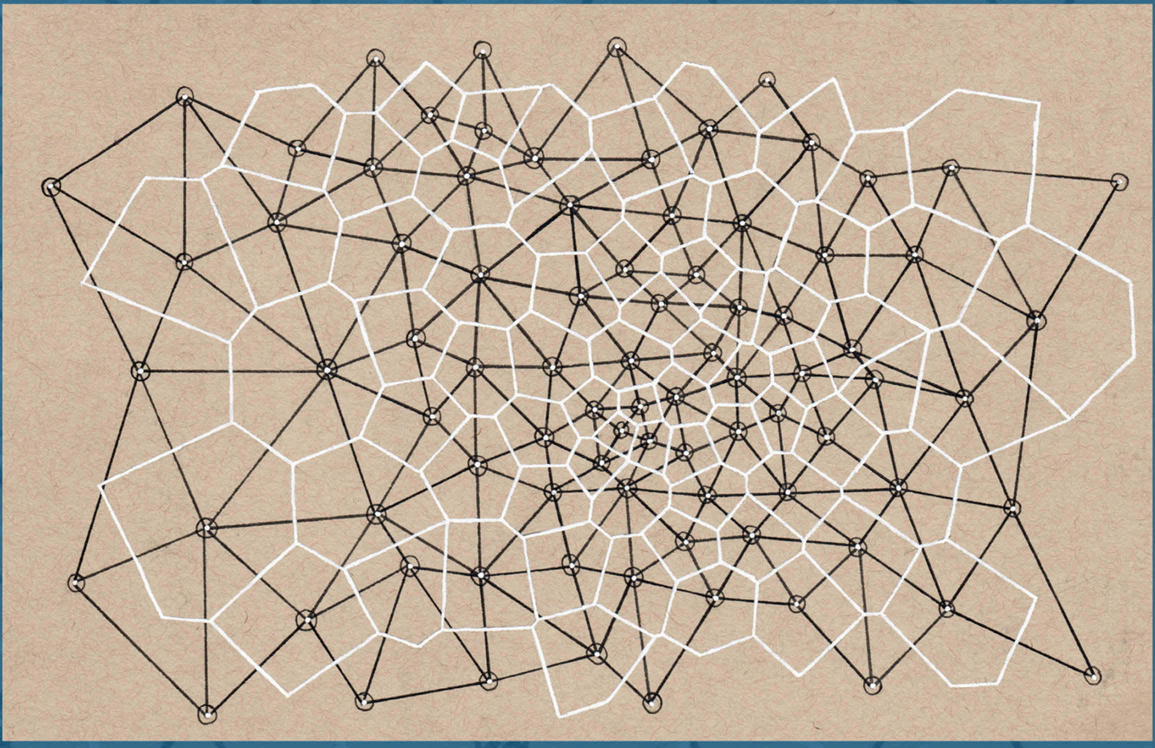
TEXTBOOKS IN MATHEMATICS

MATHEMATICAL MODELING WITH EXCEL

Second Edition



Brian Albright
William P. Fox

CRC Press
Taylor & Francis Group
A CHAPMAN & HALL BOOK

Mathematical Modeling with Excel

Second Edition

Textbooks in Mathematics

Series editors:
Al Boggess and Ken Rosen

GRAPH THEORY AND ITS APPLICATIONS, THIRD EDITION
Jonathan L. Gross, Jay Yellen and Mark Anderson

COMPLEX VARIABLES: A PHYSICAL APPROACH WITH APPLICATIONS, SECOND EDITION
Steven G. Krantz

GAME THEORY: A MODELING APPROACH
Richard Alan Gillman and David Housman

FORMAL METHODS IN COMPUTER SCIENCE
Jiacun Wang and William Tepfenhart

AN ELEMENTARY TRANSITION TO ABSTRACT MATHEMATICS
Gove Effinger and Gary L. Mullen

ORDINARY DIFFERENTIAL EQUATIONS: AN INTRODUCTION TO THE FUNDAMENTALS, SECOND EDITION
Kenneth B. Howell

SPHERICAL GEOMETRY AND ITS APPLICATIONS
Marshall A. Whittlesey

COMPUTATIONAL PARTIAL DIFFERENTIAL PARTIAL EQUATIONS USING MATLAB®, SECOND EDITION
Jichun Li and Yi-Tung Chen

AN INTRODUCTION TO MATHEMATICAL PROOFS
Nicholas A. Loehr

DIFFERENTIAL GEOMETRY OF MANIFOLDS, SECOND EDITION
Stephen T. Lovett

MATHEMATICAL MODELING WITH EXCEL, SECOND EDITION
Brian Albright and William P. Fox

https://www.crcpress.com/Textbooks-in-Mathematics/book-series/CANDHTEXBOOMTH

Mathematical Modeling with Excel

Second Edition

Brian Albright
William P. Fox

CRC Press
Taylor & Francis Group
Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business
A CHAPMAN & HALL BOOK

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed on acid-free paper

International Standard Book Number-13: 978-1-138-59707-5 (Hardback)

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface

The main goal of this text is to present different ways of building and analyzing mathematical models in a format that can be read by students, not just instructors. This is not a text on how to use Excel. Rather, Excel is seen as a tool to further the goal of building and analyzing mathematical models. No prior knowledge or experience with Excel is required to use this text.

Excel is chosen as the only software used to implement and analyze models for two main reasons:

1. It is easy to use and most everyone is familiar with it, so it takes very little time to become comfortable with the software.

2. It is everywhere. Students will have access to Excel for every mathematical modeling project they encounter inside and outside of academics.

Each section contains step-by-step instructions for building the models in Excel. These instructions were originally written for use with Office Excel 2016. Some of the instructions may be slightly different for other versions of Excel.

Pedagogical Approach

This text presents a wide variety of common types of models found in other mathematical modeling texts, as well as some new types. However, the models are presented in a very unique format. A typical section begins with a general description of the scenario being modeled. The model is then built using the appropriate mathematical tools. Then it is implemented and analyzed in Excel via step-by-step instructions. In the exercises, we ask students to modify or refine the existing model, analyze it further, or adapt it to similar scenarios.

In each section, we try to focus on the main mathematical modeling concept being illustrated and not get too bogged down in details. We also focus on the analysis of models, and in each case try to address the question, "What does this mean?"

This is not a "plug-and-chug" textbook. We do not ask students to simply plug numbers into some "black-box" Excel formula and accept the results. Rather, we discuss the mathematics behind the analysis of the models and, where appropriate, build the analytical tools in Excel from scratch.

Each section ends with several exercises of varying degree of difficulty. In addition, each chapter ends with a "For Further Reading" section which contains resources for additional information.

Audience/Prerequisites

This text is appropriate for mathematics majors (including secondary mathematics education majors) who need an introductory mathematical modeling course. Some sections require calculus, linear algebra, differential equations, or basic statistics, so this text is appropriate for use with junior or senior level students. However, many other sections require only mathematical maturity, so this text could also be used with sophomore level students.

The Flow of Material

This text contains a wide variety of modeling techniques, mathematical concepts, and types of applications. Here we give a brief overview of the highlights of each chapter.

Chapter 1 – **What is Mathematical Modeling?** This chapter begins with the definitions of the terms *model* and *mathematical modeling*. It then discusses the steps involved in mathematical modeling, and concludes with a discussion of the importance of assumptions in the process of mathematical modeling.

Chapter 2 – **Proportionality and Geometric Similarity** This chapter begins with an introduction to graphing and working with data in Excel which includes a discussion of fitting straight lines to data. Then the geometric concepts of proportionality and similarity are used to model systems such as free-falling objects. We stress the point that data are used to test the validity of the models. The chapter ends with an introduction to fitting straight lines to data, empirical modeling, and the coefficient of determination.

Chapter 3 – **Linear Algebra** This chapter begins with a brief introduction to topics in linear algebra used throughout this book including matrices, vectors, and systems of linear equations. We give a few examples of using linear equations to create models. We then use linear equations to fit polynomials, multiple regression, and spline models to data.

Chapter 4 – **Discrete Dynamical Systems** This chapter begins with the definitions of a discrete dynamical system, a solution, and an equilibrium value. We stress the point that we are usually interested in the long–term behavior of the system, not necessarily the value at a single point in time, and how equilibrium values are important in the analysis. The chapter includes several different types of applications modeled with discrete dynamical systems including population growth, predator–prey systems, and simple epidemics.

Chapter 5 – **Differential Equations** This chapter begins with a discussion of the fact that it is often easier to describe how a quantity changes over time than it is to describe the value of the quantity at any particular time. This motivates the use of differential equations for modeling dynamical systems. We focus on finding approximate numerical solutions to differential equations rather than finding exact analytical solutions. To this end, we discuss Euler's method for approximating solutions to differential equations and apply it to several applications of systems of differential equations. We also introduce Runge-Kutta methods for approximating solutions to differential equations.

Chapter 6 – **Simulations** We cover the topic of simulations more extensively than most other mathematical modeling text books. The main goal of this chapter is to illustrate several different types of simulation models including games of chance, queuing models, inventory models, and scheduling models. We also discuss how pseudo-random number generators work and how to model random variables using density functions.

Chapter 7 – Linear Optimization The main focus of this chapter is linear programming and the simplex method. We do not discuss much theory; rather we try to give students an overview of the basic ideas behind the simplex method. We introduce the assignment problem and the transportation problem as examples of linear programs and how to model with linear programs.

Chapter 8 – Nonlinear Optimization In this chapter we cover several numeric techniques for approximating solutions to nonlinear optimization problems including Newton's method, gradient methods, and Lagrange multipliers. We also discuss the inherent difficulties of solving nonlinear problems.

Selection of Material

There is more than a semester's worth of material in this text. An instructor can easily pick and choose the sections that are appropriate for the students. Most sections can easily be covered in one 50-minute class period. We give three suggestions for choosing material:

1. We suggest starting with Chapters 1, 2, and 3. Chapter 1 gives an essential overview of the mathematical modeling process which can be covered in one class period. We have embedded an introduction to working with Excel throughout Chapter 2. Chapter 2 also introduces many concepts used throughout the rest of the text. Chapter 3 is a continuation of Chapter 2 focused on linear algebra. If students have a solid background in linear algebra, Sections 3.1 and 3.2 can be skipped. Section 3.5 can be skipped without loss of continuity.

2. If the instructor wishes to focus on dynamical systems, we suggest covering Chapters 4 and 5. In Chapter 4 we model systems using discrete time increments, and in Chapter 5 we model time continuously. Sections 4.6, 5.6, 5.7, and 5.8 can be skipped without loss of continuity. Selected topics from Chapters 6 and 7 can round out the semester.

3. If the instructor wants to focus on topics from operations research, we suggest covering Chapters 6, 7, and 8. These chapters do not depend on Chapters 4 and 5. Sections 7.7, 8.6, 8.7, and 8.8 can be skipped without loss of continuity.

What's New in the Second Edition

Work on the second edition began as soon as the first edition was published. The authors would like to thank their students for finding mistakes and suggesting improvements and new exercises. Changes made from the first edition include:

- All directions for using Excel have been updated for Excel 2016.
- Numerous new exercises have been added throughout the text.
- Chapter 3 on linear algebra has been added. This chapter includes topics from Chapters 2 and 3 in the first edition as well as new material.

- Chapter 5 on differential equations has been expanded to include mixing problems and an introduction to Runge-Kutta methods.
- Chapter 6 on simulation models has been revised.
- Chapter 8 on nonlinear optimization has been added.
- Many chapters have a subsection containing project ideas. A project is defined as a problem devised by a student, the end-result of which is a short written report describing the problem, the analysis, and the conclusions. These project ideas are inspired by actual student projects and can serve as kernels of inspiration for future students.

Acknowledgements

The authors would like to thank the following reviewers who provided valuable suggestions during the development of this book:

- Philip Gustafson, Mesa State University
- Greg Grindey, United States Transportation Command
- Bruce DeGroot, Concordia University, Nebraska
- Len Cabrera, United States Air Force Academy
- Gary De Young, Dordt College

1

What is Mathematical Modeling?

Chapter Objectives

- Define the terms *model*, *mathematical model*, and *mathematical modeling*
- Understand the purpose and process of mathematical modeling
- Understand the importance and significance of assumptions behind a mathematical model

Every student of mathematics has done some "mathematical modeling" in his/her educational career. These instances of mathematical modeling are typically called "applications" and are used to illustrate how mathematics is implemented in the "real world."

In most math classes, the main goal is to learn the theory of some particular mathematical discipline. The applications are used to help achieve this goal by providing a more concrete context in which to study and understand the theory. For instance, in Calculus I, the real goal is to understand the idea of the limit and the derivative. An applied maximization problem is used to motivate the idea of the derivative and to provide practice in calculating and interpreting derivatives.

In mathematical modeling, the opposite is true. Here we will start with some "real world" problem and use mathematical theory and techniques to better understand the phenomena behind the problem.

1.1 Definitions

To define the phrase *mathematical modeling*, we will first define the term *model*. The word model is used frequently in everyday language. We talk about model airplanes, model houses, models on a runway, etc. What does the term model mean in a mathematical sense?

Lucas (Lucas, William F., The Impact and Benefits of Mathematical Modeling, in *Applied Mathematical Modeling* (D.R Shier and K.T. Wallenius eds.), Chapman and Hall/CRC, 1999, pg. 5) defines a model as "a simpler realization or idealization of some more complex reality." The real world is a very complex place. To better understand it, we need to try to simplify it to a reasonable degree, describe the simplification in ways we can understand and work with, and then study the simplification. This is what we call *modeling*.

A *mathematical model* then can be defined as a model constructed using mathematical terms, symbols, and ideas. Giordano et. al. (Frank. R. Giordano, M. D. Weir, and W. P. Fox, *A First Course in Mathematical Modeling*, Third ed., Thomson Brooks/Cole, 2003, pg. 54) defines a mathematical model as "a mathematical construct designed to study a particular real world system or phenomenon." Mathematical models can take many different forms. They may involve equations, inequalities, differential equations, matrices, logic, or any other type of "mathematical" idea.

The key idea is that we use mathematics to describe a portion of the real world. Therefore, a very simple but general definition of the *process* of mathematical modeling is:

Definition 1.1.1. *Mathematical modeling* is the application of mathematics to real world problems.

1.2 Purpose

Why do we do mathematical modeling? Since we want to answer a question about real world phenomena, we could just sit back, observe, and take notes. Suppose we put 500 bacteria in a Petri dish. The next day we count 525 in the dish, and the next we count 551.

Obviously, the number of bacteria is growing. Based on this observation, we might ask these questions:

1. How long will it be until there are 600 bacteria in the dish?
2. If we need 900 bacteria for an experiment in 3 days, how many must we put into the dish today?

We could answer each question as follows:

1. Wait until we count 600 bacteria in the dish.
2. Put 1 bacterium in a dish, 2 in a second dish, 3 in a third, etc. up to 900, wait 3 days, and determine which dish contains 900 bacteria.

These solutions only require us to make simple observations of this real world phenomenon of bacteria growing in a Petri dish. However, these solutions are obviously impractical for they might require too much time or too many resources (the second solution requires 900 Petri dishes and a total of $1 + 2 + \cdots + 900 = 405,450$ bacteria).

A much more practical approach to answering these questions is to construct a function that gives the number of bacteria in the dish in terms of time (i.e. construct a mathematical model of the bacteria growth).

In other situations, making observations may itself be a complicated ordeal. For instance, suppose we wanted to find the optimal mixture of doctors and nurses (i.e. the number of doctors and number of nurses) to staff a hospital emergency room. The concept of "optimal" may take into account several factors, including:

1. Quality of patient care. (Do they get the care they need?)
2. Patient waiting time. (Do they have to wait a long time?)
3. Time spent with patients. (Are the doctors and nurses over-worked, or do they have too much "free time?")
4. Resources. (Is there enough floor space or are people running into each other?)

One approach to finding an optimal number is to choose some mixture of doctors and nurses (say 3 doctors and 8 nurses), put them to work, and have a team of people record data for a series of weeks or months. Then choose another mixture (say 2 doctors and 7 nurses) and repeat the process. Repeat this until all possible combinations of doctors and nurses have been tried, analyze the data, and pick the optimal mixture.

This approach has many of the same problems as the bacteria growth problem. It would take too much time and be too expensive. Plus there are additional problems. If there are too few doctors and nurses on staff, patients might unnecessarily die. Plus we may not observe how the different mixtures handle infrequent events such as a bus crash that floods the ER with dozens of patients at once.

A much more practical solution would be to try to replicate the behavior of the ER on a computer (i.e. create a type of mathematical model called a simulation) where the numbers of doctors and nurses can be easily changed. Each mixture can be simulated for a long period of time under many different situations and at low cost. Plus, nobody dies.

1.3 The Process

We will illustrate the process of mathematical modeling with an example of modeling the number of bacteria in a Petri dish as described in Section 1.2.

Step 1: State the question to be answered

In many situations, this step is almost trivial; in others it is the most difficult part of the process. The question should be narrow enough to make the problem manageable, but not too narrow so that the problem is trivial. Initially we may want to focus on a narrow question, and then use the knowledge gained to broaden the question at a later time. The question should also be stated in precise mathematical terms so it can easily be translated into mathematical notation.

In this example we will answer the question "How long will it be until the number of bacteria in the dish reaches 600?"

Step 2: Select the modeling approach

In this step we determine the form of the model. In some situations this is easy to do; in others we may have several reasonable choices. Making the right choice requires at least some knowledge of all the possibilities. It also depends on the nature of the assumptions being made.

Often times this step begins with some simple observations. Note that we started with 500 bacteria. After 1 day, it increased by 25, which is 5% of 500. After a second day it increased by 26, which is approximately 5% of 525. The growth rate (or change per day) appears to be relatively constant. This suggests a simple relationship between the populations on consecutive days:

$$\text{Population on one day} = \text{Population on the previous day} + 5\%$$

This relationship indicates that we may be able to derive a simple equation to model the population.

Step 3: Define variables and parameters

Variables are quantities that could change within a problem. *Parameters* are quantities that are constant within a problem, but that could change between problems of the same type. The first part on this step is to determine what variables and parameters are involved. This may be simple and obvious, or very complicated. Often times there are potentially hundreds of quantities involved. To make the model manageable, we need to make assumptions as to which are the most important and which can be ignored. At a later time we could add additional variables and parameters to refine the model.

In this example, variables include:

1. Time
2. Population
3. Temperature
4. Amount of food present
5. Amount of available space in the dish

These are all values that change as the population grows. Since the initial observation did not give any information on temperature, food, or space, we will ignore these variables and focus on only time and population.

Possible parameters to consider include:

1. The initial population
2. Growth rate (we will assume this is constant)
3. Size of the dish
4. Initial amount of food

These are all values that are constant once we put the bacteria in the dish and allow them to grow. But if we consider a different dish with a different population of bacteria, they could change. Again, since we don't know anything about the size of the dish or the amount of food, we will ignore these parameters.

The second part of this step is to choose symbols to represent the variables and parameters. For this example, let

$$\begin{aligned} n &= \text{time in days from the present}\,(n = 0, 1, ...) \\ r &= \text{the growth rate (in decimal form)} \\ a_n &= \text{the population at the beginning of day } n \\ a_0 &= \text{the initial population} \end{aligned}$$

Step 4: State the assumptions

Making assumptions is an essential aspect of creating a valid and manageable model. Assumptions fall into many different categories. Some are used to simplify the model, such as those used to select the important variables. Some are needed to define relationships between the variables because the precise relationships are not known. Others are needed to determine the values of parameters when the exact values are not known.

Clearly stating the assumptions is an important part of interpreting and presenting the results. The results of a model are only as valid as the underlying assumptions. If the assumptions are unreasonable, then the conclusion will be unreasonable regardless of the precision of the mathematical analysis.

In this problem, we have already chosen to ignore temperature, size of the dish, and many other possible variables and parameters. This is a simplification. Furthermore, we will assume that the population growth is constant (i.e. the population will increase 5% each day).

Step 5: Formulate the model

This is where the "mathematics" starts. We have observed that the number of bacteria on day 1 is equal to the number on day 0 plus 5%. The number on day 2 is equal to the number on day 1 plus 5%, etc. In mathematical notation using our variables and parameters, we have

$$\begin{aligned} a_1 &= a_0 + r\,a_0 = (1+r)\,a_0 \\ a_2 &= a_1 + r\,a_1 = (1+r)\,a_1 \\ &\vdots \\ a_{n+1} &= a_n + r\,a_n = (1+r)\,a_n \end{aligned}$$

This forms a recursively defined sequence. To form an explicit description of a_n in terms of n, note that

$$\begin{aligned} a_1 &= (1+r)\,a_0 \\ a_2 &= (1+r)\,a_1 = (1+r)(1+r)\,a_0 = (1+r)^2 a_0 \\ a_3 &= (1+r)\,a_2 = (1+r)(1+r)^2\,a_0 = (1+r)^3 a_0 \\ &\vdots \\ a_n &= (1+r)^n a_0 \end{aligned}$$

This last equation is our model.

Step 6: Solve the model and state the solution

Here we use the term "solve" loosely. Solving a model may involve solving a single equation, as in this example, it may involve constructing a graph and qualitatively describing its behavior, or it may involve running a simulation several hundred times and summarizing the resulting data. The meaning of the term solve is relative to the type of model.

In this example, the question is "when will the population be 600?" In terms of our variables, this can be stated as "find n such that $a_n = 600$." This yields the equation

$$600 = (1 + 0.05)^n 500$$

Solving this equation using logarithms yields $n \approx 3.7$. This means that at the beginning of the fourth day we will have over 600 bacteria. This is our solution.

Often times the results of a model are used to guide decisions. In many practical situations, such as in business or the military, the person doing the modeling is not the final decision maker. The final decision maker is a CEO or officer who is not a mathematician. Therefore, the solution should be stated in as non-technical language as reasonably possible.

Step 7: Verify the model

Verification is necessary to test the reasonableness of our assumptions. Typically we verify a model by comparing it to some real world data. Let's suppose we let the bacteria grow for a total of 7 days and collect the data in Table 1.1. Next to the actual observed populations are the populations predicted by the model.

We see that on day 4, the actual population is just below 600. Even though the predicted population does not equal the actual population, our solution of 4 days is reasonable.

Note that on days 5 and 6, the actual and predicted populations differ considerably. The data indicates that the growth rate slows down. This means that our assumption of a constant growth rate is incorrect.

TABLE 1.1

Day	Actual Population	Predicted Population
0	500	500
1	525	525
2	551	551
3	575	579
4	598	608
5	610	638
6	620	670

Our model is accurate up to day 4, but inaccurate for later days. This example illustrates that we must be very cautious about using data from the past to make predictions about the future.

Step 8: Refine the model

Refine means to improve the model in some way. One way to do this is to add variables that we chose to ignore in step 3 to make a more accurate model. Another way is to generalize the model so that it can be used to solve other similar problems. Either one will require us to repeat steps 3 – 7 to some degree.

We have already noted that the data indicates the growth rate slows down over time. This could be a result of diminishing food supplies or room to grow. These are two variables we chose to ignore.

One possible refinement is to redo the model incorporating these two variables. This would require additional observations and data to determine how these variables are related to the other variables. Another possible refinement is to use the available data to model a decreasing growth rate. We will illustrate how to do this in Chapter 2.

A simple diagram that illustrates the basic process of mathematical modeling is given in Figure 1.1. The process begins in the upper left-hand corner with observations (or data). From this we get the basic problem we want to solve. We then make assumptions, construct the model, solve it using appropriate mathematical tools, and obtain a mathematical result. Then we must interpret the mathematical result in light of the assumptions to make our conclusions. We then verify the model using more observations.

This figure also illustrates the cyclic nature of mathematical modeling. We rarely stop once we answer the original question. We continually repeat the process, to some extent, to test, refine, and implement the model.

The right half of this diagram is done in the "math world" and the left half is done in the "real world." In the math world, we use the absolute certainty of mathematics. The real world contains no such certainties. Making assumptions and interpreting are necessary steps to move between these two worlds.

1.4 Assumptions

Every model is based on some set of assumptions. Sometimes those assumptions are rather trivial and obvious, but most times they are significant enough to potentially affect the validity of the model. We will never be able to describe each component of a real world

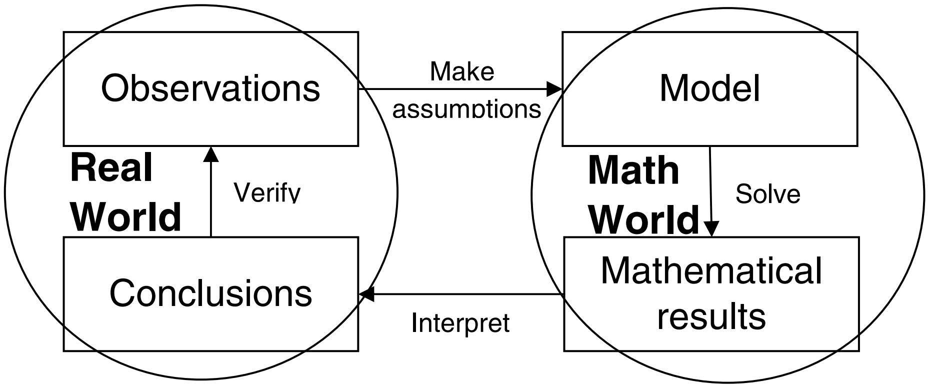


FIGURE 1.1

system exactly. Assumptions are needed to fill in these gaps, and whenever possible, the reasonableness of assumptions should be tested. In fact, assumptions are so important that any mention of a model should include the assumptions behind it.

Here are a few examples of well–known models and some of their underlying assumptions

Example 1.4.1 (Range of an Electric Car)
When considering the purchase of an electric car, the first question most consumers ask is, "what's the range?" The exact answer depends on many factors including the battery size and condition, driving style, vehicle speed, wind speed and direction, temperature, and elevation change. Any numeric answer to this question must be accompanied with many assumptions about these factors. □

Example 1.4.2 (Carbon–14 Dating)
Carbon–14 dating methods are used to date organic material, such as a piece of bone, found at archaeological sites. The underlying mathematical model requires knowledge of the proportion of Carbon–14 originally present in the sample. Obviously we cannot measure this precisely, so we assume that this proportion is the same as in a modern bone. This assumption can't be tested directly, but if a date can be confirmed via independent means, it would be an indication that the assumption is correct. □

Example 1.4.3 (Newtonian Mechanics)
Newton's second law says that the force exerted on an object is equal to its rest mass times the acceleration, or $F = ma$. This was assumed to hold at any velocity. In the 20^{th} century, it was discovered that for speeds approaching the speed of light, this rest mass must be replaced with the relativistic mass, which is larger. □

Example 1.4.4 (Relativity)
Einstein's theory of special relativity is really a mathematical model. It is based on several postulates (a type of assumption), one of which is that the speed of light in a vacuum is constant to any observer in an inertial frame of reference. These assumptions cannot be tested in every circumstance imaginable, but the results of Einstein's theory can be tested. These results have been shown to be correct, so it suggests that the underlying assumptions are also correct. □

Exercises

1.4.1 Each part below describes a calculated number. Think about how this number was calculated and identify at least one assumption behind the calculation.

a. A consumer magazine reports that a laundry detergent costs \$0.31 per load.

b. An exercise bike displays the number of calories burned.

c. A jogging pedometer measures the distance jogged.

d. A dashboard display in a car shows that it can travel 220 miles on the remaining fuel in the tank.

e. A small business owner predicts that his company will spend \$500,000 on phone bills next year.

f. A cooking magazine lists one ingredient for a recipe as 1-2 cups of shredded cheddar cheese. It then claims that each serving has 320 calories.

1.4.2 Identify at least two possible variables and two parameters involved in each of the following models. Clearly identify which is a variable and which is a parameter.

a. A biologist wants to model the population of foxes and rabbits in a forest over a period of time.

b. A consumer wants to model the monthly balance in his credit card.

c. A public health researcher wants to model the amount of alcohol in the bloodstream of a college student during an evening of partying.

d. An ecologist wants to model the amount of pollution in a lake over a period of time.

e. A market researcher wants to model the number of viewers of a TV show over the span of the season.

1.4.3 To predict the time at which the bacteria population in Section 1.3 reached 600, we solved the equation $600 = (1 + 0.05)^n 500$ and concluded that at the beginning of day 4 we will have over 600.

a. Show how this equation was solved for n.

b. A student argues that we should be more precise in our conclusion. She argues that we should say, "on day 3.736850652 there will be exactly 600 bacteria" because this is the value given by her calculator. How would you explain to her that such a level of precision is not appropriate?

1.4.4 A cashier at a supermarket can checkout, on average, 3 customers a minute (this is called the *service rate*). Customers arrive, on average, 2 a minute (this is called the *arrival rate*). The manager figures that since the service rate is greater than the arrival rate, customers will never have to wait in line. What assumptions is the manager making? Do these assumptions seem reasonable? What does this say about the validity of the conclusion?

1.4.5 A college student plans to ask 100 different girls for a date. He calculates the number who will say yes with the following reasoning:

> Since every girl can say yes or no, exactly half will say yes. Since half of 100 is 50, exactly 50 girls will say yes.

What, if anything, is faulty with this model? Explain.

1.4.6 A model for the population of the United States (in millions) for the years 1800-1960 is

$$P(t) = 0.0063t^2 + 0.0719t + 5.0747$$

where t is the number of years since 1800. A student predicts that the population in the year 2050 will be $P(250) \approx 416.8$ million. What assumption is the student making when doing this calculation? Do you think this assumption is reasonable? What does this say about the validity of the prediction?

1.4.7 Watch the Numberphile video which claims to give a "proof" that $1 + 2 + 3 + \cdots = -1/12$ (https://www.youtube.com/watch?v=w-I6XTVZXww or Google "numberphile sum of natural numbers"). Believe it or not, there are mathematically valid and rigorous ways of assigning the number $-1/12$ to the sum $1 + 2 + 3 + \cdots$, but we will think about the errors made in this proof from a Calculus II perspective.

a. At about the 2:20 mark, the presenter makes the claim (or we could call it the assumption) that $1 - 1 + 1 - 1 + \cdots = 1/2$. He does this by considering the sequence of partial sums:

$$1,\ 0,\ 1,\ \ldots$$

According to the Calculus II definition of convergent series, a series converges to a number S if and only if the sequence of partial sums converges to the number S. Does this sequence of partial sums converge to 1/2? What does this mean about the validity of the assumption that $1 - 1 + 1 - 1 + \cdots = 1/2$?

b. The rest of the proof is based on the assumption that $1 - 1 + 1 - 1 + \cdots = 1/2$. Without considering the details of the rest of the proof, what can we say about the validity of the final conclusion that $1 + 2 + 3 + \cdots = -1/12$?

For Further Reading

- For a further discussion of the modeling process and related issues, see Frank. R. Giordano, M. D. Weir, and W. P. Fox, *A First Course in Mathematical Modeling*, Third ed., Thomson Brooks/Cole, 2003, pg. 52–63.
- For more information on modeling methodology, tips on writing reports, and numerous example problems, see Dilwyn Edwards and Mike Hamson, *Guide to Mathematical Modelling*, CRC Press, 1990.
- For a discussion of the benefits of mathematical modeling, see Lucas, William F., The Impact and Benefits of Mathematical Modeling, in *Applied Mathematical Modeling* (D.R Shier and K.T. Wallenius eds.), Chapman and Hall/CRC, 1999, pg. 1–25, and the included references.

- For a lengthy discussion of how assumptions impact environmental models, see Orrin H. Pilkey and Linda Pilkey–Jarvis, *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future*, Columbia University Press, 2007.

- For a brief discussion of some of the assumptions behind global–warming models, see K. K. Tung, *Topics in Mathematical Modeling*, Princeton University Press, 2007, pg. 146–157.

2

Proportionality and Geometric Similarity

Chapter Objectives

- Introduce the basics of graphing with Excel
- Use *proportionality* and *geometric similarity* as simplifying assumptions in the modeling process
- Fit straight lines to data
- Use data to find constants of proportionality
- Use geometric similarity to construct models
- Fit various types of models to a set of data to predict values
- Use the coefficient of determination to assess how well a model fits a set of data

2.1 Introduction

One of the first steps in modeling is to make simplifying assumptions, and one of the simplest assumptions is that one variable is simply a constant multiple of the other. This type of relationship is called *proportionality.*

Definition 2.1.1. The variable y is said to be *proportional* to the variable x if there exists a nonzero constant c (called the *constant of proportionality*) such that

$$y = cx \tag{2.1}$$

The expression $y \propto x$ is used to indicate that y is proportional to x. Often times the phrase "directly proportional" is used to describe this type of relationship. We must point out that a proportionality relationship between two variables does not mean that one variable *causes* the other.

Note that if $y = cx$, then $x = \frac{1}{c}y$ so that x is proportional to y. Thus if y is proportional to x, we immediately know that x is proportional to y (i.e. a proportionality relationship is symmetric). For this reason, if y is proportional to x, we simply say that x and y are proportional.

Graphically, $y \propto x$ means that a graph of y vs. x (y on the vertical axis and x on the horizontal axis), should form a straight line through the origin. The constant of proportionality is the slope of this line.

2.2 Using Data

One famous proportionality relationship is *Hooke's Law* which relates the force applied to a spring to the distance it is stretched or compressed. Hooke's Law simply states that

$$d = kF \tag{2.2}$$

where F is the force applied to a spring, d is the distance stretched or compressed, and k is a constant related to the stiffness of the spring.

Example 2.2.1 (Bucket on a Spring)
Suppose that we hang a bucket from a spring, fill the bucket with varying amounts of sand, and measure the distance the spring is stretched. The results are recorded in Table 2.1.

TABLE 2.1

Weight (newtons)	Distance (cm)
5	1.02
10	1.86
15	3.00
20	3.94
25	4.95
30	5.82
35	6.95
40	7.80

We will plot this data to (1) verify that Hooke's law holds for this spring, and (2) find the constant of proportionality. This data comes from the real world, so it is subject to errors and uncertainty. Therefore, we cannot expect the data to lie perfectly on a straight line as predicted by the idealized Hooke's law. However, the data should lie very near a straight line. The slope of this line is the constant of proportionality.

In this and following examples we will provide instructions for producing worksheets that allow us to explore the mathematical models we will be building throughout the rest of the text. As you become more familiar with building worksheets you will be expected to be able to construct worksheets for models with fewer instructions.

1. Rename a blank worksheet "**Spring**." Format the worksheet as in Figure 2.1 and enter the rest of the data from Table 2.1.

	A	B
1	**Weight**	**Distance**
2	5	1.02
3	10	1.86
4	15	3

FIGURE 2.1

2. Follow these steps to form a scatter plot as in Figure 2.2:

   (a) Highlight the data (including the headings).

   (b) Click on the **Insert** tab.

(c) In the **Charts** section of the ribbon, click on **Scatter**, and select the type in the upper left-hand corner called **Scatter only with Markers**. This will create a chart similar to that in Figure 2.2. Next we will format it.

(d) Left-click on the legend in the graph, and press **Delete**. Do the same to the vertical lines in the chart and the chart title.

(e) To add axis titles, click on the **Layout** tab. Select **Axis Titles** in the **Labels** section of the ribbon. Select **Primary Horizontal Axis Title** and then **Title Below Axis**. Type the name of the horizontal axis and press **Enter**. For the **Primary Vertical Axis Title**, select **Rotated Title**.

(f) To change the min and max values on the axes, right-click on a number on one of the axes and select **Format Axis**. Next to **Minimum:** and **Maximum:** select **Fixed** and enter the appropriate value. Press **Close**. Do the same to the other axis.

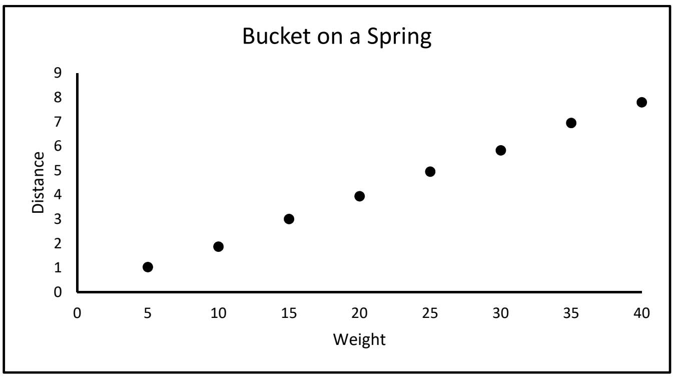


FIGURE 2.2

3. Next we need to estimate the slope of a line that "fits" these points. If we were doing this with paper and pencil, we would use a ruler or straight edge to draw a line through the origin that goes close to each data point and then find the slope. We will do the electronic version of this in Excel. Format the spreadsheet as in Figure 2.3.

	D	E	F
1	**Line**		
2	**x**	**y**	**Slope**
3	0	0	0.1
4	40	=D4*F3	

FIGURE 2.3

4. Next we will add these points in Figure 2.3 to the graph.

(a) Right-click anywhere on the graph and choose **Select Data...** Press **Add** and format the window as in Figure 2.4. Press **OK** twice. This adds the two points to the graph.

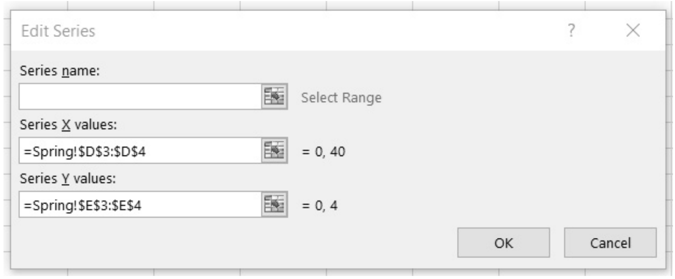


FIGURE 2.4

(b) To connect these two points with a straight line, right-click on one of the points and select **Format Data Series**. Select **Line Color**, choose **Solid Line**, and the desired color. Press **Close**. Your graph should look similar to Figure 2.5.

