

24. The sum of the diagonal entries in the matrix $A = [a_{ij}] \in M_{n \times n}$ is called the **trace** of A and is denoted by $\operatorname{trace}(A)$: $\operatorname{trace}(A) = a_{11} + a_{22} + \ldots a_{nn} = \sum_{i=1}^{n} a_{ii}$.
    - (a) Show that for any square matrix A: $\operatorname{trace}(A) = \operatorname{trace}(A^{\mathrm{T}})$ (see Exercise 19).
    - (b) Show that for any two square matrices A and B of the same dimension:
    $$\operatorname{Trace}(A + B) = \operatorname{Trace}(A) + \operatorname{Trace}(B)$$
    - (c) Show that for any two square matrices A and B of the same dimension:
    $$\operatorname{Trace}(AB) = \operatorname{Trace}(BA)$$

Exercises 25-30. (PMI) Determine A^n for the given matrix A. Use the Principle of Mathematical Induction to substantiate your claim.

25. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$

26. $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

27. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$

28. $\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$

29. $\begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$

30. $\begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix}$

31. **(PMI)** Show that if $AB = BA$, then for any positive integer n, $(AB)^n = A^n B^n$.

32. **(PMI)** Let $A \in M_{m \times m}$ and $B \in M_{m \times s}$. Show that if $AX = X$, then $A^n X = X$ for every positive integer n.

33. **(PMI)** Show that if the entries in each column of $A \in M_{n \times n}$ sum to 1, then the entries in each column of A^m also sum to 1, for any positive integer m.

34. **(PMI)** Show that if $A \in M_{n \times n}$ is a diagonal matrix, then so is A^n. (See Exercise 18.)

35. **(PMI)** Show that if $A \in M_{n \times n}$ is an idempotent matrix, then $A^n = A$ for all integers $n \geq 1$. (See Exercise 22.)

36. **(PMI)** Show that for any $A \in M_{n \times n}$, and for any positive integer n, $(A^n)^{\mathrm{T}} = (A^{\mathrm{T}})^n$. (See Exercise 19.)

37. **(PMI)** Let $A_i \in M_{m \times n}$, for $1 \leq i \leq n$. Show that:

$$(A_1 + A_2 + \cdots + A_n)^{\mathrm{T}} = A_1^T + A_2^T + \cdots + A_n^T. \text{ (See Exercise 19.)}$$

38. **(PMI)** Let $A_i \in M_{n \times n}$ for $1 \leq i \leq n$. show that:

$$(A_1 \cdot A_2 \cdots A_n)^{\mathrm{T}} = A_n^T \cdot A_{n-1}^T \cdots A_1^T. \text{ (See Exercise 19.)}$$

PROVE OR GIVE A COUNTEREXAMPLE

39. For $A \in M_{m \times n}$ and $B \in M_{n \times r}$, if $AB = 0$ then either $A = 0$ or $B = 0$.

40. Let A and B be two-by-two matrices with $A \neq 0$. If $AB = AC$, then $B = C$.

41. If A and B are square matrices of the same dimension, and if AB is idempotent, then $AB = BA$. (See Exercise 22.)

42. For all $A, B \in M_{2 \times 2}$, $A^2 - B^2 = (A + B)(A - B)$.

43. For any given matrix $A \in M_{2 \times 2}$, all entries in the matrix A^2 are nonnegative.

44. For $A \in M_{m \times n}$ and $B \in M_{n \times r}$, if A has a column consisting entirely of 0's, then so does AB.

45. For $A \in M_{m \times n}$ and $B \in M_{n \times r}$, if A has a row consisting entirely of 0's, then so does AB.

46. For $A \in M_{m \times n}$ and $B \in M_{n \times r}$, if A has two identical columns, then so does AB.

47. For $A \in M_{m \times n}$ and $B \in M_{n \times r}$, if A has two identical rows, then so does AB.

48. For $A \in M_{n \times n}$, $\{B \in M_{n \times n} | AB = 0\}$ is a subspace of $M_{n \times n}$.

49. For $A \in M_{n \times n}$, $\{B \in M_{n \times n} | AB = BA\}$ is a subspace of $M_{n \times n}$.

50. For $A \in M_{n \times n}$, $\{B \in M_{n \times n} | \mathrm{Trace}(AB) \geq 0\}$ is a subspace of $M_{n \times n}$. (See Exercise 24.)

51. If A is a nilpotent matrix, then so is A^2. (See Exercise 23.)

52. A is idempotnet if and only if A^{T} is idempotent. (See Exercise 22 and 19.)

§2. INVERTIBLE MATRICES

Since all identity matrices are square we can get away by specifying just one of its dimensions, as with:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

instead of $I_{3 \times 3}$.

The **identity matrix** of dimension n, denoted by I_n, is the $n \times n$ matrix that has 1's along its main diagonal (top-left corner to bottom-right corner), and 0's elsewhere. In the event that the dimension n of the identity matrix is understood, we may simply write I rather than I_n.

Just as $1 \cdot a = a$ and $a \cdot 1 = a$ for any number a, so it is that for any $A \in M_{m \times n}$: $I_m A = A$ and $AI_n = A$. In particular, for any square matrix $A \in M_{n \times n}$: $I_n A = AI_n = A$.

INVERTIBLE MATRICES

DEFINITION5.6

INVERTIBLE MATRIX

A **square matrix** $A \in M_{n \times n}$ is said to be **invertible** if there exists a matrix $B \in M_{n \times n}$ such that:

$$AB = BA = I$$

The matrix B is then said to be the **inverse** of A, and we write $B = A^{-1}$. If no such B exists, then A is said to be **non-invertible**, or **singular.**

We will soon show, that a matrix ***A*** can have but one inverse.

EXAMPLE 5.4

Determine if the matrix $A = \begin{bmatrix} -5 & 2 \\ 9 & -4 \end{bmatrix}$ is invertible. If it is, find its inverse.

SOLUTION: Let us see if there exists a matrix $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $AB = I_2$:

$$\begin{bmatrix} -5 & 2 \\ 9 & -4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{-5a+2c} & -5b+2d \\ 9a-4c & 9b-4d \end{bmatrix} = \begin{bmatrix} \mathbf{1} & 0 \\ 0 & 1 \end{bmatrix}$$

Equating corresponding entries leads us to the following pair of systems of two equations in two unknowns:

$$\left.\begin{aligned} \mathbf{-5a+2c} &= \mathbf{1} \\ 9a-4c &= 0 \end{aligned}\right\} \qquad \Bigg| \qquad \left.\begin{aligned} -5b+2d &= 0 \\ 9b-4d &= 1 \end{aligned}\right\}$$

If you take the time to solve the above systems, you will find that the one on the left has solution $a = -2$, $c = -\frac{9}{2}$; and the one on the right has solution $b = -1$, and $d = -\frac{5}{2}$; which is to say:

$$B = \begin{bmatrix} -2 & -1 \\ -\frac{9}{2} & -\frac{5}{2} \end{bmatrix}$$

We leave it for you to verify that both AB and AB equal I and that therefore $A^{-1} = \begin{bmatrix} -2 & -1 \\ -\frac{9}{2} & -\frac{5}{2} \end{bmatrix}$ (see margin).

EXAMPLE 5.5 Determine if the matrix $A = \begin{bmatrix} 2 & 3 \\ -4 & -6 \end{bmatrix}$ is invertible. If it is, find its inverse.

SOLUTION: As we did in the previous example, we again try to find a matrix $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $AB = I_2$:

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

$$\begin{bmatrix} 2a+3c & 2b+3d \\ -4a-6c & -4b-6d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Equating corresponding entries of the two matrices leads us to the following two systems of equations:

$$\left.\begin{aligned} 2a+3c &= 1 \\ -4a-6c &= 0 \end{aligned}\right\} \qquad \left.\begin{aligned} 2b+3d &= 0 \\ -4b-6d &= 1 \end{aligned}\right\}$$

The system of equation on the right also has no solution.

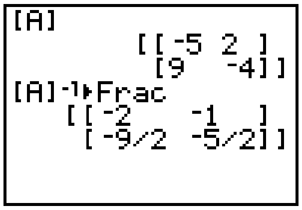
Multiplying the equation $2a+3c = 1$ in the system on the left by 2, and adding it to the bottom equation, leads to an absurdity:

$$\begin{aligned} 4a+6c &= 2 \\ -4a-6c &= 0 \\ \hline \text{add:} \quad 0 &= 2 \end{aligned}$$

We have just observed that there does not exist a matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that $\begin{bmatrix} 2 & 3 \\ -4 & -6 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, which tells us that $\begin{bmatrix} 2 & 3 \\ -4 & -6 \end{bmatrix}$ is a singular matrix.

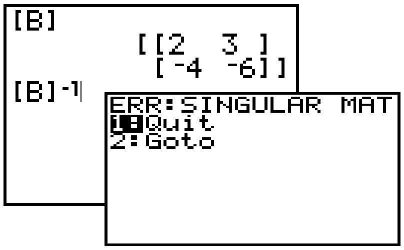
GRAPHING CALCULATOR GLIMPSE 5.2

Example 5.4



```
[A]
        [[-5 2 ]
         [9   -4]]
[A]-1▸Frac
    [[-2    -1  ]
     [-9/2  -5/2]]
```

Example 5.5



```
[B]
        [[2   3 ]
         [-4  -6]]
[B]-1
```

```
ERR:SINGULAR MAT
1:Quit
2:Goto
```

CHECK YOUR UNDERSTANDING 5.7

Determine if $\begin{bmatrix} 3 & -2 \\ 4 & -1 \end{bmatrix}$ is invertible. If it is, find its inverse.

Answer: See page B-19.

While a square matrix need not have an inverse (Example 5.5), if it does, then it is unique:

THEOREM 5.10 An invertible matrix has a **unique** inverse.

PROOF: We assume that B and C are inverses of the invertible matrix A and then go on to show that $B = C$:

$$B = B\underset{\uparrow}{I} = B(A\underset{\uparrow}{C}) = (B\underset{\uparrow}{A})C = \underset{\uparrow}{I}C = C$$

since C is an inverse of A since B is an inverse of A

Here are three additional results pertaining to invertible matrices:

THEOREM 5.11

(i) If A is invertible, then so is A^{-1}, and:

$$(A^{-1})^{-1} = A$$

(ii) If A is invertible and $r \neq 0$, then rA is also invertible, and:

$$(rA)^{-1} = \frac{1}{r}A^{-1}$$

(iii) If A and B are invertible matrices of the same dimension, then AB is also invertible, and:

$$(AB)^{-1} = B^{-1}A^{-1}$$

This is sometimes refered to as the ***shoe-sock theorem*** Can you guess why?

PROOF: (i) We must suppress the temptation to appeal to the familiar exponent rule $(a^{-1})^{-1} = a^{(-1)(-1)} = a$, for we are now dealing with matrices and not numbers. What we do, instead, is to turn to Definition 5.6 which tells us that:

If $A^{-1}\boxed{} = \boxed{}A^{-1}$ for some matrix in the box, then the matrix A^{-1} is invertible, and the matrix in the box is its inverse: $(A^{-1})^{-1} = \boxed{}$. Now put A in the box:

$$A^{-1}\boxed{A} = \boxed{A}A^{-1} = I$$

We have shown that $(A^{-1})^{-1} = \boxed{A}$.

(ii) Exercise 25.

(iii) Returning to our "box game," we see that:

From the given conditions, we know that B^{-1} and A^{-1} exist. What we do here is to show that the product $B^{-1}A^{-1}$ is, in fact, the inverse of AB.

$$(AB)\boxed{(B^{-1}A^{-1})} = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and: $\boxed{(B^{-1}A^{-1})}(AB) = B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I$

The above shows that $(AB)^{-1} = B^{-1}A^{-1}$. In words:

The inverse of a product of invertible matrices is the product of their inverses, in the **reverse order**.

CHECK YOUR UNDERSTANDING 5.8

Use the Principle of Mathematical Induction to show that if $A_1, A_2, \ldots, A_n$ are invertible, then so is their product, and:

$$(A_1A_2\cdots A_n)^{-1} = A_n^{-1}A_{n-1}^{-1}\ldots A_1^{-1}$$

Answer: See page B-19.

As you know, for any nonzero number a:

$$\text{(i) } a^0 = 1 \text{ and (ii) } a^{-n} = \frac{1}{a^n} \text{ (for } a \neq 0\text{)}$$

While we can adopt the notation in (i) for invertible matrices, we cannot do the same for (ii), since the expression $\frac{1}{A}$ is simply **not defined** for a matrix A. We can however rewrite (ii) in the form $a^{-n} = (a^{-1})^n$, and use that form to define A^{-n}:

DEFINITION 5.7

A^0 and A^{-n}

For A, an invertible matrix, and n a positive integer, we define:

(i) $A^0 = I$

(ii) $A^{-n} = (A^{-1})^n$

ELEMENTARY MATRICES

Elementary row operations were introduced on page 3.

DEFINITION 5.8
ELEMENTARY MATRIX

An **elementary matrix** is a matrix that is obtained by performing an elementary row operation on an identity matrix.

Here are some examples of elementary matrices:

$$E_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad E_3 = \begin{bmatrix} 1 & 6 \\ 0 & 1 \end{bmatrix}$$

Interchange the first and second row of I_3

Multiply the second row of I_4 by 5

Add 6 times the second row of I_2 to the first row of I_2

CHECK YOUR UNDERSTANDING 5.9

Show that the above three elementary matrices, E_1, E_2 and E_3 are invertible, and that the inverse of each is itself an elementary matrix.

Answer: See page B-19.

The situation in the above Check Your Understanding box is not a fluke. You are invited to establish the following theorem in the exercises:

THEOREM 5.12 Every elementary matrix is invertible, and its inverse is also an elementary matrix. Indeed:

(a) If E is the elementary matrix obtained by interchanging rows i and j of the identity matrix I, then E^{-1} is the elementary matrix obtained by again interchanging rows i and j of I.

(b) If E is obtained by multiplying row i of I by $c \neq 0$, then E^{-1} is obtained by multiplying row i of I by $\frac{1}{c}$.

(c) If E is obtained by adding c times row i to row j of I, then E^{-1} is obtained by adding $-c$ times row i to row j.

Consider the row operation that is performed in the first and second line of Figure 5.2. As you can see in line 3, the matrix resulting from performing the elementary row operation on A coincides with the product matrix EA:

multiply row 3 by -2 and add it to the first row

$$A = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 1 & 3 & 2 & 6 \\ 3 & 3 & 4 & 1 \end{bmatrix} \xrightarrow{-2R_3 + R_1 \to R_1} \begin{bmatrix} \mathbf{-4} & \mathbf{-5} & \mathbf{-8} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{2} & \mathbf{6} \\ \mathbf{3} & \mathbf{3} & \mathbf{4} & \mathbf{1} \end{bmatrix}$$

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow{-2R_3 + R_1 \to R_1} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = E$$

SAME

$$EA = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 & 3 \\ 1 & 3 & 2 & 6 \\ 3 & 3 & 4 & 1 \end{bmatrix} = \begin{bmatrix} \mathbf{-4} & \mathbf{-5} & \mathbf{-8} & \mathbf{1} \\ \mathbf{1} & \mathbf{3} & \mathbf{2} & \mathbf{6} \\ \mathbf{3} & \mathbf{3} & \mathbf{4} & \mathbf{1} \end{bmatrix}$$

Figure 5.2

The following theorem, a proof of which is relegated to the exercises, tells us that the situation depicted in Figure 5.1 holds in general:

THEOREM 5.13 The matrix obtained by performing an elementary row operation on a matrix $A \in M_{m \times n}$ equals that matrix EA, where E is the matrix obtained by performing the same elementary row operation on the identity matrix I_m.

CHECK YOUR UNDERSTANDING 5.10

(a) Switch the first and third rows of the matrix A of Figure 5.1 to arrive at a matrix B, and then find the elementary matrix E such that $EA = B$.

(b) Multiply the second row of the matrix A of Figure 5.1 by 2 to arrive at a matrix B, and then find the elementary matrix E such that $EA = B$.

Answer: See page B-19.

We remind you that two matrices are equivalent if one can be derived from the other via a sequence of elementary row operations (Definition 1.1, page 3). The following theorem asserts that if a matrix A is invertible, then every matrix equivalent to A is also invertible.

THEOREM 5.14 If A and B are equivalent square matrices, then **A is invertible if and only if B is invertible**.

PROOF: Since A and B are equivalent, there exist elementary matrices $E_1, E_2, \ldots, E_s$ such that:

$$B = E_s \cdots E_2 E_1 A$$

Theorem 5.11 tells us that each E_i is invertible. Consequently, if A is invertible then so is B, for it is a product of invertible matrices. By symmetry we also have that if B is invertible then so is A.

You can quickly determine whether or not a matrix is invertible by looking at its row-reduced-echelon form:

THEOREM 5.15 $A \in M_{n \times n}$ is invertible if and only if:

$$\text{rref}(A) = I_n$$

Any matrix A is equivalent to $\text{rref}(A)$.

PROOF: If $\text{rref}(A) = I_n$ then, sine I is invertible, A is also invertible (Theorem 5.14).

We establish the converse by showing that if $\text{rref}(A) \neq I_n$, then A is not invertible:

Let $\text{rref}(A) = C \neq I_n$. Since $C = [c_{ij}]$ has less than n leading ones (otherwise it would be I_n), its last row must consist entirely of zeros. This being the case, for any given $n \times n$ matrix $D = [d_{ij}]$, the product matrix $CD = [e_{ij}]$ cannot be the identity matrix, as its lower-right-corner entry is not 1:

$$e_{nn} = \sum_{\alpha = 1}^{n} c_{n\alpha} d_{\alpha n} = \sum_{\alpha = 1}^{n} 0 \cdot d_{\alpha n} = 0 \neq 1$$

Since there does not exist a matrix D such that $CD = I$, $\text{rref}(A) = C$ is not invertible. It follows, from Theorem 5.14, that A is not invertible.

The following theorem provides a systematic method for finding the inverse of an invertible matrix:

THEOREM 5.16 If a sequence of elementary row operations reduces the invertible matrix A to I, then applying the same sequence of elementary row operations on I will yield A^{-1}.

PROOF: Let $E_1, E_2, \ldots, E_s$ be the elementary matrices corresponding to elementary row operations which take A to I:

$$E_s \cdots E_2 E_1 A = I$$

Since elementary matrices are invertible, their product is also invertible and we have:

$$A = (E_s \ldots E_2 E_1)^{-1} I = (E_s \ldots E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} \ldots E_s^{-1}$$

Thus:

We have shown that $E_s \ldots E_2 E_1 I$ is the inverse of the matrix A; to put it another way:

$$A^{-1} = (E_1^{-1}E_2^{-1}\ldots E_s^{-1})^{-1} = (E_s^{-1})^{-1}\ldots(E_2^{-1})^{-1}(E_1^{-1})^{-1}$$

$$= E_s\ldots E_2E_1 = E_s\ldots E_2E_1I$$

The exact same sequence of row operations that transform A to I can also be used to transform I to A^{-1}.

We now illustrate how the above theorem can effectively be used to find, by hand, the inverse of an invertible matrix:

EXAMPLE 5.6

Determine if $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{bmatrix}$ is invertible. If it is, find its inverse.

SOLUTION: We know that A is invertible if and only if $\text{rref}(A) = I$. Rather than obtaining the row-reduced-echelon form of A, we will find the row-reduced echelon form of the matrix $[A|I]$ which is that matrix obtained by "adjoining" the 3×3 identity matrix to A:

$$[A|I] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right]$$

This will feed two birds with one seed:

First bird: A is invertible if and only if $\text{rref}([A|I]) = [I|\ \ ]$

(By Theorem 5.14)

Second bird: If A is invertible, then $\text{rref}([A|I]) = [I|A^{-1}]$

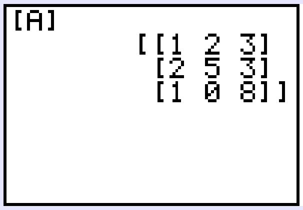
(By Theorem 5.16)

Using the Gauss-Jordan Elimination Method (page 10), we have:

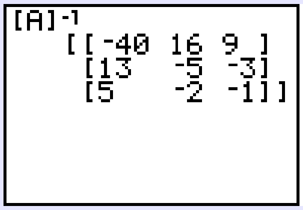
$$[A|\boldsymbol{I}] = \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{array}\right] \xrightarrow[-1R_1+R_3 \to R_3]{-2R_1+R_2 \to R_2} \left[\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{array}\right] \xrightarrow[2R_2+R_3 \to R_3]{-2R_2+R_1 \to R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{array}\right]$$

$$\xrightarrow{-1R_3 \to R_3} \left[\begin{array}{ccc|ccc} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] \xrightarrow[3R_3+R_2 \to R_2]{-9R_3+R_1 \to R_1} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{array}\right] = [\boldsymbol{I}|A^{-1}]$$

Oh well:



```
[A]
        [[1 2 3]
         [2 5 3]
         [1 0 8]]
```



```
[A]⁻¹
   [[-40 16 9 ]
    [13  -5 -3]
    [5   -2 -1]]
```

CHECK YOUR UNDERSTANDING 5.11

Determine if $A = \begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 1 & 4 \\ 1 & 1 & 1 & 0 \\ 0 & 3 & 1 & 1 \end{bmatrix}$ is invertible. If it is, find its inverse.

Answer: See page B-19.

Theorem 5.15 says that a matrix A is invertible if and only $\text{rref}(A) = I$. Here are some other ways of saying the same thing:

THEOREM 5.17 Let $A \in M_{n \times n}$. The following are equivalent:

(i) A is invertible.

(ii) $AX = B$ has a unique solution for every $B \in M_{n \times 1}$.

(iii) $AX = \mathbf{0}$ has only the trivial solution.

(iv) $\text{rref}(A) = I$.

(v) A is a product of elementary matrices.

PROOF: We show $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (iv) \Rightarrow (v) \Rightarrow (i)$:

$(i) \Rightarrow (ii)$: Let A be invertible. For any given equation $AX = B$, we have:

$$\begin{aligned} AX &= B \\ A^{-1}(AX) &= A^{-1}B \\ (A^{-1}A)X &= A^{-1}B \\ IX &= A^{-1}B \\ X &= A^{-1}B \longleftarrow \text{unique solution} \end{aligned}$$

$(ii) \Rightarrow (iii)$: If $AX = B$ has a unique solution for any B, then $AX = \mathbf{0}$ has a unique solution. But $AX = \mathbf{0}$ has the trivial solution $X = \mathbf{0}$. It follows that $AX = \mathbf{0}$ has only the trivial solution.

$(iii) \Rightarrow (iv)$: Assume that $AX = \mathbf{0}$ has only the trivial solution. Since A and $\text{rref}(A)$ are equivalent matrices, $[\text{rref}(A)]X = \mathbf{0}$ has only the trivial solutions. This tells us that $\text{rref}(A)$ does not have a free variable, and must therefore have n leading ones. As such: $\text{rref}(A) = I$.

$(iv) \Rightarrow (v)$: If $\text{rref}(A) = I$ then:

$$E_s \cdots E_2 E_1 A = I$$

for some elementary matrices $E_1, E_2, \ldots, E_s$. Since elementary matrices are invertible, their product is also invertible and we have:

$$A = (E_s \ldots E_2 E_1)^{-1} I = (E_s \ldots E_2 E_1)^{-1} = \underset{\substack{\uparrow \\ \text{a product of elementary matrices} \\ \text{(see CYU 5.9)}}}{E_1^{-1} E_2^{-1} \ldots E_s^{-1}}$$

$(v) \Rightarrow (i)$: If A is a product of elementary matrices, then it is a product of invertible matrices, and is therefore invertible [CYU 5.9].

We defined a matrix A to be invertible if there exists a matrix B such that $AB = BA = I$. As it turns out, B need only "work" on the left (or right side) of A:

THEOREM 5.18 Let $A \in M_{n \times n}$. If there exists $B \in M_{n \times n}$ such that $BA = I$, then A is invertible and $A^{-1} = B$.

PROOF: Assume that $BA = I$, and consider the equation $AX = \mathbf{0}$.

$$\begin{aligned} AX &= \mathbf{0} \\ B(AX) &= B\mathbf{0} \\ (BA)X &= B\mathbf{0} \\ IX &= \mathbf{0} \\ X &= \mathbf{0} \end{aligned}$$

Since the equation $AX = \mathbf{0}$ has only the trivial solution, the matrix A is invertible (Theorem 5.17). Upon multiplying both sides of $BA = I$

$$(BA)A^{-1} = IA^{-1}$$

on the right by A^{-1}, we have: $B(AA^{-1}) = A^{-1}$

$$B = A^{-1}$$

In a similar fashion:

CHECK YOUR UNDERSTANDING 5.12

Let $A \in M_{n \times n}$. Show that if there exists $B \in M_{n \times n}$ such that $AB = I$, then B is invertible and $B^{-1} = A$.

Answer: See page B-20.

EXERCISES

Exercises 1-6. Find the inverse of the given invertible matrix A, and then check directly that $AA^{-1} = A^{-1}A = I$.

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

4. $\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 1 & 2 & 3 \end{bmatrix}$ 5. $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{bmatrix}$ 6. $\begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 3 \\ 2 & 0 & 4 \end{bmatrix}$

Exercises 7-15. Determine if the given matrix A is invertible and if so, find its inverse.

7. $\begin{bmatrix} 2 & 3 \\ -4 & -6 \end{bmatrix}$ 8. $\begin{bmatrix} 3 & 2 \\ 3 & 1 \end{bmatrix}$ 9. $\begin{bmatrix} -1 & 4 \\ 5 & -20 \end{bmatrix}$

10. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 3 & 3 & 3 \end{bmatrix}$ 11. $\begin{bmatrix} 1 & 4 & 0 \\ 2 & -2 & 3 \\ 13 & 2 & 12 \end{bmatrix}$ 12. $\begin{bmatrix} 1 & 3 & 0 \\ 2 & -2 & 3 \\ 1 & 2 & 1 \end{bmatrix}$

13. $\begin{bmatrix} 1 & 0 & 1 & 2 \\ 0 & 2 & 1 & 4 \\ 1 & 1 & 2 & 2 \\ 0 & 3 & 1 & 8 \end{bmatrix}$ 14. $\begin{bmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 1 & 0 \\ 1 & -1 & 2 & 0 \\ 0 & 3 & 1 & 1 \end{bmatrix}$ 15. $\begin{bmatrix} 4 & 0 & 4 & 8 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 \\ 2 & 1 & 1 & -2 \end{bmatrix}$

Exercises 16-18. (a) Find elementary matrices E_1 and E_2 such that $E_2E_1I = A$ for the given matrix A. (b) Find two elementary row operations which will take A to I.

16. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ 17. $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 5 & 0 & 0 \end{bmatrix}$ 18. $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 7 & 1 & 0 \end{bmatrix}$

Exercises 19-24. Assume that A, B, and X are square matrices of the same dimension and that A and B are invertible. (a) Solve the given matrix equation for the matrix X.

(b) Challenge your answer in (a) using $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 1 \\ 3 & 1 \end{bmatrix}$.

19. $2XA = AB$ 20. $2XA = BA$ 21. $AXB^{-1} = BA$

22. $BXA = B^2$ 23. $BXB = (BAB)^2$ 24. $AXB = A^2(BA)^{-1}$

25. Prove Theorem 5.11(ii).
26. Prove Theorem 5.12.
27. Prove Theorem 5.13.
28. Prove that if A is invertible, then $A^{-2}A^{-3} = A^{-5}$.
29. Let $A \in M_{m \times n}$. Prove that there exists an invertible matrix $C \in M_{m \times m}$ such that $CA = \text{rref}(A)$.
30. Let $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s\}$ be a linearly independent set of vectors in $M_{n \times n}$, and let $A \in M_{n \times n}$ be invertible. Show that $\{A\boldsymbol{A}_1, A\boldsymbol{A}_2, \ldots, A\boldsymbol{A}_s\}$ is linearly independent.
31. Let $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_{n^2}\}$ be a basis for $M_{n \times n}$, and let $A \in M_{n \times n}$ be invertible. Show that $\{A\boldsymbol{A}_1, A\boldsymbol{A}_2, \ldots, A\boldsymbol{A}_{n^2}\}$ is also a basis.
32. Show that a (square) matrix that has a row consisting entirely of zeros cannot be invertible.
33. Show that a (square) matrix that has a columns consisting entirely of zeros cannot be invertible.
34. Show that if a row of a (square) matrix is a multiple of one of its other rows, then it is not invertible.
35. State necessary and sufficient conditions for a diagonal matrix to be invertible. (See Exercise 18, page 161.)
36. Prove that $A \in M_{n \times n}$ is invertible if and only if the rows of A constitute a basis for $\Re^n$.
37. Prove that the transpose A^{T} of an invertible matrix A is invertible, and that $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$. (See Exercise 19, page 161.)
38. Prove that $A \in M_{n \times n}$ is invertible if and only if the columns of A constitute a basis for $\Re^n$.
39. Prove that if a symmetric matrix is invertible, then its inverse is also symmetric. (See Exercise 20, page 161.)
40. Prove that if $A \in M_{n \times n}$ is an idempotent invertible matrix, then $A = I_n$. (See Exercise 22, page 162.)
41. Prove that every nilpotent matrix is singular. (See Exercise 23, page 162.)
42. (a) Prove that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible if and only if $ad - bc \neq 0$.

    (b) Assuming that $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is invertible, find A^{-1} (in terms of a, b, c, and d).
43. Let $A \in M_{n \times n}$ be such that $A^2 - 2A + I = 0$. Show that A is invertible.
44. Let $A \in M_{n \times n}$ be such that $A^2 + sA + tI = 0$, with $t \neq 0$. Show that A is invertible.

45. Let $A \in M_{n \times n}$ be such that $A^3 - 3A + 2I = 0$. Show that A is invertible.

46. **(PMI)** Show that if A is invertible, then so is A^n for every positive integer n.

47. **(PMI)** Let A and B be invertible matrices of the same dimension with $AB = BA$. Sow that:

    (a) $(B^{-1})^n A^{-1} = A^{-1}(B^{-1})^n$ for every positive integer n.

    (b) $(AB)^{-n} = A^{-n}B^{-n}$ for every positive integer n.

PROVE OR GIVE A COUNTEREXAMPLE

48. If A is invertible, then so is $-A$, and $(-A)^{-1} = -A^{-1}$.

49. If $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s\}$ is a linearly independent set in the vector space $M_{n \times n}$, and if $A \in M_{n \times n}$ is not the zero vector, then $\{A\boldsymbol{A}_1, A\boldsymbol{A}_2, \ldots, A\boldsymbol{A}_s\}$ is linearly independent.

50. Let A be an $n \times n$ invertible matrix, and $B \in M_{n \times m}$. If $AB = 0$, then $B = 0$.

51. Let $A \in M_{n \times n}$ be invertible, and $B \in M_{m \times n}$. If $BA = 0$, then $B = 0$

52. If A and B are $n \times n$ invertible matrices, then $A + B$ is also invertible, and $(A + B)^{-1} = A^{-1} + B^{-1}$.

53. If $A \in M_{n \times n}$ and $A^2 = 0$, then A is not invertible.

54. If a square matrix A is singular, then $A^2 = 0$.

55. If A and B are $n \times n$ matrices, and if AB is invertible, then both A and B are invertible.

56. If A and B are $n \times n$ matrices, and if AB is singular, then both A and B are singular.

§3. MATRIX REPRESENTATION OF LINEAR MAPS

The main objective of this section is to associate to a general linear transformation $T: V \to W$, where $\dim(V) = n$ and $\dim(W) = m$, a matrix $A \in M_{m \times n}$ which can be used to find the value of $T(\boldsymbol{v})$, for every $\boldsymbol{v} \in V$. The importance of all of this is that, in a way, the linear transformation T is "linked to a finite object:" an $m \times n$ matrix A.

The notion of an **ordered basis** for a vector space plays a role in the current development. The only difference between a basis and an ordered basis is that, in the latter, the listing of the vectors is of consequence. For example, while $\{(\mathbf{1}, \mathbf{2}), (\mathbf{3}, \mathbf{5})\}$ and $\{(\mathbf{3}, \mathbf{5}), (\mathbf{1}, \mathbf{2})\}$ are one-and-the-same bases, they are **not** the same **ordered** bases (different ordering of the elements).

Choosing an ordered basis for a vector space V of dimension n enables us to associate vertical n-tuple to each vector of V:

DEFINITION 5.9

COORDINATE VECTOR

Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be an ordered basis for a vector space V, and let

$$\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$$

We define the **coordinate vector** of v relative $\boldsymbol{\beta}$ to be the column vector:

$$[\boldsymbol{v}]_\beta = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

EXAMPLE 5.7 In Example 3.10, page 94, we showed that $\beta = \{(\mathbf{1}, \mathbf{3}, \mathbf{0}), (\mathbf{2}, \mathbf{0}, \mathbf{4}), (\mathbf{0}, \mathbf{1}, \mathbf{2})\}$ is a basis for $\Re^3$. Find the coordinate vector of $(\mathbf{1}, \mathbf{11}, \mathbf{8})$ with respect to β.

SOLUTION: As is often the case, the problem boils down to that of solving a system of linear equations:

If: $(\mathbf{1}, \mathbf{11}, \mathbf{8}) = a(\mathbf{1}, \mathbf{3}, \mathbf{0}) + b(\mathbf{2}, \mathbf{0}, \mathbf{4}) + c(\mathbf{0}, \mathbf{1}, \mathbf{2})$, then:

$$\mathbf{S}: \left.\begin{array}{rcl} a + 2b + 0c &=& 1 \\ 3a + 0b + c &=& 11 \\ 0a + 4b + 2c &=& 8 \end{array}\right\} \xrightarrow{\mathbf{aug}(S)} \begin{array}{c} \begin{matrix} a & b & c & \end{matrix} \\ \left[\begin{array}{ccc|c} 1 & 2 & 0 & 1 \\ 3 & 0 & 1 & 11 \\ 0 & 4 & 2 & 8 \end{array}\right] \end{array} \xrightarrow{\mathbf{rref}} \begin{array}{c} \begin{matrix} a & b & c & \phantom{-\frac{1}{2}} \end{matrix} \\ \left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 5 \end{array}\right] \end{array} \to \begin{array}{rcl} \boldsymbol{a} &=& \mathbf{2} \\ \boldsymbol{b} &=& -\frac{\mathbf{1}}{\mathbf{2}} \\ \boldsymbol{c} &=& \mathbf{5} \end{array} \Rightarrow [(\mathbf{1}, \mathbf{11}, \mathbf{8})]_\beta = \begin{bmatrix} \mathbf{2} \\ -\frac{\mathbf{1}}{\mathbf{2}} \\ \mathbf{5} \end{bmatrix}$$

CHECK YOUR UNDERSTANDING 5.13

Find the coordinate vector of $(\mathbf{-1, 1, 2}) \in \Re^3$ with respect to:

$$\beta = \{(\mathbf{1, 3, 0}), (\mathbf{2, 0, 4}), (\mathbf{0, 1, 2})\}$$

Answer: $\begin{bmatrix} -1/4 \\ -3/8 \\ 7/4 \end{bmatrix}$

THEOREM 5.19 If $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for a vector space V, then the function $L_\beta\colon V \to \overline{\Re}^n$ given by $L_\beta(\mathbf{v}) = [v]_\beta$ (a vertical n-tuple) is linear.

[Indeed, it is an isomorphism (Exercise 32).]

Throughout this section the term "basis" will be understood to mean "**ordered** basis."

We remind you that we are using $\overline{\Re}^n$ to denote $M_{n \times 1}$

PROOF: For $\mathbf{v} = \sum_{i=1}^{n} a_i \mathbf{v}_i$ and $\mathbf{v}' = \sum_{i=1}^{n} b_i \mathbf{v}_i$ in V, and $r \in \Re$, we have:

$$L_\beta(r\mathbf{v} + \mathbf{v}') = L_\beta\left(r\sum_{i=1}^{n} a_i\mathbf{v}_i + \sum_{i=1}^{n} b_i\mathbf{v}_i\right) = L_\beta\left(\sum_{i=1}^{n}(ra_i + b_i)\mathbf{v}_i\right)$$

$$= \begin{bmatrix} ra_1 + b_1 \\ \vdots \\ ra_n + b_n \end{bmatrix} = \begin{bmatrix} ra_1 \\ \vdots \\ ra_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = r\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

$$= rL_\beta\left(\sum_{i=1}^{n} a_i\mathbf{v}_i\right) + L_\beta\left(\sum_{i=1}^{n} b_i\mathbf{v}_i\right) = rL_\beta(\mathbf{v}) + L_\beta(\mathbf{v}')$$

$$\begin{tikzcd} L_\beta(r\mathbf{v} + \mathbf{v}') \arrow[d, bend right] \\ = rL_\beta(\mathbf{v}) + L_\beta(\mathbf{v}') \end{tikzcd}$$

Consider a linear transformation $T\colon V \to W$, with $\dim(V) = n$ and $\dim(W) = m$. If we fix a basis $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ for V and a basis $\gamma = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_m\}$ for W, then for every $\mathbf{v} \in V$ we can determine the coefficient matrix $[\mathbf{v}]_\beta$ of $\mathbf{v}$ with respect to β, as well as the coefficient matrix $[T(\mathbf{v})]_\gamma$ of $T(\mathbf{v})$ with respect to γ. The following definition provides an important relation between $[\mathbf{v}]_\beta$ and $[T(\mathbf{v})]_\gamma$:

γ (gamma) is the Greek letter c.

DEFINITION 5.10

MATRIX REPRESENTATION OF A LINEAR MAP

Let $T: V \to W$ be a linear map from a vector space V of dimension n to a vector space W of dimension m. Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\gamma = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ be bases for V and W, respectively. We define the **matrix representation** of T with respect to β and γ to be the matrix $[T]_{\gamma\beta} \in M_{m \times n}$ whose i^{th} column is $[T(\boldsymbol{v}_i)]_\gamma$.

The above definition looks intimidating, but appearances can be misleading. Consider the following example.

EXAMPLE 5.8

Let $T: \Re^3 \to P_2$ be the linear map given by:

$$T(a, b, c) = 2ax^2 + (a+b)x + c$$

Determine the matrix representation $[T]_{\gamma\beta}$ of T with respect to the bases:

$$\beta = \{(1, 3, 0), (2, 0, 4), (0, 1, 2)\}$$

and

$$\gamma = \{2x^2 + x, 3x^2 - 1, x\}$$

SOLUTION: Definition 5.10 tells us that the first column of $[T]_{\gamma\beta}$ is the coefficient matrix of $T(1, 3, 0)$ with respect to γ [entries are the values of a, b, c stemming from (1) below], the second column is $[T(2, 0, 4)]_\gamma$ [values of a, b, c stemming from (2)], and the third column is $[T(0, 1, 2)]_\gamma$ [values stemming from (3)]:

$$\text{(1):}\quad T(1, 3, 0) = \mathbf{2}x^2 + \mathbf{4}x + \mathbf{0} = \boldsymbol{a}(\mathbf{2x^2+x}) + \boldsymbol{b}(\mathbf{3x^2-1}) + \boldsymbol{c}(\mathbf{x})$$

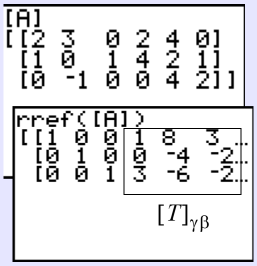
$$\text{(2):}\quad T(2, 0, 4) = \mathbf{4}x^2 + \mathbf{2}x + \mathbf{4} = \boldsymbol{a}(\mathbf{2x^2+x}) + \boldsymbol{b}(\mathbf{3x^2-1}) + \boldsymbol{c}(\mathbf{x})$$

$$\text{(3):}\quad T(0, 1, 2) = \mathbf{0}x^2 + \mathbf{1}x + \mathbf{2} = \boldsymbol{a}(\mathbf{2x^2+x}) + \boldsymbol{b}(\mathbf{3x^2-1}) + \boldsymbol{c}(\mathbf{x})$$

Equating coefficients in each of the above we come to the following three systems of equations:

$$\text{(1):}\ \begin{array}{ccc|c} \overset{2x^2+1x+0}{\downarrow} & \overset{3x^2+0x-1}{\downarrow} & \overset{0x^2+1x+0}{\downarrow} & \\ 2 & 3 & 0 & \mathbf{2} \\ 1 & 0 & 1 & \mathbf{4} \\ 0 & -1 & 0 & \mathbf{0} \end{array} \qquad \text{(2):}\ \begin{array}{ccc|c} a & b & c & \\ 2 & 3 & 0 & \mathbf{4} \\ 1 & 0 & 1 & \mathbf{2} \\ 0 & -1 & 0 & \mathbf{4} \end{array} \qquad \text{(3):}\ \begin{array}{ccc|c} a & b & c & \\ 2 & 3 & 0 & \mathbf{0} \\ 1 & 0 & 1 & \mathbf{1} \\ 0 & -1 & 0 & \mathbf{2} \end{array}$$

We want the solutions of the above systems of three equations in the three unknowns a, b, c. Noting that the coefficient matrix of all three systems are one and the same, we can “compress” all three systems into a single matrix form:



```
[A]
[[2 3  0 2 4 0]
 [1 0  1 4 2 1]
 [0 -1 0 0 4 2]]
rref([A])
[[1 0 0 1 8  3 …
 [0 1 0 0 -4 -2…
 [0 0 1 3 -6 -2…
```

$[T]_{\gamma\beta}$

$$\begin{array}{ccc|ccc} & & & \text{system (1)} & \text{system (2)} & \text{system (3)} \\ & & & a & b & c \\ 2 & 3 & 0 & \mathbf{2} & \mathbf{4} & \mathbf{0} \\ 1 & 0 & 1 & \mathbf{4} & \mathbf{2} & \mathbf{1} \\ 0 & -1 & 0 & \mathbf{0} & \mathbf{4} & \mathbf{2} \end{array} \xrightarrow{\text{rref}} \begin{array}{ccc|ccc} & & & \text{solutions of system (1)} & \text{solutions of system (2)} & \text{solutions of system (3)} \\ 1 & 0 & 0 & \mathbf{1} & \mathbf{8} & \mathbf{3} \\ 0 & 1 & 0 & \mathbf{0} & \mathbf{-4} & \mathbf{-2} \\ 0 & 0 & 1 & \mathbf{3} & \mathbf{-6} & \mathbf{-2} \end{array} \Rightarrow [T]_{\gamma\beta} = \begin{bmatrix} \mathbf{1} & \mathbf{8} & \mathbf{3} \\ \mathbf{0} & \mathbf{-4} & \mathbf{-2} \\ \mathbf{3} & \mathbf{-6} & \mathbf{-2} \end{bmatrix}$$

Figure 5.2

CHECK YOUR UNDERSTANDING 5.14

Let $T: M_{2\times 2} \to \Re^3$ be the linear map given by:

$$T = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = (b, a, c+d)$$

Determine $[\boldsymbol{T}]_{\gamma\beta}$ for:

$$\beta = \left\{ \begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix}, \begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 5 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 6 \\ 5 & 9 \end{bmatrix} \right\}$$

and $\gamma = \{(\mathbf{1}, \mathbf{3}, \mathbf{0}), (\mathbf{2}, \mathbf{0}, \mathbf{4}), (\mathbf{0}, \mathbf{1}, \mathbf{2})\}$.

Answer:

$$\begin{bmatrix} 1 & 2 & -3/4 & 1/4 \\ 1 & 0 & -1/8 & 23/8 \\ 0 & 1 & 13/4 & 5/4 \end{bmatrix}$$

To know the coordinates of a vector $\boldsymbol{v}$ in a finite dimensional vector space $\boldsymbol{V}$ with respect to a fixed ordered basis for that space is the same as knowing the vector $\boldsymbol{v}$ itself. This being the case, the following theorem tells us that the action of a linear transformation $T: V \to W$ can, in effect, be realized by means of a matrix multiplication.

THEOREM 5.20 Let $T: V \to W$ be a linear map from a vector space V of dimension n to a vector space W of dimension m. Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ and $\gamma = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}$ be bases for V and W, respectively. Then, for every $\boldsymbol{v} \in V$:

$$[T(\boldsymbol{v})]_\gamma = [T]_{\gamma\beta}[\boldsymbol{v}]_\beta$$

Note that the dimensions match up:

$$\underset{\in M_{m\times 1}}{[T(\boldsymbol{v})]_\gamma} = \underset{\in M_{m\times n}}{[T]_{\gamma\beta}}\underset{\in N_{n\times 1}}{[\boldsymbol{v}]_\beta}$$

PROOF: Since $[T]_{\gamma\beta}[\boldsymbol{v}]_\beta: V \to \overline{\Re^m}$ and $[T(\boldsymbol{v})]_\gamma: V \to \overline{\Re^m}$ are both linear, it suffices to show that $[T]_{\gamma\beta}[\boldsymbol{v}_i]_\beta = [T(\boldsymbol{v}_i)]_\gamma$ for each $1 \le i \le n$ (Theorem 4.6, page 115). They are:

$[T]_{\gamma\beta}[\boldsymbol{v}_i]_\beta$ is the i^{th} column of $[T]_{\gamma\beta}$, namely: $[T(\boldsymbol{v_i})]_\gamma$ (see margin and Exercise 10, page 160). By Definition 5.10, $[T(\boldsymbol{v_i})]_\gamma$ is also the i^{th} column of the matrix $[T]_{\gamma\beta}$.

Since $v_i = 0v_1 + \cdots + 1v_i + \ldots + 0v_n$,

$$[v_i] = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow i^{\text{th}}\text{position} \\ \\ \\ \end{matrix}$$

EXAMPLE 5.9 Let $T: \Re^2 \to \Re^3$ be the linear map given by:

$$T(\mathbf{a}, \mathbf{b}) = (\mathbf{a} + \mathbf{b}, \mathbf{b}, 2\mathbf{a})$$

Determine $[T]_{\gamma\beta}$ of T with respect to the bases:

$$\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{3}, \mathbf{0})\}$$

and

$$\gamma = \{(\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$$

and then show directly that:

$$[T(\mathbf{5}, \mathbf{7})]_\gamma = [T]_{\gamma\beta}[(\mathbf{5}, \mathbf{7})]_\beta$$

SOLUTION: Proceeding as in Figure 5.2 of Example 5.8 we find $[T]_{\gamma\beta}$. Since $T(\mathbf{1}, \mathbf{2}) = (\mathbf{3}, \mathbf{2}, \mathbf{2})$ and $T(\mathbf{3}, \mathbf{0}) = (\mathbf{3}, \mathbf{0}, \mathbf{6})$:

$[T]_{\gamma\beta}$

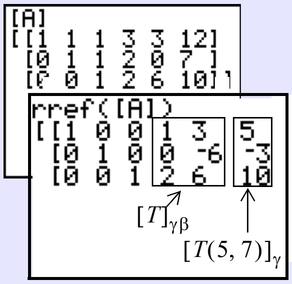
$$(*)\ \begin{array}{c} \begin{matrix} a & b & c \end{matrix} \\ \left[\begin{array}{ccc|cc} 1 & 1 & 1 & 3 & 3 \\ 0 & 1 & 1 & 2 & 0 \\ 0 & 0 & 1 & 2 & 6 \end{array}\right] \end{array} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|cc} 1 & 0 & 0 & 1 & 3 \\ 0 & 1 & 0 & 0 & -6 \\ 0 & 0 & 1 & 2 & 6 \end{array}\right]$$

The three vectors in γ

Solving the two systems of equations:

$$a(1, 0, 0) + b(1, 1, 0) + c(1, 1, 1) = \begin{cases} (3, 2, 2) \\ (3, 0, 6) \end{cases}$$

INCIDENTALLY, noting that the coefficient matrix of system (*) is identical to that of (**) we could save a bit of time by doing this



```
[A]
[[1 1 1 3 3 12]
 [0 1 1 2 0 7 ]
 [0 0 1 2 6 10]]
rref([A])
[[1 0 0 1 3 5]
 [0 1 0 0 -6 -3]
 [0 0 1 2 6 10]]
```

Finding $[T(\mathbf{5}, \mathbf{7})]_\gamma$. Since $T(\mathbf{5}, \mathbf{7}) = (\mathbf{12}, \mathbf{7}, \mathbf{10})$:

$[T(5, 7)]_\gamma$

$$(**)\ \left[\begin{array}{ccc|c} 1 & 1 & 1 & 12 \\ 0 & 1 & 1 & 7 \\ 0 & 0 & 1 & 10 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 5 \\ 0 & 1 & 0 & -3 \\ 0 & 0 & 1 & 10 \end{array}\right]$$

Finding $[(5, 7)]_\beta$:

$[(\mathbf{5}, \mathbf{7})]_\beta$

$$\left[\begin{array}{cc|c} 1 & 3 & 5 \\ 2 & 0 & 7 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|c} 1 & 0 & \frac{7}{2} \\ 0 & 1 & \frac{1}{2} \end{array}\right]$$

The two vectors in β

We leave it for you to verifying that

$$[T(\mathbf{5}, \mathbf{7})]_\gamma = [T]_{\gamma\beta}[(\mathbf{5}, \mathbf{7})]_\beta\text{; that is: } \begin{bmatrix} 5 \\ -3 \\ 10 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 0 & -6 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} \frac{7}{2} \\ \frac{1}{2} \end{bmatrix}.$$

CHECK YOUR UNDERSTANDING 5.15

Referring to the situation described in CYU 5.15, verify directly that:

$$\left[T\left(\begin{bmatrix}1 & 3\\2 & 0\end{bmatrix}\right)\right]_{\gamma} = [T]_{\gamma\beta}\left[\begin{bmatrix}1 & 3\\2 & 0\end{bmatrix}\right]_{\beta}$$

Answer: See page B-20.

The next result is particularly nice — it tells us that a matrix of a composite of linear transformations is a product of the matrices of those transformations:

$$\begin{tikzcd} V_{\beta} \arrow[r, "T"] \arrow[rr, bend right, "L\circ T"'] & W_{\gamma} \arrow[r, "L"] & Z_{\delta} \end{tikzcd}$$

$$[L\circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$$

THEOREM 5.21

THE COMPOSITION THEOREM

Let $T: V \to W$ and $L: W \to Z$ be linear maps, and let β, γ, and δ be bases for the finite dimensional spaces V, W, and Z, respectively. Then:

$$[L\circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$$

PROOF: For any $\boldsymbol{v} \in V$ we have:

$$\begin{aligned}[L\circ T]_{\delta\beta}[v]_{\beta} &\overset{\text{Theorem 5.20}}{=} [(L\circ T)(\boldsymbol{v})]_{\delta}\\ &= [L(T(\boldsymbol{v}))]_{\delta} \overset{\text{Theorem 5.20}}{=} [L]_{\delta\gamma}[T(\boldsymbol{v})]_{\gamma} \overset{\text{Theorem 5.20}}{=} [L]_{\delta\gamma}[T]_{\gamma\beta}[\boldsymbol{v}]_{\beta}\end{aligned}$$

The result now follows from Exercise 34 which asserts that if A is a matrix such that $[L\circ T]_{\delta\beta}[v]_{\beta} = A[v]_{\beta}$ for every $\boldsymbol{v} \in V$, then the matrices $[L\circ T]_{\delta\beta}$ and A must be one and the same.

EXAMPLE 5.10

Let $T: \Re^2 \to M_{2\times 2}$ and $L: M_{2\times 2} \to P_2$ be the linear maps given by:

$$T(\boldsymbol{a}, \boldsymbol{b}) = \begin{bmatrix}a & a+b\\0 & 2b\end{bmatrix}$$

and

$$L\left(\begin{bmatrix}a & b\\c & d\end{bmatrix}\right) = ax^2 + bx + (c+d)$$

Show directly that:

$$[L\circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$$

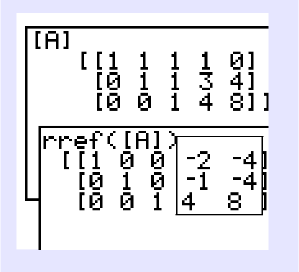
For bases:

$$\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{0}, \mathbf{4})\}$$

$$\gamma = \left\{\begin{bmatrix}1 & 1\\0 & 0\end{bmatrix}, \begin{bmatrix}0 & 2\\1 & 0\end{bmatrix}, \begin{bmatrix}0 & 0\\1 & 1\end{bmatrix}, \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right\}$$

$$\delta = \{x^2, x^2 + x, x^2 + x + 1\}$$

SOLUTION: Determining $[T]_{\gamma\beta}$:



```
[A]
    [[1 1 1 1 0]
     [0 1 1 3 4]
     [0 0 1 4 8]]
rref([A])
  [[1 0 0 -2 -4]
   [0 1 0 -1 -4]
   [0 0 1 4  8 ]]
```

$$T(\mathbf{0},\mathbf{4}) = \begin{bmatrix}0&4\\0&8\end{bmatrix} \qquad T(\mathbf{1},\mathbf{2}) = \begin{bmatrix}1&3\\0&4\end{bmatrix}$$

$$\begin{array}{cccc|cc} \begin{bmatrix}1&1\\0&0\end{bmatrix} & \begin{bmatrix}0&2\\1&0\end{bmatrix} & \begin{bmatrix}0&0\\1&1\end{bmatrix} & \begin{bmatrix}0&0\\0&1\end{bmatrix} & T(\mathbf{1},\mathbf{2}) & T(\mathbf{0},\mathbf{4}) \\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow \\ 1&0&0&0&1&0\\ 1&2&0&0&3&4\\ 0&1&1&0&0&0\\ 0&0&1&1&4&8 \end{array} \qquad \left[\begin{array}{cccc|cc} 1&0&0&0&1&0\\ 0&1&0&0&1&2\\ 0&0&1&0&-1&-2\\ 0&0&0&1&5&10 \end{array}\right] \Rightarrow [T]_{\gamma\beta} = \begin{bmatrix}1&0\\1&2\\-1&-2\\5&10\end{bmatrix}$$

Determining $[L]_{\delta\gamma}$:

$$L\left(\begin{bmatrix}1&1\\0&0\end{bmatrix}\right) = 1x^2+1x+0 \qquad L\left(\begin{bmatrix}0&2\\1&0\end{bmatrix}\right) = 0x^2+2x+1 \qquad L\left(\begin{bmatrix}0&0\\1&1\end{bmatrix}\right) = 0x^2+0x+2 \qquad L\left(\begin{bmatrix}0&0\\0&1\end{bmatrix}\right) = 0x^2+0x+1$$

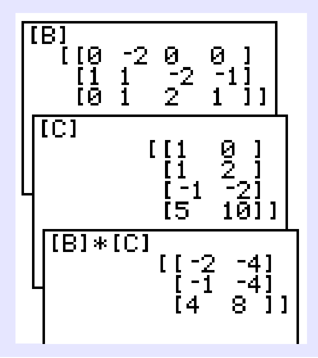
$$\begin{array}{ccc|cccc} 1x^2+0x+0 & 1x^2+1x+0 & 1x^2+1x+1 & 1x^2+1x+0 & 0x^2+2x+1 & 0x^2+0x+2 & 0x^2+0x+1 \\ \downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow&\downarrow \\ 1&1&1&1&0&0&0\\ 0&1&1&1&2&0&0\\ 0&0&1&0&1&2&1 \end{array} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|cccc} 1&0&0&0&-2&0&0\\ 0&1&0&1&1&-2&-1\\ 0&0&1&0&1&2&1 \end{array}\right] \Rightarrow [L]_{\delta\gamma} = \begin{bmatrix}0&-2&0&0\\1&1&-2&-1\\0&1&2&1\end{bmatrix}$$

Determining $[L\circ T]_{\delta\beta}$:

$$L\circ T(\mathbf{0},\mathbf{4}) = L[T(\mathbf{0},\mathbf{4})] = L\left(\begin{bmatrix}0&4\\0&8\end{bmatrix}\right) = 0x^2+4x+8$$

$$L\circ T(\mathbf{1},\mathbf{2}) = L[T(\mathbf{1},\mathbf{2})] = L\left(\begin{bmatrix}1&3\\0&4\end{bmatrix}\right) = x^2+3x+4$$

$$\begin{array}{ccc|cc} 1x^2+0x+0 & 1x^2+1x+0 & 1x^2+1x+1 & & \\ & & & \downarrow & \downarrow \\ 1&1&1&1&0\\ 0&1&1&3&4\\ 0&0&1&4&8 \end{array} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|cc} 1&0&0&-2&-4\\ 0&1&0&-1&-4\\ 0&0&1&4&8 \end{array}\right] \Rightarrow [L\circ T]_{\delta\beta} = \begin{bmatrix}-2&-4\\-1&-4\\4&8\end{bmatrix}$$



```
[B]
  [[0 -2 0  0 ]
   [1 1  -2 -1]
   [0 1  2  1 ]]
[C]
        [[1  0 ]
         [1  2 ]
         [-1 -2]
         [5  10]]
[B]*[C]
         [[-2 -4]
          [-1 -4]
          [4  8 ]]
```

And it does all fit nicely together, as you can easily check:

$$\begin{array}{ccccc} [L\circ T]_{\delta\beta} & & [L]_{\delta\gamma} & [T]_{\gamma\beta} \\ \downarrow & & \downarrow & \downarrow \\ \begin{bmatrix}-2&-4\\-1&-4\\4&8\end{bmatrix} & = & \begin{bmatrix}0&-2&0&0\\1&1&-2&-1\\0&1&2&1\end{bmatrix} & \begin{bmatrix}1&0\\1&2\\-1&-2\\5&10\end{bmatrix} \end{array}$$

CHECK YOUR UNDERSTANDING 5.16

Let $T: \Re^3 \to P_2$ and $L: (P_2 \to \Re^2)$ be the linear maps given by:

$$T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = bx^2 + ax + c$$

and

$$L(ax^2 + bx + c) = (a, a + b + c)$$

Show directly that:

$$[L \circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$$

for bases:

$$\beta = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0})\}$$

$$\gamma = \{x^2, x, 2\}$$

$$\delta = \{(\mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1})\}$$

Answer: See page B-21.

A proof of the following result is relegated to the exercises:

We recall that I_n denotes the identity matrix of dimension n, and that I_V denotes the identity map from V to V.

THEOREM 5.22 Let β be a basis for a vector space V of dimension n. Then:

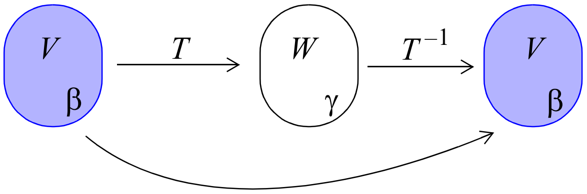
$$[I_V]_{\beta\beta} = I_n$$

As might be expected:

THEOREM 5.23 Let $T: V \to W$ be an isomorphism. Let β and γ be bases for V and W, respectively. Then:

$$[T^{-1}]_{\beta\gamma} = [T]_{\gamma\beta}^{-1}$$

PROOF: We have:



$$T^{-1} \circ T = I_V$$

$$[T^{-1} \circ T]_{\beta\beta} = [I_V]_{\beta\beta}$$

Theorems 5.21 and 22: $[T^{-1}]_{\beta\gamma}[T]_{\gamma\beta} = I$

$$[T^{-1}]_{\beta\gamma} = [T]_{\gamma\beta}^{-1}$$

CHECK YOUR UNDERSTANDING 5.17

Show that if the matrix representation of $T: V \to W$ with respect to any chosen basis β for V and γ for W is invertible, then the linear map T is itself invertible (an isomorphism).

Answer: See page B-21.

EXERCISES

Exercises 1-3. Let $T: \Re^2 \to \Re^2$ be the linear operator given by $T(\mathbf{a}, \mathbf{b}) = (\mathbf{a} + \mathbf{b}, 2\mathbf{b})$. Find $[(\mathbf{2}, \mathbf{3})]_\beta$ and $[T(\mathbf{2}, \mathbf{3})]_\beta$ for the given (ordered) basis β.

1. $\beta = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$
2. $\beta = \{(\mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{0})\}$
3. $\beta = \{(\mathbf{1}, \mathbf{2}), (-\mathbf{2}, \mathbf{1})\}$

Exercises 4-5. Let $T: \Re^3 \to \Re^3$ be the linear operator given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} + \mathbf{b}, \mathbf{b} - \mathbf{a}, \mathbf{c})$. Find $[(\mathbf{2}, \mathbf{3}, \mathbf{1})]_\beta$ and $[T(\mathbf{2}, \mathbf{3}, \mathbf{1})]_\beta$ for the given basis β.

4. $\beta = \{(\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{0}, \mathbf{1})\}$
5. $\beta = \{(-\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (-\mathbf{2}, \mathbf{1}, \mathbf{1})\}$

Exercises 6-7. Let $T: R^2 \to P_2$ be the linear map given by $T(\mathbf{a}, \mathbf{b}) = ax^2 - bx + (a - b)$. Find $[(\mathbf{1}, \mathbf{2})]_\beta$ and $[T(\mathbf{1}, \mathbf{2})]_\gamma$ for the given bases β and γ.

6. $\beta = \{(\mathbf{2}, \mathbf{1}), (\mathbf{1}, \mathbf{0})\}, \gamma = \{2x^2, -x, 2\}$
7. $\beta = \{(\mathbf{3}, -\mathbf{1}), (\mathbf{1}, \mathbf{2})\}, \gamma = \{2x^2 + 1, 2x, x^2 + x + 1\}$

Exercises 8-9. Let $T: P_3 \to M_{2 \times 2}$ be the linear map given by $T(ax^3 + bx^2 + cx + d) = \begin{bmatrix} a & -d \\ b & c \end{bmatrix}$. Find $[x^2 + x + 1]_\beta$ and $[T(x^2 + x + 1)]_\gamma$ for the given bases β and γ.

8. $\beta = \{x^3, 2x^2, x + 1, x - 1\}, \gamma = \left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$
9. $\beta = \{x^3 + x^2, x^2 + x, x + 1, 1\}, \gamma = \left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$

Exercises 10-11. Let $T: \Re^3 \to \Re^2$ be the linear map given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{a} + \mathbf{b}, 2\mathbf{c})$. Find $[T]_{\gamma\beta}$ with respect to the given bases β and γ, and show directly that $[T(\mathbf{1}, \mathbf{2}, \mathbf{1})]_\gamma = [T]_{\gamma\beta}[(\mathbf{1}, \mathbf{2}, \mathbf{1})]_\beta$.

10. $\beta = \{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}, \gamma = \{(\mathbf{2}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}$
11. $\beta = \{(\mathbf{0}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}, \gamma = \{(\mathbf{0}, \mathbf{1}), (\mathbf{2}, \mathbf{2})\}$

Exercises 12-13. Let $T: \Re^2 \to P_2$ be the linear map given by $T(\mathbf{a}, \mathbf{b}) = ax^2 - bx + (a - b)$. Find $[T]_{\gamma\beta}$ with respect to the given bases β and γ, and show directly that $[T(\mathbf{1}, \mathbf{2})]_\gamma = [T]_{\gamma\beta}[(\mathbf{1}, \mathbf{2})]_\beta$.

12. $\beta = \{(\mathbf{2}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}, \gamma = \{x, x^2, x + 1\}$
13. $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{1})\}, \gamma = \{x, x^2 + x, x^2 + x + 1\}$

Exercises 14-15. Let $T\colon P_1 \to P_2$ be the linear map given by $T(p(x)) = xp(x) + p(0)$. Show directly that $[T(2x+1)]_\gamma = [T]_{\gamma\beta}[2x+1]_\beta$ for the given bases β and γ.

14. $\beta = \{4, 2x\}, \gamma = \{4, 2x, 4x^2\}$
15. $\beta = \{x+1, 2x+3\}, \gamma = \{x, x^2+x, x^2+x+1\}$

Exercises 16-17. Let $I_{\Re^3}\colon \Re^3 \to \Re^3$ be the identity map on $\Re^3$. Find $[I_{\Re^3}]_{\beta'\beta}$ with respect to the given bases β and β', and show directly that $[I_{\Re^3}(\mathbf{1, 2, 1})]_{\beta'} = [I_{\Re^3}]_{\beta'\beta}[(\mathbf{1, 2, 1})]_\beta$

16. $\beta = \{(\mathbf{1, 1, 0}), (\mathbf{1, 0, 1}), (\mathbf{1, 1, 1})\}, \beta' = \{(\mathbf{1, 1, 0}), (\mathbf{1, 0, 1}), (\mathbf{1, 1, 1})\}$
17. $\beta = \{(\mathbf{1, 1, 0}), (\mathbf{1, 0, 1}), (\mathbf{1, 1, 1})\}, \beta' = \{(\mathbf{1, 0, 1}), (\mathbf{1, 1, 1}), (\mathbf{1, 1, 0})\}$

Exercises 18-19. Let $T\colon M_{2\times 2} \to M_{2\times 2}$ be the linear map given by $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Find $[T]_{\beta'\beta}$ for the given bases β and β', and show directly that $\left[T\left(\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}\right)\right]_{\beta'} = [T]_{\beta'\beta}\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}_\beta$.

18. $\beta = \left\{\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}, \beta' = \left\{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}$

19. $\beta = \left\{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}\right\}, \beta' = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}\right\}$

Exercises 20-22. Let $T\colon \Re^2 \to \Re^2$ be the linear operator given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a} + \boldsymbol{b}, \mathbf{2}\boldsymbol{b})$. Find $[T]_{\beta\beta}$ with respect to the given basis β, and show directly that $[T(\mathbf{1, 3})]_\beta = [T]_{\beta\beta}[(\mathbf{1, 3})]_\beta$.

20. $\beta = \{(\mathbf{1, 0}), (\mathbf{0, 1})\}$ 21. $\beta = \{(\mathbf{0, 1}), (\mathbf{1, 0})\}$ 22. $\beta = \{(\mathbf{1, 2}), (\mathbf{-2, 1})\}$

23. **(Calculus Dependent)** Let $T\colon P_2 \to P_3$ be the linear map given by $T(p(x)) = xp(x)$ and let $D\colon P_3 \to P_2$ be the differentiation linear function: $D(p(x)) = p'(x)$. Determine the given matrices for the basis $\beta = \{1, x, x^2\}$ of P_2, and the basis $\gamma = \{1, x, x^2, x^3\}$ of P_3.

    (a) $[T]_{\gamma\beta}$ (b) $[D]_{\beta\gamma}$ (c) $[D \circ T]_{\beta\beta}$
    (d) $[T \circ D]_{\gamma\gamma}$ (e) $[T \circ D \circ T]_{\gamma\beta}$ (f) $[D \circ T \circ D]_{\beta\gamma}$

24. **(Calculus Dependent)** Let V be the subspace of $F(\Re)$ spanned by the three vectors 1, $\sin x$, and $\cos x$. Let $D\colon V \to V$ be the differentiation operator. Determine $[D]_{\beta\beta}$ for $\beta = \{1, \sin x, \cos x\}$, and show directly that $[D(5 + 2\sin x)]_\beta = [D]_{\beta\beta}[5 + 2\sin x]_\beta$.

25. **(Calculus Dependent)** Let $D\colon P_3 \to P_3$ be the differentiation operator. Determine $[D]_{\beta\beta}$ for $\beta = \{1, 2x, 3x^2, 4x^3\}$, and show directly that $[D(5x^3 + 3x^2)]_\beta = [D]_{\beta\beta}[5x^3 + 3x^2]_\beta$.

26. Find the linear function $T: \Re^2 \to \Re^2$, if $[T]_{\beta\beta} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ for $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{0})\}$.

27. Find the linear function $T: \Re^2 \to \Re^2$ if $[T]_{\beta'\beta} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ for $\beta' = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{0})\}$ and $\beta = \{(\mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{2})\}$.

28. Find the linear function $T: \Re^2 \to P_2$ if $[T]_{\gamma\beta} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}$ for $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{0})\}$ and $\gamma = \{2, 2x+1, x^2\}$.

29. Find the linear function $T: M_{2\times 2} \to M_{2\times 2}$ if $[T]_{\beta'\beta} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ for

$$\beta = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\} \text{ and } \beta' = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}.$$

30. Let $T: \Re^2 \to M_{2\times 2}$ and $L: M_{2\times 2} \to \Re^3$ be the linear maps given by: $T(\boldsymbol{a}, \boldsymbol{b}) = \begin{bmatrix} a & -b \\ -a & 0 \end{bmatrix}$ and $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (b, a, c+d)$. Show directly that $[L \circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$ for bases:

$$\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}, \gamma = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}, \delta = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0})\}$$

31. Let $T: \Re^2 \to \Re^3$ and $L: \Re^3 \to P_2$ be the linear maps given by: $T(\boldsymbol{a}, \boldsymbol{b}) = (-\boldsymbol{a}, \mathbf{0}, \boldsymbol{a} + \boldsymbol{b})$ and $L(a, b, c) = bx^2 - cx + a$. Show directly that $[L \circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$ for bases:

$$\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}, \gamma = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0})\}, \delta = \{x^2, x+1, 3\}$$

32. Prove that the linear function of Theorem 5.21 is an isomorphism.

33. Prove Theorem 5.24.

34. Let $T: V \to W$ be a linear map from a vector space V of dimension n to a vector space W of dimension m. Let β and γ be bases for V and W, respectively. Show that if $A \in M_{m\times n}$ is such that $A[\boldsymbol{v}]_\beta = [T(\boldsymbol{v})]_\gamma$ for every $\boldsymbol{v} \in V$, then $A = [T]_{\gamma\beta}$.

Exercises 35-36. Prove that the function $T: \Re^2 \to P_1$ given by $T(\boldsymbol{a}, \boldsymbol{b}) = ax + 2b$ is an isomorphism. Find the linear map $T^{-1}: P_1 \to \Re^2$. Determine $[T]_{\gamma\beta}$ and $[T^{-1}]_{\beta\gamma}$ for the given bases β and γ, and show directly that $[T^{-1}]_{\beta\gamma} = [T]_{\gamma\beta}^{-1}$.

35. $\beta = \{(\mathbf{2}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}, \gamma = \{x, x+1\}$

36. $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{1})\}, \gamma = \{x, 2\}$

Exercises 37-38. Prove that the function $T: \Re^4 \rightarrow M_{2\times 2}$ given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = \begin{bmatrix} a & a+b \\ d & c-d \end{bmatrix}$ is an isomorphism. Find the linear map $T^{-1}: M_{2\times 2} \rightarrow R^4$. Determine $[T]_{\gamma\beta}$ and $[T^{-1}]_{\beta\gamma}$ for the given bases β and γ, and show directly that $[T^{-1}]_{\beta\gamma} = [T]_{\gamma\beta}^{-1}$.

37. $\beta = \{(\mathbf{1, 1, 0, 0}), (\mathbf{0, 1, 1, 0}), (\mathbf{0, 0, 1, 1}), (\mathbf{1, 0, 0, 0})\}, \gamma = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right\}$

38. $\beta = \{(\mathbf{1, 1, 1, 1}), (\mathbf{1, 1, 1, 0}), (\mathbf{1, 1, 0, 0}), (\mathbf{1, 0, 0, 0})\}, \gamma = \left\{ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$

39. Let $L: M_{3\times 2} \rightarrow M_{2\times 3}$ be given by $L(A) = A^T$ (See Exercise 19, page 161). Let:

$$\beta = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \right\}, \gamma = \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \right\}.$$

(a) Determine $[L]_{\gamma\beta}$.

(b) Show that L is an isomorphism, and use Theorem 5.25 to find $[L^{-1}]_{\beta\gamma}$.

40. Let $T: V \rightarrow V$ be a linear operator. A nontrivial subspace S of V is said to be **invariant** under T if $T(S) = \{T(\boldsymbol{v}) | \boldsymbol{v} \in S\} \subseteq S$. Assume that $\dim(V) = n$ and $\dim(S) = m$. Show that there exists a basis β for V such that $[T]_{\beta\beta} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$, where $\mathbf{0}$ is the zero $(n-m) \times m$ matrix.

41. Let $T: V \rightarrow W$ be a linear function and let β and γ be bases for the finite dimensional vector spaces V and W, respectively. Let $A = [T]_{\gamma\beta}$. Show that:

(a) $\boldsymbol{v} \in \text{Ker}(T)$ if and only if $[\boldsymbol{v}]_\beta \in \text{null}(A)$.

(b) $\boldsymbol{w} \in \text{Im}(T)$ if and only if $[\boldsymbol{w}]_\gamma$ is in the column space of A.

42. Let V and W be vector spaces of dimensions n and m, respectively. Prove that the vector space $L(V, W)$ of Exercise 35, page 122, is isomorphic to $M_{m\times n}$.
Suggestion: Let β and γ be bases for V and W, respectively. Show that the function $\phi: L(V, W) \rightarrow M_{m\times n}$ given by $\phi(T) = [T]_{\gamma\beta}$ is an isomorphism.

43. **(PMI)** Let $V_1, V_2, \ldots, V_n$ be vector spaces and let β_i be a basis for V_i, $1 \le i \le n$. Let $T_i: V_i \rightarrow V_{i+1}$ be a linear map, $1 \le i \le n-1$. Use the Principle of Mathematical Induction to show that $[T_{n-1} \circ T_{n-2} \circ \cdots \circ T_1]_{\beta_n\beta_1} = [T_{n-1}]_{\beta_n\beta_{n-1}} [T_{n-2}]_{\beta_{n-1}\beta_{n-2}} \cdots [T_2]_{\beta_3\beta_2} [T_1]_{\beta_2\beta_1}$.

PROVE OR GIVE A COUNTEREXAMPLE

44. Let $T: V \to W$ be an isomorphism, and let $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ be a basis for V. Then, for every $\boldsymbol{v} \in V$, $[\boldsymbol{v}]_\beta = [T(\boldsymbol{v})]_\gamma$, where $\gamma = \{T(\boldsymbol{v_1}), T(\boldsymbol{v_2}), \ldots, T(\boldsymbol{v_n})\}$.

45. Let $T: V \to W$ be linear, and let β and γ be bases for V and W, respectively. Let $rT: V \to W$ be defined by $(rT)(\boldsymbol{v}) = r(T(v))$. Then: $[rT]_{\gamma\beta} = r[T]_{\gamma\beta}$.

46. Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for V, and let $\beta' = \{2\boldsymbol{v}_1, 2\boldsymbol{v}_2, \ldots, 2\boldsymbol{v}_n\}$. If $T: V \to V$ is a linear operator on V, then $[T]_{\beta\beta} = 2[T]_{\beta'\beta'}$.

47. If $Z: V \to W$ is the zero transformation from the n-dimensional vector space V to the m-dimensional vector space W, then $[Z]_{\gamma\beta}$ is the $m \times n$ zero matrix for every pair of bases β and γ for V and W, respectively.

48. Let $I_V: V \to V$ be the identity map on a space V of dimension n, and let β and β' be (ordered) basis for V. Then $[I_V]_{\beta'\beta} = I_n$ if and only if $\beta = \beta'$.

49. Let $T: \Re^2 \to \Re^2$ be given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a} + \boldsymbol{3b}, \boldsymbol{2a} + \boldsymbol{2b})$. There exists a basis β such that $[T]_{\beta\beta}$ is a diagonal matrix (See Exercise 18, page 161).

50. Let $T: \Re^2 \to \Re^2$ be given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a} + \boldsymbol{3b}, -\boldsymbol{2a} + \boldsymbol{2b})$. There exists a basis β such that $[T]_{\beta\beta}$ is a diagonal matrix (See Exercise 18, page 161).

51. For $T: \Re^n \to \Re^n$ and $T: \Re^n \to \Re^n$ and any basis β for $\Re^n$: $[T + L]_{\beta\beta} = [T]_{\beta\beta} + [L]_{\beta\beta}$.

§4. CHANGE OF BASIS

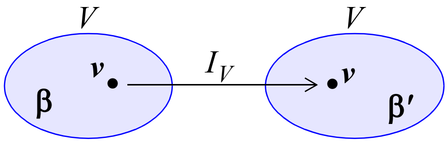
In the previous section we observed that by choosing a basis β in an n-dimensional vector space V one can associate to each vector its coordinate vector relative to β (Definition 5.9, page 177). The following result tells us how the coordinate vector of $\boldsymbol{v} \in V$ changes when switching from one basis to another:

THEOREM 5.24 For β and β' bases for the finite dimensional vector space V, and for $\boldsymbol{v} \in V$ we have:

$$[\boldsymbol{v}]_{\beta'} = [I_V]_{\beta'\beta}[\boldsymbol{v}]_{\beta}$$

where I_V denotes the identity map from V to V.

PROOF: Consider the identity map I_V: $V \to V$ along with accompanying chosen bases β and β':



Applying Theorem 5.20, page 180 we have our result:

$$[\boldsymbol{v}]_{\beta'} = [I_V(\boldsymbol{v})]_{\beta'} = [I_V]_{\beta'\beta}[I_V(\boldsymbol{v})]_{\beta} = [I_V]_{\beta'\beta}[\boldsymbol{v}]_{\beta}$$

CHANGE OF BASE MATRIX

The above matrix $[I_V]_{\beta'\beta}$ is called the **change of base matrix** from β to β'.

EXAMPLE 5.11 Find the change-of-base matrix $[I_{P_2}]_{\beta'\beta}$ for the basis $\beta = \{x^2, x, 1\}$, $\beta' = \{1+x, x+x^2, x\}$ of P_2, and then verify directly that $[\boldsymbol{v}]_{\beta'} = [I_{P_2}]_{\beta'\beta}[\boldsymbol{v}]_{\beta}$ for $\boldsymbol{v} = 2x^2+3$.

SOLUTION: Definition 5.10, page 179, tells us that the first column of $[I_{P_2}]_{\beta'\beta}$ is the coefficient matrix of $I_{P_2}(x^2)$ with respect to β', the second column is $[I_{P_2}(x)]_{\beta'}$, and the third column is $[I_{P_2}(x)]_{\beta'}$; bringing us to the following three vector equations:]

$$I_{P_2}(x^2) = x^2 = a(1+x)+b(x+x^2)+c(x)$$

$$I_{P_2}(x) = x = a(1+x)+b(x+x^2)+c(x)$$

$$I_{P_2}(1) = 1 = a(1+x)+b(x+x^2)+c(x)$$

Noting that we also have to find $[2x^2+3]_{\beta'}$ we might as well throw in the fourth vector equation:

$$2x^2+3 = a(1+x)+b(x+x^2)+c(x)$$

Equating coefficients in each of the above four vector equations we come to the following four systems of equations in the three unknowns a, b, and c:

$$x = 0x^2+1x+0 \qquad x+x^2 = 1x^2+1x+0 \qquad 1+x = 0x^2+1x+1$$

$$I_{P_2}(x^2) = x^2 = 1x^2+0x+0$$

$$I_{P_2}(x) = x = 0x^2+1x+0$$

$$I_{P_2}(1) = 1 = 0x^2+0x+1$$

$$2x^2+3 = 2x^2+0x+3$$

$$\left[\begin{array}{ccc|cccc} 0 & 1 & 0 & 1 & 0 & 0 & 2 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{rref}} \begin{array}{ccc|c|c} a & b & c & & \\ \hline 1 & 0 & 0 & 0 \quad 0 \quad 1 & 3 \\ 0 & 1 & 0 & 1 \quad 0 \quad 0 & 2 \\ 0 & 0 & 1 & \underbrace{-1 \quad 1 \quad -1}_{[I_{P_2}]_{\beta'\beta}} & \underset{[2x^2+3]_{\beta'}}{\underset{\uparrow}{5}} \end{array}$$

At this point we have two of the three matrices in the equation

$$[2x^2+3]_{\beta'} = [I_{P_2}]_{\beta'\beta}[2x^2+3]_{\beta}$$

As for $[2x^2+3]_{\beta}$, it is simply $\begin{bmatrix} 2 \\ 0 \\ 3 \end{bmatrix}$, since $\beta = \{x^2, x, 1\}$.

We leave it for you to verify that, indeed:

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

CHECK YOUR UNDERSTANDING 5.18

Let $V = \Re^2$, $\beta = \{(\mathbf{1},\mathbf{2}),(\mathbf{2},\mathbf{1})\}$, and $\beta' = \{(\mathbf{0},\mathbf{3}),(\mathbf{2},\mathbf{-1})\}$. Determine the change-of-base matrix $[I_{\Re^2}]_{\beta'\beta}$ and verify directly that $[(\mathbf{2},\mathbf{3})]_{\beta'} = [I_{\Re^2}]_{\beta'\beta}[(\mathbf{2},\mathbf{3})]_{\beta}$.

Answer: See page B-21.

EXAMPLE 5.12 Find the coordinates of the point $P = (1, 3)$ with respect to the coordinate axes obtained by rotating the standard axes by 45° in a counterclockwise direction.

SOLUTION: We are to find the coordinate vector of the point $(1, 3)$ with respect to vectors of length 1 (unit vectors) in the direction of the x'- and y'-axis depicted in Figure 5.3.

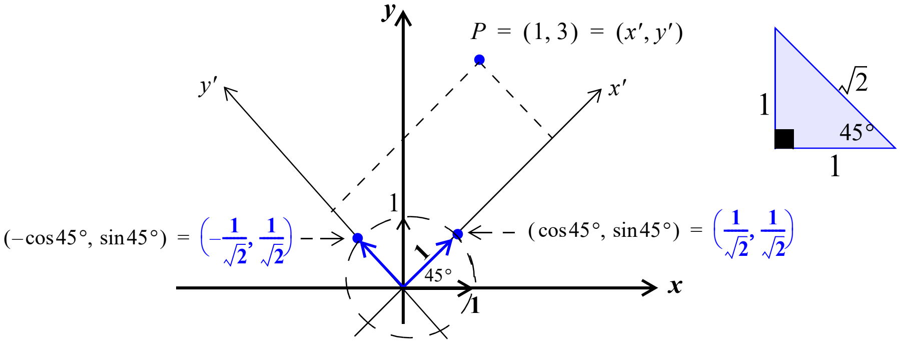


Figure 5.3

We begin by finding the change-of-base matrix $[I_{\Re^2}]_{\beta'\beta}$, for $\beta = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$ and $\beta' = \left\{\left(\frac{\mathbf{1}}{\sqrt{\mathbf{2}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}\right), \left(-\frac{\mathbf{1}}{\sqrt{\mathbf{2}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}\right)\right\}$:

$$I_{\Re^2}(\mathbf{1}, \mathbf{0}) = (\mathbf{1}, \mathbf{0}) \qquad I_{\Re^2}(\mathbf{0}, \mathbf{1}) = (\mathbf{0}, \mathbf{1})$$

$$\overbrace{\begin{array}{cc} \downarrow & \downarrow \end{array}}^{\beta'}\begin{array}{cc} \downarrow & \downarrow \end{array}$$

$$\left[\begin{array}{cc|cc} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 1 & 0 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 & 1 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc} 1 & 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ 0 & 1 & -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{array}\right]$$

Applying Theorem 5.20, page 180, we have:

$$[(\mathbf{1}, \mathbf{3})]_{\beta'} = [I]_{\beta'\beta}[(\mathbf{1}, \mathbf{3})]_{\beta} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix}$$

Conclusion: In the x', y'-coordinate system of Figure 5.3:

$$P = (2\sqrt{2}, \sqrt{2})$$

Check:

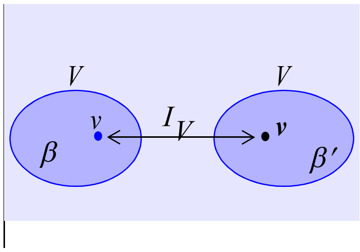
$$2\sqrt{2}\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) + \sqrt{2}\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) = (2 - 1, 2 + 1) = (2, 3)$$

CHECK YOUR UNDERSTANDING 5.19

Find the coordinates (x', y') of the point $P = (1, 3)$ with respect to the coordinate axes obtained by rotating the standard axes by $60°$ in a clockwise direction.

Answer: See page B-21.

The adjacent identity map is pointing in two directions. The left-to-right direction gives rise to the change-of-base matrix $[I_V]_{\beta'\beta}$, while the right-to-left directions brings us to $[I_V]_{\beta\beta'}$. Are $[I_V]_{\beta'\beta}$ and $[I_V]_{\beta\beta'}$ related? Yes:



In other words: $[I_V]_{\beta'\beta}$ and $[I_V]_{\beta\beta'}$ are invertible, with each being the inverse of the other.

THEOREM 5.25 Let β and β′ be two bases for a finite dimensional vector space V. Then:

$$[I_V]_{\beta'\beta} = ([I_V]_{\beta\beta'})^{-1}$$

PROOF: A direct consequence of Theorem 5.23, page 184, with $I_V\colon V_\beta \to V_{\beta'}$ playing the role of $T\colon V_\beta \to W_\gamma$ (note that $I_V^{-1} = I_V$).

We now turn our attention to the matrix representations of a linear operator $T\colon V \to V$.

A generalization of this result appears in Exercise 24.

THEOREM 5.26
CHANGE OF BASIS

Let $T\colon V \to V$ be a linear operator on a finite dimensional space V. Let β and β′ be two bases for V. Then:

$$[T]_{\beta'\beta'} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'}$$

PROOF: Consider the following figure:

In reading the composition of functions, you kind of have to read from **right to left**: the right-most function being performed first.

