This book is concerned with ways to organize and analyze complex systems. From the time of the ancient Greeks until the middle of the twentieth century, scientists concentrated on problems involving a small number of variables. The calculus of Newton and Leibnitz dealt with functions of a single variable, later generalized to several variables. Although the functions studied in calculus can display very complex behavior, the amount of input data is usually quite small.

Today, scientists face problems involving large amounts of data. Consider the following examples of modern complex systems.

1. A mathematical model of the U.S. economy that considers the interactive effects of supplies and demands of various goods. The model may involve thousands of variables and equations.

2. The task of routing long-distance telephone calls. Every second, thousands of calls must be instantly routed from various origins to destinations through many intermediate switching stations. The system doing the routing procedure must look for circuitous indirect routes when more direct pathways are saturated.

3. Statistical studies of factors implicated in the spread of some new disease. Hundreds of causative agents must be analyzed for interactive effects; whereas neither effect A alone nor effect B alone may make a person susceptible to the illness, effects A and B together make a person very susceptible.
4. A mathematical model simulating the airflow around a jet aircraft whose design requires thousands of parameters to describe.

Although this book does not contain 100-, much less 1000-variable problems (5 variables are usually sufficient for realistic examples), it is concerned with mathematical techniques that can readily be applied to such large problems.

In this opening chapter we present some basic models that are used throughout the book to illustrate concepts of matrix theory and associated computations. In Chapter 2 we develop the basic tools of matrix algebra. In Chapter 3 we present various ways to solve a system of linear equations. In Chapter 4 we use matrix algebra to analyze a collection of models in greater depth. In Chapter 5 we discuss the theory of solutions to systems of equations.

Intuitively, the more information we put into a model, the more accurate should be the analysis obtained. The problem is, how do we handle all this information? How do we construct sensible models using hundreds of inputs when we do not really know the underlying mechanism by which the input variables affect the model? In the "old" days when scientists studied simpler systems, very accurate mathematical models were obtained, say, to describe a spherical body's rate of fall based on three critical parameters: the time elapsed since the body was dropped, the body's density, and the density of air. When hundreds of interdependent variables are involved, there is little chance of obtaining a precise mathematical model.

If we do not understand well the system we are modeling, then the structure of our mathematical model should be simple. But the model must still be useful—tell us things about the system that we could not otherwise easily find out. We shall see as we work through this book, that a linear mathematical model is often the best choice. Before defining a linear model, let us state what we mean by a mathematical model in general.

A mathematical model is a mathematical formulation of some class of real-world problems. The formulation may be a set of equations, it may be the minimization of some function, or it may involve integrals. The model may embody various constraints on its variables or it may be a combination of other mathematical models. Part of the modeling process involves input values that vary from one instance of the problem to another. These values are coefficients and constants for equations in the model.

Let us consider five simple mathematical models. The first is a physics model derived with calculus. The other four are standard high school algebra problems.

Example 1. Falling Objects

In physics, the height $H$ of an object dropped off a building is modeled by the formula

$$H = -16T^2 + H_0$$

(1)

where $T$ is time (in seconds) elapsed since the object was dropped,
and $H_0$ is the height (in feet) of the building. (The numbers 16 and $H_0$ are the input values of the model.) One can derive this formula with calculus (if drag from air resistance is ignored).

Suppose that the building is 100 feet tall. Figure 1.1 gives the graph of the equation $H = -16T^2 + 100$. When 2 seconds has elapsed, the object’s current height can be computed to be

$$H = -16(2)^2 + 100 = -64 + 100 = 36$$

To determine the time until the object hits the ground, we set $H = 0$ and solve for $T$:

$$0 = -16T^2 + 100 \rightarrow 16T^2 = 100$$

$$\rightarrow T^2 = \frac{100}{16} = \frac{25}{4}$$

$$\rightarrow T = \sqrt{\frac{25}{4}} = \frac{5}{2}$$

Example 2. Elementary Algebra:
A Problem of Relative Ages

Consider the following word problem. Michael is three times the age of his sister, but in 6 years he will be only twice his sister’s age. How old is Michael now?

We want to model the information in the word problem with algebraic equations. If $M$ is Michael’s current age and $S$ is his sister’s age, we have

$$M = 3S \tag{2}$$

$$M + 6 = 2(S + 6) \tag{3}$$

This model is simply an algebraic restatement of information given in words. We were told that certain quantities were equal; for example, Michael’s current age equals three times his sister’s current age, and we expressed these equalities in symbolic (algebraic) form.
There are two variables in this problem, but the relation between the two variables \( M \) and \( S \) is so simple—one is triple the other—that one can easily rewrite the second equation (3) in terms of one of the variables. If we write (3) in terms of \( S \) by substituting \( 3S \) for \( M \), we obtain

\[
(3S) + 6 = 2(S + 6)
\]

Solving equation (3'), we obtain \( S = 6 \), so \( M = 3S = 18 \).

We now consider a slightly more complicated age problem.

**Example 3. Another Problem of Relative Ages**

Alice is currently twice as old as her brother Bill. If twice the sum of their current ages is equal to the product of their ages 4 years ago, how old is Bill?

If \( A \) represents Alice’s current age and \( B \) Bill’s current age, the given information can be modeled by the system of two equations

\[
\begin{align*}
A &= 2B \\
2(A + B) &= (A - 4)(B - 4)
\end{align*}
\]

Expanding and simplifying the second equation, we obtain

\[
A = 2B \\
AB - 6A - 6B + 16 &= 0
\]

Note that this model involves a term of “degree 2,” the product \( AB \) in the second equation.

Substituting \( 2B \) for \( A \) in the second equation yields

\[
(2B)B - 6(2B) - 6B + 16 = 0 \quad \text{or} \quad B^2 - 9B + 8 = 0
\]

Factoring the right equation in (6) yields

\[
(B - 1)(B - 8) = 0
\]

and thus \( B = 1 \) or \( B = 8 \).

The solution \( B = 1 \) is not possible in terms of the problem statement (if Bill is currently 1 year old, what was he 4 years ago?). Thus the answer is: Bill is 8 (and Alice is 16).

Although both \( (B = 1, A = 2) \) and \( (B = 8, A = 16) \) satisfy algebraic system (4), only the second pair of values makes sense in the original real-world problem. This illustrates an important aspect of modeling that is frequently assumed in this book: namely, interpreting a mathematical solution
in terms of the original real-world problem to see that the solution makes sense. Figure 1.2 provides a picture of the major steps in the modeling process.

**Example 4. Speed of a Canoe**

Consider another classic word problem. When Mary paddles a canoe up a river (against the river's flow), the canoe goes 3 miles per hour, and when she paddles downstream the canoe goes 11 miles per hour. How fast would the canoe be going if she were paddling on a still lake?

If $C$ is the canoe’s speed in still water and $R$ is the speed of the river, then algebra books model the problem with the two equations

$$C + R = 11$$
$$C - R = 3$$  \hspace{1cm} (8)$$

This model expresses the upstream and downstream speeds as functions of $C$ and $R$ by using the intuitive physical principle that the net upstream speed is the canoe speed minus the river speed, and that the net downstream speed is the canoe speed plus the river speed.

Figure 1.3 has a graph of the two equations in (8). We see that they intersect at the point $C = 7$, $R = 4$. This solution can also be obtained by solving (8) algebraically.

**Figure 1.3** Graphical solution of canoe problem.
Examples 2 and 4 are called **linear models** because the mathematical equations in the model are linear, that is, equations of lines. Examples 1 and 3 are nonlinear models because they contain nonlinear equations. A linear equation involves sums and differences of variables or multiples of variables, such as

\[
y = 2x - 6,
\]

\[
x = \frac{y}{2} + 3
\]

\[
4x - 2y = 12
\]

Note that the three preceding equations are all equivalent.

The left sides in (8) and (9b) are called **linear combinations** of variables: sums and differences of variables or of multiples of variables. A linear combination is the simplest type of expression of variables that one can build.

The widespread use of linear models results primarily from the ease of computing linear expressions as well as the existence of a powerful theory for analyzing linear models. Even when a problem is nonlinear, a linear model will often be used as an approximation. For example, for small values of \(x\), \(\sin x\) is often approximated by \(x\). However, the reader should not expect that every solution has a satisfactory linear model, as the next example shows.

**Example 5. Speed of Canoe with Sail**

Suppose that the canoe has a sail mounted at its front. The canoe is on a lake with a wind of \(W\) miles per hour. We assume that downwind speed (moving with the wind) is again \(C + W\), the sum of the canoe’s speed (paddled by Mary) plus the wind’s speed. However, boats with sails can move upwind by an aerodynamic principle, the same principle that holds up airplanes. So we try a linear combination of the \(C\) and \(W\) of the form \(C + kW\), where \(k\) is a constant to be determined. If \(U\) is the upwind speed and \(D\) the downwind speed, we obtain the equations

\[
\text{Downwind:} \quad C + W = D
\]

\[
\text{Upwind:} \quad C + kW = U
\]

Let us consider a numerical example. Suppose that our downwind speed \(D\) is 7 and our upwind speed \(U\) is 5. We try our model—the equations in (10)—with \(k\) values of \(k = .75\) and \(k = 1\).

\[
k = .75: \quad \begin{align*} C + W &= 7 \\
C + .75W &= 5 \end{align*}
\]

\[
k = 1: \quad \begin{align*} C + W &= 7 \\
C + W &= 5 \end{align*}
\]
Figure 1.4 Equations for canoe with sail.

Obviously, the two equations in (12) can never be satisfied simultaneously. In (11) with $k = .75$, there are also difficulties. Subtracting the second equation in (11) from the first, we obtain $.25W = 2$ or $W = 8$. Since $C + W = 7$, then $C = 7 - W = 7 - 8 = -1$. A negative canoe speed is impossible (we assume that Mary is not cheating by paddling backwards).

A good way to see what is happening is with a graph. In Figure 1.4, we have plotted the lines in (11). Because these two lines do not meet inside the positive quadrant, there is no feasible solution. If we change our estimate for the constant $k$ slightly from $.75$ to $.5$, then we obtain a solution $W = 6$, $C = 1$, which might be close to the true value of $W$ and $C$.

Instead of assigning specific values to $k$, $U$ and $D$, let us solve the two equations of (10) for $C$ and $W$ in terms of these general parameters. If we subtract the upwind equation from the downwind equation, we have

$$
C + W = D \\
C + kW = U \\
(1 - k)W = D - U \\
W = \frac{D - U}{1 - k}
$$

If we substitute the formula for $W$ found in (13) into the downwind equation, we obtain

$$
C + \frac{D - U}{1 - k} = D \\
$$

or

$$
C = D - \frac{D - U}{1 - k} = \frac{D - U}{1 - k} = \frac{(1 - k)D - (D - U)}{1 - k} = \frac{U - kD}{1 - k}
$$
Observe that whereas originally $D$ and $U$ were expressed as linear combinations of $C$ and $W$, we have now expressed $C$ and $W$ as linear combinations of $D$ and $U$. However, the critical factor here is $k$. Recall that the parameter $k$ is supposed to allow us to express the upwind speed with a sail as a linear combination of $C$ and $W$. In (13) and (14), $C$ and $W$ depend on $k$ in a nonlinear fashion: When $k$ approaches 1, the denominator $1 - k$ in (13) and (14) approaches 0, so the values of $C$ and $W$ blow up.

This sensitivity to small changes in $k$ near 1 makes this model inherently poor. More generally, it appears that a linear combination such as $C + kW$ of $C$ and $W$ is unable to model properly the upwind speed of a canoe with sail.

Example 5 illustrates the possible inadequacy of a linear model. It also points out that one must always be careful about the accuracy of coefficients in linear equations, because the solution depends on them in a nonlinear fashion. A system of equations is called **ill-conditioned** if a large change in the answer can be produced by a small error in the value of a coefficient (or by a roundoff error in computation, such as writing $\frac{1}{3}$ as .33). When $k$ was near 1, the system of equations in the canoe-with-sail model was very ill-conditioned. In Section 3.5 we learn how to calculate the condition number of a system of equations, which tells how poorly conditioned a system is.

**Section 1.1 Exercises**

**Summary of Exercises**

These exercises examine the five models presented in this section and ask the reader to make mathematical models of similar problems. Exercises 1–3 are based on Example 1; Exercises 4–8 on Example 2; Exercises 9–15 on Example 4; Exercises 16–19 on Example 5. All the exercises require only first-year high school algebra.

1. In Example 1 about a falling object (dropped from the top of a 100-foot building), what height will the object have after 1 second?

2. If the object in Example 1 were dropped from a 400-foot building, how high would it be after 2 seconds? When would it hit the ground? What is the relation between the time of impact in this problem and the time of impact in the original 100-foot problem in Example 1? Guess the time of impact if the object were dropped from the top of a 1600-foot building, and verify that this guess is right.

3. If the object in Example 1 were dropped from a building of height $H_0$, solve equation (1) for the time when the object hits the ground (your answer will involve $H_0$ and the constant 16). Make a graph plotting the time when the object hits the ground as a function of building height.

4. Suppose that John is twice Mary’s age but 4 years ago he was three times her age. Express this information as a pair of linear equations
involving \( J \) (John’s current age) and \( M \) (Mary’s current age). Solve for \( J \) and \( M \).

5. Store \( A \) expects to sell three times as many books as store \( B \) this month. But next month store \( B \)'s sales are scheduled to double while store \( A \)'s sales stay the same. Over the 2-month period, the two stores together are expected to sell a total of 4500 books. How many books would each store sell this month?