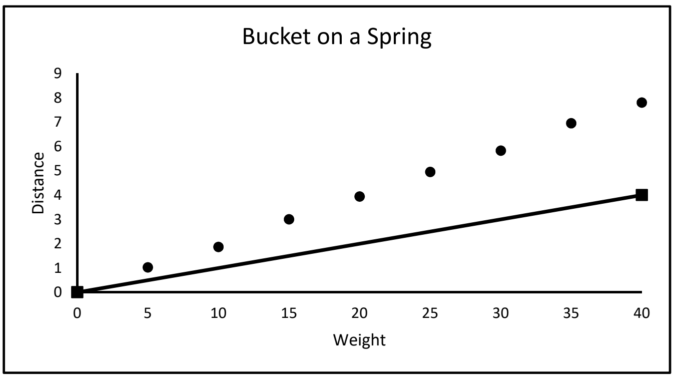


FIGURE 2.5

5. Change the value of the slope in cell **F3**. Notice that the graph of the line automatically changes.

6. Next we will "move" the line so it better fits the data and find the slope of this line. Follow these steps:

   (a) Insert a scroll bar by selecting the **Developer** tab. In the **Controls** section of the ribbon, select **Insert**. Under **ActiveX Controls**, select the **Scroll Bar** on the right-hand side of the first row and draw a horizontal bar anywhere on the worksheet.

   (b) Right-click on the scroll bar and select **Properties**. Set the **LinkedCell** to **G2** and the **Max** to 1000. (For more details on creating scroll bars, see Appendix A.4.)

   (c) Add the formula in Figure 2.6 to calculate the slope of the line using the scroll bar. This allows us to vary the slope between 0 and 1 in increments of 0.001.

	F
3	=G2/1000

FIGURE 2.6

(d) Slide the scroll bar back and forth until the line "fits" the data.

(e) Your graph should look similar to Figure 2.7 and the value of the slope in cell **F3** should be approximately 0.195.

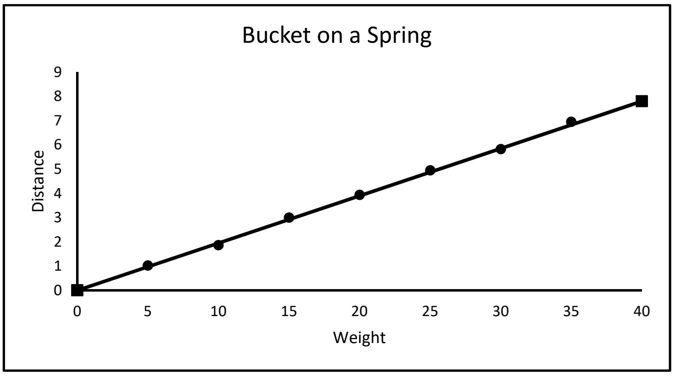


FIGURE 2.7

Conclusion: We see that the data do indeed lie very near a straight line, so Hooke's law is verified (at least in this example). The constant of proportionality is approximately 0.195. Therefore, if we know the amount of weight w on the spring we can approximate the distance the spring has been stretched, d, by

$$d = 0.195w$$

Likewise, if we know the distance the spring has been stretched we can approximate the amount of weight on the spring by

$$w = \frac{1}{0.195}d \approx 5.128\,d$$

□

Example 2.2.2 (No Correlation)
A mathematics professor claims that the grade on the final exam is directly proportional to the amount of time spent studying. To test this claim, he asked each student how much time he/she studied and recorded it along with the grade. A graph of the collected data is shown in Figure 2.8.

Notice that the points do not lie close to a straight line through the origin, so the grade is not directly proportional to time spent studying. In fact, the points appear to be randomly scattered. This indicates that there is not much of a relationship at all between the two variables. □

Example 2.2.3 (Boyle's Law)
Another well-known proportionality relationship is Boyle's law which relates the pressure of a gas to its volume at a constant temperature,

$$V = \frac{k}{P}$$

where V denotes the volume of the gas, P denotes its pressure, and k is a constant. To test Boyle's law, a student measures the pressure of a gas at different volumes while keeping the temperature of the gas constant. The resulting data is shown in Table 2.2.

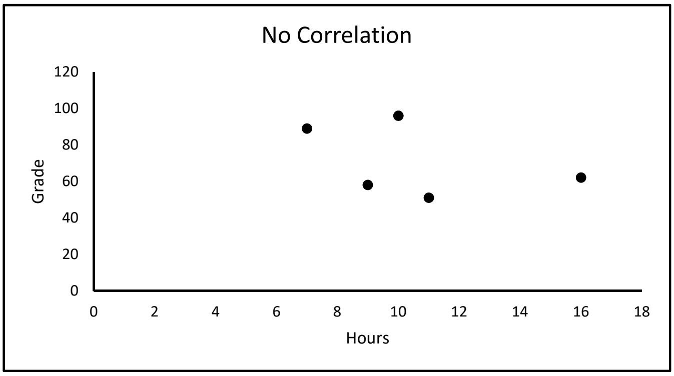


FIGURE 2.8

TABLE 2.2

Volume (V)	50	45	40	35	30	25	20	15	10
Pressure (P)	27.24	30.36	34.01	38.73	45.5	54.35	68.03	90.53	135.73

We will use this data to verify Boyle's law and find the constant of proportionality for this gas. Note that the law does not say that V is proportional to P. It says that V is proportional to $\frac{1}{P}$. Therefore we will plot V vs. $\frac{1}{P}$ and fit a straight line through the origin to this *transformed* data.

1. Rename a blank worksheet "**Boyle**" and format it as in Figure 2.9. Enter the data from Table 2.2 in columns **A** and **C**. Left-click on cell **B2**. Double-click on the small box in the lower right-hand corner of the border. The formula in **B2** will be copied down to row 10.

	A	B	C
1	**P**	**1/P**	**V**
2	27.24	=1/A2	50

FIGURE 2.9

2. Use the transformed data in column **B** and the original data in column **C** to form a scatter plot as in Figure 2.10.

3. We see in Figure 2.10 that the data do lie very near a straight line as expected. To estimate the slope of this line, we will simply pick one of the data points and calculate the slope of the line through the origin and the point. Let's choose the right-most point $(0.036711, 50)$. The slope of a line through this point and the origin is

$$\text{slope} = \frac{50 - 0}{0.036711 - 0} \approx 1362$$

4. To examine how well a line through the origin with this slope fits the data, format the spreadsheet as in Figure 2.11.

5. Add the line to the scatter plot as in Example 2.2.1. Your graph should look similar to Figure 2.12.

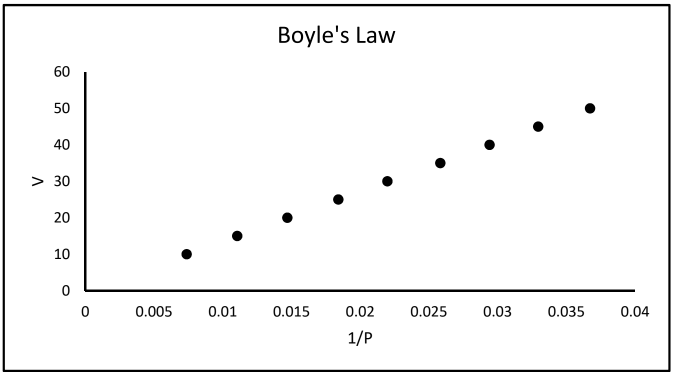


FIGURE 2.10

	E	F	G
1	**Line**		
2	**x**	**y**	**Slope**
3	0	0	1362
4	0.036711	=G3*E4	

FIGURE 2.11

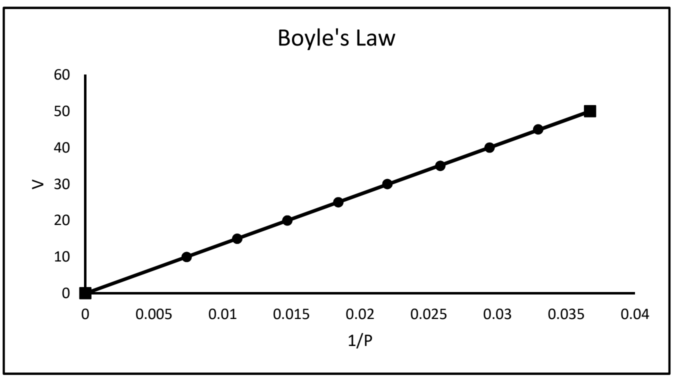


FIGURE 2.12

Conclusion: We see that V is clearly proportional to $\frac{1}{P}$ and the constant of proportionality is near 1362. Note that the constant of proportionality may be different for a different gas. □

Exercises

2.2.1 For each of the data sets below, determine if it is reasonable to assume that y is proportional to x. If it is, approximate the constant of proportionality. If it is not, describe why this assumption is not reasonable.

a.

x	1	1.1	1.2	1.3	1.4	1.5	1.6	1.7
y	1	1.21	1.44	1.69	1.96	2.25	2.56	2.89

b.

x	1	5	7	2	10	12	3	6
y	0.79	10.89	14.37	5.75	23.36	26.29	3.76	16.12

c.

x	2	6	9	15	7	25	39	4
y	26	20	18	26	6	19	20	13

2.2.2 For NBA teams, it seems reasonable to assume that if the percentage of 3-point shots made increases, the winning percentage should also increase. That is, it seems reasonable to assume that 3-point percentage is proportional to winning percentage. The table below shows the 3-point and winning percentages of 22 randomly selected NBA teams between the years 1996 and 2016 (data collected by Philip Yox and Seth Euken, 2016). Create a graph of winning percentage vs. 3-point percentage. Based on these data, does the assumption of proportionality seem reasonable? Briefly explain.

3 pt %	Winning %	3 pt %	Winning %
0.358	0.500	0.332	0.512
0.355	0.488	0.345	0.451
0.341	0.476	0.347	0.439
0.381	0.329	0.354	0.402
0.347	0.605	0.381	0.646
0.342	0.427	0.379	0.634
0.368	0.512	0.361	0.740
0.356	0.561	0.339	0.512
0.349	0.476	0.364	0.744
0.350	0.402	0.339	0.488
0.362	0.280	0.376	0.512

2.2.3 Suppose you drive your car on a perfectly flat road at a constant speed with no wind. In this case, the amount of fuel, y, (in gallons) needed is directly proportional to the distance traveled, x (in miles).

a. If the distance traveled increases, what can we say about the amount of fuel needed?

b. If the relationship is given by $y = 0.04x$ and x increases by 50 miles, how much does y increase?

c. Now, suppose it takes 12 gallons of fuel to travel 282 miles. Find the constant of proportionality.

d. In words, describe the meaning of this constant of proportionality.

2.2.4 A variable y is said to be *inversely* proportional to x if there exists a constant c such that $y = \frac{c}{x}$.

a. If y is inversely proportional to x, sketch what a graph of y vs. x would look like. What would a graph of y vs. $1/x$ look like?

b. If y is inversely proportional to x, show that x is inversely proportional to y.

c. If y is inversely proportional to x and x increases, what happens to the value of y?

2.2.5 For the data set below, determine if it is reasonable to assume that y is inversely proportional to x. If it is, approximate the constant of proportionality. If it is not, describe why this assumption is not reasonable.

x	1	1.2	1.4	1.6	1.8	2.0	2.2	2.4
y	6.85	6.21	4.24	4.32	3.92	3.18	2.93	2.96

2.2.6 Newton's Law of Universal Gravitation states that the force of attraction F between two objects with masses m_1 and m_2 is proportional to the product of the masses and inversely proportional to the square of the distance d between them. In mathematical notation,

$$F = k\frac{m_1 m_2}{d^2}$$

For two given objects, if m_1 and m_2 are constant, we may combine them with the constant of proportionality k to describe the relationship by

$$F = \frac{C}{d^2}$$

where C is a constant. If one of the two objects is a planet, the distance d is the distance from the center of the planet to the second object.

a. The radius of the Earth is approximately 4,000 miles. A satellite weighs 15 tons on the surface of the Earth (i.e. the force of attraction between the Earth and the satellite is 15 tons at the surface of the Earth). Use this information to calculate the constant of proportionality C.

b. Find the weight of the satellite (in tons) at an altitude of 500 miles above the surface of the Earth.

c. Use Excel to graph the weight of the satellite vs. altitude for values of altitude between 0 and 4,000 miles (ignore the effects of all other celestial bodies like the Moon).

2.2.7 Kepler's third law of planetary motion states that the cube of the semi-major axis of the orbit of a planet is directly proportional to the square of its orbital period. The table below shows the semi-major axis l (in astronomical units) and the orbital period p (in terrestrial years) of the planets in our solar system (data from the Wolfram Alpha Knowledgebase, 2019). Use this data to determine if Kepler's third law appears valid. If so, estimate the constant of proportionality.

Planet	l	p
Mercury	0.3871	0.2408
Venus	0.7233	0.6152
Earth	1.0000	1.0000
Mars	1.5237	1.8808
Jupiter	5.2034	11.8624
Saturn	9.5371	29.5015
Uranus	19.1913	84.0154
Neptune	30.0696	164.7884

2.2.8 Determine which of the following models best fits the data below by transforming the data appropriately and fitting a straight line to the transformed data. Find the constant of proportionality for this model. Explain why your choice is the best model.

$$y \propto x, \quad y \propto \frac{1}{x}, \quad y \propto x^2, \quad y \propto \sqrt{x}, \quad y \propto \frac{1}{x^3}$$

$\boldsymbol{x}$	0.5	0.7	0.9	1.2	1.5
$\boldsymbol{y}$	7.8	3.5	2.2	0.85	0.36

2.2.9 An online braking distance calculator allows the user to input the speed of a car (in miles per hour) and outputs the braking deceleration distance (in feet) (see https://nacto.org/docs/usdg/vehicle_stopping_distance_and_time_upenn.pdf). The table below gives several different speeds x and their resulting distances y. Determine which, if any, of the models in Exercise 2.2.8 were used to calculate the distances. Briefly explain your choice.

$\boldsymbol{x}$	10	20	30	40	50	60	70	80	90
$\boldsymbol{y}$	5	19	43	76	119	172	234	305	386

2.2.10 When a satellite has been orbiting a planet for a long time (such as the Moon orbiting the Earth), the satellite can become in *tidal lock* where the satellite's orbital period equals its rotational period. The Moon is in tidal lock with the Earth explaining why we can only see one side of the Moon. Assuming a constant satellite mass and orbital distance, a simple model for the amount of time needed for the satellite to become in orbital lock, t_l, is

$$t_l \propto \frac{1}{m_p^2}$$

where m_p is the mass of the planet. The table below shows the mass of four planets and the value of t_l calculated using this model (data collected by Joshua Hendrickson, 2019). Use the data to estimate the constant of proportionality in the model.

Planet	$\boldsymbol{m_p}$ **($\times 10^{24}$ kg)**	$\boldsymbol{t_l}$ **($\times 10^{7}$ years)**
Jupiter	1898	0.003812
Saturn	568	0.042562
Uranus	86.8	1.822561
Neptune	102	1.31984

2.2.11 In Figures 2.7 and 2.12 we drew the straight line fit to the data (or the transformed data) so that the line went through the origin.

a. If $y \propto x$, explain why the straight line fit to the data *should* go through the origin.

b. If $y \propto u(x)$, where $u(x)$ is some function of x, explain why the straight line fit to graph of the transformed data (y vs. $u(x)$) *should* go through the origin.

2.2.12 A common assumption regarding the metabolism of a drug in the blood stream is that the rate of change of the concentration of the drug is proportional to the concentration. That is, if $A(t)$ is the concentration of the drug at time t, then $dA/dt \propto A$. The data below shows the concentration of a certain drug in a person's blood stream, A, taken t hours after the drug was administered.

t	0	1.2	2.4	3.5	4.8	6.4	9.4	10.5	12.4	15.6	18.8
A	10.2	8.6	7.3	6.2	5.2	4.2	2.7	2.3	1.8	1.2	0.7

a. The change in the concentration at hour t is approximately

$$\frac{dA}{dt} \approx \frac{A(t+\Delta t) - A(t)}{\Delta t}$$

where Δt is some small change in t. Use the data to approximate dA/dt for the given values of t.

b. Graph dA/dt vs. A. Does it appear reasonable to assume that $dA/dt \propto A$? If so, approximately what is the constant of proportionality?

2.2.13 The kinetic energy (in J) of a moving object is $K_E = 0.5mv^2$ where m is the mass of the object (in kg) and v is the velocity (in m/s). Suppose a researcher measures the kinetic energy of an object at different velocities as shown below. Use the data to estimate the mass of the object by transforming the data appropriately and fitting a straight line through the origin to the transformed data.

v (m/s)	1.3	2.5	3.4	5.2	8.4	9.1	10.1	10.9
K_E ($\times 10^{-5}$ J)	0.132	0.502	0.891	2.22	5.63	6.41	8.15	9.84

2.2.14 Fit a model of the form $y \propto \sin(3x+1)/\left(x^2-2\right)$ to the data below. Approximate the constant of proportionality

x	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
y	-2.42	-2.54	-2.39	-2.14	-1.64	-0.98	-0.15	0.95	2.15	3.84

2.2.15 Prove each of the following properties of proportionality:

a. If $ab \propto ac$ and $a \neq 0$, then $b \propto c$.

b. If $a^m \propto ac$ and $a \neq 0$, then $a^{m-1} \propto c$.

c. If $a \propto c^m$, then $a^{1/m} \propto c$.

d. If $y \propto (x/z)$, $x \propto h^n$, and $z \propto h^m$ where n, $m > 0$, then $y \propto h^{n-m}$.

2.2.16 A snow-cone seller at a county fair wants to model the number of cones he will sell, C, in terms of the daily attendance a, the temperature T, the price p, and the number of other food vendors n. He makes the following assumptions:

a. C is directly proportional to a.

b. C is also directly proportional to the difference between T and 85 °F.

c. C is inversely proportional to p and n.

Derive a model for C consistent with these assumptions. For what values of T is this model valid?

2.3 Modeling with Proportionality

One important observation about a proportionality relationship is that if one of the variables increases, so does the other, and if one variable decreases, so does the other. This can be seen in Figure 2.2 and Figure 2.10. Whenever we encounter a situation where two variables increase or decrease at the same time, we should consider a proportionality relationship.

Example 2.3.1 (Population Growth)
In many situations involving populations, the larger the population, the faster it grows. This suggests a proportionality relationship between the population and the rate of growth. Table 2.3 shows the population (in thousands) of bacteria in a Petri dish at different points in time. The third column contains the change in population between time periods.

TABLE 2.3

Day n	**Actual Population** p_n	**Change in Population** $\Delta p_n = p_{n+1} - p_n$
0	10.3	6.9
1	17.2	9.8
2	27	18.3
3	45.3	34.9
4	80.2	45.1
5	125.3	50.9
6	176.2	79.4
7	255.6	

Observe that as n increases, p_n increases, and so does Δp_n. This suggests that p_n is proportional to Δp_n. A graph of Δp_n vs. p_n is shown in Figure 2.13.

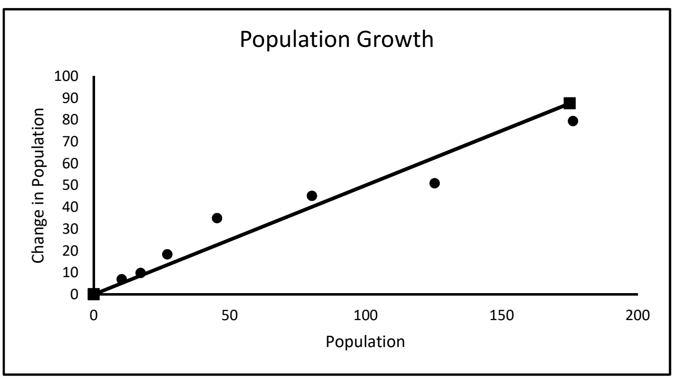


FIGURE 2.13

We see that the points do fall near a straight line through the origin, suggesting that proportionality is a reasonable assumption. The slope of this line is approximately 0.5. This gives a model that relates the population at one day, p_n, to the population at the next, p_{n+1}:

$$\Delta p_n = p_{n+1} - p_n = 0.5 p_n \quad \Rightarrow \quad p_{n+1} = 1.5 p_n$$

This model predicts that the population grows by about 50% each time period, which means the population will grow without bound. This seems unreasonable, so the model needs to be refined. We will do just this in Chapter 4.

□

Example 2.3.2 (Radioactive Decay)
One-half of the amount of a radioactive substance decays after each half-life. Radioactive Carbon-14 (^{14}C) has a half-life of 5715 years. If we start with 10g of ^{14}C, Table 2.4 shows the amount of material remaining after each half-life along with the rate of change between time periods.

TABLE 2.4

Time (years)	Amount	Change
0	10	5
5,715	5	2.5
11,430	2.5	1.25
17,145	1.25	0.625
22,860	0.625	

Note that as the amount of ^{14}C decreases, the rate at which it decreases also changes. This suggests a proportionality relationship between the amount of ^{14}C and the rate at which it decreases.

If we let $y(t)$ represent the amount of ^{14}C at time t, this proportionality relationship gives the differential equation

$$\frac{dy}{dt} = ky.$$

Solving this differential equation yields the exponential model for growth $y(t) = Ce^{kt}$ where C is the initial amount of material.

□

In the previous examples we used data of some type to suggest a proportionality relationship. Often times, as in the next example, we simply use logic to assume a proportionality relationship.

Example 2.3.3 (Free-falling Object)
An object in free–fall encounters two basic forces. The first is its weight due to gravity. The second is air resistance which slows the rate of fall. Air resistance is typically negligible at low speeds so it is often not modeled. If we ignore air resistance, then the only force acting on the object comes from acceleration due to gravity. This leads to the simple differential equation

$$\frac{dv}{dt} = g$$

where $v(t)$ = velocity of the object at time t and $g = 9.8$ m/sec^2 (the acceleration due to gravity). Solving this differential equation yields the model $v(t) = gt + v_0$ where v_0 is the initial velocity. This model predicts that the velocity grows without bound, which is inaccurate. Physics tells us that a free-falling object reaches a "terminal velocity" where the deceleration due to air resistance equals the acceleration due to gravity. At this point the object remains at a fairly constant velocity.

To refine this model for velocity, we can incorporate a simple model for air resistance. It seems reasonable to assume that as velocity increases, the force due to air resistance increases. This implies a proportionality relationship between velocity and the force due to air resistance:

$$\text{Force due to air resistance} = kv$$

This force acts upward on the object. There is also a force acting downward on the object due to its mass m:

$$\text{Downward force} = mg$$

Now, by Newton's second law,

$$F = ma = m\frac{dv}{dt} \tag{2.3}$$

Also,

$$F = \text{Downward Force - Upward Force} \tag{2.4}$$

Putting Equations (2.3) and (2.4) together we get,

$$m\frac{dv}{dt} = mg - kv \quad \Rightarrow \quad \frac{dv}{dt} + \frac{k}{m}v = g$$

Solving this last differential equation gives the model

$$v(t) = \frac{mg}{k}\left(1 - e^{-kt/m}\right)$$

Note that in this model,

$$\lim_{t\to\infty} v(t) = \frac{mg}{k}$$

This suggests a terminal velocity, so this model is more realistic.

□

Proportionality relationships satisfy the following transitive property:

Theorem 2.3.1. *If $a \propto b$ and $b \propto c$, then $a \propto c$*

Proof. By definition, $a \propto b$ and $b \propto c$ mean that $a = k_1 b$ and $b = k_2 c$ for some non-zero constants k_1 and k_2. So, substituting we get

$$a = k_1 k_2 c$$

but $k_1 k_2$ is a non-zero constant, so $a \propto c$ by definition. □

The next example illustrates an application of this property.

Example 2.3.4 (Work Done by a Train)
If a constant force is applied to an object moving it some distance, the work done is defined to be

$$\text{Work} = \text{Force} \times \text{Distance}$$

Suppose a train engine pulls a car along a flat stretch of track until it runs out of fuel, and assume that the force needed is constant. We want to model the total work done by the engine in terms of the amount of fuel it carries.

Since the force is constant, work is proportional to the distance pulled. Let W denote work and D denote distance. In standard notation,

$$W \propto D$$

Now, the total distance the train can pull the car is related to the amount of fuel. The more fuel, the farther it can pull the car. So distance pulled is proportional to the amount of fuel. If A denotes the amount of fuel, we have

$$D \propto A$$

Combining these two proportionality relationships, we arrive at the model:

$$W \propto A$$

Although we do not know the constant of proportionality, we can use this relationship to find relative values. For instance, if the constant of proportionality were 10 and the tank holds 1000 gallons, with a full tank the train could perform

$$W = 10\,(1000) = 10000$$

units of work. If the fuel tank were one–quarter full, it could perform

$$W = 10\,(250) = 2500$$

units of work, which is exactly one–quarter as much work as if the tank were full.

□

Exercises

2.3.1 A very simple assumption about the population of rabbits in a forest is that it grows at a rate proportional to the size of the population and the rabbits die at a rate proportional to the number of foxes in the forest. If p_n denotes the population of rabbits at time n, and F denotes the (constant) number of foxes, this assumption yields a model for the change in the rabbit population:

$$\Delta p_n = p_{n+1} - p_n = k_1 p_n - k_2 F$$

where $k_1,\, k_2 \geq 0$.

a. Solve this model for p_{n+1}.

b. If $p_0 = 500$, $k_1 = 0.15$, and $k_2 = 0.25$, algebraically find the values of F for which the population of rabbits is decreasing, for which it is increasing, and for which it is constant for all $n \geq 0$. (**Hint:** If the population is constant, $p_1 = p_0$.)

c. Create a spreadsheet to numerically and graphically support your answer above.

2.3.2 Refer back to Example 2.3.1. We will analyze a refined model. Assume that Δp_n is proportional to the product of the population and its difference from 621, that is,

$$\Delta p_n = p_{n+1} - p_n = k p_n (621 - p_n)$$

a. Use the data in Table 2.3 to test this assumption by plotting Δp_n vs. $p_n(621 - p_n)$. Use the graph to estimate the constant of proportionality.

b. Starting with $p_0 = 10.3$, use this refined model to predict the population for days 1 through 20.

c. Does this model seem more or less reasonable than the original one? Why or why not?

2.3.3 Refer again to Example 2.3.1 and consider the assumption that Δp_n is proportional to p_n^2. Use the data in Table 2.3 to test this assumption. Does this assumption appear to be reasonable? Why or why not?

2.3.4 Answer the following questions about proportionality:

a. If the amount of calories you consume at a meal is proportional to the length of the meal (in minutes), how does the calorie consumption of a meal lasting 5 minutes compare to that of a meal lasting 15 minutes (i.e. do you consume twice as many calories, half as many calories, etc.)?

b. If the surface area of a raindrop is proportional to the square of its radius, how does surface area of a raindrop with radius 2 compare to the area of a raindrop with radius 1?

c. If the volume of a potato is proportional to the cube of its length, how does the volume of potato A compare to potato B if the length of A is three times that of B?

d. If the maximum velocity of a car is inversely proportional to its mass, how much faster will a car go if you cut its mass in half?

2.3.5 A student wishes to determine how the size of a cloud affects the speed of a falling raindrop. She makes the following assumptions:

1. The speed is proportional to the weight of the raindrop.
2. The weight is proportional to the size, or volume, of the raindrop.
3. The size of the raindrop is proportional to the size of the cloud.

Use these assumptions to model the speed in terms of the size of the cloud. Don't forget to define the variables. Does this model seem reasonable? Why or why not?

2.3.6 Consider a charger for an electric vehicle. Assume the time needed to charge the battery from no charge to full charge is inversely proportional to the power of the charger. The power of the charger is the product of the voltage and the current supplied to the charger.

a. Model the charging time in terms of the voltage and current. Don't forget to define the variables.

b. If it takes 3 hours to charge the battery with a 240 volt, 72 amp charger, how long will it take to charge the battery with a 120 volt, 12 amp charger?

2.3.7 Prove the following statement: If a is inversely proportional to b and b is directly proportional to c, then a is inversely proportional to c.

2.3.8 Martin wants to determine how the amount of money he has in his wallet will affect his grade in Math Modeling class. Consider the following assumptions:

1. His grade is directly proportional to the amount of time studied.
2. The amount of time studied is directly proportional to the amount of free time he has.
3. The amount of free time he has is inversely proportional to the amount of time he spends going out with his girlfriend.
4. The amount of time he spends going out with his girlfriend is directly proportional to the amount of money he has in his wallet.

Use these assumptions to model his grade in terms of the amount of money in his wallet. Don't forget to define the variables. If he wants a high grade, should he have more or less money in his wallet?

2.3.9 The amount of money A in a savings account with an interest rate of $100\,r\%$ that is *continuously compounded* is

$$A = Pe^{rt}$$

where P is the amount of the initial deposit and t is time from the initial deposit.

a. A man puts \$5000 in an account with an interest rate of 0.4% a month so that $r = 0.004$ and t is measured in months. Find a model for A in terms of t.

b. Suppose at the same time the man opens the account, he goes on a diet and expects to lose 8 pounds a month. If w represents the total amount of weight lost by month t, find a simple model for w in terms of t.

c. Combine the models in part a. and b. to find a model for A in terms of w.

d. At the time the man has lost a total of 26 pounds, find the amount of money in the account.

e. When the account has \$5125, find the total amount of weight lost.

2.3.10 The variable y is said to be *exponentially proportional* to the variable x if there exist non-zero constants k and a such that

$$y = ka^x.$$

a. For the savings account in Exercise 2.3.9 part a., show that A is exponentially proportional to t and find the values of k and a.

b. Show that if y is exponentially proportional to x, and x is directly proportional to z, then y is exponentially proportional to z.

2.3.11 When modeling the spread of an infectious disease, the population is often divided into two categories: *susceptible* (those capable of getting the disease) and *infected* (those who have the disease). Let S_n and I_n denote the numbers of susceptible and infected people, respectively, at the beginning of time period n. One simple assumption is that the change in the number of infected people, $\Delta I_n = I_{n+1} - I_n$, is proportional to the product of S_n and I_n. Write down a model for ΔI_n based on this assumption and solve it for I_{n+1}. If we consider the product of S_n and I_n as modeling the "interaction" of susceptible and infected people, do you think the constant of proportionality is positive or negative? Explain.

2.3.12 Consider a forest containing foxes and rabbits where the foxes eat the rabbits. Let F_n and R_n denote the numbers of foxes and rabbits in the forest at the beginning of time period n. Ignoring all factors except the interaction of foxes and rabbits, a simple assumption about the change in the number of rabbits, $\Delta R_n = R_{n+1} - R_n$, is proportional to the product of R_n and F_n. Write down a model for ΔR_n based on this assumption and solve it for R_{n+1}. Do you think the constant of proportionality is positive or negative? Explain.

2.3.13 Translate each of the following descriptions of a physical scenario into a differential equation. Be sure to define the variables.

a. The rate of change of a population at time t is proportional to the population at time t.

b. The rate of change of the volume of a raindrop at time t as it falls through a cloud is proportional to the square of the volume at time t.

c. The velocity at time t of an object moving in a straight line is proportional to the third power of the object's distance from its starting point.

2.3.14 Let $A_1(t)$ and $A_2(t)$ denote the sizes of two opposing armies engaged in battle. One assumption we could make about these variables is that the rate of change of A_1 is proportional to A_2. Set up a differential equation that is consistent with this assumption. Ignoring reinforcements of the armies, is the constant of proportionality positive or negative? Explain.

2.3.15 When a hot cup of coffee is set on a desk, it initially cools very quickly. As its temperature decreases, it does not cool as quickly. This suggests a proportionality relationship. Newton's law of cooling states that the rate at which a hot object (such as a hot cup of coffee) cools is proportional to the difference in the room temperature and the temperature of the object (assuming the room temperature stays constant). Define variables for the temperature of the object and the room temperature and set up a differential equation to model Newton's law of cooling (do not solve the equation).

2.3.16 A simple assumption for the spread of a contagious disease is that the rate at which the number of infected individuals changes is proportional to the product of the total number infected and the number not yet infected. Assume that initially one resident carries the flu viruses into a dorm with n residents. Let $x(t)$ represent the number of residents that are infected at time t. Set up a differential equation to model the spread of the flu through the dorm (do not solve the equation).

2.3.17 To model the force due to air resistance in Example 2.3.3, we assumed that the force is proportional to velocity. Briefly explain how you might collect data and use the data to determine if this assumption is reasonable.

2.4 Fitting Straight Lines Analytically

As we have seen, modeling with proportionality often requires us to fit a straight line to a set of data. In earlier sections we used a graphical approach, which can be rather subjective. In this section we will look at different definitions of a "best fit" line and find formulas for the slope and y-intercept of a best-fit line in terms of the x- and y-coordinates of the data points. This will give an objective approach to fitting a straight line.

The first step is to define criteria for a good-fitting line. The line in the left graph in Figure 2.14 fits the data "better" than the line in the right graph. What's the difference between these two lines? There are probably many ways to answer this question.

We see that in the right graph, the line is very close to the right-most point, but further from the other two points than the line in the left graph. We might say that the line in the left graph is "closer" to the data points in general than the line in the right graph. The idea of minimizing the distance between the line and the points will form the basis of the definition of a best-fit line.

These distances (also called *errors*) are illustrated in Figure 2.15 with the dashed lines. If the coordinates of the points are given by

$$(x_i, y_i) \text{ for } i = 1, 2, \ldots, n$$

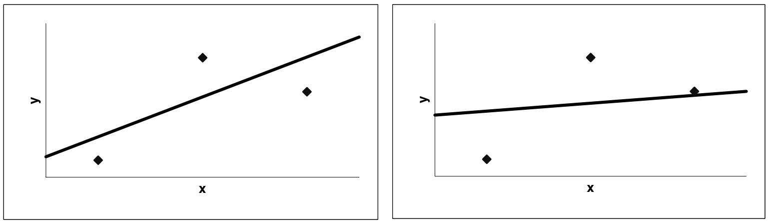


FIGURE 2.14

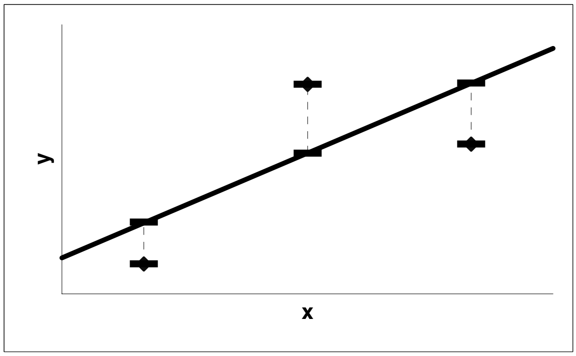


FIGURE 2.15

and the line is described by the function $f(x) = mx + b$, then the values of the distances are

$$|y_i - f(x_i)| \text{ for } i = 1, 2, \ldots, n.$$

There are many ways to define how the best-fit line minimizes these distances. One way, called *Chebyshev's* criterion, is based on the idea that the best-fit line should make the largest of these distances as small as possible. In more technical terms, this criterion says that the function $f(x) = mx+b$ giving the best-fit line is the one that minimizes the number

$$C = \text{Maximum of } \{|y_i - f(x_i)| \ : i = 1, 2, \ldots, n\}.$$

Our goal is to find formulas for the slope m and y-intercept b. Since we want to minimize something, we might think about using derivatives. Chebyshev's criterion makes logical sense, but it's not obvious how to take the derivative or use it to find simple formulas for m and b.

Another criterion is based on the idea that the best–fit line should minimize the sum of the distances. In mathematical notation, $f(x) = mx + b$ should minimize the number

$$A = \sum_{i=1}^{n} |y_i - f(x_i)|.$$

This criterion is also very logical, but the absolute values make the derivative difficult to calculate. To make the derivative simpler, we might consider getting rid of the absolute values in the above summation altogether and add the criterion that the sum must be non–negative. This, however, would make some of the terms in the summation positive and some negative. So, there might be some large positive values that cancel out some large negative values resulting in a small sum, but a poor-fitting line.

The most widely used criterion, called the *least–squares* criterion, uses squares rather than absolute values to make all the terms positive. In mathematical notation, this criterion says that the function $f(x) = mx + b$ should minimize the number

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2.$$

This number is called the *sum of squares.*

Example 2.4.1 (Illustrating the Least-Squares Criterion)
To illustrate the idea of finding a straight line that best fits a set of data and what this means in terms of the least-squares criterion, consider the data in columns **A** and **B** of Figure 2.16.

1. Rename a blank worksheet "**Least-Squares**" and format it as in Figure 2.16. Cells **B10** and **B11** store the slope and y-intercept of the line we will fit to the data. Note that we are not claiming that the values shown in the figure are the best values. We will experiment with changing these values.

	A	B	C	D	E
1	**x**	**y**	**f(x$_i$)**	**(y$_i$-f(x$_i$))^2**	**Sum of Squares**
2	0.8	2			
3	2.5	4.2			
4	3.5	3.5			
5	4.2	5.3			
6	5.8	4.5			
7	7.5	7.5			
8			**Fitted Line**		
9			**x**	**y**	
10	**m =**	0.5	0	=B11	
11	**b =**	2	8	=C11*B10+B11	

FIGURE 2.16

2. Create a scatterplot with the data in the range **A2:B7**. Add the data in the range **C10:D11** to the resulting graph and connect these two data points with a straight line. Format the graph to resemble Figure 2.17.