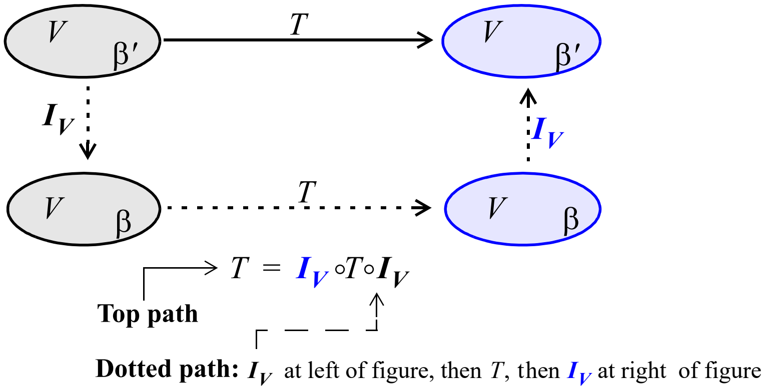


Figure 5.4

Since the identity map I_V does not move anything, we have:

$$T = I_V \circ T \circ I_V$$

and therefore: $[T]_{\beta'\beta'} = [I_V \circ T \circ I_V]_{\beta'\beta'}$

Applying Theorem 5.21, page 182, to the above equation, we have:

$$\overset{\substack{\textbf{top path in Figure 5.4}\\ \downarrow}}{[T]_{\beta'\beta'}} = \overbrace{\underset{\substack{\uparrow\\ \text{and finally this function}}}{[I_V]_{\beta'\beta}}\,\underset{\substack{\uparrow\\ \text{then this function}}}{[T]_{\beta\beta}}\,\underset{\substack{\uparrow\\ \textbf{this function first}}}{[I_V]_{\beta\beta'}}}^{\substack{\textbf{dotted path in Figure 5.4}\\ \downarrow}}$$

EXAMPLE 5.13 Verify, directly, that Theorem 5.26 holds for the linear operator $T\colon \Re^3 \to \Re^3$ given by:

$$T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{2a}, \mathbf{b} + \mathbf{c}, \mathbf{0})$$

and bases:

$$\beta = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0})\}$$
$$\beta' = \{(\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{0}, \mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{0}, \mathbf{1})\}$$

SOLUTION: We determine the four matrices:

$$[T]_{\beta'\beta'} \qquad [T]_{\beta\beta} \qquad [I_{\Re^3}]_{\beta'\beta} \qquad [I_{\Re^3}]_{\beta\beta'}$$

For $[T]_{\beta'\beta'}$

$T(\mathbf{1}, \mathbf{0}, \mathbf{1}) = (\mathbf{2}, \mathbf{1}, \mathbf{0})$

$T(\mathbf{0}, \mathbf{1}, \mathbf{0}) = (\mathbf{0}, \mathbf{1}, \mathbf{0})$

$T(\mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{0}, \mathbf{1}, \mathbf{0})$

β'

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -2 & 0 & 0 \end{array}\right]$$

For $[T]_{\beta\beta}$

$T(\mathbf{1}, \mathbf{1}, \mathbf{1}) = (\mathbf{2}, \mathbf{2}, \mathbf{0})$

$T(\mathbf{1}, \mathbf{1}, \mathbf{0}) = (\mathbf{2}, \mathbf{1}, \mathbf{0})$

$T(\mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{2}, \mathbf{0}, \mathbf{0})$

β

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 2 & 2 & 2 \\ 1 & 1 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 2 \end{array}\right]$$

For $[I_{\Re^3}]_{\beta'\beta}$

$I_{\Re}(\mathbf{1}, \mathbf{1}, \mathbf{1}) = (\mathbf{1}, \mathbf{1}, \mathbf{1})$

$I_{\Re}(\mathbf{1}, \mathbf{1}, \mathbf{0}) = (\mathbf{1}, \mathbf{1}, \mathbf{0})$

$I_{\Re}(\mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{1}, \mathbf{0}, \mathbf{0})$

β'

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & -1 & -1 \end{array}\right]$$

For $[I_{\Re^3}]_{\beta\beta'}$

$I_{\Re}(\mathbf{1, 0, 1}) = (\mathbf{1, 0, 1})$

$I_{\Re}(\mathbf{0, 1, 0}) = (\mathbf{0, 1, 0})$

$I_{\Re}(\mathbf{0, 0, 1}) = (\mathbf{0, 0, 1})$

β

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & -1 & 1 & -1 \\ 0 & 0 & 1 & 1 & -1 & 0 \end{array}\right]$$

As advertised:

$$[I]_{\beta'\beta}[T]_{\beta\beta}[I]_{\beta\beta'} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & -1 & -1 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 1 \\ -2 & 0 & 0 \end{bmatrix} = [T]_{\beta'\beta'}$$

CHECK YOUR UNDERSTANDING 5.20

Verify, directly, that Theorem 5.26 holds for the linear operator $T: P_2 \to P_2$ given by $T(ax^2 + bx + c) = bx^2 - ax + 2c$, and bases:

$$\beta = \{x^2, x^2 + x, x^2 + x + 1\}, \beta' = \{x^2 + 1, x^2 - x, 1\}$$

Answer: See page B-21.

DEFINITION 5.11

SIMILAR MATRICES

$A, B \in M_{n \times n}$ are **similar** if there exists an invertible matrix $P \in M_{n \times n}$ such that $B = P^{-1}AP$.

Theorem 5.26 tells us that if T: V $\to V$ is a linear operator on a vector space of dimension n, and if β and β' are basis for V, then $[T]_{\beta'\beta'}$ and $[T]_{\beta\beta}$ are similar matrices. The following result kind of goes in the opposite direction:

THEOREM 5.27

Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for V, and let T be a linear operator on V. If A is similar to $[T]_{\beta\beta}$, then there exists a basis β' for V, such that $A = [T]_{\beta'\beta'}$.

PROOF: Since A is similar to $[T]_{\beta\beta}$, there exist a matrix $P = [p_{ij}]$ such that $A = P^{-1}[T]_{\beta\beta}P$. In Exercise 22 you are asked to verify that $\beta' = \{\boldsymbol{v}_1', \boldsymbol{v}_2', \ldots, \boldsymbol{v}_n'\}$, where $\boldsymbol{v}_i' = p_{1i}\boldsymbol{v}_1 + p_{2i}\boldsymbol{v}_2 + \ldots + p_{ni}\boldsymbol{v}_n$, is a basis for V. Applying Theorem 5.26 we have:

$$[T]_{\beta'\beta'} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'}.$$

The i^{th} column of $[I_V]_{\beta\beta'}$, namely:
$$[I_V(v_i')]_\beta = [v_i']_\beta$$
equals the i^{th} column of P, since:
$$\boldsymbol{v}_i' = p_{1i}\boldsymbol{v}_1 + p_{2i}\boldsymbol{v}_2 + \dots + p_{ni}\boldsymbol{v}_i$$

By its very construction: $[I_V]_{\beta\beta'} = P$ (see margin). Moreover: $[I_V]_{\beta'\beta} = P^{-1}$ (Theorem 5.23, page 184). Hence:

$$[\boldsymbol{T}]_{\boldsymbol{\beta'\beta'}} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'} = P^{-1}[T]_{\beta\beta}P = \boldsymbol{A}$$

EXAMPLE 5.14

Let $T: \Re^2 \to \Re^2$ be the linear map given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{3a} + \boldsymbol{6b}, \boldsymbol{6a})$.

(a) Find $[T]_{\beta\beta}$ for $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{1})\}$.

(b) Show that $\begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix}$ is similar to $[T]_{\beta\beta}$.

(c) Find a basis β' for $\Re^2$ such that $[T]_{\beta'\beta'} = \begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix}$

$$\textbf{Solution: } \text{(a): } \left[\begin{array}{cc|cc} \overset{\beta}{\overbrace{1 \quad 2}} & & \overset{T(\mathbf{1},\,\mathbf{2}) = (\mathbf{15},\,\mathbf{6})}{\downarrow} & \overset{T(\mathbf{2},\,\mathbf{1}) = (\mathbf{12},\,\mathbf{12})}{\downarrow} \\ 1 & 2 & 15 & 12 \\ 2 & 1 & 6 & 12 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc} 1 & 0 & -1 & 4 \\ 0 & 1 & 8 & 4 \end{array}\right]^{\overset{[T]_{\beta\beta}}{\downarrow}}$$

(b) We determine $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ such that

$$\begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}\begin{bmatrix} -1 & 4 \\ 8 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

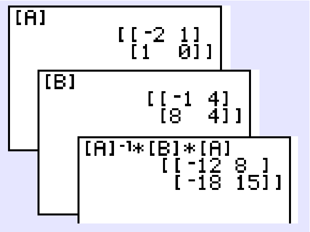
$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix} = \begin{bmatrix} -1 & 4 \\ 8 & 4 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

The above leads us to a homogeneous system of four equations in four unknowns:

$$\left.\begin{aligned} -12a - 18b &= -a + 4c \\ 8a + 15b &= -b + 4d \\ -12c - 18d &= 8a + 4c \\ 8c + 15d &= 8b + 4d \end{aligned}\right\} \Rightarrow \left.\begin{aligned} -11a - 18b - 4c + 0d &= 0 \\ 8a + 16b + 0c - 4d &= 0 \\ -8a + 0b - 16c - 18d &= 0 \\ 0a - 8b + 8c + 11d &= 0 \end{aligned}\right\}$$

$$\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} -11 & -18 & -4 & 0 \\ 8 & 16 & 0 & -4 \\ -8 & 0 & -16 & -18 \\ 0 & -8 & 8 & 11 \end{bmatrix} \end{array} \xrightarrow{\text{rref}} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 1 & 0 & 2 & \frac{9}{4} \\ 0 & 1 & -1 & -\frac{11}{8} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array} \quad \textbf{(*)}$$

Matrix (*) has two free variables, telling us that there are infinitely many $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ for which $\begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} \begin{bmatrix} -1 & 4 \\ 8 & 4 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Letting $d = 0$ and $c = 1$, we arrive at a solution, namely: $a = -2, b = 1, c = 1, d = 0$. And so we have:



```
[A]
      [[-2 1]
       [1  0]]
[B]
      [[-1 4]
       [8  4]]
[A]⁻¹*[B]*[A]
      [[-12 8 ]
       [-18 15]]
```

$$\begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} -1 & 4 \\ 8 & 4 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}$$

(c) Following the procedure spelled out in the proof of Theorem 5.27, with $P = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}$, and $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{1})\}$, we determine a basis $\beta' = \{v_1', v_2'\}$ such that $[T]_{\beta'\beta'} = \begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix}$:

$$v_1' = -2(\mathbf{1}, \mathbf{2}) + (\mathbf{2}, \mathbf{1}) = (\mathbf{0}, -\mathbf{3})$$
$$v_2' = 1(\mathbf{1}, \mathbf{2}) + 0(\mathbf{2}, \mathbf{1}) = (\mathbf{1}, \mathbf{2})$$

Let's verify that $[T]_{\beta'\beta'} = \begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix}$ for $\beta' = \{(\mathbf{0}, -\mathbf{3}), (\mathbf{1}, \mathbf{2})\}$:

$T(\mathbf{0}, -\mathbf{3}) = (-\mathbf{18}, \mathbf{0})$

$T(\mathbf{1}, \mathbf{2}) = (\mathbf{15}, \mathbf{6})$

$$\left[\begin{array}{cc|cc} \multicolumn{2}{c|}{\beta'} & \downarrow & \downarrow \\ 0 & 1 & -18 & 15 \\ -3 & 2 & 0 & 6 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc} & & \multicolumn{2}{c}{[T]_{\beta'\beta'}} \\ 1 & 0 & -12 & 8 \\ 0 & 1 & -18 & 15 \end{array}\right]$$

CHECK YOUR UNDERSTANDING 5.21

Referring to Example 5.14, determine a basis, β'' distinct from $\beta' = \{(\mathbf{0}, -\mathbf{3}), (\mathbf{1}, \mathbf{2})\}$, for which $[T]_{\beta''\beta''} = \begin{bmatrix} -12 & 8 \\ -18 & 15 \end{bmatrix}$.

Answer: See page B-21.

EXERCISES

Exercises 1-7. Verify directly that $[\boldsymbol{v}]_{\beta'} = [I_V]_{\beta'\beta}[\boldsymbol{v}]_{\beta}$ holds for the given vector space V, the vector $\boldsymbol{v} \in V$, and the bases β and β':

1. $V = \Re^2$, $\boldsymbol{v} = (\mathbf{2, 5})$, $\beta = \{(\mathbf{1, 0}), (\mathbf{0, 1})\}$, and $\beta' = \{(\mathbf{1, 2}), (\mathbf{-2, 1})\}$.
2. $V = \Re^2$, $\boldsymbol{v} = (\mathbf{3, -1})$, $\beta = \{(\mathbf{1, 0}), (\mathbf{0, 1})\}$, and $\beta = \{(\mathbf{1, 2}), (\mathbf{-2, 1})\}$.
3. $V = \Re^3$, $\boldsymbol{v} = (\mathbf{2, 3, -1})$, $\beta = \{(\mathbf{1, 1, 0}), (\mathbf{1, 0, 1}), (\mathbf{1, 1, 1})\}$, and
   $\beta' = \{(\mathbf{1, 0, 0}), (\mathbf{0, 1, 0}), (\mathbf{0, 0, 1})\}$.
4. $V = \Re^3$, $\boldsymbol{v} = (\mathbf{3, 0, -2})$, $\beta = \{(\mathbf{1, 0, 2}), (\mathbf{0, 0, 1}), (\mathbf{1, 1, 2})\}$, and
   $\beta' = \{(\mathbf{1, 0, 1}), (\mathbf{0, 1, 1}), (\mathbf{1, 1, 0})\}$,
5. $V = P_2$, $\boldsymbol{v} = 2x^2 + x + 1$, $\beta = \{x^2, x, x + 1\}$, and $\beta' = \{x, x^2 + x, x^2 + x + 1\}$.
6. $V = P_2$, $\boldsymbol{v} = x^2 + 1$, $\beta = \{2, 2x, 2x^2\}$, and $\beta' = \{1, x + 1, x^2 + x\}$.
7. $V = M_{2\times 2}$, $\boldsymbol{v} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, $\beta = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$, and
   $\beta' = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$.
8. Find the coordinates of the point $P = (x, y)$ in the xy-plane with respect to the coordinate axes obtained by rotating the standard axes $\theta°$ in a counterclockwise direction. (See Example 5.12.)

Exercises 9-13. Verify directly that $[T]_{\beta'\beta'} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'}$ holds for the given vector space V, the linear operator T, and the bases β and β':

9. $V = \Re^2$, $T\colon V \to V$ given by $T(\boldsymbol{a, b}) = (\boldsymbol{-b, a})$, $\beta = \{(\mathbf{1, 0}), (\mathbf{0, 1})\}$, and
   $\beta' = \{(\mathbf{1, 2}), (\mathbf{-2, 1})\}$.
10. $V = \Re^3$, $T\colon V \to V$ given by $T(\boldsymbol{a, b, c}) = (\boldsymbol{-b, a, c})$,
    $\beta = \{(\mathbf{0, 1, 0}), (\mathbf{1, 0, 1}), (\mathbf{1, 1, 1})\}$, and $\beta' = \{(\mathbf{1, 2, 0}), (\mathbf{0, -2, 1}), (\mathbf{1, 0, 1})\}$.
11. $V = P_2$, $T\colon V \to V$ given by $T[ax^2 + bx + c] = cx^2 + b$, $\beta = \{x^2, x + 1, 1\}$, and
    $\beta' = \{2, x, 1 + x^2\}$.
12. $V = M_{2\times 2}$, $T\colon V \to V$ given by $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} -b & c \\ d & -a \end{bmatrix}$,
    $\beta = \left\{ \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \right\}$, and $\beta' = \left\{ \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$.

13. **(Calculus Dependent)** $V = P_3$, $T\colon V \to V$ given by $T(p(x)) = p'(x)$, $\beta = \{1, x, x^2, x^3\}$, and $\beta' = \{1, x, 2x^2, 3x^3\}$.

14. Let $T\colon \Re^2 \to \Re^2$ be the linear operator given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{2a}, \boldsymbol{-b})$. Find a basis β' for $\Re^2$ such that $[T]_{\beta'\beta'} = P^{-1}[T]_{\beta\beta}P$, where $\beta = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$ and $P = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

15. Let $T\colon \Re^3 \to \Re^3$ be a linear operator. Find the basis β' for $\Re^3$ such that $[T]_{\beta'\beta'} = P^{-1}[T]_{\beta\beta}P$, where: $\beta = \{(\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$ and $P = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$.

16. Let $T\colon P_2 \to P_2$ be a linear operator. Find the basis β' for P_2 such that $[T]_{\beta'\beta'} = P^{-1}[T]_{\beta\beta}P$, where $\beta = \{1, x+1, x^2+x\}$ and $P = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$.

17. Show that $\begin{bmatrix} 2 & 0 \\ -2 & -1 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$ are similar.

18. Show that $\begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix}$ and $\begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}$ are not similar.

19. Find all matrices that are similar to the identity matrix I_n.

20. Let $T\colon \Re^2 \to \Re^2$ be the linear map given by $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{a} + \boldsymbol{b}, \boldsymbol{b})$.
    (a) Find $[T]_{\beta\beta}$ for $\beta = \{(\mathbf{1}, \mathbf{2}), (\mathbf{2}, \mathbf{1})\}$.
    (b) Show that $\begin{bmatrix} 20 & 8 \\ -38 & -17 \end{bmatrix}$ is similar to $[T]_{\beta\beta}$.
    (c) Find a basis β' for $\Re^2$ such that $[T]_{\beta'\beta'} = \begin{bmatrix} 20 & 8 \\ -38 & -17 \end{bmatrix}$.

21. Show that "*similar*" is an equivalence relation on $M_{n \times n}$. (See Exercises 37-39, page 147 for the definition of an equivalence relation).

22. Show that $\beta' = \{\boldsymbol{v}_1', \boldsymbol{v}_2', \ldots, \boldsymbol{v}_n'\}$ in the proof of Theorem 5.27 is a basis for V.

23. Let $A, B \in M_{n \times n}$ be similar. Show that there exists a linear operator $T\colon \Re^n \to \Re^n$ and bases β and β' for $\Re^n$ such that $A = [T]_{\beta\beta}$ and $B = [T]_{\beta'\beta'}$.

24. **(A generalization of Theorem 5.26)** Let $T\colon V \to W$ be a linear map from the finite dimensional vector space V to the finite dimensional vector space W. Let β and β' be bases for V, and let γ and γ' be bases for W. Prove that: $[T]_{\gamma'\beta'} = [I_W]_{\gamma'\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'}$.

Exercises 25-29. Referring to Exercise 24, show directly that $[T]_{\gamma'\beta'} = [I_W]_{\gamma'\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'}$ holds for the given linear transformation $T: V \to W$, the bases β and β' for V, and the bases γ and γ' for W.

25. $V = \Re^2, W = \Re^3$, $T(\boldsymbol{a}, \boldsymbol{b}) = (-\boldsymbol{b}, \boldsymbol{a}, \boldsymbol{a} + \boldsymbol{b})$, $\beta = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$,
$\beta' = \{(\mathbf{1}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}$, $\gamma = \{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (-\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$, and
$\gamma' = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{0}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}$

26. $V = \Re^3, W = \Re^2$, $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a} + \boldsymbol{b}, -\boldsymbol{c})$, $\beta = \{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (-\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$,
$\beta' = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{0}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}$, $\gamma = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$, and $\gamma' = \{(\mathbf{1}, \mathbf{2}), (\mathbf{0}, \mathbf{1})\}$.

27. $V = \Re^3, W = P_2$, $T(a, b, c) = bx^2 + cx - a$, $\beta = \{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (-\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$,
$\beta' = \{(\mathbf{1}, \mathbf{1}, \mathbf{1}), (\mathbf{0}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}$, $\gamma = \{2, 1 + x, 2 - x^2\}$, and $\gamma' = \{1, x, x^2\}$.

28. $V = P_2, W = \Re^3$, $T(ax^2 + bx + c) = (\boldsymbol{a} + \boldsymbol{b}, \mathbf{0}, -\boldsymbol{c})$, $\beta = \{1, x, x^2\}$,
$\beta' = \{2, 1 + x, 2 - x^2\}$, $\beta = \{2, 1 + x, 2 - x^2\}$, $\gamma = \{(\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{0}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0})\}$ and
$\gamma' = \{(\mathbf{1}, \mathbf{1}, \mathbf{0}), (-\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{1})\}$.

29. $V = P_2, W = P_1$, $T(ax^2 + bx + c) = bx + (a + c)$, $\beta = \{x^2, x, x + 1\}$,
$\beta' = \{2, 1 + x, 2 - x^2\}$, $\gamma = \{x, x - 1\}$ and $\gamma' = \{x, 2x + 1\}$.

30. Let $T: V \to W$ be linear. Let β, β' be bases for the n-dimensional space V, and let γ, γ' be bases for the m-dimensional space W. Prove that there exists an invertible matrix $Q \in M_{m \times m}$ and an invertible matrix $P \in M_{n \times n}$ such that $[T]_{\gamma'\beta'} = Q[T]_{\gamma\beta}P$.
Suggestion: Consider Exercise 24.

31. Let $T: V \to W$ and $L: W \to Z$ be linear maps. Let β, β' be bases for V, γ, γ' be bases for W, and δ, δ' be bases for Z. Show that $[L \circ T]_{\delta'\beta'} = [I_Z]_{\delta'\delta}[L]_{\delta\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'}$.

PROVE OR GIVE A COUNTEREXAMPLE

32. Let $T: V \to V$ be a linear operator, and let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and β' be a bases for V. If $[T]_{\beta'\beta'} = 2[T]_{\beta}\beta$, then $\beta' = \{2\boldsymbol{v}_1, 2\boldsymbol{v}_2, \ldots, 2\boldsymbol{v}_n\}$.

33. If A and B are similar matrices, then A^2 and B^2 are also similar.

34. If A and B are similar invertible matrices, then A^{-1} and B^{-1} are also similar.

35. If A and B are similar matrices, then at least one of them must be invertible.

36. If A and B are similar matrices, then so are their transpose. (See Exercise 19, page 161.)

37. If A and B are similar matrices, and if A is symmetric, then so is B. (See Exercises 20, page 161.)

38. If A and B are similar matrices, and if A is idempotent, then so is B. (See Exercises 22, page 162.)

39. If A and B are similar matrices, then $\text{Trace}(A) = \text{Trace}(B)$. (See Exercises 24, page 162.)

CHAPTER SUMMARY

MULTIPLYING MATRICES	You can perform the product $\boldsymbol{C}_{m \times n} = \boldsymbol{A}_{m \times r}\boldsymbol{B}_{r \times n}$ of two matrices, if the number of columns of ***A*** equals the number of rows of ***B***, and you get c_{ij} of the product matrix ***C*** by running across the i^{th} row of ***A*** and down the j^{th} column of ***B***: $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \cdots + a_{ir}b_{rj}$
Properties	Assuming that the matrix dimensions are such that the given operations are defined, we have: (i) $A(B+C) = AB+AC$ (ii) $(A+B)C = AC+BC$ (iii) $A(BC) = (AB)C$ (iv) $r(AB) = (rA)B = A(rB)$
A connection between matrix multiplication and linear transformations.	For $A \in M_{m \times n}$ and any positive integer z, the map $T_A : M_{n \times z} \Rightarrow M_{m \times z}$ given by $T_A(X) = AX$ is linear. In particular: For $A \in M_{m \times n}$ the map $T_A \colon \Re^n \to \Re^m$ given by $T_A(X) = AX$ is linear, where X is a vertical n-tuple and $T_A(X)$ is a vertical m-tuple.
NULL SPACE OF A MATRIX, AND NULLITY	For $A \in M_{m \times n}$, the **null space** of A is the subspace of $\boldsymbol{R}^n$ consisting of the solutions of the homogeneous linear system of equations $A\boldsymbol{X} = \boldsymbol{0}$. It is denoted by null$(A)$.
INVERTIBLE MATRIX	A **square matrix *A*** is said to be **invertible** if there exists a matrix ***B*** (necessarily of the same dimension as ***A***), such that: $\boldsymbol{AB} = \boldsymbol{BA} = \boldsymbol{I}$ The matrix ***B*** is then said to be the **inverse** of ***A***, and we write $\boldsymbol{B} = \boldsymbol{A}^{-1}$. If no such ***B*** exists, then ***A*** is said to be **non-invertible**, or **singular.**
Need only "work" on one side	Let ***A*** be a square matrix. If ***B*** is a square matrix such that **either** $\boldsymbol{AB} = \boldsymbol{I}$ **or** $\boldsymbol{BA} = \boldsymbol{I}$, then ***A*** is invertible and $\boldsymbol{A}^{-1} = \boldsymbol{B}$.
Uniqueness	An invertible matrix has a **unique** inverse.

Properties	If $\boldsymbol{A}$ is invertible and $r \neq 0$, then $r\boldsymbol{A}$ is also invertible, and: $(A^{-1})^{-1} = A$ $(r\boldsymbol{A})^{-1} = \frac{1}{r}\boldsymbol{A}^{-1}$ $(\boldsymbol{A}_1\boldsymbol{A}_2\cdots\boldsymbol{A}_n)^{-1} = \boldsymbol{A}_n^{-1}\boldsymbol{A}_{n-1}^{-1}\ldots\boldsymbol{A}_1^{-1}$

ELEMENTARY MATRIX	A matrix that is obtained by performing an elementary row operation of an identity matrix
Invertibility	Every elementary matrix is invertible.
Inverses by means of multiplication	The matrix obtained by performing an elementary row operation on a matrix $\boldsymbol{A} \in M_{m \times n}$ equals that matrix obtained by multiplying $\boldsymbol{A}$ on the left by the elementary matrix obtained by performing the same elementary row operation on the identity matrix $\boldsymbol{I_m}$.
Inverses by row-reduction	If a sequence of elementary row operations reduces the invertible matrix $\boldsymbol{A}$ to $\boldsymbol{I}$, then applying the same sequence of elementary row operations on $\boldsymbol{I}$ will yield $\boldsymbol{A}^{-1}$.
EQUIVALENCES OF INVERTIBILITY	Let $A \in M_{n \times n}$. The following are equivalent: (i) A is invertible. (ii) $A\boldsymbol{X} = \boldsymbol{B}$ has a unique solution for every $\boldsymbol{B} \in M_{n \times 1}$. (iii) $A\boldsymbol{X} = \boldsymbol{0}$ has only the trivial solution. (iv) $\text{rref}(A) = I$. (v) A is a product of elementary matrices.

COORDINATE VECTOR	$[\boldsymbol{v}]_\beta = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ where $\boldsymbol{v} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n$ and $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for V
MATRIX REPRESENTATION OF A LINEAR MAP	Let $T\colon V \to W$ be linear, $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for V, and and $\gamma = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ a basis for W. The **matrix representation** of T with respect to β and γ is that matrix $[T]_{\gamma\beta} \in M_{m \times n}$ whose i^{th} column is $[T(\boldsymbol{v}_i)]_\gamma$

The matrix representation of a linear map T describes the "action" of T.	Let $T: V \to W$ be linear. If $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for V, and $\gamma = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ is a basis for W, then: $$[T(\boldsymbol{v})]_\gamma = [T]_{\gamma\beta}[\boldsymbol{v}]_\beta$$
The matrix of a composition function is the product of matrices of those functions.	Let $T: V \to W$ and $L: W \to Z$ be linear maps, and let β, γ and δ be bases for the finite dimensional spaces V, W, and Z, respectively. Then: $$[L \circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta}$$
Relating coordinate vectors with respect to different bases.	Let β and β' be bases for V. Then: $$[\boldsymbol{v}]_{\beta'} = [I]_{\beta'\beta}[\boldsymbol{v}]_\beta$$
The matrix of the inverse of a transformation is the inverse of the matrix of that transformation.	Let $T: V \to W$ be an invertible transformation and let β and γ be bases for V and W, respectively. Then: $$[T^{-1}]_{\beta\gamma} = [T]_{\gamma\beta}^{-1}$$
Relating matrix representations of a linear operator with respect to different bases.	Let $T: V \to V$ be a linear operator, and let β and β' be two bases for V. Then: $$[T]_{\beta'} = [I_V]_{\beta'\beta}[T]_\beta[I]_{\beta\beta'}$$
SIMILAR MATRICES	The matrices $A, B \in M_{n \times n}$ are **similar** if there exists an invertible matrix $P \in M_{n \times n}$ such that $B = P^{-1}AP$.
Similar matrices represent linear maps with respect to different basis.	Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for V, and let T be a linear operator on V. If A is similar to $[T]_{\beta\beta}$, then there exists a basis β' for V, such that $A = [T]_{\beta'\beta'}$.

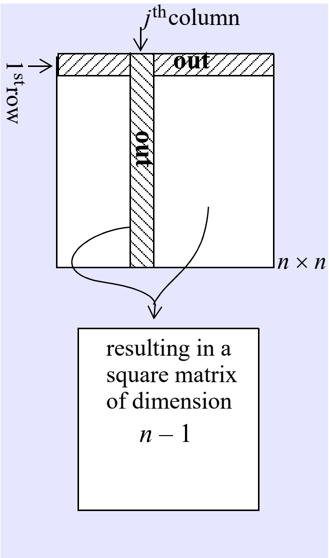
CHAPTER 6

DETERMINANTS AND EIGENVECTORS

As you know, a linear operator $T: V \to V$ has a matrix representation $[T]_{\beta\beta}$, which depends on the chosen basis β for V. A main goal of this chapter is to determine if there exists a basis β for which $[T]_{\beta\beta}$ turns out to be a diagonal matrix. Determinants, as you will see, play an essential role in that endeavor.

§1. DETERMINANTS

Using mathematical inductgion, we define a function that assigns to each square matrix a (real) number:



DEFINITION 6.1 DETERMINANT

For a 2×2 matrix:

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$$

For $A \in M_{n \times n}$, with $n > 2$, let A_{1j} denote the $(n-1) \times (n-1)$ matrix obtained by deleting the first row and j^{th} column of the matrix A (see margin). Then:

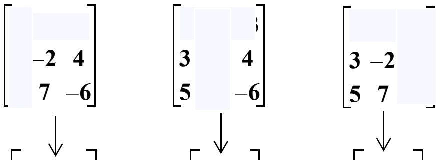
$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j})$$

IN WORDS: Multiply each entry a_{1j} in the first row of A by the determinant of the (smaller) matrix obtained by discarding the first row and the j^{th} column of A, and then sum those n products with alternating signs (starting with a + sign). L

EXAMPLE 6.1

Evaluate: $\det\begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix}$

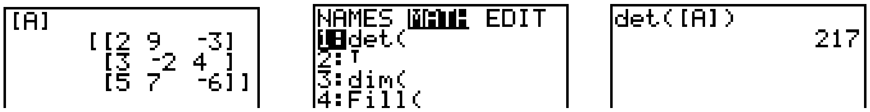
SOLUTION:



$$\det\begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix} = 2\det\begin{bmatrix} -2 & 4 \\ 7 & -6 \end{bmatrix} - 9\det\begin{bmatrix} 3 & 4 \\ 5 & -6 \end{bmatrix} - 3\det\begin{bmatrix} 3 & -2 \\ 5 & 7 \end{bmatrix}$$

$$= 2[(-2)(-6) - (4 \cdot 7)] - 9[3(-6) - (4 \cdot 5)] - 3[3 \cdot 7 - 5(-2)] = 217$$

GRAPHING CALCULATOR GLIMPSE 6.1



```
[A]
      [[2 9  -3]
       [3 -2 4 ]
       [5 7  -6]]
```

```
det([A])
                217
```

Definition 6.1 defines the determinant of a matrix by an expansion process involving the first row of the given matrix. The next theorem, known as the Laplace Expansion Theorem, enables one to expand along any row or column of the matrix. A proof of this important result is offered at the end of the section.

Note that the sign of the $(-1)^{i+j}$ has an alternating checkerboard pattern

+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	−
−	+	−	+	−	+	−	+

THEOREM 6.1

For given $A \in M_{n \times n}$, A_{ij} will denote the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the i^{th} row and j^{th} column of A. We then have:

EXPANDING ALONG THE i^{th} ROW

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

and:

EXPANDING ALONG THE j^{th} COLUMN

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$$

Note: $\det(A_{ij})$ is called the **minor** of a_{ij}, and $C_{ij} = (-1)^{i+j} a_{ij} \det(A_{ij})$ is called the $[i, j]^{\text{th}}$ **cofactor** of A

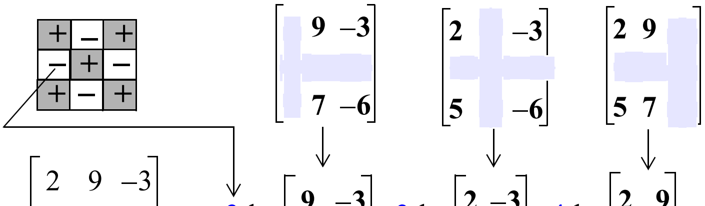
EXAMPLE 6.2

Evaluate:

$$\det \begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix}$$

by expanding about its second row.

SOLUTION:



$$\det \begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix} = -3\det \begin{bmatrix} 9 & -3 \\ 7 & -6 \end{bmatrix} - 2\det \begin{bmatrix} 2 & -3 \\ 5 & -6 \end{bmatrix} - 4\det \begin{bmatrix} 2 & 9 \\ 5 & 7 \end{bmatrix}$$

same as in Example 6.1

$$= -3[9(-6) - 7(-3)] - 2[2(-6) - 5(-3)] - 4[2 \cdot 7 - (5 \cdot 9)] = 217$$

CHECK YOUR UNDERSTANDING 6.1

Evaluate $\det\begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix}$ expanding along:

(a) The third row. (b) The second column.

Answer: See page B-23.

While it was not so bad to calculate the 3×3 determinant of Example 6.1, the task gets increasingly more tedious as the dimension of the matrix increases. If we were only interested in calculating determinants of matrices with numerical entries, then we could avoid the whole mess entirely and simply use a calculator. But this will not always be the case.

In any event, one can easily calculate the determinant of any **upper triangular** matrix (see margin):

An **upper triangular** matrix is a square matrix with zero entries below its main diagonal. For example:

$$\begin{bmatrix} 2 & 5 & 0 & 1 & -2 \\ 0 & 1 & 7 & -2 & 0 \\ 0 & 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 9 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

A **lower triangular** matrix is a square matrix with zero entries above its main diagonal. For example:

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

THEOREM 6.2 The determinant of an upper diagonal matrix equals the product of the entries along its diagonal.

PROOF: By induction on the dimension, n, of $M_{n \times n}$.

I. Claim holds for $n = 2$: $\det\begin{bmatrix} a & b \\ 0 & d \end{bmatrix} = ad - b \cdot 0 = ad$.

II. Assume claim holds for $n = k$.

III. We establish validity at $n = k + 1$:

Let $A = [a_{ij}] \in M_{(k+1) \times (k+1)}$. Since all entries in the first column below its first entry a_{11} is zero, expanding about the first column of A we have $\det(A) = a_{11}\det(A_{11})$ (where A_{11} is the k by k upper triangular matrix obtained by removing the first row and first column from the matrix A. As such, by II: $\det(A_{11}) = a_{22}a_{33}\cdots a_{(k+1)(k+1)}$. Consequently:

$$\det(A) = a_{11}\det(A_{11}) = a_{11}a_{22}a_{33}\cdots a_{(k+1)(k+1)}$$

COROLLARY For any n: $\det(I_n) = 1$.

CHECK YOUR UNDERSTANDING 6.2

Prove that the determinant of a lower diagonal matrix equals the product of the entries along its diagonal.

Answer: See page B-23.

ROW OPERATIONS AND DETERMINANTS

Since it is easy to find the determinant of an upper triangular matrix, ad since any square matrix can be transformed into an upper triangular matrix by means of elementary row operations, it would be nice to have relations between the determinant of a matrix and that obtained by performing an elementary row operations on that matrix. Niceness is at hand:

THEOREM 6.3

(a) If two rows of $A \in M_{n \times n}$ are interchanged, then the determinant of the resulting matrix is $-\det(A)$.

(b) If one row of A is multiplied by a constant c, then the determinant of the resulting matrix is $c[\det(A)]$.

(c) If a multiple of one row of A is added to another row of A, then the determinant of the resulting matrix is $\det(A)$.

PROOF: (a) By induction on the dimension of the matrix A. For $n = 2$:

$$\det\begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \quad \text{and} \quad \det\begin{bmatrix} c & d \\ a & b \end{bmatrix} = cb - da$$

negative of each other

Assume the claim holds for matrices of dimension $k > 2$ (the induction hypothesis).

Let $A = [a_{ij}]$ be a matrix of dimension $k + 1$, and let $B = [b_{ij}]$ denote the matrix obtained by interchanging rows p and q of A. Let i be the index of a row **other than** p and q. Expanding about row i we have:

$$\det(A) = \sum_{j=1}^{k+1} (-1)^{i+j} a_{ij} \det(A_{ij}) \quad \text{and} \quad \det B = \sum_{j=1}^{k+1} (-1)^{i+j} b_{ij} \det(B_{ij})$$

Since rows p and q were switched to go from A to B, row i of B still equals that of A, and therefore: $b_{ij} = a_{ij}$. Since B_{ij} is the matrix A_{ij} with two of its rows interchanged, and since those matrices are of dimension k, we have: $\mathbf{det}\boldsymbol{B_{ij}} = \mathbf{-det}\boldsymbol{A_{ij}}$ (the induction hypothesis).

Consequently:

$$\mathbf{det}(\boldsymbol{B}) = \sum_{j=1}^{k+1} (-1)^{i+j} b_{ij} \mathbf{det}(\boldsymbol{B_{ij}}) = \sum_{j=1}^{k+1} (-1)^{i+j} a_{ij} [\mathbf{-det}(\boldsymbol{A_{ij}})]$$

$$= -\sum_{j=1}^{k+1} (-1)^{i+j} a_{ij} \det(A_{ij}) = -\mathbf{det}(\boldsymbol{A})$$

(b) Let B denote the matrix obtained by multiplying row i of matrix A by c. Expanding both matrices about the i^{th} row, we have:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \text{ and } \det(B) = \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(B_{ij})$$

Matrix A and B differ only in the i^{th} row, and that row has been removed from both A and B to arrive at the matrices A_{ij} and B_{ij}.

Since $b_{ij} = ca_{ij}$, and since $B_{ij} = A_{ij}$ (margin), we have:

$$\mathbf{det}(\boldsymbol{B}) = \sum_{j=1}^{n} (-1)^{i+j} b_{ij} \det(B_{ij}) = \sum_{j=1}^{n} (-1)^{i+j} (ca_{ij}) \det(A_{ij})$$

pull out that common factor, c: $= c \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) = \boldsymbol{c}[\mathbf{det}(\boldsymbol{A})]$

(c) Let B be the matrix obtained by multiplying row r of A by c and adding it to row i. If $i = r$, then the result follows from (b). Assume $i \neq r$. Expanding B about its i^{th} row, we have:

$$\mathbf{det}(\boldsymbol{B}) = \sum_{j=1}^{n} (-1)^{i+j} (a_{ij} + ca_{rj}) \det(A_{ij})$$

$$= \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) + c \sum_{j=1}^{n} (-1)^{i+j} a_{rj} \det(A_{ij})$$

$$= \mathbf{det}(\boldsymbol{A}) + \boldsymbol{c} \sum_{j=1}^{n} (-1)^{i+j} a_{rj} \det(A_{ij}) = \mathbf{det}(\boldsymbol{A}) + \boldsymbol{c}\,0 = \mathbf{det}(\boldsymbol{A})$$

$\sum_{j=1}^{n} (-1)^{i+j} a_{rj} \det(A_{ij})$ is the determinant of a matrix with two equal rows: the i^{th} and r^{th} row

The result now follows from CYU 6.3 below

CHECK YOUR UNDERSTANDING 6.3

Show that if two rows of a matrix A are identical, then $\det(A) = 0$.

Answer: See page B-23.

The following example illustrates how Theorem 6.3 can effectively be used to calculate determinants.

EXAMPLE 6.3

Evaluate: $\det\begin{bmatrix}1&2&6&8\\1&0&0&2\\2&1&3&5\\-1&1&1&1\end{bmatrix}$

SOLUTION:

$$\det\begin{bmatrix}1&2&6&8\\1&0&0&2\\2&1&3&5\\-1&1&1&1\end{bmatrix}\overset{\text{Theorem 6.3(a)}}{=}-\det\begin{bmatrix}1&0&0&2\\1&2&6&8\\2&1&3&5\\-1&1&1&1\end{bmatrix}\overset{\substack{\text{Theorem 6.3(c)}\\\text{(applied three times)}}}{=}-\det\begin{bmatrix}1&0&0&2\\0&2&6&6\\0&1&3&1\\0&1&1&3\end{bmatrix}$$

$$\overset{\text{Theorem 6.3(b)}}{=}-2\det\begin{bmatrix}1&0&0&2\\0&1&3&3\\0&1&3&1\\0&1&1&3\end{bmatrix}\underset{\substack{\text{Theorem 6.3(c)}\\\text{(applied two times)}}}{=}-2\det\begin{bmatrix}1&0&0&2\\0&1&3&3\\0&0&0&-2\\0&0&-2&0\end{bmatrix}\underset{\text{Theorem 6.3(a)}}{=}2\det\begin{bmatrix}1&0&0&2\\0&1&3&3\\0&0&-2&0\\0&0&0&-2\end{bmatrix}\underset{\text{Theorem 6.2}}{=}8$$

CHECK YOUR UNDERSTANDING 6.4

Evaluate:

$$\det\begin{bmatrix}2&1&0&1\\0&1&2&2\\1&0&1&4\\4&1&1&3\end{bmatrix}$$

Answer: See page B-23.

To establish any of the following three elementary matrix results, you need but substitute the identity matrix I_n for A in Theorem 6.3:

Note that $\det I_n = 1$

THEOREM 6.4

(a) If $E \in M_{n \times n}$ is obtained by interchanging two rows of I_n, then $\det(E) = -1$.

(b) If $E \in M_{n \times n}$ is obtained by multiplying a row of I_n by $c \neq 0$, then $\det(E) = c$.

(c) If $E \in M_{n \times n}$ is obtained by adding a multiple of one row of I_n to another row, then $\det(E) = 1$.

The restriction $c \neq 0$ is imposed in (b) since we are concerned with elementary row operations (see page 3).

Soon, we will be in a position to show that the determinant of a product of any two $n \times n$ matrices is equal to the product of their determinants. For now:

THEOREM 6.5 For any $B \in M_{n \times n}$ and any $n \times n$ elementary matrix E:

$$\det(EB) = \det(E)\det(B)$$

PROOF: Let $E \in M_{n \times n}$ be an elementary matrix obtained by interchanging two rows of I_n. By Theorem 5.13, page 169:

(*) EB is the matrix obtained by interchanging the same two rows in the matrix B.

Consequently:

$$\det(EB) \underset{\uparrow}{=} -\det(B) \underset{\uparrow}{=} \det(E)\det(B)$$

(*) and Theorem 6.4(a)

$\det(E) = -1$ [Theorem 6.4(a)]

As for the rest of the proof:

CHECK YOUR UNDERSTANDING 6.5

(a) Establish Theorem 6.5, for the elementary matrix $E \in M_{n \times n}$:

(i) Obtained by multiplying a row of I_n by $c \neq 0$.

(ii) Obtained by adding a multiple of one row of I_n to another row of I_n.

(b) Let $B \in M_{n \times n}$ and $E_1, E_2, \ldots, E_s \in M_{n \times n}$ be elementary matrices. Use the Principle of Mathematical Induction to show that:

$$\begin{aligned}\det(E_s \cdots E_2 E_1 B) &= \det(E_s \cdots E_2 E_1)\det(B)\\ &= \det(E_s)\cdots\det(E_2)\det(E_1)\det(B)\end{aligned}$$

Answer: See page B-24.

We now come to one of the most important results of this section:

THEOREM 6.6 A matrix $A \in M_{n \times n}$ is invertible if and only if $\det(A) \neq 0$.

You can add this result to the list of equivalences for invertibility appearing in Theorem 5.17, page 172:

(vi) $\det(A) \neq 0$

PROOF: Let $E_1, E_2, \ldots, E_s$ be a sequence of elementary matrices such that $E_s \cdots E_2 E_1 A = \text{rref}(A)$ (Theorem 5.13, page 169). Appealing to CYU 6.5(b), we have:

$$\det(E_s)\cdots\det(E_2)\det(E_1)\det(A) = \det[\text{rref}(A)]$$

By Theorem 6.4, $\det(E_s)\ldots\det(E_2)\det(E_1) \neq 0$. Consequently:

$\det[\text{rref}(A)] \neq 0$ if and only if $\det(A) \neq 0$

margin: if and only if $\text{rref}(A) = I$

Theorem 5.17(iv), page 172: if and only if A is invertible

If rref$A = I$, then:

$$\det[\text{rref}(A)] = 1 \neq 0$$

If $\text{rref}(A) \neq I$, then its last row consists entirely of zeros, and

$$\det[\text{rref}(A)] = 0$$

We are now in a position to establish another powerful result, attributed to Cauchy:

Austin Cauchy, a prolific French mathematician (1789-1857).

THEOREM 6.7 For $A, B \in M_{n \times n}$:

$$\det(AB) = \det(A)\det(B)$$

PROOF:

Case 1: A is invertible. By Theorem 5.17, page 172, A can be expressed as a product of elementary matrices:

$$A = E_s \ldots E_2 E_1$$

Then: $\mathbf{det}(\boldsymbol{AB}) = \det(E_s \ldots E_2 E_1 B)$

CYU 6.5(b): $= \det(E_s \ldots E_2 E_1)\det(B) = \mathbf{det}(\boldsymbol{A})\mathbf{det}(\boldsymbol{B})$

Case 2: A is not invertible. AB is not invertible; for:

$$AB \text{ invertible } \Rightarrow \exists C \ni (AB)C = I$$
$$\Rightarrow A(BC) = I \Rightarrow A \text{ invertible -- a contradiction.}$$

It follows, from Theorem 6.6, that both $\det(AB)$ and $\det(A)$ are zero, and we again have equality:

$$\mathbf{det}(\boldsymbol{AB}) = 0 = 0[\det(B)] = \mathbf{det}(\boldsymbol{A})\mathbf{det}(\boldsymbol{B})$$

CHECK YOUR UNDERSTANDING 6.6

Prove that if A is invertible, then:

$$\det(A^{-1}) = \frac{1}{\det(A)}$$

Answer: See page B-24.

For the brave at heart:

PROOF OF THE LAPLACE EXPANSION THEOREM

The column-expansion part of the theorem is relegated to the exercises.

We use induction on n to show that the determinant of $A \in M_{n \times n}$ can be evaluated by expanding about any row of A:

The claim is easily seen to hold when $n = 2$:

$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22} - a_{12}a_{21} = -a_{21}a_{12} + a_{11}a_{22}$$

Assume that the claim holds for $n = k$ (the induction hypothesis).

Let $A \in M_{(k+1)(k+1)}$. We show that for any $1 < t \leq k+1$:

This will show that the expansion about **any** row equals that of expanding about the first row.

$$\underbrace{\sum_{j=1}^{k+1} (-1)^{1+j} a_{1j} \det(A_{1j})}_{(*)} = \underbrace{\sum_{s=1}^{k+1} (-1)^{t+s} a_{ts} \det(A_{ts})}_{(**)}$$

expanding about first row ↑

↑ expanding about row t

Working with (*), we employ the induction hypothesis and evaluate the determinant of each $k \times k$ matrix A_{1j} along its $t-1$ row, which is row t of A (see Figure 6.1). In so doing, we need to keep in mind that just as each A_{1j} has a row and a column removed from A (1^{st} row and j^{th} column), so then will each submatrix in the expansion of $\det(A_{ij})$ have two rows and two columns of A removed.

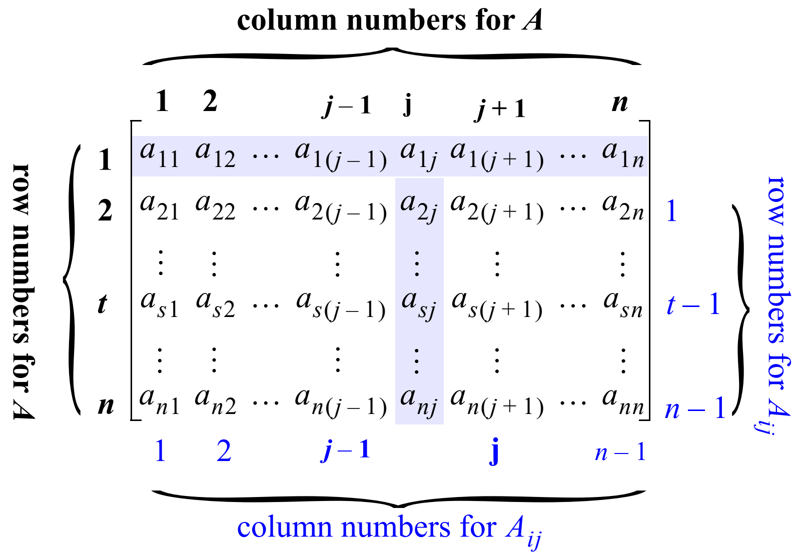


Figure 6.1

Let $A_{1t,js}$ denote the submatrix of A with rows 1 and t, and columns j and s of A removed. Using the induction hypothesis we can obtain the determinant of A_{1j} by expanding about its $t-1$ row (which is the t^{th} row of A), breaking that sum into two pieces the "before-j" piece, and the "after-j" piece we have:

$$\det(A_{1j}) = \sum_{s=1}^{j-1} (-1)^{(t-1)+s}\ a_{ts}\det(A_{1t,js}) + \sum_{s=j+1}^{k+1} (-1)^{(t-1)+(s-1)}\ a_{ts}\ \det(A_{1t,js})$$

Bringing us to:

$$\sum_{j=1}^{k+1} (-1)^{1+j} a_{1j}\det(A_{1j})$$

$$= \sum_{j=1}^{k+1} (-1)^{1+j}\ a_{1j}\left(\sum_{s=1}^{j-1} (-1)^{(t-1)+s}\ a_{ts}\det(A_{1t,js}) + \sum_{s=j+1}^{k+1} (-1)^{(t-1)+(s-1)}\ a_{ts}\ \det(A_{1t,js})\right)$$

$$= \sum_{s<j} \underset{\overset{\uparrow}{\boxed{1+J+(t-1)+s}}}{(-1)^{j+t+s}} a_{1j}a_{ts}\det(A_{1t,js}) + \sum_{s>j} \underset{\overset{\uparrow}{\boxed{1+j+(t-1)+(s-1)}}}{(-1)^{j+t+s-1}} a_{1j}a_{ts}\det(A_{1t,js}) \quad \textbf{(A)}$$

Turning to **(**)**, we again appeal to the induction hypothesis, and expand about the first row to calculate each $\det(A_{ts})$:

$$\sum_{s=1}^{k+1}(-1)^{t+s}a_{ts}\det(A_{ts})$$

$$= \sum_{s=1}^{k+1}(-1)^{t+s}a_{ts}\left(\sum_{j=1}^{s-1}(-1)^{1+j}\ a_{1j}\det(A_{1t,js}) + \sum_{j=s+1}^{k+1}(-1)^{1+(j-1)}\ a_{1j}\ \det(A_{1t,js})\right)$$

$$= \sum_{j<s}(-1)^{t+s+j+1}\ a_{ts}a_{1j}\det(A_{1t,js}) + \sum_{j>s}(-1)^{t+s+j}\ a_{ts}a_{1j}\det(A_{1t,js}) \qquad \textbf{(B)}$$

To complete the proof, we observe that the left summation in **(A)** is equal to the right summation in **(B)**, and that the right summation in **(A)** equals the left in **(B)**:

$$\sum_{s<j}(-1)^{j+t+s}a_{1j}a_{ts}\det(A_{1t,js}) = \sum_{j>s}(-1)^{t+s+j}\ a_{ts}a_{1j}\det(A_{1t,js})$$

$$\sum_{s>j}(-1)^{j+t+s-1}a_{1j}a_{ts}\det(A_{1t,js}) = \sum_{j<s}(-1)^{t+s+j+1}\ a_{ts}a_{1j}\det(A_{1t,js})$$

note that $(-1)^{j+t+s-1} = (-1)^{t+s+j+1}$

EXERCISES

Exercises 1-8. Compute the determinant of the given matrix.

1. $\begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 5 & 3 \\ -1 & 0 & 1 \\ 4 & 2 & 9 \end{bmatrix}$

3. $\begin{bmatrix} 2 & 2 & 4 \\ 5 & 7 & 11 \\ 3 & 6 & 9 \end{bmatrix}$

4. $\begin{bmatrix} 1 & 2 & 4 \\ -9 & 1 & 9 \\ 4 & 6 & 1 \end{bmatrix}$

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

6. $\begin{bmatrix} 0 & 3 & 0 & 1 \\ -1 & 2 & 4 & 2 \\ 4 & 0 & 4 & 0 \\ 1 & 2 & 3 & 5 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 3 & 0 & 4 \\ -1 & 2 & 4 & 6 \\ 0 & 0 & 4 & 2 \\ 1 & 0 & 3 & 5 \end{bmatrix}$

8. $\begin{bmatrix} 6 & 3 & 3 & 9 \\ 3 & -3 & 3 & 6 \\ 3 & 0 & 3 & 12 \\ 9 & 0 & 3 & 6 \end{bmatrix}$

Exercises 9-14. Given that $\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = 9$, find:

9. $\det\begin{bmatrix} g & h & i \\ a & b & c \\ d & e & f \end{bmatrix}$

10. $\det\begin{bmatrix} a & b & c \\ 2d & 2e & 2f \\ -3g & -3h & -3i \end{bmatrix}$

11. $\det\begin{bmatrix} a+2d & b+2e & c+2f \\ d & e & f \\ g & h & i \end{bmatrix}$

12. $\det\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

13. $\det\begin{bmatrix} a & b & c \\ g & h & i \\ d & e & f \end{bmatrix}^2$

14. $\det\begin{bmatrix} a+d-g & b+e-h & c+f-i \\ d & e & f \\ g & h & i \end{bmatrix}$

Exercises 15-18. Find all values of k for which the given matrix is invertible.

15. $\begin{bmatrix} k & 1 \\ k & k \end{bmatrix}$

16. $\begin{bmatrix} k & -1 & 0 \\ 0 & k & -1 \\ 2 & 4 & 1 \end{bmatrix}$

17. $\begin{bmatrix} k & k^2 & 0 \\ 0 & k & k^2 \\ k^2 & k & 0 \end{bmatrix}$

18. $\begin{bmatrix} k & 0 & 0 & 1 \\ 0 & k & 1 & 0 \\ 1 & 0 & k & 0 \\ 0 & 1 & 1 & k \end{bmatrix}$

Exercises 19-22. Verify:

19. $\det\begin{bmatrix} x & -1 & 0 \\ 0 & x & -1 \\ c & b & a \end{bmatrix} = ax^2+bx+c$

20. $\det\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix} = (a-b)(b-c)(c-a)$

21. $\det\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{bmatrix} = (a-b)(b-c)(c-a)(a+b+c)$

22. $\det\begin{bmatrix} 1+a & 1 & 1 \\ 1 & 1+b & 1 \\ 1 & 1 & 1+c \end{bmatrix} = abc+bc+ac+bc$

23. While one can certainly find matrices $A, B \in M_{n \times n}$ such that $AB \neq BA$, prove that one can not find matrices $A, B \in M_{n \times n}$ such that $\det(AB) \neq \det(BA)$.

24. Show that the matrix $\begin{bmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{bmatrix}$ is invertible if and only if the numbers a, b, and c, are all distinct.

25. Prove that if a matrix A contains a row (or column) consisting entirely of zeros, then $\det(A) = 0$.

26. If $D = [d_{ij}] \in M_{n \times n}$ is a diagonal matrix and if $X_i \in M_{n \times 1}$ is the column matrix whose i^{th} entry is 1 and all other entries are 0, then $DX_i = d_{ii}X_i$.

27. Let $A \in M_{n \times n}$. Prove that $\det(A) = \det(A^T)$, where A^T denotes the transpose of A (see Exercise 19, page 161).

28. Prove that if $A \in M_{n \times n}$ is skew-symmetric, then $\det(A) = (-1)^n \det(A)$ (see Exercise 21, page 162). What conclusion can you draw from this result?

29. For $A \in M_{n \times n}$, let B be obtained from A by interchanging pairs of rows of A m times. Express $\det(B)$ as a function of m and $\det(A)$.

30. Let A be similar to B (see Definition 5.11, page 195). Prove that:

    (a) $\det(A) = \det(B)$ (b) $\det(A - cI) = \det(B - cI)$ for every $c \in \Re$.

    Suggestion: Consider Theorem 5.1, page 153.

31. Show that $\det\begin{bmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{bmatrix} = 0$ is an equation of the line passing through the points (x_1, y_1) and (x_2, y_2) in $\Re^2$.

32. Show that $\det\begin{bmatrix} x & y & z & 1 \\ x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \\ x_3 & y_3 & z_3 & 1 \end{bmatrix} = 0$ is an equation of the plane passing through the points (x_1, y_1, z_1), (x_2, y_2, z_2), and (x_3, y_3, z_3) in $\Re^3$.

33. Show that the area of the triangle with vertices at (x_1, y_1), (x_2, y_2), and (x_3, y_3) is given by $\pm\frac{1}{2}\det\begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{bmatrix}$, where the sign (±) is chosen to yield a positive number.

34. (Cramer's Rule) If $AX = B$ is a system of n equations in n unknowns, with A invertible, then the system has a unique solution $(x_1, x_2, \ldots, x_n)$ [Theorem 5.17(ii), page 172]. Cramer's rule asserts that:

    $$x_1 = \frac{\det(A_1)}{\det(A)}, x_2 = \frac{\det(A_2)}{\det(A)}, \ldots, x_n = \frac{\det(A_n)}{\det(A)}$$

    where A_i is the matrix obtained by replacing the i^{th} column of A with B.

    Use Cramer's rule to solve the system of:

    (a) Example 1.3, page 9. (b) CYU 1.3, page 10.

35. Prove the “column-expansion-part” of Theorem 6.3 (Laplace Expansion Theorem).

Exercises 36-39. Use the Principle of Mathematical Induction to show that:

36. For any m and $A_i \in M_{n \times n}$, $\det(A_1 \cdot A_2 \cdots A_m) = \det(A_1)\det(A_2)\cdots\det(A_m)$.

37. For any $A \in M_{n \times n}$ and $c \in R$: $\det(cA) = c^n \det(A)$.

38. Prove that for $A \in M_{n \times n}$ and any positive integer m: $\det(A^m) = [\det(A)]^m$.

39. If $A \in M_{n \times n}$ is of the form $A = \begin{bmatrix} cI & X \\ 0 & Y \end{bmatrix}$, where I is the $r \times r$ identity matrix, 0 is the $(n-r) \times r$ zero matrix, and X and Y are $r \times (n-r)$ and $(n-r) \times (n-r)$ matrices, respectively, then: $\det(A) = c^r \det(Y)$.

40. If $A \in M_{n \times n}$ is of the form $A = \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix}$, where X and Z are square matrices and 0 is a zero matrix, then: $\det(A) = \det(X)\det(Z)$.

PROVE OR GIVE A COUNTEREXAMPLE

41. For $A \in M_{n \times n}$, if $\det A = 1$, then $A = I$.

42. For $A \in M_{n \times n}$, if $\det(A) = 0$, then A is the zero matrix.

43. For $A, B \in M_{n \times n}$, if $\det(AB) = 1$, then both A and B are invertible and $A = B^{-1}$.

44. For any $A, B \in M_{n \times n}$, $\det(A+B) = \det(A) + \det(B)$.

45. For any $A \in M_{n \times n}$, $\det(-A) = -\det(A)$.

46. If $A \in M_{n \times n}$ is nilpotent, then $\det(A) = 0$ (see Exercise 23, page 162).

47. If $A \in M_{4 \times 4}$ and $X, Y, Z, W \in M_{2 \times 2}$, and if $A = \begin{bmatrix} X & Y \\ Z & W \end{bmatrix}$, then:

$\det(A) = \det(X)\det(W) - \det(Y)\det(Z)$.

§2. EIGENSPACES

The German word *eigen* translates to: *characteristic*. At one time, eigenvalues were called *latent values*, and it is for this reason that λ (lamba), the Greek letter for "*l*" is used.

We remind you that we use $\overline{\Re^n}$ to denote $M_{n \times 1}$, and that $\bar{v} \in \overline{\Re^n}$ is the vector $v \in \Re^n$ in "column form."

We begin by defining eigenvalues and eigenvectors for matrices:

DEFINITION 6.2

EIGENVALUES AND EIGENVECTORS (FOR MATRICES)

An **eigenvalue** of a matrix $A \in M_{n \times n}$ is a <u>scalar</u> $\lambda \in \Re$ (which may be zero) for which there exists a <u>nonzero</u> vector $X \in \overline{\Re^n}$ such that:

$$A(X) = \lambda X$$

Any such vector X is said to be an **eigenvector** corresponding to the eigenvalue λ.

EXAMPLE 6.4 Show that $\overline{(2, 1)}$ and $\overline{(-1, 3)}$ are eigenvectors of the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$.

SOLUTION: Since $\begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 8 \\ 4 \end{bmatrix} = 4\begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $\overline{(2, 1)}$ is an eigenvector of A corresponding to the eigenvalue 4. By the same token, $\overline{(-1, 3)}$ is an eigenvector corresponding to the eigenvalue -3:

$$\begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}\begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix} = -3\begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Since the set of eigenvectors corresponding to an eigenvalue λ of a matrix $A \in M_{n \times n}$ does not contain the zero vector, it cannot be a subspace of $\overline{\Re^n}$. If you throw in the zero vector, however, you do end up with a subspace:

Recall that $\text{null}(A)$ denotes the solution set of the homogeneous system of equations $AX = \mathbf{0}$.

$$\overset{\substack{\text{eigenvectors, along with the zero vector}\\ \downarrow}}{\{\mathbf{X}|\mathbf{AX} = \boldsymbol{\lambda}\mathbf{X}\}} = \{X|AX - \lambda X = \mathbf{0}\} = \{X|AX - \lambda \overset{\substack{n\text{-by-}n\text{ identity matrix}\\ \downarrow}}{I_n} X = \mathbf{0}\}$$

$$= \{(A - \lambda I_n)X = \mathbf{0}\} = \underset{\substack{\uparrow\\ \text{a subspace of } \Re^n \text{ (Theorem 5.4, page 159)}}}{\mathbf{null}(\boldsymbol{A} - \boldsymbol{\lambda}\boldsymbol{I_n})}$$

Bringing us to:

DEFINITION 6.3 The **eigenspace** associated with an eigenvalue λ of a matrix $A \in M_{n \times n}$, denoted by $E[\lambda]$, is given by:

$$E[\lambda] = \text{null}(A - \lambda I_n)$$

EXAMPLE 6.5 Find a basis for the eigenspace $E[4]$ of the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$ of Example 6.4.

$\text{null}\left(\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\right)$ is the solution set of the homogeneous system:

$$\left.\begin{aligned} -x + 2y &= 0 \\ 3x - 6y &= 0 \end{aligned}\right\}$$

SOLUTION: $E[4] = \text{null}\left(\begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix} - 4\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}\right)$

From $\begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} \mathbf{1} & -2 \\ 0 & 0 \end{bmatrix}$ we see that $E[4] = \{\overline{(2r, r)} \mid r \in R\}$ with basis $\{\overline{(2, 1)}\}$.

CHECK YOUR UNDERSTANDING 6.7

Find a basis for the eigenspace $E[-3]$ of the matrix $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$ of Example 6.4.

Answer: See page B-24.

CHARACTERISTIC POLYNOMIALS

At this point, there is a gap in our eigenvector development; namely:

How does one go about finding the eigenvalues of a matrix?

The answer hinges on the following objects:

DEFINITION 6.4
CHARACTERISTIC POLYNOMIAL (FOR MATRICES)

For $A \in M_{n \times n}$, the n-degree polynomial $\det(A - \lambda I_n)$ is said to be the **characteristic polynomial** of A, and $\det(A - \lambda I_n) = 0$ is said to be the **characteristic equation** of A.

We are now in the position to state the main theorem of this section:

THEOREM 6.8 The eigenvalues of $A \in M_{n \times n}$ are the solutions of the characteristic equation $\det(A - \lambda I_n) = 0$.

PROOF: To say that λ is an eigenvalue of A is to say that there exists a **nonzero** vector $\boldsymbol{X} \in \overline{\Re^n}$ such that:

$$\begin{aligned} A\boldsymbol{X} &= \lambda\boldsymbol{X} \\ A\boldsymbol{X} - \lambda\boldsymbol{X} &= \mathbf{0} \\ A\boldsymbol{X} - (\lambda I_n)\boldsymbol{X} &= \mathbf{0} \\ (A - \lambda I_n)\boldsymbol{X} &= \mathbf{0} \end{aligned}$$

But, to say that $(A - \lambda I_n)\boldsymbol{X} = \mathbf{0}$ has a **nontrivial** solution is to say that $\det(A - \lambda I_n) = 0$ (Theorem 6.6, page 212, and Theorem 5.19(iii), page 172).

EXAMPLE 6.6 Find the eigenvalues and corresponding eigenspaces of the matrix:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

SOLUTION: The eigenvalues are the solutions of the equation:

$$\det(A - \lambda I_3) = \det\begin{bmatrix} 1-\lambda & 0 & 1 \\ 2 & 2-\lambda & 1 \\ \mathbf{1} & \mathbf{0} & \mathbf{1-\lambda} \end{bmatrix} = 0$$

Expanding about the third row, we have:

A better choice is to expand about the second column. If you do, pay particular attention to the checkerboard sign pattern of page 206.

$$1 \cdot \mathbf{det}\begin{bmatrix} \mathbf{0} & \mathbf{1} \\ \mathbf{2-\lambda} & \mathbf{1} \end{bmatrix} - 0 \cdot \det\begin{bmatrix} 1-\lambda & 1 \\ 2 & 1 \end{bmatrix} + (1-\lambda)\mathbf{det}\begin{bmatrix} \mathbf{1-\lambda} & \mathbf{0} \\ \mathbf{2} & \mathbf{2-\lambda} \end{bmatrix} = 0$$

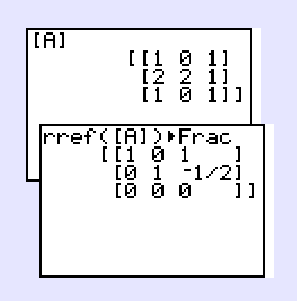
$$\searrow \qquad\qquad\qquad\qquad \swarrow$$

$$\begin{aligned} -(\mathbf{2-\lambda}) + (1-\lambda)(\mathbf{1-\lambda})(\mathbf{2-\lambda}) &= 0 \\ (2-\lambda)[-1 + (1-\lambda)^2] &= 0 \\ (2-\lambda)(-1 + 1 - 2\lambda + \lambda^2) &= 0 \\ (2-\lambda)(-\lambda)(2-\lambda) &= 0 \\ -\lambda(2-\lambda)^2 &= 0 \\ \lambda = 0, \lambda &= 2 \end{aligned}$$

We now determine the eigenspaces for the two eigenvalues, **0** and **2**.

Finding $E[0]$:

$$E[0] = \text{null}\left(\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} - 0\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}\right)$$

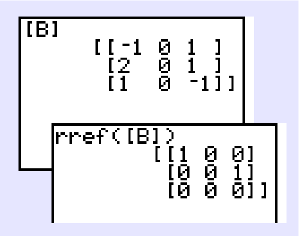

```
[A]
        [[1 0 1]
         [2 2 1]
         [1 0 1]]
rref([A])▶Frac
     [[1 0 1   ]
      [0 1 -1/2]
      [0 0 0   ]]
```

$$\begin{array}{ccc} x & y & z \\ \hline 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{array} \xrightarrow{\textbf{rref}} \begin{array}{ccc} x & y & z \\ \hline \mathbf{1} & 0 & 1 \\ 0 & \mathbf{1} & -\frac{1}{2} \\ 0 & 0 & 0 \end{array}$$

Setting the free variable z equal to r, we have:

$$E[0] = \left\{\overline{\left(-r, \frac{r}{2}, r\right)} \,\middle|\, r \in \Re\right\} = \{\overline{(-\mathbf{2}\boldsymbol{r}, \boldsymbol{r}, \mathbf{2}\boldsymbol{r})} \mid r \in R\}$$

with basis: $\{\overline{(-\mathbf{2}, \mathbf{1}, \mathbf{2})}\}$.



```
[B]
    [[-1  0  1 ]
     [2   0  1 ]
     [1   0  -1]]
rref([B])
    [[1 0 0]
     [0 0 1]
     [0 0 0]]
```

Finding $E[2]$:

$$E[2] = \text{null}\left(\begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix}\right)$$

$$\begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

Setting the free variable y equal to r, we have:

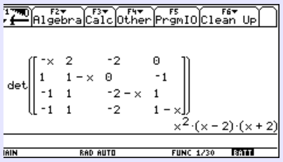
$$E[2] = \{\overline{(\mathbf{0}, \boldsymbol{r}, \mathbf{0})} \mid r \in \Re\}$$

with basis $\{\overline{(\mathbf{0}, \mathbf{1}, \mathbf{0})}\}$.

EXAMPLE 6.7 Find the eigenvalues and corresponding eigenspaces of the matrix:

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

A TI-92 teaser:



```
det([[-x  2    -2    0  ]
     [1   1-x  0     -1 ]
     [-1  1    -2-x  1  ]
     [-1  1    -2    1-x]])
                          x²·(x-2)·(x+2)
```

SOLUTION: To find the eigenvalues of A we need to solve the characteristic equation:

$$\det(A - \lambda I_4) = \det\begin{bmatrix} -\lambda & 2 & -2 & 0 \\ 1 & 1-\lambda & 0 & -1 \\ -1 & 1 & -2-\lambda & 1 \\ -1 & 1 & -2 & 1-\lambda \end{bmatrix} = 0$$

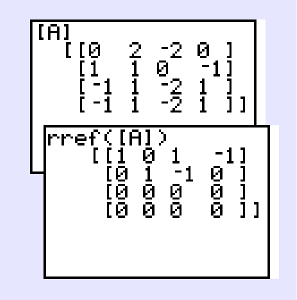
In this endeavor, we first use Theorem 6.3, page 208, to express $\det(A - \lambda I_4)$ as the determinant of a matrix in upper triangular form, and then take advantage of Theorem 6.2, page 207, to finish the job:

$$\det\begin{bmatrix}-\lambda & 2 & -2 & 0\\ 1 & 1-\lambda & 0 & -1\\ -1 & 1 & -2-\lambda & 1\\ -1 & 1 & -2 & 1-\lambda\end{bmatrix} \overset{\text{switch the first two rows}}{=} -\det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ -\lambda & 2 & -2 & 0\\ -1 & 1 & -2-\lambda & 1\\ -1 & 1 & -2 & 1-\lambda\end{bmatrix} \overset{(\lambda R_1+R_2\to R_2);(R_1+R_3\to R_3);(R_1+R_4\to R_4)}{=} -\det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ 0 & 2+\lambda-\lambda^2 & -2 & -\lambda\\ 0 & 2-\lambda & -2-\lambda & 0\\ 0 & 2-\lambda & -2 & -\lambda\end{bmatrix}$$

$$\overset{\text{switch rows 2 and 4 (introduces another minus sign)}}{=} \det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ 0 & 2-\lambda & -2 & -\lambda\\ 0 & 2-\lambda & -2-\lambda & 0\\ 0 & 2+\lambda-\lambda^2 & -2 & -\lambda\end{bmatrix} \overset{-1R_2+R_3\to R_3}{=} \det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ 0 & 2-\lambda & -2 & -\lambda\\ 0 & 0 & -\lambda & \lambda\\ 0 & 2+\lambda-\lambda^2 & -2 & -\lambda\end{bmatrix} \overset{(-1-\lambda)R_2+R_4\to R_4}{=} \det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ 0 & 2-\lambda & -2 & -\lambda\\ 0 & 0 & -\lambda & \lambda\\ 0 & 0 & 2\lambda & \lambda^2\end{bmatrix}$$

$$\overset{2R_3+R_4\to R_4}{=} \det\begin{bmatrix}1 & 1-\lambda & 0 & -1\\ 0 & 2-\lambda & -2 & -\lambda\\ 0 & 0 & -\lambda & \lambda\\ 0 & 0 & 0 & \lambda^2+2\lambda\end{bmatrix} \underset{\text{Theorem 6.2, page 207}}{=} (1)(2-\lambda)(-\lambda)(\lambda^2+2\lambda) = \lambda^2(\lambda-2)(\lambda+2)$$

Setting $\det(A-\lambda I_4) = \lambda^2(\lambda-2)(\lambda+2)$ to 0, we see that there are three distinct eigenvalues: $\lambda = 0, \lambda = 2$, and $\lambda = -2$. As for their corresponding eigenspaces:

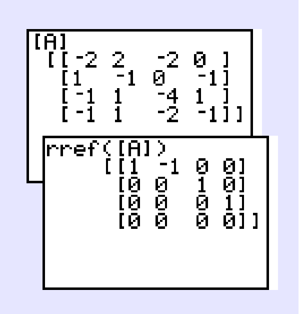


```
[A]
  [[0  2 -2 0 ]
   [1  1 0  -1]
   [-1 1 -2 1 ]
   [-1 1 -2 1 ]]
rref([A])
    [[1 0 1  -1]
     [0 1 -1 0 ]
     [0 0 0  0 ]
     [0 0 0  0 ]]
```

$$E[0] = \text{null}\left(\begin{bmatrix}0 & 2 & -2 & 0\\ 1 & 1 & 0 & -1\\ -1 & 1 & -2 & 1\\ -1 & 1 & -2 & 1\end{bmatrix} - 0\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{bmatrix}\right) = \text{null}\left(\begin{bmatrix}0 & 2 & -2 & 0\\ 1 & 1 & 0 & -1\\ -1 & 1 & -2 & 1\\ -1 & 1 & -2 & 1\end{bmatrix}\right)$$

$$\begin{array}{cccc} x & y & z & w\\ \hline 0 & 2 & -2 & 0\\ 1 & 1 & 0 & -1\\ -1 & 1 & -2 & 1\\ -1 & 1 & -2 & 1\end{array} \xrightarrow{\text{rref}} \begin{array}{cccc} x & y & z & w\\ \hline \mathbf{1} & 0 & 1 & -1\\ 0 & \mathbf{1} & -1 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\end{array}$$

Setting the free variables z and w equal to r and s respectively, we find that $E[0] = \{\overline{(\mathbf{-r+s, r, r, s})}\,|\,r, s \in \Re\}$. By letting $r = 1, s = 0$, and then $r = 0, s = 1$ we arrive at the basis $\{\overline{(\mathbf{-1, 1, 1, 0})}, \overline{(\mathbf{1, 0, 0, 1})}\}$.

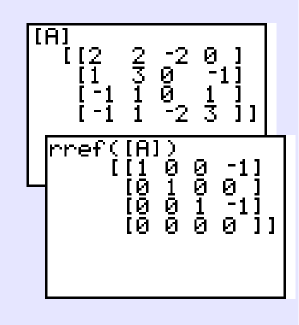


```
[A]
 [[-2 2   -2 0 ]
  [1   -1 0   -1]
  [-1 1   -4 1 ]
  [-1 1   -2 -1]]
rref([A])
        [[1 -1 0 0]
         [0 0  1 0]
         [0 0  0 1]
         [0 0  0 0]]
```

$$E[2] = \text{null}\left(\begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix} - 2\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} -2 & 2 & -2 & 0 \\ 1 & -1 & 0 & -1 \\ -1 & 1 & -4 & 1 \\ -1 & 1 & -2 & -1 \end{bmatrix}\right)$$

$$\begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} -2 & 2 & -2 & 0 \\ 1 & -1 & 0 & -1 \\ -1 & 1 & -4 & 1 \\ -1 & 1 & -2 & -1 \end{bmatrix} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} \mathbf{1} & -1 & 0 & 0 \\ 0 & 0 & \mathbf{1} & 0 \\ 0 & 0 & 0 & \mathbf{1} \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Setting the free variable y equal to r, we have $E[2] = \{\overline{(\boldsymbol{r}, \boldsymbol{r}, \mathbf{0}, \mathbf{0})} | r \in R\}$ with basis $\{\overline{(\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0})}\}$.



```
[A]
   [[2   2  -2 0 ]
    [1   3  0   -1]
    [-1 1  0   1 ]
    [-1 1  -2 3 ]]
rref([A])
        [[1 0 0 -1]
         [0 1 0 0 ]
         [0 0 1 -1]
         [0 0 0 0 ]]
```

$$E[-2] = \text{null}\left(\begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix} + 2\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} 2 & 2 & -2 & 0 \\ 1 & 3 & 0 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & -2 & 3 \end{bmatrix}\right)$$

$$\begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} 2 & 2 & -2 & 0 \\ 1 & 3 & 0 & -1 \\ -1 & 1 & 0 & 1 \\ -1 & 1 & -2 & 3 \end{bmatrix} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{cccc} x & y & z & w \end{array} \\ \begin{bmatrix} \mathbf{1} & 0 & 0 & -1 \\ 0 & \mathbf{1} & 0 & 0 \\ 0 & 0 & \mathbf{1} & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{array}$$

Setting the free variable w equal to r, we find that $E[-2] = \{\overline{(r, 0, r, r)} | r \in \Re\}$ with basis $\{\overline{(1, 0, 1, 1)}\}$.

CHECK YOUR UNDERSTANDING 6.8

Find the eigenvalues and corresponding eigenspaces of the matrix:

$$A = \begin{bmatrix} 16 & 3 & 2 \\ -4 & 3 & -8 \\ -2 & -6 & 11 \end{bmatrix}$$

Answer: See page B-24.

TURNING TO LINEAR OPERATORS

Shifting our attention from matrices to linear operators we have:

Compare with Definition 6.3, page 218.

DEFINITION 6.5

EIGENVALUES AND EIGENVECTORS
(FOR LINEAR OPERATORS)

Let $T: V \to V$ be a linear operator. An **eigenvalue** of T is a scalar $\lambda \in \Re$ (which may be zero) for which there exists a nonzero vector $v \in V$ such that:

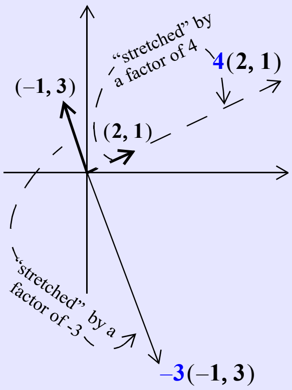
$$T(\boldsymbol{v}) = \lambda \boldsymbol{v}$$

Any such $\boldsymbol{v}$ is then said to be an **eigenvector** corresponding to the eigenvalue λ.

EXAMPLE 6.8 Show that $(\mathbf{2}, \mathbf{1})$ and $(\mathbf{-1}, \mathbf{3})$ are eigenvectors of the linear operator $T: \Re^2 \to \Re^2$ given by:

$$T(\boldsymbol{a}, \boldsymbol{b}) = (\mathbf{3}\boldsymbol{a} + \mathbf{2}\boldsymbol{b}, \mathbf{3}\boldsymbol{a} - \mathbf{2}\boldsymbol{b})$$

Note that the linear map T stretches the eigenvector $(2, 1)$ by its eigenvalue **4**, and $(-1, 3)$ by **−3**:



SOLUTION:

(a) $T(\mathbf{2}, \mathbf{1}) = (\mathbf{6} + \mathbf{2}, \mathbf{6} - \mathbf{2}) = (\mathbf{8}, \mathbf{4}) = \mathbf{4}(\mathbf{2}, \mathbf{1})$

and $T(\mathbf{-1}, \mathbf{3}) = (\mathbf{-3} + \mathbf{6}, \mathbf{-3} - \mathbf{6}) = (\mathbf{3}, \mathbf{-9}) = \mathbf{-3}(\mathbf{-1}, \mathbf{3})$.

We see that $(\mathbf{2}, \mathbf{1})$ is an eigenvector corresponding to the eigenvalue **4**, and that $(\mathbf{-1}, \mathbf{3})$ is an eigenvector corresponding to the eigenvalue **−3**.

Since the set of eigenvectors corresponding to an eigenvalue λ of a linear operator $T: V \to V$ does not contain the zero vector, it cannot be a subspace of V. As it was with matrices, however, if you throw in the zero vector, then you do end up with a subspace, for:

eigenvectors, along with the zero vector

$$\begin{aligned} \{\boldsymbol{v} \mid \boldsymbol{T}(\boldsymbol{v}) = \boldsymbol{\lambda}\boldsymbol{v}\} &= \{\boldsymbol{v} \mid T(\boldsymbol{v}) - \lambda\boldsymbol{v} = \mathbf{0}\} \\ &= \{\boldsymbol{v} \mid (T - \lambda I_V)\boldsymbol{v} = 0\} = \boldsymbol{Ker}(\boldsymbol{T} - \boldsymbol{\lambda}\boldsymbol{I}_V) \end{aligned}$$

a subspace of V (Theorem 4.8, page 126)

Bringing us to:

Compare with Definition 6.4, page 219.

DEFINITION 6.6 The **eigenspace** associated with an eigenvalue λ of a linear operator $T: V \to V$, denoted by $E[\lambda]$, is given by:

$$E[\lambda] = \operatorname{Ker}(T - \lambda I_V)$$

Compare with Example 6.5.

EXAMPLE 6.9 Find a basis for the eigenspace $E[4]$ of the linear operator:

$$T(\boldsymbol{a}, \boldsymbol{b}) = (\mathbf{3}\boldsymbol{a} + \mathbf{2}\boldsymbol{b}, \mathbf{3}\boldsymbol{a} - \mathbf{2}\boldsymbol{b})$$

of Example 6.8.

SOLUTION: The kernel of the linear operator:

$$\begin{aligned}(T-4I_{\Re^2})(\boldsymbol{a},\boldsymbol{b}) &= T(\boldsymbol{a},\boldsymbol{b})-4(\boldsymbol{a},\boldsymbol{b}) = (\boldsymbol{3a+2b-4a, 3a-2b-4b})\\ &= (\boldsymbol{-a+2b, 3a-6b})\end{aligned}$$

is, by definition, the set:

$$\{(a,b)|(\boldsymbol{-a+2b, 3a-6b}) = (\boldsymbol{0,0})\}$$

Equating coefficients, we have:

$$\left.\begin{aligned}-a+2b &= 0\\ 3a-6b &= 0\end{aligned}\right\} \longrightarrow \begin{array}{c}\begin{array}{cc} a & b\end{array}\\ \begin{bmatrix}-1 & 2\\ 3 & -6\end{bmatrix}\end{array} \xrightarrow{\textbf{rref}} \begin{array}{c}\begin{array}{cc} a & b\end{array}\\ \begin{bmatrix}\mathbf{1} & -2\\ 0 & 0\end{bmatrix}\end{array}$$

Setting the free variable b equal to r, we have $E[4] = \{(\boldsymbol{2r,r})|r \in \Re\}$ with basis $\{(\boldsymbol{2,1})\}$

CHECK YOUR UNDERSTANDING 6.9

Find a basis for the eigenspace $E[-3]$ of the linear operator $T(\boldsymbol{a},\boldsymbol{b}) = (\boldsymbol{3a+2b, 3a-2b})$ of Example 6.8.

Answer: See page B-25.

How does one go about finding the eigenvalues of a linear operator? Like this:

THEOREM 6.9 The eigenvalues of a linear operator $T: V \to V$ on a vector space of dimension n are the eigenvalues of the matrix $[T]_{\beta\beta} \in M_{n \times n}$, where β is **any** basis for V.

PROOF: We show that $\boldsymbol{v}$ is an eigenvector for the linear operator T corresponding to the eigenvalue λ, if and only if $[\boldsymbol{v}]_\beta$ is an eigenvector for the matrix $[T]_{\beta\beta}$ corresponding to the eigenvalue λ:

$$\begin{aligned}T(\boldsymbol{v}) &= \lambda\boldsymbol{v} \quad (\text{with } \boldsymbol{v} \neq \boldsymbol{0})\\ [T(\boldsymbol{v})]_\beta &= [\lambda\boldsymbol{v}]_\beta\\ \text{Theorem 5.22, page 180:}\quad [T]_{\beta\beta}[\boldsymbol{v}]_\beta &= \lambda[\boldsymbol{v}]_\beta \quad (\text{with } [\boldsymbol{v}]_\beta \neq \boldsymbol{0})\end{aligned}$$

Note that
$[v]_\beta = \boldsymbol{0} \Rightarrow \boldsymbol{v} = \boldsymbol{0}$
(Why?)