6. Suppose that Mary and Nancy are sisters and the sum of their ages equals their older brother Bill’s age, that Nancy is 4 years older than Mary, and that Bill is 6 years older than Nancy. Express this information in three linear equations in \( B \) (Bill’s age), \( M \) (Mary’s age), and \( N \) (Nancy’s age). Express \( B \) and \( N \) in terms of \( M \), and then solve for the ages of the three children.

7. A rectangle is twice as high as it is wide. The sum of the height and width of this rectangle is equal to one-half the area of the rectangle. Find the height and width.

8. A company has a budget of $280,000 for computing equipment. Three types of equipment are available: microcomputers at $2000 each, terminals at $500 each, and word processors at $5000 each. There should be five times as many terminals as microcomputers, and two times as many microcomputers as word processors. How many machines of each type should be purchased?

9. Suppose that the canoe in Example 4 went 5 miles per hour upstream and 9 miles per hour downstream. Solve the resulting system of linear equations algebraically to determine the speed of the canoe (in still water) and the speed of the river.

10. Suppose that the canoe in Example 4 goes \( U \) miles per hour upstream and \( D \) miles per hour downstream. Find general formulas in terms of \( U \) and \( D \) for the speed of the canoe (in still water) and the speed of the river.

11. A company has $36,000 to hire a mathematician and his or her secretary. Out of respect for the mathematician’s training, the mathematician will be paid $8000 more than the secretary. How much will each be paid?

12. Cook \( A \) cooks 2 steaks and 6 hamburgers in half an hour. Cook \( B \) cooks 4 steaks and 3 hamburgers in half an hour. If there is a demand for 16 steaks and 21 hamburgers, how many half-hour periods should cook \( A \) work and how many half-hour periods should cook \( B \) work to fill this demand?

13. We have two oil refineries. Refinery \( A \) produces 20 gallons of heating
oil and 8 gallons of diesel oil out of each barrel of petroleum it refines. Refinery B produces 6 gallons of heating oil and 15 gallons of diesel oil out of each barrel of petroleum it refines. There is a demand of 500 gallons of heating oil and 750 gallons of diesel oil. How many barrels of petroleum should be refined at each refinery to equal this demand?

14. The sum of John’s weight and Sally’s weight is 20 pounds more than four times the difference between their two weights (John is the heavier). Twice Sally’s weight is 40 pounds more than John’s weight. Write down these two facts in two linear equations. Simplify the first equation and solve the two equations to find the weights of John and Sally.

15. Two ferries travel across a lake 50 miles wide. One ferry goes 5 miles per hour slower than the other. If the slower ferry leaves 1 hour before the faster and arrives at the opposite shore at the same time as the other ferry, what is the speed of the slower ferry?

16. In Example 5, re-solve the downwind—upwind equations of (10) when \( k = .9 \) with \( D = 7 \) and \( U = 5 \). Solve again with \( k = .9 \) but now \( D = 6.5 \) and \( U = 5 \). Does this small change in \( D \) result in an equally small change in \( C \)?

17. In Example 5, suppose that \( D = 7 \) and \( U = 5 \), as in equations (10) and (11). Then formula (14) for \( C \) becomes

\[
C = \frac{5 - 7k}{1 - k}
\]

Plot \( C \) as a function of \( k \) in the interval \( 0 \leq k \leq 1 \). Is \( C \) equal to +infinity or equal to -infinity when \( k = 1 \)? Explain your answer.

18. In the downwind—upwind system of equations in Example 5, suppose that Mary were not paddling, so that \( C = 0 \). If \( k = .75 \) and \( U \) (upwind speed) = 6, what is \( W \) and what is \( D \)? In this case, what must be the relation between \( U \) and \( D \)? (That is, write \( U \) as a function of \( D \).)

19. In the system of equations for our canoe-with-sail model,

\[
\begin{align*}
C + kW &= 8 \\
C + W &= 12
\end{align*}
\]

pick the unknown \( k \) so that \( C = 0 \) when this system is solved. What is \( W \)?

20. A refinery produces 8 gallons of gasoline and 6 gallons of heating oil from each barrel of petroleum it refines. There is a demand for 400 gallons of gasoline and 200 gallons of heating oil. Can you set up a linear equation to determine how many barrels of petroleum should be
refined? Explain your answer and tell how many barrels you would refine if you were the manager of the refinery.

21. Suppose we estimate that a child’s IQ is the average of the parents’ IQs and that the income the child has when grown to the age of 40 is 200 times the father’s IQ plus 100 times the mother’s IQ. If the child has an IQ of 120 and earns $48,000 at age 40, what are the parents’ IQs? Does this model give a reasonable answer?

22. Suppose that factory A produces 12 tables and 6 chairs an hour while factory B produces 8 tables and 4 chairs an hour. How many hours should each factory work to produce 48 tables and 24 chairs? How many different solutions are there to this problem?

23. We estimate that Jack can do 3 chemistry problems and 6 math problems in an hour, while Paula can do 4 chemistry problems and 7 math problems in an hour. There are 11 chemistry problems and 17 math problems. Set up and solve a system of two linear equations to determine how long Jack and Paula should work to do these problems. What is the matter with the solution? Propose a solution that makes more sense.

Section 1.2 Systems of Linear Equations

In Section 1.1 we gave a quadratic formula \( H = -16T^2 + H_0 \) to model mathematically the height of an object dropped from a building \( H_0 \) feet tall. From this formula we could determine how long it took for an object to hit the ground (see Example 1 of Section 1.1). The model involved a single nonlinear equation with one variable to be determined. Nonlinear equations in one variable are not difficult to derive and solve.

When many variables need to be determined, then almost surely the mathematical model will be a system of linear equations. There are four basic reasons for using linear models.

1. There is a rich theory for analyzing and solving systems of linear equations. There is limited theory and no general solution techniques for systems of nonlinear equations.
2. Systems of nonlinear equations involving several variables exhibit very complex behavior and we rarely understand real-world phenomena well enough to use such complex models.
3. Small changes in coefficients in nonlinear systems can cause huge changes in the behavior of the systems, yet precise values for these coefficients are rarely known.
4. All nonlinear phenomena are approximately linear over small intervals; that is, a complicated curve can be approximated by a collection of many short line segments (see Figure 1.5).
For these reasons we have the following general principle that is true for numerical problems in all fields: physics, statistics, economics, for instance.

**General Mathematical Principle for Multivariable Problems.** Any problem involving several unknowns is normally solved by recasting the problem as a system of linear equations.

The two canoe problems, with and without a sail, in Section 1.1 were simple examples of systems of linear equations. Here are two more examples. These examples, together with those in Section 1.3, are used dozens of times throughout the book to motivate and illustrate theory and numerical methods. It is *very important* for the reader to gain familiarity with these examples through working some of the numerical exercises at the end of each section.

---

**Example 1. Oil Refinery Model**

A company runs three oil refineries. Each refinery produces three petroleum-based products: heating oil, diesel oil, and gasoline. Suppose that from 1 barrel of petroleum, the first refinery produces 20 gallons of heating oil, 10 gallons of diesel oil, and 5 gallons of gasoline. The second and third refineries produce different amounts of these three products as described in the following table.

<table>
<thead>
<tr>
<th>Product</th>
<th>Refinery 1</th>
<th>Refinery 2</th>
<th>Refinery 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heating oil:</td>
<td>20</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Diesel oil:</td>
<td>10</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>Gasoline:</td>
<td>5</td>
<td>5</td>
<td>12</td>
</tr>
</tbody>
</table>

(1)
Let \( x_i \) be the number of barrels of petroleum used by the \( i \)th refinery. Then the total amount of each product produced by the refineries is given by the linear expressions

\[
\begin{align*}
\text{Heating oil:} & \quad 20x_1 + 4x_2 + 4x_3 \\
\text{Diesel oil:} & \quad 10x_1 + 14x_2 + 5x_3 \\
\text{Gasoline:} & \quad 5x_1 + 5x_2 + 12x_3
\end{align*}
\]

(2)

Suppose that the demand is 500 units of heating oil, 850 units of diesel oil, and 1000 units of gasoline. What values \( x_1, x_2, x_3 \) are needed to produce these amounts? We require the \( x_i \) to satisfy the following system of linear equations.

\[
\begin{align*}
20x_1 + 4x_2 + 4x_3 &= 500 \\
10x_1 + 14x_2 + 5x_3 &= 850 \\
5x_1 + 5x_2 + 12x_3 &= 1000
\end{align*}
\]

(3)

Later in this book we shall learn several ways to solve systems of linear equations. For now, let us use the tried-and-true method of trial and error. As an initial guess, try \( x_1 = 25, x_2 = 25, x_3 = 25 \). Using these values in (3), we get

\[
\begin{align*}
20(25) + 4(25) + 4(25) &= 700 \\
10(25) + 14(25) + 5(25) &= 725 \\
\end{align*}
\]

(4)

(It is helpful to have a programmable calculator or microcomputer for calculations with systems of linear equations.)

We need to alter the \( x_i \) values to make the first expression (heating oil) smaller and the last expression (gasoline) larger. Since \( x_1 \) makes the largest contribution to the first expression and \( x_3 \) makes the largest contribution to the last expression, we decrease \( x_1 \) and increase \( x_3 \). Suppose that we try \( x_1 = 10, x_2 = 25, x_3 = 70 \).

\[
\begin{align*}
20(10) + 4(25) + 4(70) &= 580 \\
10(10) + 14(25) + 5(70) &= 800 \\
5(10) + 5(25) + 12(70) &= 1015
\end{align*}
\]

(5)

Although these \( x \) values are much better than the original ones, let us try to do better. We need to decrease the first expression and increase the second expression (without changing the third expression). To decrease the first expression, we should decrease \( x_1 \). Similarly, to increase the second expression, we increase \( x_2 \). Trying the following values, we obtain the result that production levels
Ch. 1 Introductory Models

\[ x_1 = 5, \quad x_2 = 35, \quad x_3 = 70 \]

yield

heating oil = 520, \quad diesel oil = 890, \quad gasoline = 1040

This is getting quite close to our production goals. The overproduction of 20 to 40 gallons in each product might be a reasonable safety margin that is actually desirable.

To get closer, we should decrease \( x_2 \) and \( x_3 \) a little.

\[ x_1 = 5, \quad x_2 = 33, \quad x_3 = 68 \]

yield

heating oil = 504, \quad diesel oil = 852, \quad gasoline = 1006

This is an excellent fit. We have been a bit lucky. To do better, we would probably have to use fractional values. (In the Exercises the reader may need more tries to get this close.)

We next consider a slightly more complicated supply–demand model. This model has the balancing advantage that trial-and-error calculations to estimate a solution are easier. The reader is warned that it takes a little while to get a feel for all the numbers in this model.

---

**Example 2. A Model of General Economic Supply–Demand**

We present a linear model due to W. Leontief, a Nobel Prize–winning economist. The model seeks to balance supply and demand throughout a whole economy. For each industry, there will be one supply–demand equation. In practical applications, Leontief economic models can have hundreds or thousands of specific industries. We consider an example with four industries.

The left-hand side of each equation is the supply, the amount produced by the \( i \)th industry. Call this quantity \( x_i \); it is measured in dollars. On the right-hand side, we have the demand for the product of the \( i \)th industry. There are two parts to the demand. The first part is demand for the output by other industries (to create other products requires some of this product as input). The second part is consumer demand for the product.

For a concrete instance, let us consider an economy of four general industries: energy, construction, transportation, and steel. Suppose that the supply–demand equations are
Sec. 1.2 Systems of Linear Equations

Industrial Demands

<table>
<thead>
<tr>
<th>Supply</th>
<th>Energy</th>
<th>Construct.</th>
<th>Transport.</th>
<th>Steel</th>
<th>Consumer Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy:</td>
<td>$x_1 = .4x_1 + .2x_2 + .2x_3 + .2x_4 + 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Construct.:</td>
<td>$x_2 = .3x_1 + .3x_2 + .2x_3 + .1x_4 + 50$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transport.:</td>
<td>$x_3 = .1x_1 + .1x_2 + .2x_4 + 100$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Steel:</td>
<td>$x_4 = .1x_2 + .1x_3 + 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The first equation, for energy, has the supply of energy $x_1$ on the left. The terms on the right of this equation are the various demands that this supply must meet. The first term on the right, $.4x_1$, is the input of energy required to produce our $x_1$ dollars of energy (.4 units of energy input for one unit of energy output). Also, the second term of $.2x_2$ is the input of energy needed to make $x_2$ dollars of construction. Similarly, terms $.2x_3$ and $.2x_4$ are energy inputs required for transportation and steel production. The final term of 100 is the fixed consumer demand.

Each column gives the set of input demands of an industry. For example, the third column tells us that to produce the $x_3$ dollars of transportation requires as input $.2x_3$ dollars of energy, $.2x_3$ dollars of construction, and $.1x_3$ dollars of steel. In the previous refinery model, the demand for each product was a single constant quantity. In the Leontief model, there are many unknown demands that each industry’s output must satisfy. There is an ultimate consumer demand for each output, but to meet this demand, industries generate input demands on each other. Thus the demands are highly interrelated: Demand for energy depends on the production levels of other industries, and these production levels depend in turn on the demand for their outputs by other industries, and so on.

When the levels of industrial output satisfy these supply–demand equations, economists say that the economy is in equilibrium.