3. Note that this line in Figure 2.17 fits the data, but it appears that we could find a better-fitting line by adjusting the slope. To do this, add a scroll bar, set the **min** and **max** to 0 and 1000, and set the **linked cell** to **B12**. (See Appendix A for more information on adding scroll bars.) Add the formula in Figure 2.18 to change the value of m between 0 and 1 in an increment of 0.001.

4. As we change the value of m, observe that the line fits the data better for some values of m than others. To understand what this means in terms of the least-squares criterion, add the formulas in Figure 2.19. Copy the range **C2:D2** down to row 7.

Use the scroll bar to change the value of m. Observe that when the line graphically fits the data better, the sum of squares is smaller. The sum of squares takes its minimum value of about 3.4993 when $m = 0.634$. Therefore, when the y-intercept is $b = 2$, we would say the best-fit line has a slope of $m = 0.634$. □

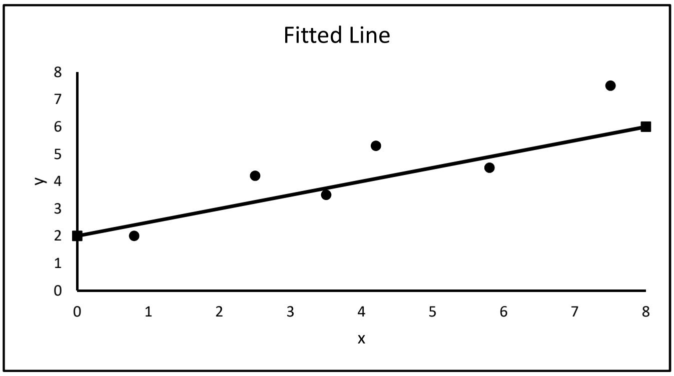


FIGURE 2.17

	B
10	=B12/1000

FIGURE 2.18

	C	D	E
1	**f(x_i)**	**(y_i-f(x_i))^2**	**Sum of Squares**
2	=$B$10*A2+$B$11	=(B2-C2)^2	=SUM(D2:D7)

FIGURE 2.19

In Example 2.4.1, we might be able to get an even better-fitting straight line (meaning get a smaller sum of squares) by adjusting the value of b. But if we adjust b, we would then need to adjust the value of m. This could get rather complicated. So instead, we use calculus to find formulas for m and b in terms of the data values that minimize the sum of squares.

We want to find values of m and b that minimize the quantity

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2.$$

Note that S is really a function of the variables m and b. A necessary condition for optimality is that the partial derivatives with respect to each of these variables is zero. This gives the equations

$$\frac{\partial S}{\partial m} = \sum 2(y_i - mx_i - b)(-x_i) = -2\sum (y_i - mx_i - b)\, x_i = 0 \tag{2.5}$$

$$\frac{\partial S}{\partial b} = \sum 2(y_i - mx_i - b)(-1) = -2\sum (y_i - mx_i - b) = 0 \tag{2.6}$$

where all summations are from 1 to n. Rewriting these equations and solving for m and b yield the formulas

$$m = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - \left(\sum x_i\right)^2} \quad \text{and} \quad b = \bar{y} - m\bar{x} \tag{2.7}$$

where $\bar{x} = \sum x/n$ and $\bar{y} = \sum y/n$ are the averages of the x- and y-values, respectively. The line with the slope and y-intercept calculated with these formulas is called the *least-squares regression line*.

Example 2.4.2 (Calculating a Least-Squares Regression Line)
Implementing formulas (2.7) is relatively easy. Add the formulas in Figure 2.20 to the worksheet **Least-Squares**.

	F	G	H
1	**Sums**		**Averages**
2	**x**	=SUM(A2:A7)	=AVERAGE(A2:A7)
3	**y**	=SUM(B2:B7)	=AVERAGE(B2:B7)
4	**x^2**	=SUMSQ(A2:A7)	
5	**xy**	=SUMPRODUCT(A2:A7,B2:B7)	
6			
7	**n =**	6	
8	**m =**	=(G7*G5-G2*G3)/(G7*G4-G2^2)	
9	**b =**	=H3-G8*H2	

FIGURE 2.20

These formulas give $m \approx 0.694$ and $b \approx 1.689$ so that the least-squares regression line is $y = 0.694x + 1.689$. If we change the values of m and b in cells **B10** and **B11** to 0.694 and 1.689, respectively, we see that the sum of squares is approximately 3.3699. This is a much smaller value than found in Example 2.4.1, illustrating the fact that the least-squares regression line minimizes the sum of squares.

Formulas (2.7) are also built into Excel. To implement these, add the formulas in Figure 2.21. Note that these built-in formulas give the same values of m and b as formulas (2.7).

	F	G
11		**Built-in Formulas**
12	**m =**	=SLOPE(B2:B7,A2:A7)
13	**b =**	=INTERCEPT(B2:B7,A2:A7)

FIGURE 2.21

There is yet another way to calculate the least-squares regression line with built-in formulas. On the scatterplot of the data created in step 1 of Example 2.4.1, right-click on one of the data points and select **Add Trendline...** Under **Trend/Regression Type** choose **Linear**. Check the box next to **Display Equation on chart** and press **Close**. The graph should resemble Figure 2.22.

□

Exercises

2.4.1 To determine if there is a relationship between shoe length and height of a person, the author had ten of his students measure their shoe length (to the nearest quarter of

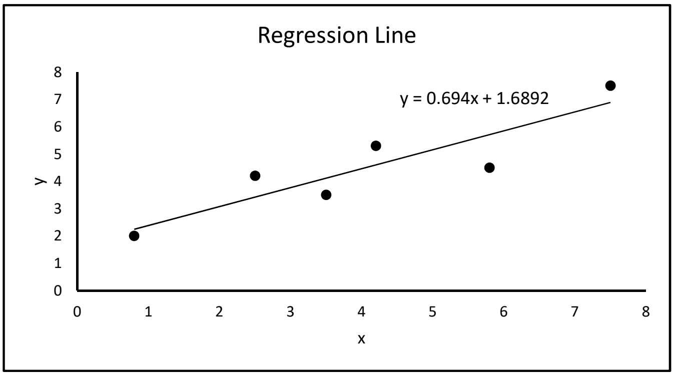


FIGURE 2.22

an inch) and height (to the nearest half inch). The resulting data are shown below. Graph Height vs. Shoe Length and fit a straight line. How well does this model fit the data?

Shoe Length	9	10	10.5	11	11.5	11.75	12	12.5	12.75	13
Height	62	64	64.5	69	70	73	72	75	74	77

2.4.2 The table below shows the price (in dollars) of several name-brand items and their generic equivalents at a large retail store (data collected by Amanda Schroeder, 2010).

Name	7.68	9.56	7.62	1.58	2.98	3.58	6.54	8.97	4.06	6.46
Generic	4.47	2.97	3.34	0.67	3.12	2.98	5.34	6.97	1.76	4.12

a. For each product, calculate the percent savings if the generic product is purchased instead of the name-brand.

b. Fit a straight line to the graph of percent savings vs. price of name-brand product. How well does this model fit the data?

2.4.3 Suppose a biologist records the number of pulses per second of the chirps of a cricket at different temperatures (in °F). The data collected is shown below.

Temperature	72	73	89	75	93	85	79	97	86	91
Pulses/sec	16	16.2	21.2	16.5	20	18	16.75	19.25	18.25	18.5

a. Fit a straight line to this data (where temperature is on the x-axis). How well does the model fit the data?

b. What is the slope of this line? What does the sign of the slope tell you about the relationship between pulses/sec and temperature?

2.4.4 The table below gives the average corn and soybean yields (in bushels per acre) on a Nebraska farm for the years 1995 - 2018 where Year 1 corresponds to 1995 and a − indicates a missing data value (data collected by Ross Briggs, 2018).

Year	Corn	Beans	Year	Corn	Beans
1	149.53	47.51	13	187.55	–
2	172.95	–	14	194.14	70.97
3	196.8	61.54	15	221.79	71.38
4	163.38	57.57	16	204.19	57.41
5	186.94	56.28	17	189.9	62.23
6	189.22	57.19	18	172.94	57.46
7	153.17	59.12	19	215.31	56.76
8	144.25	60	20	217.91	67.76
9	187.84	61.65	21	202.28	65.43
10	206.95	61.51	22	209.8	68.61
11	200.2	61.6	23	230.04	66.54
12	174.82	69.64	24	238.05	66.71

a. For each crop, fit a straight line to the graph of yield vs. year.

b. Comment on how well each line fits the set of data. Could we use the line for soybeans to predict the missing data values? Briefly explain.

c. What does the slope of each line mean about the quality of the farming practices used on this farm? Do they appear to be increasing or decreasing? Briefly explain.

2.4.5 When the temperature of a fixed volume of gas decreases, the pressure of the gas also decreases. The table below gives the pressures (in bars) and corresponding temperatures (in degrees C) of a gas at a fixed volume (data collected by Zachary Klatt, 2012).

Pressure	1.01	1.04	1.06	1.10	1.13	1.16	1.20
Temperature	12.6	20.0	30.0	40.0	50.0	60.0	70.0

a. Fit a straight line to the graph of temperature vs. pressure.

b. *Absolute zero* can be thought of as the temperature at which the pressure is 0. According to the model in part a., what is the predicted value of this temperature?

c. The accepted value of absolute zero is $-273\,^\circ$C. How close is your prediction?

2.4.6 The table below gives the win/loss percentage and other key statistics of 15 NCAA Division 1 women's basketball teams (data collected by Quinn Wragge, 2019).

a. Create graphs of win/loss percentage vs. each of the other statistics (one graph per statistic) and fit a straight line to the data.

b. Comment on how well each line fits the data. Which line appears to fit the data the best?

c. Comment on the sign of the slope of each line. Does the sign make sense? Briefly explain.

Win/Loss Percent	**Rebounds per Game**	**Turnovers per Game**	**Steals per Game**	**3 pt Shots Made/Game**	**Field Goal Percent**
93.8	48.8	13.1	7.4	3.3	51.5
52.6	32.89	13.4	4.6	10.3	42.9
89.5	39.89	14.2	10.2	4.8	45.3
77.8	42.67	14.6	8.1	8.9	45.2
68.8	36.19	13.8	8.8	7.7	41.2
100	45.53	14.3	5.6	8	46.3
50	40.82	15.7	7.8	7.2	42.9
94.7	44.17	14.5	8.7	4.1	51.7
68.4	38.26	13.7	7.4	7.5	46.6
70.6	46	16.8	9.7	6.2	42.9
83.3	43.28	16.4	7.6	5.6	45.5
70.6	38.41	15.9	9.6	6.2	43.3
94.1	42.13	11.8	7.3	7.9	49.1
33.3	38.29	16.7	7.4	5.4	39
85	38.84	13.4	9.5	8.2	44.4

2.4.7 Starting with Equations (2.5) and (2.6), derive the formula for m in Equation (2.7). (**Hint**: Multiply Equation (2.5) by $-n$ and (2.6) by $\sum x_i$. Add the resulting equations and solve this result for m.)

2.4.8 As we saw in Section 2.2, often times we want to fit a straight line that goes through the origin to a set of data. This means we want the y-intercept of the straight line to be 0. So the equation of the line is simply $y = mx$. In this case, the least–squares criterion says that we want to minimize

$$S = \sum_{i=1}^{n} (y_i - mx_i)^2.$$

a. Take the derivative of this equation with respect to m, set it equal to 0, and solve it for m to find a formula for m in terms of the x- and y-coordinates.

b. Implement this formula in Excel and use it to fit a straight line through the origin to the data in Table 2.1 on page 12. How does the slope of this line compare to the slope of the line found in Example 2.2.1?

c. This formula can be implemented by adding a trendline to the data and selecting **Set intercept =** under **Options**. Do this to the data and compare the slope of the trendline to the slope calculated by your formula. Are they indeed equal?

d. If the data are modeled by $y = mx$, then $y_i \approx mx_i$ for each i. Algebraically show that

$$m \approx \frac{\sum y_i}{\sum x_i}.$$

e. Use the data in Table 2.1 to demonstrate that the formula in part d. gives similar results as the formula in part a.

2.4.9 Generalize your result in 2.4.8. Suppose you have specified a value of b, say $b = b_0$, so that the model is now $y = mx + b_0$. In this case, the least-squares criterion says that we want to minimize

$$S = \sum_{i=1}^{n} [y_i - (mx_i + b_0)]^2.$$

a. Take the derivative of this equation with respect to m, set it equal to 0, and solve it for m to find a formula for m in terms of the x- and y-coordinates and b_0.

b. Suppose you hang a bucket from a spring that stretches it 15 mm from its natural length. You then fill the bucket with different weights of sand and record the total distance stretched as recorded below.

Weight (N)	10	15	20	25	30	35
Distance (mm)	57.4	67.6	71.6	104.5	106.9	136.0

Because the spring was stretched before the first amount of sand was added, Hooke's law predicts that the relationship between the distance stretched, D, and the weight of sand, W, is

$$D = mW + 15.$$

Create a graph of the data to test if this prediction is reasonable. If it is, use your formula from part a. to estimate the value of m.

c. We could also estimate the value of m by transforming the data by subtracting 15 from each distance measurement, and then fitting a straight line through the origin to the transformed data. The slope of this line is m. Do this and compare the value of m to that found in part b.

2.4.10 In Exercise 2.4.9 we dealt with the problem of fixing the y-intercept of a linear model and fitting it to a set of data. Suppose we instead fix the slope. That is, we want to fit a model of the form $y = m_0x + b$ to the data where m_0 is some given value. Show that when using the least-squares criterion, the appropriate formula for the parameter b is $b = \bar{y} - m_0\bar{x}$ where $\bar{y}$ and $\bar{x}$ are the means of the y- and x-values, respectively.

2.4.11 Consider the problem of fitting a model of the form $y = a$ to a set of data. Show that when using the least-squares criterion, the appropriate formula for the parameter a is $a = \frac{1}{n}\sum_{i=1}^{n} y_i$.

2.4.12 The sum of squares can be calculated using the Excel formula **SUMXMY2**. Research this formula, then modify the worksheet **Least-Squares** to calculate the sum of squares using it. Verify that you get the same results.

2.4.13 When using Chebyshev's criterion to fit a straight line to data, we find the values of m and b that minimize the number

$$C = \text{Maximum of } \{|y_i - (mx_i + b)| \ : i = 1, 2, \ldots, n\}.$$

Consider the problem of fitting a straight line to the data in Figure 2.16, but set $m = 0.694$. Use a scroll bar to estimate the value of b that minimizes the value of C in Chebyshev's criterion. Graph the resulting line on the data and compare it to the line found using the least-squares criterion. Does Chebyshev's criterion give the exact same value of b as the least-squares criterion?

2.4.14 When using the sum of the distances criterion to fit a straight line to data, we find the values of m and b that minimize the number

$$A = \sum_{i=1}^{n} |y_i - (mx_i + b)|.$$

Consider the problem of fitting a straight line to the data in Figure 2.16, but set $b = 1.69$. Use a scroll bar to estimate the value of m that minimizes the value of A in this criterion. Graph the resulting line on the data and compare it to the line found using the least-squares criterion. Does this criterion give the exact same value of m as the least-squares criterion?

2.4.15 Sports Manufacturing Inc. manufactures footballs, basketballs, and soccer balls. Each week the company manufactures a different type of ball in varying quantities. Manufacturing costs fall into two different categories: start-up and unit. Start-up costs are costs necessary to begin production of a particular product (retool machinery, etc.). Unit costs (\$/unit) are the costs associated with manufacturing individual units (labor, materials, etc.).

The table below shows data for 15 weeks of production. The column "Total Cost" gives the total cost to produce the given number of units of that type of ball in a week. Your goal is to model the total cost, estimate the start-up and unit costs for each product, and implement the model.

Footballs		**Basketballs**		**Soccer balls**	
Units	**Total Cost**	**Units**	**Total Cost**	**Units**	**Total Cost**
2222	3125	962	1520	2481	4300
2263	3250	2246	2850	1825	3190
1267	1955	2430	2990	2238	3930
2177	3120	1395	1920	949	1890
2266	3090	2405	2750	1250	2350

a. Define variables and create a model for the total cost for producing a given number of units of each product in terms of the start-up and unit costs. List the assumptions you make.

b. Estimate the start-up and unit costs for each product.

c. Create a spreadsheet in which a user can easily input production data, such as that shown above, along with the number of units of a product that are planned for production in a given week and see an estimated total production cost for that week. Make sure the spreadsheet is logical and easy to use.

2.5 Geometric Similarity

Shapes such as circles and rectangles are easy to work with. We can calculate the area and volume of objects with these shapes using very simple formulas. Real world objects rarely come in these simple forms. This necessitates some simplifying assumptions. Geometric similarity is one such assumption.

Definition 2.5.1. Two objects are *geometrically similar* if the following two conditions are met:

1. There is a one-to-one correspondence between points of the objects (i.e. the two objects have the same "shape").
2. The ratio of distances between corresponding points is the same for all pairs of points.

In simpler terms, two objects are geometrically similar if one is a scaled up or down version of the other.

What does geometric similarity allow us to do? Let's start with a very simple example. Consider a rectangle of length 3 cm and height 2 cm. Its area is 6 cm^2. Note that this area is proportional to the square of the length of the rectangle since

$$6 = \frac{6}{3^2}\left(3^2\right) = \frac{2}{3}\left(3^2\right)$$

Now consider another rectangle of length 5 cm and height 4 cm. Its area is 20 cm^2 which is again proportional to the square of the length since

$$20 = \frac{20}{5^2}\left(5^2\right) = \frac{4}{5}\left(5^2\right)$$

Consider a third rectangle of length 6 cm and height 4 cm. This rectangle is a scaled up version of the first rectangle (its dimensions are simply 2 times the dimensions of the first one). In other words, these two rectangles are geometrically similar. Its area is 24 cm^2 which is again proportional to the square of the length since

$$24 = \frac{24}{6^2}\left(6^2\right) = \frac{2}{3}\left(6^2\right)$$

Notice that the constants of proportionality are the same for these two geometrically similar rectangles. The second rectangle is not geometrically similar to the other two rectangles, and its constant of proportionality is not the same as the others.

Let's generalize this example. Suppose we have a rectangle of length $3k$ cm and width $2k$ cm where k is some positive number. This rectangle is geometrically similar to the first rectangle. Its area is $6k^2$ cm^2, which is proportional to the square of the length since

$$6k^2 = \frac{6k^2}{(3k)^2}(3k)^2 = \frac{2}{3}(3k)^2$$

Again note that the constant of proportionality is the same as the other geometrically similar rectangles. What does this mean? Suppose we have a bag of rectangles each of length $3k$ cm and width $2k$ cm where $k > 0$ is different for each rectangle. If we reached into the bag and pulled out one rectangle and wanted to know its area, we wouldn't need to measure *both* the length and the height. We could simply measure the length, square it, and take it times $2/3$. This simplifies the process of finding the area.

The length is an example of a *characteristic dimension*. A characteristic dimension is simply a dimension of the object that is easy to measure. We could have chosen height as the characteristic dimension and done the same analysis as above, but the constant of proportionality would have been different.

This generalization illustrates the first important property of geometrically similar objects.

Theorem 2.5.1. *Suppose H is a set of geometrically similar objects. Let A denote the surface area of an object and l denote a characteristic dimension. Then*

$$A \propto l^2,$$

and the constant of proportionality is the same for every object in H.

□

Notice that no certain shape, or dimension, of the objects is assumed in Theorem 2.5.1. This property allows us to simplify the modeling of the area of complex shapes.

Example 2.5.1 (Wool From a Sheep)
A shepherd wants to predict the volume of wool he will get from a sheep, V, in terms of the girth of the sheep (the distance around the fattest part of the belly). The volume is the thickness of the wool times the area from which it is shaved. This area is not a simple shape, so we will use geometric similarity to simplify the model. Consider the following assumptions:

1. The thickness of the wool is the same for every sheep.
2. The area on each sheep from which the wool is shaved is geometrically similar.

The first assumption allows us to model

$$V \propto A \tag{2.8}$$

where A is the area from which the wool is shaved. The second assumption allows us to model

$$A \propto l^2 \tag{2.9}$$

where l is some characteristic dimension. We will choose the girth (the distance around the belly of the sheep) to be this dimension. Combining (2.8) and (2.9) using transitivity, we get

$$V \propto l^2 \tag{2.10}$$

To find the constant of proportionality, and test the assumptions, we would need to collect data of volume and girth. □

Now let's consider a three-dimensional object. Specifically consider a rectangular box with width 4 cm, height 3 cm, and depth 2 cm. Its volume is 24 cm^3, which is proportional to the height cubed since

$$24 = \frac{24}{3^3}\left(3^3\right) = \frac{8}{9}\left(3^3\right).$$

Consider a geometrically similar box with width $4k$ cm, height $3k$ cm, and depth $2k$ cm where $k > 0$. Its volume is $24k^3$ cm^3, which again is proportional to the height cubed since

$$24k^3 = \frac{24k^3}{(3k)^3}(3k)^3 = \frac{8}{9}(3k)^3.$$

Note that the constant of proportionality is the same. This generalization illustrates the second important property of geometrically similar objects.

Theorem 2.5.2. *Suppose H is a set of geometrically similar objects. Let V denote the volume of an object and l denote a characteristic dimension. Then*

$$V \propto l^3,$$

and the constant of proportionality is the same for every object in H. □

As in the first property, no special shape of the objects is assumed. This property allows us to relate the volume of an object to some characteristic dimension, and combining this with the first property we can relate volume to surface area.

Example 2.5.2 (Surface Area of a Potato)
Suppose we want to fix a large batch of the recipe "Crispy Potato Skins" for an appetizer at our Super Bowl party. This recipe requires only the skin from a potato, so when we buy the potatoes, we want to get the maximum surface area for our money.

At the supermarket we have the choice of several different sizes of potatoes. We have to decide if we want to buy several small potatoes or a few large ones (we are assuming that we can choose individual potatoes). Let's restrict ourselves to the following problem:

Should we buy 8 small baking potatoes weighing 0.25 lbs each, or 4 large baking potatoes weighing 0.5 lbs each?

To answer this question, we want to relate the surface area of a potato, A, to its weight, W. Consider the following assumptions:

1. Potatoes have a constant density.

2. Potatoes are geometrically similar.

The first assumption seems very reasonable. The veracity of the second is arguable. However, potatoes have a very irregular shape, so we need some sort of simplifying assumption to model their surface. Similarity is a *reasonable* assumption.

Now, Weight = density × volume, so the first assumption allows us to model

$$W \propto V \tag{2.11}$$

where W is the weight and V is the volume. The second assumption allows us to model

$$V \propto l^3 \tag{2.12}$$

where l is *any* characteristic dimension (such as length). Combining (2.11) and (2.12) we get

$$W \propto l^3. \tag{2.13}$$

The second assumption also allows us to model

$$A \propto l^2. \tag{2.14}$$

where A is the surface area of a potato and l is the same characteristic dimension used in (2.13). Rewriting (2.13) and combining it with (2.14), we get

$$l \propto W^{1/3} \rightarrow A \propto \left(W^{1/3}\right)^2 = W^{2/3} \tag{2.15}$$

Since potatoes are sold by the pound, each choice in the original problem will cost the same amount, so we want the choice with the largest surface area. If A_S and A_L represent the total surface area of the small and large potatoes, respectively, then (2.15) gives

$$A_S = 8k\,(0.25)^{2/3} \text{ and } A_L = 4k\,(0.5)^{2/3}$$

where k is some constant (note k is the same for both A_S and A_L). Thus

$$\frac{A_S}{A_L} = \frac{8k\,(0.25)^{2/3}}{4k\,(0.5)^{2/3}} \approx 1.26 \quad \Rightarrow \quad A_S \approx 1.26A_L$$

Therefore, the surface area of the small potatoes is approximately 26% greater than the surface area of the larger potatoes. So we should buy the smaller potatoes. □

Exercises

2.5.1 Answer the following questions related to geometric similarity and proportionality:

a. If water bottles are geometrically similar, how much more water will a bottle that is 30 cm tall hold than one that is 10 cm tall?

b. If air resistance is proportional to the surface area of a falling object at a given velocity, how much more air resistance will a sphere of diamater 5 cm encounter than one with a diameter of 1 cm?

c. If gymnasts have a constant density, then weight is proportional to volume. If we further assume that they are geometrically similar, how much less would a gymnast who is 5 ft. tall weigh than one who is 5.5 ft. tall?

d. If hearts are geometrically similar and the volume of blood pumped in one beat is proportional to the volume of the heart, how much more blood will a heart 4 cm wide pump in one beat than a heart that is 1 cm wide?

e. If the amount of heat lost by a submarine over a unit of time is proportional to its surface area, how much more heat will a submarine that is 50 ft. long lose over a given period of time than a scale model that is 10 ft. long?

f. If objects are geometrically similar and have a constant density, we saw in Example 2.5.2 that $A \propto W^{2/3}$ where $A =$ Surface Area and $W =$ Weight. If the weight of one such object is 5 times the weight of another, how much larger is the surface area?

g. Suppose that an ice cube melts so that at any point in time, the remaining cube is geometrically similar to the initial cube (i.e. before it started melting). At one point in time, the length is half the initial length. What fraction of the initial volume has melted?

h. Suppose that two elephants are geometrically similar and have a constant density. If one elephant weighs 3,500 lbs and another weighs 7,000 lbs, how much more surface area does the large elephant have than the small elephant?

2.5.2 Derive proportionality models in the following scenarios. Be sure to define variables in each model.

a. If some right triangles drawn on a page are geometrically similar, model the area of a triangle in terms of the length of its hypotenuse.

b. If some cubes in a bag are geometrically similar, model the volume of a cube in terms of the distance between opposite corners.

c. If some rocks in a bag are geometrically similar and have a constant density, model the surface area of a rock in terms of its weight.

d. If the sides of certain hollow aluminum tubes are 1 mm thick and the outer surface of the tubes are geometrically similar, model the volume of aluminum used to make a tube in terms of its length.

e. If some cargo airplanes are geometrically similar, the lift capacity of the wings is proportional to the surface area of the wings, and the cargo capacity of an airplane is directly proportional to the lift capacity of its wings, model the cargo capacity of an airplane in terms of its wing span.

f. If some books on a shelf are geometrically similar with a constant density and the thickness of a book is directly proportional to the number of pages, model the weight of a book in terms of its number of pages.

2.5.3 Consider two geometrically similar objects where the ratio of distances between corresponding points is some constant $k > 1$ called the *scaling factor*. This means that if l_s is a characteristic dimension of the "small" object, then the corresponding characteristic dimension of the "big" object is $l_b = k \cdot l_s$.

a. Prove that the areas of the small and big objects are related by $A_b = k^2 A_s$.

b. Prove that the volumes of the small and big objects are related by $V_b = k^3 V_s$.

c. Suppose that two elephants are geometrically similar and have a constant density. If one elephant weighs 3,500 lbs and another weighs 7,000 lbs, how much more surface area does the large elephant have than the small elephant?

2.5.4 When a pineapple is peeled, the resulting edible fruit has approximately the shape of a right–circular cylinder. Let d denote the diameter of this cylinder.

a. Assuming pineapples are geometrically similar and have a constant density, derive a model for the weight of a pineapple's edible fruit in terms of d.

b. Consider 6 small pineapples each of diameter 5 in, and 6 large pineapples each of diameter 5.5 in. In terms of a percentage, how much more total fruit do the large pineapples have than the small pineapples?

2.5.5 Grapes are approximately spherical in shape. Let d denote the diameter of this sphere.

a. Assuming grapes are geometrically similar and have a constant density, derive a model for the weight of a grape in terms of d. Then derive a model for the surface area in terms of d.

b. Consider 10 large grapes each of diameter 1 cm and 20 small grapes each of diameter 0.5 cm. Which set of grapes has more total weight? How much more?

c. Which set of grapes has more total surface area? How much more?

2.5.6 An asparagus spear has approximately the shape of a right–circular cone. Let d and V denote the diameter of the base and volume of this cone, respectively. Assume that all asparagus spears are geometrically similar. Some parts of a spear are tender and very edible, but some are tough and not edible. Typically spears with a larger diameter are tougher than those with a smaller diameter. In parts a–c, model the edible volume of a spear in terms of d under the given assumption.

a. The edible volume is some fraction (or proportion) of V.

b. The edible volume is proportional to V/d.

c. The edible volume is proportional to $V/d^{3.5}$.

d. Which, if any, of the above assumptions seems most reasonable to you? Explain.

e. Consider the assumption in part c. Suppose you have the option of ordering 10 "small" spears each of diamater 0.3 in, or 10 "large" spears each of diamater 0.4 in. Which option will have more edible volume? Which option would you choose?

2.5.7 Suppose it takes 0.75 fluid ounces of sunscreen to cover all exposed areas of a 3-foot tall child. Derive a model for the volume of sunscreen needed to cover a person in terms of height. Use the model to estimate the amount of sunscreen it would take to cover a 6-foot tall man. List all assumptions you make.

2.5.8 Ace manufacturing company owns a warehouse in the shape of a rectangular box that measures 100 ft. wide by 200 ft. long by 10 ft. high. It has a furnace with an output of 500,000 BTU's. They want to build a larger warehouse measuring 200 ft. by 400 ft. by 20 ft. and are trying to determine what size of furnace needs to be installed.

a. If they assume that the size of the furnace is proportional to the volume of the building, what size furnace should be installed?

b. If they assume that the size of the furnace is proportional to the total surface area of the building (including the floor), what size furnace should be installed?

c. If they assume that the size of the furnace is proportional to the surface area of the walls and roof *only* (meaning the floor is not a factor), what size furnace should be installed?

2.5.9 In Example 2.5.2 we assumed potatoes are geometrically similar and of constant density. This yielded the model $W \propto l^3$ where W is weight and l is some characteristic dimension. To test these assumptions, a student measured the length (inches) and weight (pounds) of several yellow and Idaho russet potatoes as shown in the table below (data collected by Brennan DeForest, 2019).

a. Use the data to determine if the model $W \propto l^3$ is reasonable for the yellow potatoes.

b. Repeat part a. for the Idaho russet potatoes.

c. Combine the two types of potatoes into one large sample and repeat part a.

d. What do these results suggest about the validity of the assumptions?

Yellow		**Idaho Russet**	
Length	**Weight**	**Length**	**Weight**
2.5	0.2	4.75	0.55
3.5	0.4	5.5	0.55
3	0.3	5	0.7
2.75	0.2	5.25	0.7
3.5	0.4	4.75	0.4
3.25	0.35	5.5	0.7
2.5	0.25	5.75	0.75
3	0.3	5	0.55
3.25	0.4		
3.25	0.4		
3.25	0.35		
2	0.25		

2.5.10 The table below gives the overall length (inches) and weight (pounds) of several male black bears (data collected by Brett Troyer, 2011). If we assume male black bears are geometrically similar, then we would expect that Weight $\propto$ Length^3. Use this data in the table to determine if geometric similarity is a reasonable assumption.

Length	138	166	180	129.5	150	132	148	140	137	149
Weight	60	155	220	105	110	75	105	90	75	115
Length	102	173	104.5	138	144.5	164	129	158	150	142
Weight	35	220	33	90	80	180	77	120	100	100

2.5.11 A rowing shell is a slim, needle–like boat built for speed that is powered by one, two, four, or eight rowers. We want to build a model that predicts the boat's speed, v, in terms of the number of rowers, r. The table below contains data on boat length and winning race times at four world championships (data collected from Science News, 2008).

		Time for 2,000 m (min)			
Rowers, r	**Length, l**	**I**	**II**	**III**	**IV**
8	18.28	5.87	5.92	5.82	5.73
4	11.75	6.33	6.42	6.48	6.13
2	9.76	6.87	6.92	6.95	6.77
1	7.93	7.16	7.25	7.28	7.17

Consider the following assumptions:

1. $l \propto r^{1/3}$ where l is the length of the boat.
2. The only drag slowing down the boat is from the water, which is proportional to Sv^2 where S is the wetted surface area of the boat.
3. The boats travel at a constant speed. This means that the force applied by the rowers equals the force of drag.
4. All boats are geometrically similar with regard to S.
5. The power available to the boat is proportional to r. (Power = Force × Speed where Force is the force applied by the rowers)

Is the first assumption reasonable? Test it with the given data. Build a model that predicts v in terms of r and test it with the given data. Here are a few suggestions:

1. Start with the relationship Power = Force × Speed and substitute the proportionality relationships from the other assumptions.
2. Remember, velocity = distance/time.
3. Use the average time for the 2,000 m race to calculate the velocity.

2.5.12 Consider a raindrop falling from a cloud. Ignoring any effects of wind, it is influenced by two forces, weight due to gravity and air resistance. At some point, the raindrop reaches a terminal velocity, v_t, where these two forces are equal. Consider the following assumptions:

1. All raindrops are geometrically similar.
2. All raindrops have the same density.
3. Air resistance is proportional to the product of its surface area and the square of its speed.

Use these assumptions to model the terminal velocity of a raindrop in terms of its weight.

2.5.13 Agility is an important characteristic for athletes in sports such as gymnastics. Anyone who has watched gymnastics has noted that there are very few tall gymnasts. Why is this? The goal of this exercise is to model agility in terms of height to help answer this question. Consider the following assumptions:

1. Agility $\propto \dfrac{\text{Strength}}{\text{Weight}}$ and
2. Strength $\propto$ height^2.

The first assumption is basically a mathematical "definition" of agility. The second assumption is supported by the physiological argument that a muscle's strength is proportional to its cross-sectional area.

a. Use these assumptions to model agility in terms of height. What additional assumptions do you use?

b. According to your model do shorter athletes or taller athletes tend to be more agile?

c. Critique the given assumptions. Do you think they are reasonable? How might you modify them?

2.6 Linearizable Models

In previous sections we used theory of one form or another to construct models and then used data to determine the values of parameters within the model. This process is called *model fitting* and the resulting models are called *analytical models*. The model never fit the data perfectly, but we were willing to accept some error because the model helps *explain* the behavior of the system. In this section, and the next chapter, we build models guided solely by data. We will not even attempt to use theory to explain behavior. Rather, we will find a model that captures the trend of the data and use it to *predict* values rather than explain the behavior. These models are called *empirical models*. Many of the topics related to empirical modeling are closely related to the field of statistics, and particularly the topic of regression.

Linearizable models are those which can be fit to a set of data by making an appropriate transformation and then fitting a linear model to the transformed data. Listed below are the three common types of linearizable models:

Logarithmic	**Power**	**Exponential**
$y = a + b\ln(x)$	$y = ax^b$	$y = ae^{bx}$

The variable x is called the *predictor* variable while the variable y is called the *response* variable. Graphs of these different types of models are shown in Figure 2.23. Note the "shape" of the different graphs. Each one of these different functions increases as x increases, but they increase at different rates. The logarithmic function and the power functions with exponents less than 1 increase much slower than the other types of functions. They almost appear to "level off" whereas the other types grow very quickly. Being able to recognize the shapes of the different graphs will help us to select an appropriate type of model.

To illustrate how to use these models, consider the data in Table 2.5 which gives the number of people per physician and male life expectancy (in years) for various countries around the world (data from *World Almanac Book of Facts*, 1992, Pharos Books). Our goal is to predict life expectancy in terms of the number of people per physician.

Obviously there are many factors involved with life expectancy; the number of people per physician is only one of them. It seems reasonable to believe that as the number of people per physician increases (meaning fewer physicians per person), life expectancy decreases since people would not have as easy access to health care. We do not claim that the number of people per physician *causes* life expectancy, but there is probably a relationship between the two variables. It is not at all clear how the variables of people per physician and life expectancy are related theoretically, so we will not even attempt to construct an analytical

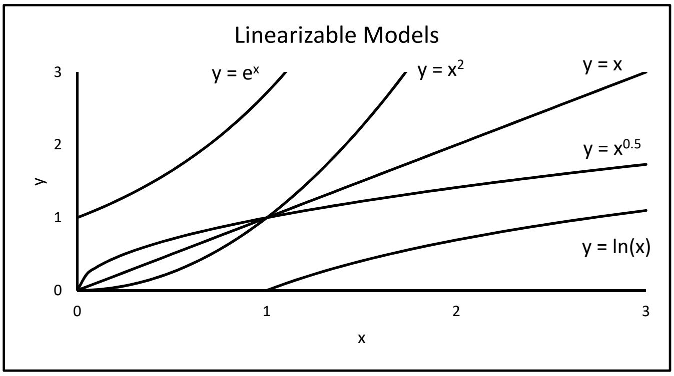


FIGURE 2.23

TABLE 2.5

Country	People/Physician P	Life Expectancy L
Spain	275	74
United States	410	72
Canada	467	73
Romania	559	67
China	643	68
Taiwan	1,010	70
Mexico	1,037	67
South Korea	1,216	66
India	2,471	57
Morocco	4,873	62
Bangladesh	6,166	54
Kenya	7,174	59

model. We will construct an empirical model by fitting various linearizable models to this data and analyzing how well each one fits.

When constructing any type of empirical model, the first step is plot y vs. x (L vs. P in this case) and look at the shape. The second step is to select an appropriate type of model and fit it to the data.

A graph of the data is shown in Figure 2.24. Notice that as the number of people per physician increases, the life expectancy decreases, agreeing with our intuition. Also note that the points seems to form a curve that initially decreases rapidly, but then levels off. This suggests that a logarithmic or power model might be appropriate.

Example 2.6.1 (Logarithmic Model)
We will fit a curve of the form $L = a + b \ln(P)$ to the data by graphing L vs. $\ln(P)$ and fitting a straight line. The value of b is the slope of this line and the value of a is the y-intercept.