The above theorem leads us to the following definition:

DEFINITION 6.7
CHARACTERISTIC POLYNOMIAL (FOR LINEAR OPERATORS)

Let $T: V \to V$ be a linear operator on a vector space V of dimension n. The **characteristic polynomial** of T is the n-degree polynomial $\det([T]_{\beta\beta} - \lambda I_n)$ where β is any basis for V, and $\det([T]_{\beta\beta} - \lambda I_n) = 0$ is said to be the **characteristic equation** of T.

Theorem 5.28, page 193, and Exercise 30(b), page 216, tell us that
$$\det([T]_{\beta\beta} - \lambda I_n) = \det([T]_{\beta'\beta'} - \lambda I_n)$$
for **any** bases β and β'

Embedding the above terminology in the statement of Theorem 6.9, we come to:

Compare with Theorem 6.8

THEOREM 6.10 Let V be a vector space of dimension n. The eigenvalues of the linear operator $T: V \to V$ are the solutions of the characteristic equation $\det([T]_{\beta\beta} - \lambda I_n) = 0$, where β is any basis for V.

EXAMPLE 6.10 Find the eigenvalues and corresponding eigenspaces of the linear operator $T: P_2 \to P_2$ given by:

$$T(ax^2 + bx + c) = (a + c)x^2 + (2a + 2b + c)x + (a + c)$$

PROOF: With respect to the basis $\beta = \{x^2, x, 1\}$ in P_2, we have:

$$\begin{aligned} T(x^2) &= \mathbf{1}x^2 + \mathbf{2}x + \mathbf{1} \\ T(x) &= 0x^2 + 2x + 0 \\ T(1) &= x^2 + x + 1 \end{aligned} \qquad [T]_{\beta\beta} = \begin{bmatrix} \mathbf{1} & 0 & 1 \\ \mathbf{2} & 2 & 1 \\ \mathbf{1} & 0 & 1 \end{bmatrix}$$

Theorem 6.10 tells us that the eigenvalues are the solutions of the equation:

$$\det([T]_{\beta\beta} - \lambda I_3) = \det\begin{bmatrix} 1-\lambda & 0 & 1 \\ 2 & 2-\lambda & 1 \\ 1 & 0 & 1-\lambda \end{bmatrix} \overset{\text{see Example 6.6}}{=} -\lambda(2-\lambda)^2 = 0$$

We now determine the eigenspaces associated with the two eigenvalues, $\lambda = 0$ and $\lambda = 2$.

Finding $E[0]$:

$$E[0] = \ker(T - 0I_{P_2}) = \ker(T) = \{ax^2 + bx + c \mid \underbrace{T(ax^2 + bx + c) = 0}\}$$

$$(a + c)x^2 + (2a + 2b + c)x + (a + c) = 0x^2 + 0x + 0$$

Equating coefficients brings us to the following homogeneous system of equations:

$$\left.\begin{aligned} a + c &= 0 \\ 2a + 2b + c &= 0 \\ a + c &= 0 \end{aligned}\right\} \to \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix} \end{array} \xrightarrow{\text{rref}} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} \mathbf{1} & 0 & 1 \\ 0 & \mathbf{1} & -\frac{1}{2} \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

As shown in Example 6.6, $\{(-2r, r, 2r) | r \in \Re\}$ is the solution set of the above system of equations. Thus:

$$E[0] = \{-2rx^2 + rx + 2r | r \in \Re\} \text{ with basis } \{-2x^2 + x + 2\}$$

Finding $E[2]$:

$$E[2] = \ker(T - 2I_{P_2}) = \{ax^2 + bx + c | \underbrace{(T - 2I_{P_2})(ax^2 + bx + c)} = 0\}$$

$$(a + c)x^2 + (2a + 2b + c)x + (a + c) - 2(ax^2 + bx + c) = 0$$

$$= (-a + c)x^2 + (2a + c)x + (a - c) = 0x^2 + 0x + 0$$

Equating coefficients:

$$\left.\begin{aligned} -a + c &= 0 \\ 2a + c &= 0 \\ a - c &= 0 \end{aligned}\right\} \rightarrow \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} -1 & 0 & 1 \\ 2 & 0 & 1 \\ 1 & 0 & -1 \end{bmatrix} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{c} \begin{array}{ccc} a & b & c \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \end{array}$$

As shown in Example 6.6, $\{(0, r, 0) | r \in \Re\}$ is the solution set of the above system of equations. Thus:

$$E[2] = \{rx | r \in \Re\} \text{ with basis } \{x\}$$

CHECK YOUR UNDERSTANDING 6.10

Find the eigenspaces of the linear operator of Example 6.10 using $\beta = \{x^2 + x + 1, x + 1, 1\}$ instead of $\beta = \{x^2, x, 1\}$.

Answer: See page B-25.

EXERCISES

Exercises 1-14. Determine the eigenvalues and corresponding eigenspaces of the given matrix.

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

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

3. $\begin{bmatrix} 7 & -1 \\ 6 & 2 \end{bmatrix}$

4. $\begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

5. $\begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ -1 & -2 & 3 \end{bmatrix}$

6. $\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$

7. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 3 & 1 & -1 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$

8. $\begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 1 \\ 0 & 0 & 5 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}$

A factorization for the characteristic polynomial in the next six exercises can be obtained with the help of the following result:

A **zero of a polynomial** $p(x)$ is a number which when substituted for the variable x yields zero. For example, -1 is a zero of the polynomial $p(x) = x^3 - 3x - 2$, since $p(-1) = 0$. One can show that: c is a zero of a polynomial if and only if $(x - c)$ is a factor of the polynomial.

The adjacent example illustrates how the above result can be used to factor certain polynomials.

Since -1 is a zero $p(x) = x^3 - 7x - 6$, $x - (-1) = x + 1$ must be a factor, and we have:

$$\begin{array}{rl} & \,x^2 - x - 6 \\ x+1 & \overline{\big)\,x^3 \qquad - 7x - 6} \\ & \,x^3 + x^2 \\ & \qquad -x^2 - 7x - 6 \\ & \qquad -x^2 - x \\ & \qquad\qquad -6x - 6 \\ & \qquad\qquad -6x - 6 \end{array}$$

So: $x^3 - 7x - 6 = (x+1)(x^2 - x - 6)$
$= (x+1)(x+2)(x-3)$

9. $A = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 1 \end{bmatrix}$

10. $A = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ -1 & 0 & 2 \end{bmatrix}$

11. $A = \begin{bmatrix} 1 & 2 & -2 \\ -2 & 5 & -2 \\ -6 & 6 & -3 \end{bmatrix}$

12. $A = \begin{bmatrix} 3 & 2 & -2 \\ -3 & -1 & 3 \\ 1 & 2 & 0 \end{bmatrix}$

13. $A = \begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & -1 \end{bmatrix}$

14. $A = \begin{bmatrix} 4 & 2 & -2 & 2 \\ 1 & 3 & 1 & -1 \\ 0 & 0 & 2 & 0 \\ 1 & 1 & -3 & 5 \end{bmatrix}$

Exercises 15-31. Determine the eigenvalues and corresponding eigenspaces of the given linear operator.

15. $T: \Re \to \Re$ given by $T(x) = -5x$.
16. $T: \Re^2 \to \Re^2$ where $T(1, 0) = (2, 0)$ and $T(0, 1) = (1, 1)$.
17. $T: \Re^2 \to \Re^2$ given by $T(a, b) = (8a - 6b, 12a - 19b)$.
18. $T: \Re^3 \to \Re^3$ given by $T(a, b, c) = (0, a + c, 3b - c)$.
19. $T: \Re^3 \to \Re^3$ given by $T(a, b, c) = (a - 9b + 9c, a - 5b + 3c, 2a - 6b + 4c)$.
20. $T: \Re^4 \to \Re^4$, where $T(a, b, c, d) = (a, d, b, c)$.
21. $T: \Re^4 \to \Re^4$, where $T(a, b, c, d) = (2a, a - b, 3a + 2c - d, a - b + c - 2d)$.
22. $T: P_1 \to P_1$, where $T(ax + b) = (a + b)x - b$.
23. $T: P_1 \to P_1$, where $T(1) = x$ and $T(x) = 1$.
24. $T: P_2 \to P_2$ given by $P(ax^2 + bx + c) = -bx - 2c$.
25. $T: P_2 \to P_2$ given by $P(ax^2 + bx + c) = -cx^2 + bx - a$.
26. $T: P_2 \to P_2$, if $T(x^2) = 3x^2 - 2x + 4$, $T(x) = 7x - 8$, and $T(1) = 1$.
27. $T: P_3 \to P_3$, if $T(ax^3 + bx^2 + cx + d) = (a + d)x^3 + (-2a - c + d)x^2 + (2c - 2d)x - b + d$.
28. $T: M_{2 \times 2} \to M_{2 \times 2}$ given by $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} -c & 0 \\ a & -d \end{bmatrix}$.
29. $T: M_{2 \times 2} \to M_{2 \times 2}$, where $T\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$, $T\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} -9 & -2 \\ 0 & 0 \end{bmatrix}$, $T\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -2 & 1 \end{bmatrix}$, and $T\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -7 & 2 \end{bmatrix}$.
30. $I: V \to V$, where I is the identity map: $I(\boldsymbol{v}) = \boldsymbol{v}$.
31. $Z: V \to V$, where Z is the zero map: $Z(\boldsymbol{v}) = \boldsymbol{0}$.

32. **(Calculus Dependent)** Let V be the vector space of differentiable functions, and let $D: V \to V$ be the derivative operator. Show that $\boldsymbol{v} = e^{2x}$ is an eigenvector for D.

33. Prove that a square matrix A is invertible if and only if 0 is not an eigenvalue of A.

34. Let A be an invertible matrix with eigenvalue $\lambda \neq 0$ and corresponding eigenvector $\boldsymbol{v}$. Prove that $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with corresponding eigenvector $\boldsymbol{v}$.

35. Let λ_1 and λ_2 be distinct eigenvalues of $A \in M_{n \times n}$. Prove that $E[\lambda_1] \cap E[\lambda_2] = \{\boldsymbol{0}\}$

36. (a) Show that similar matrices have equal characteristic polynomials (see Definition 5.11, page 195).

    (b) Let $T: V \to V$ be a linear operator on a finite dimensional space V. Show that if β and β' are bases for V then: $[T]_{\beta\beta}$ and $[T]_{\beta'\beta'}$ have equal characteristic polynomials.

37.Let $A, P \in M_{n \times n}$, with P invertible. Prove that if $\boldsymbol{v}$ is an eigenvector of A, then $P^{-1}\boldsymbol{v}$ is an eigenvector of $P^{-1}AP$.

38. Let $A = \begin{bmatrix} a & b \\ 0 & c \end{bmatrix}$. Prove that a, and c are eigenvalues of A.

39. For $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, find necessary and sufficient conditions for A to have:

    (a) Two eigenvectors. (b) One eigenvector. (c) No eigenvector.

40. Let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$. Prove that a_{11}, a_{22}, and a_{33} are eigenvalues of A.

41. (a) Let $T: V \to V$ be a linear operator with eigenvalue λ. Prove that:

$$E[\lambda] = \{\boldsymbol{v} | \boldsymbol{v} \text{ is an eigenvector of } T\} \cup \{\boldsymbol{0}\}$$

    (b) Let $A \in M_{n \times n}$ with eigenvalue λ. Prove that:

$$E[\lambda] = \{\boldsymbol{v} | \boldsymbol{v} \text{ is an eigenvector of } A\} \cup \{\boldsymbol{0}\}$$

42. For $A \in M_{n \times n}$, show that $\text{null}(A - \lambda I) = \ker(T_A - \lambda I)$.

43. Prove that 0 is an eigenvalue for a linear operator $T: V \to V$ if and only if $\ker(T) \neq \{\boldsymbol{0}\}$.

44. Show that if $\boldsymbol{v}$ is an eigenvector for the linear operator $T: V \to V$. then so is $r\boldsymbol{v}$ for any $r \neq 0$.

45. Let $T: V \to W$ be an isomorphism. Show that $\boldsymbol{v}$ is an eigenvector in V if and only if $T(\boldsymbol{v})$ is an eigenvector in W.

46. Let $\boldsymbol{v}$ be an eigenvector for the linear operators $T: V \to V$ and $L: V \to V$. Show that $\boldsymbol{v}$ is also an eigenvector for the linear operator $L \circ T: V \to V$. Find a relation between the eigenvalues corresponding to $\boldsymbol{v}$ for T, L, and $L \circ T$.

47. Show that if λ_1 and λ_2 are distinct eigenvalues of a linear operator $T: V \to V$, then $E[\lambda_1] \cap E[\lambda_2] = \{\boldsymbol{0}\}$.

48. Let $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ be eigenvectors corresponding to distinct eigenvalues λ_1 and λ_2 of a linear operator $T: V \to V$. Show that $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ is a linearly independent set.

49. Let $T: V \to V$ be a linear operator on a vector space V of dimension n, and let $L: V \to \Re^n$ be an isomorphisms. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of the matrix $A = [L \circ T \circ L^{-1}]_S$, where S is the standard basis of $\Re^n$, and that $E[\lambda] = \{L^{-1}(\boldsymbol{v}) | \boldsymbol{v} \in \text{null}(A - \lambda I)\}$.

50. Let $T: V \to W$ be an isomorphism. Show that if $\boldsymbol{v}$ is an eigenvector of the linear operator $L: V \to V$, then $T(\boldsymbol{v})$ is an eigenvector of the linear operator $(T \circ L \circ T^{-1}): W \to W$.

51. Let β be a basis for a space V of dimension n, and $L: V \to V$ a linear operator. Prove that if $\boldsymbol{v} \in V$ is an eigenvector of T with eigenvalue λ, then $[\boldsymbol{v}]_\beta$ is an eigenvector of $T_{[L]_{\beta\beta}}: R^n \to R^n$ with eigenvalue λ.

52. Show that if λ is an eigenvalue of $A \in M_{n \times n}$ then λ is also an eigenvalue of A^{T}. (See Exercise 19, page 162)

53. Show that if $A \in M_{n \times n}$ is nilpotent, then 0 is the only eigenvalue of A. (See Exercise 23, page 163.)

54. Show that the characteristic polynomial of $A \in M_{2 \times 2}$ can be expressed in the form $\lambda^2 - \text{Trace}(A)\lambda + \det(A)$, where Trace($A$) denotes the trace of A (see Exercise 24, page 163).

55. Let $A \in M_{2 \times 2}$. Prove that the characteristic polynomial of A is of the form $p(\lambda) = \lambda^2 + b\lambda + \det(A)$, and that $A^2 + bA + \det(A)I = 0$. (This is the Cayley-Hamilton Theorem for square matrices of dimension 2.)

56. **(PMI)** Let $A \in M_{n \times n}$. Use the Principle of Mathematical Induction to show that the coefficient of the leading term of the characteristic polynomial of A is ± 1.

57. **(PMI)** Let $A \in M_{n \times n}$. Show that the constant term of the characteristic polynomial of A is $\det A$.

58. **(PMI)** Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of A for $A \in M_{m \times m}$. Prove that $\lambda_1^n, \lambda_2^n, \ldots, \lambda_k^n$ are the distinct eigenvalues of A^n.

59. **(PMI)** Let A be a square matrix with eigenvalue λ and corresponding eigenvector $\boldsymbol{v}$. Show that for any positive integer n, λ^n is an eigenvalue of A^n with corresponding eigenvalue $\boldsymbol{v}$.

60. **(PMI)** Let A be a square matrix with eigenvalue λ and corresponding eigenvector $\boldsymbol{v}$. Show that for any integer n, λ^n is an eigenvalue of A^n with corresponding eigenvalue $\boldsymbol{v}$.

61. **(PMI)** Let λ be an eigenvalue for a linear operator $T: V \to V$. Use the Principle of Mathematical Induction to show that λ^n is an eigenvalue for $T^n: V \to V$, where T^n is defined inductively as follows: $T^1 = T$, and $T^{k+1} = T \circ T^k$.

PROVE OR GIVE A COUNTEREXAMPLE

62. If λ is an eigenvalue for $T: V \to V$ then it is also an eigenvalue for $(kT): V \to V$, where $(kT)\boldsymbol{v} = kT(\boldsymbol{v})$.

63. If λ is an eigenvalue for the two operators $T: V \to V$ and $L: V \to V$, then it is also an eigenvalue for the operator $(T+L): V \to V$, where $(T+L)(\boldsymbol{v}) = T(\boldsymbol{v}) + L(\boldsymbol{v})$.

64. For $A, B \in M_{2 \times 2}$, if λ_A and λ_B are eigenvalues of A and B, respectively, then $\lambda_A + \lambda_B$ is an eigenvalue of T_{A+B}.

65. For $A, B \in M_{2 \times 2}$, if λ_A and λ_B are eigenvalues for A and B, respectively, then $\lambda_A \lambda_B$ is an eigenvalue for AB.

66. If λ is an eigenvalue of the linear operator $T: V \to V$, then λ^2 is an eigenvalue of $(T \circ T): V \to V$.

67. If λ_T and λ_L are eigenvalues for the linear operators $T: V \to V$ and $L: V \to V$, respectively, then $\lambda_A \lambda_B$ is an eigenvalue for $(T \circ L): V \to V$.

68. If λ_T and λ_L are eigenvalues for the linear operators $T: V \to V$ and $L: V \to V$, respectively, then $\lambda_A + \lambda_B$ is an eigenvalue for $(T+L): V \to V$.

69. If $\boldsymbol{v}$ is an eigenvector for $T: V \to V$ and $L: V \to V$, then $\boldsymbol{v}$ is also an eigenvector for $(T+L): V \to V$.

70. If $T: V \to V$ is a linear operator with eigenvector $\boldsymbol{v}$, then $\boldsymbol{v} + r\boldsymbol{v}$ is also an eigenvector of T for every $r \in \Re$.

71. For $A \in M_{2 \times 2}$, $A^2 = \mathbf{0}$ if and only if 0 is the only eigenvalue of T_A.

72. Let T be a linear operator on a vector space V of dimension n. Let λ be an eigenvalue for T and let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ be a basis for $E[\lambda]$. Then, for any $r \in \Re$, $\lambda + r$ is an eigenvalue for $T + rI_n$, and $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is a basis for $E[\lambda + r]$.

§3. DIAGONALIZATION

We begin with:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & a_{nn} \end{bmatrix}$$

DEFINITION 6.8
DIAGONAL MATRIX

A square matrix $A = [a_{ij}]_{n \times n}$ for which $a_{ij} = 0$ if $i \neq j$ is said to be a **diagonal matrix** (see margin).

DIAGONALIZABLE OPERATOR

A linear operator $T: V \to V$ on a finite dimensional vector space V is said to be **diagonalizable** if there exists a basis β for which $[T]_{\beta\beta}$ is a diagonal matrix.

There is an intimate connection between diagonalizable matrices and eigenvectors. Focusing first on linear operators, we have:

THEOREM 6.11

Let $T: V \to V$ be a linear operator on a finite dimensional vector space. Then:

T is diagonalizable if and only if there exists a basis for V consisting of eigenvectors of T.

PROOF: Assume that T is diagonalizable. Let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ be such that $[T]_{\beta\beta}$ is a diagonal matrix: $[T]_{\beta\beta} = [a_{ij}]$ with $a_{ij} = 0$ for $i \neq j$. Since the i^{th} column of $[T]_{\beta\beta}$ consists of the coefficients of the vector $T(\boldsymbol{v}_i)$ with respect to the basis β, we have:

$$\boldsymbol{T(v_i)} = 0\boldsymbol{v}_1 + \dots + 0\boldsymbol{v}_{i-1} + \boldsymbol{a_{ii}v_i} + 0\boldsymbol{v}_{i+1} + \dots + 0\boldsymbol{v}_n = \boldsymbol{a_{ii}v_i}$$

From the above we see that $\boldsymbol{v_i}$ is an eigenvector for T corresponding to the eigenvalue a_{ii}.

Conversely, let $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ be a basis for V consisting of eigenvectors, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues corresponding to $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n$. From:

$$\boldsymbol{T(v_i)} = \boldsymbol{\lambda_i v_i} = \mathbf{0}\boldsymbol{v}_1 + \mathbf{0}\boldsymbol{v}_2 + \cdots + \boldsymbol{\lambda_i v}_i + \dots + \mathbf{0}\boldsymbol{v}_n$$

we have (see Definition 5.10, page 179):

$$[T]_{\beta\beta} = \begin{bmatrix} \boldsymbol{\lambda_1} & \mathbf{0} & \dots & \dots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\lambda_2} & \dots & \dots & \mathbf{0} \\ \dots & \dots & \dots & \dots & \dots \\ \mathbf{0} & \mathbf{0} & \dots & \boldsymbol{\lambda_{n-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \dots & \dots & \boldsymbol{\lambda_n} \end{bmatrix}$$

The λ_i ‘s can be zero and need not be distinct (several of the eigenvectors in β may share a common eigenvalue).

CHECK YOUR UNDERSTANDING 6.11

Let $T: \Re^3 \to \Re^3$ be the linear map given by:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{3a - b - c}, \boldsymbol{2a - 2c}, \boldsymbol{2a - b - c})$$

Show that $\beta = \{(\mathbf{1}, \mathbf{2}, \mathbf{0}), (\mathbf{1}, \mathbf{2}, \mathbf{2}), (\mathbf{2}, \mathbf{1}, \mathbf{1})\}$ is a basis for $\Re^3$ consisting of eigenvectors of T. Determine $[T]_{\beta\beta}$ and show that its diagonal elements are eigenvalues of T.

Answer: See page B-25.

In our quest for bases consisting of eigenvectors, we note that:

THEOREM 6.12 If $\lambda_1, \lambda_2, \dots, \lambda_m$ are distinct eigenvalues of a linear operator $T: V \to V$, and if $\boldsymbol{v}_i$ is any eigenvector corresponding to λ_i, for $1 \le i \le m$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$ is a linearly independent set.

PROOF: By induction on m:

If $m = 1$, then $\{\boldsymbol{v}_1\}$ consists of a single nonzero vector and is therefore linearly independent (Exercise 33, page 92).

Assume the assertion holds for $m = k$ (the induction hypothesis).

Let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{k+1}\}$ be a set of eigenvector corresponding to distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_{k+1}$. We are to show that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_{k+1}\}$ is a linearly independent set. With this in mind, we consider the linear combination:

$$a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \dots + a_k\boldsymbol{v}_k + a_{k+1}\boldsymbol{v}_{k+1} = \mathbf{0} \qquad \textbf{(*)}$$

Applying T to both sides, we have:

Since $\boldsymbol{v}_i$ is an eigenvector corresponding to λ_i:

$$T(\boldsymbol{v}_i) = \lambda_i\boldsymbol{v}_i$$

$$T(a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \dots + a_k\boldsymbol{v}_k + a_{k+1}\boldsymbol{v}_{k+1}) = T(\mathbf{0})$$

$$a_1T(\boldsymbol{v}_1) + a_2T(\boldsymbol{v}_2) + \dots + a_kT(\boldsymbol{v}_k) + a_{k+1}T(\boldsymbol{v}_{k+1}) = \mathbf{0}$$

$$a_1\lambda_1\boldsymbol{v}_1 + a_2\lambda_2\boldsymbol{v}_2 + \dots + a_k\lambda_k\boldsymbol{v}_k + a_{k+1}\lambda_{k+1}\boldsymbol{v}_{k+1} = \mathbf{0} \qquad \textbf{(**)}$$

Multiply both sides of (*) by λ_{k+1}:

$$a_1\lambda_{k+1}\boldsymbol{v}_1 + a_2\lambda_{k+1}\boldsymbol{v}_2 + \dots + a_k\lambda_{k+1}\boldsymbol{v}_k + a_{k+1}\lambda_{k+1}\boldsymbol{v}_{k+1} = \mathbf{0} \qquad \textbf{(***)}$$

Subtract **(***)** from **(**)**:

$$a_1(\lambda_1 - \lambda_{k+1})\boldsymbol{v}_1 + a_2(\lambda_2 - \lambda_{k+1})\boldsymbol{v}_2 + \dots + a_k(\lambda_k - \lambda_{k+1})\boldsymbol{v}_k = \mathbf{0}$$

By the induction hypothesis, the k eigenvectors $\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_k$ corresponding to the distinct eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_k$ are linearly independent. Consequently:

$$a_1(\lambda_1 - \lambda_{k+1}) = a_2(\lambda_2 - \lambda_{k+1}) = \dots = a_k(\lambda_k - \lambda_{k+1}) = 0$$

Since the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{k+1}$ are distinct, none of the above $(\lambda_i - \lambda_{k+1})$ is equal to 0. Hence:

$$a_1 = a_2 = \ldots = a_k = 0$$

Returning to **(*)**, we then have:

$$a_{k+1}\boldsymbol{v}_{k+1} = \boldsymbol{0}$$

Being an eigenvector, $\boldsymbol{v}_{k+1} \neq \boldsymbol{0}$, and therefore $a_{k+1} = 0$ (Theorem 2.8, page 54).

We have just observed that if you take eigenvectors corresponding to different eigenvalues you will end up with a linearly independent set of vectors. More can be said:

THEOREM 6.13 Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be distinct eigenvalues of a linear operator $T: V \to V$, and let $S_i = \{\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, \ldots, \boldsymbol{v}_{ir_i}\}$ be any linearly independent subset of $E[\lambda_i]$. Then:

$$S = S_1 \cup S_2 \cup \ldots \cup S_m$$

is a linearly independent set.

linear combination = **0**

each coefficient is 0

PROOF: Consider the vector equation:

The r_1 vectors in S_1 The r_m vectors in S_m

$$(a_{11}\boldsymbol{v}_{11} + \cdots + a_{1r_1}\boldsymbol{v}_{1r_1}) + \ldots + (a_{m1}\boldsymbol{v}_{m1} + \cdots + a_{mr_m}\boldsymbol{v}_{mr_m}) = \boldsymbol{0} \quad \textbf{(*)}$$

(we will show that every coefficient must be zero)

For $1 \le i \le m$, let $\boldsymbol{v}_i = a_{i1}\boldsymbol{v}_{i1} + \ldots + a_{ir_i}\boldsymbol{v}_{ir_i}$. Assume, without loss of generality, that $\boldsymbol{v}_i \neq \boldsymbol{0}$ for $1 \le i \le t$ and that the rest are zero vectors.

As for any of the zero vectors, $\boldsymbol{v}_i = a_{i1}\boldsymbol{v}_{i1} + \ldots + a_{ir_i}\boldsymbol{v}_{ir_i} = \boldsymbol{0}$, its coefficients must be zero, as $\{\boldsymbol{v}_{i1}, \boldsymbol{v}_{i2}, \ldots, \boldsymbol{v}_{ir_i}\}$ is given to be a linearly independent set.

As for the nonzero vectors, we begin by rewriting (*) in the form:

$$\boldsymbol{v}_1 + \boldsymbol{v}_2 + \cdots + \boldsymbol{v}_t = \boldsymbol{0} \quad (**)$$

Since the nonzero vectors $\{\boldsymbol{v}_1, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_t\}$ are eigenvectors associated with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_t$, they are linearly independent (Theorem 6.12). It follows, from (**), that each $\boldsymbol{v}_i$ must be **0**:

$$\boldsymbol{v}_i = a_{i1}\boldsymbol{u}_{i1} + \ldots + a_{ir_i}\boldsymbol{u}_{ir_i} = \boldsymbol{0}$$

Using, again, the fact that each $S_i = \{\boldsymbol{u}_{i1}, \boldsymbol{u}_{i1}, \ldots, \boldsymbol{u}_{ir_i}\}$ is a linearly independent set, we conclude that each scalar a_{ij} must be zero

CHECK YOUR UNDERSTANDING 6.12

Let $T: V \to V$ be a linear operator on a space of dimension n. Prove that if T has n distinct eigenvalues, then T is diagonalizable.

Answer: See page B-26.

RETURNING TO MATRICES

To say that $T: V \to V$ is **diagonalizable** is to say that V contains a basis consisting of eigenvectors of T (Theorem 6.11). Let's modify this characterization to accommodate matrices:

DEFINITION 6.9
DIAGONALIZABLE MATRIX

A matrix $A \in M_{n \times n}$ is **diagonalizable** if there exists a basis for $\overline{\Re^n}$ consisting of eigenvectors of A.

Here is a link between diagonalizable matrices and diagonalizable linear operators:

THEOREM 6.14 $A \in M_{n \times n}$ is diagonalizable if and only if the linear map $T_A: \overline{\Re^n} \to \overline{\Re^n}$ given by $T_A X = AX$ is diagonalizable.

(X is a vertical n-tuple: a column matrix.)

PROOF: The linear map $T_A: \overline{\Re^n} \to \overline{\Re^n}$ is diagonalizable

if and only if:

there exists a basis $\{X_1, X_2, \ldots, X_n\}$ of $\overline{\Re^n}$, and scalars $\lambda_1, \lambda_2, \ldots, \lambda_n$, such that $T_A(X_i) = \lambda_i X_i$

if and only if:

$AX_i = \lambda_i X_i$ (Definition of T_A)

if and only if

A is diagonalizable (Definition 6.9).

Definition 6.9 is okay, but how do we go from a diagonalizable matrix to a specific diagonal matrix? Like this::

THEOREM 6.15 Let $A \in M_{n \times n}$ be diagonalizable. Let $\beta = \{X_1, X_2, \ldots, X_n\}$ be any basis for $\overline{\Re^n}$ consisting of eigenvectors of A, with associated eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. If $P \in M_{n \times n}$ is the matrix whose i^{th} column is X_i, then:

$D = P^{-1}AP$ where $D = [d_{ij}]$ is the diagonal matrix, with diagonal entry $d_{ii} = \lambda_i$.

This theorem asserts that any diagonalizable matrix is similar to a diagonal matrix. The converse also holds (Exercise 37). And so we have:

$A \in M_{n \times n}$ is diagonalizable if and only if it is similar to a diagonal matrix.

PROOF: Let $S_n = \{\overline{e_1}, \overline{e_2}, \ldots, \overline{e_n}\}$ denote the standard basis for $\overline{\Re^n}$ (see page 94). Employing Theorem 5.26 (page 193), and Theorem 5.23 (page 184) to the linear map $T_A: \overline{\Re^n} \to \overline{\Re^n}$ given by $T_A(X) = AX$, we have:

$$[T_A]_{\beta\beta} = P^{-1}[T_A]_{S_nS_n}P \quad (*)$$

$$\begin{tikzcd}
\Re^n\ \beta \arrow[r, "T_A"] \arrow[d, dashed, "I"'] & \Re^n\ \beta \\
\Re^n\ S_n \arrow[r, dashed, "T_A"] & \Re^n\ S_n \arrow[u, dashed, "I"']
\end{tikzcd}$$

(See page 193)

where $P = [I]_{S_n\beta}$ (see margin). Since $I(X_i) = X_i$, the i^{th} column of P is simply the vector X_i (recall that S_n is the standard basis). Now:

The i^{th} column of $[T_A]_{S_nS_n}$ is:

$$[T_A(\overline{e_i})]_{S_n} = [A\overline{e_i}]_{S_n} = \text{the } i^{\text{th}} \text{ column of } A$$

Hence: $\boldsymbol{[T_A]_{S_nS_n} = A}$.

Since β is a basis of eigenvectors:

$$T_A(X_i) = \lambda_i X_i = 0X_1 + 0X_2 + \ldots + \lambda_i X_i + \ldots + 0X_n$$

It follows that $[T_A]_{\beta\beta}$ is the diagonal matrix $D = [d_{ij}]$ with $d_{ii} = \lambda_i$.

Putting all of this together we have (see *):

$$D = P^{-1}AP, \text{ or: } A = PDP^{-1}$$

EXAMPLE 6.11 Show that the matrix:

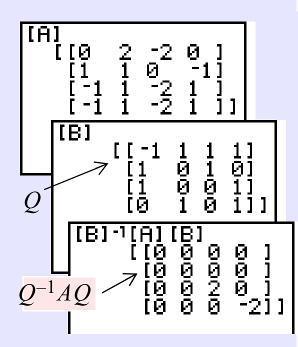
$$A = \begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}$$

is diagonalizable, and find a matrix P such that $P^{-1}AP$ is a diagonal matrix.

SOLUTION: In Example 6.7, page 221, we found that 0, 2, and -2 are the eigenvalues of A. We also observed that $\{\overline{\mathbf{(-1, 1, 1, 0)}}, \overline{\mathbf{(1, 0, 0, 1)}}\}$ is a basis for $E[0]$, $\{\overline{\mathbf{(1, 1, 0, 0)}}\}$ is a basis for $E[2]$, and that $\{\overline{\mathbf{(1, 0, 1, 1)}}\}$ is basis for $E[-2]$. It is easy to see that the four eigenvectors $\{\overline{\mathbf{(-1, 0, 1, 0)}}, \overline{\mathbf{(1, 1, 0, 1)}}, \overline{\mathbf{(1, 1, 0, 0)}}, \overline{\mathbf{(1, 0, 1, 1)}}\}$ are linearly independent, and therefore constitute a basis for $\Re^4$. Taking P to be the matrix with columns the above four eigenvectors:

$$P = \begin{bmatrix} -1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

we have:



```
[A]
  [[0  2 -2 0 ]
   [1  1 0  -1]
   [-1 1 -2 1 ]
   [-1 1 -2 1 ]]
[B]
     [[-1 1 1 1]
      [1  0 1 0]
      [1  0 0 1]
      [0  1 0 1]]
[B]-1[A][B]
     [[0 0 0 0 ]
      [0 0 0 0 ]
      [0 0 2 0 ]
      [0 0 0 -2]]
```

$$P^{-1}\begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}$$

CHECK YOUR UNDERSTANDING 6.13

Determine if the given matrix is diagonalizable. If it is, use Theorem 6.15 to find a matrix P such that $P^{-1}AP$ is a diagonal matrix.

(a) $A = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 3 & 0 \\ -4 & 13 & -1 \end{bmatrix}$ (b) $A = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ -1 & -2 & 3 \end{bmatrix}$

Answer: See page B-26.

The next result plays an important role in many eigenvector applications:

THEOREM 6.16 If $A \in M_{m \times m}$ is diagonalizable with $A = PDP^{-1}$, then $A^n = PD^nP^{-1}$.

PROOF: (By induction on n) For $n = 1$ we have:

$$P^{-1}AP = D^1 \Rightarrow A = PDP^{-1}$$

Assume that $A^k = PD^kP^{-1}$ (the induction hypothesis). Then:

$$A^{k+1} = AA^k \underset{\substack{\uparrow \\ \text{induction hypothesis}}}{=} A(PD^kP^{-1}) \overset{\substack{P^{-1}AP = D \Rightarrow A = PDP^{-1} \\ \downarrow}}{=} (PDP^{-1})(PD^kP^{-1})$$

$$= P(DD^k)P^{-1} = PD^{k+1}P^{-1}$$

CHECK YOUR UNDERSTANDING 6.14

Calculate A^{10} for the diagonalizable matrix A of Example 6.11.

Answer: See page B-27.

ALGEBRAIC AND GEOMETRIC MULTIPLICITY OF EIGENVALUES

An eigenvalue λ_0 of a matrix $A \in M_{n \times n}$ (or of a linear operator T on a vector space of dimension n) has **algebraic multiplicity** k if $(\lambda_0 - \lambda)^k$ is a factor of A's (or T's) characteristic polynomial, and $(\lambda_0 - \lambda)^{k+1}$ is not. We also define the **geometric multiplicity** of λ_0 to be the dimension of $E[\lambda_0]$ (the eigenspace corresponding to λ_0).

EXAMPLE 6.12 Find the algebraic and geometric multiplicity of the eigenvalues of the matrix:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 2 & 2 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

SOLUTION: In Example 6.6, page 220, we showed that $-\lambda(2-\lambda)^2$ is the characteristic polynomial of A. It follows that the eigenvalue 0 has algebraic multiplicity 1, and that the eigenvalue 2 has algebraic multiplicity 2. Since both of the eigenspaces $E[0]$ and $E[2]$ were seen to have dimension 1, the geometric multiplicity of both eigenvalues is 1.

The above example illustrates the fact that the geometric multiplicity of an eigenvalue can be less than its algebraic multiplicity; it cannot go the other way around:

THEOREM 6.17 If λ_0 is an eigenvalue of $A \in M_{n \times n}$ with algebraic multiplicity m_a and geometrical multiplicity m_g, then $m_g \le m_a$.

PROOF: (By contradiction) Assume that $m_a < m_g$ and let $\{\overline{\mathbf{v}_1}, \overline{\mathbf{v}_2}, \dots, \overline{\mathbf{v}_{m_g}}\}$ be a basis for $(E[\lambda_0])$. Expand $\{\overline{\mathbf{v}_1}, \overline{\mathbf{v}_2}, \dots, \overline{\mathbf{v}_{m_g}}\}$ to a basis $\beta = \{\overline{\mathbf{v}_1}, \overline{\mathbf{v}_2}, \dots, \overline{\mathbf{v}_{m_g}}, \overline{\mathbf{v}_{m_g+1}}, \dots, \overline{\mathbf{v}_n}\}$ for $\Re^n$. Since, for $1 \le i \le m_g$:

$$\begin{aligned} T_A(\overline{\mathbf{v}_i}) &= A[\overline{v_i}] = \lambda_0 \overline{\mathbf{v}_i} \\ &= 0\overline{\mathbf{v}_1} + \dots + 0\overline{\mathbf{v}_{i-1}} + \lambda_0 \overline{\mathbf{v}_i} + 0\overline{\mathbf{v}_{i+1}} + \dots + 0\overline{\mathbf{v}_n} \end{aligned}$$

the matrix $[T_A]_{\beta\beta}$ is of the following form:

$$[T_A]_{\beta\beta} = \left[\begin{array}{cc} \overset{m_g}{\boxed{\begin{matrix} \lambda_0 & 0 & \dots & 0 \\ 0 & \lambda_0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \lambda_0 \end{matrix}}}\left.\vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}}\right\}m_g & \boxed{\quad X \quad} \\ \boxed{\begin{matrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 \end{matrix}} & \boxed{\quad Y \quad} \end{array}\right]_{n \times n}$$

Recall that $T_A: \mathbf{R}^n \to \mathbf{R}^n$ is the linear operator given by:
$$T_A(\mathbf{v}) = A\mathbf{v}$$

Recall that the i^{th} column of $[T_A]_\beta$ consists of the coefficients of the vector $T_A(v_i)$ with respect to the basis $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$.

In the proof of Theorem 6.15 we observed that $[T_A]_{S_nS_n} = A$ (where S_n is the standard basis in $\Re^n$). It follows, from Exercise 36(a), page

230, that the characteristic polynomial of A equals that of $[T_A]_{\beta\beta}$, namely, $\det([T_A]_{\beta\beta}-\lambda I)$:

$$\det\begin{bmatrix} cI & X \\ 0 & Y \end{bmatrix} = c^r\det(Y)$$

$$\det\left[\begin{array}{cccc|c} \multicolumn{4}{c|}{m_g} & \\ \lambda_0-\lambda & 0 & \cdots & 0 & \\ 0 & \lambda_0-\lambda & \cdots & 0 & X \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & \lambda_0-\lambda & \\ \hline 0 & 0 & \cdots & 0 & \\ 0 & 0 & \cdots & 0 & Y \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & 0 & \end{array}\right]_{n\times n} = (\lambda_0-\lambda)^{m_g}\det(Y)$$

Exercise 40, page 217 (see margin)

This leads to a contradiction, for the factor $(\lambda_0-\lambda)$ cannot appear with exponent greater than m_a in the characteristic polynomial of A (remember that m_a is the algebraic multiplicity of λ_0).

In certain cases, the algebraic and geometric multiplicities of a linear operator can be used to determine if the operator is diagonalizable:

THEOREM 6.18 Assume that the characteristic polynomial of a linear operator $T: V\to V$ (or of a matrix), can be factored into a product of linear factors (with real coefficients). Then T is diagonalizable if and only if the algebraic multiplicity of each eigenvalue of T *is* equal to its geometric multiplicity.

PROOF: Assume that T is diagonalizable. By Theorem 6.11, there exists a basis $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ for V consisting of eigenvectors of T. Let $\{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ be the set of T's (several of the $\boldsymbol{v_i}$'s may correspond to the same eigenvalue). If necessary, reorder β so that $\beta = S_1\cup S_2\cup\ldots\cup S_k$, where S_j consists of the eigenvectors in β corresponding to λ_j.

Since the vectors in S_j are linearly independent, and since their combined sum equals the dimension of V, it follows, from Theorem 6.13, that the number of vectors in S_j must equal $g_j = \dim E[\lambda_j]$, the geometric dimension of λ_j. Hence: $n = g_1+g_2+\cdots+g_k$.

We are given that the characteristic polynomial of T can be factored into a product of linear factors:

$$\det([T]_\beta-\lambda I) = (\lambda-\lambda_1)^{a_1}(\lambda-\lambda_2)^{a_2}\ldots(\lambda-\lambda_k)^{a_k} \qquad (*)$$

It follows that:

$$n = g_1 + g_2 + \cdots + g_k \underset{\substack{\uparrow \\ \text{Theorem 6.17}}}{\le} a_1 + a_2 + \cdots + a_k \underset{\substack{\uparrow \\ \text{degree of the characteristic polynomial (*)}}}{=} n$$

Consequently:

$$a_1 + a_2 + \cdots + a_k = g_1 + g_2 + \cdots + g_k$$

$$\text{Or: } (a_1 - g_1) + (a_2 - g_2) + \ldots + (a_k - g_k) = 0$$

Knowing that $a_i \ge g_i$ (Theorem 6.17), we conclude that $a_i = g_i$ for each $1 \le i \le k$.

Conversely, assume that the multiplicity of each eigenvalue λ_j, $1 \le j \le k$, is equal to its degree: $a_j = \dim(E[\lambda_j]) = g_j$. Let S_j be a basis for $E[\lambda_j]$. By Theorem 6.13, the set $S_1 \cup S_2 \cup \ldots \cup S_k$, which contains $g_1 + g_2 + \cdots + g_k$ vectors, is linearly independent. It is in fact a basis, since it contains n vectors (Theorem 3.11, page 99):

$$g_1 + g_2 + \cdots + g_k = a_1 + a_2 + \cdots + a_k \overset{\substack{\text{degree of the characteristic polynomial} \\ \downarrow}}{=} n$$

Possessing a basis of eigenvectors, T is diagonalizable (Theorem 6.11).

EXAMPLE 6.13 Appeal to the previous theorem to show that the linear operator $T\colon \Re^4 \to \Re^4$ given by:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = (\boldsymbol{2b - 2c}, \boldsymbol{a + b - d}, \boldsymbol{-a + b - 2c + d}, \boldsymbol{-a + b - 2c + d})$$

is diagonalizable. Find a basis β for which $[T]_{\beta\beta}$ is a diagonal matrix.

SOLUTION: For S the standard basis of $\Re^4$, we have:

$T(\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}) = (\mathbf{0}, \mathbf{1}, \mathbf{-1}, \mathbf{-1})$

$$[T]_{SS} = \begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}$$

The above matrix was encountered in Example 6.7, page 221, where we found its characteristic polynomial to be $\lambda^2(\lambda - 2)(\lambda + 2)$. We also showed that:

The eigenspace corresponding to the eigenvalue 0, of multiplicity 2, has dimension 2 (with basis $\{(\mathbf{-1}, \mathbf{1}, \mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{1})\}$).

The eigenspace corresponding to the eigenvalue 2, of multiplicity 1, has dimension 1 (with basis $\{(\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0})\}$.)

The eigenspace corresponding to the eigenvalue $-\mathbf{2}$, of multiplicity 1, has dimension 1 (with basis $\{(\mathbf{1, 0, 1, 1})\}$).

Since the algebraic multiplicity of each eigenvalue equals its geometric multiplicity, the linear operator is diagonalizable. Moreover, since

$$\beta = \{(\mathbf{-1, 1, 1, 0}), (\mathbf{1, 0, 0, 1}), (\mathbf{1, 1, 0, 0}), (\mathbf{1, 0, 1, 1})\}$$

is a basis for V of eigenvectors, we know that $[T]_{\beta\beta}$ will be a diagonal matrix with diagonal entries equal to the eigenvalues corresponding to the eigenvalues of β; namely:

$$[T]_{\beta\beta} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{2} & 0 \\ 0 & 0 & 0 & -\mathbf{2} \end{bmatrix}$$

CHECK YOUR UNDERSTANDING 6.15

Verify that the algebraic multiplicity of each eigenvalue of the diagonalizable matrix of CYU 6.13(b) equals that of its geometric multiplicity.

Answer: See page B-27.

EXERCISES

Exercises 1-19. Determine if the given linear operator $T: V \to V$ is diagonalizable. If it is, find a basis β for V such that $[T]_{\beta\beta}$ is a diagonal matrix.

(Note: To factor the characteristic polynomial of the given operator, you may need to use the division process discussed above Exercise 9 on page 228.)

1. $T: \Re^2 \to \Re^2$ given by $T(\mathbf{a}, \mathbf{b}) = (2\mathbf{a}, 3b)$.
2. $T: \Re^2 \to \Re^2$ given by $T(\mathbf{a}, \mathbf{b}) = (7\mathbf{a} - \mathbf{b}, 6\mathbf{a} + 2\mathbf{b})$.
3. $T: \Re^2 \to \Re^2$ given by $T(\mathbf{a}, \mathbf{b}) = (2a, -a + 3b)$.
4. $T: \Re^2 \to \Re^2$ where $T(\mathbf{1}, \mathbf{0}) = (\mathbf{4}, \mathbf{-1})$ and $T(\mathbf{0}, \mathbf{1}) = (\mathbf{1}, \mathbf{2})$.
5. $T: \Re^2 \to \Re^2$ where $T(\mathbf{1}, \mathbf{1}) = (\mathbf{1}, \mathbf{2})$ and $T(\mathbf{0}, \mathbf{1}) = (\mathbf{2}, \mathbf{0})$.
6. $T: \Re^3 \to \Re^3$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (-\mathbf{a}, \mathbf{2c}, \mathbf{2b} + \mathbf{3c})$.
7. $T: \Re^3 \to \Re^3$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{13a} - \mathbf{4b}, \mathbf{8b} - \mathbf{2c}, \mathbf{5c})$.
8. $T: \Re^3 \to \Re^3$ where $T(\mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{0}, \mathbf{-1}, \mathbf{-1})$, $T(\mathbf{0}, \mathbf{1}, \mathbf{0}) = (\mathbf{1}, \mathbf{4}, \mathbf{5})$, and $T(\mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{1}, \mathbf{-1}, \mathbf{-2})$.
9. $T: \Re^3 \to \Re^3$ where $T(\mathbf{1}, \mathbf{1}, \mathbf{1}) = (\mathbf{1}, \mathbf{0}, \mathbf{1})$, $T(\mathbf{0}, \mathbf{1}, \mathbf{1}) = (\mathbf{1}, \mathbf{1}, \mathbf{0})$, and $T(\mathbf{0}, \mathbf{0}, \mathbf{2}) = (\mathbf{1}, \mathbf{1}, \mathbf{0})$.
10. $T: \Re^3 \to \Re^3$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{2a} + \mathbf{4b} - \mathbf{4c}, -\mathbf{3b} + \mathbf{5c}, -\mathbf{6b} + \mathbf{8c})$.
11. $T: \Re^4 \to \Re^4$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) = (\mathbf{b}, \mathbf{a}, \mathbf{d}, \mathbf{c})$.
12. $T: \Re^4 \to \Re^4$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}) = (-\mathbf{a}, \mathbf{0}, \mathbf{a} + \mathbf{2c}, \mathbf{-2b})$.
13. $T: \Re^4 \to \Re^4$ where $T(\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0}) = (\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{0})$, $T(\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{3}, \mathbf{-2}, \mathbf{0}, \mathbf{0})$, $T(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0}) = (\mathbf{2}, \mathbf{-1}, \mathbf{4}, \mathbf{0})$, and $T(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{6}, \mathbf{0}, \mathbf{5}, \mathbf{-3})$.
14. $T: \Re^4 \to \Re^4$ where $T(\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{1})$, $T(\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{1}) = (\mathbf{0}, \mathbf{0}, \mathbf{1}, \mathbf{0})$, $T(\mathbf{1}, \mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{1})$, and $T(\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{1}, \mathbf{1}, \mathbf{1}, \mathbf{1})$.
15. $T: \Re^5 \to \Re^5$ given by $T(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{e}) = (\mathbf{2a}, \mathbf{a} - \mathbf{b}, \mathbf{4c} + \mathbf{5d}, \mathbf{3d}, \mathbf{c} + \mathbf{8e})$.
16. $T: P_2 \to P_2$ given by $T[p(x)] = p(x+1)$.
17. $T: P_2 \to P_2$ given by $T(ax^2 + bx + c) = (3a - 2b + c)x^2 + (2b - c)x + b$.
18. $T: P_2 \to P_2$ where $T(x^2) = x^2$, $T(x) = 2x + 1$ and $T(1) = x^2 - 1$.
19. $T: M_{2\times 2} \to M_{2\times 2}$ given by $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a & -b \\ 3c & 0 \end{bmatrix}$.

Exercises 20-34. Determine if the given matrix A is diagonalizable. If it is, find a matrix P such that $P^{-1}AP$ is a diagonal matrix.

20. $A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

21. $A = \begin{bmatrix} 1 & -6 \\ -1 & 2 \end{bmatrix}$

22. $A = \begin{bmatrix} 2 & -4 \\ 3 & -3 \end{bmatrix}$

23. $A = \begin{bmatrix} 1 & 3 & 5 \\ 0 & -2 & 6 \\ 0 & 0 & 4 \end{bmatrix}$

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

25. $A = \begin{bmatrix} 2 & 3 & 4 \\ 2 & 3 & 0 \\ 0 & 0 & 5 \end{bmatrix}$

26. $A = \begin{bmatrix} 2 & 0 & 3 \\ -1 & 0 & 3 \\ 1 & 0 & 2 \end{bmatrix}$

27. $A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 3 \\ -1 & 2 & -3 \end{bmatrix}$

28. $A = \begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix}$

29. $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 5 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

30. $A = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 \\ -1 & 9 & 1 & 0 \\ 3 & 3 & 5 & 7 \end{bmatrix}$

31. $A = \begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}$

32. $A = \begin{bmatrix} 2 & 5 & 3 & -6 \\ -2 & 4 & 0 & 2 \\ -4 & -10 & -6 & 12 \\ 1 & 7 & 3 & -5 \end{bmatrix}$

33. $A = \begin{bmatrix} 3 & 1 & 2 & -2 & 4 \\ 0 & 1 & 4 & 4 & 2 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$

34. $A = \begin{bmatrix} 2 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 1 \\ 0 & 0 & 5 & 3 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{bmatrix}$

35. Let $A \in M_{n \times n}$ be such that $A^2 = I$. Show that:
    (a) If λ is an eigenvalue of A, then $\lambda = 1$ or $\lambda = 0$.
    (b) A is diagonalizable.

36. Let $A \in M_{n \times n}$ be diagonalizable. Prove that the rank of A is equal to the number of nonzero eigenvalues of A.

37. Prove that if $A \in M_{n \times n}$ is similar to a diagonal matrix, then A is diagonalizable.

38. Let $A \in M_{n \times n}$. Prove that A and its transpose A^T have the same eigenvalues, and that they occur with equal algebraic multiplicity (see Exercise 19, page 161).

39. Let $A \in M_{n \times n}$. Prove that if λ is an eigenvalue of A with geometric multiplicity d, then λ is an eigenvalue of its transpose A^T with geometric multiplicity d (see Exercise 19, page 161).

40. Let $L: V \to W$ be an isomorphism on a finite dimensional vector space. Prove that:
    (a) The linear operator $T: V \to V$ and $L \circ T \circ L^{-1}: W \to W$ have equal characteristic polynomials.
    (b) The eigenspace corresponding to an eigenvalue λ of $L \circ T \circ L^{-1}$ is isomorphic to the eigenspace corresponding to that eigenvalue of T.
    (c) T is diagonalizable if and only if $L \circ T \circ L^{-1}$ is diagonalizable.

PROVE OR GIVE A COUNTEREXAMPLE

41. Let $T: V \to V$ be a linear operator on a space of dimension n. If $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct eigenvalues of T, and if there exists a basis β for V such that $[T]_\beta$ is a diagonal matrix, then $m = n$.

42. Let $\lambda_1, \lambda_2, \ldots, \lambda_k$ be the distinct eigenvalues of a linear operator $T: V \to V$ on a vector space V of dimension n. The operator T is diagonalizable if and only if $k = n$.

43. If $A, B \in M_{n \times n}$ are both diagonalizable, then so is AB.

44. If $A, B \in M_{n \times n}$ are such that AB is diagonalizable, then both A and B are diagonalizable.

§4. APPLICATIONS

Applications of eigenvectors surface in numerous fields. In this section we focus on recurrence relations, and on differential equations.

FIBONACCI NUMBERS AND SYSTEMS OF DIFFERENTIAL EQUATIONS

The Fibonacci sequence is that sequence whose first two terms are 1, and with each term after the second being obtained by summing its two immediate predecessors:

$$1, 1, \overset{1+1\,\downarrow}{2}, \underset{1+2\,\uparrow}{3}, 5, \mathbf{8}, \mathbf{13}, \overset{\mathbf{8+13}\,\downarrow}{21}, 34, 55, \underset{\mathbf{34+55}\,\uparrow}{\ldots}$$

Leonardo Fibonacci (Italian; circa 1170 - 1250), is considered by many to be the best mathematician of the Middle Ages. The sequence bearing his name evolved from the following question he posed and resolve in 1220:

Assume that pairs of rabbits do no produce offspring during their first month of life, but will produce a new pair of offspring each month thereafter. Assuming that no rabbit dies, how many pairs of rabbits will there be after k months?

What is the 100^{th} Fibonacci number? Given enough time, we could keep generating the numbers of the sequence, eventually arriving at its 100^{th} element. There is a better way:

Letting s_k denote the k^{th} Fibonacci number we have $s_1 = s_2 = 1$, and, for $k \geq 3$:

$$\boldsymbol{s_k = s_{k-1} + s_{k-2}}$$

We observe that $\boldsymbol{s_k}$ is the top entry of the matrix product:

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} s_{k-1} \\ s_{k-2} \end{bmatrix} = \begin{bmatrix} \boldsymbol{s_{k-1} + s_{k-2}} \\ s_{k-1} \end{bmatrix} = \begin{bmatrix} \boldsymbol{s_k} \\ s_{k-1} \end{bmatrix} \qquad (*)$$

Letting $\boldsymbol{F} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and $\boldsymbol{S_k} = \begin{bmatrix} s_k \\ s_{k-1} \end{bmatrix}$ we can express (*) in the form:

$$\boldsymbol{S_k = FS_{k-1}}$$

In particular

$$\boldsymbol{S_3 = FS_2 = F\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ S_4 = FS_3 = F \cdot F\begin{bmatrix} 1 \\ 1 \end{bmatrix} = F^2\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ldots, \ S_k = F^{k-2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$$

Note that the k^{th} Fibonacci number $\boldsymbol{s_k}$ in (*) is the **top entry** in the matrix $\boldsymbol{S_k} = \begin{bmatrix} \boldsymbol{s_k} \\ s_{k-1} \end{bmatrix} = \boldsymbol{F^{k-2}\begin{bmatrix} 1 \\ 1 \end{bmatrix}}$, which is simply the sum of the entries in the first row of $\boldsymbol{F^{k-2}}$ (see margin). But this is of little benefit unless we can readily find the powers of the matrix $F = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$. We can:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{a+b} \\ c+d \end{bmatrix}$$

Since the characteristic polynomial of F is $\lambda^2-\lambda-1$ (margin), the matrix F has eigenvalues $\lambda_1 = \frac{1-\sqrt{5}}{2}$ and $\lambda_2 = \frac{1+\sqrt{5}}{2}$. Let's find eigenvectors associated with those eigenvectors:

$$\det\left(\begin{bmatrix}1&1\\1&0\end{bmatrix}-\lambda\begin{bmatrix}1&0\\0&1\end{bmatrix}\right)=0$$

$$\det\begin{bmatrix}1-\lambda&1\\1&-\lambda\end{bmatrix}=0$$

$$\lambda^2-\lambda-1=0$$

$$\lambda=\frac{1\pm\sqrt{1+4}}{2}$$

$$\begin{bmatrix}1&1\\1&0\end{bmatrix}\begin{bmatrix}a\\b\end{bmatrix}=\left(\frac{1+\sqrt{5}}{2}\right)\begin{bmatrix}a\\b\end{bmatrix}\Rightarrow\begin{bmatrix}a+b\\a\end{bmatrix}=\begin{bmatrix}\frac{1+\sqrt{5}}{2}a\\[4pt]\frac{1+\sqrt{5}}{2}b\end{bmatrix}\Rightarrow\underbrace{\begin{aligned}\frac{1-\sqrt{5}}{2}a+b&=0\\a-\frac{1+\sqrt{5}}{2}b&=0\end{aligned}}_{\text{homogeneous system of equations}}$$

$$\begin{array}{c}\begin{array}{cc}a & \qquad b\end{array}\\\begin{bmatrix}\frac{1-\sqrt{5}}{2}&1\\[4pt]1&-\frac{1+\sqrt{5}}{2}\end{bmatrix}\end{array}\xrightarrow{\text{rref}}\begin{array}{c}\begin{array}{cc}a & \quad b\end{array}\\\begin{bmatrix}1&-\frac{1+\sqrt{5}}{2}\\[4pt]0&0\end{bmatrix}\end{array}\overset{\text{solution}}{\Rightarrow} a=\frac{1+\sqrt{5}}{2}b$$

Setting the free variable b to 1 we find that $a = \frac{1+\sqrt{5}}{2}$. It follows that $\begin{bmatrix}a\\b\end{bmatrix}=\begin{bmatrix}\frac{1+\sqrt{5}}{2}\\1\end{bmatrix}$ is an eigenvector associated with the eigenvalue λ_1.

In the same manner one can show that $\begin{bmatrix}\frac{1-\sqrt{5}}{2}\\1\end{bmatrix}$ is an eigenvector associated with the eigenvalue λ_2 — a fact that is verified in the margin.

$$\begin{bmatrix}1&1\\1&0\end{bmatrix}\begin{bmatrix}\frac{1-\sqrt{5}}{2}\\1\end{bmatrix}=\begin{bmatrix}\frac{1-\sqrt{5}}{2}+1\\[4pt]\frac{1-\sqrt{5}}{2}\end{bmatrix}$$

$$=\begin{bmatrix}\frac{3-\sqrt{5}}{2}\\[4pt]\frac{1-\sqrt{5}}{2}\end{bmatrix}$$

$$=\frac{1-\sqrt{5}}{2}\begin{bmatrix}\frac{1-\sqrt{5}}{2}\\1\end{bmatrix}$$

Theorem 6.15, page 236, tell us that:

$$D=\underbrace{\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}^{-1}\begin{bmatrix}\mathbf{1}&\mathbf{1}\\\mathbf{1}&\mathbf{0}\end{bmatrix}\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}}_{P^{-1}\boldsymbol{F}P}=\overset{\text{a diagonal matrix}}{\begin{bmatrix}\mathbf{\frac{1+\sqrt{5}}{2}}&\mathbf{0}\\\mathbf{0}&\mathbf{\frac{1-\sqrt{5}}{2}}\end{bmatrix}}$$

Leading us to:

$$\boldsymbol{F}=P\boldsymbol{D}P^{-1}=\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}\begin{bmatrix}\mathbf{\frac{1+\sqrt{5}}{2}}&\mathbf{0}\\\mathbf{0}&\mathbf{\frac{1-\sqrt{5}}{2}}\end{bmatrix}\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}^{-1}$$

Applying Theorem 6.16, page 238, we have:

$$F^k=PD^kP^{-1}=\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}\begin{bmatrix}\left(\frac{1+\sqrt{5}}{2}\right)^k&0\\0&\left(\frac{1-\sqrt{5}}{2}\right)^k\end{bmatrix}\begin{bmatrix}\frac{1+\sqrt{5}}{2}&\frac{1-\sqrt{5}}{2}\\1&1\end{bmatrix}^{-1}$$

You are invited to show that the first row of the above matrix product can be expressed in the following form:

$$F^k = \frac{1}{\sqrt{5}}\begin{bmatrix} \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} & \left(\frac{1+\sqrt{5}}{2}\right)^{k} - \left(\frac{1-\sqrt{5}}{2}\right)^{k} \\ ********************* & ***************** \end{bmatrix}$$

Recalling that the k^{th} Fibonacci number is the sum of the entries in the first row of F^{k-2}, we have:

$$\boldsymbol{s_k} = \frac{\mathbf{1}}{\sqrt{\mathbf{5}}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-1} + \left(\frac{1+\sqrt{5}}{2}\right)^{k-2} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-2}\right]$$

$$= \frac{\mathbf{1}}{\sqrt{\mathbf{5}}}\left\{\left[\left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-2}\right] - \left[\left(\frac{1-\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-1}\right]\right\}$$

$$= \frac{\mathbf{1}}{\sqrt{\mathbf{5}}}\left[\left(\frac{\mathbf{1}+\sqrt{\mathbf{5}}}{\mathbf{2}}\right)^{\boldsymbol{k}} - \left(\frac{\mathbf{1}-\sqrt{\mathbf{5}}}{\mathbf{2}}\right)^{\boldsymbol{k}}\right] \text{ (see margin)}$$

$$\left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1+\sqrt{5}}{2}\right)^{k-2}$$

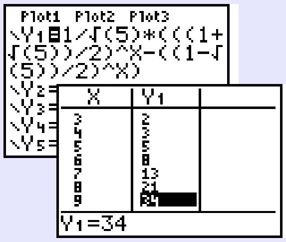
$$= \left(\frac{1+\sqrt{5}}{2}\right)^{k-2}\left[\left(\frac{1+\sqrt{5}}{2}\right) - 1\right]$$

$$= \left(\frac{1+\sqrt{5}}{2}\right)^{k-2}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{2}\right]$$

$$= \left(\frac{1+\sqrt{5}}{2}\right)^{k-2}$$

Looking at the above "$\sqrt{5}$-expression" from a strictly algebraic point of view, one would not expect to find that each s_k is an integer. Being a Fibonacci number, that must be the case. In particular, the 100^{th} Fibonacci number is:

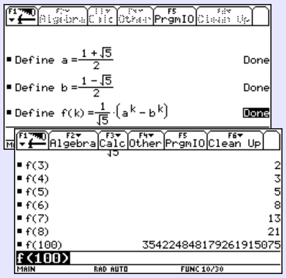
$$s_{100} = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{100} - \left(\frac{1-\sqrt{5}}{2}\right)^{100}\right] = 354{,}224{,}848{,}179{,}261{,}915{,}075$$



```
Plot1 Plot2 Plot3
\Y1■1/√(5)*(((1+
√(5))/2)^X-((1-√
(5))/2)^X)
\Y2=
\Y3=
\Y4=
\Y5=
```

```
  X  |  Y1  |
  3  |  2
  4  |  3
  5  |  5
  6  |  8
  7  |  13
  8  |  21
  9  |  34
Y1=34
```

Using the TI-92:

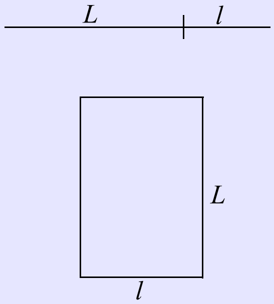


```
■ Define a = (1 + √5)/2                Done
■ Define b = (1 − √5)/2                Done
■ Define f(k) = 1/√5 · (a^k − b^k)     Done
```

```
■ f(3)                                    2
■ f(4)                                    3
■ f(5)                                    5
■ f(6)                                    8
■ f(7)                                   13
■ f(8)                                   21
■ f(100)           354224848179261915075
f(100)
MAIN          RAD AUTO          FUNC 10/30
```

The number $\phi = \frac{1+\sqrt{5}}{2}$ has an interesting history dating back to the time of Pythagoras (c. 500 B.C.). It is called the golden ratio (ϕ is the first letter in the Greek spelling of Phydias, a sculptor who used the golden ratio in his work).

Basically, and for whatever aesthetic reason, it is generally maintained that the most "visually appealing" partition of a line segment into two pieces is that for which the ratio of the length of the longer piece L to the length of the sorter piece l equals the ratio of the entire line segment to that of the longer piece, leading us to:



$$\frac{L}{l} = \frac{L+l}{L}$$

$$L^2 - lL - l^2 = 0$$

$$L = \frac{l + l\sqrt{5}}{2} \Rightarrow \frac{L}{l} = \frac{1+\sqrt{5}}{2}$$

RECURSIVE RELATION

The formula $s_k = s_{k-1} + s_{k-2}$ for the k^{th} element of the Fibonacci sequence describes each element of the sequence in terms of previous elements. It is an example of a **recurrence relation**. You are invited to consider addition recurrence relations in the exercises, and in the following Check Your Understanding box.

CHECK YOUR UNDERSTANDING 6.16

Find a formula for the k^{th} term of the sequence $s_1, s_2, s_3, \ldots,$ if $s_1 = 2$, $s_2 = 3$, and $s_k = s_{k-1} + 2s_{k-2}$ for $k \geq 3$.

Answer: See page B-27.

SYSTEMS OF DIFFERENTIAL EQUATIONS (CALCULUS DEPENDENT)

We begin by extending the concept of a matrix to allow for function entries; as with:

$$A(x) = \begin{bmatrix} 3x & e^{2x} \\ 4 & \sin x \end{bmatrix} \text{ and } B(x) = \begin{bmatrix} 0 & \ln x \\ x-5 & 2x \end{bmatrix}$$

The arithmetic of such matrices mimics that of numerical matrices. For example:

$$\begin{bmatrix} 3x & e^{2x} \\ 4 & \sin x \end{bmatrix} + \begin{bmatrix} 0 & \ln x \\ x-5 & 2x \end{bmatrix} = \begin{bmatrix} 3x & e^{2x} + \ln x \\ x-1 & \sin x + 2x \end{bmatrix}$$

and:

$$\begin{bmatrix} 3x & e^{2x} \\ 4 & \sin x \end{bmatrix} \begin{bmatrix} 0 & \ln x \\ x-5 & 2x \end{bmatrix} = \begin{bmatrix} e^{2x}(x-5) & 3x\ln x + 2xe^{2x} \\ \sin x(x-5) & 4\ln x + 2x\sin x \end{bmatrix}$$

We also define the derivative of a function-matrix to be that matrix obtained by differentiating each of its entry. For example:

$$\text{If } A(x) = \begin{bmatrix} 3x & e^{2x} \\ 4 & \sin x \end{bmatrix}, \text{ then } A'(x) = \begin{bmatrix} (3x)' & (e^{2x})' \\ (4)' & (\sin x)' \end{bmatrix} = \begin{bmatrix} x & 2e^{2x} \\ 0 & \cos x \end{bmatrix}$$

In the exercises, you are invited to show that the following familiar derivative properties:

$$[f(x) + g(x)]' = f'(x) + g'(x)$$
$$[f(x)g(x)]' = f(x)g'(x) + g(x)f'(x)$$
$$[cf(x)]' = cf'(x)$$

extend to matrices:

THEOREM 6.19 Let the entries of the matrices $A(x)$ and $B(x)$ be differentiable function, and let C be a matrix with scalar entries (real numbers). Then:

(i) $[A(x)+B(x)]' = A'(x)+B'(x)$

(ii) $[A(x)B(x)]' = A(x)B'(x)+B(x)A'(x)$

(iii) $[CA(x)]' = CA'(x)$

(assuming, of course, that the matrix dimensions are such that the operations are defined)

Differential equations of the form:

$$f'(x) = af(x) \qquad (\text{or } y' = ay)$$

play an important role in the development of this subsection. As you may recall:

THEOREM 6.20 The solution set of $f'(x) = af(x)$, conists of all functions of the form $f(x) = ce^{ax}+d$ for any constants c and d.

PROOF:

$$f'(x) = (ce^{ax}+d)' = a(ce^{ax}) = af(x)$$

At this point, we know that every function of the form $y = ce^{ax}$ is a solution of the differential equation $f'(x) = af(x)$. Moreover, if $f(x)$ is **any** solution of $f'(x) = af(x)$, then the derivative of the function $g(x) = \frac{f(x)}{e^{ax}}$ is zero:

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{g(x)f'(x)-f(x)g'(x)}{[g(x)]^2}$$

$$g'(x) = \frac{e^{ax}f'(x)-f(x)ae^{ax}}{e^{2ax}} = \frac{[f'(x)-af(x)]e^{ax}}{e^{2ax}} = \frac{\overset{\text{since } f'(x) = af(x)}{0}}{e^{ax}} = 0$$

If the derivative of a function is zero, then the function must be constant.

It follows that $g(x) = \frac{f(x)}{e^{ax}} = c$ for some constant c, or that $f(x) = ce^{ax}$.

We now turn our attention to systems of linear differential equation of the form:

$$\left.\begin{aligned} f_1'(x) &= a_{11}f_1(x)+a_{12}f_2(x)+\ldots+a_{1n}f_n(x) \\ f_2'(x) &= a_{21}f_1(x)+a_{22}f_2(x)+\ldots+a_{2n}f_n(x) \\ &\vdots \\ f_n'(x) &= a_{n1}f_1(x)+a_{n2}f_2(x)+\ldots+a_{nn}f_n(x) \end{aligned}\right\}$$

where the coefficients a_{ij} are real numbers. As it is with systems of linear equations, the above system can be expressed in the form:

$$F'(x) = AF(x)$$

were $F(x) = [f_i(x)] \in M_{n\times 1}$ and $A = [a_{ij}] \in M_{n\times n}$.

In the event that A is a diagonal matrix, the system $F'(x) = AF(x)$ is easily solved:

$$\left.\begin{array}{l} f_1'(x) = \lambda_1 f_1(x) \\ f_2'(x) = \lambda_2 f_2(x) \\ \\ f_n'(x) = \lambda_n f_n(x) \end{array}\right\}$$

THEOREM 6.21 If $\begin{bmatrix} f_1'(x) \\ f_2'(x) \\ \vdots \\ f_n'(x) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{bmatrix}$ then: $\begin{bmatrix} f_1(x) \\ f_2(x) \\ \\ f_n(x) \end{bmatrix} = \begin{bmatrix} c_1 e^{\lambda_1 x} \\ c_2 e^{\lambda_2 x} \\ \\ c_n e^{\lambda_n x} \end{bmatrix}$

where $c_1, c_2, \dots c_n \in \Re$.

PROOF: Simply apply Theorem 6.20 to each of the n differential equations: $f_i'(x) = \lambda_i f_i(x)$.

We now consider systems of differential equations of the form:

$$F'(x) = AF(x) \qquad (*)$$

where A is a diagonalizable matrix. In accordance with Theorem 6.15, page 236, we know that for any chosen basis $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ of eigenvectors of A:

$$A = PDP^{-1}$$

where the i^{th} column of $P \in M_{n \times n}$ is the eigenvector $\boldsymbol{v_i} \in (a_{1i}, a_{2i}, \dots, a_{ni})$ with eigenvalue λ_i, and $D = [d_{ij}]$ is the diagonal matrix with $d_{ii} = \lambda_i$. Substituting in (*), we have:

$$F'(x) = (PDP^{-1})F(x)$$

Multiply both sides by P^{-1}: $$P^{-1}F'(x) = DP^{-1}F(x)$$

Theorem 6.19(iii): $$[\boldsymbol{P^{-1}F(x)}]' = D\boldsymbol{P^{-1}F(x)}$$

Letting $\boldsymbol{G(x)} = \boldsymbol{P^{-1}F(x)}$, brings us to:

$$\boldsymbol{G'(x)} = D\boldsymbol{G(x)}$$

Applying Theorem 6.21, we have:

$$\boldsymbol{G(x)} = \begin{bmatrix} c_1 e^{\lambda_1 x} \\ c_2 e^{\lambda_2 x} \\ \\ c_n e^{\lambda_n x} \end{bmatrix} \Rightarrow \boldsymbol{P^{-1}F(x)} = \begin{bmatrix} c_1 e^{\lambda_1 x} \\ c_2 e^{\lambda_2 x} \\ \\ c_n e^{\lambda_n x} \end{bmatrix} \Rightarrow F(x) = P\begin{bmatrix} c_1 e^{\lambda_1 x} \\ c_2 e^{\lambda_2 x} \\ \\ c_n e^{\lambda_n x} \end{bmatrix}$$

At this point we have:

$$F(x) = \begin{array}{c} \begin{array}{cccc} \boldsymbol{v_1} & \boldsymbol{v_2} & & \boldsymbol{v_n} \\ \downarrow & \downarrow & & \downarrow \end{array} \\ \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \end{array} \begin{bmatrix} c_1 e^{\lambda_1 x} \\ c_2 e^{\lambda_2 x} \\ \\ c_n e^{\lambda_n x} \end{bmatrix}$$

Appealing to Theorem 5.3, page 154, we conclude that:

$$F(x) = c_1 e^{\lambda_1 x}\mathbf{v}_1 + c_2 e^{\lambda_2 x}\mathbf{v}_2 + \dots + c_n e^{\lambda_n x}\mathbf{v}_n$$

Summarizing, we have:

THEOREM 6.22 Let $A \in M_{n \times n}$ be diagonalizable, and let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for $\Re^n$ consisting entirely of eigenvectors of A with corresponding eigenvalues λ_i. Then, the general solution of:

$$F'(x) = AF(x)$$

is of the form:

$$c_1 e^{\lambda_1 x}\boldsymbol{v}_1 + c_2 e^{\lambda_2 x}\boldsymbol{v}_2 + \cdots + c_n e^{\lambda_n x}\boldsymbol{v}_n$$

for $c_1, c_2, \ldots c_n \in \Re$.

In alternate notation form:

$$\left.\begin{aligned} f_1{}'(x) &= -3f_1(x) + f_2(x) \\ f_2{}'(x) &= 6f_1(x) + 2f_2(x) \end{aligned}\right\}$$

EXAMPLE 6.14 Find the general solution for:

$$\left.\begin{aligned} y_1{}' &= -3y_1 + y_2 \\ y_2{}' &= 6y_1 + 2y_2 \end{aligned}\right\}$$

SOLUTION: Our first order of business is to find (if possible) a basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ of $\Re^2$ consisting of eigenvectors of $A = \begin{bmatrix} -3 & 1 \\ 6 & 2 \end{bmatrix}$.

From: $\det\begin{bmatrix} -3-\lambda & 1 \\ 6 & 2-\lambda \end{bmatrix} = (-3-\lambda)(2-\lambda) - 6$

$$= \lambda^2 + \lambda - 12 = (\lambda + 4)(\lambda - 3)$$

we see that the matrix $A = \begin{bmatrix} -3 & 1 \\ 6 & 2 \end{bmatrix}$ is diagonalizable, with eigenvalues -4 and 3. Here are their corresponding eigenspaces:

$$E[-4] = \text{null}\left(\begin{bmatrix} -3 & 1 \\ 6 & 2 \end{bmatrix} - \begin{bmatrix} -4 & 0 \\ 0 & -4 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} 1 & 1 \\ 6 & 6 \end{bmatrix}\right)$$

homogeneous system of equations:

$$\begin{array}{cc} x & y \\ \hline 1 & 1 \\ 6 & 6 \end{array} \xrightarrow{\text{rref}} \begin{array}{cc} x & y \\ \hline \mathbf{1} & 1 \\ 0 & 0 \end{array}$$

Setting the free variable b equal to r, we have:

$$E[-4] = \{(-r, r) | r \in R\}$$

And:

$$E[3] = \text{null}\left(\begin{bmatrix} -3 & 1 \\ 6 & 2 \end{bmatrix} - \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} -6 & 1 \\ 6 & -1 \end{bmatrix}\right)$$

$$\begin{array}{cc} a & b \\ \hline -6 & 1 \\ 6 & -1 \end{array} \xrightarrow{\text{rref}} \begin{array}{cc} a & b \\ \hline \mathbf{1} & -\frac{1}{6} \\ 0 & 0 \end{array}$$

Bringing us to:

$$E[3] = \{(r, 6r) | r \in R\}$$

Any other two eigenvectors corresponding to the two eigenvalues will do just as well.

Choosing the eigenvector $\boldsymbol{v_1} = (-1, 1)$ for the eigenvalue -4 and the eigenvector $\boldsymbol{v_2} = (1, 6)$ for the eigenvalue 3, we obtain a basis of $\Re^2$ consisting of eigenvectors for A. Applying Theorem 6.22, we conclude that the general solution of the given system of differential equations is give by:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = c_1e^{-4x}\begin{bmatrix} -1 \\ 1 \end{bmatrix} + c_2e^{3x}\begin{bmatrix} 1 \\ 6 \end{bmatrix} = \begin{bmatrix} -c_1e^{-4x} + c_2e^{3x} \\ c_1e^{-4x} + 6c_2e^{3x} \end{bmatrix}$$

Which is to say:

$$y_1 = -c_1e^{-4x} + c_2e^{3x} \quad \text{and} \quad y_2 = c_1e^{-4x} + 6c_2e^{3x}$$

Let's check our result in the given system $\left.\begin{aligned} \boldsymbol{y_1}' &= -3\boldsymbol{y_1} + \boldsymbol{y_2} \\ \boldsymbol{y_2}' &= 6\boldsymbol{y_1} + 2\boldsymbol{y_2} \end{aligned}\right\}$:

$$\boldsymbol{y_1}' = (-c_1e^{-4x} + c_2e^{3x})' = \mathbf{4c_1e^{-4x} + 3c_2e^{3x}}$$

and:

$$-\,\mathbf{3}\boldsymbol{y_1} + \boldsymbol{y_2} = -3(-c_1e^{-4x} + c_2e^{3x}) + (c_1e^{-4x} + 6c_2e^{3x}) = \mathbf{4c_1e^{-4x} + 3c_2e^{3x}}$$

Similarly: $\boldsymbol{y_2}' = (c_1e^{-4x} + 6c_2e^{3x})' = -4c_1e^{-4x} + 18c_2e^{3x}$

$$= 6(-c_1e^{-4x} + c_2e^{3x}) + 2(c_1e^{-4x} + 6c_2e^{3x}) = \mathbf{6}\boldsymbol{y_1} + \mathbf{2}\boldsymbol{y_2}$$

CHECK YOUR UNDERSTANDING 6.17

Find the general solution for:

$$\left.\begin{aligned} y_1' &= 2y_2 - 2y_3 \\ y_2' &= y_1 + y_2 - y_4 \\ y_3' &= -y_1 + y_2 - 2y_3 + y_4 \\ y_4' &= -y_1 + y_2 - 2y_3 + y_4 \end{aligned}\right\}$$

Suggestion: Consider Example 6.11, page 237.

Answer: See page B-28.

Let us return momentarily to the system of equations of Example 6.14:

$$\left.\begin{aligned} y_1' &= -3y_1 + y_2 \\ y_2' &= 6y_1 + 2y_2 \end{aligned}\right\} \text{ with general solution: } \begin{aligned} y_1 &= -c_1e^{-4x} + c_2e^{3x} \\ y_2 &= c_1e^{-4x} + 6c_2e^{3x} \end{aligned} \quad \textbf{(*)}$$

To arrive a particular or specific solution for the system, we need some additional information. Suppose, for example, that we are given the initial condition $Y(0) = \begin{bmatrix} y_1(\mathbf{0}) \\ y_2(\mathbf{0}) \end{bmatrix} = \begin{bmatrix} \mathbf{-2} \\ \mathbf{3} \end{bmatrix}$. Substituting in **(*)**, we then have:

$$\left.\begin{aligned} -2 &= -c_1e^{-4\cdot\mathbf{0}}+c_2e^{3\cdot\mathbf{0}} \\ 3 &= c_1e^{-4\cdot\mathbf{0}}+6c_2e^{3\cdot\mathbf{0}} \end{aligned}\right\} \Rightarrow \left.\begin{aligned} -2 &= -c_1+c_2 \\ 3 &= c_1+6c_2 \end{aligned}\right\} \Rightarrow \begin{aligned} c_1 &= \frac{15}{7} \\ c_2 &= \frac{1}{7} \end{aligned}$$

Solution:
$$\begin{aligned} y_1 &= -\tfrac{1}{7}e^{-4x}+\tfrac{15}{7}e^{3x} \\ y_2 &= \tfrac{1}{7}e^{-4x}+\tfrac{90}{7}e^{3x} \end{aligned}$$

CHECK YOUR UNDERSTANDING 6.18

Find the specific solution of the system in CYU 6.17, if

$$y_1(0)=0, y_2(0)=1, y_3(0)=2, \text{ and } y_4(0)=3$$

Answer: See page B-28.

EXAMPLE 6.15 In the forest of Illtrode lived a small peaceful community of 50 elves, when they were suddenly invaded by 25 trolls. The wizard Callandale quickly determined that:

$$\begin{aligned} \frac{d}{dt}T(t) &= \tfrac{1}{2}T(t)-\tfrac{1}{9}E(t) \\ \frac{d}{dt}E(t) &= -T(t)+\tfrac{1}{2}E(t) \end{aligned}$$

where $T(t)$ and $E(t)$ represents the troll and elves populations t years after the troll invasion. Analyze the nature of the two populations as time progresses.

SOLUTION: To find the general solution of the system:

$$\begin{bmatrix} T'(t) \\ E'(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{9} \\ -1 & \frac{1}{2} \end{bmatrix}\begin{bmatrix} T(t) \\ E(t) \end{bmatrix}$$

we first find the eigenvalues of the above 2×2 matrix:

$$\det\begin{bmatrix} \frac{1}{2}-\lambda & -\frac{1}{9} \\ -1 & \frac{1}{2}-\lambda \end{bmatrix} = (\tfrac{1}{2}-\lambda)^2-\frac{1}{9} = 0 \Rightarrow \tfrac{1}{2}-\lambda = \pm\frac{1}{3} \Rightarrow \begin{aligned} \lambda &= \tfrac{5}{6} \\ \lambda &= \tfrac{1}{6} \end{aligned}$$

Then: $$E\left[\frac{5}{6}\right] = \text{null}\left(\begin{bmatrix} \frac{1}{2} & -\frac{1}{9} \\ -1 & \frac{1}{2} \end{bmatrix} - \begin{bmatrix} \frac{5}{6} & 0 \\ 0 & \frac{5}{6} \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} -\frac{1}{3} & -\frac{1}{9} \\ -1 & -\frac{1}{3} \end{bmatrix}\right)$$

homogeneous system of equations:
$$\begin{array}{cc} x & y \\ \hline -\frac{1}{3} & -\frac{1}{9} \\ -1 & -\frac{1}{3} \end{array} \xrightarrow{\textbf{rref}} \begin{array}{cc} x & y \\ \hline \mathbf{1} & \frac{1}{3} \\ 0 & 0 \end{array}$$

Setting the free variable y to 3, we obtain the eigenvector $(-1, 3)$ for the eigenvalue $\frac{5}{6}$. In a similar fashion, you can show that $(1, 3)$ is an eigenvector for $\frac{1}{6}$. This leads us to the general solution:

$$\begin{bmatrix} T(t) \\ E(t) \end{bmatrix} = c_1 e^{\frac{5}{6}t} \begin{bmatrix} -1 \\ 3 \end{bmatrix} + c_2 e^{\frac{1}{6}t} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -c_1 e^{\frac{5}{6}t} + c_2 e^{\frac{1}{6}t} \\ 3c_1 e^{\frac{5}{6}t} + 3c_2 e^{\frac{1}{6}t} \end{bmatrix}$$

Turning to the initial conditions, we have:

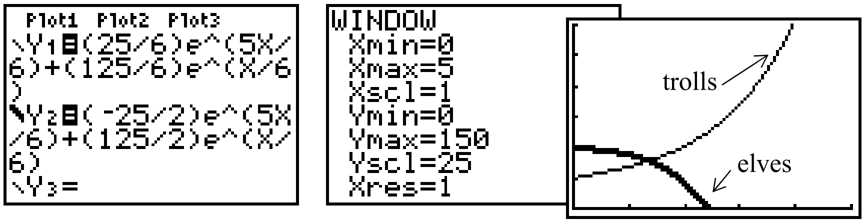
$$\begin{bmatrix} T(0) \\ E(0) \end{bmatrix} = \begin{bmatrix} 25 \\ 50 \end{bmatrix} \Rightarrow \left.\begin{aligned} 25 &= -c_1 + c_2 \\ 50 &= 3c_1 + 3c_2 \end{aligned}\right\} \Rightarrow c_1 = -\frac{25}{6} \text{ and } c_2 = \frac{125}{6}$$

Leading us to the specific solution:

$$T(t) = \frac{25}{6} e^{\frac{5}{6}t} + \frac{125}{6} e^{\frac{1}{6}t}$$

$$E(t) = -\frac{25}{2} e^{\frac{5}{6}t} + \frac{125}{2} e^{\frac{1}{6}t}$$

A consideration of the graphs of the two functions reveals that while the troll population will continue to flourish in the region, the poor elves vanish around two-and-a-half years following the invasion:



```
Plot1 Plot2 Plot3
\Y1■(25/6)e^(5X/
6)+(125/6)e^(X/6
)
\Y2■(-25/2)e^(5X
/6)+(125/2)e^(X/
6)
\Y3=
```

```
WINDOW
 Xmin=0
 Xmax=5
 Xscl=1
 Ymin=0
 Ymax=150
 Yscl=25
 Xres=1
```

CHECK YOUR UNDERSTANDING 6.19

Turtles and frogs are competing for food in a pond, which currently contains 120 turtles and 200 frogs. Assume that the turtles' growth rate and the frogs' growth rate are given by

$$T'(t) = \frac{5T(t)}{2} - \frac{F(t)}{4} \quad \text{and} \quad F'(t) = \frac{5F(t)}{2} - T(t)$$

respectively; where $T(t)$ and $F(t)$ denote the projected turtle and frog population t years from now. Find, to one decimal place, the number of years it will take for the turtle population to equal that of the frog population.