As in the refinery model, let us try to solve this system of equations by trial-and-error. As a first guess, let us set the production levels at twice the consumer demand (the doubling tries to account for the interindustry demands). So $x_1 = 200$, $x_2 = 100$, $x_3 = 200$, and $x_4 = 0$; these are our supplies. Given these production levels, we can compute the demands from (6).

<table>
<thead>
<tr>
<th>Supply</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy:</td>
<td>$.4(200) + .2(100) + .2(200) + .2(0) + 100 = 240</td>
</tr>
<tr>
<td>Construct.:</td>
<td>$.3(200) + .3(100) + .2(200) + .1(0) + 50 = 180</td>
</tr>
<tr>
<td>Transport.:</td>
<td>$.1(200) + $.1(100) + .2(0) + 100 = 130</td>
</tr>
<tr>
<td>Steel:</td>
<td>$.1(100) + $.1(200) = 30</td>
</tr>
</tbody>
</table>

(7)
For our next approximations, let us try supply levels halfway between the supply and demand values in (7). That is, \( x_1 = \frac{1}{2}(200 + 240) = 220 \), and similarly, \( x_2 = 140 \), \( x_3 = 165 \), and \( x_4 = 15 \).

<table>
<thead>
<tr>
<th>Supply</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy: 220</td>
<td>( .4(220) + .2(140) + .2(165) + .2(15) + 100 = 252 )</td>
</tr>
<tr>
<td>Construct.: 140</td>
<td>( .3(220) + .3(140) + .2(165) + .1(15) + 50 = 192.5 )</td>
</tr>
<tr>
<td>Transport.: 165</td>
<td>( .1(220) + .1(140) + .2(15) + 100 = 139 )</td>
</tr>
<tr>
<td>Steel: 15</td>
<td>( + .1(140) + .1(165) = 30.5 )</td>
</tr>
</tbody>
</table>

The second approximation is only moderately better. The interaction effects between different industries are hard to predict. Adjusting production levels was much easier in the refinery problem, where the demand for each product was constant.

Let us stop trying to be clever and just use the simple-minded approach of setting production levels (i.e., supply levels) equal to the previous demand levels. So from (8), we try

<table>
<thead>
<tr>
<th>Supply</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy: 252</td>
<td>( .4(252) + .2(192) + .2(139) + .2(30) + 100 = 273 )</td>
</tr>
<tr>
<td>Construct.: 192</td>
<td>( .3(252) + .3(192) + .2(139) + .1(30) + 50 = 214 )</td>
</tr>
<tr>
<td>Transport.: 139</td>
<td>( .1(252) + .1(192) + .2(30) + 100 = 150 )</td>
</tr>
<tr>
<td>Steel: 30</td>
<td>( + .1(192) + .1(139) = 33 )</td>
</tr>
</tbody>
</table>

The demand values here have been rounded to whole numbers. The supplies and demands are getting a little closer together in (9). Repeating the process of setting the new supply levels equal to the previous demand levels (i.e., the demands on the right side in (9)) yields

<table>
<thead>
<tr>
<th>Supply</th>
<th>Demand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy: 273</td>
<td>( .4(273) + .2(214) + .2(150) + .2(33) + 100 = 289 )</td>
</tr>
<tr>
<td>Construct.: 214</td>
<td>( .3(273) + .3(214) + .2(150) + .1(33) + 50 = 229 )</td>
</tr>
<tr>
<td>Transport.: 150</td>
<td>( .1(273) + .1(214) + .2(33) + 100 = 155 )</td>
</tr>
<tr>
<td>Steel: 33</td>
<td>( + .1(214) + .1(150) = 36 )</td>
</tr>
</tbody>
</table>

Repeating this process again, we have
Observe that in successive rounds (9), (10), (11), supplies are rising. This is because as we produce more, we need more input which requires us to produce still more, and so on. It may be that this iteration will go on forever, and no equilibrium exists. On the other hand, the gap between supplies and demands is decreasing.

Leontief proposed a constraint on the input costs that we shall show (in Section 3.4) guarantees that an equilibrium exists. The constraint is

\textit{Input Constraint.} Every industry is profitable: Every industry must require less than $1$ of inputs to produce $1$ of output.

In mathematical terms, this means that the sum of the coefficients in each column must be less than 1. Our data in (6) satisfy this constraint, so an equilibrium does exist for this four-industry economy. Moreover, the iteration process of repeatedly setting production levels equal to the previous demands will converge to this equilibrium. The reader should check that the following numbers are equilibrium values (rounded to the nearest integer).

Equilibrium: energy = 325, construction = 265, transportation = 168, steel = 43

Note that any system of linear equations can be rewritten in the form of supply–demand equations with $x_i$ appearing alone on the left side of the $i$th equation, as in the Leontief supply–demand model (6). It is standard practice to solve large systems of linear equations by some sort of iterative method. The nature of the supply–demand equations suggested the iterative scheme we used here, letting the demands from one round be the production levels of the next round.

\section*{Section 1.2 Exercises}

\textbf{Summary of Exercises}

Exercises 1–5 are based on the refinery model. Exercises 6–8 are other problems involving a system of three linear equations in three unknowns. Exercises 9–12 are based on the Leontief economic model. Exercises 13 and 14 involve converting the refinery problem into a system of Leontief-type equations.
Exercises 1–3 refer to the refinery model in Example 1.

1. Suppose that refinery 1 processes 15 barrels of petroleum, refinery 2 processes 20 barrels, and refinery 3 processes 60 barrels. With this production schedule, for which product does production deviate the most from the set of demands 500, 850, 1000?

2. Suppose that the demand for heating oil grows to 800 gallons, while other demands stay the same. Find production levels of the three refineries to meet approximately this new set of demands (by "approximately" we mean with no product off by more than 30 gallons).

3. Suppose that refinery 3 is improved so that each barrel of petroleum yields 8 gallons of heating oil, 10 gallons of diesel oil, and 20 gallons of gasoline. Find production levels of the three refineries to meet approximately the demands (by "approximately" we mean with no product off by more than 30 gallons).

4. Consider the following refinery model. There are three refineries 1, 2, and 3 and from each barrel of crude petroleum, the different refineries produce the following amounts (measured in gallons) of heating oil, diesel oil, and gasoline.

<table>
<thead>
<tr>
<th></th>
<th>Refinery 1</th>
<th>Refinery 2</th>
<th>Refinery 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heating oil</td>
<td>6</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Diesel oil</td>
<td>4</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Gasoline</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
</tbody>
</table>

Suppose that we have the following demand:

280 gallons of heating oil,
350 gallons of diesel oil, and
350 gallons of gasoline.

(a) Write a system of equations whose solution would determine production levels to yield the desired amounts of heating oil, diesel oil, and gasoline. As in Example 1, let \( x_i \) be the number of barrels processed by the \( i \)th refinery.

(b) Find an approximate solution to this system of equations with no product off by more than 30 gallons from its demand.

5. Repeat the refinery model in Exercise 4 with new demand levels of 500 gallons heating oil, 300 gallons diesel oil, and 600 gallons gasoline. Try to find an approximate solution (within 30 gallons) with this set of demands. Something is going wrong and there is no valid set of pro-
duction levels to attain this set of demands. What is invalid about the solution to this refinery problem?

*Extra Credit:* Try to explain in words why this set of demands is unattained while the demands in Exercise 4 were attainable.

6. The staff dietician at the California Institute of Trigonometry has to make up a meal with 600 calories, 20 grams of protein, and 200 milligrams of vitamin C. There are three food types to choose from: rubbery jello, dried fish sticks, and mystery meat. They have the following nutritional content per ounce.

<table>
<thead>
<tr>
<th></th>
<th>Jello</th>
<th>Fish Sticks</th>
<th>Mystery Meat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Calories</td>
<td>10</td>
<td>50</td>
<td>200</td>
</tr>
<tr>
<td>Protein</td>
<td>1</td>
<td>3</td>
<td>.2</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>30</td>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Make a mathematical model of the dietician’s problem with a system of three linear equations.
(b) Find an approximate solution (accurate to within 10%).

7. A furniture manufacturer makes tables, chairs, and sofas. In one month, the company has available 300 units of wood, 350 units of labor, and 225 units of upholstery. The manufacturer wants a production schedule for the month that uses all of these resources. The different products require the following amounts of the resources.

<table>
<thead>
<tr>
<th></th>
<th>Table</th>
<th>Chair</th>
<th>Sofa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wood</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Labor</td>
<td>3</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Upholstery</td>
<td>2</td>
<td>0</td>
<td>4</td>
</tr>
</tbody>
</table>

(a) Make a mathematical model of this production problem.
(b) Find an approximate solution (accurate to within 10%).

8. A company has a budget of $280,000 for computing equipment. Three types of equipment are available: microcomputers at $2000 a piece, terminals at $500 a piece, and word processors at $5000 a piece. There should be five times as many terminals as microcomputers and two times as many microcomputers as word processors. Set this problem up as a system of three linear equations. Determine approximately how many machines of each type there should be by solving by trial-and-error.
Note: Check your answer by expressing the numbers of terminals and microcomputers in terms of the number of word processors and solving the remaining single equation in one unknown.

Exercises 9–11 are based on the Leontief model in Example 2.

9. If we produced 300 units of energy, 250 units of construction, 160 units of transportation, and 40 units of steel, what would be the largest deviation between supply and demand among the four commodities?

10. Start the iteration procedure followed in (9), (10), and (11) with an initial set of supplies equal to the consumer demands, that is, \( x_1 = 100, x_2 = 50, x_3 = 100, x_4 = 0 \). Compute the right sides of the equations in (6) with this set of \( x_i \)'s and let the resulting numbers be the new values for \( x_1, x_2, x_3, x_4 \); compute the right sides again with these new \( x_i \)'s, and so on. Do this iteration five times. Do the successive sets of \( x_i \)'s appear to be converging toward the equilibrium values given at the end of Example 2?

11. This exercise explores the effect on all industries of changes in one industry. Quadrupling the price of petroleum had a widespread effect on all industrial sectors in the 1970s. But smaller changes in one seemingly unimportant industry can also result in important changes in many other industries.

(a) Change the system of equations in the Leontief model in (6) by decreasing the coefficient of \( x_1 \) (energy) in the construction equation from .3 to .2 (this is the result of new energy efficiencies in construction equipment). We want to know how this change affects our economy. Iterate five times, as in Exercise 10, with this altered system using as starting \( x_i \)'s the equilibrium values for the original model: \( x_1 = 325, x_2 = 265, x_3 = 168, x_4 = 43 \).

(b) Repeat part (a), but now decrease the coefficient of \( x_4 \) (steel) in the transportation equation from .2 to .1.

(c) Repeat part (a), but now increase the coefficient of \( x_2 \) (construction) in the energy equation from .2 to .3.

12. Consider the Leontief system

\[
\begin{align*}
    x_1 &= .4x_1 + .3x_2 + .3x_3 + 100 \\
    x_2 &= .3x_1 + .4x_2 + .3x_3 + 100 \\
    x_3 &= .3x_1 + .3x_2 + .4x_3 + 100
\end{align*}
\]

Here the column sums are 1, violating the Leontief input constraint given in the text. Show that this system cannot have a solution.

Hint: Add the three equations together.

Extra Credit: Try to explain in economic terms why no solution exists.
13. (a) Rewrite the refinery model's equations in (3) to look like Leontief equations as follows. Divide the first equation by 20 (the coefficient of $x_1$ in that equation), divide the second equation by 14, and divide the third equation by 12. Next move the $x_2$ and $x_3$ terms in the new first equation over to the right side, leaving just $x_1$ on the left (the new equation should be $x_1 = -0.2x_2 - 0.2x_3 + 25$); similarly, in the second equation leave just $x_2$ on the left and in the third equation leave just $x_3$ on the left. Note that the Leontief input constraint about column sums is not satisfied by this system.

(b) Use the iteration method introduced in equations (9), (10), and (11) (see also Exercise 10) to get an approximate solution to the refinery problem (do five iterations, starting with the "consumer demands" of 25, 50, 100).

14. (a) Rewrite the refinery model's equations in (3) to look somewhat like Leontief equations as follows. In the first equation, move the $x_2$ and $x_3$ terms to the right side and also move 19 of the 20 units of the $x_1$ term to the right, leaving just $x_1$ on the left (the equation is now $x_1 = -19x_1 - 4x_2 - 4x_3 + 500$). Similarly, in the second equation leave just $x_2$ on the left; move everything else to the right side. In the third equation leave just $x_3$ on the left. Note that Leontief's input constraint about column sums is far from satisfied by this system.

(b) Try using the iteration method introduced in equations (9), (10), and (11) (see also Exercise 10) to get an approximate solution to the refinery problem (do five iterations starting with the guess of $x_1 = 25$, $x_2 = 50$, $x_3 = 100$). Does the iteration process seem to be converging?

Section 1.3 Markov Chains and Dynamic Models

The refinery and economic supply–demand models of Section 1.2 were static in the sense that we solved them once and that was it. There was one set of production levels required, not a sequence of levels that would be needed to describe an economy changing over time. A model that tries to predict the behavior of a system over a period of time is called a dynamic model. In this section we examine two dynamic linear models.

The first dynamic model we consider involves probability. This model, called a Markov chain, will arise over and over in this book, so it is important to understand the model well. The concepts of probability we need for this model are simple and intuitive.

A Markov chain is a probabilistic model that describes the random movement over time of some activity. At each period of time, the activity is in one of several possible states. States might be amounts won in gam-
bling, different weather conditions (e.g., sunny, snowy), or numbers of jobs waiting to be processed by a computer. The model specifies probabilities that tell how the activity changes states from period to period. Consider a simple two-state Markov chain.