1. Name a blank worksheet "**ln**" and format it as in Figure 2.25. Enter the rest of the data from Table 2.5 in columns **A** and **C**. Copy cell **B2** down to row 13.

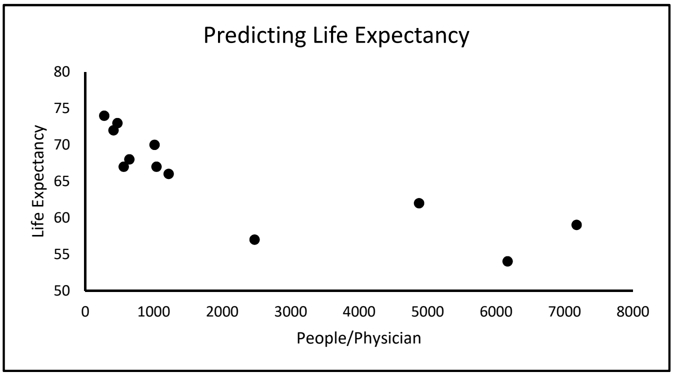


FIGURE 2.24

	A	B	C
1	**P**	**ln(P)**	**L**
2	275	=LN(A2)	74

FIGURE 2.25

2. Create a graph of L vs. $\ln(P)$, add a linear trendline, and display the equation of the line as in Figure 2.26. Using the slope and y-intercept of this line, we get our model: $L = 103.4 - 5.2376 \ln(P)$

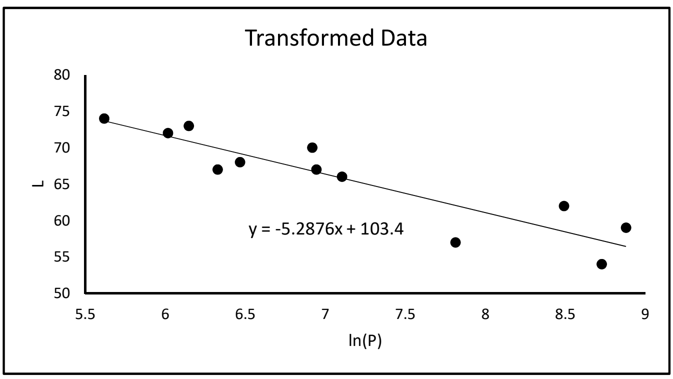


FIGURE 2.26

3. Next we need to compare the model to the original data. Add the formula in Figure 2.27 and copy cell **D2** down to row 13.

	D
1	**Predicted**
2	=103.4-5.2876*LN(A2)

FIGURE 2.27

4. Create a graph to compare the observed values of L to the predicted ones as in Figure 2.28. Notice that the two sets of values are fairly close together, indicating that we have a good model.

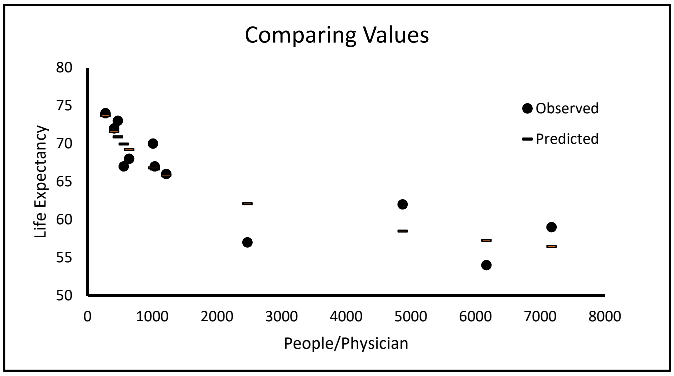


FIGURE 2.28

5. To further analyze how well the model fits the data, for each data point define

$$\text{Residual} = (\text{Observed value}) - (\text{Predicted value})$$

Note that a positive residual means that the predicted value is less than the observed value. A negative value means that the predicted value is greater than the observed value. To calculate the residual for each data point, add the formula in Figure 2.29 and copy cell **E2** down to row 13.

	E
1	**Residual**
2	=C2-D2

FIGURE 2.29

6. Create a graph of Residual vs. P as in Figure 2.30. Note that roughly half the residuals are positive and half are negative. This indicates that the model does not tend to over predict or under predict the values of L. Also note that the magnitudes of the residuals (i.e. the absolute values) are all relatively small, less than 6, and that there is no "pattern" to the residuals. These three observations indicate that this model fits relatively well.

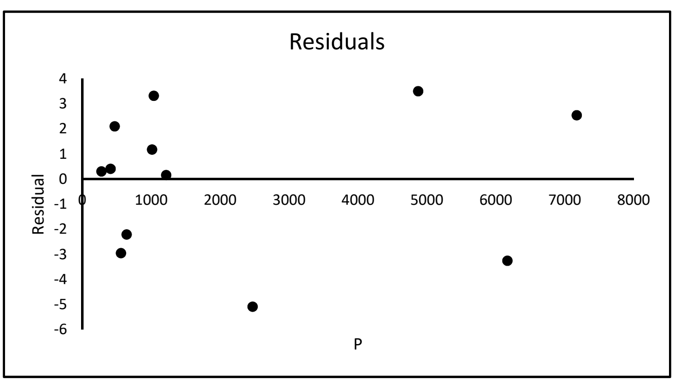


FIGURE 2.30

□

Example 2.6.2 (Power Model)
We will fit a curve of the form $L = aP^b$ to the data. To find the values of a and b, we take the natural logarithm of both sides of the model to get

$$\ln L = \ln(aP^b) = \ln a + \ln P^b = \ln a + b \ln P.$$

Thus a straight line fit to the graph of $\ln L$ vs. $\ln P$ will have a slope of b and a y-intercept of $\ln a$.

1. Name a blank worksheet "**Power**" and format it as in Figure 2.31. Copy the rest of the data from worksheet **ln** into columns **A** and **B**. Copy the range **C2:D2** down to row 13.

	A	B	C	D
1	**P**	**L**	**ln(P)**	**ln(L)**
2	275	74	=LN(A2)	=LN(B2)

FIGURE 2.31

2. Graph $\ln L$ vs. $\ln P$, fit a straight line, and display the equation on the graph as in Figure 2.32.

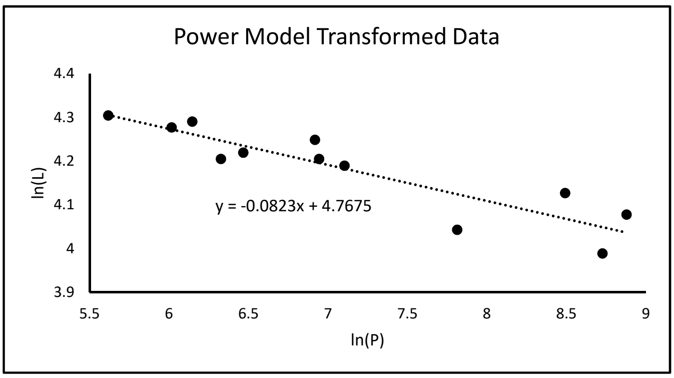


FIGURE 2.32

The equation of the line is $y = -0.0823x + 4.7675$, so $b = -0.0823$ and

$$\ln a = 4.7675 \quad \Rightarrow \quad a = e^{4.7675} = 117.62$$

Therefore, the model is $L = 117.62P^{-0.0823}$.

3. Add the formulas in Figure 2.33 to calculate the predicted values and the residuals. Copy row 2 down to row 13.

4. Create a graph of Residual vs. P as shown in Figure 2.34. Note that again roughly half of the residuals are positive and half are negative, they all have magnitudes less than 5, and there is no pattern. This indicates a good model.

□

	E	F
1	**Predicted**	**Residual**
2	=117.62*A2^-0.0823	=B2-E2

FIGURE 2.33

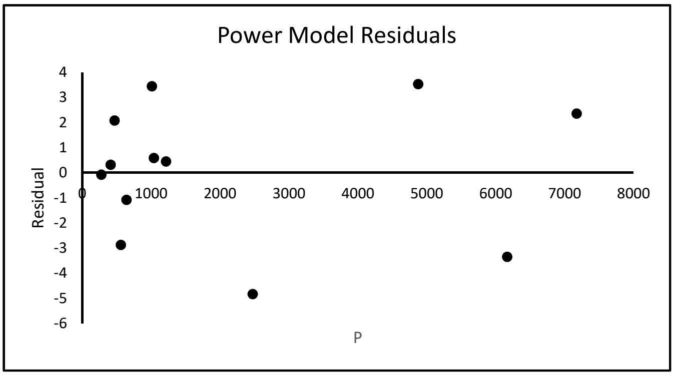


FIGURE 2.34

Example 2.6.3 (Exponential Model)
The plot of the data does not resemble the graph of an exponential model, so this type of model may not be the best. However, we will fit one to the data to illustrate the process. The exponential model has the form $L = ae^{bP}$. To find the values of a and b, we take the natural logarithm of both sides of the model to get

$$\ln(L) = \ln(ae^{bP}) = \ln a + \ln(e^{bP}) = \ln a + bP.$$

Thus a straight line fit to the graph of $\ln L$ vs. P will have a slope of b and a y-intercept of $\ln a$.

1. Name a blank worksheet "**Exponential**" and format it as in Figure 2.35. Copy the rest of the data from the worksheet **Power** and copy cell **C2** down to row 13.

	A	B	C
1	**P**	**L**	**ln(L)**
2	275	74	=LN(B2)

FIGURE 2.35

2. Create a graph of $\ln L$ vs. P and fit a straight line to the data as in Figure 2.36. Display the equation on the chart. Notice that this line does not fit the data as well as with the other two models. This is an indication that an exponential model does not fit the data as well as the others.

   The equation of this line is $y = -0.00003x + 4.2561$, so $b = -0.00003$ and

   $$\ln a = 4.2561 \quad \Rightarrow \quad a = e^{4.2561} = 70.53$$

   Therefore, the model is $L = 70.53\, e^{-0.00003P}$.

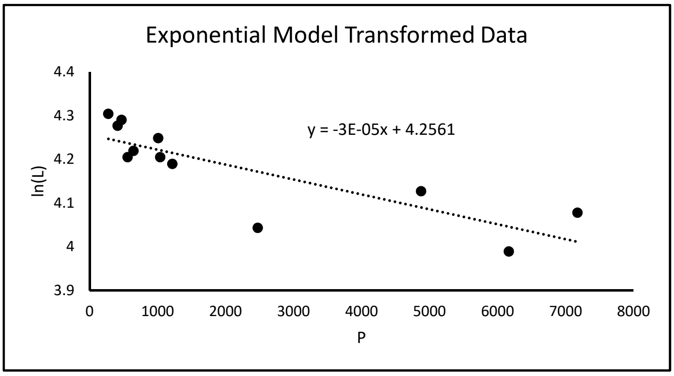


FIGURE 2.36

	D	E
1	**Predicted**	**Residual**
2	=70.53*EXP(-0.00003*A2)	=B2-D2

FIGURE 2.37

3. Add the formulas in Figure 2.37 to calculate the predicted values and the residuals. Copy row 2 down to row 16.

4. Create a graph of Residual vs. P as shown in Figure 2.38. Notice that all the residuals are at least 1 in magnitude and that one is almost -9. This indicates that the model doesn't predict any of the values of L very accurately. Therefore, this is not the best fitting model, as previously suspected.

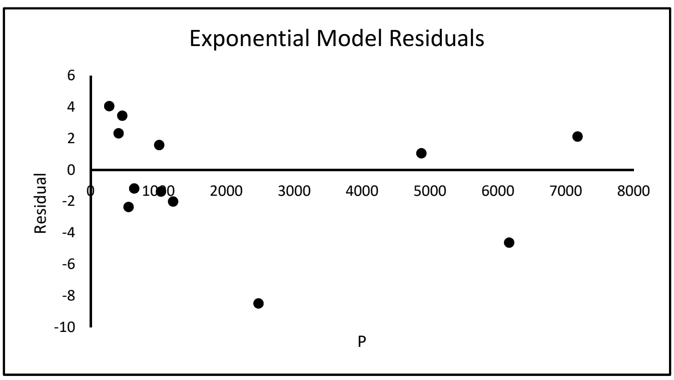


FIGURE 2.38

□

Example 2.6.4 (Trendlines)
Excel will automatically calculate these different models for us. Create a graph of L vs. P, add a trendline, select the type of model you want under **Type**, and display the equation on the chart. The results are shown in Figure 2.39. Note that these are exactly the same models we derived. Also note that the graphs of the logarithmic and power models capture

the trend of the data very well while the exponential model does not. This confirms our conclusions based on the graphs of the residuals.

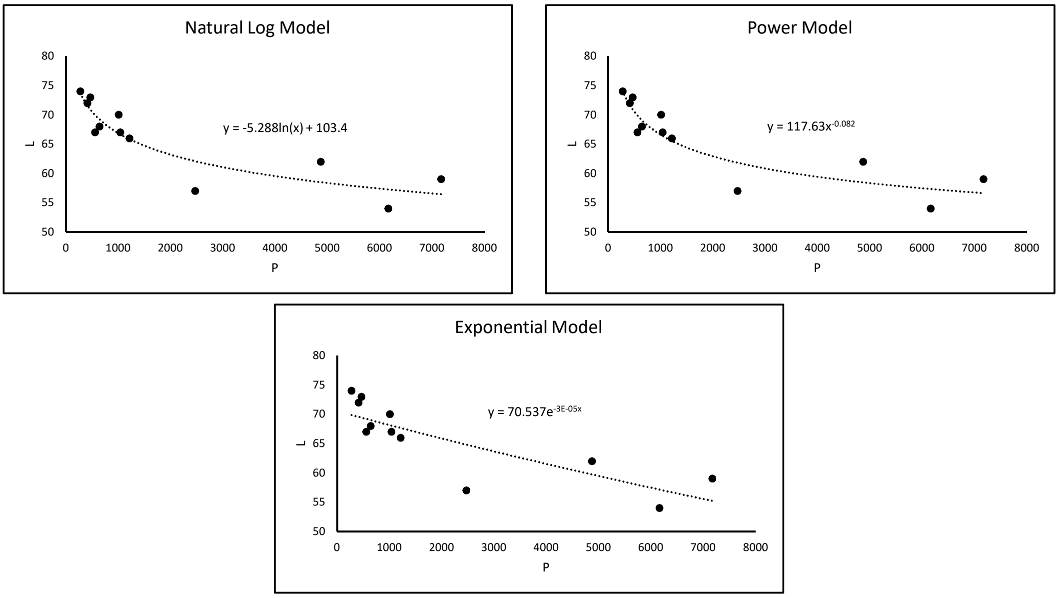


FIGURE 2.39

□

Out of the three models, the logarithmic and power models are the "best" based on an analysis of the graph of the models and of the residuals. In the next section we will look at more analytical techniques for measuring how well a model fits a set of data.

Now that we have two good–fitting models for the data, what can we do with them? There are at least two uses. First of all, the graphs of the models help us to see the trend of the data. The graphs decrease from left to right, helping us to illustrate the point that as the number of people per physician increases, life expectancy decreases. The plot of the data also shows this, but a curve helps exemplify the relationship.

Second of all, we can use the models to *predict* life expectancy if we know the number of people per physician. For instance, suppose a country has 3,500 people per physician. The logarithmic model predicts that life expectancy is

$$L = 103.4 - 5.2876 \ln(3500) \approx 60.25 \text{ years}$$

while the power model gives

$$L = 117.63\,(3500)^{-0.0823} \approx 60.10 \text{ years.}$$

We certainly could plug $P = 3,500$ into the exponential model and get

$$L = 70.537e^{-0.00003(3500)} \approx 63.51 \text{ years.}$$

However, we saw that the exponential model did not fit the data very well, so it would not be appropriate to use it for making predictions. This illustrates the first caution when using empirical models: **If a model does not fit a set of data, do not use it for making predictions.**

Now suppose a country has 100 people per physician. We could plug $P = 100$ into the logarithmic model and get

$$L = 103.4 - 5.2876 \ln(100) \approx 79.04 \text{ years.}$$

However, note that $P = 100$ is outside the range of the original data. We do not know the trend of the data for values of P less than 275. It could change or stay the same; we simply do not know. Therefore, it would be inappropriate to use any of these models to predict values of L for P less than 275. This value of 79.04 years may or may not be accurate, so we should not report this "prediction." This illustrates the second caution when using empirical models: **Only use values of the predictor variable that are within the range of the original set of data.**

Exercises

2.6.1 For each data set below, determine which model - exponential, power, or logarithmic - best fits the data. Briefly explain your reasoning.

a.

x	1	2	3	4	5	6
y	1.66	2.41	6.04	9.89	17.31	31.54

b.

x	1	2	3	4	5	6
y	2.68	5.61	8.71	9.83	11.16	11.03

2.6.2 The table below contains the total length and weight of 20 black bears (data collected by Brett Troyer, 2011). Graph weight vs. length, fit different linearizable models to the data, and select the one that best fits the data. Briefly explain your reasoning.

Length	139	138	139	120.5	149	141	150	166	180	129.5
Weight	110	60	90	60	85	95	85	155	220	105
Length	150	142	162	148	140	134	137	149	102	151.5
Weight	110	115	255	105	90	75	75	115	35	140

2.6.3 The table below contains data on overall fuel economy (measured in miles per gallon, MPG) and acceleration (measured in seconds to accelerate from 0 to 60 miles per hour) of 20 gasoline powered passenger cars (data from https://www.consumerreports.org/cro/news/2013/06/fuel-economy-vs-performance/index.htm, accessed 5/16/2018):

MPG	33	31	29	38	26	25	24	35	35	33
Acc.	9.9	10	9.8	7.6	6.3	6.9	9.2	9	10.3	10
MPG	33	30	29	27	25	25	28	27	26	24
Acc.	10.9	7	7.3	7.2	6.4	5.2	7.5	7.4	7.5	6.6

a. Plot acceleration vs. MPG and fit different linearlizable models. Do any of them fit the data particularly well? What, if anything, can we say about the general relationship between MPG and acceleration time?

b. The table below contains similar data on six electric passenger cars. The overall fuel economy is measured in miles-per-gallon equivalent (MPGe). Plot these data on the same graph as part a. Is there any model that fits all these data points? What can we say about the relationship between gasoline powered and electric passenger cars?

MPGe	139	107	106	105	94	87
Acc.	7.5	10.2	10.3	8	8.1	3.5

2.6.4 For each set of data below, fit a model of the given form by transforming the data appropriately and fitting a straight line to the transformed data. Graph the resulting model on top of the data and analyze how well the model fits the data.

a. Model: $y = ax^2 + b$

$\boldsymbol{x}$	1	2	3	4	5	6
$\boldsymbol{y}$	16.3	23.1	37.4	46.9	58.7	91.0

b. Model: $y = a\sin(x) + b$

$\boldsymbol{x}$	1	2	3	4	5	6
$\boldsymbol{y}$	1.34	1.61	-0.98	-3.80	-4.55	-2.30

c. Model: $y = a\dfrac{x^2+1}{\ln(x)} + b$

$\boldsymbol{x}$	2	3	4	5	6	7
$\boldsymbol{y}$	3.30	5.63	9.52	14.31	19.84	26.061

2.6.5 Consider the data below:

$\boldsymbol{x}$	0	2	4	6	8	10
$\boldsymbol{y}$	3.000	3.061	3.122	3.186	3.250	3.316

a. Fit a linear model to the data. How well does the model appear to fit the data? Create a graph of the residuals. What do you notice about the pattern of the residuals?

b. Fit an exponential model to the data. How well does the model appear to fit the data? Calculate and graph the residuals. What does this tell you about how well this model fits the data?

2.6.6 Use algebra or the natural logarithm to derive the given linearization of each of the following models:

	Model	Linearization
a.	$y = \dfrac{1}{a+bx}$	$\dfrac{1}{y} = a + bx$
b.	$y = \dfrac{x}{a+bx}$	$\dfrac{1}{y} = \dfrac{a}{x} + b$
c.	$y = 1 + ae^{bx}$	$\ln(y-1) = \ln a + bx$
d.	$y = 1 - e^{-x^a/b}$	$\ln\left[\ln\left(\dfrac{1}{1-y}\right)\right] = a\ln x + \ln\dfrac{1}{b}$

2.6.7 Growing populations typically have a maximum size called the *carrying capacity*. The size of a population y at time t can often be modeled with the *logistic model*

$$y = \frac{L}{1 + e^{a+bt}}$$

where L is the carrying capacity and a and b are parameters.

a. Linearize this model with simple algebra and the natural logarithm by showing that

$$\ln\left(\frac{L-y}{y}\right) = a + bt.$$

b. The data below show the size of a bacteria population at certain points in time. Assuming the carrying capacity is $L = 625$, fit a logistic model to the data using the linearization in part a.

t	0	1	2	3	4	5	6	7	8	9
y	10.3	17.2	27	45.3	80.2	125.3	176.2	255.6	330.8	390.4
t	10	11	12	13	14	15	16	17	18	19
y	440	520.4	560.4	600.5	610.8	614.5	618.3	619.5	620	621

c. Create a graph of the residuals for the model found in part b. What does this say about the quality of the fit of this model?

2.6.8 In 1965, Intel co-founder Gordon Moore predicted that the number of transistors in integrated circuits would double approximately every 24 months. This prediction has become known as *Moore's law*. In the data below, y denotes the number of transistors (in millions) in various integrated circuits between years 1971 and 2014 (data collected by Jonathan Grant, 2018).

Year	1971	1972	1974	1976	1978	1979	1982	1988	1989	1993	1995	1998
y	0.0023	0.0035	0.0045	0.0065	0.029	0.029	0.134	0.25	1.18	3.1	5.5	7.5
Year	1999	2000	2001	2002	2003	2004	2006	2008	2010	2011	2012	2014
y	27.4	42	45	220	410	592	1700	1900	2300	2600	5000	5560

Mathematically speaking, we can think of Moore as predicting that y is described by a model of the form

$$y = c_0 2^{t/r}$$

where c_0 is the initial number of transistors (corresponding to 1971 in the data), t is the number of months since 1971, and r is the number of months needed for y to double.

a. Linearize this model.

b. Use the linearization to fit the model to the data and estimate a value of r.

c. Based on these data, does Moore's law appear to be accurate?

2.6.9 This problem will illustrate that we have to be careful when working with big numbers. Consider the problem of fitting an exponential model to the data in the table below.

$\boldsymbol{x}$	1947	1957	1965	1977
$\boldsymbol{y}$	21500	13500	8000	7500

a. Graph the data and fit an exponential trendline. Graphically, how well does this model appear to fit the data?

b. Use the exponential model found in part a. to predict the values of y (note that the symbol 2E+35 means 2×10^{35}). How good are these predictions?

c. We might wonder if the poor predictions are caused by rounding errors in the parameters. To investigate this, right-click on the trendline label and select **Format Trendline Label**. Under **Category** choose **Scientific** and set the number of **Decimal places** to 6. Use the resulting model to predict the values of y. Are these predictions any better?

d. Another way to work with big numbers is to scale them to smaller numbers. For instance, consider dividing each y-value by 1,000 and subtracting 1947 from each x-value. This yields the modified model

$$\frac{y}{1000} = ae^{b(x-1947)}.$$

Transform the data by subtracting 1947 from each x-value and dividing each y-value by 1,000, plot the transformed data, and fit an exponential trendline to the transformed data. The parameters in this exponential trendline match the parameters in the modified model.

e. Solve the modified model for y and use the parameters found in part d. to predict the values of y. How good are these predictions?

2.6.10 As we have seen, logarithms can be used to linearize models. They can also help us work with big or small numbers. Consider the problem of finding the constant of proportionality in the model

$$t_l \propto \frac{1}{m_p^2}$$

where t_l is the amount of time needed for the satellite to become in orbital lock and m_p is the mass of the planet, as described in Exercise 2.2.10. The table below shows the values for the eight planets in our solar system (data collected by Joshua Hendrickson, 2019).

Planet	m_p ($\times 10^{24}$ kg)	t_l ($\times 10^7$ years)
Mercury	0.33	126094
Venus	4.87	578.98
Earth	5.97	385.28
Mars	0.642	33315.9
Jupiter	1898	0.003812
Saturn	568	0.042562
Uranus	86.8	1.822561
Neptune	102	1.31984

Using the approach introduced in this section, we could graph t_l vs. $1/m_p^2$ and fit a straight line through the origin the slope of which is the constant of proportionality. However, some of the numbers are so much bigger than the others that many of the data points are so close together that they cannot be distinguished from each other. To help solve this problem, we can take the natural log of both sides of the model $t_l = k/m_p^2$ where k is the constant of proportionality and algebraically rewrite yielding

$$\ln(t_l) = \ln(k) + \ln\left(\frac{1}{m_p^2}\right).$$

Thus a graph of $\ln(t_l)$ vs. $\ln\left(1/m_p^2\right)$ should have a slope of 1 and a y-intercept of $\ln(k)$. Such a graph is called a *log-log plot*.

a. Calculate $\ln(t_l)$ and $\ln\left(1/m_p^2\right)$ for all the planets.

b. Graph $\ln(t_l)$ vs $\ln\left(1/m_p^2\right)$ and fit a straight line to the data.

c. Use the straight line to estimate the constant of proportionality k.

Directions for Exercises 2.6.11 - 2.6.13: In this section, we found the values of the parameters in the models by transforming the data appropriately and then fitting a straight line to the transformed data. Another approach is to find formulas for the parameters using the least-squares criterion. That is, if the model is $y = f(x)$, then the parameters should minimize the quantity

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2.$$

This can be done, in principle, by taking the partial derivatives of S with respect to each of the parameters, setting the derivatives equal to 0, and solving for the parameters, much like we did in Section 2.4.

In Exercises 2.6.11 - 2.6.13 you are given a type of linearizable model and formulas for the parameters. For each exercise:

a. Design a spreadsheet to show, by example, that the given formulas give the same values of the parameters as we would get if we transformed the data appropriately and fit a straight line to the transformed data.

b. Derive the given formulas using the approach described above.

2.6.11 Model: $y = ax^3$

$$a = \frac{\sum y_i x_i^3}{\sum x_i^6}$$

2.6.12 Model: $y = ax^2 + b$

$$a = \frac{-n\sum x_i^2 y_i + \sum y_i \sum x_i}{\left(\sum x_i^2\right)^2 - n\sum x_i^4}, \quad b = \frac{\sum y_i - a\sum x_i^2}{n}$$

2.6.13 Model: $y = a + b\ln x$

$$b = \frac{\sum y_i \sum \ln x_i - n\sum y\ln x_i}{\left(\sum \ln x_i\right)^2 - n\sum\left(\ln x_i\right)^2}, \quad a = \frac{\sum y_i - b\sum \ln x_i}{n}$$

2.7 Coefficient of Determination

The coefficient of determination, denoted R^2, is a numerical measure of how well a line fits a set of data. To illustrate the fundamental concepts, we will generate some hypothetical data according to the relationship $y = 3 + 2x$.

1. Rename a blank worksheet "**R2**" and format it as in Figure 2.40. Copy row 3 down to row 11 to generate 10 data points.

	A	B
1	**x**	**y**
2	0	=3+2*A2
3	=A2+1	=3+2*A3

FIGURE 2.40

2. Create a graph of the data and add a trendline as in Figure 2.41. Note that the line goes through each data point and the equation of this line (also called the *regression equation*) is exactly what was used to generate the data.

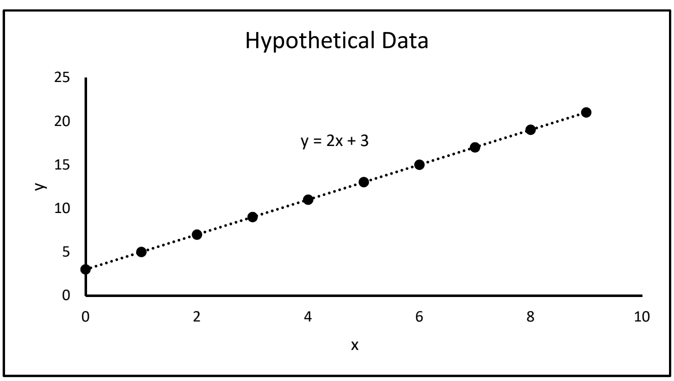


FIGURE 2.41

One purpose of fitting a line to data is to use it for predicting the value of y when x is known. If we did not have a graph of the data or the regression equation and we were given a value of x, the best guess as to the corresponding value of y would be the mean of the y-values. This mean is denoted by $\bar{y}$ and equals 12 in this case.

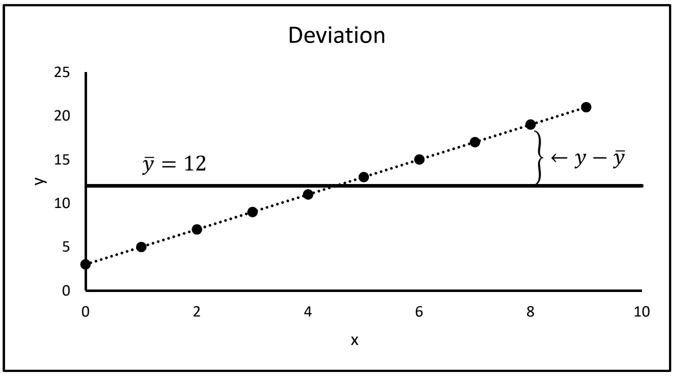


FIGURE 2.42

We test this simple prediction strategy by examining the "error" it would cause for the given data points. This error, or deviation, $y - \bar{y}$, is illustrated in Figure 2.42.

If a y-value predicted by the regression equation is denoted $\hat{y}$, we measure how well a regression equation fits the data by comparing the deviation to the difference between $\hat{y}$ and $\bar{y}$, $\hat{y} - \bar{y}$. In this case the regression line goes through each data point, so $\hat{y} = y$. So,

$$\hat{y} - \bar{y} = y - \bar{y} \quad \Rightarrow \quad \frac{\hat{y} - \bar{y}}{y - \bar{y}} = 1$$

Thus we give this regression equation an R^2 value of 1. This value is often interpreted by saying that the regression equation "explains" 100% of the deviation. The definition of the coefficient of determination is based on this idea of comparing the deviation to the difference between $\hat{y}$ and $\bar{y}$ to measure the percentage of deviation "explained" by the regression equation.

This set of data is highly idealized because it was generated exactly according to the linear relationship $y = 3 + 2x$. In reality, data never conforms to an exact relationship like this. Real data with a linear relationship satisfies an equation of the form

$$y = \beta_0 + \beta_1 x + \varepsilon$$

where ε is some "noise." This noise may be due to measurement error, sampling variation, or some other random event outside of our control.

The relation $y = \beta_0 + \beta_1 x$ is called the "true" linear trend of the population while the regression equation has the generic form $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$. The "hats" indicate that the parameters $\hat{\beta}_0$ and $\hat{\beta}_1$ are estimates of the population values based on sample data. Likewise, $\hat{y}$ is an estimate of y.

Example 2.7.1 (Including Noise)
To generate data with some noise modify the worksheet **R2** as in Figure 2.43. Copy cell **B2** down to row 11. Here our noise is given by NORMINV(RAND(),0,2) which is a normally distributed pseudorandom variable with mean 0 and standard deviation 2 (see Chapter 6, particularly Section 6.6, for more details on how this formula works).

The graph of this noisy data should resemble Figure 2.44. Note that your graph will probably look different due to the random noise. Also note that the regression equation is *not* $y = 3 + 2x$.

□

	B
1	**y**
2	=3+2*A2+NORMINV(RAND(),0,2)

FIGURE 2.43

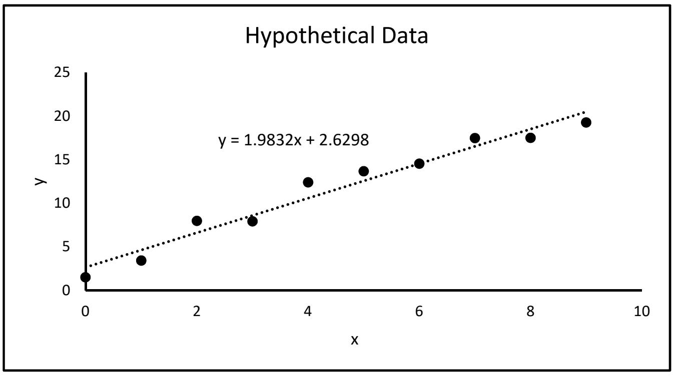


FIGURE 2.44

To define the R^2 value for a line fit to noisy data, we introduce three different types of deviation, *explained*, *unexplained*, and *total* deviation. Figure 2.45 illustrates these types of deviations. We see from the figure that

$$\begin{array}{ccccc} (\text{total deviation}) & = & (\text{explained deviation}) & + & (\text{unexplained deviation}) \\ (y-\bar{y}) & = & (\hat{y}-\bar{y}) & + & (y-\hat{y})\,. \end{array}$$

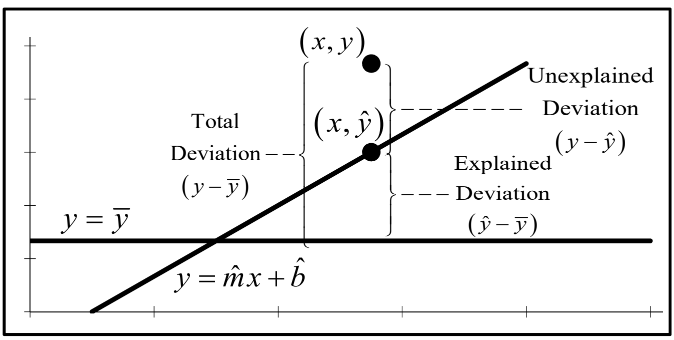


FIGURE 2.45

If we square these deviations and add them together for all data points, we get amounts of *variation* (measures of the total deviation). The total variation is called the *total sum of squares* and is given by

$$SS_{Tot} = \sum (y_i - \bar{y})^2$$

The explained variation is called the *regression sum of squares* and is given by

$$SS_{Reg} = \sum (\hat{y}_i - \bar{y})^2$$

The unexplained variation is called the *residual sum of squares* and is given by

$$SS_{Res} = \sum (y_i - \hat{y}_i)^2$$

Analogous to the relationship between the deviations we get

$$\begin{array}{ccccc}
\text{(total variation)} & = & \text{(explained variation)} & + & \text{(unexplained variation)} \\
\sum (y_i - \bar{y})^2 & = & \sum (\hat{y}_i - \bar{y})^2 & + & \sum (y_i - \hat{y}_i)^2
\end{array}$$

(proving this relationship is true is not a trivial matter). Rewriting this equation we get

$$\begin{array}{ccccc}
\text{(explained variation)} & = & \text{(total variation)} & - & \text{(unexplained variation)} \\
\sum (\hat{y}_i - \bar{y})^2 & = & \sum (y_i - \bar{y})^2 & - & \sum (y_i - \hat{y}_i)^2 \\
SS_{Reg} & = & SS_{Tot} & - & SS_{Res}.
\end{array}$$

Therefore, the "percentage" of the total variation explained by the regression equation is

$$R^2 = \frac{SS_{Tot} - SS_{Res}}{SS_{Tot}}. \tag{2.16}$$

This is the definition of the coefficient of determination . The closer R^2 is to 1, the better the line fits the data. An R^2 value close to 0 indicates a very poor fitting line.

Example 2.7.2 (Calculating R^2)
To use formula (2.16) to calculate R^2 for the regression line fit to the noisy data, follow these steps:

1. Modify the worksheet **R2** as in Figure 2.46.

	A	B
14	**Average =**	=AVERAGE(B2:B11)
15	**Slope =**	=SLOPE(B2:B11,A2:A11)
16	**y-int =**	=INTERCEPT(B2:B11,A2:A11)

FIGURE 2.46

2. Add the formulas in Figure 2.47 and copy row 2 down to row 11.

	C	D	E
1	**Predicted**	**SS_{Tot}**	**SS_{Res}**
2	=A2*$B$15+$B$16	=(B2-$B$14)^2	=(B2-C2)^2

FIGURE 2.47

3. Add the formulas in Figure 2.48.

	C	D	E
13	**Sum =**	=SUM(D2:D11)	=SUM(E2:E11)
14		**R^2 =**	=(D13-E13)/D13

FIGURE 2.48

4. Add a linear trendline to the graph of the data. Under **Options**, select "Display R-squared value on chart." Note that this R^2 value is equal to what we calculated.

5. Press the **F9** key several times. Each time you press it, the random numbers in the noise are regenerated and a new set of data is created. Your R^2 value in cell **E14** should equal the one automatically generated by Excel each time.

   Notice that the R^2 value is not exactly one. That is because our data has some noise, so the underlying linear relationship $y = 3 + 2x$, or any other linear relationship, does not account for all of the variation from the mean. The R^2 value is, however, very close to 1. This indicates that the regression line does fit the data very well.

6. To add more noise to the data, modify the formulas in the worksheet **R2** as in Figure 2.49 and copy cell **B2** down to row 11. Now the noise is normally distributed with mean 0 and standard deviation 4, so it is more "spread out" than before.

	B
1	y
2	=3+2*A2+NORMINV(RAND(),0,4)

FIGURE 2.49

   Notice that the R^2 value is less than before. The straight line cannot account for as much of the variation from the mean because of the greater noise.

As mentioned above, one purpose for finding a regression equation is to estimate the values of β_0 and β_1 in the relationship $y = \beta_0 + \beta_1 x + \varepsilon$. The R^2 value is not a measure of the accuracy of these estimates. It is simply a measure of how well the regression line fits the observed data. □

Example 2.7.3 (Applying R^2 to Linearizable Models)
Consider the linearizable models fit to the life expectancy data in Section 2.6. We can compare how well the different models fit the data by calculating the R^2 value for each model and then comparing the R^2 values. To calculate the R^2 value for a linearizable model, we calculate the R^2 value for the straight line fit to the *transformed* data.