Answer: See page B-29.

EXERCISES

Exercises 1-8. Find a formula for the k^{th} term of the sequence $s_1, s_2, s_3, \ldots$, if:

1. $s_1 = 2$, $s_2 = 2$, and $s_k = s_{k-1} + s_{k-2}$ for $k \geq 3$.

2. $s_1 = a$, $s_2 = a$, and $s_k = s_{k-1} + s_{k-2}$ for $k \geq 3$.

3. $s_1 = 1$, $s_2 = 2$, and $s_k = s_{k-1} + s_{k-2}$ for $k \geq 3$.

4. $s_1 = 1$, $s_2 = 6$, and $s_k = 6s_{k-1} - 9s_{k-2}$ for $k \geq 3$.

5. $s_1 = 1$, $s_2 = 4$, and $s_k = 3s_{k-1} - 2s_{k-2}$ for $k \geq 3$.

6. $s_1 = 1$, $s_2 = 2$, and $s_k = as_{k-1} - bs_{k-2}$ for $k \geq 3$ and $a^2 - 4b > 0$.

7. $s_1 = 1$, $s_2 = 2$, $s_3 = 3$, and $s_k = 2s_{k-1} + s_{k-2} - 2s_{k-3}$ for $k \geq 4$.

   Hint: Note that $S_4 = \begin{bmatrix} \mathbf{2} & \mathbf{1} & \mathbf{-2} \\ \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} & \mathbf{0} \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$.

8. $s_1 = 1$, $s_2 = 2$, $s_3 = 3$, and $s_k = -2s_{k-1} + s_{k-2} + 2s_{k-3}$ for $k \geq 4$.

9. **(PMI)** Let s_k denote the k^{th} Fibonacci number. Prove that $s_k s_{k-2} - (s_{k-1})^2 = (-1)^{k+1}$, for $k \geq 3$.

   Suggestion: Use the Principle of Mathematical Induction to show that for $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ and $k \geq 3$, $A^k = \begin{bmatrix} s_k & s_{k-1} \\ s_{k-1} & s_{k-2} \end{bmatrix}$.

10. Let s_0 and s_1 be the first two elements of a sequence and let $s_k = as_{k-1} + bs_{k-2}$ be a recurrence relation which defines the remaining elements of the sequence. Prove that if the quadratic equation $\lambda^2 - a\lambda - b = 0$ has two distinct solutions, λ_1 and λ_2, then $s_k = c_1\lambda_1^n + c_2\lambda_2^n$ for some $c_1, c_2 \in \Re$.

    Suggestion: Replace the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$ in the development of the Fibonacci sequence with the matrix $\begin{bmatrix} a & b \\ 1 & 0 \end{bmatrix}$.

11. Let the entries of the matrices $A(x)$ and $B(x)$ be differentiable function, and let C be a matrix with scalar entries (real numbers). Given that the dimensions of the matrices are such that the operations can be performed, prove that:
    (i) $[A(x)+B(x)]' = A'(x)+B'(x)$
    (ii) $[A(x)B(x)]' = A(x)B'(x)+B(x)A'(x)$
    (iii) $[CA(x)]' = CA'(x)$

Exercises 12-17. Find the general solution of the given system of differential equations, then check your answer by substitution.

12. $$\left.\begin{aligned} f_1'(x) &= 2f_1(x) \\ f_2'(x) &= 3f_1(x)-f_2(x) \end{aligned}\right\}$$

13. $$\left.\begin{aligned} y_1' &= y_1-y_2 \\ y_2' &= 2y_1+4y_2 \end{aligned}\right\}$$

14. $$\left.\begin{aligned} y_1' &= 3y_1+2y_2 \\ y_2' &= 6y_1-y_2 \end{aligned}\right\}$$

15. $$\left.\begin{aligned} f'(x) &= f(x)+2g(x)-h(x) \\ g'(x) &= 2f(x)+4g(x)-2h(x) \\ h'(x) &= -f(x)-2g(x)+h(x) \end{aligned}\right\}$$

16. $$\left.\begin{aligned} x' &= 4x+3y+3z \\ y' &= -x+3y+8z \\ z' &= -6x+8y+6z \end{aligned}\right\}$$

17. $$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} -1 & 2 & 6 \\ -1 & 3 & 5 \\ 0 & 2 & 5 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Exercises 18-21. Solve the given initial-value problem.

18. $$\begin{bmatrix} y_1' \\ y_2' \end{bmatrix} = \begin{bmatrix} 3 & -1 \\ 6 & -2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \begin{bmatrix} y_1(0) \\ y_2(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

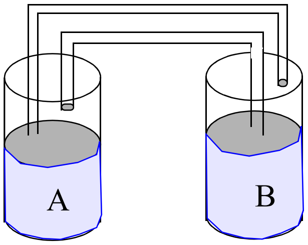
19. $$\begin{bmatrix} f_1'(x) \\ f_2'(x) \end{bmatrix} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix}, \begin{bmatrix} f_1(0) \\ f_2(0) \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

20. $$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 0 & -2 & 1 \\ 0 & 0 & 3 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}, \begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}$$

21. $$\begin{bmatrix} f'(x) \\ g'(x) \\ h'(x) \end{bmatrix} = \begin{bmatrix} -1 & 2 & 6 \\ -1 & 3 & 5 \\ 0 & 2 & 5 \end{bmatrix}\begin{bmatrix} f(x) \\ g(y) \\ h(z) \end{bmatrix}, \begin{bmatrix} f(0) \\ g(0) \\ h(0) \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

22. Given enough space and nourishment, the rate of growth of plants A and B are given by $A'(t) = \frac{3}{2}A(t)$ and $B'(t) = \frac{3}{2}B(t)$, respectively, where t denotes the number of months after planting. One year, 50 of A and 30 of B were planted, and in such a fashion that the rates of growth of each of the two plants were compromised by the presence of the other; in accordance with: $A'(t) = \frac{3}{2}A(t) - B(t)$ and $B'(t) = \frac{3}{2}B(t) = \frac{1}{4}A'(t)$. Analyze the nature of the two plant populations as time progresses.

23. Assume that initially, tank A contains 20 gallons of a liquid solution that is 10% alcohol, and that tank B contains 30 gallons of a solution that is 20% alcohol. At time $t = 0$, the mixture in A is pumped to B at a rate of 1 gallons/minute, while that of B is pumped to A at a rate of 1.5 gallons/minute. Find the percentage of alcohol concentration in each tank t minutes later.



§5. MARKOV CHAINS

Stochos: Greek for "guess." **Stochastices**: Greek for "one who predicts the future." **Andrei Markov**: Russian Mathematician (1856-1922).

Certain systems can occupy a number of distinct states. The transmission in a car, for example, may be in the neutral state, or reverse (state), or first gear (state), etc. When chance plays a role in determining the current state of a system, then the system is said to be a **stochastic process**, and if the probabilities of moving from one state to another remain constant, then the stochastic process is said to be a **Markov process**, or **Markov chain**.

Here is an example of a two-state Markov process:

State **Y**: Person x was involved in an automobile accident within the previous 12 month period.

State **N**: Person x was not involved in an automobile accident within the previous 12 month period.

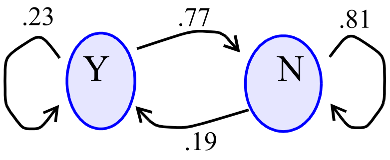
Let's move things along a bit by citing the following study:

$$\left.\begin{array}{l}\text{Probability of } x \text{ being involved}\\ \text{in an accident within the next}\\ \text{12 month period}\end{array}\right\} = \begin{cases} .23 & \text{if x is in } \mathbf{Y} \\ .19 & \text{if x is in } \mathbf{N} \end{cases}$$

Transition matrices are also called probability matrices.

The above information is reflected in Figure 6.2(a) (called a **transition diagram**), wherein each arrow is annotated with the probability of taking that path. The same information is also conveyed in the **transition matrix**, $A = [a_{ij}]$, of Figure 6.2(b), where t_{ij} represents the probability of moving from the j^{th} state to the i^{th} state in the next move. The 0.19 in the upper right-hand corner of the matrix, for example, gives the probability of moving from state **N** to state **Y** in the next move, while the entry 0.81 is the probability of remaining in state **N**.

Since the entries in the transition matrix are probabilities, they **must lie between 0 and 1** (inclusive). Moreover, since the entries down either column account for all possible outcomes (staying in **Y**, or leaving **Y**, for example), their sum **must equal 1**. In particular, since there is a 0.23 probability that a person in state **Y** returns to state **Y**, there has to be a 0.77 probability that the person will leave that state and, consequently, move to state **N**.



Transition Diagram
(a)

current state

$$A: \begin{array}{cc|l} \mathbf{Y} & \mathbf{N} & \\ \hline .23 & .19 & \mathbf{Y} \\ .77 & .81 & \mathbf{N} \end{array}$$

next state

Transition Matrix
(b)

Figure 6.2

Assume that, initially, 25% of the population was involved in an automobile accident within the previous 12 month period (and 75% was not). This given condition brings us to the so called initial state matrix of the system: $S_0 = \begin{bmatrix} \mathbf{.25} \\ \mathbf{.75} \end{bmatrix}$.

Utilizing matrix multiplication we can arrive at the next state, S_1:

$$(*)\quad \overset{T}{\begin{bmatrix} .23 & .19 \\ .77 & .81 \end{bmatrix}}\overset{S_0}{\begin{bmatrix} \mathbf{.25} \\ \mathbf{.75} \end{bmatrix}} = \begin{bmatrix} (.23)(.25)+(.19)(.75) \\ (.77)(.25)+(.81)(.75) \end{bmatrix} = \overset{S_1}{\begin{bmatrix} \mathbf{.20} \\ \mathbf{.80} \end{bmatrix}}\begin{matrix} \mathbf{Y} \\ \mathbf{N} \end{matrix}$$

The above tells us that there is a 0.20 probability that a person will be involved in an accident in the first 12 month period.

To get to the next state matrix, we replace S_0 with S_1 in (*):

$$\overset{T}{\begin{bmatrix} .23 & .19 \\ .77 & .81 \end{bmatrix}}\overset{S_1}{\begin{bmatrix} .20 \\ .80 \end{bmatrix}} = \begin{bmatrix} (.23)(.20)+(.19)(.80) \\ (.77)(.20)+(.81)(.80) \end{bmatrix} = \overset{S_2}{\begin{bmatrix} .198 \\ .802 \end{bmatrix}}\begin{matrix} \mathbf{Y} \\ \mathbf{N} \end{matrix}$$

The above tells us that there is a 0.198 probability that a person will be involved in an accident in the next (second) 12 month period.

$$\text{Similarly:}\quad \overset{T}{\begin{bmatrix} .23 & .19 \\ .77 & .81 \end{bmatrix}}\overset{S_2}{\begin{bmatrix} .198 \\ .802 \end{bmatrix}} = \overset{S_3}{\begin{bmatrix} .19792 \\ .80208 \end{bmatrix}}\begin{matrix} \mathbf{Y} \\ \mathbf{N} \end{matrix}$$

Working backwards, we find that we can also arrive at S_3 by multiplying the initial state matrix S_0 by T^3:

$$S_3 = TS_2 = T(TS_1) = T^2\boldsymbol{S_1} = T^2(\boldsymbol{TS_0}) = T^3\boldsymbol{S_0}$$

Generalizing, we have:

THEOREM 6.23 If T is the transition matrix of a Markov process with **initial-state matrix** S_0, then the n^{th} state matrix in the chain is given by:

$$S_n = T^n S_0$$

CHECK YOUR UNDERSTANDING 6.20

Of the 1560 freshmen at Bright University, 858 live in the dorms. There is a 0.8 probability that a freshman, sophomore, or junior currently living in the dorms will do so in the following year, and a 0.1 probability that a currently commuting student will live on campus next year. Assuming (big assumption) that all 1560 freshmen will graduate, determine (to the nearest integer) the number of the current freshman that will be living in the dorms in their sophomore, junior, and senior years.

Answer: 757, 686, and 636 of the current freshmen will live in the dorm in their sophomore, junior, and senior year, respectively.

POWERS OF THE TRANSITION MATRIX

Let us formally define the concept of a transition matrix:

DEFINITION 6.10
TRANSITION MATRIX

A **transition matrix** $T \in M_{n \times n}$ is a matrix that satisfies the following two properties:

(1) T contains no negative entry.

(2) The sum of the entries in each column of T equals 1.

Consider the three-state Markov process with transition matrix:

$$T\colon \quad \begin{array}{c|ccc|c} & \text{I} & \text{II} & \text{III} & \\ \hline & .3 & \mathbf{.1} & .6 & \text{I} \\ & .2 & \mathbf{.1} & .4 & \text{II} \\ & .5 & \mathbf{.8} & 0 & \text{III} \end{array}$$

Assume that at the start of the process we are in state II:

$$S_0 = \begin{array}{c|c|c} & 0 & \text{I} \\ & \mathbf{1} & \text{II} \\ & 0 & \text{III} \end{array}$$

Observe that: $S_1 = TS_0 = \begin{array}{ccc} \text{I} & \text{II} & \text{III} \\ \end{array} \begin{bmatrix} .3 & \mathbf{.1} & .6 \\ .2 & \mathbf{.1} & .4 \\ .5 & \mathbf{.8} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{1} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{.1} \\ \mathbf{.1} \\ \mathbf{.8} \end{bmatrix} \begin{array}{c} \text{I} \\ \text{II} \\ \text{III} \end{array}$

same

And that: $$S_2 = T^2 S_0 = \begin{bmatrix} .3 & .1 & .6 \\ .2 & .1 & .4 \\ .5 & .8 & 0 \end{bmatrix}^2 \begin{bmatrix} 0 \\ \mathbf{1} \\ 0 \end{bmatrix} = \begin{bmatrix} .41 & \mathbf{.52} & .22 \\ .28 & \mathbf{.35} & .16 \\ .31 & \mathbf{.13} & .62 \end{bmatrix} \begin{bmatrix} 0 \\ \mathbf{1} \\ 0 \end{bmatrix} = \begin{bmatrix} \mathbf{.52} \\ \mathbf{.35} \\ \mathbf{.13} \end{bmatrix} \begin{array}{c} \text{I} \\ \text{II} \\ \text{III} \end{array}$$

same

In general:

THEOREM 6.24

Let T denote the transition matrix of a Markov chain. If the process **starts in state *j***, then the element in the i^{th} row of the j^{th} column of T^m represents the probability of ending up at state i after m steps.

EXAMPLE 6.16

Analyze the nature of the second column of T^2, T^4, and T^8 for the transition matrix:

$$T = \begin{array}{c|ccc|c} & \text{I} & \text{II} & \text{III} & \\ \hline & .2 & .6 & .4 & \text{I} \\ & .3 & .1 & .3 & \text{II} \\ & .5 & .3 & .3 & \text{III} \end{array}$$

Given that the system is initially in state II.

SOLUTION:

$$T^2 = \begin{bmatrix} .2 & .6 & .4 \\ .3 & .1 & .3 \\ .5 & .3 & .3 \end{bmatrix}\begin{bmatrix} .2 & .6 & .4 \\ .3 & .1 & .3 \\ .5 & .3 & .3 \end{bmatrix} = \begin{array}{l} \begin{bmatrix} .42 & \mathbf{.30} & .38 \\ .24 & \mathbf{.28} & .24 \\ .34 & \mathbf{.42} & .38 \end{bmatrix} \end{array}\begin{array}{l} \text{I} \\ \text{II} \\ \text{III} \end{array}$$

The second column of T^2 tells us that if you start in state II, then there is a 0.30, 0.28, and 0.42 probability that you will end up at states I, II, and III, respectively, after two steps.

$$T^4 = T^2 \cdot T^2 = \begin{bmatrix} .42 & .30 & .38 \\ .24 & .28 & .24 \\ .34 & .42 & .38 \end{bmatrix}\begin{bmatrix} .42 & .30 & .38 \\ .24 & .28 & .24 \\ .34 & .42 & .38 \end{bmatrix} = \begin{bmatrix} .3776 & \mathbf{.3696} & .3760 \\ .2496 & \mathbf{.2512} & .2496 \\ .3728 & \mathbf{.3792} & .3744 \end{bmatrix}\begin{array}{l} \text{I} \\ \text{II} \\ \text{III} \end{array}$$

The second column of P^4 tells us that if you start in state II, then there is a 0.3696, 0.2512, and 0.3792 probability that you will end up at states A, B, and C, respectively, after four steps.

$$T^8 = T^4 \cdot T^4 = \begin{bmatrix} .3776 & .3696 & .3760 \\ .2496 & .2512 & .2496 \\ .3728 & .3792 & .3744 \end{bmatrix}\begin{bmatrix} .3776 & .3696 & .3760 \\ .2496 & .2512 & .2496 \\ .3728 & .3792 & .3744 \end{bmatrix} = \begin{bmatrix} .3750 & \mathbf{.3750} & .3750 \\ .2500 & \mathbf{.2500} & .2500 \\ .3750 & \mathbf{.3750} & .3750 \end{bmatrix}$$

Whoa! The three columns of T^8 are identical (to four decimal places). Moreover, if you take higher powers if *T*, you will find that you will again end up at the above T^8 matrix. This suggests that eventually there is, to four decimal places, a 0.375, 0.250, and 0.375 probability, respectively, that you will end up at states I, II, and III, **independently of whether you start at state I, or II, or III!**

Indeed, no matter what initial state you start with, say the state $S_0 = \begin{bmatrix} .2 \\ .5 \\ .3 \end{bmatrix}$, it looks like you will still end up at the same situation:

$$T^8\begin{bmatrix} .2 \\ .5 \\ .3 \end{bmatrix} = \begin{bmatrix} .3750 & .3750 & .3750 \\ .2500 & .2500 & .2500 \\ .3750 & .3750 & .3750 \end{bmatrix}\begin{bmatrix} .2 \\ .5 \\ .3 \end{bmatrix} = \begin{bmatrix} .3750 & .3750 & .3750 \\ .2500 & .2500 & .2500 \\ .3750 & .3750 & .3750 \end{bmatrix}\begin{array}{l} \text{A} \\ \text{B} \\ \text{C} \end{array}$$

It appears that for this Markov chain, there is a probability of 0.375 that you will eventually end up in state *A*, a probability of 0.250 that you will end up in state *B*, and a probability of 0.375 that you will end up in state *C*, **independently of the initial state of the process!** Even more can be said; but first, a definition:

DEFINITION 6.11

$S_F \in \Re^n$ is a **fixed state** for a transition matrix $T \in M_{n \times n}$ if $TS_F = S_F$.

EXAMPLE 6.17

Show that the transition matrix $T = \begin{bmatrix} .2 & .6 & .4 \\ .3 & .1 & .3 \\ .5 & .3 & .3 \end{bmatrix}$ of Example 6.16 has a fixed state.

SOLUTION: We are to show that there exists a state $S_F = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $TS_F = S_F$:

$$\begin{bmatrix} .2 & .6 & .4 \\ .3 & .1 & .3 \\ .5 & .3 & .3 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \begin{bmatrix} .2x + .6y + .4z \\ .3x + .1y + .3z \\ .5x + .3y + .3z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Equating entries brings us to a system of three equations in three unknowns:

$$\left.\begin{aligned} .2x + .6y + .4z &= x \\ .3x + .1y + .3z &= y \\ .5x + .3y + .3z &= z \end{aligned}\right\} \Rightarrow \left.\begin{aligned} -.8x + .6y + .4z &= 0 \\ .3x - .9y + .3z &= 0 \\ .5x + .3y - .7z &= 0 \end{aligned}\right\}$$

It can be shown, however, that for any transition matrix T, the system of equations stemming from $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ will always have more than one solution (Exercise 22). By adding the equation $x + y + z = 1$ (the sum of the entries in any state matrix of the system must equal 1) to the system, we do end up with a unique solution:

$$\left.\begin{aligned} -.8x + .6y + .4z &= 0 \\ .3x - .9y + .3z &= 0 \\ .5x + .3y - .7z &= 0 \\ x + y + z &= 1 \end{aligned}\right\} \Rightarrow \left[\begin{array}{ccc|c} -.8 & .6 & .4 & 0 \\ .3 & -.9 & .3 & 0 \\ .5 & .3 & -.7 & 0 \\ 1 & 1 & 1 & 1 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3/8 \\ 0 & 1 & 0 & 1/4 \\ 0 & 0 & 1 & 3/8 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

We see that the matrix T has a unique fixed state; namely: $S_F = \begin{bmatrix} 3/8 \\ 1/4 \\ 3/8 \end{bmatrix} = \begin{bmatrix} .3750 \\ .2500 \\ .3750 \end{bmatrix}$, which, to four decimal places, coincides with the columns of the matrix T^8 in Example 6.14 (margin).

$$T^8 = \begin{bmatrix} .3750 & .3750 & .3750 \\ .2500 & .2500 & .2500 \\ .3750 & .3750 & .3750 \end{bmatrix}$$

The above rather surprising result, as you will soon see in Theorem 6.26, actually holds for the following important class of Markov chains:

DEFINITION 6.12

REGULAR MARKOV CHAIN

A Markov chain with transition matrix T is said to be **regular** if T^k consists solely of positive entries for some integer k. The transition matrix of a regular Markov chain is said to be a **regular transition matrix**.

For example:

$$T = \begin{bmatrix} 0 & .3 & .2 \\ .8 & .4 & .5 \\ .2 & .3 & .3 \end{bmatrix}$$

is regular, since:

$$T^2 = \begin{bmatrix} .28 & .18 & .21 \\ .42 & .55 & .51 \\ .30 & .27 & .28 \end{bmatrix}$$

Note that it is possible to eventually go from any state to any other state in a regular Markov chain (see Theorem 6.24).

In other words, A^{T} is that matrix obtained by interchanging the rows and columns of A. For example:

$$\text{If } A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 4 & 5 \end{bmatrix},\ A^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \\ 3 & 5 \end{bmatrix}$$

DEFINITION 6.13
TRANSPOSE

The **transpose** of a matrix $A = [a_{ij}] \in M_{m \times n}$ is the matrix $A^{\mathrm{T}} = [\bar{a}_{ij}] \in M_{n \times m}$, where $\bar{a}_{ij} = a_{ji}$.

The following results will be called upon within the proof of Theorem 6.25 below:

A-1: If $A \in M_{m \times n}$ and $B \in M_{n \times r}$, then $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{T}$. [Exercise 19(f), page 161.]
A-2: If λ is an eigenvalue of $A \in M_{n \times n}$ then λ is also an eigenvalue of A^{T}. (Exercise 52, page 231.)

THEOREM 6.25

$\lambda = 1$ is an eigenvalue of every transitional matrix $T \in M_{n \times n}$.

PROOF: Let $\boldsymbol{w} \in \Re^n$ be the n-tuple with every entry equal to 1. Being a transition matrix, the columns of $T = [t_{ij}]$ sum to 1. Hence:

$$\boldsymbol{w}T = (1, 1, \dots 1)\begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1n} \\ t_{21} & t_{22} & \cdots & t_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ t_{n1} & t_{n2} & \cdots & t_{nn} \end{bmatrix} = (1, 1, \dots 1) = \boldsymbol{w}$$

$$t_{11} + t_{21} + \dots + t_{n1} = 1$$

Note that $\boldsymbol{w}^T$ is an eigenvector of the transpose of T, and not necessarily of T.

Taking the transpose of $\boldsymbol{w}T = \boldsymbol{w}$, we have $T^{\mathrm{T}}\boldsymbol{w}^{\mathrm{T}} = \boldsymbol{w}^{\mathrm{T}}$ (**A-1**), and this tells us that $\boldsymbol{w}^{T}$ is an eigenvector of T^{T} corresponding to the eigenvalue $\lambda = 1$. Applying **A-2** we conclude that $\lambda = 1$ is also an eigenvalue of T.

In a sense, independently of its initial state:

> The fixed state of a regular transition matrix is also the final state of the matrix

THEOREM 6.26
FUNDAMENTAL THEOREM OF REGULAR MARKOV CHAINS

Every regular transition matrix $T \in M_{n \times n}$ has a unique fixed state vector $S_F = \begin{bmatrix} r_1 \\ r_2 \\ \\ r_n \end{bmatrix}$, and:

$$\lim_{s \to \infty} T^s = \begin{bmatrix} r_1 & r_1 & \dots & r_1 \\ r_2 & r_2 & \dots & r_2 \\ & & & \\ r_n & r_n & \dots & r_n \end{bmatrix}$$

(each column of the matrix T^s approaches S_F as s increases).

Proof: Assume that $T = [t_{ij}]$ consists of positive entries. (You are invited, in Exercise 27, to establish the result under the assumption that T^k consists solely of positive entries for some integer $k > 1$.)

That "(*s*)" in $M_i^{(s)}$ is not an exponent; it is there to indicate that we are considering the matrix T^s

Consider the matrix $T^s = [c_{ij}^{(s)}]$, which must also consist of positive entries. Let $M_i^{(s)} = c_{ij_M}^{(s)}$ and $m_i^{(s)} = c_{ij_m}^{(s)}$ denote the largest and smallest entry in the i^{th} row of T^s. We will show that $\lim\limits_{s \to \infty} (M_i^{(s)} - m_i^{(s)}) = 0$. This will tell us that all entries in the i^{th} row of $\lim\limits_{s \to \infty} T^s$ are equal, which is the same as saying that the columns of $\lim\limits_{s \to \infty} T^s$ are all equal.

From $T^{s+1} = [c_{ij}^{(s+1)}] = T^s T = [c_{ij}^{(s)}][t_{ij}]$ we have:

$$c_{ij}^{(s+1)} = \sum_{\alpha = 1}^{n} c_{i\alpha}^{(s)} t_{\alpha j} = c_{ij_m}^{(s)} t_{j_m j} + \sum_{\alpha \neq j_m}^{n} c_{i\alpha}^{(s)} t_{\alpha j}$$

$$= m_i^{(s)} t_{j_m j} + \sum_{\alpha \neq j_m}^{n} c_{i\alpha}^{(s)} t_{\alpha j}$$

$$\leq m_i^{(s)} t_{j_m j} + M_i^{(s)} \sum_{\alpha \neq j_m}^{n} t_{\alpha j}$$

The entries in the j^{th} column of the transition matrix T sum to 1:

$$= m_i^{(s)} t_{j_m j} + M_i^{(s)} (1 - t_{j_m j})$$

We have shown that for **every** entry $c_{ij}^{(s+1)}$ in the i^{th} row of T^{s+1}:

$$c_{ij}^{(s+1)} \leq m_i^{(s)} t_{j_m j} + M_i^{(s)} (1 - t_{j_m j})$$

In particular, for the largest entry in that row we have:

$$M_i^{(s+1)} \leq m_i^{(s)} t_{j_m j} + M_i^{(s)} (1 - t_{j_m j})$$

A similar argument (Exercise 26) can be used to show that for the smallest entry $m_i^{(s+1)}$ in the i^{th} row of T^{s+1} we have:

$$m_i^{(s+1)} \geq M_i^{(s)} t_{j_M j} + m_i^{(s)} (1 - t_{j_M j})$$

Consequently:

$$M_i^{(s+1)} - m_i^{(s+1)}$$

$$\leq m_i^{(s)} t_{j_m j} + M_i^{(s)} (1 - t_{j_m j}) - [M_i^{(s)} t_{j_M j} + m_i^{(s)} (1 - t_{j_M j})]$$

$$= (M_i^{(s)} - m_i^{(s)}) (\mathbf{1} - t_{j_m j} - t_{j_M j})$$

Let t be the smallest entry in T. Since T consists of positive entries, and since the entries in every column of T sums to 1, we have $0 < t \le \frac{1}{n}$, and, in particular that $(1 - t_{j_m j} - t_{j_M j}) \le (1 - 2t)$. Hence:

$$M_i^{(s+1)} - m_i^{(s+1)} \le (M_i^{(s)} - m_i^{(s)})(1 - 2t)$$

Leading us to:

$$\begin{aligned} M_i^{(s)} - m_i^{(s)} &\le (1 - 2t)(M_i^{(s-1)} - m_i^{(s-1)}) \\ &\le (1 - 2t)^2(M_i^{(s-2)} - m_i^{(s-2)}) \\ &\ \vdots \\ &\le (1 - 2t)^{s-1}(M_i - m_i) \end{aligned}$$

Since $0 \le (1 - 2t) < 1$, $(1 - 2t)^{s-1} \to 0$ as $s \to \infty$, and this tells us that the elements in the i^{th} row of the matrix T^s must get arbitrarily close to each other as s tends to infinity. In turn, the columns of T^s must all tend to a common vector $\begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix}$. We complete the proof by showing that $S_F = \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix}$ is the unique fixed state of T:

Employing Theorem 6.25, we start with an eigenvector $\boldsymbol{X} = \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix}$ of T, of eigenvalue 1. Since $T\boldsymbol{X} = \boldsymbol{X}$, $T^s\boldsymbol{X} = \boldsymbol{X}$ for all k. Hence:

$$\lim_{s \to \infty} T^s\boldsymbol{X} = \begin{bmatrix} r_1 & r_1 \cdots r_1 \\ r_2 & r_2 \cdots r_2 \\ \\ r_n & r_n \cdots r_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix}$$

From Theorem 5.4, page 156:

$$\begin{bmatrix} r_1 & r_1 \cdots r_1 \\ r_2 & r_2 \cdots r_2 \\ \\ r_n & r_n \cdots r_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} r_1 \\ r_2 \\ \\ r_n \end{bmatrix} + x_2 \begin{bmatrix} r_1 \\ r_2 \\ \\ r_n \end{bmatrix} + \ldots + x_1 \begin{bmatrix} r_1 \\ r_2 \\ \\ r_n \end{bmatrix} = \left(\sum_{i=1}^{n} x_i \right) \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix}$$

Since $T(X) = X$:

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \left(\sum_{i=1}^{n} x_i \right) \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix} \qquad \textbf{(**)}$$

Since $\boldsymbol{X} \neq 0$ (it is an eigenvector), $c = \sum_{i=1}^{n} x_i \neq 0$, and dividing both sides of (**) by c brings us to:

$$S_F = \frac{1}{c} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix}$$

We then have:

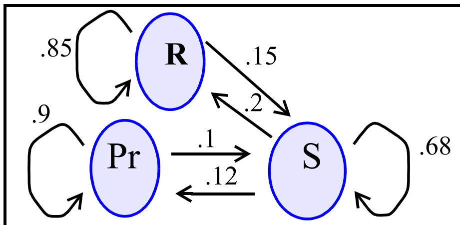
$$T \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix} = T \left(\frac{1}{c} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} \right) = \frac{1}{c} T \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \frac{1}{c} \begin{bmatrix} x_1 \\ x_2 \\ \\ x_n \end{bmatrix} = \begin{bmatrix} \boldsymbol{r_1} \\ \boldsymbol{r_2} \\ \\ \boldsymbol{r_n} \end{bmatrix}$$

The above argument also establishes the uniqueness of the fixed state vector, for if $\boldsymbol{X}$ is to be a (fixed) state vector, then c must equal one.

EXAMPLE 6.18

An automobile insurance company classifies its customers as **Pr**eferred, **S**atisfactory, or **R**isk. Each year, 10% of those in the Preferred category are downgraded to Satisfactory, while 12% of those in the Satisfactory category move to Preferred. Twenty percent of Satisfactory drop to Risk, while 15% of Risk goes to Satisfactory. No customer is moved more than one slot in either direction in a single year. Find the fixed state of the system.

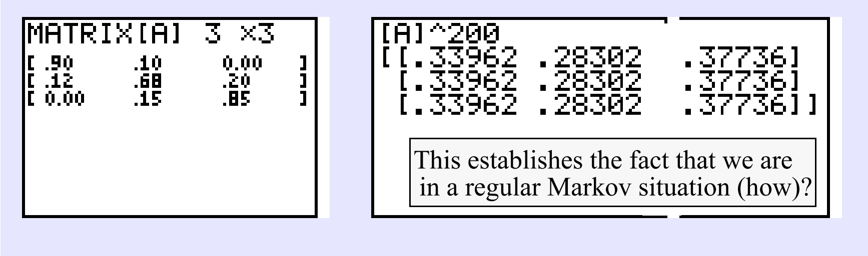
SOLUTION:



$$T = \begin{array}{c} \begin{array}{ccc} \text{Pr} & \text{S} & \text{R} \end{array} \\ \begin{bmatrix} .9 & .12 & 0 \\ .1 & .68 & .15 \\ 0 & .2 & .85 \end{bmatrix} \end{array} \begin{array}{l} \\ \text{Pr} \\ \text{S} \\ \text{R} \end{array}$$

How large is large enough? If the rows look different, then take a higher power.

The easiest way to go, is to take a “large” power of the transition matrix, and let any of its rows represent an approximation for the fixed state matrix of the regular Markov process:



```
MATRIX[A]  3 ×3
[ .90      .10      0.00     ]
[ .12      .68      .20      ]
[ 0.00     .15      .85      ]
```

```
[A]^200
[[.33962 .28302  .37736]
 [.33962 .28302  .37736]
 [.33962 .28302  .37736]]
```

This establishes the fact that we are in a regular Markov situation (how)?

We conclude that roughly 34% of the company’s clients will (eventually) fall in its Preferred category; 28% in its Satisfactory category; and 38% in its Risk category. But that is but an approximation, for:

$$\begin{bmatrix} .9 & .12 & 0 \\ .1 & .68 & .15 \\ 0 & .2 & .85 \end{bmatrix}\begin{bmatrix} .34 \\ .28 \\ .38 \end{bmatrix} = \begin{bmatrix} .3396 \\ .2814 \\ .3790 \end{bmatrix}$$

You can, however, find the exact steady state by the method of Example 6.17:

$$\begin{bmatrix} .9 & .12 & 0 \\ .1 & .68 & .15 \\ 0 & .2 & .85 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \Rightarrow \left.\begin{aligned} .9x + .12y + 0z &= x \\ .1x + .68y + .15z &= y \\ 0x + .2y + .85z &= z \end{aligned}\right\}$$

$$\left.\begin{aligned} -.1x + .12y + 0z &= 0 \\ .1x - .32 + .15z &= 0 \\ 0x + .2y - .15 &= 0 \\ x + y + z &= 1 \end{aligned}\right\} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & \frac{18}{53} \\ 0 & 1 & 0 & \frac{15}{53} \\ 0 & 0 & 1 & \frac{20}{53} \\ 0 & 0 & 0 & 0 \end{array}\right]$$

see solution of Example 6.15.

We found $\left(\frac{18}{53}, \frac{15}{53}, \frac{20}{53}\right)$ to be the (exact) steady state of the given Markov chain; telling us that the longer the process, the closer it will be that $\frac{18}{53}\%$ of the customers, for example, will be in the preferred category.

CHECK YOUR UNDERSTANDING 6.21

The transition matrix T below represents the probabilities that an individual that voted the **D**emocratic, **R**epublican, or **G**reen party ticket in the last election will vote **D**, **R**, or **G** in the next election.

$$T: \begin{array}{c} \\ \\ \\ \end{array}\begin{array}{ccc} \mathbf{D} & \mathbf{R} & \mathbf{G} \\ \end{array}$$

$$T: \begin{array}{cccc} \mathbf{D} & \mathbf{R} & \mathbf{G} & \\ .73 & .32 & .09 & \mathbf{D} \\ .21 & .61 & .04 & \mathbf{R} \\ .06 & .07 & .87 & \mathbf{G} \end{array}$$

Determine the eventual percentage of the population in each of the three category.

Answer: Approximately 41%, 26%, 33% of the population, will vote democratic, republican, green, respectively.

EXERCISES

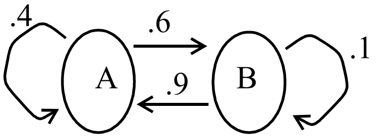
Exercises 1-6. Indicate whether or not the given matrix represents a transition matrix for a Markov Process. If not, state why not. If so, indicate whether or not the given transition matrix is regular.

1. $\begin{bmatrix} .2 & .1 \\ .8 & .9 \end{bmatrix}$ 2. $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ 3. $\begin{bmatrix} 0 & .1 \\ -1 & .9 \end{bmatrix}$

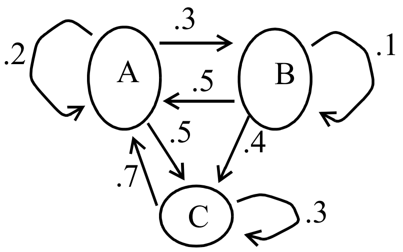
4. $\begin{bmatrix} .2 & .4 & .1 \\ .7 & 0 & .3 \\ .1 & .6 & .6 \end{bmatrix}$ 5. $\begin{bmatrix} 1 & .4 & .1 \\ 0 & 0 & .3 \\ 0 & .6 & .6 \end{bmatrix}$ 6. $\begin{bmatrix} 0 & .3 & .4 \\ .5 & .3 & .6 \\ .5 & .1 & 0 \end{bmatrix}$

Exercises 7-8. Determine the transition matrix associated with the given transition diagram.

7.



8.



Exercises 9-11. Determine a transition diagram associated with the given transition matrix.

9. $\begin{array}{cc} \begin{array}{cc} A & B \end{array} & \\ \begin{bmatrix} .3 & .4 \\ .7 & .6 \end{bmatrix} & \begin{array}{c} A \\ B \end{array} \end{array}$

10. $\begin{array}{cc} \begin{array}{ccc} A & B & C \end{array} & \\ \begin{bmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} & \begin{array}{c} A \\ B \\ C \end{array} \end{array}$

11. $\begin{array}{cc} \begin{array}{cccc} A & B & C & D \end{array} & \\ \begin{bmatrix} 0 & 1 & .2 & .6 \\ .5 & 0 & .5 & .4 \\ 0 & 0 & 0 & 0 \\ .5 & 0 & .3 & 0 \end{bmatrix} & \begin{array}{c} A \\ B \\ C \\ D \end{array} \end{array}$

12. Determine the probability of ending up at states A and B after two steps of the Markov chain associated with the transition matrix in Exercise 9, given that you are initially in state:

(a) A (b) B

13. Determine the probability of ending up at states A, B and C after two steps of the Markov chain associated with the transition matrix in Exercise 10, given that you are initially in state:

(a) A (b) B (c) C

14. Determine the probability of ending up at states A, B, C and D after two steps of the Markov chain associated with the transition matrix in Exercise 11, given that you are initially in state:

(a) A (b) B (c) C (d) D

Exercises 15-20. (a) Proceed as in Example 6.15 to find the stationary state matrix of the given regular transition matrix.

(b) Use Theorem 6.26 and a graphing utility to check your answer in (a).

15. $\begin{bmatrix} .7 & .2 \\ .3 & .8 \end{bmatrix}$ 16. $\begin{bmatrix} .1 & .6 \\ .9 & .4 \end{bmatrix}$ 17. $\begin{bmatrix} .5 & .3 \\ .5 & .7 \end{bmatrix}$

18. $\begin{bmatrix} .8 & .5 & 0 \\ .2 & .1 & .6 \\ .0 & .4 & .4 \end{bmatrix}$ 19. $\begin{bmatrix} .5 & .3 & 0 \\ .1 & .7 & .6 \\ .4 & 0 & .4 \end{bmatrix}$ 20. $\begin{bmatrix} .6 & .2 & .1 \\ .3 & .5 & .5 \\ .1 & .3 & .4 \end{bmatrix}$

21. Show that the matrix $A = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ is not a regular matrix, by:

    (a) Demonstrating that for each k, A^k will contain a row that does not consist solely of positive entries.

    (b) Showing that A does not have a fixed state vector.

22. Show that for any transition matrix T, the system of equations stemming from $T\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ has infinitely many solutions.

    Suggestion: Use the fact that the sum of the elements in each column of T sum to 1.

23. Let $T \in M_{n \times n}$ be a regular transition matrix. Prove that $(x - 1)$ is a factor of the characteristic polynomial of T.

24. Show that if the entries in each column of $A \in M_{n \times n}$ sum to k, then k is an eigenvalue of A.

25. Referring to the proof of Theorem 6.26, show that:
$$m_i^{(k+1)} \geq M_i^{(k)} b_{j_M j} + m_i^{(k)}(1 - b_{j_M j})$$

26. Establish Theorem 6.26 for an arbitrary transitional matrix T.

    Suggestion: Let r be such that $A = T^r$ consists of positive entries, and consider the matrix TA.

27. Prove that if λ is any eigenvalue of a regular transition matrix, then $|\lambda| \leq 1$.

28. Show that if λ is an eigenvalue of a regular transition matrix, then $\lambda \neq -1$.

29. **(Rapid Transit)** A study has shown that in a certain city, if a daily (including Saturday and Sunday) commuter uses rapid transit on a given day, then he will do so again on his next commute with probability 0.85, and that a commuter who does not use rapid transit will do so with probability 0.3. Assume that on Monday 57% of the commuters use rapid transit. Determine, to two decimal places, the probability that a commuter will use rapid transit on:

    (a) Tuesday (b) Wednesday (c) Sunday

30. **(Dental Plans)** A company offers its employees 3 different dental plans: A, B, and C. Last year, 550 employees were in plan A, 340 in plan B, and 260 were in plan C. This year, there are 500 employees in plan A, 360 in plan B, and 290 in plan C. Assuming that the number of employees in the company remains at 1150, and that the current trend continues, determine the number of employees in each of the three plans:

    (a) A year from now. (b) Two years from now.

    (c) In 4 years (Suggestion: use the square of the 2-year matrix).

    (d) In 8 years (Suggestion: use the square of the 4-year matrix).

    (e) In 12 years (Suggestion: use the product of 4-year and the 8-year matrix).

31. **(Campus Life)** The following transition matrix gives the probabilities that a student living in the **D**orms, at **H**ome, or **O**ff-campus (but not at home), will be living in the Dorms, at Home, or Off-campus (but not at home) next year (assume that all freshmen will graduate from the college in four years).

$$\begin{array}{cc} \begin{array}{ccc}\mathbf{D} & \mathbf{H} & \mathbf{O}\end{array} & \\ \left[\begin{array}{ccc} .62 & .23 & .25 \\ .11 & .64 & .09 \\ .27 & .13 & .66 \end{array}\right] & \begin{array}{c}\mathbf{D}\\ \mathbf{H}\\ \mathbf{O}\end{array} \end{array}$$

Currently, 55%, 24%, and 21% of the freshman class are living in the Dorms, at Home, and Off-campus (but not at home), respectively. Determine (to two decimal places) the probability that a current freshman will, three years from now, be living in the:

(a) Dorms (b) Home (c) Off-campus.

32. (**Higher Learning**) The transition matrix below represents the probabilities that a female child will receive a **D**octorate, a **M**asters, or a **B**achelors (terminal degree), or **N**o degree; given that her mother received a **D**, **M**, **B** (terminal degree), or **N**o degree.

$$\begin{array}{ccc} \text{mother} & & \\ \begin{array}{cccc}\mathbf{D} & \mathbf{M} & \mathbf{B} & \mathbf{N}\end{array} & & \\ \left[\begin{array}{cccc} .31 & .24 & .11 & .06 \\ .25 & .26 & .09 & .05 \\ .37 & .42 & .52 & .49 \\ .07 & .08 & .28 & .40 \end{array}\right] & \begin{array}{c}\mathbf{D}\\ \mathbf{M}\\ \mathbf{B}\\ \mathbf{N}\end{array} & \text{daughter} \end{array}$$

Given the initial state matrix $S_0 = [.05\quad .09\quad .39\quad .47]$ (in column form), determine the probability that:

(a) A granddaughter will receive a Bachelors degree.

(b) A great granddaughter will earn a Doctorate.

(c) A fifth generation daughter will receive no degree.

33. **(HMO Plans)** A company offers its employees 5 different HMO health plans: A, B, C, D, and E. An employee can switch plans in January of each year, resulting in the following transition matrix:

$$\begin{array}{cc}
\begin{array}{ccccc} & & \text{this year} & & \\ \mathbf{A} & \mathbf{B} & \mathbf{C} & \mathbf{D} & \mathbf{E} \end{array} & \\
\left[\begin{array}{ccccc}
.54 & .13 & .08 & .10 & .06 \\
.11 & .61 & .17 & .12 & .18 \\
.17 & .10 & .56 & .17 & .15 \\
.06 & .05 & .08 & .49 & .10 \\
.12 & .11 & .11 & .12 & .51
\end{array}\right] & \begin{array}{c} \mathbf{A} \\ \mathbf{B} \\ \mathbf{C} \\ \mathbf{D} \\ \mathbf{E} \end{array} \ \text{next year}
\end{array}$$

Given the initial state matrix $S_0 = [.11 \quad .20 \quad .31 \quad .14 \quad .24]$ (in column form), determine, to three decimal places, the probability that:

(a) An employee will chose plan B in the next enrollment period.

(b) An employee will chose plan B two enrollment periods from now.

(c) An employee will chose plan B three enrollment periods from now.

(d) Determine to 5 decimal places, the fixed state of the system.

(e) Repeat (a) through (d) with initial state matrix $S_0 = [.24 \quad .31 \quad .0 \quad .26 \quad .19]$

34. **(Mouse in Maze)** On Monday, a mouse is placed in a maze consisting of paths A and B. At the end of path A is a cheese treat, and at the end of path B there is bread. Experience has shown that if the mouse takes path A, then there is a 0.9 probability that it will take path A again, on the following day. If it takes path B, then there is a 0.6 probability that it will take that path again, the next day. The mouse takes path B on Monday. Determine the probability that the mouse will take path A on:

(a) Tuesday (b) Wednesday (c) Sunday

(d) Answer parts (a), (b), and (c), under the assumption that the mouse takes path A on Monday.

(e) Show that the transition matrix is regular, and then proceed as in Example 6.14 to determine the exact stationary state of that matrix.

(f) Indicate the long-term state of the system (the probability that the mouse will take path A and the probability that the mouse will take path B, at the n^{th} step of the process, for n "large").

35. **(Cities, Suburbs, and Country)** Within the period of a year, 2% of a population currently residing in cities will move to the suburbs, while 2% of them will move to the country. 4% of those living in the suburbs will move to the cities, while 3% of them will move to the country. One percent of the country folks will move to the cities, while 2% of them will go to the suburbs. Currently, 65% of the population are in cities, and 20% are in the suburbs. Determine, to two decimal places, the percentage of city dwellers:

(a) Next year. (b) Two years from now. (c) Four years from now.

(d) Answer parts (a), (b), and (c), under the assumption that 50% of the population are in cities, and 35% are in the suburbs.

(e) Determine to 5 decimal places, the fixed state of the system.

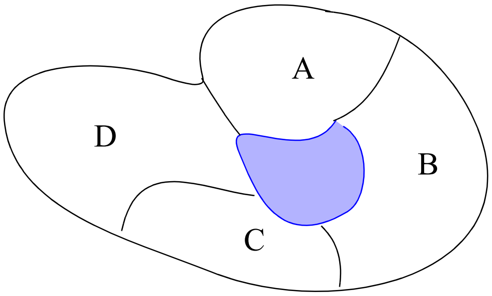
36. **(Crop Rotation)** A farmer rotates a field between crops of beans, potatoes and carrots. If she grows beans this year, then next year she will grow potatoes or carrots, each with 0.5 probability. If she grows carrots, then she will grow beans with probability 0.2, potatoes with probability 0.5 (and carrots with probability 0.3). If she grows potatoes, then she will grow beans with probability 0.5, and potatoes with probability 0.25. If she grows beans this year, what is the probability that she will grow beans again:

(a) Next year? (b) Two years from now? (c) Three years from now?

(d) Answer parts (a), (b), and (c) under the assumption that she grows potatoes this year.

(e) Determine to 5 decimal places, the fixed state of the system.

37. **(Wolf Pack)** A wolf pack hunts on one of four regions: A, B, C, and D:



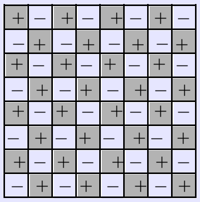
If the pack hunts in any given region one day, then it is as likely to hunt there again the next day as it is for it to hunt in either of its neighboring regions. On Monday, it hunted in region A.

(a) Determine, to two decimal places, the probability that the pack will hunt in Region B on Tuesday.

(b) Determine, to two decimal places, the probability that the pack will hunt in Region B on Sunday.

(c) Determine the fixed state of the system.

CHAPTER SUMMARY

DETERMINANTS	The **determinant** of an $n \times n$ matrix A, denoted det(A), is defined inductively as follows: For a 1×1 matrix $A = [a_{11}]$, $\det(A) = a_{11}$. For a $n \times n$ matrix A, with $n > 1$, let A_{ij} denote the $(n-1) \times (n-1)$ matrix obtained by deleting the i^{th} row and j^{th} column of the matrix A; Then: $\det(A) = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} \det(A_{1j})$
Laplace's Theorem 	For $A \in M_{n \times n}$ and any $1 \le i \le n$: $\det(A) \underset{\uparrow}{=} \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$ and $\det A \underset{\uparrow}{=} \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij})$ Expanding along the i^{th} row Expanding along the j^{th} row
Determinants of diagonal and upper triangular matrices	The determinant of a diagonal matrix or of an upper triangular matrix is the product of the entries in its diagonal.
Determinants and row operations	(a) If two rows of $A \in M_{n \times n}$ are interchanged, then the determinant of the resulting matrix is $-\det(A)$. (b) If one row of A is multiplied by a constant c, then the determinant of the resulting matrix is $c[\det(A)]$. (c) If a multiple of one row of A is added to another row of A, then the determinant of the resulting matrix is $\det(A)$
Invertibility	A matrix $A \in M_{n \times n}$ is invertible if and only if $\det(A) \neq 0$.
Product Theorem	For $A, B \in M_{n \times n}$: $\det(AB) = \det(A)\det(B)$

EIGENVALUE AND EIGENVECTOR	An **eigenvalue** of a linear operator $T: V \to V$ is a scalar $\lambda \in \boldsymbol{R}$ for which there exists a nonzero vector $\boldsymbol{v} \in V$ such that: $$T(\boldsymbol{v}) = \lambda \boldsymbol{v}$$ Any such $\boldsymbol{v}$ is then said to be an **eigenvector** corresponding to the eigenvalue λ. An **eigenvalue** of a matrix $A \in M_{n \times n}$ is a scalar $\lambda \in \boldsymbol{R}$ for which there exists a nonzero vector $\boldsymbol{X} \in \boldsymbol{R}^n$ such that: $$A(\boldsymbol{X}) = \lambda \boldsymbol{X}$$ Any such $\boldsymbol{X}$ is then said to be an **eigenvector** corresponding to the eigenvalue λ.
EIGENSPACE	The **eigenspace** of an eigenvalue λ of a matrix $A \in M_{n \times n}$ is denoted by $E[\lambda]$ and is given by: $$E[\lambda] = \text{null}(A - \lambda I)$$ The **eigenspace** of an eigenvalue λ of a linear operator $T: V \to V$ is denoted by $E[\lambda]$ and is given by: $$E[\lambda] = \ker(T - \lambda I)$$
CHARACTERISTIC POLYNOMIAL AND CHARACTERISTIC EQUATION	For $A \in M_{n \times n}$, the n-degree polynomial $\det(A - \lambda I)$ is said to be the **characteristic polynomial** of A, and $\det(A - \lambda I) = 0$ is said to be the **characteristic equation** of A. For $T: V \to V$ a linear operator on a vector space V of dimension n, the n-degree polynomial $\det([T]_\beta - \lambda I)$, where β is a basis for V, is said to be the **characteristic polynomial** of T, and $\det([T]_\beta - \lambda I) = 0$ is said to be the **characteristic equation** of T.
Finding Eigenvalues	The eigenvalues of $A \in M_{n \times n}$ are the solutions of the characteristic equation $\det(A - \lambda I) = 0$. The eigenvalues of a linear operator $T: V \to V$ on a vector space of dimension n are the eigenvalues of the matrix $[T]_\beta \in M_{n \times n}$, where β is any basis for V.
DIAGONAL MATRIX	A **diagonal matrix** is a square matrix $A = [a_{ij}]$ with $a_{ij} = 0$ for $i \neq j$.

DIAGONALIZABLE MATRICES AND LINEAR OPERATORS	A matrix $A \in M_{n \times n}$ is **diagonalizable** if A is similar to a diagonal matrix. A linear operator $T\colon V \to V$ on a finite dimensional vector space V is said to be **diagonalizable** if there exists a basis β for which $[T]_\beta$ is a diagonal matrix.
Diagonalization Theorem	Let $T\colon V \to V$ be a linear operator on a finite dimensional vector space. The following are equivalent: (i) T is diagonalizable. (ii) $[T]_\beta$ is a diagonalizable matrix, for any basis β of V. (iii)There exists a basis for V consisting of eigenvectors of T.
Eigenvectors corresponding to different eigenvalues are linearly independent.	If $\lambda_1, \lambda_2, \ldots, \lambda_m$ are distinct eigenvalues of a linear operator $T\colon V \to V$, and if $\boldsymbol{v}_i$ is an eigenvector corresponding to λ_i, for $1 \le i \le m$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ is a linearly independent set.
The union of linearly independent subsets of different eigenspaces is again a linearly independent set.	Let $\lambda_1, \lambda_2, \ldots, \lambda_m$ be distinct eigenvalues of a linear operator $T\colon V \to V$, and let $S_i = \{v_{i1}, v_{i1}, \ldots, v_{ir_i}\}$ be a linearly independent subset of $E(\lambda_i)$. Then: $$S = S_1 \cup S_2 \cup \ldots \cup S_m$$ is a linearly independent set.
ALGEBRAIC AND GEOMETRIC MULTIPLICITY OF EIGENVALUES	An eigenvalue λ_0 of a matrix $A \in M_{n \times n}$ (or of a linear operator T on a vector space of dimension n) has **algebraic multiplicity** k if $(\lambda - \lambda_0)^k$ is a factor of A's (or T's) characteristic polynomial, and $(\lambda - \lambda_0)^{k+1}$ is not. We also define the **geometric multiplicity** of λ_0 to be the dimension of $E[\lambda_0]$ (the eigenspace corresponding to λ_0)
The geometric multiplicity cannot exceed the algebraic multiplicity.	If λ_0 is an eigenvalue of $A \in M_{n \times n}$ with algebraic multiplicity m_a and geometrical multiplicity m_g, then $m_g \le m_a$.
Another Diagonalization Theorem.	Assume that the characteristic polynomial of a linear operator $T\colon V \to V$ (or of a matrix), can be factored into a product of linear factors. Then T is diagonalizable if and only if the algebraic multiplicity of each eigenvalue of T is equal to its geometric multiplicity.

Diagonalizing a Matrix	Let $A \in M_{n \times n}$ be diagonalizable. Let the columns of $P \in M_{n \times n}$ be any basis for $\boldsymbol{R}^n$ consisting of eigenvectors of A. Then $D = P^{-1}AP$ is a diagonal matrix. Moreover, the diagonal entry $d_{ii} \in D$ is the eigenvalue λ_i corresponding to the i^{th} column of P.
Power Theorem for diagonalizable matrices	Let $A \in M_{m \times m}$ be diagonalizable with $P^{-1}AP = D$. Then for any n: $$A^n = PD^nP^{-1}.$$
FIBONACCI SEQUENCE	The Fibonacci sequence is that sequence whose first two terms are 1, and with each term after the second being obtained by summing its two immediate predecessors. The k^{th} Fibonacci number is given by $\boldsymbol{\frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1-\sqrt{5}}{2}\right)^k\right]}$.
SYSTEMS OF DIFFERENTIAL EQUATIONS	The solution set of $f'(x) = af(x)$, consists of those functions of the form $f(x) = ce^{ax}$ for some constant c. Let $A \in M_{n \times n}$ be diagonalizable, and let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ be a basis for $\Re^n$ consisting entirely of eigenvectors of A with corresponding eigenvalues λ_i. Then, the general solution of: $$F'(x) = AF(x)$$ is of the form: $$c_1e^{\lambda_1 x}\boldsymbol{v}_1 + c_2e^{\lambda_2 x}\boldsymbol{v}_2 + \cdots + c_ne^{\lambda_n x}\boldsymbol{v}_n$$ for $c_1, c_2, \ldots c_n \in \Re$.
MARKOV CHAINS	**Markov chain**: When the probabilities of moving from one state of a system to another remain constant.
TRANSITION MATRIX	A **transition matrix** $T \in M_{n \times n}$ is a matrix that satisfies the following two properties: (1) T contains no negative entry. (2) The sum of the entries in each column of T equals 1.
FIXED STATE	$S_F \in \Re^n$ is a **fixed state** for a transition matrix $T \in M_{n \times n}$ if $TS_F = S_F$. If T is the transition matrix of a Markov process with initial-state matrix S_0, then the n^{th} state matrix in the chain is given by: $$S_n = T^nS_0$$
FUNDAMENTAL THEOREM OF REGULAR MARKOV CHAINS	Every regular transition matrix $T \in M_{n \times n}$ has a unique fixed state vector and each column of the matrix T^s approaches S_F as s increases.

CHAPTER 7
INNER PRODUCT SPACES

Basically, an inner product space is a vector space augmented with an additional structure, one that will enable us to generalize the familiar concepts of distance and angles in the plane to general vector spaces.

§1. DOT PRODUCT

We begin by introducing a function which assigns a real number to each pair of vectors in $\Re^n$:

DEFINITION 7.1
DOT PRODUCT

The **dot product** of $\mathbf{u} = (\mathbf{u_1, u_2, \dots, u_n})$ and $\mathbf{v} = (\mathbf{v_1, v_2, \dots, v_n})$, denoted by $\mathbf{u} \cdot \mathbf{v}$, is the real number:

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

For example:

$$(\mathbf{2, 4, -3, 1}) \cdot (\mathbf{5, 0, 7, -1}) = 2 \cdot 5 + 4 \cdot 0 + (-3) \cdot 7 + 1 \cdot (-1) = -12$$

The following four properties will lead us to the definition of an inner product space in the next section, much in the same way that the eight properties of Theorem 2.1, page 36, lead us to the definition of an abstract vector pace on page 40.

THEOREM 7.1

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \Re^n$, and $r \in \Re$. Then:

positive-definite property: (i) $\mathbf{v} \cdot \mathbf{v} \ge 0$, and $\mathbf{v} \cdot \mathbf{v} = 0$ only if $\mathbf{v} = \mathbf{0}$

commutative property: (ii) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

homogeneous property: (iii) $r\mathbf{u} \cdot \mathbf{v} = r(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot r\mathbf{v}$

distributive property: (iv) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$

PROOF: We establish (iii) and invite you to verify the remaining three properties in the exercises:

$$r\mathbf{u} \cdot \mathbf{v} = r(\mathbf{u_1, u_2, \dots, u_n}) \cdot (\mathbf{v_1, v_2, \dots, v_n})$$

scalar multiplication: $= (\mathbf{ru_1, ru_2, \dots, ru_n}) \cdot (\mathbf{v_1, v_2, \dots, v_n})$

definition 7.1: $= (ru_1)v_1 + (ru_2)v_2 + \dots + (ru_n)v_n$

associative property: $= r(u_1v_1) + r(u_2v_2) + \dots + r(u_nv_n)$

distributive property: $= r(u_1v_1 + u_2v_2 + \dots + u_nv_n)$

definition 7.1: $= r(\mathbf{u} \cdot \mathbf{v})$

CHECK YOUR UNDERSTANDING 7.1

Let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$, and $r \in \Re$. Prove that:

$$r\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot r\boldsymbol{v}$$

Answer: See page B-30.

DEFINITION 7.2

NORM IN $\Re^n$

The **norm** of a vector $\boldsymbol{v} = (\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n})$, denoted by $\|\boldsymbol{v}\|$, is given by:

$$\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$$

For $\boldsymbol{v} = (\boldsymbol{v_1}, \boldsymbol{v_2}) \in \Re^2$, $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}} = \sqrt{v_1^2 + v_2^2}$ represents the length (magnitude) of $\boldsymbol{v}$ [Figure 7.1(a)], and the same can be said about $\|\boldsymbol{v}\|$ for $\boldsymbol{v} = (\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}) \in \Re^3$ [Figure 7.1(b)].

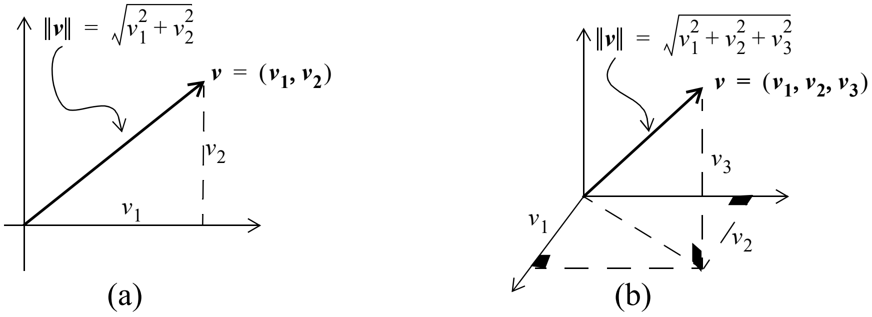


Figure 7.1

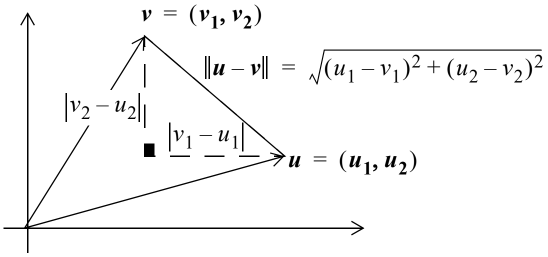
In general, for $\boldsymbol{v} \in \Re^n$:

$\|\boldsymbol{v}\|$ is defined to be the length of $\boldsymbol{v}$.

Moreover:

$\|\boldsymbol{u} - \boldsymbol{v}\|$ is defined to be the distance between $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$.

In particular, for $\boldsymbol{u} = (\boldsymbol{u_1}, \boldsymbol{u_2})$ and $\boldsymbol{v} = (\boldsymbol{v_1}, \boldsymbol{v_2})$ in $\Re^2$:



CHECK YOUR UNDERSTANDING 7.2

Prove that for $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$ and $c \in \Re$:

(a) $\|c\boldsymbol{v}\| = |c|\|\boldsymbol{v}\|$

(b) $\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - 2\boldsymbol{u}\cdot\boldsymbol{v} + \|\boldsymbol{v}\|^2$

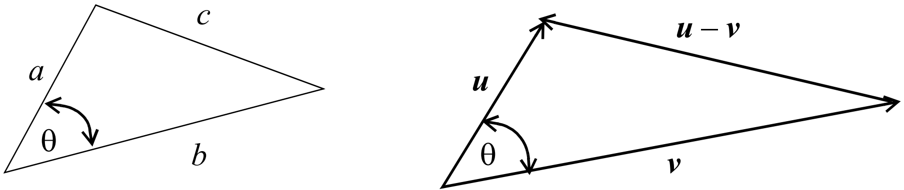
[Reminiscent of: $(a-b)^2 = a^2 - 2ab + b^2$]

Answer: See page B-30.

ANGLE BETWEEN VECTORS

Applying the law of cosines [Figure 7.2(a)] to the vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^2$ in Figure 7.2(b), we see that:

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



$c^2 = a^2 + b^2 - 2ab\cos\theta$

Law of Cosines

(a)

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

(b)

Figure 7.2

From CYU 7.2, we also have:

$$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 - 2\boldsymbol{u}\cdot\boldsymbol{v} + \|\boldsymbol{v}\|^2$$

Thus:

$$\begin{aligned}\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta &= \|\boldsymbol{u}\|^2 - 2\boldsymbol{u}\cdot\boldsymbol{v} + \|\boldsymbol{v}\|^2\\ -2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta &= -2\boldsymbol{u}\cdot\boldsymbol{v}\\ \cos\theta &= \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\end{aligned}$$

For any $-1 \le x \le 1$, $\cos^{-1}x$ is defined to be that angle $0 \le \theta \le \pi$ whose cosine is x.

see margin: $\theta = \cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$

Leading us to:

In Exercise 44 you are asked to verify that

$$\left|\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right| \le 1$$

Assuring us that:

$\cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$ exists.

DEFINITION 7.3
ANGLE BETWEEN VECTORS

The **angle** θ between two nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$ is given by:

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

EXAMPLE 7.1 Determine the angle between the vectors $\boldsymbol{u} = (\mathbf{1, 2, 0, -2})$ and $\boldsymbol{v} = (\mathbf{-1, 3, 1, 2})$.

SOLUTION:

$$\theta = \cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right) = \cos^{-1}\left(\frac{(\mathbf{1, 2, 0, -2}) \cdot (\mathbf{-1, 3, 1, 2})}{\sqrt{1+4+4}\sqrt{1+9+1+4}}\right)$$

$$= \cos^{-1}\left(\frac{1}{3\sqrt{15}}\right) \approx 85°$$

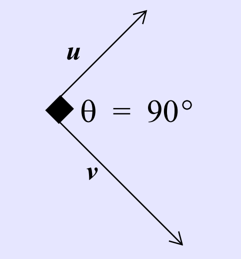
CHECK YOUR UNDERSTANDING 7.3

Determine the angle between the vectors $\boldsymbol{u} = (\mathbf{1, 2, 0})$ and $\boldsymbol{v} = (\mathbf{-1, 3, 1})$.

$\cos^{-1}\left(\frac{5}{\sqrt{55}}\right) \approx 83°$

ORTHOGONAL VECTORS IN R^n

The angle θ between the vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^2$ depicted in the adjacent figure has a measure of $90°$ ($\frac{\pi}{2}$ radians), and we say that those vectors are perpendicular or orthogonal. Appealing to Definition 7.3 we see that:



$$\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right) = 90° \quad \text{or} \quad \boldsymbol{u} \cdot \boldsymbol{v} = 0$$

(see margin)

We remind you that, for any $-1 \le x \le 1$, $\cos^{-1}x$ is that angle $0 \le \theta \le \pi$ such that $\cos\theta = x$.

So, if $\cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right) = 90°$,

then $\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} = 0$, or: $\boldsymbol{u} \cdot \boldsymbol{v} = 0$

.

Bringing us to:

DEFINITION 7.4
ORTHOGONAL VECTORS
Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\Re^n$ are **orthogonal** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

Note: The zero vector in $\Re^n$ is orthogonal to every vector in $\Re^n$.

CHECK YOUR UNDERSTANDING 7.4

Let $\boldsymbol{v} \in \Re^n$. Show that the **set** $\boldsymbol{v}^\perp$ of vectors perpendicular to $\boldsymbol{v}$:

$$\boldsymbol{v}^\perp = \{\boldsymbol{u} \in \Re^n | \boldsymbol{u} \cdot \boldsymbol{v} = 0\}$$

is a subspace of $\Re^n$.

Answer: See page B-31.

It is often useful to decompose a vector $\boldsymbol{v} \in \Re^n$ into a sum of two vectors: one parallel to a given nonzero vector $\boldsymbol{u}$, and the other perpendicular to $\boldsymbol{u}$. The parallel-vector must be of the form $c\boldsymbol{u}$ for some scalar c (see Figure 7.3).

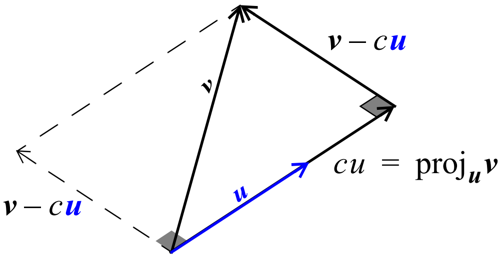


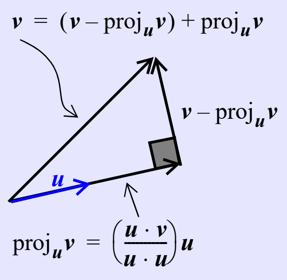
Figure 7.3

ORTHOGONAL PROJECTION

The vector $c\boldsymbol{u}$ in Figure 7.3 is said to be the **orthogonal projection of $\boldsymbol{v}$ onto $\boldsymbol{u}$** and is denoted by $\text{proj}_{\boldsymbol{u}}\boldsymbol{v}$. To determine the value of c, we note that for $\boldsymbol{v} - c\boldsymbol{u}$ to be orthogonal to $\boldsymbol{u}$, we must have:

$$\begin{aligned} \text{Theorem 7.1(iv):} \quad (\boldsymbol{v} - c\boldsymbol{u}) \cdot \boldsymbol{u} &= 0 \\ \text{Theorem 7.1(iii):} \quad \boldsymbol{v} \cdot \boldsymbol{u} - (c\boldsymbol{u}) \cdot \boldsymbol{u} &= 0 \\ \boldsymbol{v} \cdot \boldsymbol{u} - c(\boldsymbol{u} \cdot \boldsymbol{u}) &= 0 \\ c &= \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}} = \frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\|\boldsymbol{u}\|^2} \end{aligned}$$

Summarizing, we have:



THEOREM 7.2
VECTOR DECOMPOSITION

Let $\boldsymbol{v} \in \Re^n$ and let $\boldsymbol{u}$ be any nonzero vector in $\Re^n$. Then:

$$\boldsymbol{v} = (\boldsymbol{v} - \text{proj}_{\boldsymbol{u}}\boldsymbol{v}) + \text{proj}_{\boldsymbol{u}}\boldsymbol{v}$$

where:

$$\text{proj}_{\boldsymbol{u}}\boldsymbol{v} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right)\boldsymbol{u} \text{ and } (\boldsymbol{v} - \text{proj}_{\boldsymbol{u}}\boldsymbol{v}) \cdot \text{proj}_{\boldsymbol{u}}\boldsymbol{v} = 0$$

The vector $\text{proj}_{\boldsymbol{u}}\boldsymbol{v}$ is said to be the **vector component of $\boldsymbol{v}$ along $\boldsymbol{u}$**, and the vector $\boldsymbol{v} - \text{proj}_{\boldsymbol{u}}\boldsymbol{v}$ is said to be the **vector component of $\boldsymbol{v}$ orthogonal to $\boldsymbol{u}$**.

EXAMPLE 7.2 Express the vector $(\mathbf{2, 1, -3})$ as a sum of a vector parallel to $(\mathbf{1, 4, 0})$ and a vector orthogonal to $(\mathbf{1, 4, 0})$.

SOLUTION: For $\boldsymbol{v} = (\mathbf{2, 1, -3})$ and $\boldsymbol{u} = (\mathbf{1, 4, 0})$ we have:

$$\text{proj}_{\boldsymbol{u}}\boldsymbol{v} = \left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right)\boldsymbol{u} = \left[\frac{(\mathbf{1, 4, 0}) \cdot (\mathbf{2, 1, -3})}{(\mathbf{1, 4, 0}) \cdot (\mathbf{1, 4, 0})}\right](\mathbf{1, 4, 0}) = \frac{6}{17}(\mathbf{1, 4, 0})$$

$$\text{and} \quad \boldsymbol{v} - \text{proj}_{\boldsymbol{u}}\boldsymbol{v} = (\mathbf{2, 1, -3}) - \left(\frac{\mathbf{6}}{\mathbf{17}}, \frac{\mathbf{24}}{\mathbf{17}}, \mathbf{0}\right) = \left(\frac{\mathbf{28}}{\mathbf{17}}, -\frac{\mathbf{7}}{\mathbf{17}}, \mathbf{-3}\right)$$

Check: $\left(\frac{28}{17}, -\frac{7}{17}, -3\right) + \left(\frac{6}{17}, \frac{24}{17}, 0\right) = (\mathbf{2, 1, -3}) = \boldsymbol{v}$

and $\left(\frac{28}{17}, -\frac{7}{17}, -3\right) \cdot \left(\frac{6}{17}, \frac{24}{17}, 0\right) = \left(\frac{28}{17}\right)\left(\frac{6}{17}\right) + \left(-\frac{7}{17}\right)\left(\frac{24}{17}\right) = 0$

CHECK YOUR UNDERSTANDING 7.5

Express the vector $(\mathbf{3, 0, 1, -1})$ as the sum of a vector parallel to $(\mathbf{0, 2, 4, 1})$ and a vector orthogonal to $(\mathbf{0, 2, 4, 1})$.

Answer: See page B-31.

EXAMPLE 7.3 Find the distance from the point $P = (3, 1, 3)$ to the line L in $\Re^3$ which passes through the points $(1, 0, 2)$ and $(3, 1, 6)$.

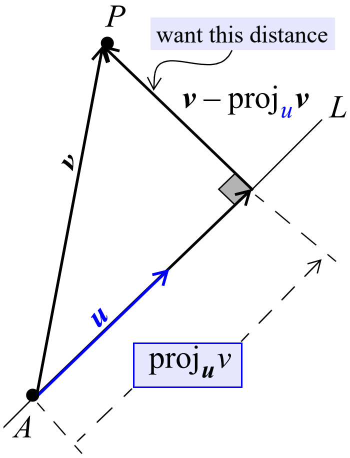
SOLUTION: We first find a direction vector for the given line (see Theorem 2.17, page 70):

$$\boldsymbol{u} = (\mathbf{3, 1, 6}) - (\mathbf{1, 0, 2}) = (\mathbf{2, 1, 4}).$$

Choosing the point $A = (1, 0, 2)$ on L we determine the vector $\boldsymbol{v}$ from A to P:

$$\boldsymbol{v} = (\mathbf{3, 1, 3}) - (\mathbf{1, 0, 2}) = (\mathbf{2, 1, 1})$$

Applying Theorem 7.2, we have:



$$\begin{aligned}\text{proj}_{\boldsymbol{u}}\boldsymbol{v} &= \left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\boldsymbol{u}\cdot\boldsymbol{u}}\right)\boldsymbol{u}\\ &= \left(\frac{(\mathbf{2, 1, 4})\cdot(\mathbf{2, 1, 1})}{(\mathbf{2, 1, 4})\cdot(\mathbf{2, 1, 4})}\right)(\mathbf{2, 1, 4})\\ &= \frac{9}{21}(\mathbf{2, 1, 4}) = \left(\frac{\mathbf{6}}{\mathbf{7}}, \frac{\mathbf{3}}{\mathbf{7}}, \frac{\mathbf{12}}{\mathbf{7}}\right)\end{aligned}$$

Thus:

$$\begin{aligned}\|\boldsymbol{v} - \text{proj}_{u}\boldsymbol{v}\| &= \left\|(\mathbf{2, 1, 1}) - \left(\frac{\mathbf{6}}{\mathbf{7}}, \frac{\mathbf{3}}{\mathbf{7}}, \frac{\mathbf{12}}{\mathbf{7}}\right)\right\|\\ &= \left\|\left(\frac{\mathbf{8}}{\mathbf{7}}, \frac{\mathbf{4}}{\mathbf{7}}, -\frac{\mathbf{5}}{\mathbf{7}}\right)\right\| \underset{\substack{\uparrow\\ \text{CYU 7.2(a)}}}{=} \frac{1}{7}\|(\mathbf{8, 4, -5})\| = \frac{1}{7}\sqrt{64+16+25} = \frac{\sqrt{105}}{7}\end{aligned}$$

CHECK YOUR UNDERSTANDING 7.6

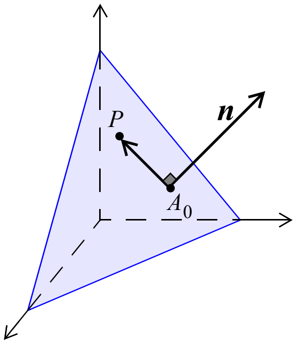
(a) Find the distance from the point $P = (2, 5)$ to the line L in $\Re^2$ passing through the points $(1, -2)$ and $(2, 4)$.

(b) Find the distance from the point $P = (1, 0, 1, 3)$ to the line L in $\Re^4$ passing through the points $(1, 2, 0, 1)$ and $(1, 2, 2, 1)$.

(a) $\frac{1}{\sqrt{37}}$ (b) 4

PLANES REVISITED

We now offer an alternative representation for a plane in $\Re^3$ than that given in Theorem 2.19, page 71.



Just as a line in $\Re^2$ is determined by a point on the line and its slope, so then is a plane in $\Re^3$ determined by a point on the plane and a nonzero vector orthogonal to the plane (a **normal vector** to the plane). To be more specific, suppose we want the equation of the plane with normal vector $\boldsymbol{n} = (a, b, c)$ that contains the point $A_0 = (x_0, y_0, z_0)$. For any point $P = (x, y, z)$ on the plane we have:

$$\boldsymbol{n} \cdot \overrightarrow{A_0P} = 0$$

or: $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) \cdot (\boldsymbol{x} - \boldsymbol{x_0}, \boldsymbol{y} - \boldsymbol{y_0}, \boldsymbol{z} - \boldsymbol{z_0}) = 0$ normal form

or: $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$ scalar form

or: $ax + by + cz = d$, where $d = ax_0 + by_0 + cz_0$ general form

Note that a normal to the plane can easily be spotted from any of the above forms. For example, $\boldsymbol{n} = (\mathbf{2}, \mathbf{5}, \mathbf{-4})$ is a normal to the plane:

$$(\mathbf{2}, \mathbf{5}, \mathbf{-4}) \cdot (\boldsymbol{x} - \mathbf{1}, \boldsymbol{y} - \mathbf{3}, \boldsymbol{z} + \mathbf{2}) = 0$$
$$\mathbf{2}(x - 1) + \mathbf{5}(y - 3) - \mathbf{4}(z + 2) = 0$$
$$\mathbf{2}x + \mathbf{5}y - \mathbf{4}z = 25$$

EXAMPLE 7.4 Find a normal, scalar, and general form equation of the plane passing through the point $(1, 3, -2)$ with normal vector $\boldsymbol{n} = (\mathbf{4}, \mathbf{-1}, \mathbf{5})$.

SOLUTION:

normal: $(\mathbf{4}, \mathbf{-1}, \mathbf{5}) \cdot (\boldsymbol{x} - \mathbf{1}, \boldsymbol{y} - \mathbf{3}, \boldsymbol{z} + \mathbf{2}) = 0$

scalar: $4(x - 1) - 1(y - 3) + 5(z + 2) = 0$

general: $4x - y + 5z = -9$

CHECK YOUR UNDERSTANDING 7.7

Find an equation of the plane passing through the point $(1, 3, -2)$ with normal parallel to the line passing through the points $(1, 1, 0), (0, 2, 1)$.

Answer: See page B-31.

EXAMPLE 7.5 Express the plane $3x + y - 2z = 6$ in the vector form of Theorem 2.19, page 71.

SOLUTION: In order to write the plane in vector form we need two direction vectors and a translation vector. We chose $(\mathbf{2, 0, 0})$ as our translation vector (corresponding to the vector $\boldsymbol{w} = (\boldsymbol{x_0}, \boldsymbol{y_0}, \boldsymbol{z_0})$ in Figure 2.9, page 71).

$(\mathbf{3, 1, -2}) \cdot (\mathbf{0, 2, 1}) = 0$ and $(\mathbf{3, 1, -2}) \cdot (\mathbf{2, 0, 3}) = 0$

Any two linearly independent vectors orthogonal to $\boldsymbol{n} = (\mathbf{3, 1, -2})$ can serve as direction vectors, say $(\mathbf{0, 2, 1})$ and $(\mathbf{2, 0, 3})$ [corresponding to $\boldsymbol{u} = \overrightarrow{AB}$ and $\boldsymbol{v} = \overrightarrow{AC}$ in Figure 2.9]. This leads us to the following vector form equation of the plane:

$$\{(\mathbf{2, 0, 0}) + r(\mathbf{0, 2, 1}) + s(\mathbf{2, 0, 3}) \mid r, s \in R\}$$

CHECK YOUR UNDERSTANDING 7.8

With reference to Example 7.5, show, directly that:

$$\{(x, y, z) \mid 3x + y - 2z = 6\} = \{(2, 0, 0) + r(0, 2, 1) + s(2, 0, 3) \mid r, s \in R\}$$

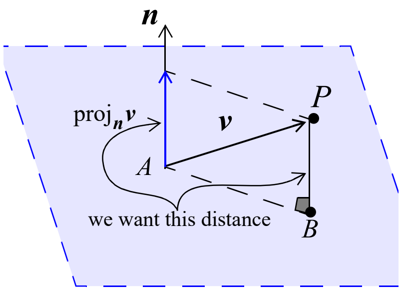
Answer: See page B-31.

EXAMPLE 7.6

Find the distance from the point $P = (4, 3, 2)$ to the plane $2x - 3y + z = 5$.

Any point (x, y, z) satisfying the equation $2x - 3y + z = 5$ would do just as well.

SOLUTION: We begin by choosing the point $A = (0, 0, 5)$ on the plane (note that $2 \cdot 0 - 3 \cdot 0 + 5 = 5$). We position the normal vector $\boldsymbol{n} = (\mathbf{2, -3, 1})$ so that its initial point is at A, and then determine the vector $\boldsymbol{v}$ from A to P:



$$\boldsymbol{v} = (\mathbf{4, 3, 2}) - (\mathbf{0, 0, 5}) = (\mathbf{4, 3, -3})$$

Applying Theorem 7.2, we have:

$$\begin{aligned}\text{proj}_{\boldsymbol{n}}\boldsymbol{v} &= \left(\frac{\boldsymbol{v} \cdot \boldsymbol{n}}{\boldsymbol{n} \cdot \boldsymbol{n}}\right)\boldsymbol{n} \\ &= \left(\frac{(\mathbf{4, 3, -3}) \cdot (\mathbf{2, -3, 1})}{(\mathbf{2, -3, 1}) \cdot (\mathbf{2, -3, 1})}\right)(\mathbf{2, -3, 1}) \\ &= \frac{-4}{14}(\mathbf{2, -3, 1}) = \left(-\frac{4}{7}, \frac{6}{7}, -\frac{2}{7}\right)\end{aligned}$$

Hence:

$$\|\text{proj}_{\boldsymbol{n}}\boldsymbol{v}\| \underset{\substack{\uparrow \\ \text{CYU 7.2(a)}}}{=} \frac{1}{7}\|(\mathbf{-4, 6, -2})\| = \frac{1}{7}\sqrt{16 + 36 + 4} = \frac{\sqrt{56}}{7} = \frac{2\sqrt{14}}{7}$$

CHECK YOUR UNDERSTANDING 7.9

Prove that the distance D from a point (x_0, y_0, z_0) to the plane $ax + by + cz = d$ is given by the formula:

$$D = \frac{|ax_0 + by_0 + cz_0 - d|}{\sqrt{a^2 + b^2 + c^2}}$$

Answer: See page B-32.