*Example 1. Markov Chain for Weather*

Suppose that we have two states of the weather: sunny or cloudy. If it is sunny today, the probability is \( \frac{3}{4} \) that it will be sunny tomorrow, and \( \frac{1}{4} \) that it will be cloudy tomorrow. If it is cloudy today, then the probability is \( \frac{1}{2} \) that it will be sunny tomorrow and \( \frac{1}{2} \) that it will be cloudy tomorrow. It is convenient to display these probabilities in an array.

<table>
<thead>
<tr>
<th>Today</th>
<th>Sunny</th>
<th>Cloudy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tomorrow</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sunny</td>
<td>( \frac{3}{4} )</td>
<td>( \frac{1}{4} )</td>
</tr>
<tr>
<td>Cloudy</td>
<td>( \frac{1}{4} )</td>
<td>( \frac{1}{2} )</td>
</tr>
</tbody>
</table>

The probabilities in this array are called *transition probabilities*, and the array is called a *transition matrix*. The probabilities in each column of the transition matrix must add up to 1. A convenient way to display the information in a Markov chain is with a transition diagram. The diagram for the weather Markov chain is drawn in Figure 1.6. There is a node for each state and arrows between states. Beside the arrow from state \( A \) to state \( B \) we write the transition probability of going from state \( A \) to state \( B \).

The transition probabilities of a Markov chain tell us the chances of being in different states one period later. We need a formula from probability theory to be able to calculate the probabilities of where an activity will be after several periods. As in weather forecasting, *it is predictions many periods into the future that are most interesting*.

To state this probability formula, we need to introduce some notation. Let \( p_1, p_2, \ldots, p_n \) be the probabilities of being in state 1, state 2, \ldots, state \( n \). This set of probabilities is called a *probability distribution*. Let \( a_{ij} \) be the transition probability of going from state \( j \) to state \( i \). In the weather Markov chain, if state 1 is sunny and state 2 is cloudy, the transition probabilities are

\[
\begin{bmatrix}
  a_{11} = \frac{3}{4} & a_{12} = \frac{1}{2} \\
  a_{21} = \frac{1}{4} & a_{22} = \frac{1}{2}
\end{bmatrix}
\]

*Weather*  
*Markov chain*
Formula for Distribution of Next States in Markov Chain. If \( p_1, p_2, \ldots, p_n \) is the current probability distribution of the activity, then the probability distribution \( p'_1, p'_2, \ldots, p'_n \) for the next period is given by

\[
\begin{align*}
p'_1 &= a_{11}p_1 + a_{12}p_2 + a_{13}p_3 + \cdots + a_{1n}p_n \\
p'_2 &= a_{21}p_1 + a_{22}p_2 + a_{23}p_3 + \cdots + a_{2n}p_n \\
p'_3 &= a_{31}p_1 + a_{32}p_2 + a_{33}p_3 + \cdots + a_{3n}p_n \\
&\vdots \quad \vdots \quad \vdots \quad \vdots \\
p'_n &= a_{n1}p_1 + a_{n2}p_2 + a_{n3}p_3 + \cdots + a_{nn}p_n
\end{align*}
\]

(1)

To illustrate this formula with the weather Markov chain, let \( p_1 = \frac{3}{4} \) and \( p_2 = \frac{1}{4} \). Then tomorrow's distribution \( p'_1, p'_2 \) is

\[
\begin{align*}
p'_1 &= a_{11}p_1 + a_{12}p_2 = \frac{3}{4} \cdot \frac{3}{4} + \frac{1}{4} \cdot \frac{1}{4} = \frac{11}{16} \\
p'_2 &= a_{21}p_1 + a_{22}p_2 = \frac{1}{4} \cdot \frac{3}{4} + \frac{3}{4} \cdot \frac{1}{4} = \frac{11}{16}
\end{align*}
\]

(2)

We explain the formulas in (2) intuitively as follows. We can be in state 1 (sunny) tomorrow either because we are in state 1 today and then stay in state 1—this is the probability \( a_{11}p_1 \)—or because we are in state 2 today and then switch to state 1—this is the probability \( a_{12}p_2 \). [To compute the probability of a sequence of two events, such as (i) the probability \( p_2 \) of now being in state 2, and (ii) the probability \( a_{12} \) of switching from state 2 to state 1, we multiply these two probabilities together, to get \( a_{12}p_2 \).]

Let us next consider a larger Markov chain that models the action of a popular video arcade game called Frogger.

Example 2. Frogger Markov Chain

We model the behavior of a frog jumping around on a four-lane highway. The possible states range from 1 = left side of highway, to 6 = right side of highway. See Figure 1.7a. Suppose that the following array gives the transition probabilities that the frog, if now in state \( j \), will be in state \( i \) one minute later. The transition diagram for this Markov chain is given in Figure 1.7b.

<table>
<thead>
<tr>
<th>State in Current Period</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>State in Current Period</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.5</td>
<td>.25</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>in</td>
<td>2</td>
<td>.5</td>
<td>.5</td>
<td>.25</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Next</td>
<td>3</td>
<td>0</td>
<td>.25</td>
<td>.5</td>
<td>.25</td>
<td>0</td>
</tr>
<tr>
<td>Period</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>.25</td>
<td>.5</td>
<td>.25</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.25</td>
<td>.5</td>
<td>.5</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.25</td>
<td>.5</td>
</tr>
</tbody>
</table>
The formulas for next-period probabilities with the frog Markov chain are

\[
\begin{align*}
\pi'_1 &= .50p_1 + .25p_2 \\
\pi'_2 &= .50p_1 + .50p_2 + .25p_3 \\
\pi'_3 &= .25p_2 + .50p_3 + .25p_4 \\
\pi'_4 &= .25p_3 + .50p_4 + .25p_5 \\
\pi'_5 &= .25p_4 + .50p_5 + .50p_6 \\
\pi'_6 &= .25p_5 + .50p_6
\end{align*}
\]

(4)

Suppose that the frog starts in state 1 (left side of highway). Then its probability distribution after 1 minute is given by the probabilities in the first column of (4), since initially \( p_1 = 1 \) and other \( p_i = 0 \):

\[
\begin{align*}
p_1 &= .5, \\
p_2 &= .5, \\
\text{other } p_i &= 0
\end{align*}
\]

(5)

Let us use (4) to compute the probability distribution for the frog after 2 minutes (remember that only \( p_1 \) and \( p_2 \) are nonzero):

\[
\begin{align*}
\pi'_1 &= .50p_1 + .25p_2 = .5 \times .5 + .25 \times .5 = .375 \\
\pi'_2 &= .50p_1 + .50p_2 = .5 \times .5 + .5 \times .5 = .5 \\
\text{other } \pi'_i &= 0
\end{align*}
\]

(6)

Other \( \pi'_i = 0 \). Note that the sum of the probabilities in (6) equals 1, as it should.

We can continue iterating with formula (4) to find the distribution after 3 minutes, after 4 minutes, and so on (it helps to let a computer
Table 1.1

<table>
<thead>
<tr>
<th>Minutes</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>.50</td>
<td>.50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>.375</td>
<td>.50</td>
<td>.125</td>
<td>.375</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>.312</td>
<td>.469</td>
<td>.188</td>
<td>.469</td>
<td>.312</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>.273</td>
<td>.438</td>
<td>.219</td>
<td>.438</td>
<td>.273</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>.246</td>
<td>.410</td>
<td>.234</td>
<td>.410</td>
<td>.246</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>.226</td>
<td>.387</td>
<td>.242</td>
<td>.387</td>
<td>.226</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>.176</td>
<td>.320</td>
<td>.241</td>
<td>.320</td>
<td>.176</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>.144</td>
<td>.272</td>
<td>.226</td>
<td>.272</td>
<td>.144</td>
<td>0</td>
</tr>
<tr>
<td>20</td>
<td>.126</td>
<td>.244</td>
<td>.217</td>
<td>.244</td>
<td>.126</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>.1</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.1</td>
<td>.1</td>
</tr>
<tr>
<td>200</td>
<td>.1</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.1</td>
<td>.1</td>
</tr>
<tr>
<td>1000</td>
<td>.1</td>
<td>.2</td>
<td>.2</td>
<td>.2</td>
<td>.1</td>
<td>.1</td>
</tr>
</tbody>
</table>

do this). Table 1.1 gives these probabilities, assuming that the frog started in state 1 (left side of highway).

Observe that the probabilities converge to the distribution $p_1 = .1, p_2 = .2, p_3 = .2, p_4 = .2, p_5 = .2, p_6 = .1$ (and then stay the same forever). Very interesting!

Would this long-term distribution evolve from any starting distribution? Do all Markov chains exhibit this type of long-term distribution? Can the long-term distribution be computed more simply than iterating the equations in (4) 100 times? (Answer: Definitely yes.)

Markov chains are a very useful type of linear model—linear because formula (1) for next-state probabilities is a system of linear equations. Part of their usefulness is due to results in matrix algebra that provide simple answers to all the questions just posed and many more.

Next we look at a simpler dynamic model, an ecological model that traces the sizes of populations of rabbits and foxes. The simplicity of the model allows us to experiment more, changing the values of the coefficients to exhibit a variety of different long-term trends.

**Example 3. Growth Model for Rabbits and Foxes Model**

Consider the following model for the monthly growth of populations of foxes and rabbits. If $R$ and $F$ are the numbers of rabbits and foxes...
this month, let \( R' \) and \( F' \) be the numbers next month given in the equations

\[
R' = R + bR - eF \\
F' = F - dF + e'R
\]

where \( b \) is the birthrate of rabbits, \( d \) the death rate of foxes, and \( e \) and \( e' \) are eating rates. The term \(-eF\) in the rabbit equation is negative and the term \(+e'R\) in the fox equation is positive because foxes eat rabbits.

The results \( R', F' \) after 1 month can become new values for \( R \) and \( F \) to project the populations 2 months hence, and so on, as we did in the frog Markov chain.

Normally, the (positive) constants would be estimated for us by ecologists. But let us make up some reasonable-sounding values for these constants and see what sort of behavior this model predicts.

Suppose that we try

\[
R' = R + .2R - .3F \\
F' = F - .1F + .1R
\]

and start with \( R = 100, F = 100 \). Then using (8) repeatedly to compute the populations in successive months, we get

<table>
<thead>
<tr>
<th>Month</th>
<th>Rabbits</th>
<th>Foxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>1</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>4</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>6</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>7</td>
<td>105</td>
<td>55</td>
</tr>
</tbody>
</table>

A negative number means that the rabbits became extinct.

If there are no rabbits, then the fox equation in (8) becomes

\[ F' = F - .1F \rightarrow F' = .9F \]

and the foxes will eventually die out, too. This behavior of foxes killing off the rabbits and then starving to death is reasonable.

Let us try new starting values for \( R \) and \( F \) that will allow the rabbits to increase in size. The term \(+.2R -.3F\) in the rabbit equation is the amount the rabbit population changes from this month to the next. For this term to be positive we require \(.2R\) to be more than \(.3F\).

Suppose that we choose \( R = 100, F = 50 \):

<table>
<thead>
<tr>
<th>Month</th>
<th>Rabbits</th>
<th>Foxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>1</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>2</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>3</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>4</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>5</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>6</td>
<td>105</td>
<td>55</td>
</tr>
<tr>
<td>7</td>
<td>105</td>
<td>55</td>
</tr>
</tbody>
</table>
While the rabbits increased initially, so did the foxes (since they fed off the rabbits). After 7 months there were enough foxes so that they were eating rabbits faster than new rabbits were being born and the rabbit population began to decline. After 13 months (when there are fewer rabbits than foxes), the foxes begin to decline, also. Now we are in the same situation as before, in (9).

It appears that the equations in our model (8) make it inevitable that the foxes will grow to a level where they eat rabbits faster than rabbits are born, causing the rabbits to decline to extinction. Then the foxes become extinct, too. The reader is asked in the Exercises to try other values for \( b \), \( d \), \( e \), and \( e' \) in this model and to explore the resulting behavior.

Akin to the stable probability distribution in the frog Markov chain, let us try to determine values of the coefficients in our model that will permit the rabbit and fox populations to stabilize, that is, to remain the same forever. This means that \( R' = R \) and \( F' = F \).