1. In the worksheet **Power** from Section 2.6, display the R^2 value for the linear trendline on the graph of $\ln L$ vs. $\ln P$ by right-clicking on the trendline, and selecting **Format Trendline...** → **Options** → **Display R-squared value on chart**. Note this value is 0.8152 indicating a good fitting model.

2. Also in the worksheet **Power**, right-click on the power trendline on the graph of L vs. P. Select **Format Trendline...** → **Options** → **Display R-squared value on chart**. This is the same R^2 value.

Repeat this process for the logarithmic and exponential models derived in Examples 2.6.1 and 2.6.3. We can now compare how well these models fit the data by comparing their R^2 values:

Logarithmic	**Power**	**Exponential**
0.8255	0.8152	0.6864

We see that the power and logarithmic models fit the data very well with the logarithmic model being slightly better. The exponential model does not fit as well. Thus, based strictly on the R^2 values, we would conclude that the logarithmic model fits the data the best. This agrees with our earlier conclusion.

□

We warn against blindly using R^2 values to choose a best model. These values should be used as only one factor when choosing a best model. Other factors that should be considered are the nature of the behavior being analyzed and the simplicity of the model.

For instance, population growth is often exponential. So if we fit curves to some data of a population, we may want to favor an exponential model over other types even if it has a lower R^2 value.

In "Modeling the U.S. Population" (*AMATYC Review*, Vol. 20, No. 2, Spring 1999, pages 17–29), Sheldon Gordon makes the point that, "The best choice (of a model) depends on the set of data being analyzed and requires an exercise in judgement, not just computation."

Exercises

2.7.1 The data below contain the weights (in lbs) and highway miles per gallon (MPG) of several cars. Calculate SS_{Reg}, SS_{Res}, SS_{Tot}, and R^2 for the linear regression equation fit to this data.

Weight (x)	3250	3425	2400	2250
MPG (y)	26	28	37	38

2.7.2 The data below contain the amounts of nitrogen applied to different corn fields (in lbs/acre) and the resulting yields (in bushels/acre). Use the definitions to calculate SS_{Reg}, SS_{Res}, SS_{Tot}, and R^2 for the linear regression equation fit to this data.

Nitrogen (x)	0	60	120	180	240
Yield (y)	78	90	140	162	210

2.7.3 The table below gives average rebounds per game (RPG) and average points scored per game (PPG) of 11 players on a university basketball team (data collected by Alexa Hopping, 2012). Calculate R^2 for the linear regression equation fit to this data. Based on the R^2 value, is RPG a very good predictor of PPG?

RPG (x)	3.1	2.8	6.7	2.5	1.5	1.4	4.8	1.4	3.7	4.1	3.6
PPG (y)	8.6	4.5	7.1	15.8	1.6	5.2	8	2.8	4	14.5	5

2.7.4 The data below contain the diameter of the trunk at chest height and volume of wood in several pine trees. Use the trendline function in Excel to model Volume in terms of Diameter with several different linearizable models and select the best one. Briefly explain how you decided which one is best.

Diameter	32	29	24	45	20	30	26	40	24	18
Volume	185	109	95	300	30	125	55	246	60	15

2.7.5 Another measure of how well a linear regression line fits a set of data is the **standard error of estimate**, denote by s_e. It is defined by

$$s_e = \sqrt{\frac{\sum (y_i - \hat{y})^2}{n-2}}$$

Equivalently, we could define it by

$$s_e = \sqrt{\frac{SS_{Res}}{n-2}}$$

The smaller s_e is, the better the fit. This quantity is also an estimate of the standard deviation of the noise ε in the relationship $y = \beta_0 + \beta_1 x + \varepsilon$.

Modify the worksheet **R2** to calculate s_e for the linear model fit to the data with some noise. How close is s_e to the true standard deviation of the noise? Try different values of the standard deviation of the noise.

2.7.6 Let $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ be the least–squares regression equation fit to a set of data $\{(x_i, y_i) : i = 1, \ldots, n\}$ and let $\hat{y}_i$ be an estimate of y_i based off this equation.

a. Modify the worksheet **R2** to illustrate these two properties:

$$\sum_{i=1}^{n} (y_i - \hat{y}_i) = 0 \text{ and } \sum_{i=1}^{n} x_i(y_i - \hat{y}_i) = 0$$

b. Also modify the worksheet **R2** to illustrate this property:

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 + \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

In other words, show that

$$SS_{Tot} = SS_{Reg} + SS_{Res}$$

c. Illustrate that none of these three properties hold if we use a slope other than $\hat{\beta}_1$ or a y–intercept other than $\hat{\beta}_0$ to calculate $\hat{y}_i$. In other words, show that these properties do not hold if we use a slope other than that given by the Excel formula **SLOPE** or a y–intercept other that that given by **INTERCEPT** to calculate the predicted value of y.

d. Use the properties in part 1 to prove the property in part 2. **Hint**: Start with

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \bar{y} + \hat{y}_i - \hat{y}_i)^2 = \sum_{i=1}^{n} ((\hat{y}_i - \bar{y}) + (y_i - \hat{y}_i))^2$$

expand the right–hand side and rewrite so that you get the terms $\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2$, $\sum_{i=1}^{n} (y_i - \hat{y})^2$, $\sum_{i=1}^{n} (y_i - \hat{y}_i)$, and $\sum_{i=1}^{n} x_i (y_i - \hat{y}_i)$.

2.7.7 Let $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$ be the least–squares regression equation fit to a set of data $\{(x_i, y_i) : i = 1, \ldots, n\}$. Modify the worksheet **R2** to illustrate this property:

$$R^2 = \hat{\beta}_1^2 \cdot \frac{n \sum x_i^2 - \left(\sum x_i\right)^2}{n \sum y_i^2 - \left(\sum y_i\right)^2}.$$

2.7.8 The R^2 value for a linearizable model is defined to be the R^2 value for the straight line fit to the transformed data. We might wonder if we could also use the definition of R^2, Equation (2.16), directly from the model. The answer is, it depends.

a. In the worksheet **ln**, use Equation (2.16) to calculate the R^2 value directly from the model. This means, use the logarithmic model to calculate $\hat{y}_i$ (the predicted values), and then use these values to calculate SS_{Res}. Calculate SS_{Tot} using the definition (note that the y-values in the definition are the values of L), and then use Equation (2.16) to calculate R^2. Does this give the same results as the power trendline function and the linear trendline fit to the transformed data?

b. Repeat part a. for the worksheet **Power**.

c. Repeat part a. for the worksheet **Exponential**.

For Further Reading

- For more examples of modeling with proportionality, see GIORDANO/WEIR/FOX, *A First Course in Mathematical Modeling*, 3e, 2003, pages 95 – 96, ©Brooks/Cole, a part of Cengage Learning, Inc.
- For examples of modeling biological systems with proportionality, see A.J. Clark, *Comparative Physiology of the Heart*, Macmillan, 1927.
- For more information on least–squares solutions and their applications to linearizable models, see Lay, David C., *Linear Algebra and its Applications*, Third edition, Pearson Addison Wesley, 2006, pg. 409 – 425. Also see the references given on page 424 of this text.

3

Linear Algebra

Chapter Objectives

- Introduce the basics of using matrices and solving systems of equations
- Model with systems of equations
- Fit polynomial models to data
- Introduce multiple regression
- Introduce spline models

3.1 Linear Algebra Basics

Elementary linear algebra deals with solving systems of linear equations and operations on matrices. As we will see throughout this chapter, systems of linear equations and matrices have a wide variety of applications. In this section we introduce some basic matrix operations and techniques for solving systems of equations used in this book. We begin with an example of solving simple systems of linear equations.

Example 3.1.1 (Systems of Linear Equations)
Consider the system of linear equations

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

A *solution* to this system is a value of x and a value of y that satisfy both equations simultaneously. To algebraically find a solution, if it exists, one approach is to multiply the equations by appropriate non-zero constants and add them in such a way that only one variable remains. For instance, we could multiply the first equation by -3:

$$\begin{aligned} -3x - 6y &= -6 \\ 3x - 4y &= 6, \end{aligned}$$

and then add the two equations together to get

$$0x - 10y = 0,$$

which can easily be solved to get $y = 0$. Then we can plug this value into one of the original equations, say the first equation, to get

$$x + 2(0) = 2 \quad \Rightarrow \quad x = 2.$$

Thus the unique solution is $x = 2$, $y = 0$. Graphically, we can think of solving a system of equations such as this as finding the point of intersection of the lines $x + 2y = 2$ and $3x - 4y = 6$ (or equivalently, $y = -x/2 + 1$ and $y = 3x/4 - 3/2$). This graph is shown in Figure 3.1. We see the lines intersect at the unique point (2, 0).

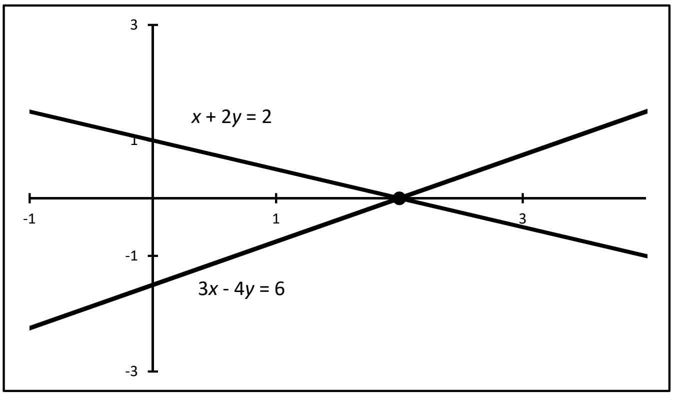


FIGURE 3.1

Not every system of linear equations has a unique solution. Consider the system

$$x + 2y = 2$$
$$x + 2y = -2.$$

The graphs of these two lines are shown in Figure 3.2. We see that the lines are parallel, so they never intersect meaning there is no solution.

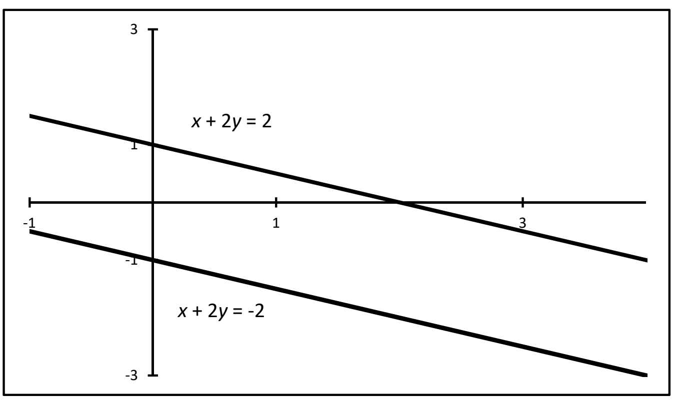


FIGURE 3.2

Other systems have infinitely many solutions. Consider the system

$$x + 2y = 2$$
$$2x + 4y = 4.$$

Note that the second equation is simply 2 times the first equation. This means that any solution to the first equation will be a solution to the second equation. Graphically, this means these two equations have the same graph. The graphs of these two lines are shown in Figure 3.3. There appears to be only one line because the lines lie on top of each other. These two lines intersect in infinitely many points, so there are infinitely many solutions to this system.

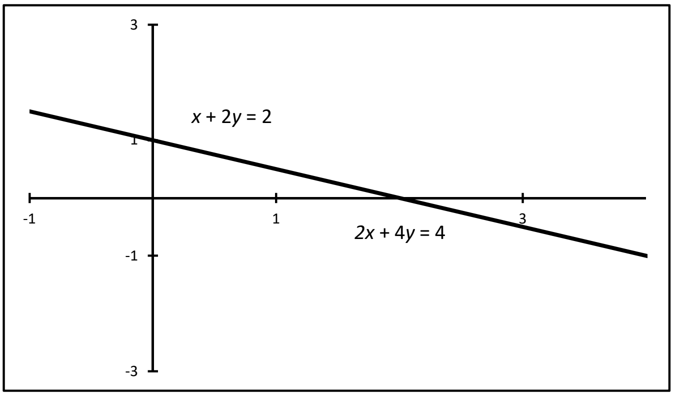


FIGURE 3.3

This analysis applies to systems with any number of variables. In general, a system of linear equations can have either one unique solution, no solution, or infinitely many solutions. □

Observe that the method for algebraically solving a system of equations presented in Example 3.1.1 involved only arithmetic and that the most important part of the process is the numbers involved. If we keep our work neatly organized, we don't really need to write the variables in each step. To this end, we write the coefficients of the variables and the constants on the right-hand sides of the equations in the following form:

$$\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 2 \\ 6 \end{bmatrix}.$$

This leads to our definition of a *matrix* and a *vector*.

Definition 3.1.1.

- A *matrix* is a two-dimensional array of numbers.
- The *size* of a matrix is denoted $a \times b$ where a is the number of rows and b is the number of columns (rows are horizontal and columns are vertical).
- A *column vector* is a matrix with one column and a *row vector* is a matrix with one row. In this book when we refer to a generic *vector*, we mean a column vector.
- The size of a vector is denoted by the number of rows and is referred to as the *dimension* of the vector.
- The symbol R^n denotes the set of all vectors of dimension n.

Matrices are usually named with capital letters, such as A, and vectors are named with bold-face lower-case letters, such as $\boldsymbol{b}$. When writing a vector by hand, we often use a lower-case letter with an overhead arrow, such as $\vec{b}$. For example,

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

is a 2×3 matrix and

$$\boldsymbol{b} = \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix}$$

is a 3×1 matrix, or a 3-dimensional vector. A vector can be written vertically using square brackets, as above, or it may be written horizontally using parentheses, as $\boldsymbol{b} = (-2, 1, 3)$.

The numbers inside a matrix are called *elements*. The position of an element is described by its row and column. For example, the element -1 in matrix A is in row 2, column 3. We say that it is in location (2, 3).

Next we present by example some operations we can do with matrices and vectors.

Example 3.1.2 (Scalar Multiplication)
In the context of linear algebra, a *scalar* is another word for a real number. So scalar multiplication means we multiply a real number by a matrix or vector. To do this, we simply multiply each element of the matrix or vector by the real number. For the matrix A and vector $\boldsymbol{b}$ defined above we can calculate, for example,

$$3A = \begin{bmatrix} 3(1) & 3(2) & 3(3) \\ 3(4) & 3(0) & 3(-1) \end{bmatrix} = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 0 & -3 \end{bmatrix} \text{ and } -2\boldsymbol{b} = \begin{bmatrix} -2(-2) \\ -2(1) \\ -2(3) \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \\ -6 \end{bmatrix}.$$

Note that we cannot do a calculation such as $3 + A$ or $-2 - \boldsymbol{b}$. In other words, we cannot do "scalar addition" or "scalar subtraction," although we can add or subtract two scalars. ☐

Example 3.1.3 (Matrix Addition)
Matrices of the same size may be added by adding corresponding entries. For example

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}.$$

Subtraction works in the same way. Note that only matrices, or vectors, of the same size may be added or subtracted. ☐

Example 3.1.4 (Matrix Multiplication)
Consider multiplying the matrix A times the vector $\boldsymbol{b}$ as defined above. This can be done as follows:

$$A\boldsymbol{b} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 0 & -1 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1(-2) + 2(1) + 3(3) \\ 4(-2) + 0(1) + (-1)(3) \end{bmatrix} = \begin{bmatrix} 9 \\ -11 \end{bmatrix}$$

Note that this calculation involves multiplying a 2×3 matrix by a 3×1 matrix and the product is a 2×1 matrix. For a product of matrices to be defined, the "inside" dimensions of the factors must be equal. The "outside" dimensions of the factors give the size of the product. More precisely, the column dimension of the first matrix must equal the row dimension of the second matrix. The row dimension of the first matrix and the column dimension of the second matrix give the size of the product. In this example the product $\boldsymbol{b}A$ is not defined because the dimensions do not match up. Also note that we do not use the symbols $\times$ or $\cdot$ to denote matrix multiplication. These symbols are reserved for other linear algebra operations.

To multiply $A\boldsymbol{b}$ in Excel, format a blank worksheet as in Figure 3.4.

Next, select the range **G2:B2**, type **=MMULT(A2:C3,E2:E4)**, and press the combination of keys **Ctrl-Shift-Enter** (this combination tells Excel to compute an array formula). The results are shown in Figure 3.5. Notice that when you select any cell in the

	A	B	C	D	E	F	G
1		**A**			**b**		**Ab**
2	1	2	3		-2		
3	4	0	-1		1		
4					3		

FIGURE 3.4

	G
1	**Ab**
2	9
3	-11

FIGURE 3.5

range **G2:G3**, the formula is in curly brackets, {...}. This indicates that an array formula has been entered. Also note that to perform a matrix calculation in Excel, you must first know the size of the result.

We can also multiply a matrix by a non-vector matrix as follows:

$$CD = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 1(-1)+2(-2) & 1(0)+2(1) \\ 3(-1)+4(-2) & 3(0)+4(1) \end{bmatrix} = \begin{bmatrix} -5 & 2 \\ -11 & 4 \end{bmatrix}$$

The product DC can be calculated,

$$DC = \begin{bmatrix} -1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} -1(1)+0(3) & -1(2)+0(4) \\ -2(1)+1(3) & -2(2)+1(4) \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ 1 & 0 \end{bmatrix},$$

but note that $CD \neq DC$. This illustrates that matrix multiplication is not commutative.

To calculate the product CD in Excel, format a blank worksheet as in Figure 3.6.

	A	B	C	D	E	F	G	H
1	**C**			**D**			**CD**	
2	1	2		-1	0			
3	3	4		-2	1			

FIGURE 3.6

Next, select the range **G2:H3**, type **=MMULT(A2:B3,D2:E3)**, and press **Ctrl-Shift-Enter**. The result is shown in Figure 3.7.

	G	H
1	**CD**	
2	-5	2
3	-11	4

FIGURE 3.7

□

Example 3.1.5 (Transpose)
To calculate the *transpose* of a matrix A, denoted A^T, we simply switch the rows and columns. Using the matrix A given above,

$$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 0 \\ 3 & -1 \end{bmatrix}$$

Note that the dimensions of A^T are the dimensions of A switched around. To perform this calculation in Excel, format a blank worksheet as in Figure 3.8.

	A	B	C	D	E	F
1		**A**			**A^T**	
2	1	2	3			
3	4	0	-1			

FIGURE 3.8

Next, select the range **E2:F4**, type **=TRANSPOSE(A2:C3)**, and press **Ctrl-Shift-Enter**. □

Example 3.1.6 (Length of a Vector)
Geometrically a 2-dimensional vector $\boldsymbol{b} = (b_1, b_2)$ can be thought of as an arrow that starts at the origin $(0, 0)$ and terminates at the point (b_1, b_2) on the $x - y$ plane. The *length* of the vector $\boldsymbol{b}$, denoted $\|\boldsymbol{b}\|$, is the distance of this point from the origin. The length is also called the *norm* or *Euclidean norm* of the vector and is calculated as

$$\|\boldsymbol{b}\| = \sqrt{b_1^2 + b_2^2}.$$

This definition can be generalized to vectors of any dimension. The *distance* between two vectors $\boldsymbol{d}$ and $\boldsymbol{c}$ of the same dimension is defined as

$$\|\boldsymbol{d} - \boldsymbol{c}\|$$

where $\boldsymbol{d} - \boldsymbol{c}$ is the component-by-component difference of the vectors. For $\boldsymbol{d} = (1, 1)$ and $\boldsymbol{c} = (2, 1)$ we have

$$\|\boldsymbol{d} - \boldsymbol{c}\| = \left\| \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 2 \\ 1 \end{bmatrix} \right\| = \left\| \begin{bmatrix} -1 \\ 0 \end{bmatrix} \right\| = \sqrt{(-1)^2 + 0^2} = 1.$$

To perform this calculation in Excel, format a blank worksheet as in Figure 3.9.

	A	B	C	D	E	F	G
1	**d**		**c**		**d - c**		**|d - c|**
2	1		2		=A2-C2		=SQRT(SUMSQ(E2:E3))
3	1		1		=A3-C3		

FIGURE 3.9

□

Next we defined an important special matrix, the *identity matrix*.

Definition 3.1.2. The *identity matrix* I_n is an $n \times n$ matrix (meaning the number of rows equals the number of columns, called a *square* matrix) with 1's along the main diagonal (from the top left corner to the bottom right corner) and 0's everywhere else. For example, the 2×2 identity matrix is

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

I_n is called an identity matrix because $I_n\boldsymbol{v} = \boldsymbol{v}$ for any n-dimensional vector $\boldsymbol{v}$.

Earlier we defined matrix multiplication, so the reader might wonder about matrix division. We cannot divide matrices, but we can multiply a matrix by its inverse, when it exists.

Definition 3.1.3. The inverse of an $n \times n$ matrix A, denoted A^{-1}, is a matrix such that

$$AA^{-1} = A^{-1}A = I_n.$$

We stress three important points about inverse matrices:

1. Only square matrices can have inverses.
2. Not every square matrix A has an inverse. If an inverse exists, A is said to be *invertible.*
3. Calculating an inverse is not a trivial matter (see Exercise 3.1.10 for one invertibility requirement and related algorithm).

See, for example, Lay, D., *Linear Algebra and its Applications*, 5^{th} edition, 2016, Pearson, for more details about matrix inverses.

Example 3.1.7 (Calculating a Matrix Inverse)

To calculate the inverse of the matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$, format a blank worksheet as in Figure 3.10.

	A	B	C	D	E
1	**A**			$\mathbf{A^{-1}}$	
2	1	2			
3	3	-4			

FIGURE 3.10

Next, select the range **C2:E3**, type **=MINVERSE(A2:B3)**, and press **Ctrl-Shift-Enter**. The result is shown in Figure 3.11.

	D	E
1	$\mathbf{A^{-1}}$	
2	0.4	0.2
3	0.3	-0.1

FIGURE 3.11

The reader should confirm by direct calculation that $AA^{-1} = A^{-1}A = I_2$. □

Example 3.1.8 (Systems of Linear Equations)
Consider again the system of linear equations

$$\begin{aligned} x + 2y &= 2 \\ 3x - 4y &= 6 \end{aligned}$$

from Example 3.1.1 which we solved algebraically and graphically. Now we show how it can be solved with matrices. First we write the system in the following matrix form:

$$\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 6 \end{bmatrix},$$

or more generically, $A\boldsymbol{x} = \boldsymbol{b}$. Matrix $A = \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$ is called the *coefficient matrix*, vector $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ is called the *unknown vector*, and vector $\boldsymbol{b} = \begin{bmatrix} 2 \\ 6 \end{bmatrix}$ is called the *constant vector*. To solve a system $A\boldsymbol{x} = \boldsymbol{b}$ where A is $n \times n$ we can multiply both sides by A^{-1} (assuming A^{-1} exists):

$$\begin{aligned} A^{-1}\left(A\boldsymbol{x} = \boldsymbol{b}\right) \\ A^{-1}A\boldsymbol{x} &= A^{-1}\boldsymbol{b} \\ I_n\boldsymbol{x} &= A^{-1}\boldsymbol{b} \\ \boldsymbol{x} &= A^{-1}\boldsymbol{b}. \end{aligned}$$

To solve this particular system in Excel, format a blank worksheet as in Figure 3.12.

	A	B	C	D
1	**A**			**b**
2	1	2		2
3	3	-4		6
4				
5	**A^{-1}**			**$A^{-1}b$**
6				
7				

FIGURE 3.12

Next, select the range **A6:B7**, type **=MINVERSE(A2:B3)**, and press **Ctrl-Shift-Enter**. Then select the range **D6:D7**, type **=MMULT(A6:B7,D2:D3)**, and press **Ctrl-Shift-Enter**. The result is shown in Figure 3.13.

	A	B	C	D
5	**A^{-1}**			**$A^{-1}b$**
6	0.4	0.2		2
7	0.3	-0.1		0

FIGURE 3.13

The results in the range D6:D7 show that the solution to the system is $x = 2$, $y = 0$, the same as is found algebraically and graphically in Example 3.1.1. Observe that the coefficient matrix in this example is invertible and there is a unique solution. This observation is true in general for systems with any number of variables. □

Example 3.1.9 (Augmented Matrices)
A linear system such as the previous example can also be solved using an *augmented matrix*. To form this augmented matrix we combine matrix A and vector $\boldsymbol{b}$ into one 2×3 matrix:

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

Next we perform *elementary row operations* on the augmented matrix. There are three possible elementary row operations:

1. Switch two rows of a matrix.

2. Multiply all elements of a row by a nonzero constant.

3. Add the elements of a row to the corresponding elements of another row.

These row operations are inspired by the algebraic solution in Example 3.1.1. Any combination of these operations is also allowed. In this example, we first multiply the first row by -3 and add it to the second row and place the result in row 2, denoted $-3R_1 + R_2 \to R_2$:

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & -10 & 0 \end{bmatrix}$$

The first row does not change. Next we perform $-(1/10)R_2 \to R_2$:

$$\begin{bmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

Lastly we perform $-2R_2 + R_1 \to R_1$:

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$$

Call this matrix B. The sequence of operations we just performed is generically called *row-reduction* or *Gauss-Jordan elimination*. Matrix B is called the *reduced row echelon form* (RREF) of the system. We say that B is *row equivalent* to the original matrix. This means that B represents a system of equations with the same set of solutions as the original system. Converting B to a system of equations yields $x = 2$ and $y = 0$, the solution to the original system.

A matrix is in RREF if the following two conditions are met:

1. The first non-zero element in each row is a 1 (called a *leading* 1).

2. Each element above or below a leading 1 is 0.

Once an augumented matrix is reduced to its RREF, we can simply read the solution off the matrix (assuming there is a unique solution). No additional work is necessary. We can show theoretically that every matrix is row equivalent to exactly one matrix in RREF. The sequence of row operations used to find the RREF of a matrix is not unique, but the final result is unique.

We can perform this row-reduction in Excel by formatting a blank worksheet as in Figure 3.14. Note that the entries in column E simply denote the operations done in each step. The ' (a single quotation mark) at the beginning of each cell prevents Excel from interpreting the entry as a formula.

	A	B	C	D	E
1	**Step 1**	1	2	2	
2		3	-4	6	
3					
4	**Step 2**	=B1	=C1	=D1	
5		=-3*B1+B2	=-3*C1+C2	=-3*D1+D2	'-3R1 + R2
6					
7	**Step 3**	=B4	=C4	=D4	
8		=-1/10*B5	=-1/10*C5	=-1/10*D5	'-(1/10)R2
9					
10	**Step 4**	=-2*B8+B7	=-2*C8+C7	=-2*D8+D7	'-2R2 + R1
11		=B8	=C8	=D8	

FIGURE 3.14

The final result is shown in Figure 3.15.

	A	B	C	D	E
10	**Step 4**	1	0	2	-2R2 + R1
11		0	1	0	

FIGURE 3.15

□

Example 3.1.10 (Solving a System with 3 Variables)
Row-reduce the augmented matrix to its RREF to solve the system

$$\begin{aligned} 2x + y + 3z &= 10 \\ x + y + z &= 6 \\ x + 3y + 2z &= 13 \end{aligned}$$

The augmented matrix is

$$\begin{bmatrix} 2 & 1 & 3 & 10 \\ 1 & 1 & 1 & 6 \\ 1 & 3 & 2 & 13 \end{bmatrix}$$

First we perform the row operations $R_1 - 2R_2 \to R_2$ and $R_1 - 2R_3 \to R_3$:

$$\begin{bmatrix} 2 & 1 & 3 & 10 \\ 0 & -1 & 1 & -2 \\ 0 & -5 & -1 & -16 \end{bmatrix}$$

Next we perform $-5R_2 + R_3 \to R_3$:

$$\begin{bmatrix} 2 & 1 & 3 & 10 \\ 0 & -1 & 1 & -2 \\ 0 & 0 & -6 & -6 \end{bmatrix}$$

Then we perform $R_1 + R_2 \to R_1$, $-R_2 \to R_2$, and $-(1/6)R_3 \to R_3$:

$$\begin{bmatrix} 2 & 0 & 4 & 8 \\ 0 & 1 & -1 & 2 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

Lastly, we perform $(1/2)R_1 - 2R_3 \to R_1$ and $R_2 + R_3 \to R_2$:

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

This final matrix is in RREF and we see the solution is $x = 2$, $y = 3$, and $z = 1$.

To do this row-reduction in Excel, format a blank worksheet as in Figure 3.16.

	A	B	C	D	E	F
1	**Step 1**	2	1	3	10	
2		1	1	1	6	
3		1	3	2	13	
4						
5	**Step 2**	=B1	=C1	=D1	=E1	
6		=B1-2*B2	=C1-2*C2	=D1-2*D2	=E1-2*E2	'R1 - 2R2
7		=B1-2*B3	=C1-2*C3	=D1-2*D3	=E1-2*E3	'R1 - 2R3
8						
9	**Step 3**	=B5	=C5	=D5	=E5	
10		=B6	=C6	=D6	=E6	
11		=-5*B6+B7	=-5*C6+C7	=-5*D6+D7	=-5*E6+E7	'-5R2 + R3
12						
13	**Step 4**	=B9+B10	=C9+C10	=D9+D10	=E9+E10	'R1 + R2
14		=-B10	=-C10	=-D10	=-E10	'-R2
15		=-1/6*B11	=-1/6*C11	=-1/6*D11	=-1/6*E11	'-(1/6)R3
16						
17	**Step 5**	=1/2*B13-2*B15	=1/2*C13-2*C15	=1/2*D13-2*D15	=1/2*E13-2*E15	'(1/2)R1 - 2R3
18		=+B14+B15	=+C14+C15	=+D14+D15	=+E14+E15	'R2 + R3
19		=B15	=C15	=D15	=E15	

FIGURE 3.16

The final result is shown in Figure 3.17.

	A	B	C	D	E	F
17	**Step 5**	1	0	0	2	(1/2)R1 - 2R3
18		0	1	0	3	R2 + R3
19		0	0	1	1	

FIGURE 3.17

□

Example 3.1.11 (A System with No Solution)
As shown in Example 3.1.1, some systems have no solution. In this example we examine what this means in terms of an augmented matrix. Consider the system

$$\begin{aligned} 3x + 6y - 3z &= 6 \\ 3x + 6y - 2z &= 10 \\ -2x - 4y - 3z &= -1 \end{aligned}$$

The augmented matrix for this system is

$$\begin{bmatrix} 3 & 6 & -3 & 6 \\ 3 & 6 & -2 & 10 \\ -2 & -4 & -3 & -1 \end{bmatrix}$$

Row-reducing this matrix with an appropriate sequence of row operations yields

$$\begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Converting this last row to an equation yields $0 = 1$, an obvious contradiction. This means there is no solution to this system. □

Example 3.1.12 (A System with Many Solutions)
As shown in Example 3.1.1, some systems have infinitely many solutions. In this example we examine what this means in terms of the reduced form of an augmented matrix. As we will see in the next section, systems of this type occur frequently in applications. Consider the system

$$\begin{aligned} x + 2y - z &= 3 \\ 2x + 4y - 2z &= 6 \\ 3x + 6y + 2z &= -1 \end{aligned}$$

The augmented matrix for this system is

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

Row-reducing this matrix with an appropriate sequence of row operations yields

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

Converting this augmented matrix back to a system of equations yields

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

The last equation is meaningless. The second equation gives us the value of z. The first equation doesn't give a specific value of any variable, but it does give us a relationship between x and y. We can rewrite this equation as $x = 1 - 2y$. The variable y is called a *free variable*, meaning it can take any value it wants. We rewrite the result in this form:

$$\begin{aligned} x &= 1 - 2y \\ y &\text{ is free} \\ z &= -2 \end{aligned}$$

This result is called the *general solution* of the system. We get a *specific solution* by choosing a value of y and calculating x. For instance, choosing $y = 1$ yields the specific solution $x = -1$, $y = 1$, $z = 0$. □

Exercises

Directions: In Exercises 3.1.1 - 3.1.6, calculate each in Excel. Make sure you label the calculations in the worksheets appropriately as done in the examples.

3.1.1 $\begin{bmatrix} 1 & 7 & -8 \\ 5 & 0 & -1 \\ 2 & 7 & 0 \end{bmatrix} \begin{bmatrix} 4 \\ 0 \\ -3 \end{bmatrix}$

3.1.2 $\begin{bmatrix} 1 & 7 & 5 & -1 \\ -5 & 0 & -3 & 2 \\ 7 & 1 & 6 & -3 \end{bmatrix} \begin{bmatrix} 5 & 7 \\ 3 & 3 \\ 5 & -5 \\ 8 & 1 \end{bmatrix}$

3.1.3 $\begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 4 & 5 \end{bmatrix}^T$

3.1.4 Find the distance between $\begin{bmatrix} 1 \\ 7 \\ -9 \end{bmatrix}$ and $\begin{bmatrix} -9 \\ 1 \\ 7 \end{bmatrix}$.

3.1.5 Let $B = \begin{bmatrix} 7 & 1 & -8 \\ 2 & 7 & -3 \\ -1 & 0 & 2 \end{bmatrix}$. Calculate B^{-1} and show that $B^{-1}B = I_3$.

3.1.6 Use matrices to solve the system

$$\begin{aligned} 2x + 3y + z &= 2 \\ -3x - 4y + 9z &= 1 \\ 8x + 6y - 9z &= 3. \end{aligned}$$

Directions: In Exercises 3.1.7 - 3.1.9, write the augmented matrix for each system of linear equations and find the solution by row-reducing.

3.1.7

$$\begin{aligned} x_1 + 5x_2 &= 7 \\ -2x_1 - 7x_2 &= -5 \end{aligned}$$

3.1.8

$$\begin{aligned} 2x_1 + 4x_2 &= -4 \\ 5x_1 + 7x_2 &= \ \ 11 \end{aligned}$$

3.1.9

$$\begin{aligned} x_1 - 3x_2 &= 5 \\ -x_1 + \ \ x_2 + 5x_3 &= 2 \\ x_2 + \ \ x_3 &= 0 \end{aligned}$$

3.1.10 One theoretical requirement for an $n \times n$ matrix A to be invertible is that it be row equivalent to the $n \times n$ identity matrix I_n. That is, we must be able to row-reduce A to I_n. This leads to the following algorithm for finding A^{-1}:

1. Form the augmented matrix $B = [A \,|\, I_n]$. That is, form an $n \times 2n$ matrix where the left half is A and the right half is I_n.

2. Row-reduce B to the form $\left[I_n \,|\, A^{-1}\right]$. That is, row-reduce B until the left half is I_n. The right half will be A^{-1}.

Use this algorithm to find the inverse of each of the following matrices. Show the row-reduction steps used and check your work by using the MINVERSE function as in Example 3.1.7.

a. $\begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}$

b. $\begin{bmatrix} 1 & -1 & -2 \\ 2 & -3 & -5 \\ -1 & 3 & 5 \end{bmatrix}$

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

3.1.11 A simple formula for the inverse of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is

$$A^{-1} = \frac{1}{(ad - bd)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

The quantity $(ad - bd)$ is called the *determinant*. A is invertible if and only if its determinant is not 0.

a. Show that

$$\frac{1}{(ad - bd)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = I_2 \text{ and } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \left(\frac{1}{(ad - bd)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \right) = I_2.$$

That is, show that the formula for A^{-1} satisfies the definition of the inverse of A. Note that $1/\,(ad - bd)$ is simply a scalar. You may do the matrix multiplication first and then multiply the product by the scalar.

b. Use this formula to find the inverse of $A = \begin{bmatrix} 4 & -6 \\ 2 & -2 \end{bmatrix}$.

3.2 Modeling with Systems of Equations

Elementary applications of linear algebra often involve describing a scenario with a system of equations and then solving the system. In this section we present four such applications: traffic flow, balancing chemical equations, and two economics applications.

Example 3.2.1 (Traffic Flow)
Consider the network of streets in Figure 3.18. The numbers and variables in the figure represent the traffic flow, measured in vehicles per hour (vph), along each street segment. The goal is to find the values of x_1 through x_4 and analyze the flow of traffic.

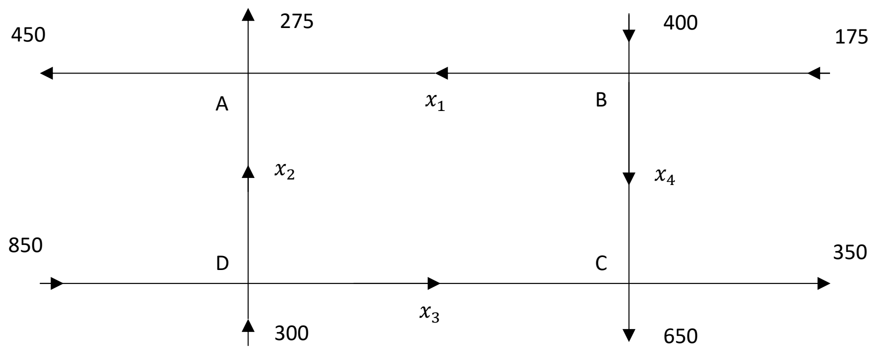


FIGURE 3.18

Our model is based on the following rather obvious assumption:

All traffic that enters an intersection must leave the intersection.