CROSS PRODUCT

Here is a handy result:

THEOREM 7.3 If $\boldsymbol{v_1} = (\boldsymbol{a_1}, \boldsymbol{a_2}, \boldsymbol{a_3}), \boldsymbol{v_2} = (\boldsymbol{b_1}, \boldsymbol{b_2}, \boldsymbol{b_3})$ are linearly independent vectors in $\Re^3$, then: $\boldsymbol{v} = (\boldsymbol{a_2 b_3 - a_3 b_2, -a_1 b_3 + a_3 b_1, a_1 b_2 - a_2 b_1})$ is perpendicular to both $\boldsymbol{v_1}$ and $\boldsymbol{v_2}$

PROOF: Rolling up our sleeves, we simply show that $\boldsymbol{v} \cdot \boldsymbol{v_1} = 0$, and leave it for you to verify that $\boldsymbol{v} \cdot \boldsymbol{v}_2$ is also zero:

$$\begin{aligned}\boldsymbol{v} \cdot \boldsymbol{v_1} &= (\boldsymbol{a_2 b_3 - a_3 b_2, -a_1 b_3 + a_3 b_1, a_1 b_2 - a_2 b_1}) \cdot (\boldsymbol{a_1, a_2, a_3}) \\ &= (a_2 b_3 - a_3 b_2)a_1 + (-a_1 b_3 + a_3 b_1)a_2 + (a_1 b_2 - a_2 b_1)a_3 \\ &= a_1 a_2 b_3 - a_1 a_3 b_2 - a_1 a_2 b_3 + a_2 a_3 b_1 + a_1 a_3 b_2 - a_2 a_3 b_1 = 0\end{aligned}$$

Handy or not, how is one to remember that complicated three-tuple $\boldsymbol{v} = (a_2 b_3 - a_3 b_2, -a_1 b_3 + a_3 b_1, a_1 b_2 - a_2 b_1)$ of Theorem 7.3? By **formally** evaluating the following determinant about its first row:

$$\det\begin{bmatrix} \boldsymbol{e_1} & \boldsymbol{e_2} & \boldsymbol{e_3} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} = \det\begin{bmatrix} a_2 & a_3 \\ b_2 & b_3 \end{bmatrix}\boldsymbol{e_1} - \det\begin{bmatrix} a_1 & a_3 \\ b_1 & b_3 \end{bmatrix}\boldsymbol{e_2} + \det\begin{bmatrix} a_1 & a_2 \\ b_1 & b_2 \end{bmatrix}\boldsymbol{e_3}$$

$$= (a_2 b_3 - a_3 b_2)\boldsymbol{e_1} - (a_1 b_3 - a_3 b_1)\boldsymbol{e_2} + (a_1 b_2 - a_2 b_1)\boldsymbol{e_3}$$

and then replacing $\boldsymbol{e_1}, \boldsymbol{e_2}$, and $\boldsymbol{e_3}$ with the standard basis vectors $(\boldsymbol{1, 0, 0}), (\boldsymbol{0, 1, 0})$, and $(\boldsymbol{0, 0, 1})$, respectively, to arrive at the three-tuple of Theorem 7.3:

$$\boldsymbol{v} = (\boldsymbol{a_2 b_3 - a_3 b_2, -a_1 b_3 + a_3 b_1, a_1 b_2 - a_2 b_1})$$

The above vector is called the **cross product of** $\boldsymbol{v_1} = (\boldsymbol{a_1, a_2, a_3})$ and $\boldsymbol{v_2} = (\boldsymbol{b_1, b_2, b_3})$, and is denote by: $\boldsymbol{v_1} \times \boldsymbol{v_2}$. Moreover, to conform with standard notation, we shall replace the standard basis vectors $\boldsymbol{e}_1, \boldsymbol{e}_2$, and $\boldsymbol{e_3}$ with the symbols ***i, j***, and ***k***, respectively; bringing us to:

DEFINITION 7.5
CROSS PRODUCT

The **cross product** of $\boldsymbol{v_1} = (\boldsymbol{a_1, a_2, a_3})$ and $\boldsymbol{v_2} = (\boldsymbol{b_1, b_2, b_3})$ is denoted by $\boldsymbol{v_1} \times \boldsymbol{v_2}$, and is expressed in the form:

$$\boldsymbol{v_1} \times \boldsymbol{v_2} = \det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}$$

For example:

$$\begin{aligned}(\mathbf{2, 3, 4}) \times (\mathbf{3, 1, -2}) &= \det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 2 & 3 & 4 \\ 3 & 1 & -2 \end{bmatrix} \\ &= \det\begin{bmatrix} 3 & 4 \\ 1 & -2 \end{bmatrix}\boldsymbol{i} - \det\begin{bmatrix} 2 & 4 \\ 3 & -2 \end{bmatrix}\boldsymbol{j} + \det\begin{bmatrix} 2 & 3 \\ 3 & 1 \end{bmatrix}\boldsymbol{k} \\ &= (-6-4)\boldsymbol{i} - (-4-12)\boldsymbol{j} + (2-9)\boldsymbol{k} = (\mathbf{-10, 16, -7})\end{aligned}$$

EXAMPLE 7.7 Find the general form equation of the plane that contains the points $A = (1, 2, -1)$, $B = (2, 3, 1)$, $C = (3, -1, 2)$.

SOLUTION: Noting that the vectors

$$\overrightarrow{AB} = (\mathbf{2, 3, 1}) - (\mathbf{1, 2, -1}) = (\mathbf{1, 1, 2})$$

$$\text{and } \overrightarrow{AC} = (\mathbf{3, -1, 2}) - (\mathbf{1, 2, -1}) = (\mathbf{2, -3, 3})$$

are parallel to the plane, we employ Theorem 7.3 to find a normal to the plane:

$$\boldsymbol{n} = \det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ 1 & 1 & 2 \\ 2 & -3 & 3 \end{bmatrix} = 9\boldsymbol{i} + \boldsymbol{j} - 5\boldsymbol{k} = (\mathbf{9, 1, -5})$$

Choosing the point $A = (1, 2, -1)$ on the plane, we proceed as in Example 7.4 to arrive at the general form equation of the plane:

$$\begin{aligned}(\mathbf{9, 1, -5}) \cdot (\boldsymbol{x} - \mathbf{1}, \boldsymbol{y} - \mathbf{2}, \boldsymbol{z} + \mathbf{1}) &= 0 \\ 9(x-1) + (y-2) - 5(z+1) &= 0 \\ 9x + y - 5z - 16 &= 0\end{aligned}$$

CHECK YOUR UNDERSTANDING 7.10

(a) Find the general form equation of the plane that contains the points $A = (3, -2, 2)$, $B = (2, 5, -3)$, $C = (4, 1, 2-3)$.

(b) Verify that your answer in (a) coincides with that of Example 2.15, page 72.

Answer: See page B-32.

EXERCISES

Exercises 1-2. Evaluate $\boldsymbol{u} \cdot \boldsymbol{v}$ for the given n-tuples.

1. $\boldsymbol{u} = (\mathbf{5, 3}), \boldsymbol{v} = (\mathbf{6, 1})$

2. $\boldsymbol{u} = (\mathbf{0, 3, 5, 7}), \boldsymbol{v} = (\mathbf{2, 1, 0, 4})$

Exercises 3-5. Determine the norm $\|\boldsymbol{v}\|$ for the given vector.

3. $\boldsymbol{v} = (\mathbf{3, 2})$

4. $\boldsymbol{v} = (\mathbf{1, 0, -5})$

5. $\boldsymbol{v} = (\mathbf{3, -1, 2, -1})$

6. Find all values of c such that $\|c(\mathbf{2, 3, 1})\| = 9$.

7. Find all values of a such that the vector $(\boldsymbol{a}, \mathbf{3})$ is orthogonal to the vector $(\mathbf{2}\boldsymbol{a}, \mathbf{-5})$.

8. Find all values of a such that the vector $(\mathbf{3}, \boldsymbol{a}, \mathbf{2}\boldsymbol{a})$ is orthogonal to the vector $(\boldsymbol{a}, \mathbf{2}, \boldsymbol{a})$.

9. Find all values of a and b such that the vector $(\boldsymbol{a}, \mathbf{3}, \boldsymbol{b})$ is orthogonal to the vector $(\boldsymbol{b}, \mathbf{3}, \boldsymbol{a})$.

Exercises 10-11. Determine the angle between the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$.

10. $\boldsymbol{u} = (\mathbf{5, 3}), \boldsymbol{v} = (\mathbf{6, 1})$

11. $\boldsymbol{u} = (\mathbf{0, 3, 5, 7}), \boldsymbol{v} = (\mathbf{2, 1, 0, 4})$

Exercises 12-13. Express the given vector $\boldsymbol{v}$ as a sum of a vector parallel to the given vector $\boldsymbol{u}$ and a vector orthogonal to $\boldsymbol{u}$.

12. $\boldsymbol{v} = (\mathbf{1, 5}), \boldsymbol{u} = (\mathbf{3, 2})$

13. $\boldsymbol{v} = (\mathbf{2, 3, -1}), \boldsymbol{u} = (\mathbf{-2, 0, 4})$

Exercises 14-15. Find a normal form, the general form, and a vector form representation (Theorem 2.20, page 72) of the plane passing through the point A_0 with given normal vector $\boldsymbol{n}$.

14. $A_0 = (2, 1, 3), \boldsymbol{n} = (\mathbf{1, 2, -1})$

15. $A_0 = (1, 2, -1), \boldsymbol{n} = (\mathbf{2, 1, 3})$

Exercises 16-18. Find both a normal form equation and a vector form representation (Theorem 2.20, page 72) for the given plane.

16. $2x - 3y + z = 2$

17. $4x + z = 1$

18. $x - 3y - 2z = -2$

19. Find the distance from the point $P = (1, 4)$ and the line L in $\Re^2$ passing through the points $(1, 2)$ and $(2, 1)$.

20. Find the distance from the point $P = (1, 4, -1)$ and the line L in $\Re^3$ passing through the points $(1, 2, 1)$ and $(2, 1, 0)$.

21. Find the distance from the point $P = (1, 4, -1, 1)$ and the line L in $\Re^4$ passing through the points $(1, 2, 1, -2)$ and $(2, 1, 0, 2)$.

22. Find the distance from the point $P = (2, 1, -2)$ to the plane $x - 4y + 2z = 3$.

23. Find the distance from the point $P = (1, 4, -2)$ to the plane $3x + y + 4z = 2$.

24. Determine the angle of intersection of the planes $x - 3y + 2z = 1$ and $2x + y - z = 2$. Suggestions: Consider the normals to those planes.

25. Find the set of vectors in $\Re^3$ orthogonal to:

    (a) the vector $(\mathbf{1, 3, 2})$. (b) the vectors $(\mathbf{1, 3, 2})$ and $(\mathbf{2, 2, -1})$.

    (c) the vectors $(\mathbf{1, 3, 2})$, $(\mathbf{2, 0, -1})$, and $(\mathbf{2, 5, -3})$.

Exercises 26-27. Find the general form equation of the plane that contains the given points.

26. (2, 1, –2), (1, 0, –1), (–1, 3, 0) 27. (0, 0, 1), (2, 0, 0), (0, 3, 0)

28. Find the angle between a main diagonal and an adjacent edge of a cube of volume 8 in.^{3} .

29. Prove Theorem 7.1(i).

30. Prove Theorem 7.1(ii).

31. Prove Theorem 7.1(iv).

32. Establish the following properties for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$ and $r \in \Re$:

    (a) $\mathbf{0} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \mathbf{0} = 0$ (b) $\boldsymbol{u}, \boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$

    (c) $\boldsymbol{u} \cdot r\boldsymbol{v} = r\boldsymbol{u} \cdot \boldsymbol{v}$ (d) $\boldsymbol{u} - \boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} - \boldsymbol{v} \cdot \boldsymbol{w}$

    (e) $\boldsymbol{u} \cdot (\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} - \boldsymbol{u} \cdot \boldsymbol{w}$ (f) $(-\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot (-\boldsymbol{w}) = -(\boldsymbol{v} \cdot \boldsymbol{w})$

33. Show that two nonzero vectors $\boldsymbol{v}$ and $\boldsymbol{u}$ are normal to a given plane if and only if each is a scalar multiple of the other.

34. **(Normal form equation of a line in R^2)** Express the line $ax + by = c$ in the form $\boldsymbol{n} \cdot \boldsymbol{v} = \boldsymbol{n} \cdot \boldsymbol{p}$, where $\boldsymbol{v} = (x, y)$, $\boldsymbol{p}$ is a point on the line, and $\boldsymbol{n} \neq \mathbf{0}$ is a normal to the line

35. Let $A \in M_{n \times n}$, and let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$. Show that $A\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot A^{\mathrm{T}}\boldsymbol{v}$. (See Exercise 19, page 161).

36. $\boldsymbol{u} \in \Re^n$. Show that the function $p_{\boldsymbol{u}}: \Re^n \to \Re$ given by $p_{\boldsymbol{u}}(\boldsymbol{v}) = u \cdot v$ is linear. What is the kernel of $p_{\boldsymbol{u}}$?

37. Let $\boldsymbol{u} \in \Re^n$. Show that if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$ for every $\boldsymbol{v} \in \Re^n$, then $\boldsymbol{u} = \mathbf{0}$.

38. **(Pythagorean Theorem in $\Re^n$)** Let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$. Show that $\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$ if and only if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.

39. **(Parallelogram Law in $\Re^n$)** Let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$. Show that: $\|\boldsymbol{u} + \boldsymbol{v}\|^2 + \|\boldsymbol{u} - \boldsymbol{v}\|^2 = 2\|\boldsymbol{u}\|^2 + 2\|\boldsymbol{v}\|^2$

40. Let $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{R}^n$. Prove that $\|\boldsymbol{u}\| = \|\boldsymbol{v}\|$ if and only if $\boldsymbol{u} + \boldsymbol{v}$ and $\boldsymbol{u} - \boldsymbol{v}$ are orthogonal.

41. Prove that if $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \subset \Re^n$ is such that $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = 0$ if $1 \le i < j \le n$, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a basis for $\Re^n$.

42. Let $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \in \Re^n$. Prove that if $\boldsymbol{u}$ is orthogonal to each $\boldsymbol{v}_i$, $1 \le i \le m$, then $\boldsymbol{u}$ is orthogonal to every $v \in \langle \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \rangle$.

43. **(Cauchy-Schwarz Inequality in $\Re^n$)** Show that if $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$, then $|\boldsymbol{u} \cdot \boldsymbol{v}| \leq \|\boldsymbol{u}\| \|\boldsymbol{v}\|$.
Suggestion: (If $\boldsymbol{u} = 0$, then equality holds). For $\boldsymbol{u} \neq 0$, use the fact that $0 \leq (r\boldsymbol{u} + \boldsymbol{v}) \cdot (r\boldsymbol{u} + \boldsymbol{v}) = (\boldsymbol{u} \cdot \boldsymbol{u})x^2 + (2\boldsymbol{u} \cdot \boldsymbol{v})x + (\boldsymbol{v} \cdot \boldsymbol{v})$ to conclude that the discriminant of the quadratic polynomial $(\boldsymbol{u} \cdot \boldsymbol{u})x^2 + (2\boldsymbol{u} \cdot \boldsymbol{v})x + (\boldsymbol{v} \cdot \boldsymbol{v})$ cannot be positive.

44. Use the above Cauchy-Schwuarz Inequality to show that for any nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$:
$$\left|\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\|\boldsymbol{u}\| \|\boldsymbol{v}\|}\right| \leq 1.$$

45. Establish the following properties for $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^3$ and $r, s \in \Re$:
(a) $(r\boldsymbol{u}) \times (s\boldsymbol{v}) = (rs)(\boldsymbol{u} \times \boldsymbol{v})$ (b) $(-\boldsymbol{u}) \times \boldsymbol{v} = \boldsymbol{u} \times (-\boldsymbol{v}) = -(\boldsymbol{u} \times \boldsymbol{v})$
(c) $\boldsymbol{u} \times (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} \times \boldsymbol{v}) + (\boldsymbol{u} \times \boldsymbol{w})$ (d) $\boldsymbol{u} \cdot (\boldsymbol{v} \times \boldsymbol{w}) = (\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$

46. **(Metric Space Structure of $\Re^n$)** Define the **distance** between two vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$ to be $d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|$. Prove that for all $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$:
    (a) $d(\boldsymbol{u}, \boldsymbol{v}) \geq 0$.
    (b) $d(\boldsymbol{u}, \boldsymbol{v}) = 0$ if and only if $\boldsymbol{u} = \boldsymbol{v}$.
    (c) $d(\boldsymbol{u}, \boldsymbol{v}) = d(\boldsymbol{v}, \boldsymbol{u})$
    (d) $d(\boldsymbol{u}, \boldsymbol{w}) \leq d(\boldsymbol{u}, \boldsymbol{v}) + d(\boldsymbol{v}, \boldsymbol{w})$ Suggestion: Use the Cauchy-Schwuarz Inequality of Exercise 41.

47. **(PMI)** Use the principle of mathematical induction to show that for any $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m \in \Re^n$ and any $a_1, a_2, \ldots, a_m \in \Re$: .
$$\boldsymbol{u} \cdot (a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_m\boldsymbol{v}_m) = a_1\boldsymbol{u} \cdot \boldsymbol{v}_1 + a_2\boldsymbol{u} \cdot \boldsymbol{v}_2 + \ldots + a_m\boldsymbol{u} \cdot \boldsymbol{v}_m$$

PROVE OR GIVE A COUNTEREXAMPLE

48. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$. If $\boldsymbol{u} \neq \boldsymbol{0}$ and if $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u} \cdot \boldsymbol{w}$, then $\boldsymbol{v} = \boldsymbol{w}$.

49. Let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$. If $\boldsymbol{v} \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w}$ for every $\boldsymbol{w} \in \Re^n$, then $\boldsymbol{u} = \boldsymbol{v}$.

50. Let $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$. If $\boldsymbol{v} = \boldsymbol{w}_1 + \boldsymbol{w}_2$ and $\boldsymbol{v} = \boldsymbol{z}_1 + \boldsymbol{z}_2$ with $\boldsymbol{w}_1$ and $\boldsymbol{z}_1$ multiples of $\boldsymbol{v}$, and if $\boldsymbol{w}_2$ and $\boldsymbol{z}_2$ are orthogonal to $\boldsymbol{u}$, then $\boldsymbol{w}_1 = \boldsymbol{z}_1$ and $\boldsymbol{w}_2 = \boldsymbol{z}_2$.

51. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{z} \in \Re^n$, with $\boldsymbol{u} \neq \boldsymbol{0}$. If $\boldsymbol{u}$ is orthogonal to both $\boldsymbol{v}$ and $\boldsymbol{z}$, then $\boldsymbol{v} = c\boldsymbol{z}$ for some $c \in \Re$.

52. The function $N\colon \Re^n \to \Re$ given by $N(\boldsymbol{v}) = \|\boldsymbol{v}\|$ is linear.

53. $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{v} \times \boldsymbol{u}$ for all $\boldsymbol{u}, \boldsymbol{v} \in \Re^3$.

§2. INNER PRODUCT

As you know, a vector space V comes equipped with but two operations: addition, and scalar multiplication. We now enrich that algebraic structure by adding another binary function on V — one inspired by the dot-product properties of Theorem 7.1, page 279:

While the scalar product $r\boldsymbol{v}$ assigns a vector to a scalar r and a vector $\boldsymbol{v}$, the inner product $\langle u, v\rangle$ assigns a real number to a pair of vectors.

DEFINITION 7.6
INNER PRODUCT

An **inner product** on a vector space V is a function which assigns a real number $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ to any two vectors $\boldsymbol{u}, \boldsymbol{v} \in V$, such that:

positive-definite axiom: (i) $\langle \boldsymbol{v}, \boldsymbol{v}\rangle \geq 0$, and $\langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0$ only if $\boldsymbol{v} = \boldsymbol{0}$

commutative axiom: (ii) $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{u}\rangle$

homogeneous axiom: (iii) $\langle r\boldsymbol{u}, \boldsymbol{v}\rangle = r\langle \boldsymbol{u}, \boldsymbol{v}\rangle$

distributive axiom: (iv) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle$

INNER PRODUCT SPACE A vector space together with an inner product is said to be an **inner product space**.

The Euclidean vector space $\Re^n$ with dot product $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{u} \cdot \boldsymbol{v}$ is called the **Euclidean inner product space** of dimension n. There are other inner products that can be imposed on $\Re^n$, among them:

Why are we requiring the c's to be positive?

EXAMPLE 7.8
WEIGHTED EUCLIDEAN INNER PRODUCT SPACE

For any positive real numbers $c_1, c_2, \ldots, c_n$:

$$\langle (\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}), (\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n})\rangle$$
$$= c_1u_1v_1 + c_2u_2v_2 + \ldots + c_nu_nv_n$$

is an inner product on $\Re^n$.

SOLUTION: We show that the distributive axiom (iv) holds and invite you to establish the remaining axioms in the exercises:

For $\boldsymbol{u} = (\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}), \boldsymbol{v} = (\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}), \boldsymbol{w} = (\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_n})$:

$$\begin{aligned}
\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}\rangle &= \langle (\boldsymbol{u_1} + \boldsymbol{v_1}, \boldsymbol{u_1} + \boldsymbol{v_1}, \ldots, \boldsymbol{u_1} + \boldsymbol{v_1}), (\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots \boldsymbol{w_n})\rangle \\
&= c_1(u_1 + v_1)w_1 + c_1(u_1 + v_1)w_1 + \ldots + c_1(u_1 + v_1)w_1 \\
&= (c_1u_1w_1 + c_2u_2w_2 + \ldots + c_nu_nw_n) + (c_1v_1w_1 + c_2v_2w_2 + \ldots + c_nv_nw_n) \\
&= \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle
\end{aligned}$$

CHECK YOUR UNDERSTANDING 7.11

Verify that

$$\langle \boldsymbol{a_2x^2 + a_1x + a_0}, \boldsymbol{b_2x^2 + b_1x + b_0}\rangle = a_2b_2 + a_1b_1 + a_0b_0$$

is an inner product on P_2.

Answer: See page B-33.

The following theorem extends the Euclidean dot-product results of Exercise 32, page 290 to general inner product spaces:

THEOREM 7.4 For every $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in an inner product space V:

(a) $\langle \boldsymbol{0}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{0}\rangle = 0$

(b) $\langle \boldsymbol{u}, \boldsymbol{v}+\boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{w}\rangle$

(c) $\langle \boldsymbol{u}, r\boldsymbol{v}\rangle = r\langle \boldsymbol{u}, \boldsymbol{v}\rangle$

(d) $\langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle - \langle \boldsymbol{v}, \boldsymbol{w}\rangle$

(e) $\langle \boldsymbol{u}, \boldsymbol{v}-\boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{u}, \boldsymbol{w}\rangle$

(f) $\langle -\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, -\boldsymbol{w}\rangle = -\langle \boldsymbol{v}, \boldsymbol{w}\rangle$

In the exercises you are asked to establish the following generalization and combination of (b) and (c).

For $\boldsymbol{u}, \boldsymbol{v_1}, \boldsymbol{v_2}\ldots, \boldsymbol{v_n} \in V$

and $\boldsymbol{c_1}, \boldsymbol{c_2}\ldots, \boldsymbol{c_n} \in \Re$:

$$\langle \boldsymbol{u}, \sum_{i=1}^{n} c_i\boldsymbol{v_i}\rangle = \sum_{i=1}^{n} c_i\langle u, v_i\rangle$$

PROOF: We verify (d), and leave it for you to establish the rest.

Definition 2.7, page 55: $\langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}+(-\boldsymbol{v}), \boldsymbol{w}\rangle$

Axiom (iv): $= \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle -\boldsymbol{v}, \boldsymbol{w}\rangle$

Theorem 2.11 (x), page 56: $= \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle -1\boldsymbol{v}, \boldsymbol{w}\rangle$

Axiom (iii): $= \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{u}, \boldsymbol{w}\rangle$

CHECK YOUR UNDERSTANDING 7.12

Prove: $\langle r\boldsymbol{u}, r\boldsymbol{v}\rangle = r^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle$

Answer: See page B-33.

DISTANCE IN AN INNER PRODUCT SPACE

In the previous section we defined the norm (or magnitude) of a vector $\boldsymbol{v}$ in $\Re^n$ in terms of the dot product: $\|\boldsymbol{v}\| = \sqrt{\boldsymbol{v}\cdot\boldsymbol{v}}$. The dot product was also used to describe the distance between two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\Re^n$: $\|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{(\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v})}$. Replacing "dot-product" with "inner product" enables us to extend the notion of magnitude and distance to any inner product space:

DEFINITION 7.7

NORM AND DISTANCE

The **norm** (or magnitude) of a vector $\boldsymbol{v}$ in an inner product space V, denoted by $\|\boldsymbol{v}\|$, is given by:

$$\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v} \rangle}$$

The **distance** between two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V is given by $\|\boldsymbol{u} - \boldsymbol{v}\|$.

CHECK YOUR UNDERSTANDING 7.13

Show that for any vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space V and any $r \in \Re$:

(a) $\|r\boldsymbol{v}\| = |r|\|\boldsymbol{v}\|$ (b) $\|\boldsymbol{u} + \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{v} \rangle + \|\boldsymbol{v}\|^2$

Answer: See page B-33.

EXAMPLE 7.9

Find the distance between the two vectors $p_1(x) = 2x^2 - x + 1$ and $p_2(x) = 3x + 4$ in the inner product space P_2 of CYU 7.11

SOLUTION: Utilizing the inner product:

$$\langle a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0 \rangle = a_2b_2 + a_1b_1 + a_0b_0$$

we have:

$$p_1(x) - p_2(x) = (2x^2 - x + 1) - (3x + 4) = 2x^2 - 4x - 3$$

$$\|\boldsymbol{p_1} - \boldsymbol{p_2}\| = \sqrt{\langle \boldsymbol{p_1} - \boldsymbol{p_2}, \boldsymbol{p_1} - \boldsymbol{p_2} \rangle} \overset{\downarrow}{=} \sqrt{\langle 2x^2 - 4x - 3, 2x^2 - 4x - 3 \rangle}$$

$$= \sqrt{2^2 + (-4)^2 + (-3)^2} = \sqrt{29}$$

CHECK YOUR UNDERSTANDING 7.14

With reference to the weighted inner product space:

$$\langle (\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}), (\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3}) \rangle = 5u_1v_1 + 2u_2v_2 + 4u_3v_3$$

on $\Re^3$ (see Example 7.8), determine:

(a) The magnitude of the vector $(\mathbf{3}, \mathbf{5}, \mathbf{-8})$.

(b) The distance between the vector $(\mathbf{3}, \mathbf{5}, \mathbf{-8})$ and the vector $(\mathbf{1}, \mathbf{0}, \mathbf{2})$.

(a) $7\sqrt{1}$ (b) $\sqrt{645}$

THE CAUCHY-SCHWARZ INEQUALITY

The following theorem will enable us to extend the concept of an angle between two vectors in $\Re^n$ to vectors in an inner product space:

THEOREM 7.5
CAUCHY-SCHWARZ

For any two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space:

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

PROOF: If either $\boldsymbol{u} = \boldsymbol{0}$ or $\boldsymbol{v} = \boldsymbol{0}$, then $|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| = 0$ and we are done. For $\boldsymbol{u} \neq \boldsymbol{0}$ and $\boldsymbol{v} \neq \boldsymbol{0}$, we first show that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle \ge -\|\boldsymbol{u}\|\|\boldsymbol{v}\|$:

The proof sketched out in Exercise 43, page 275, can also be used to establish this result.

Definition 7.5(i):
$$\left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} + \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} + \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \ge 0$$

7.5(iv):
$$\left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} + \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle + \left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} + \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \ge 0$$

7.5(iv):
$$\left\langle \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}, \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle + \left\langle \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle + \left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u}, \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle + \left\langle \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \ge 0$$

CYU 7.10:
$$\frac{\mathbf{1}}{\|\boldsymbol{u}\|^2}\langle \boldsymbol{u}, \boldsymbol{u}\rangle + \frac{\mathbf{2}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \frac{\mathbf{1}}{\|\boldsymbol{v}\|^2}\langle \boldsymbol{v}, \boldsymbol{v}\rangle \ge 0$$

$$\mathbf{1} + \frac{2}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle u, v\rangle + \mathbf{1} \ge 0$$

$$\frac{2}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle \boldsymbol{u}, \boldsymbol{v}\rangle \ge -2$$

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle \ge -\|\boldsymbol{u}\|\|\boldsymbol{v}|$$

In the following Check Your Understanding box you are asked to show that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$. Putting the two inequalities together, we come up with $-\|\boldsymbol{u}\|\|\boldsymbol{v}\| \le \langle \boldsymbol{u}, \boldsymbol{v}\rangle \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$, which is to say: $|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$

CHECK YOUR UNDERSTANDING 7.15

Verify:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

Suggestion: Begin with $\left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \ge 0$.

Answer: See page B-34.

Here are norm properties that are reminiscent of absolute value properties in $\Re$:

THEOREM 7.6 Let V be an inner product space. For all $\boldsymbol{u}, \boldsymbol{v} \in V$ and $r \in \Re$:

(a) $\|\boldsymbol{v}\| \geq 0$, and $\|\boldsymbol{v}\| = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$

(b) $\|r\boldsymbol{v}\| = |r|\|\boldsymbol{v}\|$

TRIANGLE INEQUALITY (c) $\|\boldsymbol{u}+\boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$

PROOF: (a) A consequence of Definition 7.5(i) and $\|\boldsymbol{v}\| = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$.

(b) $$\|r\boldsymbol{v}\| = \sqrt{\langle r\boldsymbol{v}, r\boldsymbol{v}\rangle} \underset{\substack{\uparrow \\ \text{CYU 7.11}}}{=} \sqrt{r^2\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = \sqrt{r^2}\sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = |r|\|\boldsymbol{v}\|$$

(c) $$\|\boldsymbol{u}+\boldsymbol{v}\|^2 \overset{\substack{\text{CYU 7.12(b)} \\ \downarrow}}{=} \|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \|\boldsymbol{v}\|^2$$

Cauchy-Schwarz inequality: $$\leq \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2$$

$$= (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2$$

Taking the square root of both sides of $\|\boldsymbol{u}+\boldsymbol{v}\|^2 \leq (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2$ yields the desired result.

We now extend the angle concept of Definition 7.3, page 281, to inner product spaces:

The Cauchy-Schwarz inequality plays a hidden role in this definition. (Where?)

DEFINITION 7.8
ANGLE BETWEEN VECTORS

The angle θ between two nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space is given by:

$$\theta = \cos^{-1}\left(\frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right)$$

EXAMPLE 7.10 Find the angle between the two vectors $p_1(x) = 2x^2 - x + 1$ and $p_2(x) = 3x + 4$ in the inner product space P_2 of CYU 7.10

SOLUTION:

While it is admirable that we were able to extend the geometrical notion of the angle between vectors in the plane to vectors in an arbitrary inner product space, the real benefit of that generalization surfaces in the next section, where the concept of orthogonality takes center stage.

$$\begin{aligned}\theta &= \cos^{-1}\left(\frac{\langle 2x^2 - x + 1, 3x + 4\rangle}{\|2x^2 - x + 1\|\|3x + 4\|}\right)\\ &= \cos^{-1}\left(\frac{2\cdot 0 + (-1)\cdot 3 + 1\cdot 4}{\sqrt{\langle 2x^2 - x + 1, 2x^2 - x + 1\rangle}\sqrt{\langle 3x + 4, 3x + 4\rangle}}\right)\\ &= \cos^{-1}\left(\frac{1}{\sqrt{2\cdot 2 + (-1)(-1) + 1\cdot 1}\ \ \sqrt{3\cdot 3 + 4\cdot 4}}\right)\\ &= \cos^{-1}\left(\frac{1}{\sqrt{6}\ \sqrt{15}}\right) \approx 84^\circ\end{aligned}$$

$\cos^{-1}\left(\frac{-49}{\sqrt{351}\ \sqrt{29}}\right) \approx 119.1^\circ$

CHECK YOUR UNDERSTANDING 7.16

With reference to the weighted inner product space:

$$\langle(\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}), (\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3})\rangle = 5u_1v_1 + 2u_2v_2 + 4u_3v_3$$

on $\Re^3$ (see Example 7.7), determine the angle between the vectors $(\mathbf{3, 5, -8})$ and $(\mathbf{1, 0, 2})$.

EXERCISES

Exercises 1-8. With reference to the weighted inner product space:

$$\langle(\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}), (\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3})\rangle = 5u_1v_1 + 2u_2v_2 + 4u_3v_3$$

of Example 7.7, determine:

1. The magnitude of the vector $(\mathbf{1, 2, -3})$.
2. The magnitude of the vector $(\mathbf{3, 2, 0})$.
3. The distance between the vectors $(\mathbf{1, 2, -3})$ and $(\mathbf{1, 0, 2})$.
4. The distance between the vectors $(\mathbf{3, 2, 0})$ and $(\mathbf{1, 2, -3})$.
5. The angle between the vectors $(\mathbf{1, 2, -3})$ and $(\mathbf{1, 0, 2})$.
6. The angle between the vectors $(\mathbf{3, 2, 0})$ and $(\mathbf{1, 2, -3})$.
7. Verify that the Cauchy-Schwarz inequality holds for the vectors $(\mathbf{1, 2, -3})$ and $(\mathbf{1, 0, 2})$.
8. Verify that the Cauchy-Schwarz inequality holds for the vectors $(\mathbf{3, 2, 0})$ and $(\mathbf{1, 2, -3})$.

Exercises 9-16. Referring to the inner product space:

$$\langle a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0\rangle = a_2b_2 + a_1b_1 + a_0b_0$$

on P_2 of CYU 7.11, determine:

9. The magnitude of the vector $2x^2 - x + 3$.
10. The magnitude of the vector $-x^2 + x - 5$.
11. The distance between the vectors $2x^2 - x + 3$ and $-x^2 + x - 5$.
12. The distance between the vectors $3x^2 + 1$ and $2x - 5$.
13. The angle between the vectors $2x^2 - x + 3$ and $-x^2 + x - 5$.
14. The angle between the vectors $3x^2 + 1$ and $2x - 5$.
15. Verify that the Cauchy-Schwarz inequality holds for the vectors $2x^2 - x + 3$ and $-x^2 + x - 5$.
16. Verify that the Cauchy-Schwarz inequality holds for the vectors $3x^2 + 1$ and $2x - 5$.

17. For $\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \boldsymbol{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$ in the vector space $M_{2 \times 2}$, define:

$$\langle \boldsymbol{A}, \boldsymbol{B}\rangle = a_{11}b_{11} + a_{12}b_{12} + a_{21}b_{21} + a_{22}b_{22}$$

Show that the above operator is an inner product on $M_{2 \times 2}$.

Exercises 18-22. Referring to the inner product space on Exercise 17, determine:

18. The magnitude of the vector $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$.

19. The magnitude of the vector $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$.

20. The distance between the vectors $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$.

21. The angle between the vectors $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$.

22. Verify that the Cauchy-Schwarz inequality holds for the vectors $\begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1 \\ 1 & 0 \end{bmatrix}$.

23. Verify that $\langle \sum_{i=0}^{n} a_i x^i, \sum_{i=0}^{n} b_i x^i \rangle = \sum_{i=0}^{n} a_i b_i$ is an inner product on the polynomial space P_n.

24. **(Calculus Dependent)** (a) Show that $C[a, b] = \{f: [a, b] \to R | f \text{ is continuous}\}$ is a subset of the function vector space $F[a, b]$ of Theorem 2.4, page 44.

    (b) Show that $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ is an inner product on $C[a, b]$ (called the **standard inner product on $C[a, b]$**).

Exercises 24-35. Calculus Dependent) Referring to the inner product space on Exercise 24, determine:

25. The magnitude of the vector $2x^2 - x + 3$ in the inner product space $C[0, 1]$.

26. The distance between the vectors $2x^2 - x + 3$ and $-x^2 + x - 5$ in the inner product space $C[0, 1]$.

27. The angle between the vectors $2x^2 - x + 3$ and $-x^2 + x - 5$ in the inner product space $C[0, 1]$.

28. The magnitude of the vector e^x in the inner product space $C[0, 1]$.

29. The distance between the vectors e^x and x in the inner product space $C[0, 1]$.

30. The angle between the vectors e^x and x in the inner product space $C[0, 1]$.

31. The magnitude of the vector $\sin x$ in the inner product space $C[-\pi, \pi]$.

32. The distance between the vectors $\sin x$ and $\cos x$ in the inner product space $C[-\pi, \pi]$.

33. The angle between the vectors $\sin x$ and $\cos x$ in the inner product space $C[-\pi, \pi]$.

34. Verify that the Cauchy-Schwarz inequality holds for the vectors e^x and x in the inner product space $C[0, 1]$.
35. Verify that the Cauchy-Schwarz inequality holds for the vectors $\sin x$ and $\cos x$ in the inner product space $C[-\pi, \pi]$.
36. Prove that ordinary multiplication in the set of real numbers ***R*** is an inner product on the vector space $\Re$.
37. Prove Theorem 7.3(a).
38. Prove Theorem 7.3(b).
39. Prove Theorem 7.3(c).
40. Prove Theorem 7.3(e).
41. Prove Theorem 7.3(f).
42. Let $\boldsymbol{u}, \boldsymbol{v} \in V$, V an inner product space. Show that $\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2 = 4\langle \boldsymbol{u}, \boldsymbol{v}\rangle$
43. Let $\boldsymbol{u}, \boldsymbol{v} \in V$, V an inner product space. Show that $\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle = \|\boldsymbol{u}\|^2 - \|\boldsymbol{v}\|^2$.
44. Let $\boldsymbol{u}, \boldsymbol{v} \in V$, V an inner product space. Show that $\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2$ if and only if $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$.
45. Let $\boldsymbol{u} \in V$, V an inner product space. Show that $\{\boldsymbol{v} \in V | \langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0\}$ is a subspace of V.
46. **(PMI)** Let V be an inner product space.Use the principle of mathematical induction to show that for any $\boldsymbol{u}, \boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ and any $a_1, a_2, \ldots, a_m \in \Re$:

$$\langle \boldsymbol{u}, a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_m\boldsymbol{v}_n\rangle = a_1\langle \boldsymbol{u}, \boldsymbol{v}_1\rangle + a_2\langle \boldsymbol{u}, \boldsymbol{v}_2\rangle + \ldots + a_m\langle \boldsymbol{u}, \boldsymbol{v}_n\rangle$$

and:

$$\langle a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \ldots + a_m\boldsymbol{v}_n, \boldsymbol{u}\rangle = a_1\langle \boldsymbol{u}, \boldsymbol{v}_1\rangle + a_2\langle \boldsymbol{u}, \boldsymbol{v}_2\rangle + \ldots + a_m\langle \boldsymbol{u}, \boldsymbol{v}_n\rangle$$

PROVE OR GIVE A COUNTEREXAMPLE

47. Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$, V an inner product space. If $\boldsymbol{u} \neq \mathbf{0}$ and if $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle$, then $\boldsymbol{v} = \boldsymbol{w}$.
48. Let $\boldsymbol{u}, \boldsymbol{v} \in V$. If $\langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle$ for every $\boldsymbol{w} \in V$, then $\boldsymbol{u} = \boldsymbol{v}$.
49. There exists an inner product on $\Re^3$ for which $\|(\mathbf{1}, \mathbf{1}, \mathbf{1})\| = 1$.
50. There exists an inner product on $\Re^3$ for which $\|(\mathbf{1}, \mathbf{1}, \mathbf{1})\| = \|(\mathbf{2}, \mathbf{2}, \mathbf{2})\|$.
51. There exists an inner product on $\Re^3$ for which $\|(\mathbf{1}, \mathbf{1}, \mathbf{1})\| > \|(\mathbf{2}, \mathbf{2}, \mathbf{2})\|$.

§3. ORTHOGONALITY

Having extended the concept of angles between vectors in $\Re^n$ to vectors in inner product spaces, we can now extend the definition of orthogonality to those spaces:

DEFINITION 7.9
ORTHOGONAL VECTORS

Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space V are **orthogonal** if $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$.

ORTHOGONAL SET

A set S of vectors in an inner product space V is an **orthogonal set** if $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ for every $\boldsymbol{u}, \boldsymbol{v} \in S$, with $\boldsymbol{u} \neq \boldsymbol{v}$.

EXAMPLE 7.11

Verify that:

$S = \{2x^2 + 2x - 1, -x^2 + 2x + 2, 2x^2 - x + 2\}$ is an orthogonal set in the inner product space P_2 of CYU 7.11, page 293, wherein

$$\langle a_2x^2 + a_1x + a_0, b_2x^2 + b_1x + b_0 \rangle = a_2b_2 + a_1b_1 + a_0b_0$$

SOLUTION: All pairs of vectors from S are orthogonal:

$$\langle 2x^2 + 2x - 1, -x^2 + 2x + 2 \rangle = 2(-1) + 2 \cdot 2 + (-1)2 = 0 \text{ (check)}$$
$$\langle 2x^2 + 2x - 1, 2x^2 - x + 2 \rangle = 2 \cdot 2 + 2(-1) + (-1)2 = 0 \text{ (check)}$$
$$\langle -x^2 + 2x + 2, 2x^2 - x + 2 \rangle = (-1)2 + 2(-1) + 2 \cdot 2 = 0 \text{ (check)}$$

You can check directly that the set S in the above example is a linearly independent in P_2. In general:

THEOREM 7.7

If $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}\}$ is an orthogonal set of nonzero vectors in an inner product space V, then $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \dots, \boldsymbol{v_n}\}$ is a linearly independent set in V.

PROOF: Let:

$$c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \dots c_n\boldsymbol{v_n} = \mathbf{0}$$

For each $\boldsymbol{v_i}$, $1 \le i \le n$, we have:

Therem 7.4(a), page 277: $\langle \boldsymbol{v_i}, c_1\boldsymbol{v_1} + \dots + c_i\boldsymbol{v_i} + \dots + c_n\boldsymbol{v_n} \rangle = \langle \boldsymbol{v_i}, \mathbf{0} \rangle = 0$

$$c_1\langle \boldsymbol{v_i}, \boldsymbol{v_1} \rangle + \dots + c_i\langle \boldsymbol{v_i}, \boldsymbol{v_i} \rangle + \dots + c_n\langle \boldsymbol{v_i}, \boldsymbol{v_n} \rangle = 0$$

$\langle \boldsymbol{v_i}, \boldsymbol{v_j} \rangle = 0$ if $i \neq j$: $c_i\langle \boldsymbol{v_i}, \boldsymbol{v_i} \rangle = 0$

Definition 7.5 (i), page 276: $c_i = 0$

$$\begin{tikzcd} c_1\boldsymbol{v_1} + c_2\boldsymbol{v_2} + \dots c_n\boldsymbol{v_n} = \mathbf{0} \arrow[d, squiggly] \\ \text{each } c_i = 0 \end{tikzcd}$$

CHECK YOUR UNDERSTANDING 7.17

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be an orthogonal set of vectors in an inner product space V, and let $\mathbf{u} \in V$ be such that $\langle \mathbf{u}, \mathbf{v}_i \rangle = 0$ for $1 \le i \le m$. Prove that $\mathbf{u}$ is orthogonal to every vector in $\text{Span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$.

Answer: See page B-34.

DEFINITION 7.10
UNIT VECTOR

A **unit vector** in an inner product space is a vector $\mathbf{v}$ of magnitude 1.

NORMALIZATION

To **normalize** a nonzero vector $\mathbf{v}$ in an inner product space simply multiply it by $\dfrac{1}{\|\mathbf{v}\|}$:

$$\left\| \frac{1}{\|\mathbf{v}\|}\mathbf{v} \right\| \underset{\uparrow}{=} \frac{1}{\|\mathbf{v}\|}\|\mathbf{v}\| = 1$$

CYU 7.13(a), page 294

DEFINITION 7.11
ORTHONORMAL SET

An **orthonormal** set of vectors in an inner product space is a set of orthogonal unit vectors.

The standard basis $S = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of page 94 can easily be shown to be an orthonormal set in the Euclidean inner product space $\Re^n$. Moreover, for any $\mathbf{v} = (c_1, \ldots, c_i, \ldots, c_n)$:

$$\mathbf{v} = (\mathbf{v}\cdot\mathbf{e}_1)\mathbf{e}_1 + \cdots + (\mathbf{v}\cdot\mathbf{e}_i)\mathbf{e}_i + \cdots + (\mathbf{v}\cdot\mathbf{e}_n)\mathbf{e}_n$$

since:

$$\mathbf{v}\cdot\mathbf{e}_i = (c_1, \ldots, \underset{\uparrow}{c_i}, \ldots, c_n)\cdot(0, \ldots, \underset{\uparrow}{1}, \ldots, 0) = c_i$$

i^{th} entry

This nicety extends to any orthonormal basis in any inner product space:

THEOREM 7.8

If $\beta = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthonormal basis in an inner product space V, then, for any $\mathbf{v} \in V$:

$$\mathbf{v} = \langle \mathbf{v}, \mathbf{v}_1 \rangle \mathbf{v}_1 + \langle \mathbf{v}, \mathbf{v}_2 \rangle \mathbf{v}_2 + \ldots + \langle \mathbf{v}, \mathbf{v}_n \rangle \mathbf{v}_n$$

PROOF: Let $\mathbf{v} = \sum_{j=1}^{n} c_j\mathbf{v}_j$. We show $c_i = \langle \mathbf{v}, \mathbf{v}_i \rangle$ for $1 \le i \le n$:

$$\langle \mathbf{v}, \mathbf{v}_i \rangle = \left\langle \sum_{j=1}^{n} c_j\mathbf{v}_j, \mathbf{v}_i \right\rangle \underset{\uparrow}{=} \sum_{j=1}^{n} c_j\langle \mathbf{v}_j, \mathbf{v}_i \rangle \underset{\uparrow}{=} c_i\langle \mathbf{v}_i, \mathbf{v}_i \rangle = c_i\|\mathbf{v}_i\|^2 \underset{\uparrow}{=} c_i$$

Exercise 46, page 300.

$\langle \mathbf{v}_j, \mathbf{v}_i \rangle = \begin{cases} 1 \text{ if } j = i \\ 0 \text{ if } j \ne i \end{cases}$

$\langle \mathbf{v}_i, \mathbf{v}_i \rangle = 1$

CHECK YOUR UNDERSTANDING 7.18

Let $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots \boldsymbol{v_n}\}$ be an orthonormal basis for an inner product space V. Show that for any $a_1\boldsymbol{v_1} + a_2\boldsymbol{v_2} + \ldots + a_n\boldsymbol{v_n}$ and

$$\boldsymbol{w} = b_1\boldsymbol{v_1} + b_2\boldsymbol{v_2} + \ldots + b_n\boldsymbol{v_n}\ ,\ \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \sum_{i=1}^{n} a_ib_i .$$

Answer: See page B-34.

The following theorem spells out a procedure that can be used to construct an orthogonal basis in any given finite dimensional inner product space. Basically, the construction process is such that each newly added vector in an evolving basis is orthogonal to all of its predecessors.

THEOREM 7.9
GRAHM-SCHMIDT PROCESS

Let $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ be a basis for an inner product space V, and let V_i be the following subspaces of V:

$$\boldsymbol{u_1} = \boldsymbol{v_1} \qquad V_1 = \text{Span}\{\boldsymbol{v_1}\}$$

$$\boldsymbol{u_2} = \boldsymbol{v_2} - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v_2}\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u_1} \qquad V_2 = \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}\}$$

$$\boldsymbol{u}_3 = \boldsymbol{v_3} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v_3}\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u_1} - \frac{\langle \boldsymbol{u_2}, \boldsymbol{v_3}\rangle}{\langle \boldsymbol{u_2}, \boldsymbol{u_2}\rangle}\boldsymbol{u_2} \qquad V_3 = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v_2}, \boldsymbol{v_3}\}$$

$$\vdots \qquad \vdots$$

$$\boldsymbol{u_n} = \boldsymbol{v_n} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v_n}\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u_1} - \frac{\langle \boldsymbol{u_2}, \boldsymbol{v_n}\rangle}{\langle \boldsymbol{u_2}, \boldsymbol{u_2}\rangle}\boldsymbol{u_2} - \ldots - \frac{\langle \boldsymbol{u_{n-1}}, \boldsymbol{v_n}\rangle}{\langle \boldsymbol{u_{n-1}}, \boldsymbol{u_{n-1}}\rangle}\boldsymbol{u_{n-1}} \qquad V_n = \text{Span}\{\boldsymbol{v}_1, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$$

Then, $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u_i}\}$ is an orthogonal basis for V_i. In particular, $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$ is an orthogonal basis for V.

PROOF: By Induction on the dimension of V:
Since $\boldsymbol{u_1} = \boldsymbol{v_1}$ and since $\{\boldsymbol{u_1}\}$ is an orthogonal set, the claim is seen to hold for V_i.

Assume that $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u}_k\}$ is an orthogonal basis for V_k, for $k < n$.

We show that $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u}_{k+1}\}$ is an orthogonal basis for V_{k+1}, where:

$$\boldsymbol{u}_{k+1} = \boldsymbol{v}_{k+1} - \frac{\langle \boldsymbol{u_1}, \boldsymbol{v}_{k+1}\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u_1} - \cdots - \frac{\langle \boldsymbol{u}_k, \boldsymbol{v}_{k+1}\rangle}{\langle \boldsymbol{u}_k, \boldsymbol{u}_k\rangle}\boldsymbol{u}_k \qquad \textbf{(*)}$$

We are assuming that $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u}_k\}$ is an orthogonal set. Consequently, to establish orthogonality of $\{\boldsymbol{u_1}, \ldots, \boldsymbol{u}_{k+1}\}$, we need but show that $\langle \boldsymbol{u_i}, \boldsymbol{u_{k+1}}\rangle = 0$ for $1 \le i \le k$:

$$\begin{aligned}
\langle \mathbf{u_i}, \mathbf{u_{k+1}}\rangle &= \left\langle \mathbf{u}_i, \mathbf{v}_{k+1} - \frac{\langle \mathbf{u_1}, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_1, \mathbf{u_1}\rangle}\mathbf{u_1} - \dots - \frac{\langle \mathbf{u}_k, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_k, \mathbf{u}_k\rangle}\mathbf{u}_k\right\rangle \\
&= \langle \mathbf{u_i}, \mathbf{v_{k+1}}\rangle - \frac{\langle \mathbf{u_1}, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u}_1, \mathbf{u_1}\rangle}\langle \mathbf{u_i}, \mathbf{u_1}\rangle - \dots - \frac{\langle \mathbf{u}_i, \mathbf{v_{k+1}}\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\langle \mathbf{u_i}, \mathbf{u_i}\rangle - \dots - \frac{\langle \mathbf{u}_i, \mathbf{v_{k+1}}\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\langle \mathbf{u_i}, \mathbf{u}_k\rangle \\
&= \langle \mathbf{u_i}, \mathbf{v_{k+1}}\rangle - \frac{\langle \mathbf{u}_i, \mathbf{v_{k+1}}\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\langle \mathbf{u_i}, \mathbf{u_i}\rangle \longleftarrow \text{since } \langle \mathbf{u_i}, \mathbf{u}_j\rangle = 0 \text{ if } i \neq j \\
&= \langle \mathbf{u_i}, \mathbf{v_{k+1}}\rangle - \frac{\langle \mathbf{u_i}, \mathbf{u_i}\rangle}{\langle \mathbf{u}_i, \mathbf{u}_i\rangle}\langle \mathbf{u}_i, \mathbf{v_{k+1}}\rangle = \langle \mathbf{u_i}, \mathbf{v_{k+1}}\rangle - 1\langle \mathbf{u}_i, \mathbf{v_{k+1}}\rangle = 0
\end{aligned}$$

Being an orthogonal set, $\{\mathbf{u_1}, \dots, \mathbf{u}_{k+1}\}$ is linearly independent. To show that Span $\{\mathbf{u_1}, \dots, \mathbf{u}_{k+1}\} = \text{Span}\{\mathbf{v}_1, \mathbf{v_2}, \dots, \mathbf{v}_{k+1}\}$ we need but show that $\text{Span}\{\mathbf{v}_1, \mathbf{v_2}, \dots, \mathbf{v}_{k+1}\} \subseteq \text{Span}\ \{\mathbf{u_1}, \dots, \mathbf{u}_{k+1}\}$ (why?). Let's do it:

$$\begin{tikzcd}
\sum\limits_{i=1}^{k+1} a_i\mathbf{v_i} \in \text{Span}\{\mathbf{v}_1, \dots, \mathbf{v}_{k+1}\} \arrow[d, squiggly] \\
\sum\limits_{i=1}^{k+1} a_i\mathbf{v_i} \in \text{Span}\{u_1, \dots, u_{k+1}\}
\end{tikzcd}$$

$$\sum_{i=1}^{k+1} a_i\mathbf{v_i} = \sum_{i=1}^{k} a_i\mathbf{v_i} + (a_{k+1}\mathbf{v_{k+1}})$$

Induction Hypothesis: $$= \sum_{i=1}^{k} b_i\mathbf{u_i} + (a_{k+1}\mathbf{v_{k+1}})$$

from (*): $$= \sum_{i=1}^{k} b_i\mathbf{u_i} + a_{k+1}\left(\mathbf{u_{k+1}} + \sum_{i=1}^{k}\frac{\langle \mathbf{u_k}, \mathbf{v}_{k+1}\rangle}{\langle \mathbf{u_k}, \mathbf{u}_k\rangle}\mathbf{u}_k\right)$$

$$\overset{\uparrow}{=} \sum_{i=1}^{k+1} c_i\mathbf{u_i}$$

where: $c_1 = b_1 + a_{k+1}\frac{\langle \mathbf{u}_1, \mathbf{v}_2\rangle}{\langle \mathbf{u}_1, \mathbf{u}_1\rangle}, c_2 = b_2 + a_{k+1}\frac{\langle \mathbf{u}_2, \mathbf{v}_3\rangle}{\langle \mathbf{u}_2, \mathbf{u}_2\rangle}, \dots, c_{k+1} = a_{k+1}$

Note: To obtain an orthonormal basis for an inner product space, simply normalize the orthogonal basis generated by the Gram-Schmidt process.

EXAMPLE 7.12

Extend $\{(\mathbf{1}, \mathbf{2}, \mathbf{0})\}$ to an orthogonal basis for the Euclidean inner product space $\Re^3$.

SOLUTION: ONE APPROACH: Extend $\{(\mathbf{1}, \mathbf{2}, \mathbf{0})\}$ to a basis for $\Re^3$:

$$\{\underbrace{(\mathbf{1}, \mathbf{2}, \mathbf{0}), (\mathbf{1}, \mathbf{0}, \mathbf{0}), (\mathbf{0}, \mathbf{0}, \mathbf{1})}_{\{\mathbf{v_1}, \mathbf{v_2}, \mathbf{v_3}\}}\}:$$

and then apply the Grand-Schmidt process to the above basis:

Multiplying any $\boldsymbol{u}_i$ in the Gram-Schmidt process by a nonzero constant will not alter that vectors "orthogonality-feature," but will simplify subsequent calculations.

$$\begin{aligned}
\boldsymbol{u}_1 &= (1, 2, 0)\\
\boldsymbol{u}_2 &= \boldsymbol{v}_2 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v}_2\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle}\boldsymbol{u}_1 = (\mathbf{1, 0, 0}) - \frac{(\mathbf{1, 2, 0})\cdot(\mathbf{1, 0, 0})}{(\mathbf{1, 2, 0})\cdot(\mathbf{1, 2, 0})}(\mathbf{1, 2, 0})\\
&= (\mathbf{1, 0, 0}) - \frac{1}{5}(\mathbf{1, 2, 0}) = \left(\frac{\mathbf{4}}{\mathbf{5}}, -\frac{\mathbf{2}}{\mathbf{5}}, \mathbf{0}\right) \underset{\uparrow}{\to} (\mathbf{4, -2, 0}) \underset{\uparrow}{\to} (\mathbf{2, -1, 0})
\end{aligned}$$

see margin

$$\begin{aligned}
\boldsymbol{u}_3 &= \boldsymbol{v}_3 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1\rangle}\boldsymbol{u}_1 - \frac{\langle \boldsymbol{u}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{u}_2, \boldsymbol{u}_2\rangle}\boldsymbol{u}_2\\
&= (\mathbf{0, 0, 1}) - \frac{(\mathbf{1, 2, 0})\cdot(\mathbf{0, 0, 1})}{(\mathbf{1, 2, 0})\cdot(\mathbf{1, 2, 0})}(\mathbf{1, 2, 0}) - \frac{(\mathbf{2, -1, 0})\cdot(\mathbf{0, 0, 1})}{(\mathbf{2, -1, 0})\cdot(\mathbf{2, -1, 0})}(\mathbf{2, -1, 0})\\
&= (\mathbf{0, 0, 1}) - \frac{0}{5}(\mathbf{1, 2, 0}) - \frac{0}{5}(\mathbf{2, -1, 10}) = (\mathbf{0, 0, 1})
\end{aligned}$$

The above process led us to the following orthogonal basis:

$$\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\} = \{(\mathbf{1, 2, 0}), (\mathbf{2, -1, 0}), (\mathbf{0, 0, 1})\}$$

ANOTHER APPROACH:

This brute force approach is not always practical. Software, such as Maple and MATLAB, include the Gram-Schmidt process as a built-in procedure. Yes, the Gram-Schmidt process works off of a basis for the inner product space, but that is not a problem: if you randomly choose n vectors in an n dimensional space, even if $n = 100$, there is little chance that those vectors end up being linearly dependent!

Since we are dealing with a vector space of dimension 3, we can simply roll up our sleeves and construct an orthogonal bases by "brute force." First, add a nonzero vector, $(\boldsymbol{a, b, c})$, to $\{(\mathbf{1, 2, 0})\}$ with:

$$\begin{aligned}
(\boldsymbol{a, b, c})\cdot(\mathbf{1, 2, 0}) &= 0\\
a + 2b + 0\cdot c &= 0
\end{aligned}$$

This can be done in many ways. One way: $(\boldsymbol{a, b, c}) = (\mathbf{0, 0, 1})$, brings us to the orthogonal set $\{(\mathbf{1, 2, 0}), (\mathbf{0, 0, 1})\}$. We still need another nonzero vector $(\boldsymbol{a, b, c})$ — one that is orthogonal to both $(\mathbf{1, 2, 0})$ and $(\mathbf{0, 0, 1})$:

$$\begin{aligned}
(\boldsymbol{a, b, c})\cdot(\mathbf{1, 2, 0}) &= 0 \quad\text{and}\quad (\boldsymbol{a, b, c})\cdot(\mathbf{0, 0, 1}) = 0\\
a + 2b + 0\cdot c &= 0 \quad\text{and}\quad 0\cdot a + 0\cdot b + c = 0\\
a + 2b &= 0 \quad\text{and}\quad c = 0
\end{aligned}$$

How about $(\boldsymbol{a, b, c}) = (\mathbf{6, 3, 0})$? Sure. Leading us to the orthogonal basis $\{(\mathbf{1, 2, 0}), (\mathbf{0, 0, 1}), (\mathbf{6, 3, 0})\}$

CHECK YOUR UNDERSTANDING 7.19

Apply the Gram-Schmidt Process to construct an orthonormal basis for the subspace S of the Euclidean inner product space of $\Re^4$ spanned by the vectors $(\mathbf{2, 1, 1, 0})$, $(\mathbf{1, 0, 1, 0})$, $(\mathbf{3, 1, 2, 0})$, $(\mathbf{0, 1, 0, 1})$.

Answer: See page B-35.

ORTHOGONAL COMPLEMENT

The set of vectors orthogonal to any subspace W of $\Re^3$, denoted by the symbol $W^\perp$, is itself a subspace of $\Re^3$. More specifically:

W	$W^\perp$
$\{\mathbf{0}\}$	$\Re^3$
Line W passing through the origin.	Plane passing through the origin with normal W.
Plane W passing through the origin.	Line passing through the origin orthogonal to W.
$\Re^3$	$\{\mathbf{0}\}$

ORTHOGONAL COMPLEMENT

In a more general setting, for W a subspace of an inner product space V we define the **orthogonal complement** of W to be:

$$W^\perp = \{\boldsymbol{u} \in V \mid \langle \boldsymbol{u}, \boldsymbol{w} \rangle = 0 \text{ for every } \boldsymbol{w} \in W\}$$

We then have:

THEOREM 7.10 If W a subspace on an inner product space V, then:

(i) $W^\perp$ is a subspace of V.

(ii) $W \cap W^\perp = \{\mathbf{0}\}$

(iii) Every vector in V can be **uniquely** expressed as a sum of a vector in W and a vector in $W^\perp$.

(iv) If β_W is a basis for W and $\beta_{W^\perp}$ is a basis for $W^\perp$, then $\beta_W \cup \beta_{W^\perp}$ is a basis for V.

PROOF: (i) For $\boldsymbol{u}_1, \boldsymbol{u}_2 \in W^\perp$, $r \in \Re$, and $\boldsymbol{w} \in W$, we have:

$$\langle r\boldsymbol{u}_1 + \boldsymbol{u}_2, \boldsymbol{w} \rangle = r\langle \boldsymbol{u}_1, \boldsymbol{w} \rangle + \langle \boldsymbol{u}_2, \boldsymbol{w} \rangle = r \cdot 0 + 0 = 0$$

The result now follows from Theorem 2.13, page 61.

(ii) Let $\boldsymbol{v} \in W \cap W^\perp$. Being in both W and $W^\perp$, $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$. It follows, from Axiom (i) of Definition 7.5 (page 287) that $\boldsymbol{v} = \mathbf{0}$.

In this part of the theorem we assume that W is finite dimensional (the result does, however, hold in general).

(iii) Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ be an orthonormal basis for W. For $\boldsymbol{v} \in V$, let:

$$\boldsymbol{v}_W = \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \ldots + \langle \boldsymbol{v}, \boldsymbol{w}_m \rangle \boldsymbol{w}_m \in W$$

We show that the vector $\boldsymbol{v}_{W^\perp} = \boldsymbol{v} - \boldsymbol{v}_W$ is in $W^\perp$ by showing that $\langle \boldsymbol{v}_{W^\perp}, \boldsymbol{w}_i \rangle = 0$, for $1 \le i \le m$ (see CYU 7.17):

$$\begin{aligned}
\langle \boldsymbol{v}_{W^\perp}, \boldsymbol{w}_i \rangle &= \langle \boldsymbol{v} - \boldsymbol{v}_W, \boldsymbol{w}_i \rangle = \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle \boldsymbol{v}_W, \boldsymbol{w}_i \rangle \\
&= \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \ldots + \langle \boldsymbol{v}, \boldsymbol{w}_m \rangle \boldsymbol{w}_m, \boldsymbol{w}_i \rangle \\
&\underset{\substack{\uparrow \\ \text{since } \langle \boldsymbol{w}_i, \boldsymbol{w}_j \rangle = 0 \text{ for } i \neq j}}{=} \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle \langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle \underset{\substack{\uparrow \\ \text{since } \langle \boldsymbol{w}_i, \boldsymbol{w}_i \rangle = 1}}{=} \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle - \langle \boldsymbol{v}, \boldsymbol{w}_i \rangle = 0
\end{aligned}$$

At this point, we have shown that $\boldsymbol{v}$ can be expressed as a sum of a vector in W and a vector in $W^{\perp}$: $\boldsymbol{v} = \boldsymbol{v}_W + \boldsymbol{v}_{W^{\perp}}$.

Uniqueness of the decomposition follows from part (ii) of this theorem, and Exercise 43, page 67.

As for (iv):

Answer: See page B-35.

CHECK YOUR UNDERSTANDING 7.20

Establish Theorem 7.10(iv)

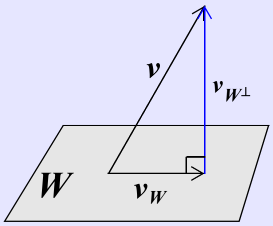
Lets highlight an important observation lurking within the proof of Theorem 7.10(iii):

Compare with Theorem 7.2, page 283.

THEOREM 7.11 Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_m\}$ be an orthonormal basis for a subspace W of an inner product space V. For any $\boldsymbol{v} \in V$ there exist unique vectors $\boldsymbol{v}_W \in W$ and $\boldsymbol{v}_{W^{\perp}} \in W^{\perp}$ such that $\boldsymbol{v} = \boldsymbol{v}_W + \boldsymbol{v}_{W^{\perp}}$, where

$$\boldsymbol{v}_W = \langle \boldsymbol{v}, \boldsymbol{w}_1\rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2\rangle \boldsymbol{w}_2 + \ldots + \langle \boldsymbol{v}, \boldsymbol{w}_m\rangle \boldsymbol{w}_m$$

and $\boldsymbol{v}_{W^{\perp}} = \boldsymbol{v} - \boldsymbol{v}_W$.



Note: $\boldsymbol{v}_W$ is said to be the **orthogonal projection** of $\boldsymbol{v}$ onto W, and we write: $\boldsymbol{v}_W = \text{proj}_W \boldsymbol{v}$.

EXAMPLE 7.13 Let $W = \text{Span}\{\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}\}$, in the Euclidean inner product space $\Re^4$.

(a) Find a basis for $W^{\perp}$.

(b) Express $(\mathbf{1, 2, 3, 4})$ as the sum of a vector in W and a vector in $W^{\perp}$.

SOLUTION: (a) Since $(\mathbf{1, 0, 0, 2})$ and $(\mathbf{1, 1, 1, 0})$ are linearly independent, they constitute a basis for W. To say that $(\boldsymbol{a, b, c, d}) \in W^{\perp}$ is to say that:

We know that $W^{\perp}$ will turn out to be of dimension 2. How?

$$(\boldsymbol{a, b, c, d}) \cdot (\mathbf{1, 0, 0, 2}) = 0 \quad\text{and}\quad (\boldsymbol{a, b, c, d}) \cdot (\mathbf{1, 1, 1, 0}) = 0$$

$$a + 2d = 0 \qquad\qquad a + b + c = 0$$

$$d = -\frac{a}{2} \qquad\qquad b = -a - c$$

Choosing $\boldsymbol{a}$ and $\boldsymbol{c}$ to be our free variables, we have:

$$W^{\perp} = \left\{\left(\boldsymbol{a}, -\boldsymbol{a} - \boldsymbol{c}, \boldsymbol{c}, -\frac{\boldsymbol{a}}{\mathbf{2}}\right) \middle| \boldsymbol{a}, c \in \Re\right\}$$

First setting $a = 0$ and $c = 1$, and then setting $c = 0$ and $a = 2$ leads us to the basis: $\{(\mathbf{0, -1, 1, 0}), (\mathbf{2, -2, 0, -1})\}$ for $W^{\perp}$.

(b) **One approach**: Simply express $(\mathbf{1, 2, 3, 4})$ as a linear combination of the basis $\{(\mathbf{1, 0, 0, 2}), (\mathbf{1, 1, 1, 0}), (\mathbf{0, -1, 1, 0}), (\mathbf{2, -2, 0, -1})\}$:

$$(\mathbf{1, 2, 3, 4}) = a(\mathbf{1, 0, 0, 2}) + b(\mathbf{1, 1, 1, 0}) + c(\mathbf{0, -1, 1, 0}) + d(\mathbf{2, -2, 0, -1})$$

$$\left.\begin{aligned} a+b+2d &= 1\\ b-c-2d &= 2\\ b+c &= 3\\ 2a-d &= 4 \end{aligned}\right\} \rightarrow a = \frac{3}{2}, b = \frac{3}{2}, c = \frac{3}{2}, d = -1$$

Bringing us to:

$$(\mathbf{1, 2, 3, 4}) = \overbrace{\left[\frac{\mathbf{3}}{\mathbf{2}}(\mathbf{1, 0, 0, 2}) + \frac{\mathbf{3}}{\mathbf{2}}(\mathbf{1, 1, 1, 0})\right]}^{\textbf{in } W} + \overbrace{\left[\frac{\mathbf{3}}{\mathbf{2}}(\mathbf{0, -1, 1, 0}) - (\mathbf{2, -2, 0, -1})\right]}^{\text{in } W^{\perp}}$$

$$= \left(\mathbf{3}, \frac{\mathbf{3}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{3}\right) + \left(\mathbf{-2}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{1}\right)$$

Another Approach: First apply the Gram-Schmidt method on $\{(\mathbf{1, 0, 0, 2}), (\mathbf{1, 1, 1, 0})\}$ to obtain an orthonormal basis for W:

$$u_1 = (\mathbf{1, 0, 0, 2})$$

$$u_2 = (\mathbf{1, 1, 1, 0}) - \frac{(\mathbf{1, 1, 1, 0}) \cdot (\mathbf{1, 0, 0, 2})}{(\mathbf{1, 0, 0, 2}) \cdot (\mathbf{1, 0, 0, 2})}(\mathbf{1, 0, 0, 2})$$

Since $\boldsymbol{u_2}$ is orthogonal to $\boldsymbol{u_1}$, so is $5\boldsymbol{u_2}$

$$= (\mathbf{1, 1, 1, 0}) - \frac{1}{5}(\mathbf{1, 0, 0, 2}) = \left(\frac{\mathbf{4}}{\mathbf{5}}, \mathbf{1}, \mathbf{1}, -\frac{\mathbf{2}}{\mathbf{5}}\right) \quad \underset{\substack{\uparrow\\ \text{see margin}}}{\text{Or}}\text{: } (\mathbf{4, 5, 5, -2})$$

Orthonormal basis for W:

$$\{\boldsymbol{w_1}, \boldsymbol{w_2}\} = \left\{\frac{1}{\sqrt{5}}(\mathbf{1, 0, 0, 2}), \frac{1}{\sqrt{70}}(\mathbf{4, 5, 5, -2})\right\}$$

Applying Theorem 7.11 we know that

$$(\mathbf{1, 2, 3, 4}) = (\mathbf{1, 2, 3, 4})_W + (\mathbf{1, 2, 3, 4})_{W^{\perp}}$$

where:

$$(\mathbf{1, 2, 3, 4})_W = [(\mathbf{1, 2, 3, 4}) \cdot \boldsymbol{w_1}]\boldsymbol{w_1} + [(\mathbf{1, 2, 3, 4}) \cdot \boldsymbol{w_2}]\boldsymbol{w_2}$$

$$= \left[(\mathbf{1, 2, 3, 4}) \cdot \left(\frac{\mathbf{1}}{\sqrt{\mathbf{5}}}, \mathbf{0}, \mathbf{0}, \frac{\mathbf{2}}{\sqrt{\mathbf{5}}}\right)\right]\left(\frac{\mathbf{1}}{\sqrt{\mathbf{5}}}, \mathbf{0}, \mathbf{0}, \frac{\mathbf{2}}{\sqrt{\mathbf{5}}}\right) + \left[(\mathbf{1, 2, 3, 4}) \cdot \left(\frac{\mathbf{4}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{-2}}{\sqrt{\mathbf{70}}}\right)\right]\left(\frac{\mathbf{4}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{-2}}{\sqrt{\mathbf{70}}}\right)$$

$$= \frac{9}{\sqrt{5}}\left(\frac{\mathbf{1}}{\sqrt{\mathbf{5}}}, \mathbf{0}, \mathbf{0}, \frac{\mathbf{2}}{\sqrt{\mathbf{5}}}\right) + \frac{21}{\sqrt{70}}\left(\frac{\mathbf{4}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{5}}{\sqrt{\mathbf{70}}}, \frac{\mathbf{-2}}{\sqrt{\mathbf{70}}}\right) = \left(\mathbf{3}, \frac{\mathbf{3}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{3}\right)$$

and:

$$(\mathbf{1, 2, 3, 4})_{W^{\perp}} = [(\mathbf{1, 2, 3, 4}) - (\mathbf{1, 2, 3, 4})_W] = (\mathbf{1, 2, 3, 4}) - \left(\mathbf{3}, \frac{\mathbf{3}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{3}\right) = \left(\mathbf{-2}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{1}\right)$$

Note that both approaches lead to the same decomposition, as must be the case:

$$(\mathbf{1, 2, 3, 4}) = \underset{\text{in } W}{\left(\mathbf{3}, \frac{\mathbf{3}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{3}\right)} + \underset{\text{in } W^{\perp}}{\left(\mathbf{-2}, \frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{3}}{\mathbf{2}}, \mathbf{1}\right)}$$

CHECK YOUR UNDERSTANDING 7.21

Find the orthogonal projection of the vector $v = (\mathbf{2}, \mathbf{0}, \mathbf{1})$ onto the subspace $W = \text{Span}\{(\mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{2}, \mathbf{0})\}$ of the Euclidean inner product space $\Re^3$.

Answer: $\left(\frac{\mathbf{7}}{\mathbf{4}}, \mathbf{1}, \frac{\mathbf{5}}{\mathbf{4}}\right)$

Consider Example 7.3, page 284.

The shortest distance between a vector $\boldsymbol{v}$ in an inner product space V and any vector $\boldsymbol{w}$ in a subspace W of V turns out to be the distance between $\boldsymbol{v}$ and $\boldsymbol{v}_W$:

THEOREM 7.12 Let W be a subspace of the inner product space V, and let $\boldsymbol{v} \in V$. Then:

$$\|\boldsymbol{v} - \underset{\boldsymbol{v}_W}{\underset{\uparrow}{\text{proj}_W \boldsymbol{v}}}\| \le \|\boldsymbol{v} - \boldsymbol{w}\| \text{ for every } \boldsymbol{w} \in W$$

PROOF: Let $\boldsymbol{v}_W = \text{proj}_W \boldsymbol{v}$. For any $\boldsymbol{w} \in W$:

$$\begin{aligned}
\|\boldsymbol{v} - \boldsymbol{w}\|^2 &= \langle \boldsymbol{v} - \boldsymbol{w}, \boldsymbol{v} - \boldsymbol{w} \rangle \\
&= \langle (\boldsymbol{v} - \boldsymbol{v}_W) + (\boldsymbol{v}_W - \boldsymbol{w}), (\boldsymbol{v} - \boldsymbol{v}_W) + (\boldsymbol{v}_W - \boldsymbol{w}) \rangle \\
&= \langle \boldsymbol{v} - \boldsymbol{v}_W, \boldsymbol{v} - \boldsymbol{v}_W \rangle + \langle \boldsymbol{v} - \boldsymbol{v}_W, \boldsymbol{v}_W - \boldsymbol{w} \rangle + \langle \boldsymbol{v}_W - \boldsymbol{w}, \boldsymbol{v} - \boldsymbol{v}_W \rangle + \langle \boldsymbol{v}_W - \boldsymbol{w}, \boldsymbol{v}_W - \boldsymbol{w} \rangle \\
&= \|\boldsymbol{v} - \boldsymbol{v}_W\|^2 + \langle \boldsymbol{v} - \boldsymbol{v}_W, \boldsymbol{v}_W - \boldsymbol{w} \rangle + \langle \boldsymbol{v}_W - \boldsymbol{w}, \boldsymbol{v} - \boldsymbol{v}_W \rangle + \|\boldsymbol{v}_W - \boldsymbol{w}\|^2
\end{aligned}$$

Theore 7.4(b), page 293:

Since $\boldsymbol{v}_W - \boldsymbol{w} \in W$ and $\boldsymbol{v} - \boldsymbol{v}_W \in W^\perp$, the two middle terms in the above expression are 0, bringing us to:

$$\|\boldsymbol{v} - \boldsymbol{w}\|^2 = \|\boldsymbol{v} - \boldsymbol{v}_W\|^2 + \|\boldsymbol{v}_W - \boldsymbol{w}\|^2$$

To complete the proof we need but note that $\|\boldsymbol{v}_W - \boldsymbol{w}\| \ge 0$.

CHECK YOUR UNDERSTANDING 7.22

Find the shortest distance between the vector $\boldsymbol{v} = 3x^2 + \sqrt{3}x$ and the subspace $W = \text{Span}\{x^2 + 1, x^3 + 1\}$ in the inner product space P_3 of CYU 7.11, page 293:

$$\langle a_0 + a_1x + a_2x^2 + a_3x^3, b_0 + b_1x + b_2x^2 + b_3x^3 \rangle = \sum_{i=0}^{3} a_ib_i.$$

Answer: 3

EXERCISES

Exercises 1-7. Determine if the given set of vectors is an orthogonal set in the given inner product space. If so, modify the set to arrive at an orthonormal set.

1. $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1}), (\mathbf{-1, 0, 1})\}$ in the Euclidean inner product space $\Re^3$.
2. $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1}), (\mathbf{-1, 0, 1})\}$ in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = 5u_1v_1 + 2u_2v_2 + 4u_3v_3$.
3. $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1}), (\mathbf{1, 0, 5})\}$ in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = 5u_1v_1 + u_2v_2 + u_3v_3$.
4. $\{x^2 + 2x + 1, 3x^2 + x - 5, 11x^2 - 8x + 5\}$ in the polynomial inner product space of CYU 7.11, page 293.
5. $\left\{\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ \frac{1}{3} & -\frac{1}{4} \end{bmatrix}\right\}$ in the inner product space of Exercise 17, page 298.
6. **(Calculus Dependent)** $\left\{x^2, -\frac{4}{3}x + 1\right\}$ in the inner product space $C[0, 1]$ of Exercise 24, page 299.
7. **(Calculus Dependent)** $\left\{x^2, -\frac{4}{3}x + 1\right\}$ in the inner product space $C[1, 2]$ of Exercise 24, page 299.
8. Use Theorem 7.8 to express $(\mathbf{3, 5, 2})$ in the Euclidean inner product space $\Re^3$ as a linear combination of the vectors in the orthonormal basis $\left\{\frac{1}{\sqrt{5}}(\mathbf{2, 1, 0}), (\mathbf{0, 0, 1}), \frac{1}{\sqrt{5}}(\mathbf{1 - 2, 0})\right\}$.
9. Use Theorem 7.8 to express $2x^2 + 3x - 1$ in the polynomial inner product space of CYU 7.11, page 293, as a linear combination of the vectors in the orthonormal basis $\left\{\frac{x^2}{\sqrt{3}} + \frac{x}{\sqrt{3}} + \frac{1}{\sqrt{3}}, \frac{x^2}{\sqrt{6}} - \frac{2x}{\sqrt{6}} + \frac{1}{\sqrt{6}}, \frac{x^2}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right\}$.
10. Find all values of a for which $\{(\mathbf{1, 3, 2}), (\mathbf{1}, \boldsymbol{a}, \mathbf{1})\}$ is an orthogonal set in the Euclidean inner product space $\Re^3$.
11. Find all values of a and b for which $\{(\mathbf{1, 3}, \boldsymbol{b}), (\mathbf{1}, \boldsymbol{a}, \mathbf{1})\}$ is an orthogonal set in the Euclidean inner product space $\Re^3$.
12. Find all values of a and b for which $\{(\mathbf{1, 1}, a), (\mathbf{-1}, 2b, \mathbf{-1}), (\mathbf{-1, 0, 1})\}$ is an orthogonal set in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = 2u_1v_1 + u_2v_2 + u_3v_3$.

13. Find all values of a and b for which $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1}), (\mathbf{-1, 0, 1})\}$ is an orthogonal set in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = au_1v_1 + bu_2v_2 + u_3v_3$.

14. Find all values of a, b, and c for which $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1}), (\mathbf{1, 0, 5})\}$ is a n orthogonal set in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = au_1v_1 + bu_2v_2 + cu_3v_3$.

15. **(Calculus Dependent)** Find all values of a and b for which $\{ax^2 + 1, -x + b\}$ is an orthogonal set in the inner product space $C[0, 1]$ of Exercise 24, page 299.

16. **(Calculus Dependent)** Find all values of a, and b for which $\{x^2, -x + 1\}$ is an orthogonal set in the inner product space $C[a, b]$ of Exercise 24, page 299.

Exercises 17-26. Find an orthonormal basis for the given inner product space.

17. Span$\{(\mathbf{2, 0, 1}), (\mathbf{-1, 2, 0})\}$ in the Euclidean inner product space $\Re^3$.

18. Span$\{(\mathbf{1, 1, 1}), (\mathbf{-1, 2, -1})\}$ in the Euclidean inner product space $\Re^3$.

19. Span$\{(\mathbf{1, 1, 1, 1}), (\mathbf{-1, 2, -1, 2})\}$ in the Euclidean inner product space $\Re^4$

20. Span$\{(\mathbf{1, 0, 1, 0}), (\mathbf{0, 2, 0, 2}), (\mathbf{2, 0, 0, 0})\}$ in the Euclidean inner product space $\Re^4$.

21. Span$\{(\mathbf{2, 0, 1}), (\mathbf{-1, 2, 0})\}$ in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = 2u_1v_1 + 3u_2v_2 + u_3v_3$.

22. $\{x^2 + 1, x - 5\}$ in the polynomial inner product space of CYU 7.11, page 293.

23. The solution set of $\left.\begin{aligned} 2x + 3y - z + w &= 0 \\ 4x + 3y - 2z - 2w &= 0 \end{aligned}\right\}$ in the Euclidean inner product space $\Re^4$.

24. The solution set of $\left.\begin{aligned} x + 3y - 2z + w &= 0 \\ 4x + 3y - 2z &= 0 \end{aligned}\right\}$ in the Euclidean inner product space $\Re^4$.

25. **(Calculus Dependent)** Span$\{x^2, 2x + 1\}$ in the inner product space $C[0, 1]$ of Exercise 24, page 299.

26. **(Calculus Dependent)** Span$\{x^3, x + 1, x^2 - 1\}$ in the inner product space $C[-1, 1]$ of Exercise 24, page 299.

27. Find an orthonormal basis for $\{(\boldsymbol{a}, \mathbf{2}\boldsymbol{a}, \mathbf{0}) | a \in R\}$ in the Euclidean inner product space $\Re^3$.

28. Find an orthonormal basis for $\{(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a} - \boldsymbol{b}, \mathbf{2}\boldsymbol{b}) | a, b \in \Re\}$ in the Euclidean inner product space $\Re^4$.

29. Find an orthonormal basis for $\{(\boldsymbol{a}, \mathbf{2}\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{a} - \boldsymbol{c}) | a, c \in \Re\}$ in the Euclidean inner product space $\Re^4$.

30. Find an orthonormal basis for $\{(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) | a + b + c = 0\}$ in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3), (\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)\rangle = -5u_1v_1 - 2u_2v_2 + u_3v_3$.

31. Find an orthonormal basis for $\{ax^3 + (a + b)x^2 + cx + 2a | a, b, c \in R\}$ in the polynomial inner product space of CYU 7.11, page 293.

Exercise 32-36. (a) Find a basis for the orthogonal complement of the given Euclidean inner product subspace W.

(b) Express the given vector $\boldsymbol{v}$ as a sum of a vector in W and a vector in $W^{\perp}$.

(c) Determine the distance from $\boldsymbol{v}$ to W.

32. $W = \text{Span}\{(\mathbf{1}, \mathbf{0}), (\mathbf{1}, \mathbf{1})\}$, $\boldsymbol{v} = (\mathbf{1}, \mathbf{3})$.
33. $W = \text{Span}\{(\mathbf{1}, \mathbf{0}, \mathbf{2})\}$, $\boldsymbol{v} = (\mathbf{1}, \mathbf{3}, \mathbf{-2})$.
34. $W = \text{Span}\{(\mathbf{1}, \mathbf{0}, \mathbf{2}), (\mathbf{1}, \mathbf{3}, \mathbf{-2})\}$, $\boldsymbol{v} = (\mathbf{1}, \mathbf{1}, \mathbf{1})$.
35. $W = \text{Span}\{(\mathbf{1}, \mathbf{0}, \mathbf{2}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{1})\}$, $\boldsymbol{v} = (\mathbf{4}, \mathbf{-1}, \mathbf{3}, \mathbf{3})$.
36. $W = \text{Span}\{(\mathbf{1}, \mathbf{0}, \mathbf{2}, \mathbf{3}), (\mathbf{2}, \mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{1}, \mathbf{1}, \mathbf{0}, \mathbf{1})\}$, $\boldsymbol{v} = (\mathbf{1}, \mathbf{3}, \mathbf{-2}, \mathbf{2})$.
37. Find a basis for the orthogonal complement of the subspace $W = \text{Span}\{(\mathbf{2}, \mathbf{0}, \mathbf{1}), (\mathbf{-1}, \mathbf{2}, \mathbf{0})\}$ in the weighted inner product space of Example 7.8, page 292, with $\langle(\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}), (\boldsymbol{v_1}, \boldsymbol{v_2}, \boldsymbol{v_3})\rangle = u_1v_1 + 2u_2v_2 + u_3v_3$, and express the vector $\boldsymbol{v} = (\mathbf{1}, \mathbf{2}, \mathbf{-1})$ as a sum of a vector in W and a vector in $W^{\perp}$.
38. Find a basis for the orthogonal complement of the subspace $W = \text{Span}\{\{x^2+1, x-5\}\}$ in the polynomial inner product space of CYU 7.11, page 293, and express the vector $\boldsymbol{v} = 2x^2 - x$ as a sum of a vector in W and a vector in $W^{\perp}$.
39. Prove that the standard basis $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ of page 94 is an orthonormal basis in the Euclidean inner product space $\Re^n$.
40. Prove that $\{x^n, x^{n-1}, \ldots, x, 1\}$ is an orthonormal basis in the polynomial inner product space of CYU 7.11, page 277.
41. Let V be an inner product space. Prove that $V^{\perp} = \{\mathbf{0}\}$ and that $\{\mathbf{0}\}^{\perp} = V$.
42. Let $W = \text{Span}\{\boldsymbol{w_1}, \boldsymbol{w_2}, \ldots, \boldsymbol{w_m}\}$ in an inner product space V. Prove that $\boldsymbol{v} \in W^{\perp}$ if and only if $\langle \boldsymbol{v}, \boldsymbol{w_i}\rangle = 0$ for all $\boldsymbol{w_i}$, $1 \le i \le m$.
43. Let $\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}, \boldsymbol{v_{k+1}}, \ldots, \boldsymbol{v_m}\}$ be an orthogonal set in an inner product space V. Show that if $\boldsymbol{w} \in \text{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_k}\}$ and $\boldsymbol{z} \in \{\boldsymbol{v_{k+1}}, \ldots, \boldsymbol{v_m}\}$, then $\langle \boldsymbol{w}, \boldsymbol{z}\rangle = 0$.
44. Let W be a subspace in an inner product space V. Prove that $(W^{\perp})^{\perp} = W$.
45. Let $\boldsymbol{w}$ be a vector in an inner product space V of dimension n. Prove that $\boldsymbol{w}^{\perp} = \{\boldsymbol{v} \in V | \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \mathbf{0}\}$ is a subspace of V of dimension $n-1$.
46. Let S be a subset of an inner product space V. Prove that $S^{\perp} = \{\boldsymbol{v} \in V | \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \mathbf{0} \text{ for all } \boldsymbol{w} \in S\}$ is a subspace of V.
47. Let S be a subspace of an inner product space V of dimension n. Prove that $\dim(S) + \dim(S^{\perp}) = n$
48. Let $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v_n}\}$ be an orthonormal basis in an inner product space V. Show that for any $\boldsymbol{u}, \boldsymbol{v} \in V$ $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = [\boldsymbol{u}]_{\beta} \cdot [\boldsymbol{v}]_{\beta}$. (See Definition 5.9, page 178.)