We return to the original general model

\[
\begin{align*}
R' &= R + bR - eF \\
F' &= F - dF + e'R
\end{align*}
\]

When \( R' = R \) and \( F' = F \), we have

\[
\begin{align*}
R &= R + bR - eF \\
F &= F - dF + e'R
\end{align*}
\]

or

\[
\begin{align*}
bR - eF &= 0 \\
e'R - dF &= 0
\end{align*}
\]
Note that in (12), the order of the terms $+e'R$ and $-dF$ in the second equation was reversed. Let us solve this pair of linear equations for $R$ and $F$. Obviously, $R = F = 0$ is a solution. But we want another solution. We use the standard method for eliminating one of the variables in (12): Multiply the first equation by $d$ and the second by $e$ and then subtract the second from the first.

\[
dbR - deF = 0 \\
-(ee'R - edF) = 0 \\
(bd - ee')R = 0
\]

If $bd - ee' = 0$, then $R$ (and $F$) need not be 0. Note that

\[
bd - ee' = 0 \iff bd = ee' \iff \frac{e}{b} = \frac{d}{e'}
\]

Suppose that $e/b = d/e'$. Then one can show that $R$ and $F$ are solutions to (11) if and only if

\[
\frac{R}{F} = \frac{e}{b} \left( = \frac{d}{e'} \right)
\]

For example, the system

\[
\begin{align*}
R' &= R + .1R - .15F \\
F' &= F - .15F + .1R
\end{align*}
\]

has stable values $R = 15, F = 10$ or $R = 6, F = 4$. In fact, any pair $(R, F)$ is stable in (15) if

\[
\text{Stable } R, F \text{ values for (15)}: \quad \frac{R}{F} = \frac{.15}{.1} \quad \text{or} \quad R = \frac{3}{2} F
\]

Further, if we start with values for $R$ and $F$ that are not stable, then over successive months the rabbit and fox populations always move toward one of these stable pairs of values, just like the Markov chain. In this model the "law of nature" is that there should be a $3:2$ ratio of rabbits to foxes. Figure 1.8 shows sample curves along which unstable values move in approaching a stable value. For example, if we start iterating (15) with $R = 50, F = 40$, we have

<table>
<thead>
<tr>
<th>Months</th>
<th>Rabbits</th>
<th>Foxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td>1</td>
<td>49</td>
<td>39</td>
</tr>
<tr>
<td>2</td>
<td>48</td>
<td>38</td>
</tr>
<tr>
<td>3</td>
<td>47</td>
<td>37</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
Figure 1.8 Stable values and trajectories to stable values in rabbit–fox growth model.

10 months: 42 rabbits, 32 foxes

20 months: 37 rabbits, 27 foxes

30 months: 34 rabbits, 24 foxes

40 months: 32.5 rabbits, 22.5 foxes

50 months: 31.5 rabbits, 21.5 foxes

75 months: 30.4 rabbits, 20.4 foxes

100 months: 30.1 rabbits, 20.1 foxes
Clearly, the populations are approaching the stable sizes of 30 rabbits and 20 foxes.

The starting pair (100, 80) goes to (60, 40); (80, 30) goes to (150, 100). If we write the starting values in the form \( R = r, F = r - k \), then the limiting stable values will always turn out to be \( R = 3k, F = 2k \). If the starting values have \( R < F \), the limiting values will both be negative, with rabbits becoming negative (extinct) first. If \( R = F \), both populations approach 0.

Why does this happen? How much of this behavior would occur if we used other values for the constants \( b, d, e, e' \) that satisfied the condition \( e/b = d/e' \)?

We conclude this section by introducing a nonlinear model for rabbit and fox populations. This nonlinear model can simulate a cyclic behavior that occurs frequently in nature.

**Example 4. Nonlinear Model for Rabbits and Foxes**

Let us consider the following nonlinear model for the monthly growth of populations of foxes and rabbits. Again, if \( R \) and \( F \) are the numbers of rabbits and foxes this month, then \( R' \) and \( F' \) are the numbers next month. Our system of equations is

\[
\begin{align*}
R' &= R + bR - eRF \\
F' &= F - dF + e'RF
\end{align*}
\]

We now use the terms \(-eRF\) and \(+e'RF\) for the effect of foxes eating rabbits because the chances of a fox catching a rabbit depend on the abundance of both species.

Let us suppose that \( b = d = .1 \) and \( e = e' = .01 \).

\[
\begin{align*}
R' &= R + .1R - .01RF \\
F' &= F - .1F + .01RF
\end{align*}
\]

If initially we let \( R = 30 \) and \( F = 30 \), then using (19) we obtain the table whose values are plotted in Figure 1.9.

<table>
<thead>
<tr>
<th>Number of Months</th>
<th>Rabbits</th>
<th>Foxes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>32</td>
<td>28</td>
</tr>
<tr>
<td>3</td>
<td>34</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>40</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>58</td>
<td>17</td>
</tr>
<tr>
<td>15</td>
<td>87</td>
<td>15</td>
</tr>
</tbody>
</table>
When we started with a small number of both rabbits and foxes, the fox population declined further for a few months (from the $-dF$ term).
term) while the rabbit population started growing (from the, $+ bR$ term); the $- eRF$ and $+ e'RF$ terms have little effect because the constants $e$ and $e'$ are so small. Soon the fox population grows and the rabbit population stops growing and starts to decline as the foxes eat the rabbits (the terms $- eRF$ and $+ e'RF$ come into effect when $R$ and $F$ are large). As the rabbit population declines, so the fox population soon declines because of less food. We eventually find the sizes of the two populations back at the levels at which we started. The cycle time in this model is about 80 months.

This cycle of sudden growth followed by sudden decline characterizes the behavior of periodic pests, such as gypsy moths. The moths are dormant for many years and then have sudden, major outbreaks when few of their natural enemies are present (for gypsy moths, the enemy is a parasitic wasp). Eventually, the large numbers of the pest stimulate the appearance of its predator. The moths are killed off by the wasps and then the wasps die for lack of food. The dormant period begins again.

Linear models cannot produce this type of behavior.

Note that after one cycle (80 months) the population sizes in (20) are 15 rabbits and 15 foxes, half the starting sizes of 30. This slippage in size is a fault in the model, due to the fact that we rounded time into units of months. If time were measured in days or seconds, the slippage would be less. Exercise 21 describes how to convert this model into units of days or seconds.

**Section 1.3 Exercises**

**Summary of Exercises**
Exercises 1–11 involve Markov chain models. Exercises 12–19 examine the model in Example 3 and similar linear growth models. Exercises 20 and 21 examine the behavior of the nonlinear model in Example 4. Exercises 11, 20, and 21 require computer programs.

1. Using the weather Markov chain in Example 1, simulate the weather over 10 days by flipping a coin to determine the chances of sunny or cloudy weather the next day according to the Markov chain’s transition probabilities. If currently sunny, flip once and a head means sunny the next day and a tail means cloudy the next day. If currently cloudy, flip twice and when either flip is a head it is sunny the next day and when both flips are tails it is cloudy the next day. To start, assume that the previous day was sunny. What fraction of the 10 days was sunny?

2. In the weather Markov chain, starting with the probability distribution $(1, 0)$ (a sunny day), compute and plot (in $p_1$, $p_2$ coordinates) the distribution over five successive days. Repeat the process starting with the probability distribution $(0, 1)$. Can you guess the value of the stable distribution to which your points are converging?
3. In the frog Markov chain, what is the probability distribution in the next period if the current distribution is
   (a) $p_3 = 1$, all other $p_i = 0$?
   (b) $p_2 = .5$, $p_3 = .5$, all other $p_i = 0$?
   (c) $p_2 = .25$, $p_3 = .25$, $p_4 = .5$, all other $p_i = 0$?
   (d) $p_1 = .1$, $p_2 = .2$, $p_3 = .2$, $p_4 = .2$, $p_5 = .2$, $p_6 = .1$?

4. The printing press in a newspaper has the following pattern of breakdowns. If it is working today, tomorrow it has 90% chance of working (and 10% chance of breaking down). If the press is broken today, it has a 60% chance of working tomorrow (and 40% chance by being broken again).
   (a) Make a Markov chain for this problem; give the matrix of transition probabilities and draw the transition diagram.
   (b) If there is a 50–50 chance of the press working today, what are the chances that it is working tomorrow?
   (c) If the press is working today, what are the chances that it is working in 2 days’ time?

5. If the local professional basketball team, the Sneakers, wins today’s game, they have a $\frac{2}{3}$ chance of winning their next game. If they lose this game, they have a $\frac{1}{2}$ chance of winning their next game.
   (a) Make a Markov chain for this problem; give the matrix of transition probabilities and draw the transition diagram.
   (b) If there is a 50–50 chance of the Sneakers winning today’s game, what are the chances that they win their next game?
   (c) If they won today, what are the chances of winning the game after the next?

6. If the stock market went up today, historical data show that it has a 60% chance of going up tomorrow, a 20% chance of staying the same, and a 20% chance of going down. If the market was unchanged today, it has a 20% chance of being unchanged tomorrow, a 40% chance of going up, and a 40% chance of going down. If the market goes down today, it has a 20% chance of going up tomorrow, a 20% chance of being unchanged, and a 60% chance of going down.
   (a) Make a Markov chain for this problem; give the matrix of transition probabilities and the transition diagram.
   (b) If there is a 30% chance that the market goes up today, a 10% chance that it is unchanged, and a 60% chance that it goes down, what is the probability distribution for the market tomorrow?

7. The following model for learning a concept over a set of lessons identifies four states of learning: $I =$ Ignorance, $E =$ Exploratory Thinking, $S =$ Superficial Understanding, and $M =$ Mastery. If now in state $I$, after one lesson you have $\frac{3}{4}$ probability of still being in $I$ and $\frac{1}{4}$ probability of being in $E$. If now in state $E$, you have $\frac{1}{4}$ probability of being in $I$, $\frac{1}{2}$ in $E$, and $\frac{1}{4}$ in $S$. If now in state $S$, you have $\frac{1}{4}$ probability of
being in \( E \), \( \frac{1}{2} \) in \( S \), and \( \frac{1}{4} \) in \( M \). If in \( M \), you always stay in \( M \) (with probability 1).

(a) Make a Markov chain model of this learning model.

(b) If you start in state \( I \), what is your probability distribution after two lessons? After three lessons?

8. (a) Make a Markov chain model for a rat wandering through the following maze if at the end of each period, the rat is equally likely to leave its current room through any of the doorways. The states of the Markov chain are the rooms.

(b) If the rat starts in room 1, what is the probability that it is in room 4 two periods later?

9. Make a Markov chain model of a poker game where the states are the number of dollars a player has. With probability .3 a player wins 1 dollar in a period, with probability .4 a player loses 1 dollar, and with probability .3 a player stays the same. The game ends if the player loses all his or her money or if the player has 6 dollars (when the game ends, the Markov chain stays in its current state forever). The Markov chain should have seven states, corresponding to the seven different amounts of money: 0, 1, 2, 3, 4, 5, or 6 dollars. If you now have $2, what is your probability distribution in the next round? In the round after that?

10. Three tanks \( A, B, C \) are engaged in a battle. Tank \( A \), when it fires, hits its target with hit probability \( \frac{1}{2} \). \( B \) hits its target with hit probability \( \frac{1}{3} \), and \( C \) with hit probability \( \frac{1}{6} \). Initially (in the first period), \( B \) and \( C \) fire at \( A \) and \( A \) fires at \( B \). Once one tank is hit, the remaining tanks aim at each other. The battle ends when there is one or no tank left. Make a Markov chain model of this battle.

**Assistance in Computing Probabilities:** Let the states of the Markov chain be the eight different possible subsets of tanks currently in action: \( ABC, AB, AC, BC, A, B, C, None \). When in states \( A \) or \( B \) or \( C \) or None, the probability of staying in the current state is 1—this simulates the battle being over. One can never get to state \( AB \). (Why?) So one only needs to determine the transition probabilities from states \( ABC, AC, \) and \( BC \). From states \( AC \) and \( BC \), the transition probabilities are products of the probability that each remaining tank hits or misses its target. For example, the probability of going from state \( AC \) to state \( A \) is the product of the probability that \( A \) hits \( C \)—\( \frac{1}{6} \)—times the probability
that C misses A—\( \frac{2}{3} \). So this probability is \( \left( \frac{1}{2} \right) \left( \frac{2}{3} \right) = \frac{1}{3} \). It takes some knowledge of probability to compute the transition probabilities from state \( ABC \). From \( ABC \) there is a \( \frac{5}{6} \) chance of remaining in state \( ABC \) (all tanks miss), a \( \frac{1}{6} \) chance of going to state \( AC \) (A hits B but B and C miss A), a \( \frac{1}{6} \) chance of going to state \( BC \) (at least one of B or C hits A and A misses B), and a \( \frac{1}{6} \) chance of going to state \( C \) (at least one of B or C hits A and A hits B).

11. Use a computer program to follow the Markov chains in the following examples and exercises for 50 periods by iterating the next-period formula (1) as done in Example 2.
   (a) Example 1, starting in state Sunny.
   (b) Example 1, starting in state Cloudy.
   (c) Example 2, starting in state 4.
   (d) Example 2, starting with \( p_1 = p_6 = .5 \), other \( p_i = 0 \).
   (e) Exercise 4, starting in state Broken.
   (f) Exercise 5, starting in state Win.
   (g) Exercise 6, starting in state Market Unchanged.
   (h) Exercise 7, starting in state I.
   (i) Exercise 8, starting in state Room 1.
   (j) Exercise 9, starting in state $2$.
   (k) Exercise 9, starting in state $3$.
   (l) Exercise 9, starting in state $4$.
   (m) Exercise 10, starting in state \( ABC \).

12. For the rabbit–fox model in equations (8), use hand calculations to verify the population sizes for months 1, 2, and 3 given in table (9). To get the sizes after 1 month, set \( R = 100 \), \( F = 100 \) (the starting sizes) and evaluate the right sides of the equations in (8). Next take the values you obtained for \( R' \) and \( F' \) and let these be the new \( R \) and \( F \). Repeat this process three times.

13. For the rabbit–fox model in equations (8), suppose that the initial population sizes are \( R = 50 \), \( F = 50 \).
   (a) Calculate by hand the population sizes after 1 month, after 2 months, and after 3 months.
   (b) Use a computer or calculator to compute the population sizes over 8 months.