We organize the flow in and out of each intersection in the following table:

Intersection	Traffic In	Traffic Out	Equation
A	$x_1 + x_2$	$450 + 275 = 725$	$x_1 + x_2 = 725$
B	$400 + 175 = 575$	$x_1 + x_4$	$575 = x_1 + x_4$
C	$x_3 + x_4$	$650 + 350 = 1000$	$x_3 + x_4 = 1000$
D	$850 + 300 = 1150$	$x_2 + x_3$	$1150 = x_2 + x_3$

Rewriting the four equations results in the final model:

$$\begin{aligned} x_1 + x_2 &= 725 \\ x_1 + x_4 &= 575 \\ x_3 + x_4 &= 1000 \\ x_2 + x_3 &= 1150 \end{aligned}$$

To solve this system, we use techniques from Section 3.1. We can write this system in the form $A\boldsymbol{x} = \boldsymbol{b}$ and try to calculate $\boldsymbol{x} = A^{-1}\boldsymbol{b}$, but it turns out the matrix A is not invertible.

Thus we must use row-reduction. We write the augmented matrix and row-reduce it:

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 725 \\ 1 & 0 & 0 & 1 & 575 \\ 0 & 0 & 1 & 1 & 1000 \\ 0 & 1 & 1 & 0 & 1150 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 1 & 575 \\ 0 & 1 & 0 & -1 & 150 \\ 0 & 0 & 1 & 1 & 1000 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

This yields the general solution

$$\begin{aligned} x_1 &= 575 - x_4 \\ x_2 &= 150 + x_4 \\ x_3 &= 1000 - x_4 \\ x_4 &\text{ is free} \end{aligned}$$

The fact that there is a free variable means that traffic can flow through this network in infinitely many ways, a result that should not be surprising. For example, if we know that on a given day, $x_4 = 250$, we would have $x_1 = 325$, $x_2 = 400$, and $x_3 = 750$ vph.

With this general solution we can do more than simply calculate specific traffic flows. From the equation $x_1 = 575 - x_4$ we see that the maximum value of x_4 is 575 vph; otherwise x_1 would be negative. Also, the equation $x_2 = 150 + x_4$ tells us the minimum value of x_2 is 150 since x_4 cannot be negative. Thus if we were planning roadwork along the road segment corresponding to x_2 resulting in restricted traffic flow, we would need to make sure the road could handle at least 150 vph. □

Example 3.2.2 (Balancing Chemical Equations)
In a chemical reaction, substances (called *reactants*, often in the form of *molecules*) composed of atoms react to form new substances (called *products*). A simple assumption behind every chemical reaction is:

Atoms cannot be created or destroyed in a reaction.

As a simple example, sulfur (S_8) reacts with oxygen (O_2) to form sulfur dioxide (SO_2). This is represented by the chemical equation

$$S_8 + O_2 \to SO_2.$$

A quick inspection of this equation reveals there are more sulfur atoms (S) on the left than the right, which violates the assumption. Thus we need to determine how many of each molecule react to form how many molecules of the product. This is called *balancing the equation*. By inspection, a balanced version of this equation is

$$S_8 + 8O_2 \to 8SO_2.$$

Not every chemical equation can be balanced so easily by inspection. For example, ethylenediamine ($C_2H_8N_2$) reacts with dinitrogen tetroxide (N_2O_4) to form nitrogen (N_2), carbon dioxide (CO_2), and water (H_2O) according to the equation

$$x_1 C_2H_8N_2 + x_2 N_2O_4 \to x_3 N_2 + x_4 CO_2 + x_5 H_2O.$$

The coefficients $x_1, \ldots, x_5$ represent the unknown number of each molecule involved. The goal is to find appropriate integer values of these unknowns. We begin by equating the number of each atom from each side of the equation:

Atom	**Equation**
C	$2x_1 = x_4$
H	$8x_1 = 2x_5$
N	$2x_1 + 2x_2 = 2x_3$
O	$4x_2 = 2x_4 + x_5$

Rewriting the four equations results in the final model:

$$\begin{aligned} 2x_1 \qquad\qquad\quad - x_4 \qquad &= 0 \\ 8x_1 \qquad\qquad\qquad\qquad - 2x_5 &= 0 \\ 2x_1 + 2x_2 - 2x_3 \qquad\qquad\quad &= 0 \\ 4x_2 \qquad - 2x_4 - x_5 &= 0 \end{aligned}$$

We solve this model by row-reduction:

$$\begin{bmatrix} 2 & 0 & 0 & -1 & 0 & 0 \\ 8 & 0 & 0 & 0 & -2 & 0 \\ 2 & 2 & -2 & 0 & 0 & 0 \\ 0 & 4 & 0 & -2 & -1 & 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 0 & -1/4 & 0 \\ 0 & 1 & 0 & 0 & -1/2 & 0 \\ 0 & 0 & 1 & 0 & -3/4 & 0 \\ 0 & 0 & 0 & 1 & -1/2 & 0 \end{bmatrix}$$

This yields the general solution

$$\begin{aligned} x_1 &= 1/4\, x_5 \\ x_2 &= 1/2\, x_5 \\ x_3 &= 3/4\, x_5 \\ x_4 &= 1/2\, x_5 \\ x_5 &\text{ is free} \end{aligned}$$

To find a particular solution, we choose the smallest positive integer value of the free variable so that all the other variables are positive integers. We choose $x_5 = 4$, yielding $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, and $x_4 = 2$ so that the balanced equation is

$$C_2H_8N_2 + 2N_2O_4 \rightarrow 3N_2 + 2CO_2 + 4H_2O.$$

□

Example 3.2.3 (Equilibrium Prices)
Consider a simple economy that consists of three sectors: Chemical, Electric, and Metal. Each sector produces some yearly output (measured in millions of dollars). Each sector also consumes a certain proportion of the output of the other sectors as illustrated in Figure 3.19. (A sector consuming a portion of its own output can be thought of as an operational expense.)

Table 3.1, called an *exchange table*, summarizes this information. Note that each column in this table adds up to exactly 1 to model the fact that each sector's output is totally consumed by the other sectors.
A reasonable assumption about a healthy economy is:

The total expenses of each sector must equal its output.

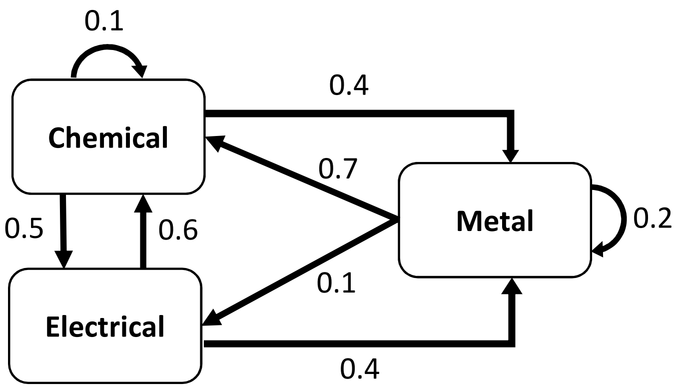


FIGURE 3.19

TABLE 3.1

Proportion of Output from			
Chemical	**Electrical**	**Metal**	**Purchased by**
0.1	0.6	0.7	**Chemical**
0.5	0	0.1	**Electrical**
0.4	0.4	0.2	**Metal**

In other words, a sector's expenses must equal its income. An economy satisfying this assumption is said to be in *equilibrium*, and the resulting outputs are called *equilibrium prices*. To find equilibrium prices, let x_c, x_e, and x_m denote the outputs from chemical, electrical, and metal, respectively. The assumption yields the system of equations

$$\begin{aligned} 0.1x_c + 0.6x_e + 0.7x_m &= x_c \\ 0.5x_c \qquad\qquad + 0.1x_m &= x_e \\ 0.4x_c + 0.4x_e + 0.2x_m &= x_m \end{aligned}$$

We rewrite this system by subtracting the quantities on the right hand side of the equation:

$$\begin{aligned} -0.9x_c + 0.6x_e + 0.7x_m &= 0 \\ 0.5x_c - \quad x_e + 0.1x_m &= 0 \\ 0.4x_c + 0.4x_e - 0.8x_m &= 0 \end{aligned}$$

Then we solve by row-reduction:

$$\left[\begin{array}{rrrr} -0.9 & 0.6 & 0.7 & 0 \\ 0.5 & -1 & 0.1 & 0 \\ 0.4 & 0.4 & -0.8 & 0 \end{array}\right] \quad \Rightarrow \quad \left[\begin{array}{rrrr} 1 & 0 & -19/15 & 0 \\ 0 & 1 & -11/15 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This yields the general solution

$$\begin{aligned} x_c &= 19/15\, x_m \\ x_e &= 11/15\, x_m \\ x_m &\text{ is free.} \end{aligned}$$

Any nonnegative choice for x_m yields a specific set of equilibrium prices. For example, if $x_m = 15$, then $x_c = 19$ and $x_3 = 11$. □

The previous applications all involved systems with infinitely many solutions. The next example involves a unique solution.

Example 3.2.4 (The Leontief Input-Output Model)
In this example we present another economics application that resembles equilibrium prices, but there are some key differences. One difference is that there will be an "open" sector which consumes products but doesn't produce anything.

Consider an open economy with three sectors: mining, electrical, and auto. To produce \$1 of mined material, the mining operation must purchase \$0.1 of its own product, \$0.3 of electricity, and \$0.1 worth of autos for its transportation. To produce \$1 of electricity, it takes \$0.25 of mined material, \$0.4 of electricity, and \$0.15 of autos. Finally, to produce \$1 worth of autos, the auto-manufacturing plant must purchase \$0.2 of mined material, \$0.5 of electricity, and consumes \$0.1 of autos. Assume also that during a period of one year, the open sector has a demand of \$50 million worth of mined material, \$75 million worth of electricity, and \$125 million worth of autos. Find the annual output of each sector in order to satisfy demand.

To solve this problem we first organize the inter-sector demands in Table 3.2 called an *input-output matrix*. This matrix looks much like an exchange table, but the entries of the matrix represent dollar amounts, not proportions. Also note that the columns do not add up to one and that the columns represent the purchaser, not the rows as in an exchange table.

TABLE 3.2

\$ Purchased by			
Mining	**Electrical**	**Auto**	**Purchased from**
0.1	0.25	0.2	**Mining**
0.3	0.40	0.5	**Electrical**
0.1	0.15	0.1	**Auto**

A reasonable assumption about this scenario is:

The total demand from each sector must equal its output.

Let x_m, x_e, and x_a denote the annual outputs from mining, electrical, and auto, respectively. The assumption yields the system of equations

$$\begin{aligned} 0.1x_m + 0.25x_e + 0.2x_a + 50 &= x_m \\ 0.3x_m + 0.40x_e + 0.5x_a + 75 &= x_e \\ 0.1x_m + 0.15x_e + 0.1x_a + 125 &= x_a \end{aligned}$$

We could solve this system by rewriting and row-reducing like we did in Example 3.2.3, but we'll take a different approach. First we'll rewrite the system in matrix form

$$\begin{bmatrix} 0.1 & 0.25 & 0.2 \\ 0.3 & 0.40 & 0.5 \\ 0.1 & 0.15 & 0.1 \end{bmatrix} \begin{bmatrix} x_m \\ x_e \\ x_a \end{bmatrix} + \begin{bmatrix} 50 \\ 75 \\ 125 \end{bmatrix} = \begin{bmatrix} x_m \\ x_e \\ x_a \end{bmatrix}$$

or more generically,

$$A\boldsymbol{x} + \boldsymbol{d} = \boldsymbol{x}$$

where A is the input-output matrix, $\boldsymbol{d}$ is the demand vector from the open sector, and $\boldsymbol{x}$ is the unknown output vector. Then we rewrite this matrix equation:

$$\begin{aligned} \boldsymbol{x} - A\boldsymbol{x} &= \boldsymbol{d} \\ (I_n - A)\,\boldsymbol{x} &= \boldsymbol{d} \\ \boldsymbol{x} &= (I_n - A)^{-1}\,\boldsymbol{d}. \end{aligned}$$

For this example,

$$(I_n - A) = \begin{bmatrix} 0.9 & -0.25 & -0.2 \\ -0.3 & 0.60 & -0.5 \\ -0.1 & -0.15 & 0.9 \end{bmatrix}$$

Using the MINVERSE function in Excel, we calculate

$$(I_n - A)^{-1} \approx \begin{bmatrix} 1.465 & 0.803 & 0.772 \\ 1.008 & 2.488 & 1.606 \\ 0.331 & 0.504 & 1.465 \end{bmatrix}$$

Thus

$$\boldsymbol{x} \approx \begin{bmatrix} 1.465 & 0.803 & 0.772 \\ 1.008 & 2.488 & 1.606 \\ 0.331 & 0.504 & 1.465 \end{bmatrix} \begin{bmatrix} 50 \\ 75 \\ 125 \end{bmatrix} = \begin{bmatrix} 229.9 \\ 437.8 \\ 237.4 \end{bmatrix}$$

So mining should produce \$229.9 million, electricity \$437.8 million, and auto \$237.4 million. One benefit of doing the calculations with matrices rather than row-reduction is that if the demand from the open sector were to change, then all we need to do is change $\boldsymbol{d}$ in the calculation $\boldsymbol{x} = (I_n - A)^{-1}\,\boldsymbol{d}$. □

Exercises

3.2.1 Consider the road network in Figure 3.20.

a. Construct a system of linear equations that models the traffic flow and find the general solution.

b. If $x_4 = 250$ vph, find the values of x_1, x_2, and x_3.

c. What is the minimum possible value of x_4?

3.2.2 Consider the road network in Figure 3.21.

a. Construct a system of linear equations that models the traffic flow and find the general solution.

b. What is the minimum possible value of x_3?

c. What is the maximum possible value of x_4?

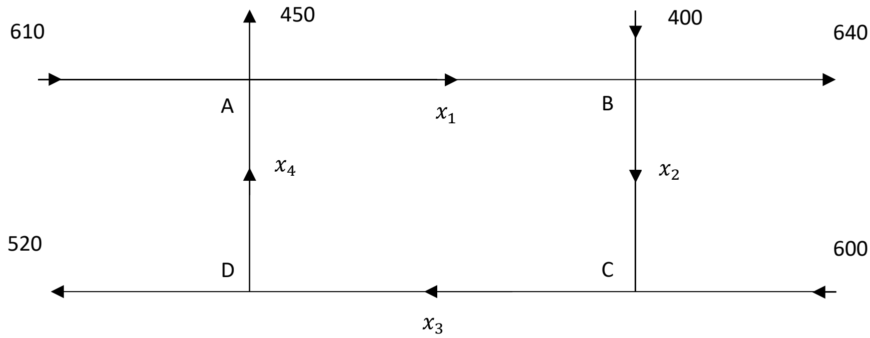


FIGURE 3.20

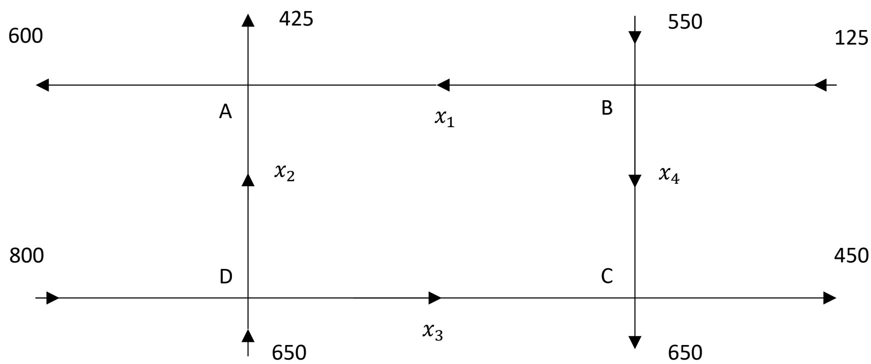


FIGURE 3.21

d. Suppose the 650 vph flowing into intersection D from the bottom were changed to 600 vph. Explain why there would be no solution to the resulting system of linear equations that models the traffic flow.

3.2.3 Consider the road network in Figure 3.22.

a. Construct a system of linear equations that models the traffic flow and find the general solution.

b. If $x_4 = 550$ vph and $x_5 = 100$ vph, find the values of x_2, x_3, and x_4.

c. What is the minimum possible value of x_4?

d. What is the minimum possible value of $x_4 - x_5$?

e. Can x_5 ever be greater than x_4? Briefly explain.

3.2.4 Balance each of the following chemical equations.

a. Propane reacts with oxygen to form carbon dioxide and water:

$$C_3H_8 + O_2 \to CO_2 + H_2O.$$

b. Potassium permanganate reacts with hydrochloric acid to form potassium chloride, manganese(II) chloride, water, and chlorine gas:

$$KMnO_4 + HCl \to KCl + MnCl_2 + H_2O + Cl_2.$$

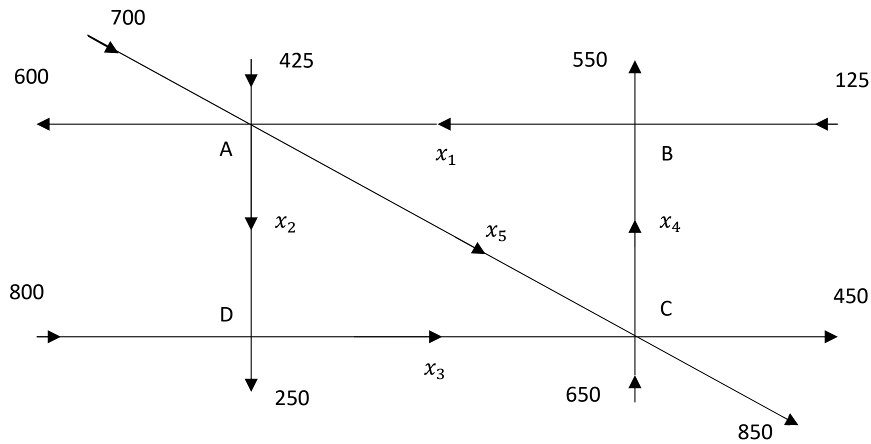


FIGURE 3.22

c. Benzoic acid reacts with oxygen to form carbon dioxide and water:

$$C_6H_5COOH + O_2 \rightarrow CO_2 + H_2O.$$

d. Sodium bicarbonate reacts with citric acid to form sodium citrate, water, and carbon dioxide:

$$NaHCO_3 + H_3C_6H_5O_7 \rightarrow Na_3C_6H_5O_7 + H_2O + CO_2.$$

3.2.5 Consider an economy with two sectors: Manufacturing and Services. Each year, Manufacturing sells 75% of its output to Services and the rest is an operational expense. Likewise, Services sells 85% of its output to Manufacturing and consumes the rest.

a. Set up an exchange table for this economy.

b. Find the equilibrium prices if the annual output from Services is $30 million.

3.2.6 Consider an economy whose exchange table is given below. Find the equilibrium prices if the annual output from Services is $46 million.

Prop. of Output from			
Auto	**Services**	**Food**	**Purchased by**
0.1	0.35	0.1	**Auto**
0.6	0.45	0.7	**Services**
0.3	0.2	0.2	**Food**

3.2.7 Consider an economy whose exchange table is given below. Find the equilibrium prices if the annual output from Chemical is $10 million.

Proportion of Output from					
Auto	**Metal**	**Electrical**	**Food**	**Chemical**	**Purchased by**
0.15	0.3	0.25	0.05	0.3	**Auto**
0.4	0.05	0.15	0.3	0.3	**Metal**
0.25	0.25	0.23	0.35	0.15	**Electrical**
0.15	0.25	0.15	0.15	0.2	**Food**
0.05	0.15	0.22	0.15	0.05	**Chemical**

3.2.8 Consider an economy consisting of two sectors, electrical and water, plus an open sector. Suppose that the production of each dollar of electricity requires \$0.3 of electricity and \$0.1 of water, and the production of each dollar of water requires \$0.2 of electricity and \$0.4 of water. The demand of the open sector is \$12 million for electricity and \$8 million for water. How much electricity and water should be produced to meet demand?

3.2.9 An economy is based on three sectors, agriculture, energy, and manufacturing, plus an open sector. Production of each dollar of agriculture requires an input of \$0.2 from the agriculture sector and \$0.4 from the energy sector. Production of each dollar of energy requires an input of \$0.2 from the energy sector and \$0.4 from the manufacturing sector. Production of each dollar of manufacturing requires an input of \$0.1 from the agriculture sector, \$0.1 from the energy sector, and \$0.3 from the manufacturing sector. Find the output from each sector that is needed to satisfy a demand of \$20 billion for agriculture, \$10 billion for energy, and \$30 billion for manufacturing from the open sector.

3.3 Polynomials

Polynomial models are often convenient to use because they are easy to differentiate and integrate. Theorem 3.3.1 is a well-known result from algebra about fitting a polynomial model to data.

Theorem 3.3.1. *Given a set of data, $\{(x_i, y_i) : i = 1, \ldots, n\}$, where $x_i \neq x_j$ for all $i \neq j$, there exists a unique polynomial $p(x)$ of degree at most $n - 1$ such that*

$$p(x_i) = y_i \text{ for all } i = 1, \ldots, n$$

□

Graphically, this theorem means that the graph of $y = p(x)$ goes through each data point (i.e. it is a perfect fitting model). This sounds like a utopia, but is it really?

Example 3.3.1 (Three Data Points)
Consider the problem of fitting a second-degree polynomial of the form $y = ax^2 + bx + c$ to the set of three data points $\{(1, 2), (2, 4), (3, 5)\}$. This means we want values of a, b, and c such that

$$\begin{aligned} a\left(1^2\right) + b(1) + c &= 2 \\ a\left(2^2\right) + b(2) + c &= 4 \\ a\left(3^2\right) + b(3) + c &= 5 \end{aligned}$$

This set of linear equations can be written in matrix form as

$$\begin{bmatrix} 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \\ 3^2 & 3 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix}$$

This matrix equation has the generic form $A\boldsymbol{x} = \boldsymbol{b}$ where $\boldsymbol{x} = (a, b, c)$ is the vector of unknowns. It can be shown that A is invertible. Thus there is a unique solution to this matrix equation, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. Doing this calculation using techniques from Section 3.1 yields the solution $\boldsymbol{x} = (-0.5, 3.5, -1)$ so that the model is $y = -0.5x^2 + 3.5x - 1$.

Excel will automatically calculate this polynomial for us using an algorithm equivalent to the one described above. Create a scatter–plot of the data in a blank worksheet and right–click on one of the data points, select **Add Trendline** and add a polynomial curve of degree 2. Under the **Options** tab, select **Display equation on chart**. The results are shown in Figure 3.23. Note that the polynomial it gives is exactly the same as what we calculated and that its graph goes through all three data points.

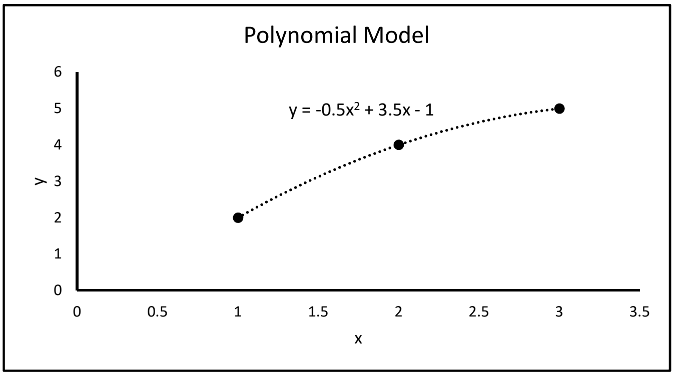


FIGURE 3.23

□

Next we add the point $(4, 2)$ to the data set and fit a second degree polynomial to these four data points. Ideally, we want to find a, b, and c such that

$$\begin{aligned} a\left(1^2\right) + b\,(1) + c &= 2 \\ a\left(2^2\right) + b\,(2) + c &= 4 \\ a\left(3^2\right) + b\,(3) + c &= 5 \\ a\left(4^2\right) + b\,(4) + c &= 2 \end{aligned}$$

This set of equations can be written in matrix form as

$$\begin{bmatrix} 1^2 & 1 & 1 \\ 2^2 & 2 & 1 \\ 3^2 & 3 & 1 \\ 4^2 & 4 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 5 \\ 2 \end{bmatrix} \tag{3.1}$$

which, as before, has the generic form $A\boldsymbol{x} = \boldsymbol{b}$. However, note that A is not square, so it is not invertible. Further analysis reveals that this equation does not even have a solution, so there is no polynomial that fits these four data points perfectly. We will have to settle for a "best–fit" polynomial model. As in Chapter 2 when we fit a linear model to a set of data, we will use a least–squares criterion to find our model. That is, we want a polynomial $p\,(x)$ that minimizes the number

$$S = \sum_{i=1}^{n} \left(y_i - p(x_i)\right)^2.$$

The resulting model is called a *least–squares* polynomial model. To find this model we will find a *least-squares solution* to the matrix Equation (3.1).

Definition 3.3.1 (Least–squares Solution)**.** Let A be an $m \times n$ matrix and $\boldsymbol{b} \in R^m$. A *least-squares solution* of $A\boldsymbol{x} = \boldsymbol{b}$ is a vector $\hat{\boldsymbol{x}} \in R^n$ such that

$$\|\boldsymbol{b} - A\hat{\boldsymbol{x}}\| \leq \|\boldsymbol{b} - A\boldsymbol{x}\| \text{ for all } \boldsymbol{x} \in R^n.$$

The idea behind this definition is that $\hat{\boldsymbol{x}} \in R^n$ gets $A\boldsymbol{x}$ as close to $\boldsymbol{b}$ as possible. The following theorem which we present without proof tells us how to find $\hat{\boldsymbol{x}}$.

Theorem 3.3.2. *Every least–squares solution of $A\boldsymbol{x} = \boldsymbol{b}$ must satisfy the Normal equation*

$$A^T A\boldsymbol{x} = A^T\boldsymbol{b}.$$

If $A^T A$ is invertible, then there is a unique least–squares solution $\boldsymbol{x}$ given by

$$\boldsymbol{x} = \left(A^T A\right)^{-1} A^T\boldsymbol{b} \tag{3.2}$$

□

Note that Theorem 3.3.2 does not say that a general matrix equation $A\boldsymbol{x} = \boldsymbol{b}$ has a unique least-squares solution. In general, there may be many least-squares solutions. However, if $A^T A$ is invertible, then the solution is unique. When fitting curves to data, $A^T A$ is usually invertible so we can use Formula (3.2) to calculate $\hat{\boldsymbol{x}}$. This technique can also be used to fit other types of models to data, as we will see later.

Example 3.3.2 (Calculating a Least–squares Solution)
To calculate $\hat{\boldsymbol{x}}$ using Formula (3.2) in Excel, we will first calculate A^T, then $A^T A$, then $\left(A^T A\right)^{-1}$, then $\left(A^T A\right)^{-1} A^T$, and then finally $\left(A^T A\right)^{-1} A^T\boldsymbol{b}$.

1. Rename a blank worksheet "**Poly 4 points**" and format it as in Figure 3.24.

	A	B	C	D	E	F	G	H
1	**Data**				**A**			**b**
2	**x**	**y**		=A3^2	=A3	1		=B3
3	1	2		=A4^2	=A4	1		=B4
4	2	4		=A5^2	=A5	1		=B5
5	3	5		=A6^2	=A6	1		=B6
6	4	2						

FIGURE 3.24

2. Format the spreadsheet as in Figure 3.25 to hold the various matrix calculations needed.

	A	B	C	D	E	F	G	H
8		$\mathbf{A^T}$					$\mathbf{A^TA}$	
9								
10								
11								
12								
13		$\mathbf{(A^TA)^{-1}}$				$\mathbf{(A^TA)^{-1}A^T}$		
14								
15								
16								
17								
18		$\mathbf{(A^TA)^{-1}A^Tb}$						

FIGURE 3.25

3. Highlight range **A9:D11**, type **=TRANSPOSE(D2:F5)**, and press **Ctrl-Shift-Enter**.

4. Highlight range **F9:H11**, type **=MMULT(A9:D11,D2:F5)**, and press **Ctrl-Shift-Enter**.

5. Highlight range **A14:C16**, type **=MINVERSE(F9:H11)**, and press **Ctrl-Shift-Enter**.

6. Highlight range **E14:H16**, type **=MMULT(A14:C16,A9:D11)**, and press **Ctrl-Shift-Enter**.

7. Highlight range **B19:B21**, type **=MMULT(E14:H16,H2:H5)**, and press **Ctrl-Shift-Enter**. The results are shown in Figure 3.26. This means our polynomial model is $y = -1.25x^2 + 6.35x - 3.25$. Such a model is called a *least–squares* 2^{nd} degree polynomial model.

	B
18	$(A^TA)^{-1}A^Tb$
19	-1.25
20	6.35
21	-3.25

FIGURE 3.26

8. Create a graph of the data points, add a 2^{nd} degree polynomial Trendline, and display the equation on the chart as in Figure 3.27. Note that this model is exactly the same as what we calculated.

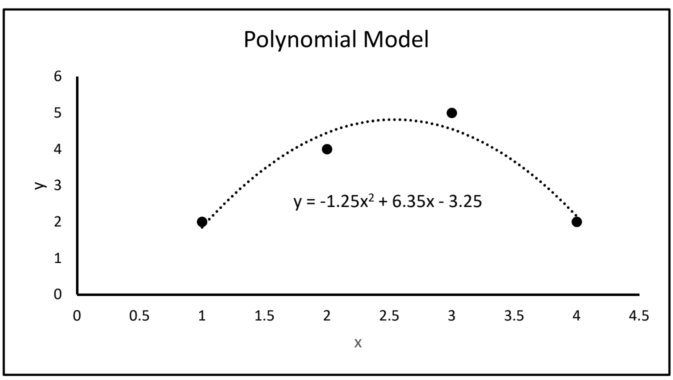


FIGURE 3.27

□

Example 3.3.3 (Calculating R^2 Value)
In Section 2.7 we defined the coefficient of variation R^2 as a measure of how well a line fits a set of data. We then applied this idea to linearizable models by calculating the R^2 value for the straight–line fit to the transformed data. Polynomial models are not linearizable, so to calculate R^2 values for these types of models, we must use the definition. The definition is

$$R^2 = \frac{SS_{Tot} - SS_{Res}}{SS_{Tot}}$$

where $SS_{Tot} = \sum (y_i - \bar{y})^2$, $SS_{Res} = \sum (y_i - \hat{y}_i)^2$, $\bar{y}$ is the mean of all the y-values in the data set, and $\hat{y}_i$ is the *predicted* value of y_i based on the model. Here we use our polynomial model $y = -1.25x^2 + 6.35x - 3.25$ to calculate $\hat{y}_i$.

1. Rename a blank worksheet "**Poly R2**" and format it as in Figure 3.28. The R^2 value is 0.9333 indicating a very good fit.

	A	B	C	D	E
1	**x**	**y**	**Predicted**	**SS$_{Tot}$**	**SS$_{Res}$**
2	1	2	=-1.25*A2^2+6.35*A2-3.25	=(B2-$B$6)^2	=(B2-C2)^2
3	2	4	=-1.25*A3^2+6.35*A3-3.25	=(B3-$B$6)^2	=(B3-C3)^2
4	3	5	=-1.25*A4^2+6.35*A4-3.25	=(B4-$B$6)^2	=(B4-C4)^2
5	4	2	=-1.25*A5^2+6.35*A5-3.25	=(B5-$B$6)^2	=(B5-C5)^2
6	**Mean =**	=AVERAGE(B2:B5)	**Totals =**	=SUM(D2:D5)	=SUM(E2:E5)
7				**R^2 =**	=(D6-E6)/D6

FIGURE 3.28

2. Excel will automatically calculate this R^2 value. On the graph of y vs. x in the worksheet **Poly 4 points**, add the R^2 value to the polynomial trendline by right-clicking on the trendline and selecting **Format Trendline...** $\rightarrow$ **Options** $\rightarrow$ **Display R-squared value on chart**. This value is equal to what we calculated.

□

Example 3.3.4 (Selecting a Best Polynomial Model)
Consider the data in Table 3.3 which gives the area A (in thousands of square miles) and the total length of railroad track R (in thousands of miles) of seven different countries (data from *The World Almanac and Book of Facts*, 2007, World Almanac Books). We want to use a polynomial to model R in terms of A.

TABLE 3.3

	Luxembourg	**Ireland**	**Azerbaijan**	**S Korea**	**Greece**	**Finland**	**Japan**
A	0.998	27.135	33.436	38.023	50.942	130.559	292.26
R	0.170	2.058	1.834	2.157	1.598	3.635	4.092

According to Theorem 3.3.1, there is a unique polynomial of degree at most 6 that fits this data perfectly. We can easily graph R vs. A and fit a sixth degree polynomial trendline. The result is shown in Figure 3.29.

Notice that the R^2 value is 1 (within round error), so the model fits the data perfectly. However, note the large oscillation between $A = 130$ and 300 and that the model predicts negative values of R for A around 100. This is totally unreasonable, so this is a terrible model even though it fits the data perfectly. The large oscillation seen in the graph of this model is typical of a high–degree polynomial model such as this. To find a better model, we can fit polynomials of degree 1 through 5 to the data, and keep track of their R^2 values. The graphs of these models are shown in Figure 3.30.

To choose the best model, we need to examine more than just the R^2 values. We also need to consider how well they will make predictions and their simplicity. The fourth and fifth degree models have the highest R^2 values, but they both have oscillations that seem unreasonable in the context of the problem, so they are not the best options. The first

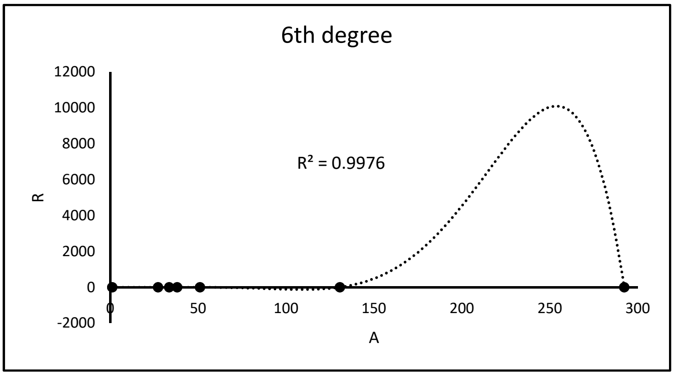


FIGURE 3.29

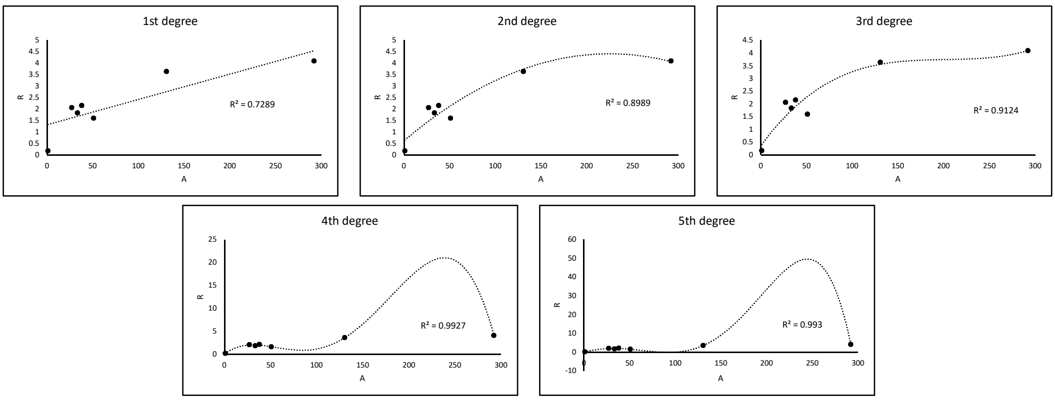


FIGURE 3.30

degree model does not capture the trend of the data very well, so it is not a good option either. The second and third degree models have very similar R^2 values and their graphs look very similar. So we will choose the simpler of the two options, the second degree model, as the best. However, one could make a case that the third degree model is the best. $\square$

This least-squares matrix approach can be applied to models other than polynomials as illustrated in the next example.

Example 3.3.5 (Deer Population)
Table 3.4 gives the number of deer in a hypothetical forest for various years between 1941 and 1982. Our goal is to model the population in terms of the year.

TABLE 3.4

Year	1941	1947	1951	1957	1962	1965	1971	1977	1982
Population	12,500	28,500	7,000	20,000	6,500	12,000	4,000	11,000	3,500

To make the data values smaller and easier to work with, we transform them by subtracting 1941 from the year and dividing the population by 1000, yielding the data in Table 3.5.

TABLE 3.5

x	0	6	10	16	21	24	30	36	41
y	12.5	28.5	7	20	6.5	12	4	11	3.5

Examining the graph of the transformed data in Figure 3.31 we see that the population fluctuates from a high to a low and back to a high in about a 10–year cycle.