Exercises 51-59 (Orthogonal Matrices) $A \in M_{n \times n}$ is an **orthogonal matrix** if the columns of A is an orthonormal set in the Euclidean inner product space $\Re^n$. (Orthogonal matrices would better have been named "orthonormal matrices," no?)

49. Sow that the following are equivalent:
    (i) $A \in M_{n \times n}$ is orthogonal.
    (ii) $A^TA = I$. (See Exercise 19, page 162).
    (iii) $\|AX\| = \|X\|$ for every $X \in \Re^n$
    (iv) $AX \cdot AY = X \cdot Y$ for every $X, Y \in \Re^n$.
50. Prove that every orthogonal matrix is invertible, and that its inverse is also orthogonal.
51. Prove that a product of orthogonal matrices (or the same dimension) is again orthogonal.
52. Prove that if A is orthogonal, then $\det(A) = \pm 1$.
53. Prove that if A is orthogonal then the rows of A also constitute an orthonormal set.
54. Prove that if A is orthogonal, and if B is equivalent to A, then B is also orthogonal.
55. Prove that every 2×2 orthogonal matrix is of the form $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ or $\begin{bmatrix} a & b \\ b & -a \end{bmatrix}$ where $\sqrt{a^2 + b^2} = 1$.
56. Show that every 2×2 orthogonal matrix is of the form $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ or $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$.
57. Show that every 2×2 orthogonal matrix corresponds to either a rotation or a reflection about a line through the origin in $\Re^2$.
58. (a) Prove that the null space of $A \in M_{m \times n}$ is the orthogonal complement of the row space of A.
    (b) Prove that the null space of $A^T \in M_{m \times n}$ is the orthogonal complement of the column space of A. (See Exercise 19, page 162.)
    (c) Verify directly that the null space of $A = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -1 & 0 & 1 & 2 \end{bmatrix}$ is the orthogonal complement of the row space of A.
    (d) Verify directly that the null space of $A^T = \begin{bmatrix} 1 & 3 & 2 & 0 \\ -1 & 0 & 1 & 2 \end{bmatrix}$ is the orthogonal complement of the column space of A. (See Exercise 19, page 162.)

59. **(Bessel's Equality)** Let $\{v_1, v_2, \ldots, v_n\}$ be an orthonormal basis for an inner product space V. Prove that for any $w \in V$: $\sum_{i=1}^{n} \langle w, v_i \rangle^2 = \|w\|^2$.

PROVE OR GIVE A COUNTEREXAMPLE

60. If $\{v_1, v_2, \ldots, v_m\}$ is an orthogonal set in an inner product space V, then $\{a_1 v_1, a_2 v_2, \ldots, a_m v_m\}$ is an orthogonal set for all scalars $a_1, a_2, \ldots, a_m$.

61. If $\{v_1, v_2, \ldots, v_m\}$ is an orthonormal set in an inner product space V, then $\{a_1 v_1, a_2 v_2, \ldots, a_m v_m\}$ is an orthonormal set for all scalars $a_1, a_2, \ldots, a_m$.

62. Let W be a subspace of an inner product space V. If $\langle w, v \rangle = 0$ with $w \in W$, then $v \in W^\perp$.

63. Let $\{v_1, v_2, \ldots, v_k, \ldots, v_n\}$ be a basis for an inner product space V such that each v_j for $k < j \le n$ is orthogonal to every v_m for $1 \le m \le k$. If $W = \text{Span}\{v_1, v_2, \ldots, v_k\}$, then $W^\perp = \text{Span}\{v_{k+1}, \ldots, v_n\}$.

64. Let $\{v_1, v_2, \ldots, v_k, \ldots, v_n\}$ be an orthogonal basis for an inner product space V. If $W = \text{Span}\{v_1, v_2, \ldots, v_k\}$hen $W^\perp = \text{Span}\{v_{k+1}, \ldots, v_n\}$.

§4. THE SPECTRAL THEOREM

We begin by recalling Definition 6.13 of page 264:

> The **transpose** of a matrix $A = [a_{ij}] \in M_{m \times n}$ is the matrix $A^{\mathrm{T}} = [b_{ij}] \in M_{n \times m}$, where $b_{ij} = a_{ji}$

In other words, the i^{th} row of A is the i^{th} column of A^{T}. For example:

$$\text{If } A = \begin{bmatrix} 1 & 0 & 3 \\ 2 & 4 & 5 \end{bmatrix}, A^{\mathrm{T}} = \begin{bmatrix} 1 & 2 \\ 0 & 4 \\ 3 & 5 \end{bmatrix}$$

DEFINITION 7.12
SYMMETRIC MATRIX

$A \in M_{n \times n}$ is **symmetric** if $A^{\mathrm{T}} = A$.

As it turns out:

THEOREM 7.13 If $A \in M_{n \times n}$ is a (real) symmetric matrix, then its eigenvalues are real numbers.

An outline of a proof for the above theorem, which involves a bit of complex number terminology, is relegated to the exercises.

We remind you that we are using $\overline{\Re^n}$ to denote $M_{n \times 1}$. For $X, Y \in \overline{\Re^n}$ we now define $X \cdot Y$ to be the dot product of the corresponding vertical n-tuples (see margin). It is easy to show that $\overline{\Re^n}$,with $\langle X, Y \rangle$ defined to be $X \cdot Y$, is an inner product space (see Definition 7.6, page 292). Note, that the above dot product can also be effected by means of matrix multiplication:

$$X \cdot Y = X^T Y \text{ (see margin)}$$

$$\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} = (2, 3, 5) \cdot (1, 4, 0)$$
$$= 2 \cdot 1 + 3 \cdot 4 + 5 \cdot 0$$
$$= 14$$

$$\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 4 \\ 0 \end{bmatrix}$$
$$= [2 \cdot 1 + 3 \cdot 4 + 5 \cdot 0]$$
$$= [14] = 14$$

THEOREM 7.14 Any two eigenvectors in the inner product space $\overline{\Re^n}$ corresponding to distinct eigenvalues of a symmetric matrix $A \in M_{n \times n}$ are orthogonal.

PROOF: Let λ_1, λ_2 be distinct eigenvalues of a symmetric matrix $A \in M_{n \times n}$, with corresponding eigenvectors X, Y, so that:

$$AX = \lambda_1 X \text{ and } AY = \lambda_2 Y$$

We show that $X \cdot Y = \mathbf{0}$:

$$\lambda_1(X \cdot Y) = (\lambda_1 X) \cdot Y = (AX)^{\mathrm{T}} Y$$

Exercise 19(f), page 162 $= (X^{\mathrm{T}} A^{\mathrm{T}}) Y$

Definition 7.12: $= (X^{\mathrm{T}} A) Y = X^{\mathrm{T}}(AY)$

$$= X^{\mathrm{T}}(\lambda_2 Y) = \lambda_2(X^{\mathrm{T}} Y)$$
$$= \lambda_2(X \cdot Y)$$

Then: $\lambda_1(X \cdot Y) = \lambda_2(X \cdot Y) \Rightarrow (\lambda_1 - \lambda_2)(X \cdot Y) = \mathbf{0}$

Since $\lambda_1 \neq \lambda_2$, $X \cdot Y = \mathbf{0}$; which is to say: X and Y are orthogonal.

EXAMPLE 7.14 Verify the result of Theorem 7.14 for the symmetric matrix:

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

SOLUTION: To determine the eigenvalues of A, we turn to Theorem 6.8, page 219, and calculate the determinant of $A - \lambda I$:

$$\det(A - \lambda I) = \det\begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix} \underset{\uparrow}{=} -\lambda(\lambda - 1)(\lambda - 3)$$

details omitted

We see that A has three distinct eignevalues: $\lambda_1 = 0$, $\lambda_2 = 1$ and $\lambda_3 = 3$, with corresponding eigenspaces:

Note:

$$\text{rref}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{rref}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\text{rref}\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E[0] = \text{null}(A - 0I) = \text{null}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \underset{\uparrow}{=} \{\overline{(a, a, a)} \mid a \in R\}$$

margin

$$E[1] = \text{null}(A - 1I) = \text{null}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} \overset{\downarrow}{=} \{\overline{(-b, 0, b)} \mid b \in R\}$$

$$E[3] = \text{null}(A - 3I) = \text{null}\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} = \{\overline{(c, -2c, c)} \mid c \in R\}$$

You can easily verify that for $1 \le i < j \le 3$, every vector in $E[\lambda_i]$ is orthogonal to every vector $E[\lambda_j]$. For example:

$$\overline{(a, a, a)} \cdot \overline{(c, -2c, c)} = ac - 2ac + ac = 0$$

CHECK YOUR UNDERSTANDING 7.23

Verify the result of Theorem 7.13 for the matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$.

Answer: See page B-37.

THEOREM 7.15 $A \in M_{n \times n}$ is symmetric **if and only if**

$$(\boldsymbol{AX}) \cdot \boldsymbol{Y} = \boldsymbol{X} \cdot (\boldsymbol{AY})$$

for all vectors $\boldsymbol{X}, \boldsymbol{Y} \in \overline{R^n}$.

PROOF: If A is symmetric, then:

$$(\boldsymbol{AX})\cdot\boldsymbol{Y} = (\boldsymbol{AX})^T\boldsymbol{Y} \underset{\uparrow}{=} (\boldsymbol{X}^T\boldsymbol{A}^T)\boldsymbol{Y} = \boldsymbol{X}^T(\boldsymbol{A}^T\boldsymbol{Y})$$

Exercise 19(f), page 161

symmetry: $= \boldsymbol{X}^T(\boldsymbol{AY}) = \boldsymbol{X}\cdot(\boldsymbol{AY})$

Conversely, assume that $A = [a_{ij}]$ is such that

$$(AX)\cdot Y = X\cdot(AY) \qquad (*)$$

Turning to the n-tuple $\overline{e_k}$, with 1 as its k^{th} entry and 0 elsewhere, we show that A is symmetric, by showing that $a_{ij} = a_{ji}$:

$$a_{ij} = (\boldsymbol{A}\overline{\boldsymbol{e_j}})\cdot\overline{\boldsymbol{e_i}} \text{ (see margin)}$$

By (*): $= \overline{\boldsymbol{e_j}}\cdot(\boldsymbol{A}\overline{\boldsymbol{e_i}})$

Theorem 7.1(ii), page 279: $= (\boldsymbol{A}\overline{\boldsymbol{e_i}})\cdot\overline{\boldsymbol{e_j}} = a_{ji}$

$$\left(\begin{bmatrix}1&3&5\\2&6&7\\4&\mathbf{5}&2\end{bmatrix}\overset{\overset{e_2}{\downarrow}}{\begin{bmatrix}0\\1\\0\end{bmatrix}}\right)\cdot\overset{\overset{e_3}{\downarrow}}{\begin{bmatrix}0\\0\\1\end{bmatrix}} = \begin{bmatrix}3\\6\\5\end{bmatrix}\cdot\begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$= \underset{\overset{\uparrow}{a_{32}}}{5}$$

CHECK YOUR UNDERSTANDING 7.24

(a) Let $A = \begin{bmatrix}1&1&1\\1&2&1\\1&1&3\end{bmatrix}$. Show directly that $(\boldsymbol{AX})\cdot\boldsymbol{Y} = \boldsymbol{X}\cdot(\boldsymbol{AY})$ for every $\boldsymbol{X}, \boldsymbol{Y} \in \overline{\Re^3}$.

(b) Write down an arbitrary non-symmetric matrix $A \in M_{3\times 3}$ and exhibit $\boldsymbol{X}, \boldsymbol{Y} \in \overline{\Re^3}$ for which $(\boldsymbol{AX})\cdot\boldsymbol{Y} \neq \boldsymbol{X}\cdot(\boldsymbol{AY})$.

Answer: See page B-37.

SYMMETRIC OPERATORS

As you know, there is an intimate relation between matrices and linear maps; bringing us to:

DEFINITION 7.13
SYMMETRIC OPERATOR

Let V be an inner product space. A linear operator $T: V \to V$ is symmetric if

$$\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle$$

for all vectors $\boldsymbol{v}, \boldsymbol{w} \in V$.

Compare with Theorem 7.15.

Here is the linear operator version of Theorem 7.14:

THEOREM 7.16 Let $T: V \to V$ be a symmetric linear operator on an inner product space V. If $\boldsymbol{v_1}$ and $\boldsymbol{v}_2$ are eigenvectors associated with distinct eigenvalues λ_1 and λ_2, then $\boldsymbol{v_1}$ and $\boldsymbol{v}_2$ are orthogonal.

PROOF:

$$\begin{array}{rr@{\;=\;}l}
 & \langle T(\mathbf{v_1}), \mathbf{v_2}\rangle - \langle T(\mathbf{v_1}), \mathbf{v_2}\rangle & 0 \\
\text{Definition 7.13:} & \langle T(\mathbf{v_1}), \mathbf{v_2}\rangle - \langle \mathbf{v_1}, T(\mathbf{v_2})\rangle & 0 \\
\text{Definition 6.5, page 224:} & \langle \lambda_1\mathbf{v_1}, \mathbf{v_2}\rangle - \langle \mathbf{v_1}, \lambda_2\mathbf{v_2}\rangle & 0 \\
\text{Theorem 7.4, page 293:} & \lambda_1\langle \mathbf{v_1}, \mathbf{v_2}\rangle - \lambda_2\langle \mathbf{v_1}, \mathbf{v_2}\rangle & 0 \\
 & (\lambda_1 - \lambda_2)\langle \mathbf{v_1}, \mathbf{v_2}\rangle & 0
\end{array}$$

Since $\lambda_1 \neq \lambda_2$, $\langle \mathbf{v_1}, \mathbf{v_2}\rangle = 0$.

The matrix representation $[T]_{\beta\beta}$ of a symmetric linear operator need not be symmetric for every basis β (see Exercise 18). However:

THEOREM 7.17 If $T: V \to V$ is a symmetric linear operator on an inner product space V, then $[T]_{\beta\beta}$ is a symmetric matrix for any orthonormal basis β of V.

PROOF: Employing Theorem 7.15, we show that for any two (column) n-tuples $\bar{\mathbf{v}} = \overline{(v_1, v_2, \dots, v_n)}$, $\bar{\mathbf{w}} = \overline{(w_1, w_2, \dots, w_n)}$ in $\Re^n$

$$([T]_{\beta\beta}\bar{\mathbf{v}}) \cdot \bar{\mathbf{w}} = \bar{\mathbf{v}} \cdot ([T]_{\beta\beta}\bar{\mathbf{w}}):$$

$$\begin{array}{rl}
 & ([T]_{\beta\beta}\bar{\mathbf{v}}) \cdot \bar{\mathbf{w}} = ([T]_{\beta\beta}[\mathbf{v}]_\beta) \cdot [\mathbf{w}]_\beta \\
\text{Theorem 5.22, page 180:} & = [T(\mathbf{v})]_\beta \cdot [\mathbf{w}]_\beta \\
\text{CYU 7.18, page 303:} & = \langle T(\mathbf{v}), \mathbf{w}\rangle \\
\text{By symmetry:} & = \langle \mathbf{v}, T(\mathbf{w})\rangle = [\mathbf{v}]_\beta \cdot [T(\mathbf{w})]_\beta = \bar{\mathbf{v}} \cdot ([T]_{\beta\beta}\bar{w})
\end{array}$$

CHECK YOUR UNDERSTANDING 7.25

For $\Re^3$ the Euclidean inner product space, let $T: \Re^3 \to \Re^3$ be the linear transformation given by:

$$T(a, b, c) = (a - b, -a + 2b - c, -b + c)$$

(a) Show that T is symmetric.

(b) Verify that $[T]_{\beta\beta}$ is symmetric for the orthonormal basis $\beta = \{(0, 1, 0), (0, 0, 1), (1, 0, 0)\}$ of $\Re^3$.

Answer: See page B-38.

Theorem 7.9, page 303, assures us that every finite dimensional inner product space contains an orthonormal basis β. It follows that every symmetric linear operator $T: V \to V$ has a symmetric matrix representation $[T]_{\beta\beta}$. Indeed, β can be chosen so that $[T]_{\beta\beta}$ is a diagonal matrix:

Note that V contains an orthonormal basis if and only if it contains a normal basis.

THEOREM 7.18
THE SPECTRAL THEOREM

A linear operator $T: V \to V$ on an inner product space V is symmetric if and only if V contains an orthonormal basis of eigenvectors of T.

PROOF: Assume that $\beta = \{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots \boldsymbol{v_n}\}$ is an orthonormal basis of eigenvectors of T with corresponding eigenvalues $\{\lambda_1, \lambda_2, \ldots \lambda_n\}$. We show that T is symmetric:

Let $\boldsymbol{v}, \boldsymbol{w} \in V$, with $\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$ and $\boldsymbol{w} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i$. Then:

$$\begin{aligned}\langle T(\boldsymbol{v}), \boldsymbol{w}\rangle &= \langle \sum_{i=1}^{n} a_i T(\boldsymbol{v}_i), \sum_{i=1}^{n} b_i \boldsymbol{v}_i \rangle \\ &= \langle \sum_{i=1}^{n} a_i \lambda \boldsymbol{v_i}, \sum_{i=1}^{n} b_i \boldsymbol{v}_i \rangle \\ &\underset{\uparrow}{=} \sum_{i=1}^{n} \lambda a_i b_i \underset{\uparrow}{=} \langle \sum_{i=1}^{n} a_i \boldsymbol{v_i}, \sum_{i=1}^{n} b_i \lambda \boldsymbol{v}_i \rangle = \langle \boldsymbol{v}, T(\boldsymbol{w})\rangle\end{aligned}$$

Since $\langle \boldsymbol{v}_i, \boldsymbol{v}_j \rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j \end{cases}$ [along with Theorem 7.4, page 293]

To establish the converse, we apply the Principle of Mathematical Induction on the dimension n of V:

I. Let $\dim(V) = 1$. For any $\boldsymbol{v} \neq \boldsymbol{0}$, $\left\{\frac{\boldsymbol{v}}{|\boldsymbol{v}|}\right\}$ is an orthonormal basis for V. Since $T\left(\frac{\boldsymbol{v}}{|\boldsymbol{v}|}\right) \in V$, and since $\left\{\frac{\boldsymbol{v}}{|\boldsymbol{v}|}\right\}$ is a basis for V, $T\left(\frac{\boldsymbol{v}}{|\boldsymbol{v}|}\right) = \lambda \frac{\boldsymbol{v}}{|\boldsymbol{v}|}$ for some λ.

II. Assume that the claim holds for $\dim(V) = k$.

III. We establish validity for $\dim(V) = k + 1$:

Applying the Grahm-Schmidt process, we can obtain an orthonormal basis β for V. By Theorem 7.17, $[T]_{\beta\beta}$ is symmetric. Let λ be a (real) eigenvalue of $[T]_{\beta\beta}$ (Theorem 7.13). Let $\bar{\boldsymbol{v}} \in \overline{\Re^n}$ be an eigenvector associated with λ, and let $\boldsymbol{v}_n$ be such that $[\boldsymbol{v}_n]_\beta = \bar{\boldsymbol{v}}$. Then:

$$[T]_{\beta\beta}\bar{v} = \lambda\bar{v} \Rightarrow [T]_{\beta\beta}[\boldsymbol{v}_n]_\beta = \lambda[\boldsymbol{v}_n]_\beta$$

Theorem 5.21, page 181: $$\Rightarrow [T(\boldsymbol{v}_n)]_\beta = \lambda[\boldsymbol{v}_n]_\beta$$

$$\Rightarrow T(\boldsymbol{v}_n) = \lambda\boldsymbol{v}_n$$

Let $W = \text{Span}\{\boldsymbol{v}_n\}$. Theorem 7.10, page 306 tells us that:

$$W^\perp = \{\boldsymbol{v} \in V | \langle \boldsymbol{v}, \boldsymbol{v}_n \rangle = 0\}$$

is a subspace of V of dimension $k - 1$.

Nothing that for any $\boldsymbol{v} \in W^\perp$:

$$\langle T(\boldsymbol{v}), \boldsymbol{v}_k \rangle = \langle \boldsymbol{v}, T(\boldsymbol{v}_k) \rangle = \langle \boldsymbol{v}, \lambda\boldsymbol{v}_k \rangle = \lambda\langle \boldsymbol{v}, \boldsymbol{v}_k \rangle = 0$$

we conclude that $T(\boldsymbol{v}) \in W^\perp$ for every $\boldsymbol{v} \in W^\perp$.

Let $T_{W^\perp}: W^\perp \to W^\perp$ denote the restriction of T to $W^\perp$. The linearity of T assures us that $T_{W^\perp}$ is linear. Moreover, since $\langle T(\boldsymbol{v}), \boldsymbol{w} \rangle = \langle \boldsymbol{v}, T(\boldsymbol{w}) \rangle$ holds for every $\boldsymbol{v}, \boldsymbol{w} \in V$, it must certainly hold for every $\boldsymbol{v}, \boldsymbol{w} \in W^\perp$. Invoking the induction hypothesis (II) to the symmetric linear operator $T_{W^\perp}: W^\perp \to W^\perp$, we let $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_{k-1}\}$ denote an orthonormal basis of eigenvectors of $T_{W^\perp}$. Since each vector in that basis is orthogonal to the eigenvector $\boldsymbol{v}_k$, the set $\left\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_{k-1}, \frac{\boldsymbol{v}_k}{\|\boldsymbol{v}_k\|}\right\}$ is seen to be an orthonormal basis for V consisting of eigenvectors.

CHECK YOUR UNDERSTANDING 7.26

Find an orthonormal basis of eigenvectors for the symmetric linear operator $T(a, b, c) = (a - b, -a + 2b - c, -b + c)$ of CYU 7.25.

Answer: See page B-38.

MATRIX VERSION OF THE SPECTRAL THEOREM

Here is the link between symmetric matrices and symmetric linear operators:

THEOREM 7.19 $A \in M_{n \times n}$ is symmetric if and only if $T_A: \overline{\Re^n} \to \overline{\Re^n}$ given by $T_A(X) = AX$ is a symmetric linear operator.

Recall that for $X, Y \in \overline{\Re^n}$:
$\langle X, Y\rangle = X \cdot Y$
(see page 307)

PROOF: Assume that $A \in M_{n \times n}$ is symmetric. For any $X, Y \in \overline{\Re^n}$:

$$\langle T_A(X), Y\rangle = T_A(X) \cdot Y = AX \cdot Y$$

$$\text{Theorem 7.15:} \quad = X \cdot AY = X \cdot T_A(Y) = \langle X, T_A(Y)\rangle$$

Conversely, if T_A is symmetric, then:

$$AX \cdot Y = T_A(X) \cdot Y = \langle T_A(X), Y\rangle$$

$$\text{Definition 7.13:} \quad = \langle X, T_A(Y)\rangle = X \cdot T_A(Y) = X \cdot AY$$

DEFINITION 7.14

ORTHOGONAL AND ORTHONORMAL MATRICES

$A \in M_{n \times n}$ is an **orthogonal matrix** if the columns of A constitute an orthogonal set in the Euclidean inner product space $\Re^n$.

A is an **orthonormal matrix** if its columns constitute an orthonormal set in $\Re^n$.

THEOREM 7.20

Every orthogonal matrix is invertible.

PROOF: If $A \in M_{n \times n}$ is orthogonal, then the columns of A constitute a linearly independent set of vectors in $\Re^n$ (Theorem 7.7, page 301), and are therefore a basis for $\Re^n$ (Theorem 3.11, page 99). The result now follows from Exercise 37, page 176.

Yes, every orthogonal matrix is invertible; but more can be said for orthonormal matrices:

THEOREM 7.21

$A \in M_{n \times n}$ is orthonormal **if and only if**

$$A^{-1} = A^{\mathrm{T}}$$

PROOF: Let $A = [a_{ij}]$, $A^{\mathrm{T}} = [\bar{a}_{ij}]$, and $AA^{\mathrm{T}} = [c_{ij}]$. Then:

$$c_{ij} = \sum_{\alpha=1}^{n} a_{i\alpha}\bar{a}_{\alpha j} = \sum_{\alpha=1}^{n} a_{i\alpha}a_{j\alpha} \uparrow$$

the dot product of the i^{th} column of A with the jth column of A

It follows that $A^{-1} = A^{\mathrm{T}}$, if and only if $AA^{\mathrm{T}} = I$, if and only if

$$\sum_{\alpha=1}^{n} a_{i\alpha}a_{j\alpha} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

if and only if the columns of A constitute an orthonormal set in $\Re^n$.

CHECK YOUR UNDERSTANDING 7.27

Prove that the product of any two orthonormal matrices in $M_{n \times n}$ is again orthonormal.

Answer: See page B-39.

Note: In the literature the term ***orthogonally diagonalizable*** is typically used to refer to what we are calling .

DEFINITION 7.15
ORTHONORMALLY DIAGONALIZABLE

$A \in M_{n \times n}$ is if there exists an orthonormal matrix P and a diagonal matrix D such that:

$$P^{-1}AP = D$$

As it turns out:

THEOREM 7.22
THE SPECTRAL THEOREM

$A \in M_{n \times n}$ is orthonormally diagonalizable if and only if it is symmetric.

PROOF: If $A \in M_{n \times n}$ is symmetric, then $T_A\colon \Re^n \to \Re^n$ given by $T_A(\boldsymbol{X}) = A\boldsymbol{X}$ is a symmetric operator (Theorem 7.19). Employing Theorem 7.18, we chose an orthonormal basis $\beta = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n\}$ of eigenvectors of T_A with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. For $S = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ the standard basis of $\Re^n$ we have:

See Theorem 5.27, page 194)

$$[T_A]_{\beta\beta} = [I]_{\beta S}[T_A]_{SS}[I]_{S\beta} \text{ (*)}$$

We now show that:

(1) $[T_A]_{\beta\beta}$ is a diagonal matrix with the $\lambda_i{}'s$ along its diagonal.

(2) $[T_A]_{SS} = A$

(3)The columns of $[I]_{S\beta}$ are the $\boldsymbol{X}_i{}'s$ — an orthonormal set.

(1), (2), (3) and (*) tell us that:
$P^{-1}AP$ is a diagonal matrix, with $P = [I]_{S\beta}$ an orthonormal matrix.
In particular:
A is !
Moreover:
$P^{-1}AP = P^{\mathrm{T}}AP$
with X_i the i^{th} column of P.

(1): A consequence of Definition 5.10, page 180, and the fact that $T_A(X_i) = AX_i = \lambda_i X_i$.

(2): A consequence of Definition 5.10 and the fact that $[T_A(e_i)]_S$ is the i^{th} column of A.

(3): A consequence of Definition 5.10 and the fact that $[I(X_i)]_S = [X_i]_S = X_i$.

Conversely, assume that A is orthogonally diagonalizable. Let P be an orthogonal matrix and D a diagonal matrix with:

$$P^{-1}AP = D$$

$$A = PDP^{-1}$$

Theorem 7.21: $A = PDP^T$

By Then:

$$A^T = (PDP^T)^T$$

Exercise 24(f), page 1643: $= (P^T)^T D^T P^T$

Exercise 24(a), page 164: $= PD^TP^{-1} \underset{\uparrow}{=} PDP^{-1} = A$

Since every diagonal matrix is symmetric

Since $A^T = A$, A is symmetric.

CHECK YOUR UNDERSTANDING 7.28

Find an orthonormal diagonalization for the symmetric matrix:

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

Answer: See page B-39.

EXERCISES

Exercises 1-4. Verify that the given matrix is orthonormal.

1. $\begin{bmatrix} -\frac{\sqrt{3}}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$
2. $\begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$
3. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$
4. $\begin{bmatrix} \frac{3}{7} & -\frac{2}{7} & \frac{6}{7} \\ -\frac{2}{7} & \frac{6}{7} & \frac{3}{7} \\ \frac{6}{7} & -\frac{2}{7} & \frac{3}{7} \end{bmatrix}$

Exercises 5-8. Verify that the given matrix $A \in M_{n \times n}$ is symmetric. Show directly that $(A\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot (A\boldsymbol{w})$ for every $\boldsymbol{v}, \boldsymbol{w} \in \Re^n$.

5. $\begin{bmatrix} 5 & 1 \\ 1 & 5 \end{bmatrix}$
6. $\begin{bmatrix} 7 & -3 \\ -3 & 4 \end{bmatrix}$
7. $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{bmatrix}$
8. $\begin{bmatrix} 2 & -1 & 0 \\ -1 & 3 & 1 \\ 0 & 1 & 2 \end{bmatrix}$

Exercises 1-4. Find an orthonormal diagonalization for the symmetric matrix of:

9. Exercise 5.
10. Exercise 6.
11. Exercise 67
12. Exercise 8.

Exercises 9-12. Verify that the given linear operator $T: \Re^n \to \Re^n$ on the Euclidean (dot product) inner product space R^n is symmetric. Determine $[T]_{S_n, S_n}$ where S_n denotes the standard basis in $\Re^n$.

13. $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{2a + b, a + 2b})$
14. $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{-a + 3b, 3a + 5b})$
15. $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a + 2b + 3c, 2a + b, 3a + 2c})$
16. $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = (\boldsymbol{a + 2b + c, 2a + 2b, a + 3c})$

17. Verify that $T(\boldsymbol{a}, \boldsymbol{b}) = (\boldsymbol{3a, 2a + b})$ is a symmetric operator on the weighted inner product space R^2 with $\langle(\boldsymbol{a}, \boldsymbol{b}), (\boldsymbol{c}, \boldsymbol{d})\rangle = 4ac + bd$. Verify that $\beta = \left\{\left(\boldsymbol{\frac{1}{2}, 0}\right), (\boldsymbol{0, 1})\right\}$ is an orthonormal basis in this inner product space, and determine $[T]_{\beta, \beta}$.

18. (a) Verify that $T(ax^2 + bx + c) = (a + 2b + c)x^2 + (2ab)x + (3a + 2c)$ is a symmetric operator on the standard inner product space P_2: $[(ax^2 + bx + c) \cdot (ax^2 + bx + c) = a\bar{a} + b\bar{b} + c\bar{c}]$.
(b) Use the Grahm-Schmidt process of page 303 on the basis $\beta = \{x^2 + x + 1, x + 1, x^2 + 1\}$ to arrive at the orthonormal basis $\bar{\beta} = \left\{\frac{1}{\sqrt{3}}x^2 + \frac{1}{\sqrt{3}}x + \frac{1}{\sqrt{3}}, -x^2 + \frac{1}{x} + \frac{1}{2}, -\frac{1}{4}x + \frac{1}{4}\right\}$. Verify that $[T]_{\beta\beta}$ is not symmetric, and that $[T]_{\bar{\beta}\bar{\beta}}$ is symmetric.

19. Let $\Re^2$ denote the standard Euclidean dot product inner product space. Find a symmetric linear operator $T\colon \Re^2 \to \Re^2$ and a basis β for which $[T]_{\beta\beta} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$.

20. Let $\Re^2$ denote the weighted inner product space R^2 with $\langle (a, b), (c, d) \rangle = 3ac + 2bd$. Find a symmetric linear operator $T\colon \Re^2 \to \Re^2$ and a basis β for which $[T]_{\beta\beta} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$.

21. Let P_1 denote the standard inner product space P_1: $[(ax + b) \cdot (\bar{a}x + \bar{b}) = a\bar{a} + b\bar{b} + c\bar{c}]$. Find a symmetric linear operator $T\colon P_1 \to P_1$ and a basis β for which $[T]_{\beta\beta} = \begin{bmatrix} 3 & 2 \\ 2 & 1 \end{bmatrix}$

22. Show that for any $\boldsymbol{A} \in M_{n \times n}$ both $\boldsymbol{A} + \boldsymbol{A^T}$ and $\boldsymbol{AA^T}$ are symmetric.

23. Show that if $\boldsymbol{A}, \boldsymbol{B} \in M_{n \times n}$ are orthonormally diagonalizable, then so is:

    (a) $c\boldsymbol{A}$ for every $c \in \Re$. (b) $\boldsymbol{A} + \boldsymbol{B}$ (c) $\boldsymbol{A}^2$

24. (PMI) Show that if $\boldsymbol{A} \in M_{m \times m}$ is orthonormally diagonalizable, then so is A^n for any positive integer *n*.

25. (PMI) Show that if $\boldsymbol{A_i} \in M_{m \times m}$ is orthonormally diagonalizable for $1 \le i \le n$, then so is $\boldsymbol{A_1} + \boldsymbol{A}_2 + \cdots + \boldsymbol{A}_n$.

26. Show that if $\boldsymbol{A} \in M_{m \times m}$ is an invertible orthonormally diagonalizable matrix, then so is $\boldsymbol{A}^{-1}$.

27. Prove that if $\boldsymbol{A}$ is a real symmetric matrix, then the eigenvalues of A are real.
    Suggestion: For $\boldsymbol{Av} = \lambda\boldsymbol{v}$, show that $\boldsymbol{A\bar{v}} = \bar{\lambda}\boldsymbol{\bar{v}}$, where here $\bar{\lambda}$ denotes the (complex) conjugate or λ and $\boldsymbol{\bar{v}}$ is the *n*-tuple obtained by taking the conjugate of each entry in the *n*-tuple $\boldsymbol{v}$. Proceed to show that $\lambda(\boldsymbol{\bar{v}}^T\boldsymbol{v}) = \bar{\lambda}(\boldsymbol{v}^T\boldsymbol{v})$.

PROVE OR GIVE A COUNTEREXAMPLE

28. If $\boldsymbol{A} \in M_{n \times n}$ is a symmetric matrices, then so is $\boldsymbol{A}^T$.

29. If $\boldsymbol{A} \in M_{n \times n}$ is a symmetric matrices, then so is $\boldsymbol{A}^{-1}$.

30. If $\boldsymbol{A}, \boldsymbol{B} \in M_{n \times n}$ are symmetric matrices, then so is $\boldsymbol{A} + \boldsymbol{B}$.

31. If $\boldsymbol{A}, \boldsymbol{B} \in M_{n \times n}$ are symmetric matrices, then so is $\boldsymbol{AB}$.

32. If $\boldsymbol{A}, \boldsymbol{B} \in M_{n \times n}$ are orthonormally diagonalizable, then so is $\boldsymbol{AB}$.

33. If $\boldsymbol{A} \in M_{n \times n}$ is orthonormally diagonalizable, then so is A^T.

34. If $\boldsymbol{A} \in M_{n \times n}$ is orthonormally diagonalizable, then so is A^{-1}.

35. Let V be an inner product space. If T: $V \to V$ is a symmetric operator, then so is cT for every $c \in \Re$.

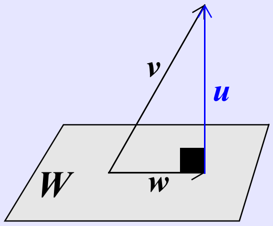
36. Let V be an inner product space. If T: $V \to V$ and L: $V \to V$ are symmetric operators, then so is $T + L$.

37. Let V be an inner product space. If T: $V \to V$ and L: $V \to V$ are symmetric operators, then so is $L \circ T$.

CHAPTER SUMMARY

DOT PRODUCT	The **dot product** of $\boldsymbol{u} = (u_1, u_2, \ldots, u_n)$ and $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$, denoted by $\boldsymbol{u} \cdot \boldsymbol{v}$, is the real number: $\boldsymbol{u} \cdot \boldsymbol{v} = u_1v_1 + u_2v_2 + \ldots + u_nv_n$
PROPERTIES	Let $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in \Re^n$, and $r \in \Re$. Then:
positive-definite property:	(i) $\boldsymbol{v} \cdot \boldsymbol{v} \geq 0$, and $\boldsymbol{v} \cdot \boldsymbol{v} = 0$ only if $\boldsymbol{v} = \boldsymbol{0}$
commutative property:	(ii) $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$
homogeneous property:	(iii) $r\boldsymbol{u} \cdot \boldsymbol{v} = r(\boldsymbol{u} \cdot \boldsymbol{v})$
distributive property:	(iv) $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$
NORM IN $\Re^n$	The **norm** of a vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_n)$, denoted by $\lVert\boldsymbol{v}\rVert$, is given by: $\lVert\boldsymbol{v}\rVert = \sqrt{\boldsymbol{v} \cdot \boldsymbol{v}}$ Denotes length of vector.
ANGLE BETWEEN VECTORS	The **angle** θ between two nonzero vectors $\boldsymbol{u}, \boldsymbol{v} \in \Re^n$ is given by: $\theta = \cos^{-1}\left(\frac{\boldsymbol{u} \cdot \boldsymbol{v}}{\lVert\boldsymbol{u}\rVert\,\lVert\boldsymbol{v}\rVert}\right)$
ORTHOGONAL VECTORS	Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in $\Re^n$ are **orthogonal** if $\boldsymbol{u} \cdot \boldsymbol{v} = 0$.
VECTOR DECOMPOSITION	Let $\boldsymbol{v} \in \Re^n$ and let $\boldsymbol{u}$ be any nonzero vector in $\Re^n$. Then: $\boldsymbol{v} = (\boldsymbol{v} - \operatorname{proj}_{\boldsymbol{u}}\boldsymbol{v}) + \operatorname{proj}_{\boldsymbol{u}}\boldsymbol{v}$ where: $\operatorname{proj}_{\boldsymbol{u}}\boldsymbol{v} = \left(\frac{\boldsymbol{v} \cdot \boldsymbol{u}}{\boldsymbol{u} \cdot \boldsymbol{u}}\right)\boldsymbol{u}$ and $(\boldsymbol{v} - \operatorname{proj}_{\boldsymbol{u}}\boldsymbol{v}) \cdot \operatorname{proj}_{\boldsymbol{u}}\boldsymbol{v} = 0$
INNER PRODUCT SPACE	An **inner product** on a vector space V is an operator which assigns to any two vectors, $\boldsymbol{u}$ and $\boldsymbol{v}$ in V, a real number $\langle \boldsymbol{u}, \boldsymbol{v} \rangle$, satisfying the following four axioms:
positive-definite axiom:	(i) $\langle \boldsymbol{v}, \boldsymbol{v} \rangle \geq 0$, and $\langle \boldsymbol{v}, \boldsymbol{v} \rangle = 0$ only if $\boldsymbol{v} = \boldsymbol{0}$
commutative axiom:	(ii) $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle$
homogeneous axiom:	(iii) $\langle r\boldsymbol{u}, \boldsymbol{v} \rangle = r\langle \boldsymbol{u}, \boldsymbol{v} \rangle$
distributive axiom:	(iv) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \boldsymbol{w} \rangle + \langle \boldsymbol{v}, \boldsymbol{w} \rangle$

PROPERTIES	For every $\boldsymbol{u}$, $\boldsymbol{v}$, and $\boldsymbol{w}$ in an inner product space V: (a) $\langle \boldsymbol{0}, \boldsymbol{v}\rangle = \langle \boldsymbol{v}, \boldsymbol{0}\rangle = 0$ (b) $\langle \boldsymbol{u}, \boldsymbol{v}+\boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{w}\rangle$ (c) $\langle \boldsymbol{u}, r\boldsymbol{v}\rangle = r\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ (d) $\langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle - \langle \boldsymbol{v}, \boldsymbol{w}\rangle$ (e) $\langle \boldsymbol{u}, \boldsymbol{v}-\boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{u}, \boldsymbol{w}\rangle$ (f) $\langle -\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, -\boldsymbol{w}\rangle = -\langle \boldsymbol{v}, \boldsymbol{w}\rangle$
NORM AND DISTANCE	The **norm** (or magnitude) of a vector $\boldsymbol{v}$ in an inner product space V, denoted by $\lVert\boldsymbol{v}\rVert$, is given by: $\lVert\boldsymbol{v}\rVert = \sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle}$ The **distance** between two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in V is given by $\lVert\boldsymbol{u}-\boldsymbol{v}\rVert$.
CAUCHY-SCHWARZ INEQUALITY	For any two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space: $\lvert\langle \boldsymbol{u}, \boldsymbol{v}\rangle\rvert \le \lVert\boldsymbol{u}\rVert\lVert\boldsymbol{v}\rVert$
PROPERTIES	Let V be an inner product space. For all $\boldsymbol{u}, \boldsymbol{v} \in V$ and $r \in \Re$: (a) $\lVert\boldsymbol{v}\rVert \ge 0$, and $\lVert\boldsymbol{v}\rVert = 0$ if and only if $\boldsymbol{v} = \boldsymbol{0}$ (b) $\lVert r\boldsymbol{v}\rVert = \lvert r\rvert\lVert\boldsymbol{v}\rVert$ (c) $\lVert\boldsymbol{u}+\boldsymbol{v}\rVert \le \lVert\boldsymbol{u}\rVert + \lVert\boldsymbol{v}\rVert$
ANGLE BETWEEN VECTORS	The angle θ between two nonzero vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space is given by: $\theta = \cos^{-1}\left(\frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}{\lVert\boldsymbol{u}\rVert\lVert\boldsymbol{v}\rVert}\right)$
ORTHOGONAL VECTORS **ORTHOGONAL SET**	Two vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ in an inner product space V are **orthogonal** if $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$. A set S of vectors in an inner product space V is an **orthogonal set** if $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$ for every $\boldsymbol{u}, \boldsymbol{v} \in S$, with $\boldsymbol{u} \ne \boldsymbol{v}$.
THEOREM	If $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is an orthogonal set of non-zero vectors in an inner product space V, then $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ is a linearly independent set in V.
UNIT VECTOR	A **unit vector** in an inner product space is a vector $\boldsymbol{v}$ of magnitude 1.
ORTHONORMAL SET	An **orthonormal** set of vectors in an inner product space is an orthogonal set of unit vectors.

THEOREM	If $\beta = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ is an orthonormal basis in an inner product space V, then, for any $\boldsymbol{v} \in V$: $\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{v}_1\rangle \boldsymbol{v}_1 + \langle \boldsymbol{v}, \boldsymbol{v}_2\rangle \boldsymbol{v}_2 + \dots + \langle \boldsymbol{v}, \boldsymbol{v}_n\rangle \boldsymbol{v}_n$
GRAHM-SCHMIDT PROCESS	An algorithm (page 303) for generating an orthogonal base in any finite dimensional inner product space.
ORTHOGONAL COMPLEMENT	The **orthogonal complement** of a subspace W of an inner product space: $W^{\perp} = \{\boldsymbol{u} \in V \mid \langle \boldsymbol{u}, \boldsymbol{w}\rangle = 0 \text{ for every } \boldsymbol{w} \in W\}$
PROPERTIES	If W a subspace on an inner product space V, then: (i) $W^{\perp}$ is a subspace of V. (ii) $W \cap W^{\perp} = \{\mathbf{0}\}$ (iii) Every vector in V can be **uniquely** expressed as a sum of a vector in W and a vector in $W^{\perp}$. (iv) If β_W is a basis for W and $\beta_{W^{\perp}}$ is a basis for $W^{\perp}$, then $\beta_W \cup \beta_{W^{\perp}}$ is a basis for V.
VECTOR DECOMPOSITION  	Let $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_m\}$ be an orthonormal basis for a subspace W of an inner product space V, and let $\boldsymbol{v} \in V$. Then, there exists a unique vector $\boldsymbol{w} \in W$ and $\boldsymbol{u} \in W^{\perp}$ such that: $\boldsymbol{v} = \boldsymbol{w} + \boldsymbol{u}$ where: $\boldsymbol{w} = \langle \boldsymbol{v}, \boldsymbol{w}_1\rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2\rangle \boldsymbol{w}_2 + \dots + \langle \boldsymbol{v}, \boldsymbol{w}_m\rangle \boldsymbol{w}_m$ and: $\boldsymbol{u} = \boldsymbol{v} - \boldsymbol{w}$.
SYMMETRIC MATRIX	$A \in M_{n \times n}$ is **symmetric** if $A^{\mathrm{T}} = A$.
THEOREMS	If $A \in M_{n \times n}$ is a (real) symmetric matrix, then its eigenvalues are real. Any two eigenvectors in the inner product space $\Re^n$ corresponding to distinct eigenvalues of a symmetric matrix $A \in M_{n \times n}$ are orthogonal. $A \in M_{n \times n}$ is symmetric **if and only if** $(A\boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{v} \cdot (A\boldsymbol{w})$ for all vectors $\boldsymbol{v}, \boldsymbol{w} \in R^n$ (in column form).

SYMMETRIC OPERATOR	Let V be an inner product space. A linear operator $T: V \to V$ is symmetric if $\langle T(\boldsymbol{v}), \boldsymbol{w} \rangle = \langle \boldsymbol{v}, T(\boldsymbol{w}) \rangle$ for all vectors $\boldsymbol{v}, \boldsymbol{w} \in V$.
THEOREMS	Let $T: V \to V$ be a symmetric linear operator on an inner product space V. If $\boldsymbol{v_1}$ and $\boldsymbol{v}_2$ are eigenvectors associated with distinct eigenvalues λ_1 and λ_2, then $\boldsymbol{v_1}$ and $\boldsymbol{v}_2$ are orthogonal. If $T: V \to V$ is a symmetric linear operator on an inner product space V, then $[T]_{\beta\beta}$ is a symmetric matrix for any orthonormal basis $\beta = \{\boldsymbol{o_1}, \boldsymbol{o_2}, \ldots \boldsymbol{o_n}\}$ of V.
SPECTRAL THEOREM	(Linear Operator) A linear operator $T: V \to V$ on an inner product space V is symmetric if and only if V contains an orthonormal basis of eigenvectors of T. (Matrix) $A \in M_{n \times n}$ is symmetric if and only if there exists a diagonal matrix D and a matrix P with columns an orthonormal set in $\Re^n$ such that $P^{-1}AP = D$.

APPENDIX A

PRINCIPLE OF MATHEMATICAL INDUCTION

We introduce a most powerful mathematical tool, the Principle of Mathematical Induction (*PMI*). Here is how it works:

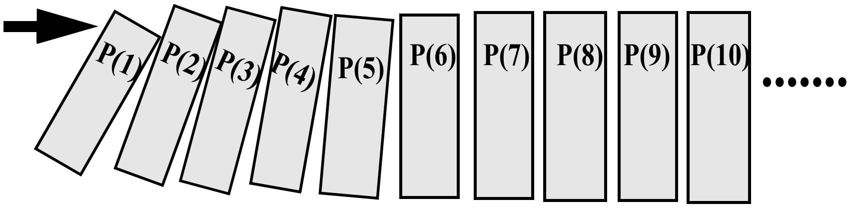
(*PMI*)

Let $P(n)$ denote a proposition that is either true or false, depending on the value of the integer n.

If:	I. $P(1)$ is True.
And if, from the **assumption** that:	II. $P(k)$ is True
one can show that:	III. $P(k+1)$ is also True.

then the proposition $P(n)$ is valid **for all integers** $n \geq 1$

Step II of the induction procedure may strike you as being a bit strange. After all, if one can assume that the proposition is valid at $n = k$, why not just assume that it is valid at $n = k + 1$ and be done with it? Well, you can assume whatever you want in Step II, but if the proposition is not valid for all n you simply are not going to be able to demonstrate, in Step III, that the proposition holds at the next value of n. Its sort of like the domino theory. Just imagine that the propositions $P(1), P(2), P(3), \ldots, P(k), P(k+1), \ldots$ are lined up, as if they were an infinite set of dominoes:



If you knock over the first domino (Step I), and if when a domino falls (Step II) it knocks down the next one (Step III), then all of the dominoes will surely fall. But if the falling k^{th} domino fails to knock over the next one, then all the dominoes will not fall.

To illustrate how the process works, we ask you to consider the sum of the first n odd integers, for $n = 1$ through $n = 5$:

The Principle of Mathematical Induction might have been better labeled a Principle of Mathematical Deduction; for: Inductive reasoning is a process used to formulate a hypotheses or conjecture, while deductive reasoning is a process used to rigorously establish whether or not the conjecture is valid.

n	Sum of the first n odd integers	Sum
1	1	1
2	1 + 3	4
3	1 + 3 + 5	9
4	1 + 3 + 5 + 7	16
5	1 + 3 + 5 + 7 + 9	25

n	Sum
1	1
2	4
3	9
4	16
5	25
6	?

Figure 1.1

Looking at the pattern of the table on the right in Figure 1.1, you can probably anticipate that the sum of the first 6 odd integers will turn out to be $6^2 = 36$, which is indeed the case. In general, the pattern certainly suggests that the sum of the first n odd integers is n^2; a fact that we now establish using the Principle of Mathematical Induction.

Let $P(n)$ be the proposition that the sum of the first n odd integers equals n^2.

I. Since the sum of the first 1 odd integers is 1^2, $P(1)$ is true.

II. **Assume** $P(k)$ is true; that is: $\mathbf{1+3+5+\cdots+}\underset{\text{see margin}}{\mathbf{(2k-1)}} = \mathbf{k^2}$

III. We show that $P(k+1)$ is true, thereby completing the proof:

$$\overbrace{\underbrace{[\mathbf{1+3+5+\cdots+(2k-1)}]}_{\text{induction hypothesis: Step II}}+(\mathbf{2k+1})}^{\text{the sum of the first } k+1 \text{ odd integers}} = \mathbf{k^2}+(\mathbf{2k+1}) = (k+1)^2$$

The sum of the first **3** odd integers is:

$$1+3+5 \longleftarrow \boxed{2\cdot\mathbf{3}-1}$$

The sum of the first **4** odd integers is:

$$1+3+5+7 \longleftarrow \boxed{2\cdot\mathbf{4}-1}$$

Suggesting that the sum of the first $\boldsymbol{k}$ odd integers is:

$$1+3+\ldots+\boxed{(2k-1)}$$

(see Exercise 1).

EXAMPLE 1.1 Use the Principle of Mathematical Induction to establish the following formula for the sum of the first n integers:

$$1+2+3+\ldots+n = \frac{n(n+1)}{2} \quad (*)$$

SOLUTION: Let $P(n)$ be the proposition:

$$1+2+3+\ldots+n = \frac{n(n+1)}{2}$$

I. $P(1)$ is true: $1 = \frac{1(1+1)}{2}$ Check!

II. **Assume** $P(k)$ is true: $\mathbf{1+2+3+\ldots+k = \dfrac{k(k+1)}{2}}$

III. We are to show that $P(k+1)$ is true; which is to say, that (*) holds when $n = k+1$:

$$1+2+3+\ldots+k+(k+1) = \frac{(k+1)[(k+1)+1]}{2} = \frac{(k+1)(k+2)}{2}$$

Let's do it:

$$1+2+3+\ldots+k+(k+1) = [\mathbf{1+2+3+\cdots+k}]+(\mathbf{k+1})$$

induction hypothesis: $$= \frac{\mathbf{k(k+1)}}{\mathbf{2}}+(\mathbf{k+1})$$

$$= \frac{k(\mathbf{k+1})+2(\mathbf{k+1})}{2} = \frac{(\mathbf{k+1})(k+2)}{2}$$

The "domino effect" of the Principle of Mathematical Induction need not start by knocking down the first domino $P(1)$. Consider the following example where domino $P(0)$ is the first to fall.

EXAMPLE 1.2 Use the Principle of Mathematical Induction to establish the inequality $n < 2^n$ for all $n \geq 0$.

SOLUTION: Let $P(n)$ be the proposition $n < 2^n$.

I. $P(0)$ is true: $0 < 2^0$, since $2^0 = 1$.

II. **Assume** $P(k)$ is true: $\boldsymbol{k < 2^k}$

III. We show $P(k+1)$ is true:

$$\boldsymbol{k} + 1 \underset{\text{II}}{<} \boldsymbol{2^k} + 1 \underset{1 \leq 2^k}{\leq} 2^k + 2^k = 2(2^k) = 2^{k+1}$$

III: We need to show that $n < 2^n$ holds for $n = k+1$; which is to say, that: $k+1 < 2^{k+1}$:

Recall that:.

$$n! = 1 \cdot 2 \cdot \cdots \cdot n$$

EXAMPLE 1.3 Use the Principle of Mathematical Induction to show that $n! > n^2$ for all integers $n \geq 4$.

SOLUTION: Let $P(n)$ be the proposition $n! > n^2$:

I. $P(4)$ is true: $4! = 1 \cdot 2 \cdot 3 \cdot 4 = 24 \geq 4^2$.

II. **Assume** $P(k)$ is true: $\boldsymbol{k! > k^2}$ (for $k \geq 4$)

III. We show $P(k+1)$ is true; namely, that $(k+1)! > (k+1)^2$:

$$(k+1)! = \boldsymbol{k!}(k+1) \underset{\text{II}}{>} \boldsymbol{k^2}(k+1)$$

Now what? Well, if we can show that $k^2(k+1) > (k+1)^2$, then we will be done. Let's do it:

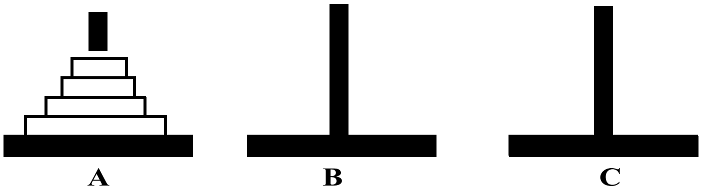
Since $k \geq 4$ (all we need here is that $k \geq 2$):

$$k^2 > k+1$$

Multiplying both sides by the positive number $(k+1)$:

$$k^2(k+1) > (k+1)^2.$$

Our next application of the Principle of Mathematical Induction involves the following **Tower of Hanoi** puzzle:

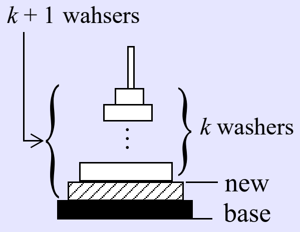
> Start with a number of washers of differing sizes on spindle A, as is depicted below:
>
> 
>
>
> The objective of the game is to transfer the arrangement currently on spindle *A* to one of the other two spindles. The rules are that you may only move one washer at a time, without ever placing a larger disk on top of a smaller one.

EXAMPLE 1.4 Show that the tower of Hanoi game is winnable for any number *n* of washers.

SOLUTION: If spindle A contains one washer, then simply move that washer to spindle B to win the game (Step I).

Assume that the game can be won if spindle A contains k washers (Step II —the induction hypothesis).

We now show that the game can be won if spindle A contains $k+1$ washers (Step III):



> Just imagine that the largest bottom washer is part of the base of spindle A. With this sleight of hand, we are looking at a situation consisting of k washers on a modified spindle A (see margin). By the induction hypothesis, we can move those k washers onto spindle B. We now take the only washer remaining on spindle A (the largest of the original $k+1$ washers), and move it to spindle C, and then think of it as being part of the base of that spindle. Applying the induction hypotheses one more time, we move the k washers from spindle B onto the modified spindle C, thereby winning the game.

APPENDIX B
CHECK YOUR UNDERSTANDING SOLUTIONS

CHAPTER 1
MATRICES AND SYSTEMS OF LINEAR EQUATIONS

CYU 1.1
$$\left[\begin{array}{ccc|c}1&0&7&13\\0&1&-3&-4\\0&0&\mathbf{15}&24\end{array}\right]\xrightarrow{\frac{1}{15}R_1\to R_3}\left[\begin{array}{ccc|c}1&0&7&13\\0&1&-3&-4\\0&0&1&\frac{8}{5}\end{array}\right]\xrightarrow{3R_3+R_2\to R_2}\left[\begin{array}{ccc|c}1&0&7&13\\0&1&0&\frac{4}{5}\\0&0&1&\frac{8}{5}\end{array}\right]\xrightarrow{-7R_3+R_1\to R_1}\left[\begin{array}{ccc|c}1&0&0&\frac{9}{5}\\0&1&0&\frac{4}{5}\\0&0&1&\frac{8}{5}\end{array}\right]$$

CYU 1.2 (a) Yes (b) No [fails (ii)] (c) Yes (d) Yes

CYU 1.3
$$\left.\begin{aligned}x+y+z&=6\\3x+2y-z&=4\\3x+y+2z&=11\end{aligned}\right\}\leftrightarrow\begin{array}{c}\begin{array}{ccc}x&y&z\end{array}\\\left[\begin{array}{ccc|c}1&1&1&6\\3&2&-1&4\\3&1&2&11\end{array}\right]\end{array}\xrightarrow{\text{rref}}\begin{array}{c}\begin{array}{ccc}x&y&z\end{array}\\\left[\begin{array}{ccc|c}1&0&0&1\\0&1&0&2\\0&0&1&3\end{array}\right]\end{array}\leftrightarrow\begin{aligned}x&=1\\y&=2\\z&=3\end{aligned}$$

CYU 1.4 (a) Inconsistent: The last row $\left[\begin{array}{ccc|c}0&0&0&2\end{array}\right]$ corresponds to the equation $0x+0y+0z=2$ which clearly has no solution.

free variable

(b)
$$\begin{array}{c}\begin{array}{ccc}x&y&z\end{array}\\\left[\begin{array}{ccc|c}\boxed{1}&0&-2&1\\0&\boxed{1}&5&4\\0&0&0&0\end{array}\right]\end{array}\leftrightarrow\left.\begin{aligned}x+0y-2z&=1\\0x+y+5z&=4\end{aligned}\right\}\longrightarrow\left.\begin{aligned}x&=1+2z\\y&=4-5z\end{aligned}\right\}:\ \{(1+2r,4-5r,r)\,|\,r\in\Re\}$$

free variables

(c)
$$\left[\begin{array}{cccc|c}\boxed{1}&2&0&1&1\\0&0&\boxed{1}&4&-2\\0&0&0&0&0\end{array}\right]\leftrightarrow\left.\begin{aligned}x+2y+0z+w&=1\\0x+0y+z+4w&=-2\end{aligned}\right\}\longrightarrow\left.\begin{aligned}x&=1-2y-w\\z&=-2-4w\end{aligned}\right\}:$$

$$\{1-2r-s,r,-2-4s,s\,|\,r,s\in R\}$$

CYU 1.5

(a)

$$\left.\begin{aligned}4x-2y+z&=a\\-2x+4y+2z&=b\\5x-y+4z&=c\end{aligned}\right\}\leftrightarrow\left[\begin{array}{ccc|c}4&-2&1&a\\-2&4&2&b\\5&-1&4&c\end{array}\right]\xrightarrow{\text{rref}}\left[\begin{array}{ccc|c}1&0&0&a+\frac{7}{18}b-\frac{4}{9}c\\0&1&0&a+\frac{11}{18}-\frac{5}{9c}\\0&0&1&-a-\frac{1}{3}b+\frac{2}{3}c\end{array}\right]:\quad\begin{aligned}x&=a+\frac{7}{18}b-\frac{4}{9}c\\y&=a+\frac{11}{18}-\frac{5}{9c}\\z&=-a-\frac{1}{3}b+\frac{2}{3}c\end{aligned}$$

Solution for all a, b, and c

(b)

$$\left.\begin{aligned}x-4y-4z&=a\\2x+8y-12z&=b\\-x+12y+2z&=c\end{aligned}\right\}\leftrightarrow\begin{bmatrix}1&-4&-4&a\\2&8&-12&b\\-1&12&2&c\end{bmatrix}\xrightarrow{\text{after two cycles}}\begin{bmatrix}1&0&-5&a+\dfrac{b-2a}{4}\\0&1&-\frac{1}{4}&\dfrac{b-2a}{16}\\0&0&0&c+3a+b\end{bmatrix}$$

The system is consistent if and only if $c+3a+b=0$

CYU 1.6

(a) $\boldsymbol{S}$:

$$\left.\begin{aligned}3x+7y-z&=a\\13x-4y+2z&=b\\2x-4y+2z&=c\end{aligned}\right\}\xrightarrow{\textbf{coef(S)}}\begin{bmatrix}3&7&-1\\13&-4&2\\2&-4&2\end{bmatrix}\xrightarrow{\textbf{rref[coef}(S)\textbf{]}}\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

does not contain a row of zeros:
system has a solution for all values of a, b, and c

(b) $\boldsymbol{S}$:

$$\left.\begin{aligned}x-3y+w&=a\\3x-y+2z-3w&=b\\x+z-5w&=c\\2x-y+3z-2w&=d\end{aligned}\right\}\xrightarrow{\textbf{coef(S)}}\begin{bmatrix}1&-3&0&1\\3&-1&2&-3\\1&0&1&-5\\2&-1&1&2\end{bmatrix}\xrightarrow{\textbf{rref[coef}(S)\textbf{]}}\begin{bmatrix}1&0&0&10\\0&1&0&3\\0&0&1&-15\\0&0&0&0\end{bmatrix}$$

contain a row of zeros:
system does not have a solution for all values of a, b, c, and d

CYU 1.7

$$\mathbf{S}:\left.\begin{array}{r} 2x+3y+4z+5w = 0\\ 3x+y+4z+w = 0\\ x+7y+4z+11w = 0\end{array}\right\}\xrightarrow{\mathbf{coef(S)}}\begin{array}{c}\begin{matrix} x & y & z & w\end{matrix}\\ \begin{bmatrix}2&3&4&5\\3&1&4&1\\1&7&4&11\end{bmatrix}\end{array}\xrightarrow{\mathbf{rref}}\begin{array}{c}\begin{matrix} x & y & z & w\end{matrix}\\ \begin{bmatrix}1&0&0&-2\\0&1&0&1\\0&0&1&\frac{3}{2}\end{bmatrix}\end{array}$$

Setting the free varialble w to r we arrive at the system:

$$\left.\begin{array}{r} x-2r = 0\\ y+r = 0\\ z+\frac{3}{2}r = 0\end{array}\right\}\text{ Solution set: }\left\{\left(2r, -r, -\frac{3}{2}r, r\right)\middle| r\in R\right\}$$

$$= \{(4r, -2r, -3r, 2r) | r \in R\}$$

CHAPTER 2
VECTOR SPACES

CYU 2.1 $r\mathbf{v}+s\mathbf{w} = 2(\mathbf{3, 2, -2})+(-3)(\mathbf{-3, 1, 0}) = (\mathbf{6, 4, -4})+(\mathbf{9, -3, 0}) = (\mathbf{15, 1, -4})$

CYU 2.2 **(iv)**: $\mathbf{v}+(-\mathbf{v}) = \mathbf{0}$ (in $\Re^2$):

$$\text{If } \mathbf{v} = (\mathbf{v_1, v_2})\text{, then: } \mathbf{v}+(-\mathbf{v}) \underset{\uparrow\,\text{Definition 2.5}}{\equiv} (\mathbf{v_1, v_2})+(-\mathbf{v_1}, -\mathbf{v_2}) \underset{\uparrow\,\text{Definition 2.3}}{\equiv} (\mathbf{v_1-v_1, v_2-v_2}) \underset{\uparrow\,\text{Definition 2.4}}{=} (\mathbf{0, 0}) \overset{\mathbf{PofR}\,\downarrow}{\equiv} \mathbf{0}$$

(in $\Re^n$): If $\mathbf{v} = (\mathbf{v_1, v_2, \ldots, v_n})$, then:

$$\mathbf{v}+(-\mathbf{v}) \equiv (\mathbf{v_1, v_2, \ldots, v_n})+(-v_1, -v_2, \ldots, -v_n) \equiv (\mathbf{v_1-v_1, v_2-v_2, \ldots, v_n-v_n}) = (\mathbf{0, 0, \ldots 0}) \equiv \mathbf{0}$$

CYU 2.3 For $\boldsymbol{A} = [a_{ij}] \in \boldsymbol{M}_{m\times n}$ and $r, s \in \Re$:

$$r(s\boldsymbol{A}) \equiv r(s[a_{ij}]) \equiv r[sa_{ij}] \equiv [r(sa_{ij})] = [(rs)a_{ij}] \equiv (rs)[a_{ij}] \equiv (rs)A$$

CYU 2.4 $\sum\limits_{i=0}^{n} a_i x^i + \sum\limits_{i=0}^{n} b_i x^i \equiv \sum\limits_{i=0}^{n} (a_i+b_i)x^i = \sum\limits_{i=0}^{n} (b_i+ai)x^i \equiv \sum\limits_{i=0}^{n} b_i x^i + \sum\limits_{i=0}^{n} a_i x^i$

CYU 2.5 $([(r+s)f](x) \equiv (r+s)[f(x)] = r[f(x)]+s[f(x)]) \equiv (rf)(x)+(sf)(x)$

CYU 2.6 **The Zero Axiom**: We need to find $\mathbf{0} = (\mathbf{a}, \mathbf{b}, \mathbf{c})$ such that for any $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in V$:

$(\mathbf{x}, \mathbf{y}, \mathbf{z}) + (\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{x}, \mathbf{y}, \mathbf{z})$, which is to say:

$$(\mathbf{x}+\mathbf{a}-\mathbf{1}, \mathbf{y}+\mathbf{b}+\mathbf{2}, \mathbf{z}+\mathbf{c}-\mathbf{3}) = (\mathbf{x}, \mathbf{y}, \mathbf{z}) \Rightarrow \begin{aligned} x+a-1 &= x \\ y+b+2 &= y \\ z+c-3 &= z \end{aligned} \Rightarrow \begin{aligned} a &= 1 \\ b &= -2 \\ c &= 3 \end{aligned}$$

Let's verify directly that $(\mathbf{1}, -\mathbf{2}, \mathbf{3}) + (\mathbf{x}, \mathbf{y}, \mathbf{z})$ does indeed equal $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ for every $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in V$: $(\mathbf{1}, -\mathbf{2}, \mathbf{3}) + (\mathbf{x}, \mathbf{y}, \mathbf{z}) = (\mathbf{1}+\mathbf{x}-\mathbf{1}, -\mathbf{2}+\mathbf{y}+\mathbf{2}, \mathbf{3}+\mathbf{z}-\mathbf{3}) = (\mathbf{x}, \mathbf{y}, \mathbf{z})$.

The Inverse Axiom: For given $(\mathbf{x}, \mathbf{y}, \mathbf{z}) \in V$ we are to find $(\mathbf{a}, \mathbf{b}, \mathbf{c})$ such that $(\mathbf{x}, \mathbf{y}, \mathbf{z}) + (\mathbf{a}, \mathbf{b}, \mathbf{c}) = \mathbf{0}$, which is to say:

$$(\mathbf{x}+\mathbf{a}-\mathbf{1}, \mathbf{y}+\mathbf{b}+\mathbf{2}, \mathbf{z}+\mathbf{c}-\mathbf{3}) = (\mathbf{1}, -\mathbf{2}, \mathbf{3}) \Rightarrow \begin{aligned} x+a-1 &= 1 \\ y+b+2 &= -2 \\ z+c-3 &= 3 \end{aligned} \Rightarrow \begin{aligned} a &= -x+2 \\ b &= -y+4 \\ c &= -z+6 \end{aligned}$$

It is easy to verify, directly that: $(\mathbf{x}+\mathbf{y}+\mathbf{z}) + (-\mathbf{x}+\mathbf{2}, -\mathbf{y}+\mathbf{4}, -\mathbf{z}+\mathbf{6}) = (\mathbf{1}, -\mathbf{2}, \mathbf{3})$.

CYU 2.7 Since $V = \{\mathbf{0}\}$ must be closed under addition and scalar multiplication we have no choice but to define: $\mathbf{0}+\mathbf{0} = \mathbf{0}$ and $r\mathbf{0} = \mathbf{0}$ for every $r \subseteq \Re$. It is easy to see that all eight axioms of Definition 2.6 hold. We establish (v): If $\mathbf{u}, \mathbf{v} \in V$, then they must both be $\mathbf{0}$. Consequently: $r(\mathbf{u}+\mathbf{v}) \equiv r(\mathbf{0}+\mathbf{0}) \equiv r\mathbf{0} \equiv \mathbf{0}$ and $r\mathbf{u}+r\mathbf{v} \equiv r\mathbf{0}+r\mathbf{0} \equiv \mathbf{0}+\mathbf{0} \equiv \mathbf{0}$. Hence: $r(\mathbf{u}+\mathbf{v}) = (r\mathbf{u})+(r\mathbf{v})$. As for (iii) and (iv), simply note that **0** is certainly the zero vector in V and that it is also its own inverse.

CYU 2.8

$$\begin{aligned} \mathbf{v}+\mathbf{z} &= \mathbf{w}+\mathbf{z} \\ (\mathbf{v}+\mathbf{z})+(-\mathbf{z}) &= (\mathbf{w}+\mathbf{z})+(-\mathbf{z}) \\ \mathbf{v}+[\mathbf{z}+(-\mathbf{z})] &= \mathbf{w}[\mathbf{z}+(-\mathbf{z})] \\ \mathbf{v}+\mathbf{0} &= \mathbf{w}+\mathbf{0} \\ \mathbf{v} &= \mathbf{w} \end{aligned}$$

CYU 2.9

$$\begin{aligned} \text{Start with: } r\mathbf{0} &= r(\mathbf{0}+\mathbf{0}) \\ r\mathbf{0} &= r\mathbf{0}+r\mathbf{0} \\ r\mathbf{0}+(-r\mathbf{0}) &= (r\mathbf{0}+r\mathbf{0})+(-r\mathbf{0}) \\ \mathbf{0} &= r\mathbf{0}+[r\mathbf{0}+(-r\mathbf{0})] \\ \mathbf{0} &= r\mathbf{0}+\mathbf{0} \\ \text{Conclusion: } \mathbf{0} &= r\mathbf{0} \end{aligned}$$

CYU 2.10 (a)

$$\begin{aligned} r\mathbf{v} &= r\mathbf{w} \\ \frac{1}{r}(r\mathbf{v}) &= \frac{1}{r}(r\mathbf{w}) \\ \left(\frac{1}{r}r\right)\mathbf{v} &= \left(\frac{1}{r}r\right)\mathbf{w} \\ 1\mathbf{v} &= 1\mathbf{w} \\ \text{Axiom (viii): } \mathbf{v} &= \mathbf{w} \end{aligned}$$

(b)

$$\begin{aligned} r\mathbf{v} &= s\mathbf{v} \\ r\mathbf{v}+[-(s\mathbf{v})] &= s\mathbf{v}+[-(s\mathbf{v})] \\ r\mathbf{v}+[-s(\mathbf{v})] &= \mathbf{0} \\ (r-s)\mathbf{v} &= \mathbf{0} \\ \text{Theorem 2.8: } r-s &= 0 \\ r &= s \end{aligned}$$

CYU 2.11 (a) $-(-\mathbf{v}) = (-1)[(-1)\mathbf{v}] = [(-1)(-1)]\mathbf{v} = 1\mathbf{v} = \mathbf{v}$

(b) $(-r)\mathbf{v} = [(-1)r]\mathbf{v} = (-1)(r\mathbf{v}) = -rv$

(c) $r(-\mathbf{v}) = r[(-1)\mathbf{v}] = [r(-1)]\mathbf{v} = (-1)[r\mathbf{v}] = -(r\mathbf{v})$

CYU 2.12 (a) $-(\mathbf{v}+\mathbf{w}) = (-1)(\mathbf{v}+\mathbf{w}) = (-1)\mathbf{v}+(-1)\mathbf{w} = -\mathbf{v}+(-\mathbf{w}) = -\mathbf{v}-\mathbf{w}$

(b) I. Claim holds for $n = 2$ $-(\mathbf{v}_1+\mathbf{v}_2) = -\mathbf{v}_1-\mathbf{v}_2$ [by part (a)].

II. Assume $-(\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k) = -\mathbf{v}_1-\mathbf{v}_2-\cdots-\mathbf{v}_k$. (The induction hypothesis)

III. We show that $-(\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k+\mathbf{v}_{k+1}) = -\mathbf{v}_1-\mathbf{v}_2-\cdots-\mathbf{v}_k-\mathbf{v}_{k+1}$:

$$\begin{aligned}
-(\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k+\mathbf{v}_{k+1}) &= -[(\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k)+\mathbf{v}_{k+1}]\\
\text{By I: } &= -(\mathbf{v}_1+\mathbf{v}_2+\cdots+\mathbf{v}_k)-\mathbf{v}_{k+1}\\
\text{By II: } &= (-\mathbf{v}_1-\mathbf{v}_2-\cdots-\mathbf{v}_n)-\mathbf{v}_{k+1} = -\mathbf{v}_1-\mathbf{v}_2-\cdots-\mathbf{v}_k-\mathbf{v}_{k+1}
\end{aligned}$$

CYU 2.13 Since $\begin{bmatrix}0&0\\0&0\end{bmatrix} \in S$, $S \neq \varnothing$. S is closed under addition:

$$\text{For any } \begin{bmatrix}a&2a\\-a&0\end{bmatrix}, \begin{bmatrix}b&2b\\-b&0\end{bmatrix} \in S\text{: } \begin{bmatrix}a&2a\\-a&0\end{bmatrix} + \begin{bmatrix}b&2b\\-b&0\end{bmatrix} = \begin{bmatrix}(a+b)&2(a+b)\\-(a+b)&0\end{bmatrix} \in S$$

S is closed under scalar multiplication:

$$\text{For any } \begin{bmatrix}a&2a\\-a&0\end{bmatrix} \in S \text{ and } r \in \Re\text{: } r\begin{bmatrix}a&2a\\-a&0\end{bmatrix} = \begin{bmatrix}(ra)&2(ra)\\-(ra)&0\end{bmatrix} \in S$$

CYU 2.14 Since $\{(\mathbf{0},\mathbf{0},\mathbf{0})\} \in S$, $S \neq \varnothing$. For any $(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}), (\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) \in S$ and $r \in \Re$: $r(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})+(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) = (\boldsymbol{rx}+\boldsymbol{a}, \boldsymbol{ry}+\boldsymbol{b}, \boldsymbol{rz}+\boldsymbol{c})$ is back in S, since:

$$(\boldsymbol{rx}+\boldsymbol{a})+(\boldsymbol{ry}+\boldsymbol{b})+(\boldsymbol{rz}+\boldsymbol{c}) = r\boxed{\boldsymbol{x}+\boldsymbol{y}+\boldsymbol{z}}+\boxed{(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c})} = r\mathbf{0}+\mathbf{0} = \mathbf{0}$$

CYU 2.15 Since the zero of that vector space is the zero function $Z\colon \Re \to \Re$ which maps every number to zero, and since $S = \{\boldsymbol{f} \in F(\Re) \mid f(9) = 0\}$, S does not contain the zero vector and is therefor not a subspace of $F(\Re)$.

CYU 2.16 Since $(\mathbf{0},\mathbf{0},\mathbf{0}) \in S$, $S \neq \varnothing$. For any $(\mathbf{16}\boldsymbol{a}-\mathbf{2}\boldsymbol{b}, \mathbf{4}\boldsymbol{a}-\mathbf{17}\boldsymbol{b}, \mathbf{11}\boldsymbol{a}, \mathbf{11}\boldsymbol{b}), (\mathbf{16}\boldsymbol{x}-\mathbf{2}\boldsymbol{y}, \mathbf{4}\boldsymbol{x}-\mathbf{17}\boldsymbol{y}, \mathbf{11}\boldsymbol{x}, \mathbf{11}\boldsymbol{y}) \in S$ and $(r \in \Re)$:

$$\begin{aligned}
&r(\mathbf{16}\boldsymbol{a}-\mathbf{2}\boldsymbol{b}, \mathbf{4}\boldsymbol{a}-\mathbf{17}\boldsymbol{b}, \mathbf{11}\boldsymbol{a}, \mathbf{11}\boldsymbol{b})+(\mathbf{16}\boldsymbol{x}-\mathbf{2}\boldsymbol{y}, \mathbf{4}\boldsymbol{x}-\mathbf{17}\boldsymbol{y}, \mathbf{11}\boldsymbol{x}, \mathbf{11}\boldsymbol{y})\\
&\quad = (\mathbf{16}[\boldsymbol{ra}+\boldsymbol{x}]-\mathbf{2}[\boldsymbol{rb}+\boldsymbol{y}], \mathbf{4}[\boldsymbol{ra}+\boldsymbol{x}]-\mathbf{17}[\boldsymbol{rb}+\boldsymbol{y}], \mathbf{11}[\boldsymbol{ra}+\boldsymbol{x}], \mathbf{11}[\boldsymbol{ra}+\boldsymbol{x}])\\
&\quad = (\mathbf{16}\boldsymbol{A}-\mathbf{2}\boldsymbol{B}, \mathbf{4}\boldsymbol{A}-\mathbf{17}\boldsymbol{B}, \mathbf{11}\boldsymbol{A}, \mathbf{11}\boldsymbol{B}), \text{ where } A = ra+x \text{ and } B = ra+x
\end{aligned}$$

CYU 2.17 False. The subsets $S = \{0, 1\}$ and $T = \{0, 2\}$ are not subspaces of $\Re$ since neither is close under addition (nor under scalar multiplication), yet $S \cap T = \{0\}$ is a subspace of $\Re$.

CYU 2.18 (a) For $\mathbf{v} = (\mathbf{2-1, 3-5}) = (\mathbf{1, -2})$ and $\mathbf{u} = (\mathbf{2, 3})$:

$$\{\mathbf{u} + r\mathbf{v} | r \in \Re\} = \{(\mathbf{2, 3}) + \boldsymbol{r}(\mathbf{1, -2})\} = \{(\mathbf{2} + \boldsymbol{r}, \mathbf{3} - \mathbf{2}\boldsymbol{r})\}$$

(b) Since $(\mathbf{1} + \boldsymbol{r}, \mathbf{5} - \mathbf{2}\boldsymbol{r}) = [\mathbf{2} + (\boldsymbol{r} - \mathbf{1}), \mathbf{3} - \mathbf{2}(\boldsymbol{r} - \mathbf{1})]$:

$$\{(\mathbf{1} + \boldsymbol{r}, \mathbf{5} - \mathbf{2}\boldsymbol{r}) | r \in \Re\} = \{(\mathbf{2} + \boldsymbol{r}, \mathbf{3} - \mathbf{2}\boldsymbol{r}) | r \in \Re\}$$

CYU 2.19 Choosing $r = -1$ and $r = 1$ in $L = \{(1, 3, 5) + r(2, 1, -1) | r \in \Re\}$ we obtain the two points $(1, 3, 5) - (2, 1, -1) = (-1, 2, 6)$, $(1, 3, 5) + (2, 1, -1) = (3, 4, 4)$ on L. Preceding as in Example 2.14 we arrive at a direction vector $\mathbf{v} = (\mathbf{3, 4, 4}) - (\mathbf{-1, 2, 6}) = (\mathbf{4, 2, -2})$. Selecting $\mathbf{u} = (\mathbf{3, 4, 4})$ as our translation vector, we have:
$\bar{L} = \{(\mathbf{3, 4, 4}) + \bar{r}(\mathbf{4, 2, -2}) | \bar{r} \in \Re\} = \{(\mathbf{3} + \mathbf{4}\boldsymbol{r}, \mathbf{4} + \mathbf{2}\boldsymbol{r}, \mathbf{4} - \mathbf{2}\boldsymbol{r}) | r \in \Re\}$, which we now show to be equal to the set $L = \{(1, 3, 5) + r(2, 1, -1) | r \in \Re\} = \{(1 + 2r, 3 + r, 5 - r) | r \in \Re\}$:

$$(3 + 4\bar{r}, 4 + 2\bar{r}, 4 - 2\bar{r}) = (1 + 2r, 3 + r, 5 - r), \text{ were } \bar{r} = \frac{r-1}{2}$$

CYU 2.20 Choosing $(4, 1, -3)$ to play the role of $\mathbf{w}$, instead of $(3, -2, 2)$ we obtain:

$$\begin{aligned}\bar{P} &= \{\mathbf{w} + \bar{r}\mathbf{u} + \bar{s}\mathbf{v} | \bar{r}, \bar{s} \in \Re\}\\ &= \{(\mathbf{4, 1, -5}) + \bar{r}(\mathbf{-1, 7, -5}) + \bar{s}(\mathbf{1, 3, -5}) | \bar{r}, \bar{s} \in \Re\}\\ &= \{(\mathbf{4} - \bar{\boldsymbol{r}} + \bar{\boldsymbol{s}}, \mathbf{1} + \mathbf{7}\bar{\boldsymbol{r}} + \mathbf{3}\bar{\boldsymbol{s}}, -\mathbf{5} - \mathbf{5}\bar{\boldsymbol{r}} - \mathbf{5}\bar{\boldsymbol{s}}) | \bar{r}, \bar{s} \in \Re\}\end{aligned}$$

which we show to be equal to the set

$$\begin{aligned}P &= \{\mathbf{w} + r\mathbf{u} + s\mathbf{v} | r, s \in \Re\} \quad :\\ &= \{(\mathbf{3, -2, 2}) + r(\mathbf{-1, 7, -5}) + s(\mathbf{1, 3, -5}) | r, s \in \Re\}\\ &= \{(\mathbf{3} - \boldsymbol{r} + \boldsymbol{s}, -\mathbf{2} + \mathbf{7}\boldsymbol{r} + \mathbf{3}\boldsymbol{s}, \mathbf{2} - \mathbf{5}\boldsymbol{r} - \mathbf{5}\boldsymbol{s}) | r, s \in \Re\}\end{aligned}$$

Equating the first two components of $\bar{P}$ and P we obtain:
$$\begin{aligned}(1)&: 4 - \bar{r} + \bar{s} = 3 - r + s\\ (2)&: 1 + 7\bar{r} + 3\bar{s} = -2 + 7r + 3s\end{aligned}.$$

Multiplying equation (1) by -3 and adding it to equation (2) we find that $\bar{r} = r$. Substituting in (1) we then find that $\bar{s} = s - 1$; bringing us to:

$$(\mathbf{4} - \bar{\boldsymbol{r}} + \bar{\boldsymbol{s}}, \mathbf{1} + \mathbf{7}\bar{\boldsymbol{r}} + \mathbf{3}\bar{\boldsymbol{s}}, -\mathbf{5} - \mathbf{5}\bar{\boldsymbol{r}} - \mathbf{5}\bar{\boldsymbol{s}}) = (\mathbf{3} - \boldsymbol{r} + \boldsymbol{s}, -\mathbf{2} + \mathbf{7}\boldsymbol{r} + \mathbf{3}\boldsymbol{s}, \mathbf{2} - \mathbf{5}\boldsymbol{r} - \mathbf{5}\boldsymbol{s})$$

CHAPTER 3
BASES AND DIMENSIONS

CYU 3.1 (a) No:

$$\mathbf{S}:\left.\begin{aligned} a+2b&=-2\\ 3a+5b&=-3\\ 8a+4b&=8\end{aligned}\right\}\xrightarrow{\mathbf{aug}(S)}\left[\begin{array}{cc|c} a & b & \\ 1 & 2 & -2\\ 3 & 5 & -3\\ 8 & 4 & 8\end{array}\right]\xrightarrow{\mathbf{rref}}\left[\begin{array}{cc|c} a & b & \\ 1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{array}\right]$$

No solution!

(b) Yes:

$$\mathbf{S}:\left.\begin{aligned} a+2b&=-2\\ 3a+5b&=-4\\ 8a+4b&=8\end{aligned}\right\}\xrightarrow{\mathbf{aug}(S)}\left[\begin{array}{cc|c} a & b & \\ 1 & 2 & -2\\ 3 & 5 & -4\\ 8 & 4 & 8\end{array}\right]\xrightarrow{\mathbf{rref}}\left[\begin{array}{cc|c} a & b & \\ 1 & 0 & 2\\ 0 & 1 & -2\\ 0 & 0 & 0\end{array}\right]$$

$$\longrightarrow (\mathbf{-2, -4-8}) = 2(\mathbf{1, 3, 8}) - 2(\mathbf{2, 5, 4})$$

CYU 3.2 (a) We are to show that for any given matrix $\begin{bmatrix} a & b\\ c & d\end{bmatrix}$ there exist scalars x, y, z, w for which

$$(*)\quad x\begin{bmatrix}1&2\\3&4\end{bmatrix}+y\begin{bmatrix}1&0\\1&0\end{bmatrix}+z\begin{bmatrix}0&1\\0&1\end{bmatrix}+w\begin{bmatrix}0&4\\2&0\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}:$$

$$\begin{bmatrix}x+y & 2x+z+4w\\ 3x+y+2w & 4x+z\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$$

$$\mathbf{S}:\left.\begin{aligned} x+y+0z+0w&=a\\ 2x+0y+z+4w&=b\\ 3x+y+0z+2w&=c\\ 4x+0y+z+0w&=d\end{aligned}\right\}\xrightarrow{\textbf{coef(S)}}\begin{bmatrix}1&1&0&0\\2&0&1&4\\3&1&0&2\\4&0&1&0\end{bmatrix}\xrightarrow{\textbf{[rref(coef(S)]}}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

Since rref[coef(S)] does not contain a row consisting entirely of zeros, the system **S** [which stems from (*)] has a solution for any given a, b, c, d.