14. Consider the following rabbit–fox models and an initial population size of \( R = 100 \), \( F = 100 \). In each case, compute the population sizes after 1 month, after 2 months, and after 3 months.
   (a) \( R' = R + .3R - .2F \)
   (b) \( R' = R + .3R - .2F \)
   \( F' = F - .2F + .1R \)
   \( F' = F - .1F + .2R \)

15. Consider the following goat–sheep models, where the two species compete for common grazing land. In each case, compute the population sizes after 1 month, after 2 months, and after 3 months if the initial population is 50 goats and 100 sheep.
16. This exercise concerns the rabbit–fox model in equations (15). For given initial population sizes, calculate the population sizes after 1 month, after 2 months, and after 3 months. Also plot the trajectory of population sizes from the starting values to the stable sizes (as in Figure 1.8). The initial sizes are
(a) \( R = 30, F = 24 \)  
(b) \( R = 8, F = 3 \)  
(c) \( R = 8, F = 10 \)  
(d) \( R = 10, F = 10 \)

17. Consider the rabbit–fox model

\[
R' = R + .1R - .15F \\
F' = F - .3F + .2R
\]

What is the equation of the line of stable population sizes? For given initial population sizes, calculate the population sizes after 1 month, after 2 months, and after 3 months. Also predict the stable sizes to which these populations are converging. Compare your numbers with the calculations in Exercise 16. The initial sizes are
(a) \( R = 30, F = 24 \)  
(b) \( R = 8, F = 3 \)  
(c) \( R = 8, F = 10 \)  
(d) \( R = 10, F = 10 \)

18. Consider the rabbit–fox model

\[
R' = R + .1R - .2F \\
F' = F - .4F + .2R
\]

On a graph plot the following:
(a) The line of stable population sizes.
(b) The trajectory of population sizes starting from \( (10, 15) \).
(c) The trajectory of population sizes starting from \( (10, 30) \).
(d) The trajectory of population sizes starting from \( (20, 10) \).

19. Consider the rabbit–fox model

\[
R' = R + 2R - 3F \\
F' = F - 3F + 2R
\]

On a graph plot the following:
(a) The line of stable population sizes.
(b) The trajectory of population sizes starting from \( (10, 15) \).
(c) The trajectory of population sizes starting from \( (1, 2) \).
(d) How do the trajectories of this model differ from those in Figure 1.8?
20. Use a computer program to follow the behavior of the nonlinear rabbit–fox model in (19) over a period of 100 months (as in Figure 1.9) with the following starting values:

(a) \( R = 6, F = 6 \)  \hspace{1cm} (b) \( 100, 100 \)  \hspace{1cm} (c) \( 10, 10 \)

21. In Example 4, if the change in \( R \) is \( .1R - .01RF \) in 1 month, then in 1 day we would expect \( \frac{1}{10} \) of such a change \( \text{[i.e., } (\frac{1}{300})R - (\frac{1}{300})RF \text{]} \); similarly for the change in foxes. Write out the full set of equations for this model with time measured in days. Starting with \( R = 3, F = 3 \), follow the populations as before for 3000 days \( (= 100 \text{ months}) \). (Use a computer program.) How do your results compare with those in table (20)?

Projects

22. Use a computer program to follow the populations for many periods in the models in Exercises 14 and 15. Try a couple of different starting population sizes. In each case describe in words the long-term trends of the populations.

23. Make a thorough analysis of long-term trends for the rabbit–fox model

\[
\begin{align*}
R &= R + bR - eF \\
F &= F - dF + e'R
\end{align*}
\]

for different values of the positive parameters \( b, d, e, e' \). That is, list all possible long-term trends and give conditions on the parameters that tell when each trend occurs. For example, one trend is that both populations become extinct, with rabbits dying out first. Determine the conditions experimentally by trying many different specific parameter values and in each case computing the population sizes over many months.

Section 1.4

Linear Programming and Models Without Exact Solutions

In this section we examine two very important variations on the problem of solving \( n \) linear equations in \( n \) unknowns. We illustrate these variations with the refinery problem from Section 1.2.

Example 1. Refinery Problem Revisited with One Refinery Broken

The original refinery problem had three refineries and three products: heating oil, diesel oil, and gasoline. We wanted the production levels, \( x_1, x_2, x_3 \), of the refineries to meet demands of 500 gallons of heating
oil, 850 gallons of diesel oil, and 1000 gallons of gasoline. The equations we got were

Heating oil: \[20x_1 + 4x_2 + 4x_3 = 500\]

Diesel oil: \[10x_1 + 14x_2 + 5x_3 = 850\] (1)

Gasoline: \[5x_1 + 5x_2 + 12x_3 = 1000\]

Suppose that the third refinery breaks down and we have to try to meet the demands with two refineries. Our system of equations is now

**Two-Refinery Production Problem**

Heating oil: \[20x_1 + 4x_2 = 500\]

Diesel oil: \[10x_1 + 14x_2 = 850\] (2)

Gasoline: \[5x_1 + 5x_2 = 1000\]

A system like (2) with more equations than unknowns is called *overdetermined* and does not normally have a solution. All one can ask for is an approximate solution. We want a "solution" of \(x_1, x_2\) values that makes the total production of the two remaining refineries *as close as possible* to the demands. In Chapter 5 we define precisely the term "as close as possible" and then show how to solve this problem.

If we take the situation in system (2) to a greater extreme, with, say, 10 or 50 equations but just two unknowns, then we have a famous estimation problem in statistics.

**Example 2. Predicting Grades in College**

A guidance counselor at Scrooge High School wants to develop a simple formula for predicting a Scrooge graduate's GPA (grade-point average) at the local state college as a function of the student's GPA at Scrooge High. The formula would be a linear model of the form

\[\text{college GPA} = q \times \text{(Scrooge GPA)} + r\] (3a)

or

\[C = qS + r\] (3b)

where \(C\) stands for college GPA, \(S\) stands for Scrooge GPA, and \(r, q\) are constants to be determined based on the performances of past Scrooge graduates. Suppose that the data from eight students are
Figure 1.10 Estimated equation for relation between Scrooge GPA and college GPA.

One should pick the constants \( q \) and \( r \) so that the predicted college GPA given by the expression \( qS + r \) will be as close as possible to the actual college GPA for these students. Using a method discussed in Chapter 4, we set \( q = 1.1 \) and \( r = -0.9 \). The predicted college GPAs with this formula are given in the last column of the table. Figure 1.10 has a plot of the \( C \) and \( S \) values from (4) along with the suggested line \( C = 1.1S - 0.9 \).

This is the same sort of problem that we faced in finding an approximate solution to the refinery problem in Example 1. The statistical name for this type of estimation problem is regression.

Suppose that the counselor wanted to break down the Scrooge grade average into various components \( G_{M/S} \), \( G_E \), and \( G_{H/L} \), representing the student's grades in the three subject areas of math/science, English, and history/languages. The counselor would give these three components separate weightings in a formula like
The solution of regression problems is discussed in Chapters 4 and 5.

Next let us consider the situation where we have more unknowns than equations. Again we use the refinery model.

**Example 3. Refinery Model Revisited**

**Without Diesel Oil**

Suppose now that there is no demand for diesel oil and the three refineries just produce heating oil and gasoline. So the system of equations to be satisfied is

Two-Product Refinery Problem

\[
\begin{align*}
\text{Heating oil:} & \\
20x_1 + 4x_2 + 4x_3 &= 500 \\
\text{Gasoline:} & \\
5x_1 + 5x_2 + 12x_3 &= 1000
\end{align*}
\]

This system of equations is called *underdetermined* in the sense that there are not enough constraints to determine each unknown uniquely. The solution we found in Section 1.2 for the refinery problem with all three equations is clearly valid with two equations: \(x_1 = 5, x_2 = 33, x_3 = 68\). But many other solutions are possible. In particular, we could shut down one of the refineries, say refinery 3, as in Example 1. In mathematical terms, we seek a solution to (6) with \(x_3 = 0\). Dropping the \(x_3\) terms from (6), we have

\[
\begin{align*}
\text{Heating oil:} & \\
20x_1 + 4x_2 &= 500 \\
\text{Gasoline:} & \\
5x_1 + 5x_2 &= 1000
\end{align*}
\]

Solving (7) for \(x_2\) and \(x_3\) yields the solution \(x_2 = 500/7 \approx 71, x_3 = 3500/16 = 218.75\). We could set \(x_2 = 0\) and get another solution. There are many more solutions in which all the refineries are running.

Which solution would we use in practice? The answer is, the solution that is most efficient. That is, the solution that is the cheapest.
Each refinery will have a cost of operation. Suppose that the costs to refine a barrel of oil (the units for the \( x_i \)'s) are

<table>
<thead>
<tr>
<th>Refinery Operation Costs</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Refinery 1</td>
<td>$30 per barrel</td>
</tr>
<tr>
<td>Refinery 2</td>
<td>$25 per barrel</td>
</tr>
<tr>
<td>Refinery 3</td>
<td>$20 per barrel</td>
</tr>
</tbody>
</table>

Then we want a solution to the two-product production problem (6) (with no \( x_i \) negative) for which the total refining cost of \( 30x_1 + 25x_2 + 20x_3 \) is minimized. The complete mathematical statement of this problem is

**Optimal Refinery Production Problem**

Minimize \( 30x_1 + 25x_2 + 20x_3 \)
subject to the constraints
\[
20x_1 + 4x_2 + 4x_3 = 500 \\
5x_1 + 5x_2 + 12x_3 = 1000 \\
x_1 \geq 0, \ x_2 \geq 0, \ x_3 \geq 0
\]

The problem of optimizing (minimizing or maximizing) a linear expression subject to constraints that are linear equations or inequalities is called **linear programming**. Linear programming is the most important mathematical tool in management science. There are thousands of different real-world problems that can be posed as linear programming problems. When scientists at Bell Laboratories recently proposed a new, more efficient way to solve linear programming problems, the announcement was a front-page story in major newspapers.

The optimal refinery production problem (10) involved a set of linear equations as constraints together with the inequalities \( x_1 \geq 0, \ x_2 \geq 0, \ x_3 \geq 0 \). As we shall see shortly, it is easier to solve linear programs in which all the constraints are inequalities. Exercise 13 shows how to convert the equations in (10) into linear inequalities. This conversion is discussed further in Section 4.6.

As an example of how linear programs with inequalities are solved, we look at a simple two-variable maximization problem.

**Example 4. A Linear Program: Optimal Production of Two Crops**

Suppose that a farmer has 200 acres on which he can plant any combination of two crops, corn and wheat. Corn requires 4 worker-days of labor and $20 of capital for each acre planted, while wheat requires 1 worker-day of labor and $10 of capital for each acre planted. Suppose
also that corn produces $60 of revenue per acre and wheat produces $40 of revenue per acre. If the farmer has $2200 of capital and 320 worker-days of labor available for the year, what is the most profitable planting strategy?

If

\[ C = \text{number of acres of corn} \]
\[ W = \text{number of acres of wheat} \]

the constraints on land, labor, and capital are given by the following system of linear inequalities:

\[
\begin{align*}
\text{Land:} & \quad C + W \leq 200 \\
\text{Labor:} & \quad 4C + W \leq 320 \\
\text{Capital:} & \quad 20C + 10W \leq 2200
\end{align*}
\]

also,

\[ W \geq 0, \quad C \geq 0 \]

Subject to these constraints, we want to determine \( C \) and \( W \) so as to maximize the total revenue.

Maximize \( 60C + 40W \) \( (12) \)

The expression to be maximized is called the \textit{objective function}. When only two variables are involved, one can plot the inequality constraints and display the region of \((x_1, x_2)\)-points that simultaneously satisfy all the constraints in (11). This region is called the \textit{feasible region} of the linear program, and its points \textit{feasible points}. See the shaded area in Figure 1.11. Recall that to find the points satisfying an inequality such as \( C + W \leq 200 \), we plot the line \( C + W = 200 \) and then shade the line and all points on the lower left side of the line.

Once we have plotted the feasible region for (11), it remains to find out which feasible point maximizes \( 60C + 40W \). The following geometric insight greatly simplifies the solution of linear programs (this theorem is proved in Section 4.6).

Theorem. A linear objective function assumes its maximum and minimum values on the boundary of the feasible region. In fact, the optimal value is achieved at a corner point of this boundary.

Now we can solve our linear program. The theorem tells us where to look for the optimal \((C, W)\)-value—at the corners of the feasible region. To find a corner that lies at the intersection of two constraint lines, we solve for a \((C, W)\) point that lies on both lines—the same old problem of solving two equations in two unknowns. Once we determine the coordinates of a corner of the feasible region, we can
evaluate the expression $60C + 40W$ at that corner. The corner that maximizes this expression is the answer to our optimal production problem. In the current problem we can see from Figure 1.11 which intersections of constraint lines form corners of the feasible region.

Table 1.2 lists the coordinates of the corners and the associated objective function values. So the optimal production schedule is to plant 20 acres of corn and 180 acres of wheat. In Section 4.6 we present a more general, systematic approach to find the maximizing corner in a linear program.