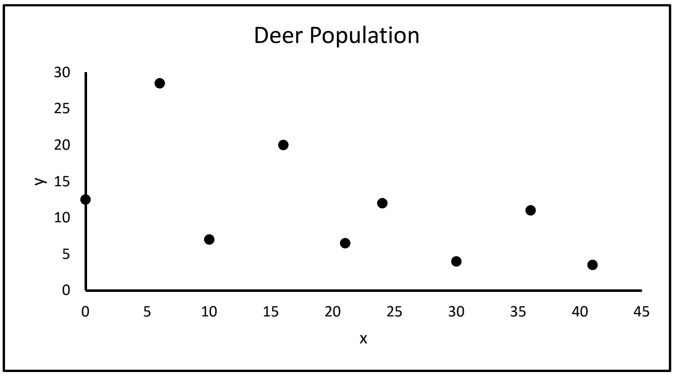


FIGURE 3.31

The periodic nature of the data suggests we use sine or cosine functions to model it. Since the period is approximately 10, our model will have terms of $\cos(\pi x/5)$ and $\sin(\pi x/5)$. Also note that the population shows a downward trend. So we will include an x term in the model with a (likely) negative coefficient. Thus our model is of the form

$$y = a + bx + c\cos\left(\frac{\pi x}{5}\right) + d\sin\left(\frac{\pi x}{5}\right).$$

Ideally we want to satisfy the system of equations

$$\begin{aligned}
a + b(0) + c\cos\left(\frac{0\pi}{5}\right) + d\sin\left(\frac{0\pi}{5}\right) &= 12.5 \\
a + b(6) + c\cos\left(\frac{6\pi}{5}\right) + d\sin\left(\frac{6\pi}{5}\right) &= 28.5 \\
&\vdots \\
a + b(41) + c\cos\left(\frac{41\pi}{5}\right) + d\sin\left(\frac{41\pi}{5}\right) &= 3.5
\end{aligned}$$

which has the matrix form

$$\begin{bmatrix} 1 & 0 & \cos(0\pi/5) & \sin(0\pi/5) \\ 1 & 6 & \cos(6\pi/5) & \sin(6\pi/5) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 41 & \cos(41\pi/5) & \sin(41\pi/5) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 12.5 \\ 28.5 \\ \vdots \\ 3.5 \end{bmatrix}$$

As before, this matrix equation has the generic form $A\boldsymbol{x} = \boldsymbol{b}$. Calculating the least-squares solution, $\hat{\boldsymbol{x}} = \left(A^T A\right)^{-1} A^T \boldsymbol{b}$, to this system as done in Example 3.3.2, gives the approximate solution $\hat{\boldsymbol{x}} = (18.965, -0.314, -5.693, -2.915)$. So the model is

$$y = 18.965 - 0.314 - 5.693 \cos\left(\frac{\pi x}{5}\right) - 2.915 \sin\left(\frac{\pi x}{5}\right).$$

Figure 3.32 shows a graph of the model on top of the data. The model appears to fit the data relatively well, but many refinements could be made. In the next section you will be asked to make one such refinement in the exercises.

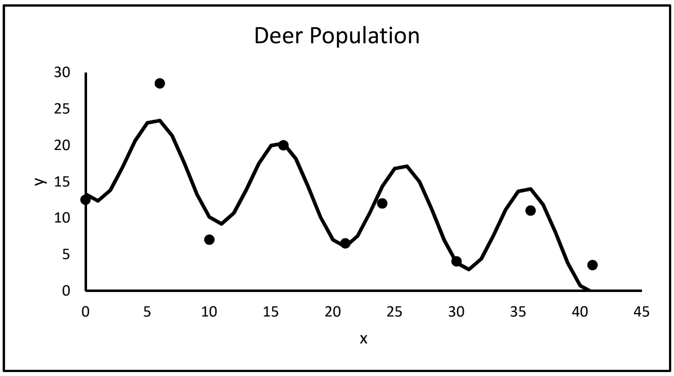


FIGURE 3.32

□

General Guidelines for Selecting a Best Model

In general when constructing empirical models, one has many different types of models from which to choose. Deciding which one is best is not easy and is often very subjective, but here are a few simple guidelines:

1. Consider the R^2 value, but don't rely solely on it.
2. Look for a pattern in the residuals. If there is a pattern, the model should be refined.
3. Consider how good the model is for making predictions between data values. If it oscillates or would give unreasonable values, look for a better model.
4. Consider "end" behavior. If the data appears to be "leveling off" at the end, but the model is increasing (or vice-versa), consider a different model.
5. Consider the simplicity of a model. In general, the fewer the terms, the better.

We must also stress that when using an empirical model such as those discussed in this section and in Chapter 2 to make predictions, the predictions are always point–*estimates* of the true values. These predictions should never be presented as precise certainties. Books on statistics and regression discuss how to use a point–estimate to form a confidence interval for the true value (a range of possible values), but this topic is beyond the scope of this book.

Exercises

3.3.1 Use the least–squares matrix approach to fit a linear model $y = a + mx$ to a randomly generated set of data with 4 points. Show by examples that this gives the same results as the formulas given in Section 2.4.

3.3.2 Fit different polynomial models to the data in the table below and select the best one. Explain how you determined which one is best.

$\boldsymbol{x}$	1.4	2.4	7.1	13.8	34.2	109.3	134
$\boldsymbol{y}$	2.7	2.27	3.31	3.39	3.81	4.88	4.62

3.3.3 Calculate the R^2 value for the model fit to deer population data in Example 3.3.5.

3.3.4 Consider the problem of fitting a 2^{nd} polynomial to a set of 4 data points as discussed in Example 3.3.2. However, suppose we require that the y-intercept of the model is some specified valued c_0. In other words, we want to fit a model of the form $y = ax^2 + bx + c_0$ where c_0 is a given number. Design a spreadsheet to fit such a model where the user can input 4 data points and specify the value of c_0.

3.3.5 One useful property of polynomials is that they are easy to differentiate and integrate. Suppose a researcher observes a particle moving in a straight line and measures the particle's velocity relative to its starting position at several points in time as shown in the table below.

$\boldsymbol{t}$	0	1.2	2.5	3.2	4.6	5.4	6.3	7.3
$\boldsymbol{v(t)}$	0	3.6	5.5	7.4	6.7	5.8	3.5	0

a. Fit several polynomial models to the data and choose the one that best models the velocity. (**Suggestion**: When choosing the best model, don't rely strictly on the R^2 value. Put a large emphasis on simplicity.)

b. The acceleration of the particle at time t is $a(t) = v'(t)$. Use your model in part a. to estimate $a(1.8)$.

c. The total distance traveled over the time interval $[a,\, b]$ is

$$\text{Total distance travled} = \int_a^b |v(t)|\; dt.$$

Use your model in part a. to estimate the total distance the particle traveled over the interval $[0,\, 7.3]$.

3.3.6 The table below contains the temperatures over one day in Seward, NE starting at midnight. The goal of this exercise is to predict the temperature y at time x. Notice that the temperature is periodic, so we will use a model of the form $y = a\sin(bx + c) + d$. Follow these steps to design a spreadsheet to implement a simple "sine regression" algorithm for fitting a model of this type to the data:

Hour	0	2	4	6	8	10	12	14	16	18	20	22
Temp	44.5	45.3	52.6	60.4	70.2	75.9	79.8	79.1	72.8	63.5	52.5	44.6

a. Enter the data in a spreadsheet, create a graph of the data, and designate cells to hold the values of a, b, c, and d.

b. The period of the function $y = a\sin(bx + c) + d$ is $2\pi/b$. Estimate the period of the data and use this to estimate the value of b.

c. Initially, let $c = 0$.

d. To find the values of a and d, create a graph of y vs. $\sin(bx + c)$ and fit a linear trendline to this transformed data. Display the R^2 value. The slope is the value of a and the y-intercept is the value of d. Use the functions **SLOPE** and **INTERCEPT** to calculate the values of a and d, respectively.

e. Create a scroll bar to vary the value of b between -1 and +1 and another scroll bar to vary the value of c between -2 and +2 (see Appendix A.4 for more information on scroll bars).

f. Use the scroll bars to find values of b and c that maximize the R^2 value.

g. Graph the model on top of the original data. How well does the model fit the data?

3.4 Multiple Regression

In previous sections we have discussed predicting the value of one response variable y with one predictor variable x. In this section we will discuss using two or more predictor variables $x_1, x_2, \ldots, x_n$. This topic is called *multiple regression*.

Consider the problem of predicting the selling price of a house. The selling price is affected by many factors including the age of the house, living area, number of bedrooms, etc. Table 3.6 lists the selling price, living area (in ft^2), acres of land, and the number of bedrooms of 10 homes in a neighborhood.

TABLE 3.6

Selling Price	Area	Acres	Bedrooms
100,000	2,205	2.5	3
93,500	2,155	0.8	3
95,650	2,600	1.1	4
75,025	1,900	0.35	3
95,000	1,200	2.5	2
80,250	2,050	1.8	3
85,250	2,250	0.9	4
121,250	2,490	1.8	3
94,575	2,390	1.6	2
109,000	3,100	1.0	4

Example 3.4.1 (Single Predictor Variable)
Consider the graphs of Selling Price vs. Area and Selling Price vs. Acres as shown in Figure 3.33 along with the linear regression equation for each.

Notice that the R^2 value for the predictor variable Area is higher than the R^2 value for Acres. This means that the regression equation for Selling Price in terms of Area will give a

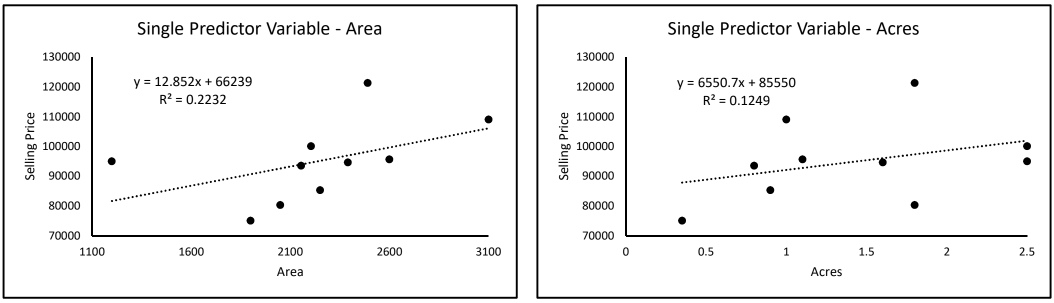


FIGURE 3.33

better predicted Selling Price than the equation in terms of Acres. We say that out of these two variables, Taxes is the best single-variable predictor of selling price.

Suppose a house has an area of 2,800 ft^2. We could use the regression equation for Selling Price in terms of Area to predict that the house would sell for 12.852(2,800) + 66,239 = \$102,224. This, of course, is only a point–estimate of the price, and probably not a very good estimate because the R^2 value for the regression equation is only 0.2232.

□

Example 3.4.2 (Multiple Predictor Variables)
Considering only one predictor variable is a bit too simple. Many variables affect the selling price, so we should consider more than one predictor variable in our regression equation. Suppose we consider both Area and Acres.

If we let y = Selling Price, x_1 = Area, and x_2 = Acres, we want to fit a model of the form

$$y = a_0 + a_1x_1 + a_2x_2$$

to the data where a_0, a_1, and a_2 are constants. As in Section 3.3, ideally we want to satisfy the system of equations

$$\begin{aligned} a_0 + a_1(2205) + 2.5a_2 &= 100000 \\ &\vdots \\ a_0 + a_1(3100) + 1a_2 &= 109000 \end{aligned}$$

We could take a matrix approach to find a least–squares solution to this system as we did in Section 3.3, but Excel will do this automatically for us.

1. Rename a blank worksheet **Homes** and format it as in Figure 3.34. Enter the rest of the data from Table 3.6 in columns **A**–**D**.

	A	B	C	D
1	**Selling Price**	**Area**	**Acres**	**Bedrooms**
2	100000	2205	2.5	3

FIGURE 3.34

2. Select **Tools** → **Data Analysis**. . . → **Regression** and press **OK** (if **Data Analysis**. . . is not available, select **Tools** → **Add-Ins**. . . → **Analysis ToolPak**, press **OK** and try selecting **Data Analysis**. . . again).

1. Next to **Input Y Range:** select **$A$1:$A$11**.
2. Next to **Input X Range:** select **$B$1:$C$11**.
3. Check the box next to **Labels**.
4. Next to **Output Range:** select **$A$13**
5. Press **OK**

Excel generates many outputs. For the purposes of this example, we focus on two subsets (see Example 3.4.3 for a description of some of the other outputs). The first subset of outputs shown in Figure 3.35 gives the coefficients in our model (i.e. the values of a_0, a_1, and a_2). We see that our model is

$$y = 36669.6 + 18.9x_1 + 11239.2x_2$$

Such a model is called a *multiple-regression equation*.

	A	B
28		*Coefficients*
29	Intercept	36669.57782
30	Area	18.86847037
31	Acres	11239.20513

FIGURE 3.35

The second subset of outputs is shown in Figure 3.36. Here we see the R^2 value which is calculated by Formula (2.16). The other important output is the **Adjusted** R^2 value which is defined by

$$\text{Adjusted } R^2 = 1 - \left[\frac{n-1}{n-(k+1)}\right](1 - R^2) \tag{3.3}$$

where n is the number of data points and k is the number of predictor variables ($n = 10$ and $k = 2$ in this case).

	A	B
15	*Regression Statistics*	
16	Multiple R	0.736297484
17	R Square	0.542133985
18	Adjusted R Square	0.411315124
19	Standard Error	10308.18525
20	Observations	10

FIGURE 3.36

The adjusted R^2 value takes into account the number of data points (the more data points, the higher the adjusted R^2 value) and the number of predictor variables (the more predictor variables, the lower the adjusted R^2 value). We want as simple a model as possible, so the fewer the variables, the better. We will compare different sets of predictor variables using adjusted R^2 values.

3. Now let's consider the combination of predictor variables Acres and Bedrooms. Repeat step 2, except select **$C$1:$D$11** for the **Input X Range:**. The R^2 and adjusted R^2 values for this combination are 0.209 and -0.017, respectively, which are lower than for the previous combination indicating that this combination does not give a better model for predicting the selling price.

The results for all the different combinations of predictor variables are shown in Table 3.7. Note that including the variable Bedrooms results in low R^2 values. This indicates that the number of bedrooms is not a good predictor of the selling price. Also note that the R^2 value for the set of all three predictor variables is the highest, but the adjusted R^2 value is lower than that for Area and Acres. This indicates that the set of three variables gives better predictions (a higher R^2 value), but the additional variable makes it more complicated, so it is less desirable as a model (a lower adjusted R^2 value).

TABLE 3.7

Predictor Variables	R^2	Adjusted R^2
Area, Acres	0.542	0.411
Area, Bedrooms	0.327	0.135
Acres, Bedrooms	0.209	-0.017
Area, Acres, Bedrooms	0.552	0.328

If we simply compare the adjusted R^2 values, we conclude that the best combination of predictor variables is Area and Acres. To refine our model we might want to collect data on other variables that might affect the selling price such as age, total number of rooms, etc. With more variables, there are many different combinations of predictor variables we could consider. The process of determining which set of variables is best is very complicated. We have presented a very simple strategy here.

□

Example 3.4.3 (Other Statistical Outputs)
Consider again the problem of predicting Selling Price in terms of Area and Acres from Example 3.4.2. In that example we focused on only two sets of the many statistics outputs. In this example we briefly discuss some of the other outputs.

Figure 3.37 shows the set of outputs from an Analysis of Variance (ANOVA) hypothesis test. For our purposes, the most important output here is the "Significance F." Informally, Significance F is a measure of how well the overall model fits the data. A smaller value means a better fitting model. More formally, Significance F is the P-value for the hypothesis test with the null and alternative hypotheses

$$H_0: \text{ All the coefficients are 0,} \qquad H_1: \text{ Not all the coefficients are 0.}$$

The value of F shown in the ANOVA table is the test statistic for this hypothesis test. The other values are part of the calculations. We can use Significance F as part of our comparison of different combinations of predictor variables. A better combination of variables has a lower Significance F.

	A	B	C	D	E	F
22	ANOVA					
23		*df*	*SS*	*MS*	*F*	*Significance F*
24	Regression	2	880705468.3	440352734.2	4.14416	0.064950789
25	Residual	7	743810781.7	106258683.1		
26	Total	9	1624516250			

FIGURE 3.37

The second set of outputs is shown in Figure 3.38. Note that this is an expanded version of the outputs shown in Figure 3.35.

	A	B	C	D	E	F	G
28		*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
29	Intercept	36669.57782	20772.38015	1.7653	0.120862	-12449.296	85788.45167
30	Area	18.86847037	7.471230069	2.52548	0.039493	1.201818568	36.53512218
31	Acres	11239.20513	5090.208538	2.208	0.062982	-797.225418	23275.63568

FIGURE 3.38

Here are some brief explanations of these outputs:

- Formally, the P-value is the P-value for the hypothesis test

$$H_0: \text{ The coefficient is 0,} \qquad H_1: \text{ The coefficient is not 0.}$$

  Informally, the P-value is a measure of how "significant" the variable is in the presence of the other variables. A smaller P-value indicates the variable is more significant. These results indicate that Area is more significant that Acres. See Exercise 3.4.11 for an example of how to use these P-values to help determine the best combination of predictor variables.

- The t Stat is the test statistic for the hypothesis test.
- The Lower 95% and Upper 95% give a 95% confidence interval estimate of the true value of the coefficient.
- The Standard Error is a number used in the calculation of the confidence interval.

For a more complete description of all the outputs see, for instance, Sincich, Terry; Levine, David M.; and Stephan, David, *Practical Statistics by Example Using Microsoft Excel and Minitab*, Second ed., Prentice Hall, 2002, pp. 602.

□

Example 3.4.4 (Polynomial Models)
Let's return to the problem of fitting a 2^{nd} degree polynomial model of the form $y = a+bx+cx^2$ to the set of 4 data points $\{(1,2), (2,4), (3,5), (4,2)\}$ as considered in Example 3.3.2. We could think of this as predicting the values of y with the "predictor" variables x and x^2, so it can be treated as a multiple regression problem.

1. Rename a blank worksheet **Polynomial** and format it as in Figure 3.39. Enter the rest of the data in columns **A** and **B** and copy the formula in **C3** down to row 6.

	A	B	C
1	**Data**		
2	**y**	**x**	**x²**
3	2	1	=B3^2

FIGURE 3.39

2. Repeat step 2 from Example 3.4.2. Select **\$A\$2:\$A\$6** as the **Input Y Range:** and **\$B\$2:\$C\$6** as the **Input X Range:**. Check the box next to **Labels** and select **\$A\$8** as the **Output Range:**. The coefficients are shown in Figure 3.40.

   These results give us the model $y = -3.25 + 6.35x - 1.25x^2$, which is exactly the same as in Example 3.3.2. Also note that the R^2 value is 0.9333, exactly the same as that calculated in Example 3.3.3.

	A	B
23		*Coefficients*
24	Intercept	-3.25
25	x	6.35
26	x2	-1.25

FIGURE 3.40

□

Example 3.4.5 (Deer Population Again)
We can use a multiple regression approach to fit a model of the form $y = a + bx + c\cos(\pi x/5) + d\sin(\pi x/5)$ to the transformed deer population data in Table 3.5. In this case we have 3 predictor variables: x, $\cos(\pi x/5)$, and $\sin(\pi x/5)$.

1. Rename a blank worksheet **Deer Population** and format it as in Figure 3.41. Enter the rest of the data from Table 3.5 in columns **A** and **B**, and copy the range **C2:D2** down to row 10.

	A	B	C	D
1	**y**	**x**	**cos**	**sin**
2	12.5	0	=COS(PI()*B2/5)	=SIN(PI()*B2/5)

FIGURE 3.41

2. Repeat step 2 from Example 3.4.2. Select **$A$1:$A$10** as the **Input Y Range:** and **$B$1:$D$10** as the **Input X Range:**. Check the box next to **Labels** and select **$A$12** as the **Output Range:**. The coefficients are shown in Figure 3.42.

	A	B
27		*Coefficients*
28	Intercept	18.96480399
29	x	-0.314217087
30	cos	-5.692825862
31	sin	-2.914769196

FIGURE 3.42

These coefficients give us the approximate model

$$y = 18.965 - 0.314 - 5.693\cos\left(\frac{\pi x}{5}\right) - 2.915\sin\left(\frac{\pi x}{5}\right),$$

which is exactly the same as in Example 3.3.5. Also note that the R^2 value is 0.8771 which should be (approximately) the same as that calculated in Exercise 3.3.3.

□

Exercises

3.4.1 In an attempt to predict the final grade of students in an Introduction to Statistics class, the professor gives each student a 20-point pretest at the beginning of the year. The table below gives the final grade, pretest score, ACT score, and year (1 = freshman, 2 = sophomore, etc.) of 10 students.

Grade	84.5	82.3	69.2	65.1	80.1	85.9	88.1	90.7	87.2	92.7
Pretest	9	8	18	10	6	8	16	11	15	19
Year	1	2	2	4	3	3	1	4	4	3
ACT	25	20	18	17	20	22	30	28	27	31

a. Find the regression equation that predicts Grade in terms of Pretest Score. Repeat using Year and then ACT. Which of these single-variable predictors is best at predicting the final grade based on the R^2 values? Does the pretest score alone appear to be a good predictor of the final grade? Explain.

b. Consider all four different combinations of two or three predictor variables. Determine which combination is best at predicting the final grade using the methods described in this section. Based on your results, does it seem worthwhile to give the pretest as a way of predicting the final grade? Does the year of the student appear to affect the final grade? Explain.

c. Use the multiple regression equation that predicts Grade in terms of Pretest and ACT to predict the grade of a student who has a pretest score of 18 and an ACT score of 28.

3.4.2 Use a multiple regression approach to fit a 7^{th} degree polynomial to the 8 data points shown in the table below. Create a graph of the resulting polynomial on top of the data points.

x	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85
y	0.90	9.00	2.10	7.80	13.50	9.30	0.40	6.20

3.4.3 Use the definition of R^2 in Formula (2.16) to calculate the R^2 value for the multiple-regression equation for predicting selling price in terms of area and acres as found in Example 3.4.2. Compare this value to the R Square value given in Figure 3.36. Also verify that the Adjusted R Square value in the figure is calculated according to Formula (3.3).

3.4.4 The sin and cos terms in the deer population model in Example 3.3.5 were included to capture the oscillating pattern of the data. One might wonder if both terms are really necessary. Fit a model of the form $y = a + bx + c\cos(\pi x/5)$ and then fit a model of the form $y = a + bx + d\sin(\pi x/5)$ to the data. Compare the R^2 value for each model to the original model. Does the inclusion of both sin and cos terms yield a significantly better model? Explain why or why not.

3.4.5 The amount of a radioactive substance remaining after time t, $y(t)$, is described by the exponential model $y(t) = Ce^{-kt}$ where C is the initial amount (the amount at time $t = 0$) and k is a constant. Suppose two radioactive substances A and B have constants $k_A = 0.03$ and $k_B = 0.05$. A mixture of these two substances contains C_A grams of A and

C_B grams of B at time $t = 0$, both of which are unknown. The total amount of the mixture at time t is modeled by

$$y(t) = C_A e^{-0.03t} + C_B e^{-0.05t} \tag{3.4}$$

A researcher measures the total amount of the mixture at several times and records the data in the table below. Estimate the values of C_A and C_B by fitting a model of the form (3.4) to the data. (**Hint**: Use a multiple regression approach to fit the model, but note that there is no intercept term in the model. Select the appropriate option in Excel.)

Time	5	6	7	8
Amount	8.8	8.6	8.2	7.9

3.4.6 Consider a refinement to the deer population model in Example 3.3.5. Note that as time increases, the difference between a high point and the next low point (the amplitude) tends to decrease. Our original model did not take this into account.

a. For the x-values 6, 16, 24, and 36, calculate the amplitude by subtracting the next y-value.

b. Create a graph of amplitude vs. x using the data in part a.

c. Fit an exponential model, $g(x) = me^{kx}$, to the data in part a.

d. Fit a model of the form

$$y = a + bx + c\,g(x) * \cos\left(\frac{\pi x}{5}\right) + d\,g(x) * \sin\left(\frac{\pi x}{5}\right)$$

to the original data. Create a graph of this model. Does this model seem to fit the data any better than the original one?

e. Compare the R^2 value for this refined model to the original model found in Example 3.3.5. How does this refined model compare to the original model?

3.4.7 Use a multiple regression approach to fit a model of the form

$$y = a + bx + c\ln(3x+1) + \frac{d}{1-x^2} + e\cos(3x)$$

to the data in the table below. Graph the resulting model on top of the data. Comment on how well the model fits the data.

x	2	3	4	5	6	7	8	9	10
y	-2.3	-4.8	11.2	4.1	16.2	11.1	20.3	17.9	22.9

3.4.8 The table below gives the poverty level, unemployment rate, high school graduation rate, and divorce rate (all in percentages) of 10 randomly selected states in 2007 (data collected by Matthew Schranz, 2011). Determine which combination of predictor variable(s) is best at predicting poverty level. Based on this data, does divorce rate appear to be related to poverty level at all?

Poverty Level	Unemployment Rate	High School Graduation Rate	Divorce Rate
14.3	4.3	69.6	3.9
9.3	3.8	71.9	3.7
8.9	3.8	86.5	2.6
12.8	5.3	81.9	3.8
5.8	3.5	81.7	3.8
15.5	5	68.6	4
12.8	5.3	73.8	3.9
9.4	2.7	82.5	3.1
10	4	88.6	3.6
11	4.5	88.5	2.9

3.4.9 The table below gives the average points scored per game (PPG), total number of turnovers, average minutes per game played, and free-throw percentage of 11 college basketball players over the course of a season (data collected by Alex Hopping, 2012). Determine which combination of predictor variable(s) is best at predicting PPG.

PPG	Turnovers	Min/Game	Free-Throw %
8.6	75	19.8	0.797
4.5	60	18.7	0.532
7.1	22	22.7	0.780
15.8	48	29.1	0.767
1.6	16	9	0.500
5.2	20	12.7	0.750
8	56	21.6	0.726
2.8	13	8.6	0.706
4	19	15.8	0.583
14.5	62	25.3	0.848
5	30	16	0.577

3.4.10 The process we presented in this section for selecting the best combination of predictor variables in a multiple regression model is to try all possible combinations of predictor variables and then pick the combination with the highest adjusted R^2 value. In general, the number of possible combinations is $2^n - 1$ if there are n predictor variables. This number grows exponentially large as n increases, making this process impractical if there are a large number of predictor variables. A more intelligent method for selecting the best combination is called *forward stepwise regression*. Here is a simplified algorithm:

1. Try all single predictor variables. Choose the one with the highest adjusted R^2 value as the best.
2. Try all combinations of two predictor variables that include the best variable from step 1. Choose the pair with the highest adjusted R^2 value as the best.
3. Try all combinations of three predictor variables that include the best pair from step 2. Choose the triplet with the highest adjusted R^2 value as the best.
4. Repeat until there are no more combinations to try, or adding another variable does not increase the adjusted R^2 value.

The data below give the average teacher salaries in 12 states along with six education-related statistics (data collected by Quinn Wragge, 2019). Use the forward stepwise regression algorithm to determine which set of statistics is best at predicting teacher salaries.

State	Avg. Salary	% of US pop.	Grad. Rate	% to College	ACT Score	Cost of Living Ind.	Diversity Score
AZ	$47,403	2.17	77.4	52.3	19.2	97.7	67.81
CA	$78,711	12.13	82	60.9	22.7	138.7	70.89
CO	$46,506	1.73	77.3	58.2	23.9	105.5	66.25
HI	$57,674	0.43	81.6	60.8	18.9	190.1	69.69
IL	$61,602	3.89	85.6	61.5	23.9	95.7	67.93
ME	$51,077	0.41	87.5	54.6	24	117.2	58.4
MI	$62,200	3.05	79.8	64	24.4	89.3	62.84
NE	$52,388	0.59	88.9	62.5	20.1	93.3	64.24
NC	$49,837	3.17	85.6	62.9	19.1	94	65.82
SD	$42,668	0.27	83.9	67.2	21.9	98.5	63.13
TX	$52,575	8.75	89	58.7	20.6	91.3	70
WY	$58,650	0.17	79.3	53.8	20	90.5	62.49

3.4.11 Another process for more intelligently selecting the best combination of predictor variables is called *backward stepwise regression*. Here is a simplified algorithm:

1. Fit a model using all predictor variables.
2. Throw out the variable with the largest P-value.
3. Fit a model using the remaining predictor variables.
4. Repeat until all the remaining predictor variables have P-values less than 0.05. The remaining variables give the best combination.

The data below give the win-loss percentage (W-L%), average runs scored (R), average runs allowed (RA), batting average (Avg), save percentage (Sv %), earned run average (ERA), and fielding percentage (Field %) of 11 MLB teams over a season (data collected by Sam Otte, 2019). Use the backward stepwise regression process to find the best combination of predictor variables for predicting win-loss percentage.

W-L%	R	RA	Avg	Sv %	ERA	Field %
0.506	4.3	4	0.235	0.59	3.72	0.988
0.290	3.8	5.5	0.239	0.61	5.18	0.982
0.414	4.3	5.1	0.254	0.67	4.63	0.984
0.558	4.8	4.6	0.256	0.65	4.33	0.988
0.494	4.5	4.5	0.242	0.57	4.15	0.987
0.391	3.7	5	0.237	0.56	4.76	0.986
0.617	5.3	4.1	0.249	0.73	3.78	0.984
0.407	3.8	4.7	0.235	0.68	4.4	0.983
0.556	4.4	4	0.258	0.7	3.74	0.986
0.451	4.4	5.1	0.244	0.65	4.85	0.983

3.5 Spline Models

Consider the data in Table 3.8 which gives the liters of milk given by a dairy cow on each of several different days after she begins producing. A graph of this data is shown in Figure 3.43. Our goal is to model Liters in terms of Day so that we can predict how much milk was given on the days not listed in the table.

TABLE 3.8

Day	5	13	25	42	50	62	75	90	100	120
Liters	10	19	30	25	22	35	50	55	40	35

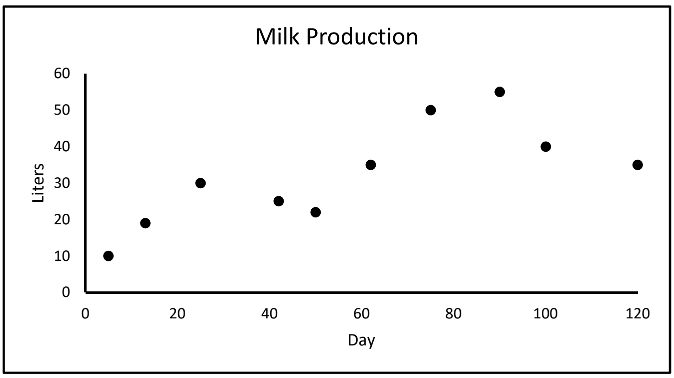


FIGURE 3.43

The graph of the data certainly does not resemble an exponential, logarithmic, or power curve, so a linearizable model does not seem appropriate. Low-degree polynomials do not capture the trend of the data, and higher-degree polynomials produce oscillations which do not seem appropriate. We will instead consider *spline* models where we simply "connect the dots" with either straight lines, forming a *linear spline* model, or with cubic polynomials, forming a *cubic spline* model.

Example 3.5.1 (Linear Spline Model)
The graph of a linear spline model is easy to form.

1. Rename a blank worksheet "**Linear**" and format it as in Figure 3.44. Enter the rest of the data from Table 3.8 in columns **A** and **B**.

	A	B
1	**Day**	**Liters**
2	5	10

FIGURE 3.44

2. Create a graph similar to Figure 3.45. The straight lines form the graph of the linear spline model.

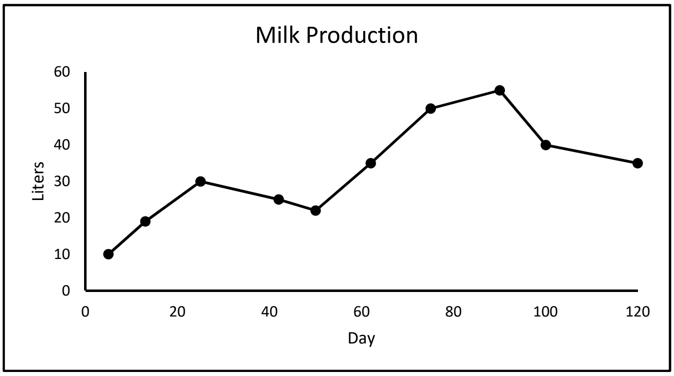


FIGURE 3.45

3. To use this model to make predictions, we need to know the slope and y-intercept of each line segment. We know two points on each line segment, call them (x_n, y_n) and (x_{n+1}, y_{n+1}) where $x_n < x_{n+1}$, so we can easily find the slope using the formula

$$\text{slope} = m = \frac{y_{n+1} - y_n}{x_{n+1} - x_n}$$

   Once we know the slope m we can find the y-intercept b using the slope-intercept form of a straight line

$$y = mx + b \quad \Rightarrow \quad b = y - mx$$

   To implement these formulas, format the spreadsheet as in Figure 3.46. Copy row 2 down to row 10.

	C	D
1	**Slope**	**Intercept**
2	=(B3-B2)/(A3-A2)	=B2-C2*A2

FIGURE 3.46

   Looking at the column **Slope**, we see that the amount of milk produced is increasing most rapidly, on average, between days 62 and 75 and decreasing most rapidly between days 90 and 100.

4. Now that we know the slopes and y-intercepts of each piece of the spline we can easily calculate a predicted value of Liters given a value of Day by

$$\text{Liters} = m(\text{Days}) + b$$

   where m and b are the slope and y-intercept, respectively, of the appropriate piece of the spline. To do this, add the formulas in Figure 3.47. (**Note**: The function **VLOOKUP** in Figure 3.47 will look down the left column of the range **A2:D11** and find the largest value less than or equal to the value in cell **F2**. It will then return the value in the

3^{rd} or 4^{th} column of that range, the slope or the y-intercept, respectively. For this to work properly it is necessary that the x-values are in ascending order.)

	F	G
1	**Day**	**Liters**
2	12	=VLOOKUP(F2,A2:D10,3)*F2+VLOOKUP(F2,A2:D10,4)

FIGURE 3.47

□

Example 3.5.2 (Cubic Spline Model)
Linear spline models are easy to calculate, but they do not form smooth curves. In fact, the curve is not differentiable at any of the data points. This type of model predicts sharp changes at each data point, which does not seem reasonable. To solve this problem, we will connect the dots with cubic polynomials instead of straight lines and put conditions on the derivatives of each segment.

To illustrate this process, we will consider a set of 3 data points

$$\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$$

where $x_1 < x_2 < x_3$. Each x-value is called a *knot*. We will connect them using 2 cubic polynomials:

$$\begin{array}{ll} p_1(x) = a_1 + b_1 x + c_1 x^2 + d_1 x^3 & \text{for } x_1 \leq x < x_2 \\ p_2(x) = a_2 + b_2 x + c_2 x^2 + d_2 x^3 & \text{for } x_2 \leq x \leq x_3 \end{array}$$

We now need to find the values of the 8 parameters a_1, b_1, c_1, d_1, a_2, b_2, c_2, and d_2. For the model to form a smooth curve that goes through each data point, we need to satisfy the following three conditions:

1. Each polynomial must pass through the two data points at the ends of the interval over which it is defined.
2. The first derivatives of two polynomials that meet must be equal at the point at which they meet.
3. The second derivatives of two polynomials that meet must be equal at the point at which they meet.

The first condition gives us the following four equations:

$$p_1(x_1) = a_1 + b_1 x_1 + c_1 x_1^2 + d_1 x_1^3 = y_1 \tag{3.5}$$

$$p_1(x_2) = a_1 + b_1 x_2 + c_1 x_2^2 + d_1 x_2^3 = y_2 \tag{3.6}$$

$$p_2(x_2) = a_2 + b_2 x_2 + c_2 x_2^2 + d_2 x_2^3 = y_2 \tag{3.7}$$

$$p_2(x_3) = a_2 + b_2 x_3 + c_2 x_3^2 + d_2 x_3^3 = y_3 \tag{3.8}$$

The first derivatives of the polynomials are:

$$p_1'(x) = b_1 + 2c_1 x + 3d_1 x^2, \quad p_2'(x) = b_2 + 2c_2 x + 3d_2 x^2$$

The second condition gives us the equation

$$p_1'(x_2) = b_1 + 2c_1 x_2 + 3d_1 x_2^2 = b_2 + 2c_2 x_2 + 3d_2 x_2^2 = p_2'(x_2) \tag{3.9}$$

The second derivatives of the polynomials are:

$$p_1''(x) = 2c_1 + 6d_1x, \quad p_2''(x) = 2c_2 + 6d_2x$$

The third condition gives us the equation

$$p_1''(x_2) = 2c_1 + 6d_1x_2 = 2c_2 + 6d_2x_2 = p_2''(x_2) \tag{3.10}$$

Equations (3.5) – (3.10) give us six equations for the eight unknowns. To uniquely determine the values of these unknowns, we need two more equations. To this end, we will specify the values of the second derivatives at the end points of the data sat (i.e. at x_1 and x_3).