(b)

$$\mathbf{S}:\left.\begin{aligned} x+y+0z+0w&=-1\\ 2x+0y+z+4w&=5\\ 3x+y+0z+2w&=1\\ 4x+0y+z+0w&=13\end{aligned}\right\}\xrightarrow{\textbf{aug(S)}}\left[\begin{array}{cccc|c}1&1&0&0&-1\\2&0&1&4&5\\3&1&0&2&1\\4&0&1&0&13\end{array}\right]\xrightarrow{\textbf{rref}}\left[\begin{array}{cccc|c}1&0&0&0&\mathbf{2}\\0&1&0&0&\mathbf{-3}\\0&0&1&0&\mathbf{5}\\0&0&0&1&\mathbf{-1}\end{array}\right]$$

from the above we see that: $\mathbf{2}\begin{bmatrix}1&2\\3&4\end{bmatrix}\mathbf{-3}\begin{bmatrix}1&0\\1&0\end{bmatrix}+\mathbf{5}\begin{bmatrix}0&1\\0&1\end{bmatrix}\mathbf{-1}\begin{bmatrix}0&4\\2&0\end{bmatrix}=\begin{bmatrix}-1&5\\1&13\end{bmatrix}$

CYU 3.3 We are to find the set of vectors $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ for which there exist scalars x, y, z, such that: $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = x(\mathbf{2, 1, 5}) + y(\mathbf{1, -2, 2}) + z(\mathbf{0, 5, 1})$:

$$\left.\begin{aligned} 2x + y + 0z &= a \\ x - 2y + 5z &= b \\ 5x + 2y + z &= c \end{aligned}\right\} \Rightarrow \left[\begin{array}{ccc|c} 2 & 1 & 0 & a \\ 1 & -2 & 5 & b \\ 5 & 2 & 1 & c \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 5 & b \\ 2 & 1 & 0 & a \\ 5 & 2 & 1 & c \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 5 & b \\ 0 & 5 & -10 & a - 2b \\ 0 & 12 & -24 & c - 5b \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & -2 & 5 & b \\ 0 & 1 & -2 & \dfrac{a - 2b}{5} \\ 0 & 0 & 0 & c - 5b - \dfrac{12(a - 2b)}{5} \end{array}\right]$$

Note that in the above we didn't bother to reduce the coefficient matrix to its row-reduced-echelon forms with 0's above and below leading ones. Rather, we obtained its row-echelon form with 0's only below the leading ones [see Exercises18-22, page 12]. This still enables us to determine $\text{Span}\{(\mathbf{2, 1, 5}), (\mathbf{1, -2, 2}), (\mathbf{0, 5, 1})\}$, for it consists of all vectors $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ for which $c - 5b - \dfrac{12(a - 2b)}{5} = 0$, which is equivalent to: $12a + b - 5c = 0$ or $b = 5c - 12a$.

Conclusion: $\text{Span}\{(\mathbf{2, 1, 5}), (\mathbf{1, -2, 2}), (\mathbf{0, 5, 1})\} = \{(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) | b = 5c - 12a\}$.

Any vector $(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c})$ for which $b \neq 5c - 12a$, say (1,2,3), will not be in the spanning set.

CYU 3.4 For given $v \in V$ we are to find scalars A, B, C such that $v = Av_1 + B(v_1 + v_2) + C(v_1 + v_2 + v_3)$. Since $\{v_1, v_2, v_3\}$ span *V*, there exist scalars *a*,*b*,*c* such that $v = av_1 + bv_2 + cv_3$. Equating these two expressions for *v* we have:

$$Av_1 + B(v_1 + v_2) + C(v_1 + v_2 + v_3) = av_1 + bv_2 + cv_3$$
$$(A + B + C)v_1 + (B + C)v_3 + Cv_3 = av_1 + bv_2 + cv_3$$

Bringing us to: $\left.\begin{aligned} A + B + C &= a \\ B + C &= b \\ C &= c \end{aligned}\right\}$ with solution: $\begin{aligned} C &= c \\ B &= b - C = b - c \\ A &= a - B - C = a - (b - c) - c = a - b \end{aligned}$

CYU 3.5
$$\begin{aligned} ax^2 + b(2x^2 + x) + c(x - 3) &= \mathbf{0} \\ (a + 2b)x^2 + (b + c)x - 3c &= \mathbf{0} \end{aligned} \longrightarrow \left.\begin{aligned} a + 2b + 0c &= 0 \\ 0a + b + c &= 0 \\ 0a + 0b - 3c &= 0 \end{aligned}\right\} \Rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & -3 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

No free variable
Linearly independent

CYU 3.6
$$a\boldsymbol{v}_1 + b(\boldsymbol{v}_1 + \boldsymbol{v}_2) + c(\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3) + d(\boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3 + \boldsymbol{v}_4) = \mathbf{0}$$

(we are to show $a = b = c = d = 0$)

$$(a + b + c + d)\boldsymbol{v}_1 + (b + c + d)\boldsymbol{v}_2 + (c + d)\boldsymbol{v}_3 + d\boldsymbol{v}_4 = \mathbf{0}$$

Since $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \boldsymbol{v}_4\}$ is linearly independent:

$$\longrightarrow \left.\begin{aligned} a + b + c + d &= 0 \\ b + c + d &= 0 \\ c + d &= 0 \\ d &= 0 \end{aligned}\right\} \Rightarrow a = b = c = d = 0$$

CYU 3.7

$$A(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4) + B(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4) + C(\mathbf{c}_1, \mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_4) + D(\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \mathbf{d}_4) + E(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4) = \mathbf{0}$$

$$\longrightarrow \begin{bmatrix} a_1 & b_1 & c_1 & d_1 & e_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 \\ a_3 & b_3 & c_3 & d_3 & e_3 \\ a_4 & b_4 & c_4 & d_4 & e_4 \end{bmatrix} \longrightarrow \begin{array}{l}\text{rref must have a free variable} \\ \text{(5 variables and 4 equations)}\end{array}$$

CYU 3.8 Linear dependent since: $\begin{bmatrix} 2 & 5 & 11 \\ 1 & 0 & 3 \\ 3 & 2 & 11 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$.

$(\mathbf{8, 4, 12}) \in \text{Span}\{(\mathbf{2, 1, 3}), (\mathbf{5, 0, 2}), (\mathbf{11, 3, 11})\}$ since: $\left[\begin{array}{ccc|c} 2 & 5 & 11 & 8 \\ 1 & 0 & 3 & 4 \\ 3 & 2 & 11 & 12 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 3 & 4 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$.

From the above rref-matrix we see that $a(\mathbf{2, 1, 3}) + b(\mathbf{5, 0, 2}) + c(\mathbf{11, 3, 11}) = (\mathbf{8, 4, 12})$ for any a,b, and c for which $\begin{array}{r} a + 3c = 4 \\ b + c = 0 \end{array}$. Letting $c = 0$ we get $a = 4$ and $b = 0$, giving us the linear combination $(\mathbf{8, 4, 12}) = 4(\mathbf{2, 1, 3}) + 0(\mathbf{5, 0, 2}) + 0(\mathbf{11, 3, 11})$. Letting $c = 1$ we get $a = 1$ and $b = -1$, giving us another linear combination $(\mathbf{8, 4, 12}) = (\mathbf{2, 1, 3}) - (\mathbf{5, 02}) + (\mathbf{11, 3, 11})$.

CYU 3.9 $x^3 + x$ and -7 are clearly linearly independent. Since x^2 cannot be “built” from those vectors, $\{x^3 + x, -7, x^2\}$ is linearly independent. Can x^3 be “built” by those three vectors? No; so $\{x^3 + x, -7, x^2, x^3\}$ is linearly independent. Just to make sure:

$$\begin{array}{r} a(x^3 + x) + b(-7) + cx^2 + dx^3 = \mathbf{0} \\ (a + d)x^3 + cx^2 + ax + b(-7) = \mathbf{0} \end{array} \nearrow \begin{array}{cccc} a & b & c & d \\ \hline 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & -7 & 0 & 0 \end{array} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(Incidentally, if you throw in any two randomly chosen vectors from P_3 into $\{x^3 + x, -7\}$ chances are really good that you will end up with a linearly independent set. Try it.)

<u>CYU 3.10</u>

For spanning: $$x\begin{bmatrix}2&1\\3&0\end{bmatrix}+y\begin{bmatrix}1&1\\2&2\end{bmatrix}+z\begin{bmatrix}-3&0\\-5&-5\end{bmatrix}+w\begin{bmatrix}0&4\\1&5\end{bmatrix}=\begin{bmatrix}a&b\\c&d\end{bmatrix}$$

For linear independence: $$x\begin{bmatrix}2&1\\3&0\end{bmatrix}+y\begin{bmatrix}1&1\\2&2\end{bmatrix}+z\begin{bmatrix}-3&0\\-5&-5\end{bmatrix}+w\begin{bmatrix}0&4\\1&5\end{bmatrix}=\begin{bmatrix}0&0\\0&0\end{bmatrix}$$

coefficient matrix $$\begin{bmatrix}2&1&-3&0\\1&1&0&4\\3&2&-5&1\\0&2&-5&5\end{bmatrix}\xrightarrow{\text{rref}}\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$$

Spans: Does not contain a row consisting entirely of zeros.

Linearly Independent: each row ha s leading one.

<u>CYU 3.11</u> If $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots \boldsymbol{v}_n\}$ is a basis, then it spans V insuring us that every vector in V can be expressed as a linear combination of the vectors in S. Being a basis, S is also linearly independent, and Theorem 3.6, page 89 insures us that the representation is unique.

Conversely, if every vector in V can uniquely be expressed as a linear combination of the vectors in S then S certainly spans V. To show that it is also linearly independent consider $c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \ldots + c_n\boldsymbol{v}_n = 0$. Since $0\boldsymbol{v}_1 + 0\boldsymbol{v}_2 + \ldots + 0\boldsymbol{v}_n = 0$, and since we have unique representation: $c_i = 0$ for $1 \le i \le n$.

<u>CYU 3.12</u> We show that $S = \{(1, 0), (0, 1)\}$ is a basis for the space V of Example 2.5.

<u>S spans V</u>: For $(\boldsymbol{a}, \boldsymbol{b}) \in V$ can we find $r, s \in \Re$ such that $(a, b) = r(1, 0) + s(0, 1)$? Yes:

$$\begin{aligned}
(a, b) &= r(1, 0) + s(0, 1)\\
\text{Since } r(x, y) = (rx - r - 1, ry + r - 1)\text{:}\quad &= (r - r + 1, r - 1) + (-s + 1, s + s - 1)\\
&= (1, r - 1) + (-s + 1, 2s - 1)\\
\text{Since } (x, y) + (x', y') = (x + x' - 1, y + y' + 1)\text{:}\quad &= (-s + 1, r - 1 + 2s - 1 + 1)\\
&= (-s + 1, 2s + r - 1)
\end{aligned}$$

equating coefficients

$$\left.\begin{aligned} a &= -s + 1\\ b &= 2s + r - 1\end{aligned}\right\}\Rightarrow s = -a + 1 \text{ and } r = b + 2a - 1$$

Check: $$\begin{aligned}
r(1, 0) + s(0, 1) &= (b + 2a - 1)(1, 0) + (-a + 1)(0, 1)\\
&= [(b - 2a - 1) - (b + 2a - 1) + 1, (b + 2a - 1) - 1] + [-(-a + 1) + 1, (-a + 1) - 1]\\
&= (1, b + 2a - 2) + (a, -2a + 1) = (1 + a - 1, (b + 2a - 2) + (a - 2a + 1) - 1) = (a, b)
\end{aligned}$$

<u>S is linearly independent</u>: Recalling that $(1, -1)$ is the zero vector in S we start with the equation $a(1, 0) + b(0, 1) = (1, -1)$ and go on to show that $a = b = 0$:

$$\begin{aligned}
a(1, 0) + b(0, 1) &= (1, -1)\\
(1, a - 1) + (-b + 1, b + b - 1) &= (1, -1)\\
(1, a - 1) + (-b + 1, 2b - 1) &= (1, -1)\\
(1 - b + 1 - 1, a - 1 + 2b - 1 + 1) &= (1, -1)\\
(1 - b, a - 1 + 2b) &= (1, -1)
\end{aligned}$$

$$\left.\begin{aligned} 1 - b &= 1\\ a - 1 + 2b &= -1\end{aligned}\right\} \text{solution: } a = b = 0$$

CYU 3.13 (a) Knowing that $M_{2\times 2}$ has dimension 4, we simply have to add two vectors to $L = \left\{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}\right\}$ without rupturing linear independence. By now, you may be convinced that if you add any two randomly chosen vectors, say $\begin{bmatrix} 1 & 5 \\ -1 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ -2 & 6 \end{bmatrix}$, chances are good that you will end up with a set of four independent vectors, and therefore a basis for $M_{2\times 2}$. Let's make sure that we do:

$$a\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + b\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} + c\begin{bmatrix} 1 & 5 \\ -1 & 1 \end{bmatrix} + d\begin{bmatrix} 0 & 1 \\ -2 & 6 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \longrightarrow \begin{array}{c} \begin{matrix} a & b & c & d \end{matrix} \\ \begin{bmatrix} 2 & 2 & 1 & 0 \\ 1 & 2 & 5 & 1 \\ 1 & 1 & -1 & -2 \\ 2 & 1 & 1 & 6 \end{bmatrix} \end{array} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(b) Here is a brute force approach to obtain a basis for Span(S). We start with the first vector in $S = \{(\mathbf{3}, -\mathbf{1}, \mathbf{2}), (-\mathbf{9}, \mathbf{3}, -\mathbf{6}), (\mathbf{1}, \mathbf{2}, -\mathbf{2}), (-\mathbf{5}, \mathbf{4}, -\mathbf{6}), (\mathbf{6}, -\mathbf{2}, \mathbf{4})\}$: $\{(\mathbf{3}, -\mathbf{1}, \mathbf{2})\}$. Since the second vector is easily seen to be a multiple of the first, we discard it and turn our attention to the third vector $(\mathbf{1}, \mathbf{2}, -\mathbf{2})$. Since that vector is not a multiple of $\{(\mathbf{3}, -\mathbf{1}, \mathbf{2})\}$, we throw it into that set to obtain the two independent vectors $\{(\mathbf{3}, -\mathbf{1}, \mathbf{2}), (\mathbf{1}, \mathbf{2}, -\mathbf{2})\}$. Can the vector $(-\mathbf{5}, \mathbf{4}, -\mathbf{6})$ be built from those two independent vectors? Yes:

$$\left[\begin{array}{cc|c} 3 & 1 & -5 \\ -1 & 2 & 4 \\ 2 & -2 & -6 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|c} 1 & 0 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right] \Rightarrow -2(\mathbf{3}, -\mathbf{1}, \mathbf{2}) + 1(\mathbf{1}, \mathbf{2}, -\mathbf{2}) = (-\mathbf{5}, \mathbf{4}, -\mathbf{6})$$

Since we are looking for a maximal independent set, we discard $(-\mathbf{5}, \mathbf{4}, -\mathbf{6})$ and turn our attention to the last remaining vector $(\mathbf{6}, -\mathbf{2}, \mathbf{4})$. Is it independent of the vectors $\{(\mathbf{3}, -\mathbf{1}, \mathbf{2}), (\mathbf{1}, \mathbf{2}, -\mathbf{2})\}$? Clearly not, since $(\mathbf{6}, -\mathbf{2}, \mathbf{4}) = 2(\mathbf{3}, -\mathbf{1}, \mathbf{2})$.

Conclusion: $\{(3, -1, 2), (1, 2, -2)\}$ is a basis for Span(S). Since $\Re^3$ is of dimension 3, S does not span $\Re^3$.

CYU 3.14:

first vector

$$\begin{bmatrix} 3 & -9 & 1 & -5 & 6 \\ -1 & 3 & 2 & 4 & -2 \\ 2 & -6 & -2 & -6 & 4 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} \boxed{1} & -3 & 0 & -2 & 2 \\ 0 & 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{Basis for Span}(S) = \{(3, -1, 2), (1, 2, -2)\}$$

third vector

CHAPTER 4
LINEARITY

CYU 4.1 The function $f\colon \Re^2 \to \Re^3$ given by $f(a, b) = (a + b, 2b, a - b)$ is linear.
f preserves sums:

$$\begin{aligned} f[(a, b) + (a', b')] &= f(a + a', b + b') \\ &= [(a + a') + (b + b'), 2(b + b'), (a + a') - (b + b')] \\ &= (a + b, 2b, a - b) + (a' + 'b, 2b', a' - b') = f(a, b) + f(a', b') \end{aligned}$$

f preserves scalar products: $f[r(a, b)] = f(ra, rb) = (ra + rb, 2rb, ra - rb)$

$$= r(a + b, 2b, a - b) = rf(a, b)$$

CYU 4.2 Since $f(0, 0) = 0x^2 + 0x + 1 = 1 \neq 0$, f is not linear.

CYU 4.3
$$\begin{aligned} f[r(a, b) + (a', b')] &= f(ra + a', rb + b') \\ &= [(ra + a') + (rb + b'), 2(rb + b'), (ra + a') - (rb + b')] \\ &= [(ra + rb) + (a' + b'), 2rb + 2b', (ra - rb) + (a' - b')] \\ &= (ra + rb, 2rb, ra - rb) + (a' + b', 2b', a' - b') \\ &= r(a + b, 2b, a - b) + (a' + b', 2b', a' - b') = rf(a, b) + f(a', b') \end{aligned}$$

CUU 4.4 A counterexample: The trivial function $T\colon \Re \to \Re$ given by $T(x) = 0$ for every $x \in \Re$ is linear. The set $S = \Re^+$ is not a subspace of $\Re$, but $f(S) = \{0\}$ is a subspace of $\Re$.

CUU 4.5 True: The proof of Theorem 4.5 makes no mention of linear independence.

CUU 4.6 (a) We first express $(3, 4, 2)$ as a linear combination of the basis $\{(1, 0, 0), (0, 2, 0), (1, 1, 1)\}$:

$$\boxed{\begin{aligned} (3, 4, 2) &= a(1, 0, 0) + b(0, 2, 0) + c(1, 1, 1) \\ &= (a + c, 2b + c, c) \end{aligned}} \longrightarrow \left.\begin{aligned} a + c &= 3 \\ 2b + c &= 4 \\ c &= 2 \end{aligned}\right\} \Rightarrow c = 2, b = 1, a = 1$$

Then:
$$\begin{aligned} T(3, 4, 2) &= T[(1, 0, 0) + (0, 2, 0) + 2(1, 1, 1)] \\ \text{By linearity:} \quad &= T(1, 0, 0) + T(0, 2, 0) + 2T(1, 1, 1) \\ &= (2x^2 + x) + (2x^2 + x) + 2(x - 5) = 4x^2 + 4x - 10 \end{aligned}$$

(b) Expressing (a, b, c) as a linear combination of the given basis we have:

$$\boxed{\begin{aligned} (a, b, c) &= A(1, 0, 0) + B(0, 2, 0) + C(1, 1, 1) \\ &= (A + C, 2B + C, C) \end{aligned}} \longrightarrow \left.\begin{aligned} A + C &= a \\ 2B + C &= b \\ C &= c \end{aligned}\right\} \Rightarrow C = c, B = \frac{b - c}{2}, A = a - c$$

Then: $$\begin{aligned} T(\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}) &= T\left[(a-c)(\mathbf{1},\mathbf{0},\mathbf{0})+\frac{b-c}{2}(\mathbf{0},\mathbf{2},\mathbf{0})+c(\mathbf{1},\mathbf{1},\mathbf{1})\right] \\ &= (a-c)T(\mathbf{1},\mathbf{0},\mathbf{0})+\frac{b-c}{2}T(\mathbf{0},\mathbf{2},\mathbf{0})+cT(\mathbf{1},\mathbf{1},\mathbf{1}) \\ &= (a-c)(2x^2+x)+\frac{b-c}{2}(2x^2+x)+c(x-5) \\ &= (2a+b-3c)x^2+\left(a+\frac{b}{2}-\frac{c}{2}\right)x-5c \end{aligned}$$

CYU 4.7 (a) $\{(\mathbf{1},\mathbf{0}),(\mathbf{1},\mathbf{1})\}$ is easily seen to be linearly independent, and therefore a basis of $\Re^2$. $\{(\mathbf{0},\mathbf{1},\mathbf{0}),(\mathbf{1},\mathbf{1},\mathbf{0}),(\mathbf{0},\mathbf{1},\mathbf{1})\}$ is easily seen to be linearly independent and therefore a basis of $\Re^3$.
(b) From the given information, we have:

$$\begin{aligned} (L\circ T)(\mathbf{1},\mathbf{0}) &= L[T(\mathbf{1},\mathbf{0})] \\ &= L(\mathbf{0},\mathbf{2},\mathbf{0}) = L[\mathbf{2}(\mathbf{0},\mathbf{1},\mathbf{0})] = 2L(\mathbf{0},\mathbf{1},\mathbf{0}) = 2(\mathbf{0},\mathbf{1}) = (\mathbf{0},\mathbf{2}) \end{aligned}$$

To determine $(L\circ T)(\mathbf{1},\mathbf{1}) = L[T(\mathbf{1},\mathbf{1})] = L(\mathbf{1},\mathbf{0},\mathbf{1})$, we need to express $(\mathbf{1},\mathbf{0},\mathbf{1})$ as a linear combination of the basis $\{(\mathbf{0},\mathbf{1},\mathbf{0}),(\mathbf{1},\mathbf{1},\mathbf{0}),(\mathbf{0},\mathbf{1},\mathbf{1})\}$. Let's do it:

$$\boxed{\begin{aligned} (\mathbf{1},\mathbf{0},\mathbf{1}) &= a(\mathbf{0},\mathbf{1},\mathbf{0})+b(\mathbf{1},\mathbf{1},\mathbf{0})+c(\mathbf{0},\mathbf{1},\mathbf{1}) \\ &= (\boldsymbol{b},\boldsymbol{a}+\boldsymbol{b}+\boldsymbol{c},\boldsymbol{c}) \end{aligned}} \rightarrow \left.\begin{aligned} b &= 1 \\ a+b+c &= 0 \\ c &= 1 \end{aligned}\right\} \Rightarrow c = 1, b = 1, a = -2$$

Then: $$\begin{aligned} (L\circ T)(\mathbf{1},\mathbf{1}) = L[T(\mathbf{1},\mathbf{1})] &= L(\mathbf{1},\mathbf{0},\mathbf{1}) \\ &= L[-2(\mathbf{0},\mathbf{1},\mathbf{0})+(\mathbf{1},\mathbf{1},\mathbf{0})+(\mathbf{0},\mathbf{1},\mathbf{1})] \\ &= -2L(\mathbf{0},\mathbf{1},\mathbf{0})+L(\mathbf{1},\mathbf{1},\mathbf{0})+L(\mathbf{0},\mathbf{1},\mathbf{1}) \\ &= -2(\mathbf{0},\mathbf{1})+(\mathbf{1},\mathbf{0})+(\mathbf{1},\mathbf{0}) = (\mathbf{2},-\mathbf{2}) \end{aligned}$$

To determine $(L\circ T)(\boldsymbol{a},\boldsymbol{b})$, we first express $(\boldsymbol{a},\boldsymbol{b})$ as a linear combination of $\{(\mathbf{1},\mathbf{0}),(\mathbf{1},\mathbf{1})\}$:
$(\boldsymbol{a},\boldsymbol{b}) = A(\mathbf{1},\mathbf{0})+B(\mathbf{1},\mathbf{1}) = (\boldsymbol{A}+\boldsymbol{B},\boldsymbol{B}) \Rightarrow B = b$ and $A = a-b$; so:

$$(\boldsymbol{a},\boldsymbol{b}) = (a-b)(\mathbf{1},\mathbf{0})+b(\mathbf{1},\mathbf{1})$$

Putting this together we find that:

$$\begin{aligned} (L\circ T)(\boldsymbol{a},\boldsymbol{b}) &= (L\circ T)[(a-b)(\mathbf{1},\mathbf{0})+b(\mathbf{1},\mathbf{1})] \\ &= (a-b)[(L\circ T)(\mathbf{1},\mathbf{0})]+b[(L\circ T)(\mathbf{1},\mathbf{0})] \\ &= (a-b)(\mathbf{0},\mathbf{2})+b(\mathbf{2},-\mathbf{2}) = (\mathbf{2}\boldsymbol{b},\mathbf{2}\boldsymbol{a}-\mathbf{4}\boldsymbol{b}) \end{aligned}$$

CYU 4.8 (a):

$$\begin{aligned} T[r(ax^2+bx+c)+(a'x^2+b'x+c')] &= T[(ra+a')x^2+(rb+b')x+(rc+c')] \\ &= \begin{bmatrix} ra+a' & rb+b' \\ rc+c' & ra+a' \end{bmatrix} = r\begin{bmatrix} a & b \\ c & a \end{bmatrix}+\begin{bmatrix} a' & b' \\ c' & a' \end{bmatrix} \\ &= rT(ax^2+bx+c)+T(a'x^2+b'x+c') \end{aligned}$$

(b) The vectors $T(x^2) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $T(x) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $T(1) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$ span $\text{Im}(T)$ (Theorem 4.9), and they are easily seen to be linearly independent. Consequently: $\text{rank}(T) = 3$.

By definition,

$$\begin{aligned} \text{Ker}(T) &= \{ax^2+bx+c \mid T(ax^2+bx+c) = \mathbf{0}\} \\ &= \left\{ax^2+bx+c \,\middle|\, \begin{bmatrix} a & b \\ c & a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right\} \Rightarrow a = b = c = 0 \end{aligned}$$

Consequently, $\text{Ker}(T) = \{\mathbf{0}\}$ and therefore $\text{nullity}(T) = 0$

CYU 4.9 We first show that $\text{Ker}(T) = \{\mathbf{0}\}$:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}) = \mathbf{0} \Rightarrow (\mathbf{2}\boldsymbol{a}, \boldsymbol{b}+\boldsymbol{c}, \boldsymbol{c}, \boldsymbol{b}) = (\mathbf{0}, \mathbf{0}, \mathbf{0}, \mathbf{0}) \Rightarrow \left.\begin{aligned} 2a &= 0 \\ b+c &= 0 \\ c &= 0 \\ b &= 0 \end{aligned}\right\} \Rightarrow a = b = c = 0$$

At this point we know that $\text{nullity}(T) = 0$ and that the kernel has no basis. Since $\text{rank}(T) + 0 = 3$, $\text{Im}(T)$ has dimension 3. It is easy to see that the three vectors $T(\mathbf{1}, \mathbf{0}, \mathbf{0}) = (\mathbf{2}, \mathbf{0}, \mathbf{0}, \mathbf{0})$, $T(\mathbf{0}, \mathbf{1}, \mathbf{0}) = (\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{1})$, $T(\mathbf{0}, \mathbf{0}, \mathbf{1}) = (\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{0})$ in the image of T are linearly independent. It follows that $\{(\mathbf{2}, \mathbf{0}, \mathbf{0}, \mathbf{0}), (\mathbf{0}, \mathbf{1}, \mathbf{0}, \mathbf{1}), (\mathbf{0}, \mathbf{1}, \mathbf{1}, \mathbf{0})\}$ is a basis for $\text{Im}(T)$.

CYU 4.10 Let $\bar{\boldsymbol{v}} \in V$ be such that $T(\boldsymbol{v}) = T(\bar{\boldsymbol{v}}) \Rightarrow \boldsymbol{v} = \bar{\boldsymbol{v}}$ (*) . We establish that T is one -to-one by showing that $\text{Ker}(T) = 0$ [see Theorem 4.11(a)]:

Assume that $T(\boldsymbol{v}) = 0$ (we want to show that $\boldsymbol{v} = \mathbf{0}$). Consider the vector $\boldsymbol{v} + \bar{v}$.

By linearity we have: $T(\boldsymbol{v} + \bar{\boldsymbol{v}}) = T(v) + T(\bar{\boldsymbol{v}}) = \mathbf{0} + T(\bar{\boldsymbol{v}}) = T(\bar{\boldsymbol{v}})$.

From (*) and the fact that $T(\boldsymbol{v} + \bar{\boldsymbol{v}}) = T(\bar{\boldsymbol{v}})$ we have: $\boldsymbol{v} + \bar{\boldsymbol{v}} = \bar{\boldsymbol{v}}$

$$\boldsymbol{v} = \mathbf{0}$$

CYU 4.11 $f^{-1}\colon Y \to X$ is one-to-one:

$$f^{-1}(a) = f^{-1}(b) \Rightarrow f[f^{-1}(a)] = f[[f^{-1}(b)]] \Rightarrow (f \circ f^{-1})(a) = (f \circ f^{-1})b \Rightarrow a = b$$

$f^{-1}\colon Y \to X$ is onto: For any $x \in X$, $f^{-1}[f(x)] = (f^{-1} \circ f)x = x$.

CYU 4.12 T is one-to-one: $T(\boldsymbol{a}, \boldsymbol{b}) = T(\boldsymbol{a}', \boldsymbol{b}') \Rightarrow (a+b)x - a = (a'+b')x - a'$

Equating coefficients: $\left.\begin{aligned} a &= a' \\ (a+b) &= a'+b' \end{aligned}\right\} \Rightarrow a = a'$ and $b = b'$

T is onto: For given $ax+b \in P_1$, we need to find $(A, B) \in \Re^2$ such that:

$$T(\boldsymbol{A}, \boldsymbol{B}) = (A+B)x - A = ax+b$$

$$\text{Equating coefficients:} \quad \left.\begin{aligned} -A &= b \\ A+B &= a \end{aligned}\right\} \Rightarrow A = -b \text{ and } B = a - A = a+b$$

$$\text{Check: } T(-\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}) = (-b+a+b)x - (-b) = ax+b.$$

Determining T^{-1}: Let $T^{-1}(ax+b) = (\boldsymbol{A}, \boldsymbol{B})$. Then:

$$\text{Applying } T \text{ to both sides:} \quad ax+b = T(\boldsymbol{A}, \boldsymbol{B}) \Rightarrow ax+b = (A+B)x - A$$

$$\text{Equating coefficients:} \quad \left.\begin{aligned} -A &= b \\ A+B &= a \end{aligned}\right\} \Rightarrow A = -b \text{ and } B = a - A = a+b$$

From the above: $T^{-1}(ax+b) = (-\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b})$. Let's show that the function $L: P_1 \rightarrow \Re^2$ given by $L(ax+b) = (-\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b})$ is linear:

$$\begin{aligned} L[r(ax+b)+(a'x+b')] &= L[(ra+a')x+(rb+b')] \\ &= [-(\boldsymbol{rb}+\boldsymbol{b'}), (\boldsymbol{ra}+\boldsymbol{a'})+(\boldsymbol{rb}+\boldsymbol{b'})] \\ &= r(-\boldsymbol{b}, \boldsymbol{a}+\boldsymbol{b}) + (-\boldsymbol{b'}, \boldsymbol{a'}+\boldsymbol{b'}) = rL(ax+b) + L(a'x+b') \end{aligned}$$

<u>CYU 4.13</u> We show that $T: \Re^4 \rightarrow M_{2\times 2}$ given by $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is an isomorphism:

T is one-to-one:

$$T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = T(\boldsymbol{a'}, \boldsymbol{b'}, \boldsymbol{c'}, \boldsymbol{d'}) \Rightarrow \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} \Rightarrow a = a', b = b', c = c', d = d'$$

T is onto: For given $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}$, $T(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$.

T is linear: $T\left(r\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}\right) = T\begin{bmatrix} ra+\bar{a} & rb+\bar{b} \\ rc+\bar{c} & rd+\bar{d} \end{bmatrix} = rT\begin{bmatrix} a & b \\ c & d \end{bmatrix} + T\begin{bmatrix} \bar{a} & \bar{b} \\ \bar{c} & \bar{d} \end{bmatrix}$.

<u>CYU 4.14</u> Let $\dim(V) = \dim(W) = n$. By Theorem 4.15, $V \cong \Re^n$ and $W \cong \Re^n$. By Theorem 4.14, $V \cong W$.

Conversely, suppose that $V \cong W$. Let $T: V \rightarrow W$ be an isomorphism, and let $\beta_V = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. We show $S = \{T(v_1), T(v_2), \ldots, T(v_n)\}$ is a basis for W, thereby establishing that V and W are of the same dimension, n.

***S* is linearly independent:**

$$\sum_{i=1}^{n} a_iT(v_i) = 0 \Rightarrow T\left(\sum_{i=1}^{n} a_iv_i\right) = 0 \underset{\substack{\uparrow\\ \text{since } T \text{ is linear} \quad \text{Theorem 4.11(a), page 129}}}{\Rightarrow} \sum_{i=1}^{n} a_iv_i = 0 \overset{\text{since } \beta_V \text{ is a linearly independent set}}{\Rightarrow} a_i = 0 \text{ for } 1 \le i \le n$$

***S* spans *W*:** Let $w \in W$. Since T is onto, there exist $v = \sum_{i=1}^{n} a_iv_i$ such that $T(v) = w$. Then:

$$w = T(v) = T\left(\sum_{i=1}^{n} a_iv_i\right) \underset{\substack{\uparrow\\ \text{since } T \text{ is linear}}}{=} \sum_{i=1}^{n} a_iT(v_i)$$

CYU 4.15 By CYU 4.15, we know that the dimension of W equals that of V. We can therefore verify that $\{L(\mathbf{v_1}), L(\mathbf{v_2}), \ldots, L(\mathbf{v_n})\}$ is a basis for W by showing that the n vectors $\{L(\mathbf{v_1}), L(v_2), \ldots, L(\mathbf{v_n})\}$ span W (see Theorem 3.11, page 99):

Let $\mathbf{w} \in W$. Since L is onto, there exist $v = \sum_{i=1}^{n} a_i\mathbf{v_i}$ such that $L(v) = \mathbf{w}$. Since L is linear: $w = L(v) = \sum_{i=1}^{n} a_iL(v_i)$.

CYU 4.16 Lets move the elements of:

$$S = \{2x^3 - 3x^2 + 5x - 1, x^3 - x^2 + 8x - 3, x^2 + 11x - 5, -x^3 + 2x^2 + 3x - 2\}$$

over to $\Re^4$ via the isomorphism $T(ax^3 + bx^2 + cx + d) = (a, b, c, d)$ to arrive at the 4-tuples:

$$T(S) = \{(\mathbf{2, -3, 5, -1}), (\mathbf{1, -1, 8, -3}), (\mathbf{0, 1, 11, -5}), (\mathbf{-1, 2, 3, -2})\}$$

Applying Theorem 3.13, page 103, we conclude that the first two vectors of $T(S)$, $(\mathbf{2, -3, 5, -1})$ and $(\mathbf{1, -1, 8, -3})$, constitute a basis for Span $T(S)$. It follows that $2x^3 - 3x^2 + 5x - 1$ and $x^3 - x^2 + 8x - 3$ is a basis for Span(S).

$$\begin{bmatrix} 2 & 1 & 0 & -1 \\ -3 & -1 & 1 & 2 \\ 5 & 8 & 11 & 3 \\ -1 & -3 & -5 & -2 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

CYU 4.17 (a) $f(x, y, z) = (2x + 1, x + y, -z)$ is one-to-one:

$$f(x_1, y_1, z_1) = f(x_2, y_2, z_2)$$

$$(2x_1 + 1, x_1 + y_1, -z_1) = (2x_2 + 1, x_2 + y_2, -z_2) \Rightarrow \begin{cases} 2x_1 + 1 = 2x_2 + 1 \\ x_1 + y_1 = x_2 + y_2 \\ -z_1 = -z_2 \end{cases} \Rightarrow \begin{cases} x_1 = x_2 \\ y_1 = y_2 \\ z_1 = z_2 \end{cases}$$

f is onto: For given $(x, y, z) \in X$ we need to find $(a, b, c) \in \Re^3$ such that:

$$f(a, b, c) = (x, y, z)$$

$$(2a+1, a+b, -c) = (x, y, z) \Rightarrow \begin{cases} 2a+1 = x \\ a+b = y \\ -c = z \end{cases} \Rightarrow \begin{cases} a = \dfrac{x-1}{2} \\ b = y-a = y-\dfrac{x-1}{2} \\ c = -z \end{cases}$$

From the above formula: $f\left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right) = (x, y, z)$, we conclude that:

$$f^{-1}(x, y, z) = \left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right)$$

(b) From Theorem 4.16 and the above formula $f^{-1}(x, y, z) = \left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right)$ we have:

$$\begin{aligned} (x_1, y_1, z_1) \oplus (x_2, y_2, z_2) &= f\left[\left(\frac{x_1-1}{2}, y_1-\frac{x_1-1}{2}, -z_1\right)+\left(\frac{x_2-1}{2}, y_2-\frac{x_2-1}{2}, -z_2\right)\right] \\ &= f\left(\frac{x_1}{2}+\frac{x_2}{2}-1, y_1+y_2-\frac{x_1}{2}-\frac{x_2}{2}+1, -z_1-z_2\right) \end{aligned}$$

Since $f(x, y, z) = (2x, x+y, -z)$:

$$\begin{aligned} &= \left[2\left(\frac{x_1}{2}+\frac{x_2}{2}-1\right), \frac{x_1}{2}+\frac{x_2}{2}-1+y_1+y_2-\frac{x_1}{2}-\frac{x_2}{2}+1, -(-z_1-z_2)\right] \\ &= (x_1+x_2-2, y_1+y_2, z_1+z_2) \end{aligned}$$

and: $r \otimes (x, y, z) = f\left[r\left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right)\right] = f\left(\frac{rx-r}{2}, ry-\frac{rx-r}{2}, -rz\right)$

Since $f(x, y, z) = (2x, x+y, -z)$:

$$\begin{aligned} &= \left(2\cdot\frac{rx-r}{2}, \frac{rx-r}{2}+ry-\frac{rx-r}{2}, -(-rz)\right) \\ &= (rx-r, ry, rz) \end{aligned}$$

(c) The zero in the space *X*: $f(0, 0, 0) = (1, 0, 0)$.

The inverse of (x, y, z) in *X*:

$$f[-f^{-1}(x, y, z)] \underset{\uparrow}{=} f\left[-\left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right)\right] = f\left(\frac{-x+1}{2}, -y+\frac{x-1}{2}, z\right) \underset{\uparrow}{=} (-x+2, -y, z)$$

$f^{-1}(x, y, z) = \left(\frac{x-1}{2}, y-\frac{x-1}{2}, -z\right)$

$f(x, y, z) = (2x+1, x+y, -z)$

CHAPTER 5
MATRICES AND LINEAR MAPS

CYU 5.1 (a) $\begin{bmatrix}3&5\\4&2\\9&0\end{bmatrix}\begin{bmatrix}6&4\\3&5\end{bmatrix} = \begin{bmatrix}3\cdot6+5\cdot3 & 3\cdot4+5\cdot5\\ 4\cdot6+2\cdot3 & 4\cdot4+2\cdot5\\ 9\cdot6+0\cdot3 & 9\cdot4+0\cdot3\end{bmatrix} = \begin{bmatrix}33&37\\30&26\\54&36\end{bmatrix}$

(b) Number of columns in A does not equal the number of rows in B.

CYU 5.2 False:

$$\left(\begin{bmatrix}1&1\\1&1\end{bmatrix}+\begin{bmatrix}1&2\\0&0\end{bmatrix}\right)^2 = \left(\begin{bmatrix}1&1\\1&1\end{bmatrix}+\begin{bmatrix}1&2\\0&0\end{bmatrix}\right)\left(\begin{bmatrix}1&1\\1&1\end{bmatrix}+\begin{bmatrix}1&2\\0&0\end{bmatrix}\right) = \begin{bmatrix}2&3\\1&1\end{bmatrix}\begin{bmatrix}2&3\\1&1\end{bmatrix} = \begin{bmatrix}7&7\\3&4\end{bmatrix}$$

While: $$\begin{bmatrix}1&1\\1&1\end{bmatrix}\begin{bmatrix}1&1\\1&1\end{bmatrix}+2\begin{bmatrix}1&1\\1&1\end{bmatrix}\begin{bmatrix}1&2\\0&0\end{bmatrix}+\begin{bmatrix}1&2\\0&0\end{bmatrix}\begin{bmatrix}1&2\\0&0\end{bmatrix} = \begin{bmatrix}2&2\\2&2\end{bmatrix}+\begin{bmatrix}2&4\\2&4\end{bmatrix}+\begin{bmatrix}1&2\\0&0\end{bmatrix} = \begin{bmatrix}5&8\\4&6\end{bmatrix}$$

CYU 5.3 $\begin{bmatrix}2&5&-3&4\\-4&-10&6&-8\\0&1&2&-4\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}1&0&-\frac{13}{2}&12\\0&1&2&-4\\0&0&0&0\end{bmatrix}$ The columns associated with the leading ones in rref(A); namely: $(2, -4, 0)$ and $(5, -10, 1)$ constitute a basis for the column space of A.

Since A and rref(A) share a row space, $\left(1, 0, -\frac{13}{2}, 12\right)$ and $(0, 1, 2, -4)$ constitutes a basis for the row space of A. Note that the first two rows of A are not linearly independent. We are assured, however, that two of A's rows will constitute a basis for its row space (either rows 1 and 3, or rows 2 and 3, will do the trick).

CYU 5.4 If $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ are solutions of the homogeneous system of m equations in n unknowns of equations, $A\boldsymbol{X} = \boldsymbol{0}$, and if $r \in \Re$, then:

$$A(r\boldsymbol{X}_1 + \boldsymbol{X}_2) = rA\boldsymbol{X}_1 + A\boldsymbol{X}_2 = r\boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$$

It follows, from Theorem2.13, page 61, that the solutions set of $A\boldsymbol{X} = \boldsymbol{0}$ is a subspace of $\Re^n$.

CYU 5.5 From $\begin{bmatrix}2&1&3&0\\1&4&-2&-7\\3&0&1&-2\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}1&0&0&-1\\0&1&0&-1\\0&0&1&1\end{bmatrix}$, we see that: $\text{null}\left(\begin{bmatrix}2&1&3&0\\1&4&-2&-7\\3&0&1&-2\end{bmatrix}\right) = \{(c, c, -c, c) | (c \in \Re)\}$

A basis: $\{(1, 1, -1, 1)\}$

CYU 5.6 Let $A \in M_{m \times n}$. By definition:

$$\text{nullity}(A) = \dim\{X | AX = 0\} = \dim\{X | T_A X = 0\} = \text{nullity}(T_A)$$

Since $\text{nullity}(A)$ equals the number of free variables in $\text{rref}(A)$ and since $\text{rank}(A)$ equals the number of leading ones in $\text{rref}(A)$: $\text{rank}(A) = n - \text{nullity}(A)$. By Theorem 4.10, page 126: $\text{rank}(T_A) = n - \text{nullity}(T_A)$. Since $\text{nullity}(A) = \text{nullity}(T_A)$: $\text{rank}(A) = \text{rank}(T_A)$.

CYU 5.7

$$\begin{bmatrix}3&-2\\4&-1\end{bmatrix}\begin{bmatrix}a&b\\c&d\end{bmatrix}=\begin{bmatrix}1&0\\0&1\end{bmatrix}\Rightarrow\left.\begin{matrix}3a-2c=1\\4a-c=0\end{matrix}\right\}\text{ and }\left.\begin{matrix}3b-2d=0\\4b-d=1\end{matrix}\right\}\Rightarrow a=-\frac{1}{5},\ b=\frac{2}{5},\ c=-\frac{4}{5},\ d=\frac{3}{5}$$

From the above, we see that the matrix $\begin{bmatrix}3&-2\\4&-1\end{bmatrix}$ is invertible, with inverse $\begin{bmatrix}-1/5&2/5\\-4/5&3/5\end{bmatrix}$.

CYU 5.8 I. Theorem 5.8 tells us that the claim is valid for $n=2$: $(A_1A_2)^{-1}=A_2^{-1}A_1^{-1}$.

II. Assume validity at $n=k$: $(A_1A_2\cdots A_k)^{-1}=A_k^{-1}A_{k-1}^{-1}\ldots A_1^{-1}$

III. We establish validity at $n=k+1$:

$$\begin{aligned}(A_1A_2\cdots A_kA_{k+1})^{-1}&=[(A_1A_2\ldots A_k)A_{k+1}]^{-1}\\ \text{By I:}\quad&=A_{k+1}^{-1}(A_1A_2\ldots A_k)^{-1}\\ \text{By II:}\quad&=A_{k+1}(A_k^{-1}A_{k-1}^{-1}\ldots A_1^{-1})=A_{k+1}A_k^{-1}A_{k-1}^{-1}\ldots A_1^{-1}\end{aligned}$$

CYU 5.9

$$\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}^{-1}=\begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}\quad\Bigg|\quad\begin{bmatrix}1&0&0&0\\0&5&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}^{-1}=\begin{bmatrix}1&0&0&0\\0&\frac{1}{5}&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}\quad\Bigg|\quad\begin{bmatrix}1&6\\0&1\end{bmatrix}^{-1}=\begin{bmatrix}1&-6\\0&1\end{bmatrix}$$

CYU 5.10 (a)

$$\begin{bmatrix}2&1&0&3\\1&3&2&6\\3&3&4&1\end{bmatrix}\xrightarrow{R_1\leftrightarrow R_3}\begin{bmatrix}3&3&4&1\\1&3&2&6\\2&1&0&3\end{bmatrix},\ \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\xrightarrow{R_1\leftrightarrow R_3}\overset{E}{\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}},\text{ and }\begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}\begin{bmatrix}2&1&0&3\\1&3&2&6\\3&3&4&1\end{bmatrix}=\begin{bmatrix}3&3&4&1\\1&3&2&6\\2&1&0&3\end{bmatrix}$$

(b)

$$\begin{bmatrix}2&1&0&3\\1&3&2&6\\3&3&4&1\end{bmatrix}\xrightarrow{2R_2\rightarrow R_2}\begin{bmatrix}3&3&4&1\\2&6&4&12\\2&1&0&3\end{bmatrix},\ \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\xrightarrow{2R_2\rightarrow R_2}\overset{E}{\begin{bmatrix}0&0&1\\0&2&0\\1&0&0\end{bmatrix}},\text{ and }\begin{bmatrix}0&0&1\\0&2&0\\1&0&0\end{bmatrix}\begin{bmatrix}2&1&0&3\\1&3&2&6\\3&3&4&1\end{bmatrix}=\begin{bmatrix}3&3&4&1\\2&6&4&12\\2&1&0&3\end{bmatrix}$$

CYU 5.11

$$[A|I]=\left[\begin{array}{cccc|cccc}\multicolumn{4}{c|}{\overbrace{\hphantom{1\ 0\ 1\ 2}}^{A}}&&&&\\1&0&1&2&1&0&0&0\\0&2&1&4&0&1&0&0\\1&1&1&0&0&0&1&0\\0&3&1&1&0&0&0&1\end{array}\right]\qquad\left[\begin{array}{cccc|cccc}&&&&\multicolumn{4}{c}{\overbrace{\hphantom{-7\ 5\ 8\ -6}}^{A^{-1}}}\\1&0&0&0&-7&5&8&-6\\0&1&0&0&-3&2&3&-2\\0&0&1&0&10&-7&-10&8\\0&0&0&1&-1&1&1&-1\end{array}\right]$$

CYU 5.12 Assume $AB=I$. Multiplying both sides of the equation $BX=\mathbf{0}$ by A we have:

$$(AB)X=A\mathbf{0}\Rightarrow IX=\mathbf{0}\Rightarrow X=\mathbf{0}\Rightarrow B\text{ is invertible (Theorem 5.17)}$$

Then: $AB=I\Rightarrow(AB)B^{-1}=B^{-1}\Rightarrow A(BB^{-1})=B^{-1}\Rightarrow A=B^{-1}\Rightarrow A^{-1}=B$.

CYU 5.13 $\left[\begin{array}{ccc|c} 1 & 2 & 0 & -1 \\ 3 & 0 & 1 & 1 \\ 0 & 4 & 2 & 2 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & -1/4 \\ 0 & 1 & 0 & -3/8 \\ 0 & 0 & 1 & 7/4 \end{array}\right] \Rightarrow \begin{bmatrix} -1 \\ 1 \\ 2 \end{bmatrix}_\beta = \begin{bmatrix} -1/4 \\ -3/8 \\ 7/4 \end{bmatrix}$

CYU 5.14

$$T\left(\begin{bmatrix} 3 & 3 \\ 2 & 2 \end{bmatrix}\right) = (3, 3, 4), T\left(\begin{bmatrix} 7 & 2 \\ 1 & 1 \end{bmatrix}\right) = (2, 7, 2), T\left(\begin{bmatrix} 1 & -1 \\ 5 & 1 \end{bmatrix}\right) = (-1, 1, 6), T\left(\begin{bmatrix} 2 & 6 \\ 5 & 9 \end{bmatrix}\right) = (6, 2, 14)$$

$$\begin{array}{c} \left[\begin{array}{ccc|cccc} 1 & 2 & 0 & 3 & 2 & -1 & 6 \\ 3 & 0 & 1 & 3 & 7 & 1 & 2 \\ 0 & 4 & 2 & 4 & 2 & 6 & 14 \end{array}\right] \\ \gamma \qquad\quad T(\beta) \end{array} \xrightarrow{\text{rref}} \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 1 & 2 & -3/4 & 1/4 \\ 0 & 1 & 0 & 1 & 0 & -1/8 & 23/8 \\ 0 & 0 & 1 & 0 & 1 & 13/4 & 5/4 \end{array}\right] \Rightarrow [T]_{\gamma\beta} = \begin{bmatrix} 1 & 2 & -3/4 & 1/4 \\ 1 & 0 & -1/8 & 23/8 \\ 0 & 1 & 13/4 & 5/4 \end{bmatrix}$$

CYU 5.15 $T\left(\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\right) = (\mathbf{3}, \mathbf{1}, \mathbf{2})$ and $\left[\begin{array}{ccc|c} 1 & 2 & 0 & \mathbf{3} \\ 3 & 0 & 1 & \mathbf{1} \\ 0 & 4 & 2 & \mathbf{2} \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 3/4 \\ 0 & 1 & 0 & 9/8 \\ 0 & 0 & 1 & -5/4 \end{array}\right] \Rightarrow \left[T\left(\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\right)\right]_\gamma = \begin{bmatrix} 3/4 \\ 9/8 \\ -5/4 \end{bmatrix}$

$$\left[\begin{array}{ccc|c|c} 3 & 7 & 1 & 2 & \mathbf{1} \\ 3 & 2 & -1 & 6 & \mathbf{3} \\ 2 & 1 & 5 & 5 & \mathbf{2} \\ 2 & 1 & 1 & 9 & \mathbf{0} \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 69/28 \\ 0 & 1 & 0 & 0 & -11/14 \\ 0 & 0 & 1 & 0 & 1/28 \\ 0 & 0 & 0 & 1 & -13/28 \end{array}\right] \Rightarrow \left[\begin{bmatrix} \mathbf{1} & \mathbf{3} \\ \mathbf{2} & \mathbf{0} \end{bmatrix}\right]_\beta = \begin{bmatrix} 69/28 \\ -11/14 \\ 1/28 \\ -13/28 \end{bmatrix}$$

$$[T]_{\gamma\beta}\left[\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\right]_\beta = \underbrace{\begin{bmatrix} 1 & 2 & -3/4 & 1/4 \\ 1 & 0 & -1/8 & 23/8 \\ 0 & 1 & 13/4 & 5/4 \end{bmatrix}}_{\text{From CYU 5.15}}\begin{bmatrix} 69/28 \\ -11/14 \\ 1/28 \\ -13/28 \end{bmatrix} = \begin{bmatrix} 3/4 \\ 9/8 \\ -5/4 \end{bmatrix} = \left[T\left(\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\right)\right]_\gamma$$

CYU 5.16

$$[T]_{\gamma\beta}\text{:}\ \left.\begin{aligned} T(\mathbf{1, 1, 1}) &= 1x^2 + 1x + 1 \\ T(\mathbf{1, 1, 0}) &= 1x^2 + 1x + 0 \\ T(\mathbf{1, 0, 0}) &= 0x^2 + 1x + 0 \end{aligned}\right\} \Rightarrow \begin{array}{c} \gamma \quad\ T(\beta) \\ \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 0 & 0 \end{array}\right] \end{array} \quad \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1/2 & 0 & 0 \end{array}\right] \Rightarrow [T]_{\gamma\beta} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1/2 & 0 & 0 \end{bmatrix}$$

$$[L]_{\delta\gamma}\text{:}\ \left.\begin{aligned} L(x^2) &= (\mathbf{1, 1}) \\ L(x) &= (\mathbf{0, 1}) \\ L(2) &= (\mathbf{0, 2}) \end{aligned}\right\} \Rightarrow \begin{array}{c} \delta \quad L(\gamma) \\ \left[\begin{array}{cc|ccc} 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 2 \end{array}\right] \end{array} \quad \left[\begin{array}{cc|ccc} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & 0 \end{array}\right] \Rightarrow [L]_{\delta\gamma} = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & 0 \end{bmatrix}$$

$$[L\circ T]_{\delta\beta}\text{:}\ \left.\begin{aligned} (L\circ T)(\mathbf{1, 1, 1}) &= L[T(\mathbf{1, 1, 1})] = L(x^2 + x + 1) = (\mathbf{1, 3}) \\ (L\circ T)(\mathbf{1, 1, 0}) &= L[T(\mathbf{1, 1, 0})] = L(x^2 + x + 0) = (\mathbf{1, 2}) \\ (L\circ T)(\mathbf{1, 0, 0}) &= L[T(\mathbf{1, 0, 0})] = L(0x^2 + x + 0) = (\mathbf{0, 1}) \end{aligned}\right\} \Rightarrow \begin{array}{c} \delta \ \ (L\circ T)(\beta) \\ \left[\begin{array}{cc|ccc} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 3 & 2 & 1 \end{array}\right] \end{array} \quad \left[\begin{array}{cc|ccc} 1 & 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{array}\right] \Rightarrow [L\circ T]_{\delta\beta} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

CYU 5.17 Since $[T]_{\gamma\beta}$ is invertible, it must be a square matrix of dimension n. It follows that V and W are both of dimension n. Let $\beta = \{v_1, v_2, \ldots, v_n\}$, $\gamma = \{w_1, w_2, \ldots, w_n\}$, and $[T]_{\gamma\beta}^{-1} = [a_{ij}]$. Consider the linear transformation $L: W \to V$ which maps w_i to the vector $a_{1i}v_1 + a_{2i}v_2 + \ldots + a_{ni}v_n$. From its very definition, we see that $[L]_{\beta\gamma} = [a_{ij}] = [T]_{\gamma\beta}^{-1}$ (note, for example that $[L(w_1)]_\beta$ is the first column of $[a_{ij}]$). Applying Theorem 5.22, we see that: $[L \circ T]_{\beta\beta} = [L]_{\beta\gamma}[T]_{\gamma\beta} = [T]_{\gamma\beta}^{-1}[T]_{\gamma\beta} = I$. It follows that $L \circ T$: $V \to V$ is the identity map, and that therefore T is invertible with inverse L.

CYU 5.18 $[I_V]_{\beta'\beta}$ and $[(\mathbf{2}, \mathbf{3})]_{\beta'}$:

$$\left[\begin{array}{cc|ccc} 0 & 2 & 1 & 2 & 2 \\ 3 & -1 & 2 & 1 & 3 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc|c} & & \multicolumn{2}{c|}{[I_V]_{\beta'\beta}} & [(2,3)]_{\beta'} \\ 1 & 0 & 5/6 & 2/3 & 4/3 \\ 0 & 1 & 1/2 & 1 & 1 \end{array}\right]$$

$$[(\mathbf{2}, \mathbf{3})]_\beta : \left[\begin{array}{cc|c} 1 & 2 & 2 \\ 2 & 1 & 3 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|c} 1 & 0 & 4/3 \\ 0 & 1 & 1/3 \end{array}\right]. \quad [I_V]_{\beta'\beta}[(\mathbf{2}, \mathbf{3})]_\beta = \begin{bmatrix} 5/6 & 2/3 \\ 1/2 & 1 \end{bmatrix}\begin{bmatrix} 4/3 \\ 1/3 \end{bmatrix} = \begin{bmatrix} 4/3 \\ 1 \end{bmatrix} = [(\mathbf{2}, \mathbf{3})]_{\beta'}.$$

CYU 5.19 Rotating the standard basis $\beta = \{(\mathbf{1}, \mathbf{0}), (\mathbf{0}, \mathbf{1})\}$ clockwise by 60° leads us to the basis $\beta' = \left\{\left(\frac{\mathbf{1}}{\mathbf{2}}, -\frac{\sqrt{\mathbf{3}}}{\mathbf{2}}\right), \left(\frac{\mathbf{1}}{\mathbf{2}}, \frac{\sqrt{\mathbf{3}}}{\mathbf{2}}\right)\right\}$. Finding $[I]_{\beta'\beta}$ and $[(1, 3)]_{\beta'}$:

$$\left[\begin{array}{cc|ccc} \frac{1}{2} & \frac{1}{2} & 1 & 0 & 1 \\ -\sqrt{3}/2 & \sqrt{3}/2 & 0 & 1 & 3 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{cc|cc|c} & & \multicolumn{2}{c|}{[I]_{\beta'\beta}} & [(1,3)]_{\beta'} \\ 1 & 0 & 1 & -\sqrt{3}/2 & 1 - \sqrt{3} \\ 0 & 1 & 1 & \sqrt{3}/2 & \sqrt{3} + 1 \end{array}\right]$$

$$\text{Check: } (1 - \sqrt{3})\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right) + (\sqrt{3} + 1)\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = (1, 3)$$

CYU 5.20 For $[T]_{\beta'\beta'}$ and $[I]_{\beta'\beta}$. Noting that $T(x^2 + 1) = -x + 2$, $T(x^2 - x) = -x^2 - x$, and $T(1) = 2$; and that $\beta = \{x^2, x^2 + x, x^2 + x + 1\}$, $\beta' = \{x^2 + 1, x^2 - x, 1\}$; leads us to:

$$\left[\begin{array}{ccc|ccc|ccc} \multicolumn{3}{c|}{\beta'} & \multicolumn{3}{c|}{T(\beta')} & \multicolumn{3}{c}{I(\beta)} \\ 1 & 1 & 0 & 0 & -1 & 0 & 1 & 1 & 1 \\ 0 & -1 & 0 & -1 & -1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 2 & 0 & 2 & 0 & 0 & 1 \end{array}\right] \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc|ccc} & & & \multicolumn{3}{c|}{[T]_{\beta'\beta'}} & \multicolumn{3}{c}{[I]_{\beta'\beta}} \\ 1 & 0 & 0 & -1 & -2 & 0 & 1 & 2 & 2 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 3 & 2 & 2 & -1 & -2 & -1 \end{array}\right]$$

Noting that $T(x^2) = -x$, $T(x^2 + x) = x^2 - x$, and $T(x^2 + x + 1) = x^2 - x + 2$ leads us to:

$$\left[\overset{\beta}{\begin{matrix}1&1&1\\0&1&1\\0&0&1\end{matrix}}\;\middle|\;\overset{T(\beta)}{\boxed{\begin{matrix}0&1&1\\-1&-1&-1\\0&0&2\end{matrix}}}\;\middle|\;\overset{I(\beta')}{\begin{matrix}1&1&0\\0&-1&0\\1&0&1\end{matrix}}\right]\xrightarrow{\text{rref}}\left[\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}\;\middle|\;\overset{[T]_{\beta\beta}}{\boxed{\begin{matrix}1&2&2\\-1&-1&-3\\0&0&2\end{matrix}}}\;\middle|\;\overset{[I]_{\beta\beta'}}{\begin{matrix}1&2&0\\-1&-1&-1\\1&0&1\end{matrix}}\right]$$

Then: $[I]_{\beta'\beta}[T]_{\beta\beta}[I]_{\beta\beta'} = \begin{bmatrix}1&2&2\\0&-1&-1\\-1&-2&-1\end{bmatrix}\begin{bmatrix}1&2&2\\-1&-1&-3\\0&0&2\end{bmatrix}\begin{bmatrix}1&2&0\\-1&-1&-1\\1&0&1\end{bmatrix} = \begin{bmatrix}-1&-2&0\\-1&-1&-1\\1&0&1\end{bmatrix} = [T]_{\beta'\beta'}$

<u>**CYU 5.21**</u> (One possible solution). Choosing to let $c = 0, d = 8$ in system (*) at the bottom of page 197 we obtain the solution $a = -18, b = 11, c = 0, d = 8$. At this point, we know that: $\begin{bmatrix}-12&8\\-18&15\end{bmatrix} = \begin{bmatrix}-18&11\\0&8\end{bmatrix}^{-1}\begin{bmatrix}-1&4\\8&4\end{bmatrix}\begin{bmatrix}-18&11\\0&8\end{bmatrix}$. Preceding as in part (c) of Example 5.13, we arrive at the basis $\beta'' = \{\boldsymbol{v_1}'', \boldsymbol{v_2}''\}$, with:

$v_1'' = -18(\mathbf{1, 2}) + 0(\mathbf{2, 1}) = (\mathbf{-18, -36})$ and $v_2'' = 11(\mathbf{1, 2}) + 8(\mathbf{2, 1}) = (\mathbf{27, 30})$.

Let's verify that $[T]_{\beta''\beta''} = \begin{bmatrix}-12&8\\-18&15\end{bmatrix}$ for $\beta'' = \{(\mathbf{-18, -36}), (\mathbf{27, 30})\}$:

$$\begin{array}{c} T(\mathbf{-18, -36}) = (\mathbf{-270, -108}) \\ T(\mathbf{27, 30}) = (\mathbf{261, 162}) \end{array}$$

$$\left[\overset{\beta''}{\begin{matrix}-18&27\\-36&30\end{matrix}}\;\middle|\;\begin{matrix}\downarrow&\downarrow\\-270&261\\-108&162\end{matrix}\right]\xrightarrow{\text{rref}}\left[\begin{matrix}1&0\\0&1\end{matrix}\;\middle|\;\underset{[T]_{\beta'\beta'}}{\boxed{\begin{matrix}-12&8\\-18&15\end{matrix}}}\right]$$

CHAPTER 6
DETERMINANTS AND EIGENVECTORS

CYU 6.1 (a-i) $\det\begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix} = 5\det\begin{bmatrix} 9 & -3 \\ -2 & 4 \end{bmatrix} - 7\det\begin{bmatrix} 2 & -3 \\ 3 & 4 \end{bmatrix} - 6\det\begin{bmatrix} 2 & 9 \\ 3 & -2 \end{bmatrix} = 5(36-6) - 7(8+9) - 6(-4-27) = 217$

(a-ii) $\det\begin{bmatrix} 2 & 9 & -3 \\ 3 & -2 & 4 \\ 5 & 7 & -6 \end{bmatrix} = -9\det\begin{bmatrix} 3 & 4 \\ 5 & -6 \end{bmatrix} - 2\det\begin{bmatrix} 2 & -3 \\ 5 & -6 \end{bmatrix} - 7\det\begin{bmatrix} 2 & -3 \\ 5 & -6 \end{bmatrix} = -9(-18-20) - 2(-12+15) - 7(8+9) = 217$

CYU 6.2 By induction on the dimension, n, of $M_{n \times n}$.

I. Claim holds for $n = 2$: $\det\begin{bmatrix} a & 0 \\ c & d \end{bmatrix} = ad - 0 \cdot c = ad$.

II. Assume claim holds for $n = k$.

III. We establish validity at $n = k+1$: Let $A = [a_{ij}] \in M_{(k+1) \times (k+1)}$. Since all entries in the first row after its first entry a_{11} is zero, expanding about the first row of A we have $\det(A) = a_{11}\det(B)$, where B is the k by k lower triangular matrix obtained by removing the first row and first column from the matrix A. As such, by II: $\det(B) = a_{22}a_{33}\cdots a_{(k+1)(k+1)}$. Consequently: $\det(A) = a_{11}\det(B) = a_{11}a_{22}a_{33}\cdots a_{(k+1)(k+1)}$

CYU 6.3 Let the i^{th} and j^{th} row of A be identical, with $i \neq j$. Multiplying row i by -1 and adding it to row j we obtain a matrix B whose j^{th} row consists entirely of zeros. As such $\det(B) = 0$ (just expand about the j^{th} row of B). Applying Theorem 6.3(c), we conclude that $\det(A) = 0$.

CYU 6.4

$$\det\begin{bmatrix} 2 & 1 & 0 & 1 \\ 0 & 1 & 2 & 2 \\ 1 & 0 & 1 & 4 \\ 4 & 1 & 1 & 3 \end{bmatrix} \overset{\substack{\text{Theorem 6.4(a)} \\ \downarrow}}{=} -\det\begin{bmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 2 \\ 2 & 1 & 0 & 1 \\ 4 & 1 & 1 & 3 \end{bmatrix} \overset{\substack{\text{Theorem 6.4(c) (two times)} \\ \downarrow}}{=} -\det\begin{bmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & -2 & -7 \\ 0 & 1 & -3 & -13 \end{bmatrix}$$

$$\longrightarrow = -\det\begin{bmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -4 & -9 \\ 0 & 0 & -5 & -15 \end{bmatrix} = -\det\begin{bmatrix} 1 & 0 & 1 & 4 \\ 0 & 1 & 2 & 2 \\ 0 & 0 & -4 & -9 \\ 0 & 0 & 0 & -\frac{15}{4} \end{bmatrix} = \underset{\substack{\uparrow \\ \text{CYU 6.2(b)}}}{(-4)\left(-\frac{15}{4}\right)} = 15$$

CYU 6.5 (a-i) Let $B \in M_{n \times n}$, and let E be the elementary matrix obtained by multiplying row i of I_n by $c \neq 0$. By Theorem 5.12, *EB* is the matrix obtained by multiplying row i of B by c. Consequently:

$$\det(EB) \underset{\text{Theorem 6.4(b)}\uparrow}{=} c\det(B) \underset{\uparrow \det(E) = c \text{ [Theorem 6.4(b)]}}{=} \det(E)\det(B)$$

(a-ii) Let $B \in M_{n \times n}$, and let E be the elementary matrix obtained by adding a multiple of the i^{th} row of I_n to its j^{th} row. By Theorem 5.12, *EB* is the matrix obtained by adding a multiple of row i of B to its j^{th} row. Consequently:

$$\det(EB) \underset{\text{Theorem 6.4(b)}\uparrow}{=} \det(B) \underset{\uparrow \det(E) = 1}{=} \det(E)\det(B)$$

(b) We use the Principle of Mathematical Induction to show that for any $B \in M_{n \times n}$ and elementary matrices $E_1, E_2, \ldots, E_s \in M_{n \times n}$:

$$\det(E_s \cdots E_2 E_1 B) = \det(E_s \cdots E_2 E_1)\det(B) = \det(E_s)\cdots\det(E_2)\det(E_1)\det(B)$$

I. Validity for $s = 1$ follows from Theorem 6.5.

II. Assume $\det(E_k \cdots E_2 E_1 B) = \det(E_x \cdots E_2 E_1)\det(B) = \det(E_k)\cdots\det(E_2)\det(E_1)\det(B)$

III.
$$\begin{aligned}\det(E_{k+1} \cdot E_k \cdots E_2 E_1 B) &= \det[(E_{k+1})(E_k \cdots E_2 E_1 B)] \\ \text{By I:} \quad &= \det(E_{k+1})\det(E_k \cdots E_2 E_1 B) \\ \text{By II:} \quad &= \det(E_{k+1})(\det(E_k)\cdots\det(E_2)\det(E_1)\det(B))\end{aligned}$$

CYU 6.6 $I = AA^{-1} \Rightarrow \det(I) = \det(AA^{-1}) \Rightarrow 1 = \det(A)\det(A^{-1}) \Rightarrow \det(A^{-1}) = \dfrac{1}{\det(A)}$

CYU 6.7 $E[-3] = \text{null}\left(\begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix} - (-3)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix}\right)$. From $\begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 1/3 \\ 0 & 0 \end{bmatrix}$ we see that $E[-3] = \left\{\left(-\frac{r}{3}, r\right) \middle| r \in R\right\}$ with basis $\{\overline{(\mathbf{1, -3})}\}$.

CYU 6.8 The eigenvalues are the solutions of the equation:

$$\det(A - \lambda I_3) = \det\begin{bmatrix} 16-\lambda & 3 & 2 \\ -4 & 3-\lambda & -8 \\ -2 & -6 & 11-\lambda \end{bmatrix} = 0$$

Which reduces to $-\lambda(x-15)^2 = 0$. It follows that 0 and 15 are the eigenvalues of A.

Then: $E[0] = \text{null}\left(\begin{bmatrix} 16 & 3 & 2 \\ -4 & 3 & -8 \\ -2 & -6 & 11 \end{bmatrix} - 0\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = \text{null}\left(\begin{bmatrix} 16 & 3 & 2 \\ -4 & 3 & -8 \\ -2 & -6 & 11 \end{bmatrix}\right)$. From $\begin{bmatrix} 16 & 3 & 2 \\ -4 & 3 & -8 \\ -2 & -6 & 11 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 1/2 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix}$

we see that $E[0] = \left\{\overline{\left(-\frac{\boldsymbol{r}}{\mathbf{2}}, \mathbf{2}\boldsymbol{r}, \boldsymbol{r}\right)} \middle| r \in \Re\right\}$ with basis $\{\overline{(\mathbf{-1, 4, 2})}\}$.

$E[15] = \text{null}\left(\begin{bmatrix}16&3&2\\-4&3&-8\\-2&-6&11\end{bmatrix} - 15\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}\right) = \text{null}\left(\begin{bmatrix}1&3&2\\-4&-12&-8\\-2&-6&-4\end{bmatrix}\right)$. From $\begin{bmatrix}1&3&2\\-4&-12&-8\\-2&-6&-4\end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix}1&3&2\\0&0&0\\0&0&0\end{bmatrix}$ we see that $E[15] = \{\overline{(-\mathbf{3}r-\mathbf{2}s, r, s)} \mid r, s \in \Re\}$ with basis $\{\overline{(-\mathbf{3}, \mathbf{1}, \mathbf{0})}, \overline{(-\mathbf{2}, \mathbf{0}, \mathbf{1})}\}$.

<u>CYU 6.9</u> The kernel of the linear operator:

$$\left[T-(-3)I_{\Re^2}\right](a, b) = T(a, b)+3(a, b) = (3a+2b+3a, 3a-2b+3b) = (6a+2b, 3a+b)$$

is, by definition, the set: $\{(a, b) \mid (6a+2b, 3a+b) = (0, 0)\}$. Equating coefficients, we have:

$\left.\begin{aligned}6a+2b &= 0\\3a+b &= 0\end{aligned}\right\} \Rightarrow \begin{bmatrix}6&2\\5&-6\end{bmatrix} \xrightarrow{\textbf{rref}} \begin{bmatrix}1&1/3\\0&0\end{bmatrix}$. It follows that $E[-3] = \{(-r, 3r) \mid r \in \Re\}$ with basis $\{(-1, 3)\}$.

<u>CYU 6.10</u> With respect to the basis $\beta = \{x^2+x+1, x+1, 1\}$:

$$T(x^2+x+1) = 2x^2+5x+2$$
$$T(x+1) = x^2+3x+1$$
$$T(1) = x^2+x+1$$

$$\begin{array}{c}\\ \underbrace{\overbrace{\begin{bmatrix}1&0&0\\1&1&0\\1&1&1\end{bmatrix}}^{\beta}\begin{bmatrix}2&1&1\\5&3&1\\2&1&1\end{bmatrix}}_{T(\beta)}\end{array} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|ccc}1&0&0&2&1&1\\0&1&0&3&2&0\\0&0&1&-3&-2&0\end{array}\right] \quad [T]_{\beta\beta} = \boxed{\begin{matrix}2&1&1\\3&2&0\\-3&-2&0\end{matrix}}$$

From $\det([T]_{\beta\beta}-\lambda I_3) = \det\begin{bmatrix}2-\lambda&1&1\\3&2-\lambda&0\\-3&(-2)&-\lambda\end{bmatrix} = -\lambda(2-\lambda)^2 = 0$ we see that the eigenvalues are 0 and 2: same as those found in the solution of Example 6.10. Since the determination of the corresponding eigenspaces only depends on T and the eigenvalues, the spaces $E[0]$ and $E[2]$ are identical to those determined in the solution of Example 6.10.

<u>CYU 6.11</u> The three vectors $(\mathbf{1}, \mathbf{2}, \mathbf{0}), (\mathbf{1}, \mathbf{2}, \mathbf{2}), (\mathbf{2}, \mathbf{1}, \mathbf{1})$ are easily seen to constitute a basis β for $\Re^3$. Since $T(\mathbf{1}, \mathbf{2}, \mathbf{0}) = (\mathbf{1}, \mathbf{2}, \mathbf{0})$, $T(\mathbf{1}, \mathbf{2}, \mathbf{2}) = -(\mathbf{1}, \mathbf{2}, \mathbf{2})$, and $T(\mathbf{2}, \mathbf{1}, \mathbf{1}) = 2(\mathbf{2}, \mathbf{1}, \mathbf{1})$,

$\beta = \{(\mathbf{1}, \mathbf{2}, \mathbf{0}), (\mathbf{1}, \mathbf{2}, \mathbf{2}), (\mathbf{2}, \mathbf{1}, \mathbf{1})\}$ consists of eigenvector with corresponding eigenvalues $1, -1$, and 2, respectively.

$$\text{From: } \left.\begin{aligned}T(\mathbf{1}, \mathbf{2}, \mathbf{0}) &= 1(\mathbf{1}, \mathbf{2}, \mathbf{0})+0(\mathbf{1}, \mathbf{2}, \mathbf{2})+0(\mathbf{2}, \mathbf{1}, \mathbf{1})\\T(\mathbf{1}, \mathbf{2}, \mathbf{2}) &= 0(\mathbf{1}, \mathbf{2}, \mathbf{0})+(-1)(\mathbf{1}, \mathbf{2}, \mathbf{2})+0(\mathbf{2}, \mathbf{1}, \mathbf{1})\\T(\mathbf{2}, \mathbf{1}, \mathbf{1}) &= 0(\mathbf{1}, \mathbf{2}, \mathbf{0})+0(\mathbf{1}, \mathbf{2}, \mathbf{2})+2(\mathbf{2}, \mathbf{1}, \mathbf{1})\end{aligned}\right\} \text{ we have: } [T]_{\beta\beta} = \begin{bmatrix}1&0&0\\0&-1&0\\0&0&2\end{bmatrix}$$

<u>**CYU 6.12**</u> By Theorem 6.12, the n eigenvalues of T are linearly independent. Since V is of dimension n, those n eigenvectors constitute a basis for V. It follows, from Theorem 6.11, that T is diagonalizable.

<u>**CYU 6.13**</u> (a) Characteristic polynomial:

$$\det(A-\lambda I) = \det\left(\begin{bmatrix} -1-\lambda & 0 & 1 \\ -1 & 3-\lambda & 0 \\ -4 & 13 & -1-\lambda \end{bmatrix}\right) = -\lambda^3+\lambda^2+\lambda+2 = -(\lambda-2)(\lambda^2+\lambda+1)$$

From the above, we see that $\lambda = 2$ is the only (real) eigenvalue of A (note that the discriminant of $\lambda^2+\lambda+1$ is negative). Determining the dimension of $E[2]$

$$E[2] = \text{null}(A-2I) = \text{null}\left(\begin{bmatrix} -1-2 & 0 & 1 \\ -1 & 3-2 & 0 \\ -4 & 13 & -1-2 \end{bmatrix}\right)$$

Turning to the homogeneous system of equations: $\begin{bmatrix} -3 & 0 & 1 \\ -1 & 1 & 0 \\ -4 & 13 & -3 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -1/3 \\ 0 & 0 & 0 \end{bmatrix}$ we conclude that $E[2] = \{\overline{(\boldsymbol{r}, \boldsymbol{r}, \mathbf{3}\boldsymbol{r})} \mid r \in \Re\}$ with basis $\{\overline{(\mathbf{1, 1, 3})}\}$. It follows that there does not exist a basis for $\Re^3$ consisting of eigenvectors of A, and that therefore A is not diagonalizable.

(b) $\det\left(\begin{bmatrix} \mathbf{3-\lambda} & 2 & -1 \\ 2 & \mathbf{6-\lambda} & -2 \\ -1 & -2 & \mathbf{3-\lambda} \end{bmatrix}\right) = -\lambda^3+\lambda^2+\lambda+2 = -(\lambda-2)^2(\lambda-8)$.

$E[\mathbf{2}] = \text{null}\left(\begin{bmatrix} \mathbf{1} & 2 & -1 \\ 2 & \mathbf{4} & -2 \\ -1 & -2 & \mathbf{1} \end{bmatrix}\right)$ and $E[\mathbf{8}] = \text{null}\left(\begin{bmatrix} \mathbf{-5} & 2 & -1 \\ 2 & \mathbf{-2} & -2 \\ -1 & -2 & \mathbf{-5} \end{bmatrix}\right)$. From $\begin{bmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 1 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

and $\begin{bmatrix} -5 & 2 & -1 \\ 2 & -2 & -2 \\ -1 & -2 & -5 \end{bmatrix} \xrightarrow{\text{rref}} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$, we see that $E[2] = \{\overline{(-\mathbf{2}\boldsymbol{r}+\boldsymbol{s}, \boldsymbol{r}, \boldsymbol{s})} \mid r, s \in \Re\}$ and $E[8] = \{\overline{(-\boldsymbol{r}, -\mathbf{2}\boldsymbol{r}, \boldsymbol{r})} \mid r \in \Re\}$, with bases $\{\overline{(\mathbf{-2, 1, 0})}, \overline{(\mathbf{1, 0, 1})}\}$ and $\overline{(\mathbf{-1, -2, 1})}$, respectively. It follows that $\{\overline{(\mathbf{-2, 1, 0})}, \overline{(\mathbf{1, 0, 1})}, \overline{(\mathbf{-1, -2, 1})}\}$ is a basis for $\Re^3$ consisting of eigenvectors of A, and that therefore A is diagonalizable. Theorem 6.15 tells us that the matrix $P^{-1}AP$ will turn out to be a diagonal matrix with eigenvalues along its diagonal, where $P = \begin{bmatrix} -2 & 1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}$.

We leave it for you to verify that: $\begin{bmatrix} -2 & 1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} -2 & 1 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$.

<u>CYU 6.14</u> From Theorem 6.19 and Example 6.11 we have:

$$\begin{bmatrix} 0 & 2 & -2 & 0 \\ 1 & 1 & 0 & -1 \\ -1 & 1 & -2 & 1 \\ -1 & 1 & -2 & 1 \end{bmatrix}^{10} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & -2 \end{bmatrix}^{10} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2^{10} & 0 \\ 0 & 0 & 0 & (-2)^{10} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1024 & 0 \\ 0 & 0 & 0 & 1024 \end{bmatrix}$$

<u>CYU 6.15</u> In the solution of CYU 6.13(b), we found the characteristic polynomial of $\begin{bmatrix} 3 & 2 & -1 \\ 2 & 6 & -2 \\ -1 & -2 & 3 \end{bmatrix}$ to be $-(\lambda - 2)^2(\lambda - 8)$ with $E[2]$ and $E[8]$ of dimensions 2 and 1, respectively.

<u>CYU 6.16</u> Let s_k denote the k^{th} element of the sequence (for $k \geq 3$), $F = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$, and $\boldsymbol{S_k} = \begin{bmatrix} s_k \\ s_{k-1} \end{bmatrix}$. From: $S_3 = FS_2 = F\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, $S_4 = FS_3 = F \cdot F\begin{bmatrix} 3 \\ 2 \end{bmatrix} = F^2\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, …, $S_k = F^{k-2}\begin{bmatrix} 3 \\ 2 \end{bmatrix}$, we see that s_k is the sum of the entries in the **<u>first row</u>** of $F^{k-2}\begin{bmatrix} 3 \\ 2 \end{bmatrix}$. We now set our sights on finding the matrix F^k, and begin by finding a diagonalization for F:

$$\det \begin{bmatrix} 1-\lambda & 2 \\ 1 & -\lambda \end{bmatrix} = (\lambda - 2)(\lambda + 1);\ E[2] = \{(2r, r) | r \in \Re\};\ E[-1] = \{(-r, r) | r \in \Re\}$$

It follows that $\{(2, 1), (-1, 1)\}$ is a basis of eigenvectors for F. Employing Theorem 6.18 we have:

$$F = PDP^{-1} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \underset{\substack{\uparrow \\ \text{steps omitted}}}{=} \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}$$

From Theorem 6.19:

$$F^k = PD^kP^{-1} = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 2^k & 0 \\ 0 & (-1)^k \end{bmatrix}\begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix}^{-1} \underset{\substack{\uparrow \\ \text{steps omitted}}}{=} \begin{bmatrix} \frac{1}{3} \cdot 2^{k+1} - \frac{1}{3}(-1)^{k+1} & \frac{1}{3} \cdot 2^{k+1} + \frac{2}{3}(-1)^{k+1} \\ ************** & *************** \end{bmatrix}$$

Thus: $F^{k-2}\begin{bmatrix}3\\2\end{bmatrix} = \begin{bmatrix}\frac{1}{3}\cdot 2^{k-1}-\frac{1}{3}(-1)^{k-1} & \frac{1}{3}\cdot 2^{k-1}+\frac{2}{3}(-1)^{k-1}\\ \text{*************} & \text{**************}\end{bmatrix}\begin{bmatrix}3\\2\end{bmatrix}$

$$= \begin{bmatrix}2^{k-1}-(-1)^{k-1}+\frac{2}{3}\cdot 2^{k-1}+\frac{4}{3}(-1)^{k-1}\\ \text{*****************************}\end{bmatrix} = \begin{bmatrix}\frac{5}{3}\cdot 2^{k-1}+\frac{1}{3}(-1)^{k-1}\\ \text{**************}\end{bmatrix}$$

Conclusion: $s_k = \frac{5}{3}\cdot 2^{k-1}+\frac{1}{3}(-1)^{k-1}$.

CYU 6.17 $\{(\mathbf{-1, 1, 1, 0}), (\mathbf{1, 0, 0, 1}), (\mathbf{1, 1, 0, 0}), (\mathbf{1, 0, 1, 1})\}$ is a basis of eigenvalues for $\begin{bmatrix}0 & 2 & -2 & 0\\ 1 & 1 & 0 & -1\\ -1 & 1 & -2 & 1\\ -1 & 1 & -2 & 1\end{bmatrix}$, with 0 the eigenvalue associated with $(\mathbf{-1, 1, 1, 0})$ and $(\mathbf{1, 0, 0, 1})$, 2 the eigenvalue associated with $(\mathbf{1, 1, 0, 0})$, and -2 the eigenvalue associated with $(\mathbf{1, 0, 1, 1})$ (Example 6.11, page 237). Applying Theorem6.25 we conclude that:

$$c_1e^{0x}\begin{bmatrix}-1\\1\\1\\0\end{bmatrix}+c_2e^{0x}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}+c_3e^{2x}\begin{bmatrix}1\\1\\0\\0\end{bmatrix}+c_4e^{-2x}\begin{bmatrix}1\\0\\1\\1\end{bmatrix} = \begin{bmatrix}-c_1+c_2+c_3e^{2x}+c_4e^{-2x}\\ c_1+c_3e^{2x}\\ c_1+c_4e^{-2x}\\ c_2+c_4e^{-2x}\end{bmatrix}$$

is the general solution of the given system of equations.

CYU 6.18 From CYU 6.16: $\begin{aligned}y_1 &= -c_1+c_2+c_3e^{2x}+c_4e^{-2x}\\ y_2 &= c_1+c_3e^{2x}\\ y_3 &= c_1+c_4e^{-2x}\\ y_4 &= c_2+c_4e^{-2x}\end{aligned}$ Since $\begin{aligned}y_1(0) &= 0\\ y_2(0) &= 1\\ y_3(0) &= 2\\ y_4(0) &= 3\end{aligned}$:

$$S:\left.\begin{aligned}-c_1+c_2+c_3+c_4 &= 0\\ c_1+c_3 &= 1\\ c_1+c_4 &= 2\\ c_2+c_4 &= 3\end{aligned}\right\}\xrightarrow{\text{aug(S)}}\begin{bmatrix}-1 & 1 & 1 & 1 & 0\\ 1 & 0 & 1 & 0 & 1\\ 1 & 0 & 0 & 1 & 2\\ 0 & 1 & 0 & 1 & 3\end{bmatrix}\xrightarrow{\text{rref}}\left[\begin{array}{cccc|c}1 & 0 & 0 & 0 & 2\\ 0 & 1 & 0 & 0 & 3\\ 0 & 0 & 1 & 0 & -1\\ 0 & 0 & 0 & 1 & 0\end{array}\right]$$

Solution: $y_1 = 1-e^{2x}, y_2 = 2-e^{2x}, y_3 = 2, y_4 = 3$.