<table>
<thead>
<tr>
<th>Corner Coordinates</th>
<th>Intersecting Constraints</th>
<th>Objective Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0)</td>
<td>$C \geq 0$ and $W \geq 0$</td>
<td>0</td>
</tr>
<tr>
<td>(0, 200)</td>
<td>$C \geq 0$ and Land</td>
<td>8000</td>
</tr>
<tr>
<td>(20, 180)</td>
<td>Land and Capital</td>
<td>8400***</td>
</tr>
<tr>
<td>(50, 120)</td>
<td>Capital and Labor</td>
<td>7800</td>
</tr>
<tr>
<td>(80, 0)</td>
<td>Labor and $W &gt; 0$</td>
<td>4800</td>
</tr>
</tbody>
</table>

Before leaving this example, we note that the farmer's constraints (Land, Labor, and Capital) determined the feasible region, and the "marketplace" (prices for corn and wheat) determined the objective function. If the market prices for corn or wheat change, or equivalently, the farmer receives a subsidy for one of the crops, the optimal solution may change.
Table 1.3

<table>
<thead>
<tr>
<th>Corner Coordinates</th>
<th>Intersecting Constraints</th>
<th>Objective Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0, 0)</td>
<td>$C \geq 0$ and $W \geq 0$</td>
<td>0</td>
</tr>
<tr>
<td>(0, 200)</td>
<td>$C \geq 0$ and Land</td>
<td>8000</td>
</tr>
<tr>
<td>(20, 180)</td>
<td>Land and Capital</td>
<td>9000</td>
</tr>
<tr>
<td>(50, 120)</td>
<td>Capital and Labor</td>
<td>9300***</td>
</tr>
<tr>
<td>(80, 0)</td>
<td>Labor and $W &gt; 0$</td>
<td>7200</td>
</tr>
</tbody>
</table>

For many years, the Federal Farm Program has offered crop subsidies in order to influence both the types of crops grown and the total number of acres planted.

To illustrate how a crop subsidy can cause land to be taken out of production, let us suppose in Example 4 that the farmer receives a subsidy for corn that increases the revenue from $60 to $90 per acre. Then the objective function is now $90C + 40W$. The values of this new objective function at the corner points of the feasible region are shown in Table 1.3. The new optimal strategy is to plant 50 acres of corn and 120 acres of wheat, for a total of 170 acres. The subsidy results in the farmer removing 30 acres from production.

Section 1.4 Exercises

Summary of Exercises
Exercises 1–5 involve overdetermined systems and regression. Exercises 6–12 involve linear programming. Exercise 13 tells how to convert a system of equations into a system of inequalities (this conversion is discussed further in Chapter 4).

1. Use a trial-and-error approach to estimate as closely as possible an approximate solution to the refinery problem in Example 1.

2. (a) Repeat Exercise 1, but now refinery 2 rather than refinery 3 is missing.
   (b) Repeat Exercise 1, but now refinery 1 rather than refinery 3 is missing.
   (c) If you had to close down one refinery, which refinery would you pick in order to meet the demand as closely as possible with the remaining two refineries?

3. Consider the following system of equations, which might represent supply–demand equations for chairs, tables, and sofas from two factories:
Find an approximate solution to this system of equations by trial and error.

4. For the following sets of x-y points, estimate a line to fit the points as closely as possible.
   (a) (1, 1), (2, 3), (3, 2), (4, 6), (5, 5)
   (b) (1, 6), (2, 4), (3, 3), (3, 2), (4, -1), (4, 0)

5. Suppose that the estimate for GPA in college in Example 2 had been
   \[ G_C = .6G_{M/S} + .3G_E + .3G_{H/L} - 1 \]
   where \( G_{M/S} \), \( G_E \), and \( G_{H/L} \) are the GPAs in mathematics/science, English, and humanities/languages. Based on this predictor, on which courses should students work hardest (if students want to improve their expected college GPA)?

6. Find a solution to the refinery problem in Example 3 in which the values of \( x_2 \) and \( x_3 \) are the same.

7. Change the labor constraint in the crop linear program of Example 4 to be \( 4C + 2W = 320 \). Now what would be the optimal solution?

8. Suppose that a Ford Motor Company factory requires 7 units of metal, 20 units of labor, 3 units of paint, and 8 units of plastic to build a car, while it requires 10 units of metal, 24 units of labor, 3 units of paint, and 4 units of plastic to build a truck. A car sells for $6000 and a truck for $8000. The following resources are available: 2000 units of metal, 5000 units of labor, 1000 units of paint, and 1500 units of plastic.
   (a) State the problem of maximizing the value of the vehicles produced with these resources as a linear program.
   (b) Plot the feasible region of this linear program.
   (c) Solve this linear program by the method in Example 4, by determining the coordinates of the corners of the feasible region and finding which corner maximizes the objective function.

   \textit{Hint:} By looking at the objective function, you should be able to tell which corners are good candidates for the maximum.

9. Suppose that a meal must contain at least 500 units of vitamin A, 1000 units of vitamin C, 200 units of iron, and 50 units of protein. A dietician has the following two foods from which to choose:
Meat: One unit of meat has 20 units of vitamin A, 30 units of vitamin C, 10 units of iron, and 15 units of protein.
Fruit: One unit of fruit has 50 units of vitamin A, 100 units of vitamin C, 1 unit of iron, and 2 units of protein.
Meat costs 50 cents a unit and fruit costs 40 cents a unit.

(a) State the problem of minimizing the cost of a meal that meets all the minimum nutritional requirements as a linear program (now you want to minimize the objective function).

(b) Plot the boundary of the feasible region for this linear program.

(c) Solve this linear program by the method in Example 4 [see Exercise 8, part (c)].

10. Consider the two-refinery problem in Example 1.

\[
\begin{align*}
\text{Heating oil:} & \quad 20x_1 + 4x_2 = 500 \\
\text{Diesel oil:} & \quad 10x_1 + 14x_2 = 850 \\
\text{Gasoline:} & \quad 5x_1 + 5x_2 = 1000 \\
\end{align*}
\]

Suppose that it costs $30 to refine a barrel in refinery 1 and $25 a barrel in refinery 2. What is the production schedule (i.e., values of \(x_1, x_2\)) that minimize the cost while producing at least the amounts demanded of each product (i.e., at least 500 gallons of heating oil, etc.)?

Hint: Solve by the method in Exercise 8, part (c).

11. Consider the following two linear programs.

(i) Maximize \(3x_1 + 3x_2\) subject to \(x_1 \geq 0, \ x_2 \geq 0\)
\[x_1 + 2x_2 \leq 10\]
\[2x_1 + x_2 \leq 8\]

(ii) Minimize \(10x_1 + 8x_2\) subject to \(x_1 \geq 0, \ x_2 \geq 0\)
\[x_1 + 2x_2 \geq 3\]
\[2x_1 + x_2 \geq 3\]

Solve them and show that the optimum values of these two objective functions are the same.

12. Set up the following problem as a linear program, but do not solve. There are two truck warehouses and two stores that sell trucks. The following table gives the cost of transporting a truck from one of the warehouses to one of the stores.

<table>
<thead>
<tr>
<th></th>
<th>Store 1</th>
<th>Store 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warehouse 1</td>
<td>$40</td>
<td>$50</td>
</tr>
<tr>
<td>Warehouse 2</td>
<td>$60</td>
<td>$40</td>
</tr>
</tbody>
</table>
Warehouse 1 has 100 trucks and warehouse 2 has 80. Store 1 needs at least 50 trucks and store 2 needs at least 100 trucks. Find the cheapest way to meet the stores’ demand.

_Hint:_ Let the variables be $x_{11}$, $x_{12}$, $x_{21}$, and $x_{22}$, where $x_{ij}$ is the amount shipped from warehouse $i$ to store $j$.

13. To convert a system of equations in which each variable must be $\geq 0$, into a system of inequalities with each variable $\geq 0$, we perform the following steps.

(i) Pick a variable in the first equation and solve that equation for the chosen variable, that is, so that the chosen variable is alone on one side of the equation. For example, in the system of equations

\[
\begin{align*}
2x + 4y + 6z &= 8 \\
x + 3y + 2z &= 6 \\
x \geq 0, \quad y \geq 0, \quad z \geq 0
\end{align*}
\]

if we pick $x$ in the first equation, then we rewrite it as

\[
(*) \quad x = -2y - 3z + 4
\]

(ii) Replace the chosen variable in the other equations by substituting in its place the right-hand side in $(*)$. So the second equation in this problem becomes

\[
(-2y - 3z + 4) + 3y + 2z = 6
\]

or

\[
y - z = 2
\]

(iii) Since the chosen variable is $\geq 0$, the right side of $(*)$ must be $\geq 0$, that is,

\[
-2y - 3z + 4 \geq 0
\]

or

\[
4 \geq 2y + 3z \quad \text{(equivalently,} \quad 2y + 3z \leq 4)\]

Now the original three-variable problem has been reduced to a two-variable problem with one equation converted into an inequality.

\[
\begin{align*}
2y + 3z &\leq 4 \\
y - z &= 2 \\
y \geq 0, \quad z \geq 0
\end{align*}
\]
(iv) Repeat the entire procedure for a variable in the second equation. If we pick \( y \), solve the second equation for \( y \) in terms of \( z \), substitute this expression involving \( z \) in place of \( y \) in the first inequality; also make this expression in \( z \) be \( \geq 0 \).

(a) Complete step iv.

(b) Convert the refinery linear program at the end of Example 3 into a linear program with inequalities, and solve the linear program.

(c) Convert the following system of equations for nonnegative variables into a system of inequalities.

\[
\begin{align*}
x_1 + 3x_2 + 2x_3 &= 10 \\
2x_1 + 5x_2 - 4x_3 &= 15 \\
x_1 &\geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0
\end{align*}
\]

Section 1.5 Arrays of Data and Linear Filtering

In the previous sections we have encountered arrays of numbers that were the coefficients of systems of equations. But not all arrays of numbers are sets of coefficients. There are many problems in which arrays of numbers are input data to be analyzed. In statistics, we study huge data sets that come in sequences, two-dimensional arrays, and more complex structures. In the field of information processing and pattern recognition, certain information must be extracted from the data, be it a coded message or a picture. Both of these fields make heavy use of linear models to process arrays of data.

The examples in this section illustrate the use of linear models in pattern recognition and encoding of information. First we consider an example in which the data to be processed are letters, not numbers.

**Example 1. Linear Models for Encoding Alphabetic Messages**

A common approach to coding an alphabetic message is to treat each letter in the message as a number between 1 and 26: A \( \rightarrow 1 \), B \( \rightarrow 2 \), C \( \rightarrow 3 \), \ldots, Z \( \rightarrow 26 \). The simplest way to encode a message is to convert each letter (number) to a different letter (number) using some simple arithmetic formula. For example, we could add 7 to each number (this shifts the corresponding letter seven places to the right in the alphabet) or multiply each number by 11. However, these operations will sometimes convert a number between 1 and 26 to a number greater than 26.

To ensure that the result of some calculation is a number between 1 and 26, we usually assume that all arithmetic is done mod 26. As an example,
7 \times 13 = 91 = 13 \pmod{26} \quad \text{since } 91 = 3 \times 26 + 13

Encoding schemes based on adding a constant to each number (letter) or multiplying by a constant are easy for an outsider to break because once one letter is guessed, the constant can be determined. For example, if \( X_L \) is the original letter and \( X_C \) the coded letter, the encoding

\[
X_C = 7X_L
\]  

(1)

transforms THE into JDI since in numbers T = 20, H = 8, and E = 5, so

\[
\begin{align*}
7 \cdot T &= 7 \cdot 20 = 140 \equiv (\text{mod } 26) \quad 10 = J \\
7 \cdot H &= 7 \cdot 8 = 56 \equiv (\text{mod } 26) \quad 4 = D \\
7 \cdot E &= 7 \cdot 5 = 35 \equiv (\text{mod } 26) \quad 9 = I
\end{align*}
\]

Now if we guess that I is the encoding of E, it is easy to compute that the constant in (1) is 7. Even a code with multiplication and addition, such as \( X_C = 9X_L + 21 \), is easy to break.

A better scheme, in which frequent words and letters are scrambled, is to encode numbers in pairs using two linear equations of the following form. Let \( L_1, L_2 \) be a pair of original letters (represented as numbers between 1 and 26), and \( C_1, C_2 \) be the coded letters (also represented as numbers) into which \( L_1, L_2 \) are transformed.

\[
\begin{align*}
C_1 &= aL_1 + bL_2 \pmod{26} \\
C_2 &= cL_1 + dL_2 \pmod{26}
\end{align*}
\]  

(2)

For example, in the scheme

\[
\begin{align*}
C_1 &= 9L_1 + 17L_2 \pmod{26} \\
C_2 &= 7L_1 + 2L_2 \pmod{26}
\end{align*}
\]  

(3)

the pair \( E, C \), represented numerically as 5, 3, would be encoded as

\[
\begin{align*}
C_1 &= 9 \times 5 + 17 \times 3 = 96 \equiv 18 \pmod{26} = R \\
C_2 &= 7 \times 5 + 2 \times 3 = 41 \equiv 15 \pmod{26} = O
\end{align*}
\]  

(4)

To use (2) for a whole message, we divide the string of \( m \) letters (ignoring punctuation) into \( m/2 \) successive pairs. Observe that if the fifth letter in the message were an E, the fifth letter in the encoded sequence would vary depending on what the sixth letter was [with which E is paired in (2)]. There are four constants \([a, b, c, d] \) in this encoding scheme, and hence \( 26^4 = 456,976 \) different possible schemes. Moreover, there are no meaningful patterns of frequently used letters to help a codebreaker.
If the codebreaker had access to a large computer, we could counter this by grouping letters in sets of 5 and replace (2) by a scheme involving five equations of linear combinations of five letters. Now there would be 25 constants yielding \(26^{25} = 10^{35}\) schemes—and we can go to sleep knowing that our code is secure.

Although it is important to keep this code from being broken, it is also important that a code not be too hard to decode by a receiver. If the receiver knows the constants in (2), he or she still has to reverse the encoding process by solving a pair of equations in two unknowns.

For example, if (3) were being used and the pair \(R, O\) \((= 18, 15)\) generated in (4) were received, the decoder would have to solve the system of equations

\[
\begin{align*}
9L_1 + 17L_2 &= 18 \pmod{26} \\
7L_1 + 2L_2 &= 15 \pmod{26}
\end{align*}
\]

For a more complex scheme of five equations in five unknowns, the decoding problem gets harder, especially since arithmetic is mod 26. Fortunately, we shall show in Chapter 3 that there exist simple formulas for decoding so that the original pair \(L_1, L_2\) (or 5-tuple) can be computed as a linear combination of the coded pair \(C_1, C_2\). For example, the decoding equations for (3) are

\[
\begin{align*}
L_1 &= 18C_1 + 3C_2 \pmod{26} \\
L_2 &= 15C_1 + 3C_2 \pmod{26}
\end{align*}
\]

The next two examples involve data analysis.

---

**Example 2. The Mean of a Data Set**

The most basic piece of statistical information about a set of data is the mean, or average, of the data. The mean is obtained by summing all data values and dividing by the number of data. For example, the mean of the sequence of numbers 2, 3, 13, 3, 7, 1, 9, 3, 4, 5 is \((2 + 3 + 13 + 3 + 7 + 1 + 9 + 3 + 4 + 5)/10 = 5\).

The mean is a linear combination of these 10 data values in which each value is multiplied by \(1/10\) (recall from Section 1.1 that a linear combination is an expression of the form \(c_1x_1 + c_2x_2 + \cdots + c_nx_n\)).

---

**Example 3. Smoothing a Time Series**

In many situations one receives a sequence of numbers recorded over time that form a pattern, but randomness in nature or in recording and transmitting the numbers has obscured the pattern. Such sequences are called time series. Consider the series of readings in Data Plot 1 taken
over a period of time. Suppose that this time series gives the numbers of people applying for welfare aid in some city in successive months (the numbers are presented in units of 100). The values might equally well have represented levels of X-rays measured in a spacecraft or the numbers of new houses started in the United States in successive months.

To help picture the data, we plot the numbers in a graph, with time measured on the vertical axis and the data values on the horizontal axis. (The axes are omitted in the graph.)

We want to try to find a long-term pattern in this time series by smoothing the data—that is, reducing the jumps in data from one period to the next. In engineering, the task of smoothing a noisy electronic signal is called filtering (the term is now also used in nonengineering settings).
The simplest way to filter a time series is by replacing the \( i \)th value \( d_i \) by the average of \( d_i \) and the two adjacent values \( d_{i-1} \) and \( d_{i+1} \). That is, we form a new time series whose \( i \)th value \( d'_{i} \) is given by

\[
d'_{i} = \frac{d_{i-1} + d_{i} + d_{i+1}}{3}
\]  

(7)

For example, we replace \( d_{15} \) by

\[
d'_{15} = \frac{d_{14} + d_{15} + d_{16}}{3} = \frac{22 + 25 + 20}{3} = 22
\]

(When the value of \( d'_i \) is fractional, we shall round to the nearest integer.) The formula in (7) is not defined for the first and last values \((i = 1 \text{ and } i = 27)\); instead, let us set \( d'_1 = \frac{d_1 + d_2}{2} \) and \( d'_{27} = \frac{d_{26} + d_{27}}{2} \). The complete system of linear equations (the linear model) for filtering is then

\[
\begin{align*}
d'_1 &= \frac{1}{2}d_1 + \frac{1}{2}d_2 \\
d'_2 &= \frac{1}{3}d_1 + \frac{1}{3}d_2 + \frac{1}{3}d_3 \\
d'_3 &= \frac{1}{3}d_2 + \frac{1}{3}d_3 + \frac{1}{3}d_4 \\
&\vdots \\
d'_{n-1} &= \frac{1}{3}d_{n-2} + \frac{1}{3}d_{n-1} + \frac{1}{3}d_n \\
d'_n &= \frac{1}{3}d_{n-1} + \frac{1}{3}d_n
\end{align*}
\]

(8)

Our new time series looks as shown in Data Plot 2. This time series is much smoother. There is a clear trend of increasing values, then decreasing, then increasing, and finishing relatively level.

To smooth these data further, we could apply the smoothing transformation (8) again to this new time series. Instead, let us smooth the original data (in Data Plot 1) by applying a weighted average of five values in which \( d_i \) is weighted more and \( d_{i-2} \) and \( d_{i+2} \) are weighted less:

\[
d''_i = \frac{d_{i-2} + 2d_{i-1} + 3d_i + 2d_{i+1} + d_{i+2}}{9}
\]

(9)

For example, now \( d_{15} \) is replaced by

\[
d''_{15} = \frac{d_{13} + 2d_{14} + 3d_{15} + 2d_{16} + d_{17}}{9} = \frac{24 + 2 \times 22 + 3 \times 25 + 2 \times 20 + 16}{9} = \frac{199}{9} = 22
\]
For $d''_i$, drop the missing terms from (8) to obtain $d''_i = (3d_1 + 2d_2 + d_3)/6$, and similarly for $d''_2$, $d''_{29}$, $d''_{30}$. The new time series obtained when transformation (8) is applied to the original data is as shown in Data Plot 3. Observe how smooth this transformed time series is as compared to the original one. The Exercises have other examples of time series for which filtering reveals important trends.

We can show that applying the three-value average (7) to Data Plot 2; or, equivalently, applying the three value average twice to the original data, produces the same time series as in Data Plot 3. In general, successively performing two (or more) linear filterings is equivalent to performing another (more complicated) filtering; see the Exercises for examples.
Example 4. Linear Filtering in Pattern Recognition

When the TV camera that serves as the eyes of a robot transmits a picture to the robot’s computer, the picture is sent as a two-dimensional array of numbers that indicate the darkness of each point in the picture. A computer program to perform pattern recognition must determine what the robot is seeing by analyzing this digital representation of the picture. Light reflecting off an object or confusing patterns in the background can make a simple object quite difficult to recognize. Transformations to filter the data and increase the level of contrast are an essential part of any pattern recognition program.
Suppose that there are nine shades of darkness of a point, represented by integers 0 through 8, 0 for white (least dark) and 8 for black:

\[0 = \ldots, \quad 1 = \ldots, \quad 2 = \ldots, \quad 3 = \ldots, \quad 4 = \ldots, \quad 5 = \ldots, \quad 6 = \ldots, \quad 7 = \ldots, \quad 8 = \ldots\]

A picture represented by the $12 \times 12$ array of darkness values shown in Data Plot 4 has been received from a robot's TV camera. (Of course, a $12 \times 12$ array would only be a small section of the full TV image.)

Let us perform a linear filtering on this array of darkness values. We replace each value by a weighted average of the value and the eight values surrounding it, as denoted by 1's and 4's in Figure 1.12.

We use a weighting in which the old value gets a weight of 16, the four neighboring values in the same row or column (denoted by 4's in Figure 1.12) get a weight of 4, and the diagonal neighboring values (denoted by 1's in Figure 1.12) get a weight of 1; we divide this weighted sum by 36. (At the edges of the array, we increase the weights on the border values to 5, 20, 5, as shown in Figure 1.12.) The transformed array is shown in Data Plot 5.
We might now start to perceive a person in the figure. But to see the person clearly we need greater contrast. To increase the contrast between light and dark, we apply the following linear function with roundoff: $f(x) = 3x - 9$; a value below 0 is rounded to 0 and a value above 8 is rounded to 8. The table for this contrast function is

<table>
<thead>
<tr>
<th>Old Value</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Value</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>6</td>
<td>8</td>
<td>8</td>
<td>8</td>
</tr>
</tbody>
</table>

With this contrast function, Data Plot 5 becomes Data Plot 6. Now the person is fairly visible, perhaps with an object by the left foot. Greater contrast would help a little more. A good computer program to recognize patterns should now be able to "see" that the object pictured is a human being.
"Adaptive" filtering schemes do a small amount of filtering and then look for "borders" between light and dark regions. The region around a border is then subjected to a contrast transformation to accentuate the border, while nonborder regions are filtered as above.

Section 1.5 Exercises

Summary of Exercises
Exercises 1–5 are associated with Example 1 about codes. Exercises 6–12 are associated with Example 3 on filtering time series; Exercises 10–12 involve algebraic composition of transformations—they require more maturity. Exercises 13–15 are associated with Example 4 about two-dimensional arrays forming "pictures."

1. Evaluate the following expressions mod 26. All answers must be positive numbers between 1 and 26.
   (a) 7 \times 7  
   (b) 12 \times 5  
   (c) -5 \times 5  
   (d) -11 \times 19  
   (e) 12 \times (14 + 17)

2. Use the encoding \( C = 5L + 7 \pmod{26} \) to encode the letters in the following words.
   (a) BE  
   (b) AT  
   (c) APE

3. Use the encoding in equations (3) to encode the following pairs of letters.
   (a) BE  
   (b) AT  
   (c) CC

4. Use the decoding in equations (6) to decode the following pairs of letters.
   (a) BG  
   (b) CC  
   (c) RD

5. Determine the value(s) of \( x \) that satisfies (satisfy) the following equations mod 26. Which equations have unique solutions?
   (a) 9x = 11  
   (b) 7x = 13  
   (c) 14x = 3  
   (d) 11x = 9

6. Apply the filtering transformation in formula (7) to the first 15 numbers in the time series in Data Plot 2. (You should get the same results as in Data Plot 3.)

7. Apply the following filtering transformations to Data Plot 1 (explain how you alter these transformations for the first and last values).
   (a) \( d'_i = \frac{d_{i-2} + d_i + d_{i+2}}{3} \)
(b) \[ d'_i = \frac{d_{i-2} + d_{i-1} + d_i + d_{i+1} + d_{i+2}}{5} \]

(c) \[ d'_i = \frac{d_{i-3} + d_{i-1} + d_{i+1} + d_{i+3}}{4} \]

8. Consider the time series 2, 10, 4, 12, 6, 14, 8, 16, 10, 18, 12, 20. Apply transformations (a), (b), and (c) from Exercise 7. Which of the transformations smooth this time series well, and which do a poor job?

9. Consider the time series 1, 4, 2, 5, 8, 6, 3, 10, 3, 12, 10, 9, 8, 12, 18, 13, 21, 16, 16. Apply transformations (a), (b), and (c) from Exercise 7. Which of the transformations smooth this time series well, and which do a poor job?

10. Show algebraically that if the transformation in formula (7) were applied to a time series and applied again to the resulting time series, then the cumulative result would be the same as the transformation in formula (9).

11. (a) Suppose that the transformation in Exercise 7, part (a) is applied twice (as described in Exercise 10). Give a formula for the cumulative transformation.

(b) Suppose that the transformation in Exercise 7, part (a) is applied to a time series and then the resulting time series is filtered by the transformation in Exercise 7, part (b). Give a formula for the cumulative transformation.

12. Suppose that \[ d'_i = a_1d_{i-1} + a_2d_i + a_3d_{i+1}, \]
\[ d''_i = b_1d_{i-1} + b_2d_i + b_3d_{i+1}. \]
Give a formula for the transformation obtained by performing the first and then the second transformation on a time series.

13. Apply the contrast function (given just before Data Plot 6) to Data Plot 6.

14. Apply the following filtering transformations to the upper 8-by-8 corner of Data Plot 4 (just the first 8 rows and first 8 columns), and then apply the contrast function given in the text. The transformations are described by a 3-by-3 square of weights, as in Figure 1.12. Explain what you did at the borders.

(a) \[
\begin{pmatrix}
1 & 1 & 1 \\
1 & 1 & 1 \\
1 & 1 & 1 \\
\end{pmatrix}
\]

(b) \[
\begin{pmatrix}
0 & 1 & 0 \\
1 & 4 & 1 \\
0 & 1 & 0 \\
\end{pmatrix}
\]

(c) \[
\begin{pmatrix}
1 & 2 & 1 \\
2 & 8 & 2 \\
1 & 2 & 1 \\
\end{pmatrix}
\]

(d) \[
\begin{pmatrix}
1 & 2 & 1 \\
2 & 0 & 2 \\
1 & 2 & 1 \\
\end{pmatrix}
\]

How effectively does each transformation help reveal the human being in the picture?
15. Consider the following "pictures," given in terms of numbers rather than darkness levels. Apply the transformation in Example 4 followed by the contrast function. Give your answer in darkness levels. What is the letter or number in each picture?

<p>| | | | | | | | | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>5 7 1 2 7</td>
<td>(b)</td>
<td>8 7 2 5 8</td>
<td>(c)</td>
<td>3 3 5 1 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 4 1 5 1</td>
<td>7 6 1 3 7</td>
<td>4 7 3 3 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 3 5 2 1</td>
<td>4 8 3 7 6</td>
<td>3 8 4 1 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 7 5 7 8</td>
<td>8 5 8 5 1</td>
<td>2 6 3 3 8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 8 6 7 5</td>
<td>7 4 8 7 3</td>
<td>6 6 1 6 6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8 4 3 4 8</td>
<td>6 8 2 8 5</td>
<td>3 8 2 4 7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 8 7 4 7</td>
<td>7 4 4 6 7</td>
<td>1 3 1 5 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5 7 6 8 6</td>
<td>8 7 1 1 8</td>
<td>5 1 6 2 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>