CYU 6.19 $\begin{bmatrix} T'(t) \\ F'(t) \end{bmatrix} = \begin{bmatrix} \frac{5}{2} & -\frac{1}{4} \\ -1 & \frac{5}{2} \end{bmatrix} \begin{bmatrix} T(t) \\ F(t) \end{bmatrix}$. det $\begin{bmatrix} \frac{5}{2} - \lambda & -\frac{1}{4} \\ -1 & \frac{5}{2} - \gamma \end{bmatrix} = 0 \Rightarrow \left(\frac{5}{2} - \lambda\right)^2 - \frac{1}{4} = 0 \Rightarrow \frac{5}{2} - \lambda = \pm\frac{1}{2} \Rightarrow \lambda = \begin{cases} 3 \\ 2 \end{cases}$

$$E[3] = \text{null}\begin{bmatrix} \frac{5}{2} - 3 & -\frac{1}{4} \\ -1 & \frac{5}{2} - 3 \end{bmatrix} = \{(-\boldsymbol{r}, \boldsymbol{2r}) | r \in \Re\} \text{ and } E[2] = \text{null}\begin{bmatrix} \frac{5}{2} - 2 & -\frac{1}{4} \\ -1 & \frac{5}{2} - 2 \end{bmatrix} = \{(\boldsymbol{r}, \boldsymbol{2r}) | r \in \Re\}$$

Choosing $(\mathbf{-1, 2})$ and $(\mathbf{1, 2})$ as eigenvectors for the eigenvalues 3 and 2, respectively, we have:

$\begin{bmatrix} T(t) \\ F(t) \end{bmatrix} = c_1 e^{3t} \begin{bmatrix} -1 \\ 2 \end{bmatrix} + c_2 e^{2t} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} -c_1 e^{3t} + c_2 e^{2t} \\ 2c_1 e^{3t} + 2c_2 e^{2t} \end{bmatrix}$. Turning to the initial conditions $\begin{bmatrix} T(t) \\ F(t) \end{bmatrix} = \begin{bmatrix} 120 \\ 200 \end{bmatrix}$:

$$\begin{bmatrix} -c_1 e^{3 \cdot 0} + c_2 e^{2 \cdot 0} \\ 2c_1 e^{3 \cdot 0} + 2c_2 e^{2 \cdot 0} \end{bmatrix} = \begin{bmatrix} 120 \\ 200 \end{bmatrix} \Rightarrow \begin{cases} -c_1 + c_2 = 120 \\ 2c_1 + 2c_2 = 200 \end{cases} \Rightarrow c_1 = -10 \text{ and } c_2 = 110$$

Bringing us to: $\begin{bmatrix} T(t) \\ F(t) \end{bmatrix} = \begin{bmatrix} 10e^{3t} + 110e^{2t} \\ -20e^{3t} + 220_2 e^{2t} \end{bmatrix}$. Setting $T(t) = F(t)$ we have:

$$\begin{aligned} 10e^{3t} + 110e^{2t} &= -20e^{3t} + 220_2 e^{2t} \\ 30e^{3t} &= 110e^{2t} \\ \ln(30e^{3t}) &= \ln(110e^{2t}) \\ \ln 30 + \ln e^{3t} &= \ln(110) + \ln e^{2t} \\ \ln 30 + 3t &= \ln 110 + 2t \Rightarrow t = \ln 110 - \ln 30 \approx 1.3\text{years} \end{aligned}$$

CYU 6.20 Let D denote the state that a student is living in the dorm, and C denote the state that the student is a commuter. Then: $T = \begin{array}{cc} \begin{array}{c} \text{current state} \\ \begin{array}{cc} \text{D} & \text{C} \end{array} \end{array} & \\ \begin{bmatrix} .8 & .1 \\ .2 & .9 \end{bmatrix} & \begin{array}{c} \text{D} \\ \text{C} \end{array} \end{array} \text{next state}$ with $S_0 = \begin{bmatrix} 858 \\ 702 \end{bmatrix}$. Then:

$$S_1 = \begin{bmatrix} .8 & .1 \\ .2 & .9 \end{bmatrix} \begin{bmatrix} 858 \\ 702 \end{bmatrix} = \begin{bmatrix} \frac{3783}{5} \\ \frac{4017}{5} \end{bmatrix} \approx \begin{bmatrix} \mathbf{757} \\ 803 \end{bmatrix} \begin{array}{c} \mathbf{D} \\ \text{C} \end{array}$$

$$S_2 = \begin{bmatrix} .8 & .1 \\ .2 & .9 \end{bmatrix} \begin{bmatrix} 757 \\ 803 \end{bmatrix} \approx \begin{bmatrix} \mathbf{686} \\ 874 \end{bmatrix} \begin{array}{c} \mathbf{D} \\ \text{C} \end{array} \text{ and } S_2 = \begin{bmatrix} .8 & .1 \\ .2 & .9 \end{bmatrix} \begin{bmatrix} 686 \\ 874 \end{bmatrix} \approx \begin{bmatrix} \mathbf{636} \\ 924 \end{bmatrix} \begin{array}{c} \mathbf{D} \\ \text{C} \end{array}$$

Conclusion: 757, 686, and 636 of the current freshmen will live in the dorm in their sophomore, junior, and senior year, respectively.

CYU 6.21

$$\begin{bmatrix}.73 & .32 & .09\\ .21 & .61 & .04\\ .06 & .07 & .87\end{bmatrix}\begin{bmatrix}x\\ y\\ z\end{bmatrix} = \begin{bmatrix}x\\ y\\ z\end{bmatrix} \Rightarrow \left.\begin{aligned}.73x + .32y + .09z &= x\\ .21x + .61y + .04z &= y\\ .06x + .07y + .87z &= z\end{aligned}\right\}$$

$$\left.\begin{aligned}-.27x + .32y + .09z &= 0\\ .21x - .39 + .04z &= 0\\ .06x + .07y - .13 &= 0\\ x + y + z &= 1\end{aligned}\right\} \xrightarrow{\text{rref}} \left[\begin{array}{ccc|c}1 & 0 & 0 & \frac{479}{1157}\\ 0 & 1 & 0 & \frac{297}{1157}\\ 0 & 0 & 1 & \frac{381}{1157}\\ 0 & 0 & 0 & 0\end{array}\right]$$

see solution of Example 6.15.

Conclusion: Eventually $\frac{479}{1157}, \frac{297}{1157}, \frac{381}{1157}$, or approximately 41%, 26%, 33% of the population, will vote democratic, republican, green, respectively.

CHAPTER 7
INNER PRODUCT SPACES

CYU 7.1 $r\boldsymbol{u}\cdot\boldsymbol{v} = r(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n)\cdot(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n) = \sum_{i=1}^{n}(ru_i)v_i = \sum_{i=1}^{n}u_i(rv_i) = \boldsymbol{u}\cdot r\boldsymbol{v}$

CYU 7.2

(a) $$\|c\boldsymbol{v}\| = \sqrt{c(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)\cdot c(\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n)} = \sqrt{(c\boldsymbol{v}_1, c\boldsymbol{v}_2, \ldots, c\boldsymbol{v}_n)\cdot(c\boldsymbol{v}_1, c\boldsymbol{v}_2, \ldots, c\boldsymbol{v}_n)}$$
$$= \sqrt{\sum_{i=1}^{n}c^2v_i^2} = \sqrt{c^2}\sqrt{\sum_{i=1}^{n}v_i^2} = |c|\|\boldsymbol{v}\|$$

(b) $$\|\boldsymbol{u}-\boldsymbol{v}\|^2 = (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) = \boldsymbol{u}\cdot(\boldsymbol{u}-\boldsymbol{v}) - \boldsymbol{v}\cdot(\boldsymbol{u}-\boldsymbol{v})$$
$$= \boldsymbol{u}\cdot\boldsymbol{u} - \boldsymbol{u}\cdot\boldsymbol{v} - \boldsymbol{v}\cdot\boldsymbol{u} + \boldsymbol{v}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|^2 - 2\boldsymbol{u}\cdot\boldsymbol{v} + \|\boldsymbol{v}\|^2$$

CYU 7.3 $\theta = \cos^{-1}\left(\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\right) = \cos^{-1}\left(\frac{(\mathbf{1, 2, 0})\cdot(\mathbf{-1, 3, 1})}{\sqrt{1+4}\sqrt{1+9+1}}\right) = \cos^{-1}\left(\frac{5}{\sqrt{55}}\right) \approx 83°$

CYU 7.4 For $\mathbf{u}, \bar{\mathbf{u}} \in \mathbf{v}^{\perp}$ and $r \in \Re$: $(r\mathbf{u} + \bar{\mathbf{u}}) \cdot \mathbf{v} = r(\mathbf{u} \cdot \mathbf{v}) + \bar{\mathbf{u}} \cdot \mathbf{v} = r(0) + 0 = 0$. We see that $r\mathbf{u} + \bar{\mathbf{u}} \in \mathbf{v}^{\perp}$. It follows, from Theorem 2.13 (page 61), that $\mathbf{v}^{\perp}$ is a subspace of $\Re^n$.

CYU 7.5 $\text{proj}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u} = \left[\frac{(\mathbf{0, 2, 4, 1}) \cdot (\mathbf{3, 0, 1, -1})}{(\mathbf{0, 2, 4, 1}) \cdot (\mathbf{0, 2, 4, 1})}\right](\mathbf{0, 2, 4, 1}) = \frac{3}{21}(\mathbf{0, 2, 4, 1}) = \left(\mathbf{0}, \frac{\mathbf{6}}{\mathbf{21}}, \frac{\mathbf{12}}{\mathbf{21}}, \frac{\mathbf{3}}{\mathbf{21}}\right)$

and: $\mathbf{v} - \text{proj}_{\mathbf{u}}\mathbf{v} = (\mathbf{3, 0, 1, -1}) - \left(\mathbf{0}, \frac{\mathbf{6}}{\mathbf{21}}, \frac{\mathbf{12}}{\mathbf{21}}, \frac{\mathbf{3}}{\mathbf{21}}\right) = \left(\mathbf{3}, -\frac{\mathbf{6}}{\mathbf{21}}, \frac{\mathbf{9}}{\mathbf{21}} - \frac{\mathbf{24}}{\mathbf{21}}\right)$.

$$(\mathbf{3, 0, 1, -1}) = \left(\mathbf{3}, -\frac{\mathbf{6}}{\mathbf{21}}, \frac{\mathbf{9}}{\mathbf{21}} - \frac{\mathbf{24}}{\mathbf{21}}\right) + \left(\mathbf{0}, \frac{\mathbf{6}}{\mathbf{21}}, \frac{\mathbf{12}}{\mathbf{21}}, \frac{\mathbf{3}}{\mathbf{21}}\right)$$

CYU 7.6 (a) A direction vector for the line L passing through $(1, -2)$ and $(2, 4)$: $\mathbf{u} = (2, 4) - (1, -2) = (\mathbf{1, 6})$. The vector from the point $(1, -2)$ on L to $P = (2, 5)$: $\mathbf{v} = (2, 5) - (1, -2) = (\mathbf{1, 7})$. Applying Theorem 7.2, we have:

$$\text{proj}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u} = \left(\frac{(\mathbf{1, 6}) \cdot (\mathbf{1, 7})}{(\mathbf{1, 6}) \cdot (\mathbf{1, 6})}\right)(\mathbf{1, 6}) = \frac{43}{37}(\mathbf{1, 6})$$

Hence: $\|\mathbf{v} - \text{proj}_{u}\mathbf{v}\| = \left\|(\mathbf{1, 7}) - \left(\frac{\mathbf{43}}{\mathbf{37}}, \frac{\mathbf{258}}{\mathbf{37}}\right)\right\| = \left\|\left(-\frac{\mathbf{6}}{\mathbf{37}}, \frac{\mathbf{1}}{\mathbf{37}}\right)\right\| = \frac{\sqrt{6^2 + 1^2}}{37} = \frac{1}{\sqrt{37}}$.

(b) $\mathbf{u} = (1, 2, 2, 1) - (1, 2, 0, 1) = (\mathbf{0, 0, 2, 0})$. $\mathbf{v} = (1, 0, 1, 3) - (1, 2, 0, 1) = (\mathbf{0, -2, 1, 2})$.

$$\text{proj}_{\mathbf{u}}\mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{u} \cdot \mathbf{u}}\right)\mathbf{u} = \left(\frac{(\mathbf{0, 0, 2, 0}) \cdot (\mathbf{0, -2, 1, 2})}{(\mathbf{0, 0, 2, 0}) \cdot (\mathbf{0, 0, 2, 0})}\right)(\mathbf{0, 0, 2, 0}) = \frac{1}{2}(\mathbf{0, 0, 2, 0}) = (\mathbf{0, 0, 1, 0})$$

Hence: $\|\mathbf{v} - \text{proj}_{u}\mathbf{v}\| = \|(0, -2, 1, 2) - (0, 0, 1, 0)\| = \|(0, -2, 0, 2)\| = \sqrt{16} = 4$.

CYU 7.7 A normal to the desired plane will have the same direction is that of the line passing through the two given points; namely: $\mathbf{n} = (0, 2, 1) - (1, 1, 0) = (\mathbf{-1, 1, 1})$. Normal form for the plane: $(-1, 1, 1) \cdot (x - 1, y - 3, z + 2) = 0$.

CYU 7.8 Let $A = \{(x, y, z) | 3x + y - 2z = 6\}$, and

$B = \{(2, 0, 0) + r(0, 2, 1) + s(2, 0, 3) | r, s \in R\} = \{(2 + 2s, 2r, r + 3s) | r, s \in R\}$.

$B \subseteq A$: $(x, y, z) = (2 + 2s, 2r, r + 3s)$.

$$\Rightarrow 3x + y - 2z = 3(2 + 2s) + 2r - 2(r + 3s) = 6 + 6s + 2r - 2r - 6s = 6$$

$A \subseteq B$: If (x, y, z) is such that $3x + y - 2z = 6$ (*), can we find $r, s \in R$ such that $x = 2 + 2s$, $y = 2r$, and $z = r + 3s$? Yes:

In order for $x = 2+2s$, $s = \frac{x-2}{2}$. In order for $y = 2r$, $r = \frac{y}{2}$. We show that for those particular values of r and s, $z = r+3s$:

$$z \underset{\substack{\uparrow \\ (*)}}{=} \frac{3x+y-6}{2} = \frac{3(2+2s)+2r-6}{2} = \frac{6+6s+2r-6}{2} = 2s+2$$

CYU 7.9 Let $A = (x, y, z)$ be on the plane $ax+by+cz = d$ with normal $\boldsymbol{n} = (a, b, c)$. Determine the vector $\boldsymbol{v}$ from A to (x_0, y_0, z_0): $\boldsymbol{v} = (x, y, z)-(x_0, y_0, z_0) = (x-x_0, y-y_0, z-z_0)$. Applying Theorem 7.2, we have:

$$\|\text{proj}_{\boldsymbol{n}}\boldsymbol{v}\| = \left\|\left(\frac{\boldsymbol{v}\cdot\boldsymbol{n}}{\boldsymbol{n}\cdot\boldsymbol{n}}\right)\boldsymbol{n}\right\| = \left\|\left(\frac{ax-ax_0+by-by_0+cz-cz_0}{a^2+b^2+c^2}\right)(a, b, c)\right\|$$

$$= \left|\frac{\mathbf{ax}-ax_0+\mathbf{by}-by_0+\mathbf{cz}-cz_0}{a^2+b^2+c^2}\right|\|(a, b, c)\|$$

$$\text{Since } ax+by+cz = d\text{:} \quad = \frac{|ax_0+by_0+cz_0-d|}{a^2+b^2+c^2}\sqrt{a^2+b^2+c^2} = \frac{|ax_0+by_0+cz_0-d|}{\sqrt{a^2+b^2+c^2}}$$

CYU 7.10

(a) $\overrightarrow{AB} = (2, 5, -3)-(3, -2, 2) = (\mathbf{-1, 7, -5})$ and $\overrightarrow{AC} = (4, 1, -3)-(3, -2, 2) = (\mathbf{1, 3, -5})$.

Here is a normal to the plane: $\boldsymbol{n} = \det\begin{bmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ -1 & 7 & -5 \\ 1 & 3 & -5 \end{bmatrix} = -20\boldsymbol{i}-10\boldsymbol{j}-10\boldsymbol{k} = (\mathbf{-20, -10, -10})$.

Here is a “nicer” normal: $\boldsymbol{n} = -\frac{1}{10}(\mathbf{-20, -10, -10}) = (\mathbf{2, 1, 1})$.

Choosing the point $A = (3, -2, 2)$ on the plane, we arrive at the general form equation of the plane: $(\mathbf{2, 1, 1})\cdot(\boldsymbol{x}-\mathbf{3}, \boldsymbol{y}+\mathbf{2}, \boldsymbol{z}-\mathbf{2}) = 0$ or: $2x+y+z-6 = 0$.

(b) In Example 2.15, page 72, we found the vector form representation for the above plane:

$$P = \{(3-r+s, -2+7r+3s, 2-5r-5s) | r, s \in \Re\}.$$

We are to show that:

$$P = \{(3-r+s, -2+7r+3s, 2-5r-5s) | r, s \in \Re\} = \{(x, y, z) | 2x+y+z = 6\} = Q.$$

If $(x, y, z) = (3-r+s, -2+7r+3s, 2-5r-5s) \in P$, then:

$$\begin{aligned} 2z+y+z &= 2(3-r+s)+(-2+7r+3s)+(2-5r-5s) \\ &= 6-2r+2s-2+7r+3s+2-5r-5s = 6 \end{aligned} \qquad \text{Thus: } P \subseteq Q.$$

If (x, y, z) is such that $\mathbf{2x+y+z = 6}$ can we find real numbers r and s such that:

$$3-r+s = x, -2+7r+3s = y, \text{ and } 2-5r-5s = z \underset{\substack{\uparrow \\ \text{since } \mathbf{2x+y+z = 6}}}{=} 6-2x-y$$

which is to say: $-r+s = x-3$, $7r+3s = y+2$, and $-5r-5s = -2x-y+4$

Turning to the system of equation:

$$\left.\begin{aligned} -r+s &= x-3\\ 7r+3s &= y+2\\ -5r-5s &= -2x-y+4 \end{aligned}\right\}\xrightarrow[\text{matrix}]{\text{augmented}}\left[\begin{array}{cc|c} r & s & \\ -1 & 1 & x-3\\ 7 & 3 & y+2\\ -5 & -5 & -2x-y+4 \end{array}\right]\xrightarrow[\substack{\uparrow\\ \text{steps ommited}}]{\text{rref}}\left[\begin{array}{cc|c} \mathbf{1} & 0 & \dfrac{-3x+y+19}{10}\\ 0 & \mathbf{1} & \dfrac{7x+y-19}{10}\\ 0 & 0 & 0 \end{array}\right]$$

We see that $r = -x+3$ and $s = \dfrac{7x+y-19}{10}$ does the trick. Thus: $Q \subseteq P$.

<u>CYU 7.11</u> (i) For $\boldsymbol{v} = ax^2+bx+c$, $\langle \boldsymbol{v}, \boldsymbol{v}\rangle = \langle ax^2+bx+c, ax^2+bx+c\rangle = a^2+b^2+c^2 \geq 0$ and $\langle \boldsymbol{v}, \boldsymbol{v}\rangle = 0$ only if $a = b = c = 0$.

(ii) For $\boldsymbol{u} = a_2x^2+a_1x+a_0$, $\boldsymbol{v} = b_2x^2+b_1x+b_0$:

$$\begin{aligned}\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \langle a_2x^2+a_1x+a_0, b_2x^2+b_1x+b_0\rangle &= a_2b_2+a_1b_1+a_0b_0\\ &= ba_2+b_1a_1+b_0a_0\\ &= \langle b_2x^2+b_1x+b_0, a_2x^2+a_1x+a_0\rangle = \langle \boldsymbol{v}, \boldsymbol{u}\rangle\end{aligned}$$

(iii) For $\boldsymbol{u} = a_2x^2+a_1x+a_0$, $\boldsymbol{v} = b_2x^2+b_1x+b_0$, and $r \in \Re$:

$$\begin{aligned}\langle r\boldsymbol{u}, \boldsymbol{v}\rangle &= \langle r(a_2x^2+a_1x+a_0), b_2x^2+b_1x+b_0\rangle\\ &= \langle ra_2x^2+ra_1x+ra_0, b_2x^2+b_1x+b_0\rangle\\ &= ra_2b_2+ra_1b_1+ra_0b_0 = r(a_2b_2+a_1b_1+a_0b_0) = r\langle \boldsymbol{u}, \boldsymbol{v}\rangle\end{aligned}$$

(iv) For $\boldsymbol{u} = a_2x^2+a_1x+a_0$, $\boldsymbol{v} = b_2x^2+b_1x+b_0$ and $\boldsymbol{z} = c_2x^2+c_1x+c_0$:

$$\begin{aligned}\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{z}\rangle &= \langle (a_2x^2+a_1x+a_0)+(b_2x^2+b_1x+b_0), c_2x^2+c_1x+c_0\rangle\\ &= \langle (a_2+b_2)x_2+(a_1+b_1)x+(a_0+b_0), c_2x^2+c_1x+c_0\rangle\\ &= (a_2+b_2)c_2+(a_1+b_1)c_1+(a_0+b_0)c_0\\ &= (a_2c_2+a_1c_1+a_0c_0)+(b_2c_2+b_1c_1+b_0c_0) = \langle \boldsymbol{u}, \boldsymbol{z}\rangle+\langle \boldsymbol{v}, \boldsymbol{z}\rangle\end{aligned}$$

<u>CYU 7.12</u> $$\langle r\boldsymbol{u}, r\boldsymbol{v}\rangle \underset{\substack{\uparrow\\ \text{Definition 7.5(iii)}}}{=} r\langle \boldsymbol{u}, r\boldsymbol{v}\rangle \underset{\substack{\uparrow\\ \text{Theorem 7.4 (c)}}}{=} r[r\langle \boldsymbol{u}, \boldsymbol{v}\rangle] = r^2\langle \boldsymbol{u}, \boldsymbol{v}\rangle$$

<u>CYU 7.13</u> (a) $$\|r\boldsymbol{v}\| = \sqrt{\langle r\boldsymbol{v}, r\boldsymbol{v}\rangle} \underset{\substack{\uparrow\\ \text{CYU 7.12}}}{=} \sqrt{r^2\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = \sqrt{r^2}\sqrt{\langle \boldsymbol{v}, \boldsymbol{v}\rangle} = |r|\langle \boldsymbol{v}, \boldsymbol{v}\rangle$$

Definition 7.5(iv)

(b) $$\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}+\boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle = \langle \boldsymbol{u}, \boldsymbol{u}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \langle \boldsymbol{v}, \boldsymbol{v}\rangle$$

Definition 7.6 and 7.5(i): $$= \|\boldsymbol{u}\|^2 + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \|\boldsymbol{v}\|^2$$

$$= \|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \|\boldsymbol{v}\|^2$$

CYU 7.14 (a) $\|(\mathbf{3}, \mathbf{5}, -\mathbf{8})\| = \sqrt{(\mathbf{3}, \mathbf{5}, -\mathbf{8}) \cdot (\mathbf{3}, \mathbf{5}, -\mathbf{8})} = \sqrt{5(3^2) + 5(5^2) + 5[(-8)^2]} = 7\sqrt{10}$.

(b) $$\|(\mathbf{3}, \mathbf{5}, -\mathbf{8}) - (\mathbf{1}, \mathbf{0}, \mathbf{2})\| = \|(\mathbf{2}, \mathbf{5}, -\mathbf{10}) \cdot (\mathbf{2}, \mathbf{5}, -\mathbf{10})\|$$
$$= \sqrt{(2, 5, -10) \cdot (2, 5, -10)} = \sqrt{5(2^2) + 5(5^2) + 5[(-10)^2]} = \sqrt{645}$$

CYU 7.15

$$\left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \geq 0 \Rightarrow \left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle - \left\langle \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} - \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \geq 0$$

$$\Rightarrow \left\langle \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}, \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle - \left\langle \frac{\mathbf{1}}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}\right\rangle - \left\langle \frac{\mathbf{1}}{\|\boldsymbol{u}\|}\boldsymbol{u}, \frac{\mathbf{1}}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle + \left\langle \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}, \frac{1}{\|\boldsymbol{v}\|}\boldsymbol{v}\right\rangle \geq 0$$

$$\Rightarrow \frac{\mathbf{1}}{\|\boldsymbol{u}\|^{\mathbf{2}}}\langle \boldsymbol{u}, \boldsymbol{u}\rangle - \frac{\mathbf{2}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \frac{\mathbf{1}}{\|\boldsymbol{v}\|^{\mathbf{2}}}\langle \boldsymbol{v}, \boldsymbol{v}\rangle \geq \mathbf{0}$$

$$\Rightarrow \mathbf{1} - \frac{\mathbf{2}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \mathbf{1} \geq \mathbf{0} \Rightarrow \mathbf{2} \geq \frac{\mathbf{2}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}\langle \boldsymbol{u}, \boldsymbol{v}\rangle \Rightarrow \langle \boldsymbol{u}, \boldsymbol{v}\rangle \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

CYU 7.16 $$\theta = \cos^{-1}\left(\frac{\langle (\mathbf{3}, \mathbf{5}, -\mathbf{8}), (\mathbf{1}, \mathbf{0}, \mathbf{2})\rangle}{\|(3, 5, -8)\|\|(1, 0, 2)\|}\right)$$
$$= \cos^{-1}\left(\frac{5(3)(1) + 2(5)(0) + 4(-8)(2)}{\sqrt{5(3)(3) + 2(5)(5) + 4(-8)(-8)}\sqrt{5(1)(1) + 2(0)(0) + 4(2)(3)}}\right)$$
$$= \cos^{-1}\left(\frac{-49}{\sqrt{351}\ \sqrt{29}}\right) \approx 119.1^\circ$$

CYU 7.17 For any $\boldsymbol{v} = \sum_{i=1}^{m} c_i\boldsymbol{v_i} \in \operatorname{Span}\{\boldsymbol{v_1}, \boldsymbol{v_2}, \ldots, \boldsymbol{v}_m\}$:

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \left\langle \boldsymbol{u}, \sum_{i=1}^{m} c_i\boldsymbol{v_i}\right\rangle = \sum_{i=1}^{m} c_i\langle \boldsymbol{u}, \boldsymbol{v_i}\rangle = 0, \text{ since } \langle \boldsymbol{u}, \boldsymbol{v_i}\rangle = 0 \text{ for } 1 \leq i \leq m.$$

CYU 7.18

Exercise 46, page 300

$$\begin{aligned}\langle \boldsymbol{v}, \boldsymbol{w}\rangle &= \langle a_1\boldsymbol{v_1} + \ldots + a_n\boldsymbol{v_n}, b_1\boldsymbol{v_1} + \ldots + b_n\boldsymbol{v_n}\rangle \\ &= \langle a_1\boldsymbol{v_1} + \ldots + a_n\boldsymbol{v_n}, b_1\boldsymbol{v_1}\rangle + \ldots + \langle a_1\boldsymbol{v_1} + a_2\boldsymbol{v_2} + \ldots + a_n\boldsymbol{v_n}, b_n\boldsymbol{v_n}\rangle \\ &= \sum_{i=1}^{n}\langle a_i\boldsymbol{v_i}, b_1\boldsymbol{v_1}\rangle + \ldots + \sum_{i=1}^{n}\langle a_i\boldsymbol{v_i}, b_n\boldsymbol{v_n}\rangle = \sum_{i=1}^{n} a_ib_1\langle \boldsymbol{v_i}, \boldsymbol{v_1}\rangle + \ldots + \sum_{i=1}^{n} a_ib_n\langle \boldsymbol{v_i}, \boldsymbol{v_n}\rangle \\ &\text{since } \langle \boldsymbol{v_i}, \boldsymbol{v_j}\rangle = \begin{cases} 1 \text{ if } i = j \\ 0 \text{ if } i \neq j\end{cases}: = a_1b_1 + \ldots + a_nb_n = \sum_{i=1}^{n} a_ib_i\end{aligned}$$

CYU 7.19 The first order of business is to determine a basis for the space S spanned by the vectors $(\mathbf{2, 1, 1, 0})$, $(\mathbf{1, 0, 1, 0})$, $(\mathbf{3, 1, 2, 0})$, $(\mathbf{0, 1, 0, 1})$. Applying Theorem 3.13, page 103 we see that first, second, and fourth of the above four vectors constitute a basis: $\boldsymbol{v_1} = (\mathbf{2, 1, 1, 0})$, $\boldsymbol{v}_2 = (\mathbf{1, 0, 1, 0})$, $\boldsymbol{v}_3 = (\mathbf{0, 1, 0, 1})$ of S: $\begin{bmatrix} 2 & 1 & 3 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1\end{bmatrix}$ $\begin{bmatrix} \boxed{1} & 0 & 1 & 0 \\ 0 & \boxed{1} & 1 & 0 \\ 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0\end{bmatrix}$. We now apply the Grahm-Schmidt Process to that basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ to generate an orthonormal basis $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \boldsymbol{u_3}\}$ for S:

$$\boldsymbol{u}_1 = \boldsymbol{v}_1 = (\mathbf{2, 1, 1, 0})$$

$$\begin{aligned}\boldsymbol{u}_2 = \boldsymbol{v_2} - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v_2}\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u_1} &= (\mathbf{1, 0, 1, 0}) - \frac{(\mathbf{2, 1, 1, 0}) \cdot (\mathbf{1, 0, 1, 0})}{(\mathbf{2, 1, 1, 0}) \cdot (\mathbf{2, 1, 1, 0})}(\mathbf{2, 1, 1, 0}) \\ &= (\mathbf{1, 0, 1}) - \frac{3}{6}(\mathbf{2, 1, 1, 0}) = \left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right)\end{aligned}$$

$$\begin{aligned}u_3 = \boldsymbol{v_3} - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u_1}\rangle}\boldsymbol{u}_1 - \frac{\langle \boldsymbol{u}_2, \boldsymbol{v}_3\rangle}{\langle \boldsymbol{u}_2, \boldsymbol{u_2}\rangle}\boldsymbol{u}_2 &= (\mathbf{0, 1, 0, 1}) - \frac{(\mathbf{2, 1, 1, 0}) \cdot (\mathbf{0, 1, 0, 1})}{(\mathbf{2, 1, 1, 0}) \cdot (\mathbf{2, 1, 1, 0})}(\mathbf{2, 1, 1, 0}) - \frac{\left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right) \cdot (\mathbf{0, 1, 0, 1})}{\left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right) \cdot \left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right)}\left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right) \\ &= (\mathbf{0, 1, 0, 1}) - \frac{1}{6}(\mathbf{2, 1, 1, 0}) - \frac{-1/2}{1/2}\left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right) = \left(-\frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}}, \mathbf{1}\right)\end{aligned}$$

Conclusion: $\left\{\frac{\mathbf{1}}{\mathbf{6}}(\mathbf{2, 1, 1, 0}), \frac{\mathbf{1}}{\mathbf{2}}\left(\mathbf{0}, -\frac{\mathbf{1}}{\mathbf{2}}, \frac{\mathbf{1}}{\mathbf{2}}, \mathbf{0}\right), \frac{\mathbf{3}}{\mathbf{4}}\left(-\frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}}, \frac{\mathbf{1}}{\mathbf{3}}, \mathbf{1}\right)\right\}$ is an orthonormal basis for $\text{Span}\{(\mathbf{2, 1, 1, 0}), (\mathbf{1, 0, 1, 0}), (\mathbf{3, 1, 2, 0}), (\mathbf{0, 1, 0, 1})\}$.

CYU 7.20 A consequence of Theorem 7.10(iii) and CYU 3.11, page 98.

CYU 7.21 We first use the Grahm-Schmidt Process to determine the orthonormal basis $\{\boldsymbol{w_1}, \boldsymbol{w_2}\}$ of W stemming from the given basis $\{\boldsymbol{v_1}, \boldsymbol{v_2}\} = \{(\mathbf{1, 0, 1}), (\mathbf{1, 2, 0})\}$:

$$\boldsymbol{u}_1 = \boldsymbol{v}_1 = (\mathbf{1, 0, 1}): \boldsymbol{w}_1 = \frac{1}{\sqrt{2}}(\mathbf{1, 0, 1})$$

$$\boldsymbol{u}_2 = \boldsymbol{v}_2 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v}_2 \rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle}\boldsymbol{u}_1 = (1, 2, 0) - \frac{(\mathbf{1, 0, 1}) \cdot (\mathbf{1, 2, 0})}{(\mathbf{1, 0, 1}) \cdot (\mathbf{1, 0, 1})}(1, 0, 1) = \left(\frac{\mathbf{1}}{\mathbf{2}}, \mathbf{2}, -\frac{\mathbf{1}}{\mathbf{2}}\right) : \boldsymbol{w}_2 = \frac{\sqrt{2}}{3}\left(\frac{\mathbf{1}}{\mathbf{2}}, \mathbf{2}, -\frac{\mathbf{1}}{\mathbf{2}}\right)$$

Turning to Theorem 7.11, we determine the orthogonal projection, $\boldsymbol{v}_W$, of the vector $v = (\mathbf{2, 0, 1})$ onto W:

$$\begin{aligned}\boldsymbol{v}_W &= \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 \\ &= \left[(\mathbf{2, 0, 1}) \cdot \frac{1}{\sqrt{2}}(\mathbf{1, 0, 1})\right]\left[\frac{1}{\sqrt{2}}(\mathbf{1, 0, 1})\right] + \left[(\mathbf{2, 0, 1}) \cdot \frac{\sqrt{2}}{3}\left(\frac{\mathbf{1}}{\mathbf{2}}, \mathbf{2}, -\frac{\mathbf{1}}{\mathbf{2}}\right)\right]\left[\frac{\sqrt{2}}{3}\left(\frac{\mathbf{1}}{\mathbf{2}}, \mathbf{2}, -\frac{\mathbf{1}}{\mathbf{2}}\right)\right] \\ &= \frac{3}{2}(\mathbf{1, 0, 1}) + \frac{1}{2}\left(\frac{\mathbf{1}}{\mathbf{2}}, \mathbf{2}, -\frac{\mathbf{1}}{\mathbf{2}}\right) = \left(\frac{\mathbf{7}}{\mathbf{4}}, \mathbf{1}, \frac{\mathbf{5}}{\mathbf{4}}\right)\end{aligned}$$

CYU 7.22 **NOTE:** It is easy to see that the function

$$f[(a_0, a_1, \ldots, a_n)] = a_0 + a_1x + a_2x^2 + \ldots + a_nx^n$$

from the standard (dot-product) inner product space $\Re^n$ of Theorem 7.1 (page 279) to the polynomial inner product space P_n of Exercise 23 (page 299) is an isomorphism which **ALSO** preserves the inner product structure of the two spaces:

$$(a_0, a_1, \ldots, a_n) \cdot (b_0, b_1, \ldots, b_n) = \langle f(a_0, a_1, \ldots, a_n), f(b_0, b_1, \ldots, b_n) \rangle = \sum_{i=0}^{n} a_ib_i$$

As such, we could translate the given P_3-problem:

> Find the shortest distance between the vector $\boldsymbol{v} = 3x^2 + \sqrt{3}x$ and the subspace $W = \text{Span}\{x^2 + 1, x^3 + 1\}$ in the inner product space P_3

into the following $\Re^4$-form:

> Find the shortest distance between the vector $\boldsymbol{v} = (0, 3, \sqrt{3}, 0)$ and the subspace $W = \text{Span}\{(0, 1, 0, 1), (1, 0, 0, 1)\}$ in the inner product space $\Re^4$.

You are invited to take the above $\Re^4$-approach. For our part, we will deal directly within the inner product space P_3:

Employing the Grahm-Schmidt Process we go from the basis $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, with $\boldsymbol{v}_1 = x^2 + 1$ and $\boldsymbol{v}_2 = x^3 + 1$, to an orthonormal bases $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ for $W = \text{Span}\{x^2 + 1, x^3 + 1\}$:

$$\boldsymbol{u}_1 = \boldsymbol{v}_1 = x^2 + 1 : \quad \boldsymbol{w}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = \frac{1}{\sqrt{2}}(x^2 + 1).$$

$$\begin{aligned}\boldsymbol{u}_2 = \boldsymbol{v}_2 - \frac{\langle \boldsymbol{u}_1, \boldsymbol{v}_2 \rangle}{\langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle}\boldsymbol{u}_1 &= (x^3 + 1) - \frac{\langle x^2 + 1, x^3 + 1 \rangle}{\langle x^2 + 1, x^2 + 1 \rangle}(x^2 + 1) \\ &= (x^3 + 1) - \frac{1}{2}(x^2 + 1) = x^3 - \frac{1}{2}x^2 + \frac{1}{2} : \boldsymbol{w}_2 = \frac{\boldsymbol{u}_2}{\|\boldsymbol{u}_2\|} = \sqrt{\frac{2}{3}}\left(x^3 - \frac{1}{2}x^2 + \frac{1}{2}\right)\end{aligned}$$

Turning to Theorems 7.11 we determine the projection $\boldsymbol{v}_W$ of $\boldsymbol{v} = 3x^2 + \sqrt{3}x$ onto W:

$$\begin{aligned}
\boldsymbol{v}_W &= \langle \boldsymbol{v}, \boldsymbol{w}_1\rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2\rangle \boldsymbol{w}_2 \\
&= \langle 3x^2 + \sqrt{3}x, \tfrac{1}{\sqrt{2}}(x^2+1)\rangle\Big(\tfrac{1}{\sqrt{2}}(x^2+1)\Big) + \langle 3x^2 + \sqrt{3}x, \sqrt{\tfrac{2}{3}}\Big(x^3 - \tfrac{1}{2}x^2 + \tfrac{1}{2}\Big)\rangle\Big(\sqrt{\tfrac{2}{3}}\Big(x^3 - \tfrac{1}{2}x^2 + \tfrac{1}{2}\Big)\Big) \\
&= \frac{3}{\sqrt{2}}\Big(\frac{1}{\sqrt{2}}(x^2+1)\Big) - \sqrt{\frac{3}{2}}\Big(\sqrt{\frac{2}{3}}\Big(x^3 - \frac{1}{2}x^2 + \frac{1}{2}\Big)\Big) = (x^3 + x^2 + 1)
\end{aligned}$$

Appealing to Theorem 7.12 we calculate the shortest distance between the vector $\boldsymbol{v} = 3x^2 + \sqrt{3}x$ and the subspace $W = \text{Span}\,\{x^2+1, x^3+1\}$:

$$\begin{aligned}
\|\boldsymbol{v} - \boldsymbol{v}_W\| &= \|(3x^2 + \sqrt{3}x) - (x^3 + x^2 + 1)\| \\
&= \|-x^3 - 2x^2 + \sqrt{3} + 1\| = \sqrt{\langle -x^3 - 2x^2 + \sqrt{3} + 1, -x^3 - 2x^2 + \sqrt{3} + 1\rangle} \\
&= \sqrt{(-1)^2 + (-2)^2 + (\sqrt{3})^2 + (1)^2} = \sqrt{9} = 3
\end{aligned}$$

<u>CYU 7.23</u> Determining the eigenvalues of A:

$$\det(A - \lambda I) = \det\begin{bmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 1 & 1 & 2-\lambda \end{bmatrix} \underset{\substack{\uparrow \\ \text{details omitted}}}{=} -(\lambda - 4)(\lambda - 1)^2 \quad \text{Eignevalues: } \lambda = 4, \lambda = 1$$

Determining the corresponding eigenspaces:

$$E[4] = \text{null}(A - 4I) = \text{null}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} = \{\overline{(a, a, a)} \mid a \in \Re\} \text{ since } \text{rref}\begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E[1] = \text{null}(A - I) = \text{null}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \{\overline{(-c-d, c, d)} \mid c, d \in \Re\} \text{ since } \text{rref}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

As is indicated in Theorem 7.14: $(a, a, a) \cdot (-c-d, c, d) = a(-c-d) + ac + ad = 0$

<u>CYU 7.24</u> (a)

$$\left(\begin{bmatrix}1&1&1\\1&2&1\\1&1&3\end{bmatrix}\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\right)\cdot\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix} = (v_1+v_2+v_3,\ v_1+2v_2+v_3,\ v_1+v_2+3v_3)\cdot(w_1, w_2, w_3)$$

$$= w_1(v_1+v_2+v_3)+w_2(v_1+2v_2+v_3)+w_3(v_1+v_2+3v_3) \leftarrow$$

$$\begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix}\cdot\left(\begin{bmatrix}1&1&1\\1&2&1\\1&1&3\end{bmatrix}\begin{bmatrix}w_1\\w_2\\w_3\end{bmatrix}\right) = (v_1, v_2, v_3)\cdot(w_1+w_2+w_3,\ w_1+2w_2+w_3,\ w_1+w_2+3w_3)$$

$$= v_1(w_1+w_2+w_3)+v_2(w_1+2w_2+w_3)+v_3(w_1+w_2+3w_3)$$

$$= w_1(v_1+v_2+v_3)+w_2(v_1+2v_2+v_3)+w_3(v_1+v_2+3v_3) \leftarrow$$

(b) One possible answer: $\left(\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}0\\1\\0\end{bmatrix}\right)\cdot\begin{bmatrix}1\\0\\0\end{bmatrix} = (1,0,0)\cdot(1,0,0) = 1$

while: $\begin{bmatrix}0\\1\\0\end{bmatrix}\cdot\left(\begin{bmatrix}1&1&1\\0&0&0\\0&0&0\end{bmatrix}\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = (0,1,0)\cdot(1,0,0) = 0$

<u>CYU 7.25</u> (a) $\langle T(\boldsymbol{a, b, c}), (\boldsymbol{A, B, C})\rangle = (\boldsymbol{a-b, -a+2b-c, -b+c})\cdot(\boldsymbol{A, B, C})$

$$= A(a-b)+B(-a+2b-c)+C(-b+c) \leftarrow$$

$$\langle(\boldsymbol{a, b, c}), \boldsymbol{T(A, B, C)}\rangle = (a, b, c)\cdot(\boldsymbol{A-B, -A+2B-C, -B+C})$$

$$= a(A-B)+b(-A+2B-C)+c(-B+C)$$

$$= A(a-b)+B(-a+2b-c)+C(-b+c) \leftarrow$$

(b) For $T(a, b, c) = (\boldsymbol{a-b, -a+2b-c, -b+c})$ and $\beta = \{(\boldsymbol{0,1,0}), (\boldsymbol{0,0,1}), (\boldsymbol{1,0,0})\}$:

$$[T]_{\beta\beta} = \begin{bmatrix}1&-1&0\\-1&2&-1\\0&-1&1\end{bmatrix}$$

<u>CYU 7.26</u> Consider the symmetric matrix $A = [T]_{\beta\beta} = \begin{bmatrix}1&-1&0\\-1&2&-1\\0&-1&1\end{bmatrix}$ of CYU 7.25(b). Employing Theorem 6.10 of page 226, we find a basis of eigenvectors for the linear operator

$T(a, b, c) = (a - b, -a + 2b - c, -b + c)$:

$$\det(A - \lambda I) = \det\begin{bmatrix} 1-\lambda & -1 & 0 \\ -1 & 2-\lambda & -1 \\ 0 & -1 & 1-\lambda \end{bmatrix} = -\lambda(\lambda - 1)(\lambda - 3) \quad \text{Eignevalues: } \lambda = 0, \lambda = 1, \lambda = 3$$

Here are the associated eigenspaces:

$$E[0] = \text{null}(A) = \text{null}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = \{\overline{(\boldsymbol{a, a, a})} | a \in \overline{\Re^3}\} \text{ since rref}\begin{bmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E[1] = \text{null}(A - I) = \text{null}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \{\overline{(\boldsymbol{-a, 0, a})} | a \in \overline{\Re^3}\} \text{ since rref}\begin{bmatrix} 0 & -1 & 0 \\ -1 & 1 & -1 \\ 0 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$E[3] = \text{null}(A - 3I) = \text{null}\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} = \{\overline{(\boldsymbol{a, -2a, a})} | a \in \overline{\Re^3}\} \text{ since rref}\begin{bmatrix} -2 & -1 & 0 \\ -1 & -1 & -1 \\ 0 & -1 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Letting $a = 1$ in each of the above eigenspaces we arrive at a normal basis for $\Re^3$ consisting of eigenvectors of T: $\{(\mathbf{1, 1, 1}), (\mathbf{-1, 0, 1}), (\mathbf{1, -2, 1})\}$ (Theorem 6.15); which is easily turned into an orthonormal basis: $\left\{\left(\frac{\mathbf{1}}{\sqrt{\mathbf{3}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}\right), \left(\frac{\mathbf{-1}}{\sqrt{\mathbf{2}}}, \mathbf{0}, \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}\right), \left(\frac{\mathbf{1}}{\sqrt{\mathbf{6}}}, \frac{\mathbf{-2}}{\sqrt{\mathbf{6}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{6}}}\right)\right\}$.

<u>CYU 7.27</u> If $A^{-1} = A^{\mathrm{T}}$ and $B^{-1} = B^{\mathrm{T}}$, then:

$$(AB)^{-1} \underset{\substack{\uparrow \\ \text{Theorem 5.12(iii), page 167}}}{=} B^{-1}A^{-1} = B^{\mathrm{T}}A^{\mathrm{T}} \underset{\substack{\uparrow \\ \text{Exercise 19(f), page 162}}}{=} (AB)^{\mathrm{T}}$$

<u>CYU 7.28</u> In CYU 7.23 we showed that the matrix $\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}$ has eigenvalues $\lambda = 4, \lambda = 1$ with $E[4] = \{(a, a, a) | a \in \Re\}$, $E[1] = \{(-a - b, a, b) | a, b \in \Re\}$, with $\{(1, 1, 1)\}$ a basis for $E[4]$, and $\{(-1, 0, 1), (-1, 1, 0)\}$ a basis for $E[1]$ (set $a = 0$ and $b = 1$, and then set $b = 0$ and $a = 1$). Applying the Grahm-Schmidt process (Theorem 7.9, page 303) to $\{(-1, 0, 1), (-1, 1, 0)\}$, we arrive at the orthogonal basis $\left\{(-1, 0, 1), \left(-\frac{1}{2}, 1, -\frac{1}{2}\right)\right\}$ of $E[1]$, and to the orthonormal basis of eigenvectors $\left\{\left(\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right), \left(-\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}}\right), \left(-\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}\right)\right\}$.

Turning to the marginal comment on page 314 we conclude that:

$$P^{\mathrm{T}}\begin{bmatrix}2 & 1 & 1\\ 1 & 2 & 1\\ 1 & 1 & 2\end{bmatrix}P = \begin{bmatrix}4 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}, \text{ where } P = \begin{bmatrix}1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6}\\ 1/\sqrt{3} & 0 & 1/\sqrt{6}\\ 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6}\end{bmatrix}$$

Appendix C
Answers to Selected Exercises

1.1 Systems of Linear Equations, page 11.

1. $\left[\begin{array}{ccc|c} 3 & -3 & 1 & 2 \\ 5 & 5 & -9 & -1 \\ -3 & -4 & 1 & 0 \end{array}\right]$ **3.** $\left.\begin{aligned} 5x+y+4z &= 6 \\ -2x-3y+z &= 4 \\ \frac{1}{2}x-y &= 0 \end{aligned}\right\}$ **5.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ **7.** $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$

9. $x = 1, y = 0, z = 2$ **11.** $x_1 = 1, x_2 = 2, x_3 = 2, x_4 = 2, x_5 = -1$

13. $x = 0, y = -4, z = 2$ **15.** $x = \frac{1}{2}, y = \frac{1}{4}, z = -\frac{1}{2}, w = \frac{3}{4}$

19. No: first non-zero entry in last row is not 1. **23.** $x = 5, y = -2, z = 2$

25. $x = 12, y = -5, z = 1, w = 0$

1.2 Consistent and Inconsistent Systems of Equations, page 23.

1. $\{(r, 2) | (r \in \Re)\}$ **3.** $\{(-2, -2, -2)\}$ **5.** $\{(-s-3t, r, 1-2t, 1-2s, s, t) | r, s, t \in \Re\}$

7. $\{(11-5r, -6+3r, r) | r \in \Re\}$ 9. $\{(7, 9, -6)\}$ 11. $\left\{2-r, \frac{1+2r}{5}, \frac{3+6r}{5}, r \middle| r \in \Re\right\}$

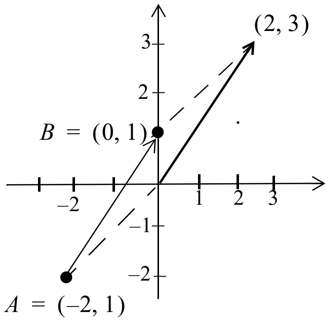
13. Yes. 15. No. Solutions if and only if a, b, c, satisfy the equation $4a - b + 2c = 0$.

17. No. **19**. Yes. **23.** $\left\{\left(-\frac{3}{2}r, r, 0\right) \middle| r \in R\right\}$ **25.** $\{(-r-11s, r+6s, r, s) | r, s \in R\}$

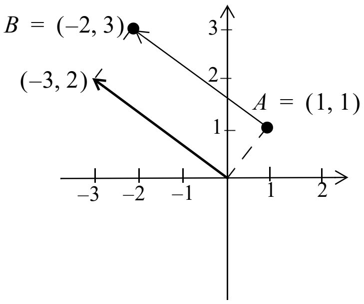
27. No. **29.** No. **31.** $a \neq \pm 1$ **33.** None. **35.** $ab \neq 1$ **37.** $ad - bc \neq 0$

2.1 Vectors in the Plane and Beyond, page 38.

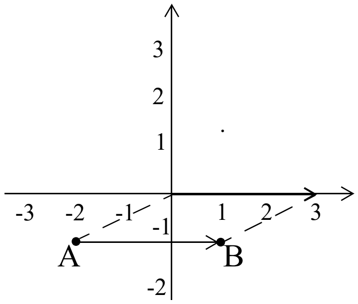
1.



3.



5.



7. $(2, 0, -2)$ **9.** $(-9, -1, 11)$ **11.** $(13, -13)$ **15.** $(-6, 6)$ **17.** $(2, -5)$

19. (a) $r = -14, s = 10$ (b) $r = -\frac{7}{5}, s = \frac{1}{10}$ (c) $r = \frac{5}{7}, s = \frac{1}{14}$

21. $r = -\frac{19}{6}, s = \frac{17}{6}, t = \frac{11}{6}$ **23.** $\left(\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$

2.2 Abstract Vectors Spaces. page 49.

Follow instructions.

2.3 Properties of Vectors Spaces, page 57.

Follow instructions

2.4 Subspaces, page 65.

1. Yes. **3.** Yes. **5.** No. **7.** Yes. **9.** No. **11.** No. **13.** Yes.
15. Yes. **17.** No. **19.** No. **21.** Yes. **23.** Yes. **25.** Yes. **27.** Yes.
29. No. **31.** Yes. **33.** No. **35.** Yes. **37.** Yes.

2.5 Lines and Planes, page 73.

1. $\{r(1, 5) | r \in \Re\}$ **3.** $\{r(5, -1) | r \in \Re\}$ **5.** $\{(1, 3) + r(1, -7) | r \in \Re\}$

7. $\{(3, 5) + r(0, 2) | r \in \Re\}$ **13.** $\{(3, 7) + r(1, 5) | r \in \Re\}$ **15.** $\{(3, 7) + r(5, -1) | r \in \Re\}$

17. $\{(3, 7) + r(1, -7) | r \in \Re\}$ **19.** $\{(3, 7) + r(0, 2) | r \in \Re\}$ **21.** $\{(3, 7) + r(5, -1) | r \in \Re\}$

23. $\{(3, 7) + r(1, 5) | r \in \Re\}$ **25.** $\{(3, 7) + r(7, 1) | r \in \Re\}$ **27.** $\{(3, 7) + r(0, 1) | r \in \Re\}$

29. $\{r(2, 4, 5) | r \in \Re\}$ **31.** $\{r(-2, 4, 0) | r \in \Re\}$ **33.** $\{(2, 4, 5) + r(1, -3, -4)\}$

35. $\{(2, 1, 0) + r(1, 3, -1)\}$ **41.** $\{(1, 2, -1) + r(2, 4, 5) | r \in \Re\}$

43. $\{(1, 2, -1) + r(-2, 4, 0) | r \in \Re\}$ **45.** $\{(1, 2, -1) + r(1, -3, -4) | r \in \Re\}$

47. $\{(1, 2, -1) + r(1, 3, -1) | r \in \Re\}$ **49.**$\{r(1, 3, 2) + s(2, 1, 1) | r, s \in \Re\}$

51. $\{r(2, 0, 0) + s(0, 2, 0) | r, s \in \Re\}$ **53.** $\{(3, 4, 1) + r(1, 3, -4) + s(2, 3, 2) | r, s \in \Re\}$

55.$\{(2, 4, -3) + r(3, -3, 8) + s(2, -3, 2) | r, s \in \Re\}$

3.1 Spanning Sets, page 84.

1. No. **3.** Yes. **5.** N0 **7.** No. **9.** No. **11.** No. **13.** $\cos 2x = \cos^2 x - \sin^2 x$

15. $\sin\left(\frac{\pi}{7} - x\right) = \left(\sin\frac{\pi}{7}\right)\cos x - \left(\cos\frac{\pi}{7}\right)\sin x$ **17.** Span.

19. Do not span. Just about any randomly chosen four-tuple will not be a linear combination of the given vectors (check it out).

21. Do not span. Just about any randomly chosen four-tuple will not be a linear combination of the given vectors (check it out).

23. Do not span. Just about any randomly chosen four-tuple will not be a linear combination of the given vectors (check it out).

25. Span. **27.** All $c \neq 0$.

3.2 Linear Independence, page 91.

1. Yes. **3.** No. **5.** No. **7.** Yes. **9.** No. **11.** No. **13.** Yes. **15.** Yes.

17. Yes. **19.** Yes. **21.** No. **23.** No. **25.** Yes. **27.** No. **29.** $a = 3$

3.3 Bases, page 104.

1. (a) $\left(-3, \frac{5}{2}\right) = -3e_1 + \frac{5}{2}e_2$ (b) $(3, \sqrt{2}, 0) = 3e_1 + \sqrt{2}e_2 + 0e_3$

3. (a) $x^2 + 3x - 1 = \frac{19}{13}(2x^2 + 3) - \frac{25}{13}(x^2 - x) + \frac{14}{13}(x - 5)$

5. No. **7.** Yes. **9.** Yes. **11.** Yes. **13.** Yes. **15.** No. **17.** No. **19.** Yes.

29. $\left\{\begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 1 & -1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}\right\}$ **31.** Do not span.

33. A basis for $\text{Span}(S)$: $\{(2, 1, 4), (-1, 3, 2)\}$. A basis for $\Re^3$: $\{(2, 1, 4), (-1, 3, 2), (1, 1, 1).\}$

35. A basis for $\text{Span}(S) = \Re^3$: $S = \{(1, 1, 3), (-1, 3, 2), (3, 2, -1)\}$

37. A basis for $\text{Span}(S) = \Re^5$:
$S = \{(1, 3, 1, 3, 2), (2, 4, 1, 4, 2), (1, 1, 2, 0, 2), (2, 2, 1, 1, 1), (1, 2, 3, 4, 5)\}$

39. A basisfor $\text{Span}(S)$: $\{5, -x^3 - x, x^4 + x^3 + x^2 + x + 1, 2x^4 - 2x^2\}$. A basis for $\Re^3$:

$\{5, -x^3 - x, x^4 + x^3 + x^2 + x + 1, 2x^4 - 2x^2, 1 + x\}$

41. $\{\sin x, \cos x, \sin^2 x, \cos^2 x, \sin 2x\}$ **43.** $\{(-1, 1, 0, 0), (1, 0, 1, 0), (-1, 0, 0, 1)\}$

45. $\{x^3 + x^2 + 2x, x + 1\}$ **47.** $c \neq 0$ **49.** $a \neq 0, b \neq 0$, and $a \neq b$

4.1 Linear Transformations, page 120.

1. Yes **3.** No **5.** No **7.** No **9.** Yes **11.** No **13.** Yes **15.** No

17. No **19.** (a) $(4, 10, -2)$ (b) $\left(\frac{a+b}{2}, 2a, -a+b\right)$

21. (a) $\begin{bmatrix} 8 & 3 \\ 13 & 8 \end{bmatrix}$ (b) $\begin{bmatrix} a+b & b \\ 2a+b & a+b \end{bmatrix}$ **27.** $b = 0$ **39.** $(L\circ T)(a, b) = (2a+2b, -a-b)$

41. $(L\circ T)(a, b) = (3a+b)x + (6a+2b)$ **43.** $(K\circ L\circ T)(a) = (-2a, 0, 2a)$

4.2 Kernel and Image, page 131.

1. Linear. Nullity: 0, Rank: 1.

3. Linear. Nullity: 0, Rank: 1. A basis for the image space: $\{(1, -1)\}$

5. Not linear. **7.** Linear. Nullity: 0, Rank: 2.

9. Linear. Nullity: 2, Rank: 2. A basis for the kernel: $\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right\}$.

A basis for the image space: $\{x-1, -x^2+1\}$. **11.** Not linear.

13. Linear. Nullity: 1, Rank: 2. A basis for the kernel: $\{x^2-x\}$.

A basis for the image space: $\{(0, 1), (1, 1)\}$. **15.** Linear. Nullity: 0, Rank: 4.

17. Linear. Nullity: 0, Rank: 3. A basis for the image space: $\left\{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}, \begin{bmatrix} 0 & -1 \\ 0 & 1 \end{bmatrix}\right\}$.

19. Nullity: 1, Rank: 3. A basis for the kernel: $\{1\}$. A basis for the image space: $\{x^2, x, 1\}$.

21. Nullity: 0, Rank: 4. A basis for the image space: $\{x^3, x^2, x, 1\}$

23. Nullity: 0, Rank: 3. A basis for the image space: $\{x^2, x, 1\}$.

25. Nullity: 2, Rank: 1. A basis for the kernel: $\left\{\frac{x^2}{3} - \frac{1}{3}, \frac{x}{2} - \frac{1}{2}\right\}$. A basis for the image space: $\{1\}$.

27. Nullity: 0, Rank: 3. A basis for the image space: $\{x^3, x^2, x\}$. **29.** (a) $(0, -3, 3)$

(b) $x^3 + x$ and 5. **31.** Nullity: 0, Rank: 3. A basis for the image space: $\left\{\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 1 & 3 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right\}$.

33. (a) $T(a, b) = (9a, a, 3a)$ (b) $T(a, b) = (9a, b, 0)$ (c) None exists.

35. $\text{nullity}(T) = n$, $\text{rank}(T) = 0$; $\text{nullity}(T) = 0$, $\text{rank}(T) = n$.

37. $\text{nullity}(T) = 1$, $\text{rank}(T) = 2$; $\text{nullity}(T) = 2$, $\text{rank}(T) = 1$;
$\text{nullity}(T) = 3$, $\text{rank}(T) = 0$.

39. $\text{nullity}(T) = 1$, $\text{rank}(T) = 4$

4.3 Isomorphisms, page 145.

1. $f^{-1}(x) = -\frac{1}{5}x$ **3.** $f^{-1}(a, b) = \left(b, -\frac{1}{2}a\right)$ **5.** $f^{-1}(ax^2 + bx + c) = (-c, a, b)$

7. Isomorphism. **9.** Not an isomorphism. **11.** Isomorphism. **13.** Not an isomorphism.
15. Isomorphism. **17.** Isomorphism. **19.** Isomorphism. **21.** Not an isomorphism.

23. $f^{-1}(x, y) = (y + 4, x - y - 7)$. Zero: $(3, -4)$ and $-(x, y) = (-x + 6, -y - 8)$.

5.1 Matrix Multiplications, page 161.

1. $\begin{bmatrix} 0 & 11 \\ 1 & -4 \\ -1 & 15 \end{bmatrix}$ **3.** $\begin{bmatrix} 0 & 0 & 2 \\ 2 & 0 & 0 \\ 0 & 9 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 6 \\ 6 & 0 & 0 \end{bmatrix}$ 5. Column: $\{(3, 10, 4), (2, 6, 2)\}$
Row: $\{(3, 2, 1, 5), (10, 6, 1, 10)\}$

7. Column: $\{(1, -1, 5, 5), (-1, 0, 3, 2), (0, -1, -2, -1)\}$
Row: $\{(1, 1, 0, 1), (-1, 0, -1, 4), (5, 3, -2, 7)\}$

25. $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}^n = \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix}$

27. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}$ 29. $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}^n = \begin{bmatrix} 1 & 0 \\ 0 & 2^n \end{bmatrix}$

5.2 Invertible Matrices, page 175.

1. $\begin{bmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{5} \end{bmatrix}$ **3.** $\begin{bmatrix} \frac{3}{4} & -\frac{1}{2} \\ -\frac{1}{4} & \frac{1}{2} \end{bmatrix}$ **5.** $\begin{bmatrix} -4 & 1 & 2 \\ 2 & 0 & -1 \\ -1 & 0 & 1 \end{bmatrix}$ **7.** Singular. **9.** Singular. **11.** $\begin{bmatrix} -1 & -\frac{8}{5} & \frac{2}{5} \\ \frac{1}{2} & \frac{2}{5} & -\frac{1}{10} \\ 1 & \frac{5}{3} & -\frac{1}{3} \end{bmatrix}$

13. Singular. **15.** $\begin{bmatrix} \frac{1}{12} & 0 & -\frac{1}{3} & \frac{1}{3} \\ 0 & 2 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ \frac{1}{12} & \frac{1}{2} & -\frac{1}{3} & -\frac{1}{6} \end{bmatrix}$ 17. $E_1 = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $E_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

19. $X = \frac{1}{2}ABA^{-1}$ **21.** $X = A^{-1}BAB$ **23.** $X = AB^2A$

5.3 Matrix Representation of Linear Maps, page 186.

1. $[(2,3)]_\beta = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $[T(2,3)]_\beta = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$

3. $[(2,3)]_\beta = \begin{bmatrix} \frac{8}{5} \\ -\frac{1}{5} \end{bmatrix}$, $[T(2,3)]_\beta = \begin{bmatrix} \frac{17}{5} \\ -\frac{4}{5} \end{bmatrix}$

5. $[(2,3,1)]_\beta = \begin{bmatrix} 4 \\ 6 \\ -3 \end{bmatrix}$, $[T(2,3,1)]_\beta = \begin{bmatrix} 7 \\ 7 \\ -6 \end{bmatrix}$

7. $[(1,2)]_\beta = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $[T(1,2)]_\gamma = \begin{bmatrix} 2 \\ 1/2 \\ -3 \end{bmatrix}$

9. $[x^2+x+1]_\beta = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$, $[T(x^2+x+1)]_\gamma = \begin{bmatrix} 0 \\ -2 \\ 2 \\ 0 \end{bmatrix}$

11. $[T]_{\gamma\beta} = \begin{bmatrix} 1 & 1 & -2 \\ \frac{1}{2} & \frac{1}{2} & 1 \end{bmatrix}$, $[T(1,2,1)]_\gamma = [T]_{\gamma\beta}[(1,2,1)]_\beta = \begin{bmatrix} -1 \\ \frac{3}{2} \end{bmatrix}$

13. $\begin{bmatrix} -3 & -3 \\ 2 & 1 \\ -1 & 1 \end{bmatrix}$, $[T(1,2)]_\gamma = [T]_{\gamma\beta}[(1,2)]_\beta = \begin{bmatrix} -3 \\ 2 \\ -1 \end{bmatrix}$

15. $[T]_{\gamma\beta} = \begin{bmatrix} 0 & 1 \\ 0 & -1 \\ 1 & 3 \end{bmatrix}$, $[T(2x+1)]_\gamma = [T]_{\gamma\beta}[2x+1]_\beta = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

17. $[I]_{\beta'\beta} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$, $[I(1,2,1)]_{\beta'} = [I]_{\beta'\beta}[(1,2,1)]_\beta = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$

19. $[T]_{\beta'\beta} = \begin{bmatrix} 2 & 0 & 3 & 2 \\ 1 & 0 & 1 & 1 \\ 5 & 1 & 6 & 4 \\ -2 & 1 & -3 & -2 \end{bmatrix}$, $\left[T\left(\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}\right)\right]_{\beta'} = [T]_{\beta'\beta}\left[\begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}\right]_\beta = \begin{bmatrix} 5 \\ 2 \\ 8 \\ -8 \end{bmatrix}$

21. $[T]_{\beta\beta} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}$, $[T(1,3)]_\beta = [T]_{\beta\beta}[(1,3)]_\beta = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$

23. (a) $[T]_{\gamma\beta} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ (b) $[D]_{\beta\gamma} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$

(c) $[D\circ T]_{\beta\beta} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (d) $[T\circ D]_{\gamma\gamma} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ (e) $[T\circ D\circ T]_{\gamma\beta} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$ (f) $[D\circ T\circ D]_{\beta\gamma} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 9 \end{bmatrix}$

25. $[D]_{\beta\beta} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $[D(5x^3+3x^2)]_\beta = [D]_{\beta\beta}[5x^3+3x^2]_\beta = \begin{bmatrix} 0 \\ 3 \\ 5 \\ 0 \end{bmatrix}$

27. $T(a,b) = (-3a+4b, 2a)$

29. T is the identity map.

31. $[L\circ T]_{\delta\beta} = [L]_{\delta\gamma}[T]_{\gamma\beta} = \begin{bmatrix} 0 & 0 \\ -3 & -1 \\ \frac{2}{3} & \frac{1}{3} \end{bmatrix}$

35. $T^{-1}(ax+b) = \left(a, \frac{b}{2}\right)$, $[T]_{\gamma\beta} = \begin{bmatrix} -2 & -2 \\ 4 & 2 \end{bmatrix}$, $[T^{-1}]_{\beta\gamma} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ -1 & -\frac{1}{2} \end{bmatrix}$

37. $T^{-1}\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a, b-a, d+c, c)$, $[T]_{\gamma\beta} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ -2 & -1 & 1 & -1 \\ 0 & 1 & -1 & 0 \end{bmatrix}$, $[T^{-1}]_{\beta\gamma} = \begin{bmatrix} -1 & 0 & -1 & -1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix}$

39. $[L]_{\gamma\beta} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & -1 & -1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

5.4 Change of Basis and Similar Matrices, page 199.

1. $[(2,5)]_{\beta'} = [I]_{\beta'\beta}[(2,5)]_{\beta} = \begin{bmatrix} \frac{12}{5} \\ \frac{1}{5} \end{bmatrix}$

3. $[(2,3,-1)]_{\beta'} = [I]_{\beta'\beta}[(2,3,-1)]_{\beta} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}$

5. $[2x^2+x+1]_{\beta'} = [I]_{\beta'\beta}[2x^2+x+1]_{\beta} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}$

7. $\left[\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\right]_{\beta'} = [I]_{\beta'\beta}\left[\begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}\right]_{\beta} = \begin{bmatrix} 0 \\ 1 \\ -2 \\ 2 \end{bmatrix}$

9. $[T]_{\beta'\beta'} = [I]_{\beta'\beta}[T]_{\beta\beta}[I]_{\beta\beta'} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

11. $[T]_{\beta'\beta'} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'} = \begin{bmatrix} -1 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & 0 \\ 2 & 0 & 1 \end{bmatrix}$

13. $[T]_{\beta'\beta'} = [I_V]_{\beta'\beta}[T]_{\beta\beta}[I_V]_{\beta\beta'} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & \frac{9}{2} \\ 0 & 0 & 0 & 0 \end{bmatrix}$

15. $\{(0,1,1), (1,1,0), (0,2,1)\}$

19. I_n

25. $[T]_{\gamma'\beta'} = [I_W]_{\gamma'\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'} = \begin{bmatrix} 0 & -\frac{1}{3} \\ 0 & \frac{1}{3} \\ 1 & \frac{1}{3} \end{bmatrix}$

27. $[T]_{\gamma'\beta'} = [I_W]_{\gamma'\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'} = \begin{bmatrix} -1 & 0 & -1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$

29. $[T]_{\gamma'\beta'} = [I_W]_{\gamma'\gamma}[T]_{\gamma\beta}[I_V]_{\beta\beta'} = \begin{bmatrix} -4 & -1 & -2 \\ 2 & 1 & 1 \end{bmatrix}$

6.1 Determinants, page 215.

1. 6 **3.** 6 **5.** 0 **7.** 86 **9.** 9 **11.** 9 **13.** 81

15. $k \neq 0$, $k \neq 1$ **17.** $k \neq 0, \pm 1$

6.2 Eigenspaces, page 228.

1. $\lambda = -1, 4$; $E[-1] = \{(-r, 2r) | r \in \Re\}$, $E[4] = \{(r, 3r) | r \in \Re\}$

3. $\lambda = 4, 5$; $E[4] = \{(r, 3r) | r \in \Re\}$, $E[5] = \{(r, 2r) | r \in \Re\}$

5. $\lambda = 2, 8$; $E[2] = \{(-2r + s, r, s) | r, s \in \Re\}$, $E[8] = \{(r, -2r, r) | r \in \Re\}$

7. $\lambda = 2, \pm 1$; $E[2] = \{(0, 3r, r, 0) | r \in \Re\}$, $E[1] = \{(2r, 0, 3r, 4r) | r \in \Re\}$
$E[-1] = \{(0, 0, r, 0) | r \in \Re\}$

9. $\lambda = 1, 2, 3$; $E[1] = \{(3, -2r, r) | r \in \Re\}$, $E[2] = \{(r, 0, 0) | r \in \Re\}$,
$E[3] = \{(0, r, 0) | r \in \Re\}$

11. $\lambda = \pm 3$; $E[3] = \{(r - s, r, s) | r, s \in \Re\}$, $E[-3] = \{(r, r, 3r) | r \in \Re\}$

13. $\lambda = 0, -1$; $E[0] = \{(r, 0, r, r) | r \in \Re\}$, $E[-1] = \{(r, 0, 0, s) | r, s \in \Re\}$

15. $\lambda = 0$, $E[0] = \Re$

17. $\lambda = 5, -16$; $E[5] = \{(2r, r) | r \in \Re\}$, $E[-16] = \{(r, 4r) | r \in \Re\}$

19. $\lambda = 4, -2$; $E[4] = \{(3r, r, 2r) | r \in \Re\}$, $E[-2] = \{(3r - 3s, r, s) | r, s \in \Re\}$

21. $\lambda = 2, -1, \pm\sqrt{3}$; $E[2] = \{(3r, r, 34r, 9r) | r \in \Re\}$, $E[-1] = \{(0, -r, r, 3r) | r \in \Re\}$
$E[\sqrt{3}] = \{(0, 0, (2 + \sqrt{3})(r, r)) | r \in \Re\}$, $E[\sqrt{3}] = \{(0, 0, (2 - \sqrt{3})(r, r)) | r \in \Re\}$

23. $\lambda = \pm 1$; $E[1] = \{rx + r | r \in \Re\}$, $E[-1] = \{-rx + r | r \in \Re\}$

25. $\lambda = \pm 1$, $E[1] = \{-rx^2 + sx + r | r, s \in \Re\}$, $E[-1] = \{rx^2 + r | r \in \Re\}$

27. $\lambda = 0, 1, 2$, $E[0] = \{rx^2 | r \in \Re\}$, $E[1] = \{-rx^3 + 2rx^2 | r \in \Re\}$,
$E[2] = \{-rx^2 + 2rx | r \in \Re\}$

29. $\lambda = 4$; $E[4] = \left\{ \begin{bmatrix} 57r & 38r \\ 84r & -72r \end{bmatrix} \middle| r \in \Re \right\}$ **31.** $\lambda = 0$; $E[0] = V$

39. (a) $a^2 + d^2 - 2ad + 4bc > 0$ (b) $a^2 + d^2 - 2ad + 4bc = 0$ (c) $a^2 + d^2 - 2ad + 4bc < 0$

6.3 Diagonalization, page 243.

1. Not diagonalizable. **3.** Diagonalizable. **5.** Diagonalizable. **7.** Diagonalizable.

9. Not diagonalizable. **11.** Diagonalizable. **13.** Diagonalizable. **15.** Diagonalizable.

17. Not diagonalizable. **19.** Diagonalizable. **21.** Diagonalizable. **23.** Diagonalizable.

25. Not diagonalizable. **27.** Diagonalizable. **29.** Not diagonalizable. **31.** Diagonalizable.

33. Not diagonalizable.

6.4 Applications, page 256.

1. $s_k = \frac{2}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^k - \left(\frac{1+\sqrt{5}}{2}\right)^k\right]$

3. $s_k = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{k-1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k-1} + 2\left(\frac{1+\sqrt{5}}{2}\right)^{k-2} - 2\left(\frac{1-\sqrt{5}}{2}\right)^{k-2}\right]$

5. $s_k = 3 \cdot 2^{k-1} - 2$ **7.** $s_k = \frac{2^{k+1}}{3} + \frac{1}{2} + \frac{1}{6}(-1)^k$ **13.** $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -c_1e^{-2x} - c_2e^{-3x} \\ c_1e^{-2x} + 2c_2e^{-3x} \end{bmatrix}$

15. $\begin{bmatrix} f(x) \\ g(x) \\ h(x) \end{bmatrix} = \begin{bmatrix} -2c_1 + c_2 - c_3e^{6x} \\ c_1 - 2c_3e^{6x} \\ c_2 + c_3e^{6x} \end{bmatrix}$ **17.** $\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} c_1e^{-7x} + c_2e^x - 7c_3e^{-x} \\ c_1e^{-7x} - 2c_2e^x - 3c_3e^{-x} \\ c_1e^{-7x} + c_2e^x + c_3e^{-x} \end{bmatrix}$ **19.** $\begin{bmatrix} f_1(x) \\ f_2(x) \end{bmatrix} = \begin{bmatrix} e^{3x} + e^x \\ -e^{3x} + e^x \end{bmatrix}$

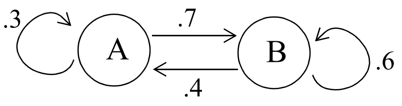
21. $\begin{bmatrix} f(x) \\ g(x) \\ h(x) \end{bmatrix} = \begin{bmatrix} -\frac{3}{8}e^{-7x} + \frac{1}{2}e^x + \frac{7}{8}e^{-x} \\ -\frac{3}{8}e^{-7x} - e^x + \frac{3}{8}e^{-x} \\ -\frac{3}{8}e^{-7x} + \frac{1}{2}e^x - \frac{1}{8}e^{-x} \end{bmatrix}$

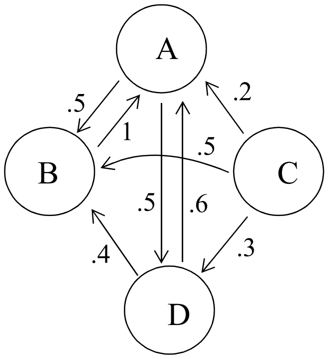
23. A-concentration of alcohol: $4.8 – 2.8e^{-\frac{5}{2}t}$

B-concentration of alcohol: $3.2 + 2.8e^{-\frac{5}{2}t}$

6.5 Markov Chains, page 270.

1. Regular **3.** No **5.** Regular **7.** $\begin{array}{c|cc|c} & \text{A} & \text{B} & \\ & .4 & .6 & \text{A} \\ & .9 & .1 & \text{B} \end{array}$ **9.**



11. 

13. (a) Probability 1 of ending up at A.
(0 probability of ending up in B or C)

(b) Probability 1 of ending up at A.
(0 probability of ending up in B or C)

(c) Probability 1 of ending up at A
(0 probability of ending up in B or C)

15. $\begin{bmatrix} \frac{2}{5} \\ \frac{3}{5} \end{bmatrix}$ **17.** $\begin{bmatrix} \frac{3}{8} \\ \frac{5}{8} \end{bmatrix}$ **19.** $\begin{bmatrix} \frac{3}{10} \\ \frac{1}{2} \\ \frac{1}{5} \end{bmatrix}$ **29.** (a) $\begin{bmatrix} .62 \\ .39 \end{bmatrix}$ (b) $\begin{bmatrix} .64 \\ .36 \end{bmatrix}$ (c) $\begin{bmatrix} .66 \\ .34 \end{bmatrix}$ **31.** $\begin{bmatrix} .40 \\ .22 \\ .38 \end{bmatrix}$

33. (a) 0.247 (b) 0.265 (c) 0.273 (d) $\begin{bmatrix} .17141 \\ .27767 \\ .24065 \\ .12204 \\ .18823 \end{bmatrix}$ (e-a) 0.281 (e-b) 0.274 (e-c) 0.274

(e-d) The initial state of the system has no bearing on the fixed or final state of the system. Independent of the initial state, eventually (to five decimal places):

17.141%, 17.767%, 24.065%, 12.204%, and 18.823% of the employees
will be enrolled in plans A, B, C, D, and E, respectively.

35. (a) 0.63 (b) 0.62 (c) 0.59 (d-a) 0.50 (d-b) 0.49 (d-c) 0.48 (e) $\begin{bmatrix} .33333 \\ .22222 \\ .44444 \end{bmatrix}$

37. (a) 0.33 (b) 0.25 (c) $\begin{bmatrix} 1/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{bmatrix}$

7.1 Dot Product, page 289.

1. 33 **3.** $\sqrt{13}$ **5.** $\sqrt{15}$ **7.** $a = \pm\sqrt{\frac{15}{2}}$ **9.** $a = c, b = \frac{9}{2c}$ for any $c \neq 0$

11. $\cos^{-1}\frac{31}{\sqrt{83}\sqrt{21}} \approx 42°$ **13.** $\left(\frac{4}{5}, 0, -\frac{8}{5}\right), \left(\frac{6}{5}, 3, \frac{3}{5}\right)$

15. A normal form: $(2, 1, 3) \cdot (x-1, y-3, z+1) = 0$; a general form: $2x + y + 3z = 1$; a vector form: $\{(0, 1, 0) + r(1, -2, 0) + s(0, -3, 1) | r, s \in \Re\}$.

17. A normal form: $(4, 0, 1) \cdot (x-1, y, z+3) = 0$; a vector form: $\{(0, 0, 1) + r(1, 0, -4) + s(0, 1, 0) | r, s \in \Re\}$. **19.** $\sqrt{2}$ **21.** $\frac{1}{19}\sqrt{5529}$

23. $\frac{9}{\sqrt{21}}$ **25** (a) $\left\{\left(a, b, \frac{-a-3b}{2}\right) \middle| a, b \in \Re\right\}$ (b) $\{(7c, -5c, 4c) | c \in \Re\}$ (c) $\{(0, 0, 0)\}$

27. $3x + 2y + 6z = 6$

7.2 Inner Product, page 298.

1. 7 **3.** $\sqrt{29}$ **5.** $\cos^{-1}\left(\frac{-19}{7\sqrt{21}}\right) \approx 126°$ **9.** $\sqrt{14}$ **11.** $\sqrt{77}$ **13.** $\cos^{-1}\left(\frac{-18}{\sqrt{378}}\right) \approx 158°$

19. $\sqrt{6}$ **21.** $\cos^{-1}\left(\frac{-1}{2\sqrt{21}}\right) \approx 96°$ **25** $\sqrt{\int_0^1 (2x^2 - x + 3)^2 dx} \approx 10.1$

27. $\cos^{-1}\left(\frac{\int_0^1 (2x^2 - x + 3)(-x^2 + x - 5) dx}{\sqrt{\int_0^1 (2x^2 - x + 3)^2 dx}\sqrt{\int_0^1 (-x^2 + x - 5)^2 dx}}\right) \approx 174°$ **29.** $\sqrt{\int_0^1 (e^x - x)^2 dx} \approx 1.2$

31. $\sqrt{\int_{-\pi}^{\pi} \sin^2 x dx} \approx 0.08$ **33.** $\cos^{-1}\left(\frac{\int_{-\pi}^{\pi} \sin x \cos x dx}{\sqrt{\int_{-\pi}^{\pi} \sin^2 x dx}\sqrt{\int_{-\pi}^{\pi} \cos^2 x dx}}\right) = 0$

7.3 Orthogonality, page 310

1. $\left\{\left(\frac{\mathbf{1}}{\sqrt{\mathbf{3}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{3}}}\right), \left(\frac{\mathbf{-1}}{\sqrt{\mathbf{6}}}, \frac{\mathbf{2}}{\sqrt{\mathbf{6}}}, \frac{\mathbf{-1}}{\sqrt{\mathbf{6}}}\right), \left(\frac{\mathbf{-1}}{\sqrt{\mathbf{2}}}, \frac{\mathbf{0}}{\sqrt{\mathbf{2}}}, \frac{\mathbf{1}}{\sqrt{\mathbf{2}}}\right)\right\}$ 3. No **5.** $\left\{\frac{1}{\sqrt{30}}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \frac{1}{\sqrt{5}}\begin{bmatrix} 2 & -1 \\ 0 & 0 \end{bmatrix}, \frac{\sqrt{134}}{5}\begin{bmatrix} 0 & 0 \\ \frac{1}{3} & -\frac{1}{4} \end{bmatrix}\right\}$

7. No **9.** $\frac{4}{\sqrt{3}}\left(\frac{x^2}{\sqrt{3}} + \frac{x}{\sqrt{3}} + \frac{1}{\sqrt{3}}\right) - \frac{5}{\sqrt{6}}\left(\frac{x^2}{\sqrt{6}} - \frac{2x}{\sqrt{6}} + \frac{1}{\sqrt{6}}\right) + \frac{3}{\sqrt{2}}\left(\frac{x^2}{\sqrt{2}} - \frac{1}{\sqrt{2}}\right)$

11. $a = r, b = -1 - 3r$ for $r \in \Re$. **13.** $a = 1, b = 2$ **15.** $b = r, a = \frac{-12r + 6}{-3 + 4r}$ for $r \in \Re$.

17. $\left\{\left(-\frac{1}{\sqrt{105}}, \frac{10}{\sqrt{105}}, \frac{2}{\sqrt{105}}\right), \left(\frac{2}{\sqrt{5}}, 0, 1\sqrt{5}\right)\right\}$ **19.** $\left\{\left(\frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{1}{\sqrt{5}}\right), \left(\frac{-7}{\sqrt{226}}, \frac{8}{\sqrt{226}}, \frac{-7}{\sqrt{226}}, \frac{8}{\sqrt{226}}\right)\right\}$

21. $\left\{\left(\frac{2}{\sqrt{5}}, 0, \frac{1}{\sqrt{5}}\right), \left(-\frac{1}{\sqrt{341}}, \frac{18}{\sqrt{341}}, \frac{4}{\sqrt{341}}\right)\right\}$ **23.** $\left\{\left(\frac{1}{\sqrt{5}}, \frac{0}{\sqrt{5}}, \frac{2}{\sqrt{5}}, \frac{0}{\sqrt{5}}\right), \left(\frac{36}{\sqrt{3256}}, \frac{-40}{\sqrt{3256}}, \frac{-18}{\sqrt{3256}}, \frac{6}{\sqrt{3256}}\right)\right\}$

25. $\left\{\sqrt{5}x^2, \frac{25}{\sqrt{31}}x^2 - \frac{12}{\sqrt{31}}x - \frac{6}{\sqrt{31}}\right\}$ **27.** $\left\{\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}, 0\right\}$ **29.** $\left\{\left(0, 0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right), \left(\frac{\sqrt{2}}{\sqrt{11}}, \frac{2\sqrt{2}}{\sqrt{11}}, \frac{\sqrt{2}}{2\sqrt{11}}, \frac{\sqrt{2}}{2\sqrt{11}}\right)\right\}$

31. $\left\{x, \frac{x^3}{\sqrt{6}} + \frac{x^2}{\sqrt{6}} + \frac{2}{\sqrt{6}}, \frac{x^3}{\sqrt{30}} - \frac{5x^2}{\sqrt{30}} + \frac{2}{\sqrt{30}}\right\}$ **33.** (a) $\{(0, 1, 0), (2, 0, -1)\}$

(b) $(1, 3, -2) = -\frac{3}{5}(1, 0, 2) + \left[\frac{4}{5}(2, 0, -1) + 3(0, 1, 0)\right]$ (c) $\frac{\sqrt{305}}{25}$

35. (a) $\{(-2, 0, 1, 0), (0, -1, 0, 1)\}$ (b) $(4, -1, 3, 3) = (2, 1, 4, 1) + (2, -2, -1, 2)$ (c) $\frac{9}{\sqrt{5}}$

37. $\{(4, 1, 8)\}$, $(1, 2, -1) = \frac{1}{23}(55, 54, 41) + \frac{1}{23}(-32, -8, -64)$

7.4 The Spectral Theorem, page 324

13. $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$ **15.** $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 3 & 0 & 2 \end{bmatrix}$ **17.** $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$