If we want these second derivatives to be some known values, say m_1 and m_2, we would add the equations

$$\begin{aligned} p_1''(x_1) &= 2c_1 + 6d_1x_1 = m_1 \\ p_2''(x_3) &= 2c_2 + 6d_2x_3 = m_2 \end{aligned}$$

The resulting model is called a *clamped* spline. If, however, we do not know the values of the second derivatives, we simply set them equal to 0 and have the equations

$$p_1''(x_1) = 2c_1 + 6d_1x_1 = 0 \tag{3.11}$$
$$p_2''(x_3) = 2c_2 + 6d_2x_3 = 0 \tag{3.12}$$

This is the approach we will take. The resulting model is called a *natural* spline. Rewriting Equations (3.5) – (3.12) so that the constants are on the right–hand side, we get the system

$$\begin{array}{rcl}
a_1 + x_1b_1 + x_1^2c_1 + x_1^3d_1 & & = y_1 \\
a_1 + x_2b_1 + x_2^2c_1 + x_2^3d_1 & & = y_2 \\
 & a_2 + x_2b_2 + x_2^2c_2 + x_2^3d_2 & = y_2 \\
 & a_2 + x_3b_2 + x_3^2c_2 + x_3^3d_2 & = y_3 \\
b_1 + 2x_2c_1 + 3x_2^2d_1 & - b_2 - 2x_2c_2 - 3x_2^2d_2 & = 0 \\
2c_1 + 6x_2d_1 & - 2c_2 - 6x_2d_2 & = 0 \\
2c_1 + 6x_1d_1 & & = 0 \\
 & 2c_2 + 6x_3d_2 & = 0
\end{array}$$

Rewriting this system in matrix form yields

$$\begin{bmatrix} 1 & x_1 & x_1^2 & x_1^3 & 0 & 0 & 0 & 0 \\ 1 & x_2 & x_2^2 & x_2^3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & x_2 & x_2^2 & x_2^3 \\ 0 & 0 & 0 & 0 & 1 & x_3 & x_3^2 & x_3^3 \\ 0 & 1 & 2x_2 & 3x_2^2 & 0 & -1 & -2x_2 & -3x_2^2 \\ 0 & 0 & 2 & 6x_2 & 0 & 0 & -2 & -6x_2 \\ 0 & 0 & 2 & 6x_1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 6x_3 \end{bmatrix} \begin{bmatrix} a_1 \\ b_1 \\ c_1 \\ d_1 \\ a_2 \\ b_2 \\ c_2 \\ d_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_2 \\ y_3 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.13}$$

which has the general form $A\boldsymbol{x} = \boldsymbol{b}$ where $\boldsymbol{x} = (a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2)$ is the vector of unknowns. We can solve this system using inverses, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. (Note that matrix A has a distinct structure and a lot of 0's. These facts can be exploited to find more computationally efficient ways to solve this system.)

To fit a cubic spline to the three data points $\{(75, 50), (95, 55), (100, 40)\}$ from Table 3.8, follow these steps:

1. Rename a blank worksheet "**Cubic**" and format it as in Figure 3.48.

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	**x**	**y**		**A**									**b**
2	75	50		1	=A2	=A2^2	=A2^3	0	0	0	0		=B2
3	90	55		1	=A3	=A3^2	=A3^3	0	0	0	0		=B3
4	100	40		0	0	0	0	1	=A3	=A3^2	=A3^3		=B3
5				0	0	0	0	1	=A4	=A4^2	=A4^3		=B4
6				0	1	=2*A3	=3*A3^2	0	=-E6	=-F6	=-G6		0
7				0	0	2	=6*A3	0	0	=-F7	=-G7		0
8				0	0	2	=6*A2	0	0	0	0		0
9				0	0	0	0	0	0	2	=6*A4		0
10				**A^{-1}**									**$A^{-1}b$**

FIGURE 3.48

2. To calculate A^{-1}, highlight the range **D11:K18**. Type **=MINVERSE(D2:K9)**, and press the combination of the keys **Ctrl-Shift-Enter**. To calculate $A^{-1}\boldsymbol{b}$, highlight the range **M11:M18**. Type **=MMULT(D11:K18,M2:M9)**, and press the combination of the keys **Ctrl-Shift-Enter**. The results show that our model is:

$$\begin{aligned} p_1(x) &= 1015 - 40.37x + 0.55x^2 - 0.0024x^3 && \text{for } 75 \leq x < 90 \\ p_2(x) &= -3440 + 108.13x - 1.1x^2 + 0.0037x^3 && \text{for } 90 \leq x \leq 100 \end{aligned}$$

3. To graph the resulting model, add the formulas in Figure 3.49. Copy row 9 down to row 58. Use these results, along with the original data, to form a graph similar to Figure 3.50.

	O	P
1	**Spline**	
2	**x**	**y**
3	=A2	=IF(O3<$A$3,$M$11+$M$12*O3+$M$13*O3^2+$M$14*O3^3,$M$15+$M$16*O3+$M$17*O3^2+$M$18*O3^3)
4	=O3+0.5	=IF(O4<$A$3,$M$11+$M$12*O4+$M$13*O4^2+$M$14*O4^3,$M$15+$M$16*O4+$M$17*O4^2+$M$18*O4^3)

FIGURE 3.49

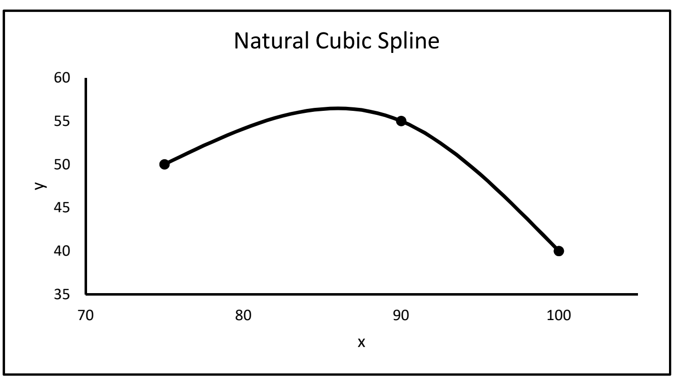


FIGURE 3.50

From the graph we see that the model forms a smooth curve that goes through each data point, as required.

□

For large sets of data the calculations for a cubic spline model can get quite complicated. For 11 data points, the model is made up of 10 separate cubic polynomials with a total of 40 unknown parameters. This results in a 40×40 matrix A. Excel can handle up to a 52×52 matrix, but setting it up is quite tedious. Fortunately, there are optional add-ins (and other software) that do the work automatically.

Exercises

3.5.1 To get the necessary 8 equations to determine the values of the 8 parameters in our cubic spline model, we added the conditions that $p_1''(x_1) = p_2''(x_3) = 0$. Setting these second derivatives equal to 0 is somewhat arbitrary. Let's experiment with changing these conditions.

a. In the worksheet **Cubic**, change the condition $p_2''(x_3) = 0$ to values other than 0 (i.e. change the number in cell **M9** to values other than 0). Does this change the shapes of the graphs of $p_1(x)$ and $p_2(x)$? (Use the data points shown in Figure 3.48 where $x_3 = 100$.)

b. In the original natural cubic spline model (with $p_2''(x_3) = 0$), the model predicts a sharp decrease in milk production around day 100 (in other words, $p_2'(x_3)$ is a large negative number). In the graph of the original data shown in Figure 3.43, we see that the milk production begins to "level off" between days 100 and 120. Find a value of $p_2''(x_3)$ that gives a model that "levels off" near $x = 100$.

c. Instead of specifying the value of $p_2''(x_3)$ in the last row of matrix A, we could replace this with any condition on $p_1(x)$ or $p_2(x)$, or their derivatives, that is independent of the other conditions. Notice that in the linear spline model the slope between days 100 and 120 is -0.5. Modify the worksheet **Cubic** by replacing the last row of matrix A with the condition that $p_2'(x_3) = -0.5$. Does the resulting model "level off" near $x = 100$?

3.5.2 Modify the worksheet **Cubic** to calculate and graph a cubic spline model fit to *four* data points.

3.5.3 Consider the problem of fitting a *quadratic* spline model to a set of 3 data points $\{(x_1, y_1), (x_2, y_2), (x_3, y_3)\}$ where $x_1 < x_2 < x_3$. The two polynomials have the general form

$$\begin{aligned} p_1(x) &= a_1 + b_1 x + c_1 x^2 \quad \text{for } x_1 \le x < x_2 \\ p_2(x) &= a_2 + b_2 x + c_2 x^2 \quad \text{for } x_2 \le x \le x_3 \end{aligned}$$

a. If these polynomials must satisfy the same 3 conditions as the cubic polynomials, find a set of 6 equations that determine the values of the parameters a_1, b_1, c_1, a_2, b_2, and c_2.

b. Design a spreadsheet to calculate the values of the parameters (use the same data points as in Example 3.5.2). What do you observe about the 2 polynomials $p_1(x)$ and $p_2(x)$?

c. Fit a 2^{nd} degree polynomial trendline to the data. How does this polynomial compare to $p_1(x)$ and $p_2(x)$? Why is this?

3.5.4 A researcher observes a particle moving in a straight line and measures its velocity at three points in time as recorded in the table below. He wants to estimate how far the particle traveled over the interval of time [0, 5].

Time (sec)	0	2	5
Velocity (m/sec)	0	60	85

In general, if $f(t)$ = velocity of an object at time t, we can calculate the distance it traveled over an interval of time $[t_0, t_1]$ by

$$\text{Distance} = \int_{t_0}^{t_1} |f(t)|\, dt$$

a. Fit a linear spline model to the data to estimate $f(t)$ and use the result to estimate the distance traveled. What is the meaning of the slope of each of the line segments?

b. Fit a cubic spline model to the data to estimate $f(t)$ and use the result to estimate the distance traveled. Also, estimate the acceleration at time $t = 1.5$.

c. Which of the two estimates of the distance traveled do you think is more accurate? Why?

Project Ideas

1. In Section 3.4 we used multiple regression models to find a good set of predictor variables for a response variable. Collect some data on a response variable of your choice and some possible predictor variables. Then apply the methods in Section 3.4 to find a good set of predictor variables. Here are some possibilities for response variables:

    1. Average life expectancy of residents of countries
    2. Golf scores of professional golfers
    3. Weight-lifting ability of college athletes
    4. College tuition rates
    5. Winning percentage of college basketball teams
    6. Mortality rates of residents of countries
    7. Cross country race times
    8. Salaries of NFL quarterbacks
    9. The collective GPA of college sports teams
    10. The population of pheasants (or any other type of wildlife) in a state
    11. Free throw percentage among NBA basketball players
    12. Fuel economy of vehicles
    13. Winning percentage of MLB baseball teams
    14. The number of shark bites in Florida per year
    15. Winning percentage of English Premier League teams

16. Crime rate of cities
17. Number of Olympic medals won by countries
18. Number of passing yards per year of NFL teams
19. Selling price of a vehicle
20. Cost of a computer
21. Personal income
22. Vertical jump height of NBA players
23. Batting averages of MLB players

2. Create a workbook that automatically fits a 19^{th} degree polynomial to 20 pairs of data inputted by the user.
3. Create and analyze an empirical model for the population growth of a city.
4. Research the topic of stepwise regression.
5. Research the topic of moving average.
6. Research the topic of standardized regression.
7. Zipf's law is a prediction of the frequencies of words in books. Research Zipf's law and apply it to one or more books of your choice.
8. Model the stopping distance of a vehicle in terms of its speed.
9. Model how a sports team winning percentage in one year affects the winning percentage in the next year.
10. Create a workbook that automatically calculates the eigenvalues and eigenvectors of a 4×4 matrix.
11. Research the details of the statistics in a multiple regression output (i.e. the confidence intervals).
12. Model the population of a country over time.
13. Model the Dow Jones Industrial Average over time.

For Further Reading

- For a good introduction to linear algebra, see Lay, D., *Linear Algebra and its Applications*, 5^{th} edition, 2016, Pearson or Williams, Gareth, *Linear Algebra with Applications*, 8^{th} edition, Jones and Bartlett, 2014.
- For more information on the 10–year fluctuations of wildlife populations, see L.G. Keith, *Wildlife's Ten Year Cycle*, University of Wisconsin Press, 1963.
- For a much more detailed theoretical discussion of most of the topics discussed in this chapter, see Hines, William W. et. al., *Probability and Statistics in Engineering*, Fourth edition, John Wiley & Sons, Inc, 2003, pg. 409 – 486.
- For a much more detailed discussion of spline models, see Bartels, Richard H., et. al., *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling*, Morgan Kaufmann Publishers, Inc., 1987.

4

Discrete Dynamical Systems

Chapter Objectives

- Define and solve discrete dynamical systems
- Analyze the long–term behavior of discrete dynamical systems numerically and graphically
- Model different scenarios with linear and nonlinear discrete dynamical systems

4.1 Introduction

A *dynamical system* is simply a system that changes over time. The bacterial growth modeled in Chapter 1 is one such example. When time is measured in discrete increments, such as in the bacterial growth model, the system is called a *discrete dynamical system*.

Dynamical systems are "easy to [model] and hard to solve" (Meerschaert, Mark M., *Mathematical Modeling*, Second ed., Academic Press, 1999, p. 127). In this chapter we will introduce basic techniques for formulating dynamical models and graphical approaches to their analysis.

In mathematical terms, a discrete dynamical system is simply a sequence of numbers. Consider a savings account that is compounded yearly and the interest is added at the end of each year. If a_n is the amount in the account at the end of year n ($n = 0, 1, \dots$) and r is the interest rate, we have the sequence

$$\begin{aligned} a_1 &= a_0 + r\, a_0 = (1+r)\, a_0 \\ a_2 &= a_1 + r\, a_1 = (1+r)\, a_1 \\ &\vdots \\ a_{n+1} &= a_n + r\, a_n = (1+r)\, a_n \end{aligned}$$

This last equation leads us to the formal definition of a dynamical system.

Definition 4.1.1 (Discrete Dynamical System)**.** A *discrete dynamical system* is a sequence of numbers $\{ a_n \,|\, n = 0, 1, ... \}$ defined by a relation of the form

$$a_{n+1} = f\,(a_n)$$

where f is some real–valued function.

The variable a_n is generically called the *state of the system*. In simpler terms, a discrete dynamical system is one in which the state of the system is determined by the previous state. In the savings account example there is only one component to the system (the amount in the account), so it is called *one-dimensional*.

When the function f has the form $f(x) = bx$ where b is a constant, the dynamical system is called *linear*.

Definition 4.1.2 (Linear Discrete Dynamical System)**.** A *linear discrete dynamical system* is a sequence of numbers $\{ a_n \,|\, n = 0, 1, ...\}$ defined by a relation of the form

$$a_{n+1} = ba_n$$

where $b \neq 0$ is a constant.

Any dynamical system that does not have this linear form is generically referred to as *nonlinear*. A *solution* of a discrete dynamical system is an explicit description of a_n in terms of n and the initial state a_0.

Theorem 4.1.1. *The solution of a linear dynamical system $a_{n+1} = ba_n$ for $b \neq 0$ is*

$$a_n = b^n a_0 \tag{4.1}$$

where a_0 is the initial state.

Proof. We first need to show that (4.1) satisfies the initial condition. Note that in (4.1),

$$a_0 = b^0 a_0 = a_0,$$

so the initial condition is satisfied. Next, we need to show that (4.1) satisfies the definition of a linear discrete dynamical system. Note that

$$a_{n+1} = b^{n+1} a_0 = b(b^n a_0) = ba_n$$

as required. □

4.2 Long–Term Behavior and Equilibria

Theorem 4.1.1 gives us an easy–to–use formula for finding the exact value of a_n. However we are usually more interested in describing the long–term behavior of the system than in finding exact values at points in time. That is, we want to know what happens to a_n for large values of n.

Example 4.2.1 (Long–term Behavior)
Let's graphically examine the long–term behavior of a linear dynamical system $a_{n+1} = ba_n$ for various values of b. For simplicity, suppose that $a_0 = 0.1$.

1. Rename a blank worksheet "**Linear**" and format it to look like Figure 4.1. Copy the formulas in **A3:B3** down to row 17. This will give the first 16 values of a_n $(0 \leq n \leq 15)$ with $b = 0.5$.

	A	B	C
1	**n**	**a_n**	**b**
2	0	0.1	0.5
3	=A2+1	=B2*$C$2	

FIGURE 4.1

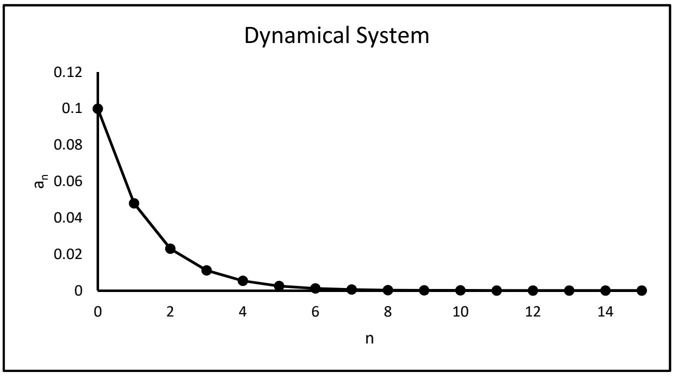


FIGURE 4.2

2. Highlight the column titled **n** and $\mathbf{a}_n$ and create a graph similar to Figure 4.2. Set the x-axis min and max to 0 and 15, respectively. Notice that for this value of b, the state of the system approaches 0 as n gets larger.

3. Next open the control toolbox by selecting **View** → **Toolbars** → **Control Toolbox**. Draw a Scroll Bar by selecting the icon on the left side of the Control Toolbox window second from the bottom. Your cursor will turn into a cross. Draw a horizontal Scroll Bar near the top of the worksheet (see Appendix A.4 for more information on drawing scroll bars).

4. Right–click on the Scroll Bar and select **Properties**. Set the **LinkedCell** to **D5**, the **max** to 100, and the **min** to 0. Close the Properties window. Exit Design Mode by selecting the icon in the upper left-hand corner of the Control Toolbox.

5. Slide the Scroll Bar left and right. The number in **D5** should change between 0 and 100. Enter a formula in **C2** as shown in Figure 4.3. This will allow us to change the value of b between -2 and $+2$ with the Scroll Bar.

	C
1	**b**
2	=-2+0.04*D5

FIGURE 4.3

6. Move the slider on the scroll bar left and right and notice how the behavior of the system changes, especially when b passes -1, 0, and +1. Our observations are summarized in Table 4.1. As we can see, the long–term behavior of the system is dramatically affected by the value of b.

□

Now consider a savings account that pays 5% interest compounded yearly. We saw in Section 4.1 that a model for an account with an interest rate r is

$$a_{n+1} = (1 + r)\, a_n.$$

In this case, we have $r = 0.05$, so our model is

$$a_{n+1} = 1.05\, a_n.$$

TABLE 4.1

Value of b	Behavior of a_n
$b < -1$	Oscillates between pos. and neg., $\lvert a_n \rvert$ grows without bound
$b = -1$	Oscillates between $-a_0$ and $+a_0$
$-1 < b < 0$	Oscillates between pos. and neg., $\lvert a_n \rvert$ approaches 0
$b = 0$	$a_n = 0$ for $n > 0$
$0 < b < 1$	a_n approaches 0
$b = 1$	$a_n = a_0$ for all n
$b > 1$	a_n grows without bound

Suppose now that we want to withdraw \$2,000 at the end of each year to supplement our income. We want to know how much money we need to deposit now so that we never run out of money.

To answer this question, we will analyze a slightly more general problem: What happens to the amount in the account in terms of the initial deposit? First we will construct our model. The amount in the account grows at 5% compounded yearly but we are withdrawing \$2,000 each year. A dynamic model that describes this scenario is

$$a_{n+1} = 1.05\, a_n - 2000.$$

As before, a_n is the amount in the account at the end of year n. We are also assuming that there is no penalty for withdrawing money each year and that we withdraw the money after the interest from the previous year has been added. This system is an example of an *affine* dynamical system.

Definition 4.2.1 (Affine Discrete Dynamical System)**.** An *affine discrete dynamical system* is a sequence of numbers $\{\, a_n \mid n = 0, 1, \ldots\}$ described by a relation of the form

$$a_{n+1} = ba_n + m$$

where $b \neq 0$.

Central to the analysis of the long–term behavior of any dynamical system are *equilibrium values* (also called *fixed points*).

Definition 4.2.2 (Equilibrium Value)**.** A number a is called an *equilibrium value* for the dynamical system $a_{n+1} = f(a_n)$ if $a_n = a$ for all n whenever $a_0 = a$.

To find equilibrium values, note that if a is an equilibrium value, we must have

$$a_{n+1} = a_n = a \quad \Rightarrow \quad f(a) = a$$

So finding equilibrium values simply requires us to solve the equation $f(a) = a$. For an affine system, we have

$$a = ba + m \quad \Rightarrow \quad a = \frac{m}{1-b}$$

In this example, $b = 1.05$ and $m = -2000$, so the equilibrium value is $a = \frac{-2000}{1-1.05} = 40,000$. Thus if we start with \$40,000 in the account and withdraw \$2,000 at the end of each year, we will always have the same amount in the account at the end of each year.

Example 4.2.2 (Savings Account)
We will take a graphical approach to analyze what happens for initial values other than the equilibrium value of \$40,000.

1. Rename a blank worksheet "**Savings**" and format it as in Figure 4.4. Copy the range **A3:B3** down to row 27 to model the account over the first 25 years.

	A	B	C	D
1	**n**	**a_n**	**r**	**m**
2	0	40000	0.05	2000
3	=A2+1	=(1+\$C\$2)*B2-\$D\$2		

FIGURE 4.4

2. Use the data in columns **A** and **B** to form a graph as in Figure 4.5. Set the y-axis min and max to 0 and 80,000, respectively, and the x-axis min and max to 0 and 25. Note that the value in the account does not change if we start with \$40,000, as expected.

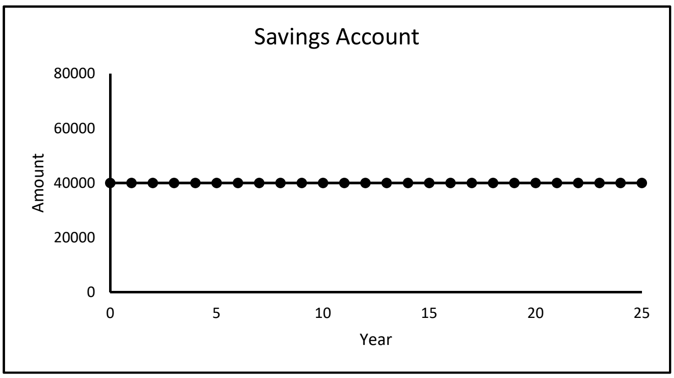


FIGURE 4.5

3. Next add a scroll bar. Set the linked cell to **B2** and the min and max to 0 and 80,000, respectively. This will allow us to vary the value of a_0 between \$0 and \$80,000 in increments of \$1.

4. Move the slider on the scroll bar left and right and observe how the long–term behavior of the system changes. Specifically, note that for a_0 less than \$40,000, the amount in the account eventually decreases to 0 and for a_0 greater than \$40,000, the amount grows without bound.

□

In Example 4.2.2 we saw that the long–term behavior of the system changed quite dramatically with a small change in a_0. In situations like this we say that the system is *sensitive to the initial condition.*

Also note that if $a_0 \neq 40{,}000$, the system either approaches 0 or increases without bound. The equilibrium value of 40,000 is an example of an *unstable* or *repelling* equilibrium.

Definition 4.2.3 (Repelling and Attracting Equilibria)**.** An equilibrium value a is *unstable* or *repelling* if there is a number ε such that

$$|a_n - a| > \varepsilon \quad \text{whenever} \quad |a_0 - a| < \varepsilon$$

for some n. The equilibrium a is *stable* or *attracting* if there is a number ε such that

$$\lim_{n\to\infty} a_n = a \quad \text{whenever} \quad |a_0 - a| < \varepsilon$$

In less technical terms, an equilibrium value a is repelling if the system starts near it, but does not approach it. This is exactly what we saw in Example 4.2.2. The equilibrium is attracting if the system starts near a and approaches it.

Example 4.2.3 (Antibiotic in the Bloodstream)
An infant is given an antibiotic to treat an ear infection. When taking an antibiotic, it is important to keep the amount of the drug in the bloodstream fairly constant. If it gets too low, the bacteria can begin to regrow. If it gets too high, it could cause other complications.

Suppose the half-life of the drug is 1 day (meaning that half the drug remains in the blood after each 1-day period) and a dosage of 0.1 mg is given at the end of each day. We want to examine what happens to the amount of the drug in the bloodstream in the long–run.

A simple affine model for this system is

$$a_{n+1} = 0.5\, a_n + 0.1$$

where $a_n =$ the amount of the drug in the blood at the end of day n. Since the problem did not specify the initial dosage, a_0, we need to experiment with different values.

1. Rename a blank worksheet "**Antibiotic**" and format it as in Figure 4.6. Copy row 3 down to row 17 to model the system from day 0 to day 15.

	A	B
1	**n**	**a_n**
2	0	0
3	=A2+1	=0.5*B2+0.1

FIGURE 4.6

2. Create a graph similar to that in Figure 4.7. Notice that even with an initial dosage of 0 mg, the amount of antibiotic in the blood appears to approach 0.2 mg at the end of each day. Note that this does not mean that the amount eventually equals 0.2 mg at every point in time, only that is equals 0.2 mg at the end of every day.

3. Next, add a scroll bar, set the min to 0, the max to 100, and the linked cell to **C1**. Add the formula in Figure 4.8 to allow us to vary the initial dosage from 0 to 1 mg in increments of 0.01 mg.

4. Move the slider on the scroll bar left and right and observe the long–term behavior of the system. Specifically note that when $a_0 = 0.2$, the system remains at 0.2, meaning that 0.2 is an equilibrium value. Also note that no matter what the value of a_0 is, the system appears to always approach 0.2. This shows that 0.2 is an attracting equilibrium.

□

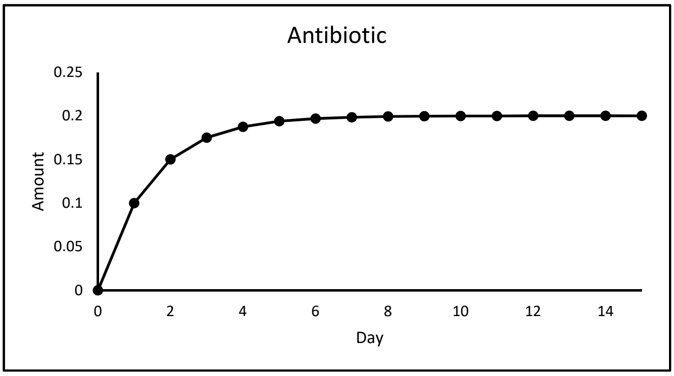


FIGURE 4.7

	B
1	**a_n**
2	=C1/100

FIGURE 4.8

Example 4.2.4 (Generalized Savings Account)
Let's generalize the model of the savings account. Suppose that we want to withdraw \$5,000 every other year (starting in year 2) instead of every year. This makes our model

$$a_{n+1} = \begin{cases} 1.05a_n - 5000 & \text{if } n+1 \text{ is a multiple of 2} \\ 1.05a_n & \text{otherwise} \end{cases}$$

This system is neither linear nor affine. We will take a strictly graphical approach to analyze the behavior in terms of the initial deposit.

1. Modify the worksheet **Savings** as in Figure 4.9 and copy the range **B3:C3** down to row 27. The **MOD** function in cell **C3** returns the remainder when the year n (cell **A3**) is divided by the number of years between withdrawals (cell **F3**). If this remainder is 0, then the year is a multiple of the years between withdrawals, and a withdrawal is taken that year. Otherwise, no withdrawal is taken. Note also that we have designed a cell to hold the parameter "Years between withdrawals" so that we can easily change its value to analyze other scenarios.

	A	B	C	D	E	F
1	**n**	**a_n**	**Withdrawal?**	**r**	**Amount of**	**Years between**
2	0	30000	0	0.05	**withdrawal**	**withdrawals**
3	=A2+1	=(1+$D$2)*B2-C3*$E$3	=IF(MOD(A3,$F$3)=0,1,0)		5000	2

FIGURE 4.9

2. Move the slider on the scroll bar left and right to change the value of a_0 and observe the long–term behavior of the system. Specifically, note that for a_0 below approximately \$48,000, the account eventually is depleted. For a_0 above \$48,000 the value tends to increase. For a_0 around \$48,000, the value fluctuates around \$48,000.

□

Exercises

4.2.1 Consider the linear dynamical system $a_{n+1} = ba_n$. Use the worksheet **Linear** to answer the following questions.

a. Verify that $a = 0$ is an equilibrium value by showing that if $a_0 = 0$, then $a_n = 0$ for all n.

b. Let's think about the stability of the equilibrium value $a = 0$. Informally, an equilibrium value is attracting if the system starts near the value and gets closer to the value. So suppose we start at $a_0 = 0.1$. Find the values of b for which the equilibrium value $a = 0$ appears to be attracting.

c. Informally, an equilibrium value is repelling if the system starts near the value, but gets further from the value. Find the values of b for which the equilibrium value $a = 0$ appears to be repelling.

4.2.2 Consider the affine system $a_{n+1} = ba_n + 1$ where $a_0 = 0$.

a. Calculate a_n for $0 \leq n \leq 15$. Graph these values. Set up a scroll bar to vary the value of b between -2 and $+2$.

b. Calculate the equilibrium value.

c. For what values of b is the equilibrium value attracting?

d. For what values of b is the equilibrium value repelling?

e. For what values of b is $\lim\limits_{n\to\infty} a_n = \infty$?

f. For what values of b is $\lim\limits_{n\to\infty} a_n = -\infty$?

4.2.3 Your grandparents have their life savings of \$750,000 in a savings account that pays 6.7% interest compounded yearly. They want to spend all of it before they die. If they plan to live 15 more years, how much should they withdraw at the end of each year to accomplish their goal?

4.2.4 In the generalized model of the savings account where we withdraw \$5,000 every two years, find a value of a_0 such that $a_2 = a_0$. If the initial deposit is this value, what is the long–term behavior of the system?

4.2.5 In the savings account examples in this section we considered only accounts that were compounded yearly, meaning that interest was added to the account at the end of each year. In this exercise we generalize this idea and allow interest to be added multiple times a year, such as monthly or daily. In this case the model is given by

$$a_{i+i} = \left(1 + \frac{r}{n}\right) a_i \tag{4.2}$$

where

- $i =$ the number of compounding periods from the time of the initial deposit,
- $r =$ the annual interest rate (in decimal form), and
- $n =$ the number of times compounded per year.

For instance, if the account were compounded daily, then i would represent days since the initial deposit and $n = 12$. Consider an account compounded monthly with an annual interest rate of 5% and an initial deposit of \$1,000.

a. Use the model (4.2) to calculate the amount in the account at the end of each of the first 24 compounding periods.

b. The formula for compound interest given in most high school algebra books is

$$A(t) = P\left(1 + \frac{r}{n}\right)^{nt}$$

where $A(t)$ in the amount in the account at the end of year t and P is the initial deposit. Verify that this formula gives the same results as part a.

4.2.6 Kayla receives a \$6,000 high school graduation present which she puts in a savings account paying 4.15% annual interest, compounded monthly (see Exercise 4.2.5). She plans to work for two years before starting college, put her earnings in the account, and then live off the money in the account during college. Here are the details of her plan:

- She will deposit the same amount into the account at the end of each month for the two years she works (starting in month 1).
- In month 25 she will stop making deposits and start withdrawing \$1,000 each month for her first year in college.
- During her second year she will withdraw \$1,250 each month. During her third and fourth years she will withdraw \$1,500 and \$1,750 each month, respectively.

Figure out how much she must deposit each month for the first two years so she has \$0 left in the account at the end of her fourth year of college. Approximately how much total interest does she earn over the six years?

4.2.7 Suppose the amount of a drug in a patient's blood stream decreases at the rate of 50% per hour and that an injection is given at the end of each hour which increases the amount of drug in the blood stream by 0.2 units.

a. Formulate a model of the amount of drug in the blood stream at the end of each hour.

b. Find all equilibrium value(s) of your model.

c. Graphically, classify each equilibrium value as attracting or repelling.

d. Suppose we give the injection every 3 hours. Describe what happens to the long–term level of the drug in the blood stream.

4.2.8 Suppose two countries are engaged in an arms race. Further suppose that the two countries have economies of similar strength and they have similar levels of distrust of each other. A simple model for T_n, the total amount of money spent by the two nations, is given by the affine system

$$T_{n+1} = (1 - r + d)\, T_n + c$$

where r is a positive constant that measures the restraint of growth due to the strength (or weakness) of the economies of the countries, d is a positive constant that measures the level of distrust between the countries, and c is a constant. If T_n eventually grows too large, then the countries will not be able to support the arms race and they must either negotiate an end or war breaks out. If T_n approaches a constant level, then they have a "stable" arms race. If T_n eventually decreases to 0, or $\lim\limits_{n\to\infty} T_n = -\infty$, then the race ends.

a. Suppose $r = 0.3$, $d = 1$, and $c = -10$. Calculate T_n for $0 \leq n \leq 15$. Graph these values. Set up a scroll bar to vary the value of T_0 between 0 and 100. For what values of T_0 does the race end?

b. Now suppose $r = 0.3$, $d = 1$, and $c = 10$. For what values of T_0 does the race end?

c. Now suppose $r = 0.3$, $d = 0.1$, and $c = 10$. What happens to the arms race now?

4.2.9 Consider the affine system $a_{n+1} = ba_n + m$ where $b \neq 0$.

a. Analytically (meaning without using Excel) show that $a_n = b^n\left(a_0 - \frac{m}{1-b}\right) + \frac{m}{1-b}$ is the solution to the system when $b \neq 1$.

b. Under what conditions is $\lim\limits_{n\to\infty} a_n = \infty$ (meaning what must be true about the values of m, b, and a_0 for which this limit holds)? Under what conditions is the limit equal to $-\infty$? In these cases, is the equilibrium value $a = \frac{m}{1-b}$ attracting or repelling?

c. Under what conditions is $\lim\limits_{n\to\infty} a_n = \frac{m}{1-b}$? In this case, is the equilibrium value $a = \frac{m}{1-b}$ attracting or repelling?

d. Consider the case where $b = 1$. Find $\lim\limits_{n\to\infty} a_n$.

4.2.10 Suppose you open a savings account that pays 5% interest compounded yearly with a \$500 initial deposit and make a \$200 deposit at the end of each year.

a. Construct a model of the amount of money in the account at the end of each year and define the variables.

b. Use this model and the solution in Exercise 4.2.9a to find the amount of time it would take to build a value of \$12,000.

4.2.11 Consider the nonlinear discrete dynamical system

$$a_{n+1} = a_n + 0.2a_n\left(1 - a_n\right)\left(2 - a_n\right).$$

a. Calculate a_n for $0 \leq n \leq 30$ and graph the system. Add a scroll bar to vary the value of a_0 between 0 and 2.1.

b. Describe how the long-term behavior of the system is related to a_0.

4.3 Discrete Logistic Equation

Table 4.2 gives the number of bacteria in a Petri dish, a_n, at the end of each hour n. This data is graphed in Figure 4.10. We want to model a_n in terms of n.

When modeling a dynamical system, it is often convenient to think about the way the variable(s) change between time periods. Specifically, we consider the change between time periods $\Delta a_n = a_{n+1} - a_n$. The values of Δa_n for the first 7 values of n are given in Table 4.3.

Notice that as a_n increases, Δa_n also increases. This suggests that Δa_n is proportional to a_n, which leads to the equation

$$\Delta a_n = a_{n+1} - a_n = r\, a_n \tag{4.3}$$

TABLE 4.2

n	0	1	2	3	4	5	6	7	8	9
a_n	10.3	17.2	27	45.3	80.2	125.3	176.2	255.6	330.8	390.4
n	10	11	12	13	14	15	16	17	18	19
a_n	440	520.4	560.4	600.5	610.8	614.5	618.3	619.5	620	621

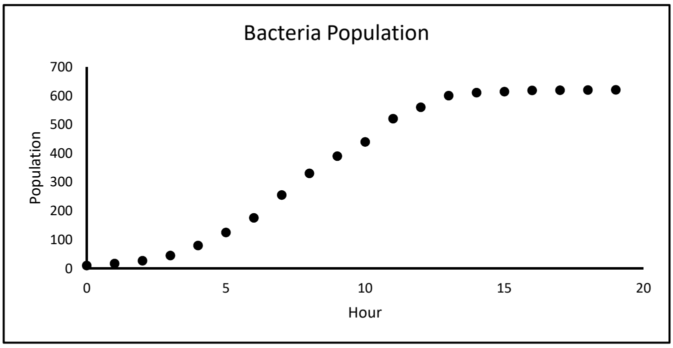


FIGURE 4.10

TABLE 4.3

n	0	1	2	3	4	5	6
a_n	10.3	17.2	27	45.3	80.2	125.3	176.2
Δa_n	6.9	9.8	18.3	34.9	45.1	50.9	79.4

where r is some positive constant. An equation describing the difference in populations between time periods, such as (4.3), is called a *difference equation*. Forming a difference equation is often the first step in modeling a discrete dynamical system. Solving this equation for a_{n+1} yields the model

$$a_{n+1} = (1 + r)\, a_n$$

The parameter r can be interpreted as the constant hourly growth rate. However, the graph of population vs. hour shows that the population does not grow at a constant rate. Also note that this constant hourly growth rate would predict a population that grows without bound, which the data do not support either.

To refine the model, note that the graph shows that the rate of growth decreases as the population nears 621. This number is called the *carrying capacity* of the system. So instead of assuming a constant growth rate r, we assume a growth rate that approaches 0 as the population approaches 621. An equation implementing this assumption is

$$\Delta a_n = a_{n+1} - a_n = b(621 - a_n)\, a_n \tag{4.4}$$

where $b > 0$ is a constant. Solving for a_{n+1} yields the model

$$a_{n+1} = a_n + b(621 - a_n)\, a_n \tag{4.5}$$

Equation (4.5) is an example of a *discrete logistic equation.*

Definition 4.3.1 (Discrete Logistic Equation)**.** A *discrete logistic equation* (also called a *logistic map* or a *constrained growth model*) is an equation of the form

$$a_{n+1} = a_n + b\,(c - a_n)\, a_n$$

where b and c are constants. This type of equation is often used to model population growth where a_n is the population at time n. The constant b is called the *intrinsic growth rate* and c is called the *carrying capacity*.

Example 4.3.1 (Bacteria Growth)
To implement the model (4.5) we need to find the value of b. Equation (4.4) predicts that $(a_{n+1} - a_n)$, is proportional to $(621 - a_n)\, a_n$. If a graph of $(a_{n+1} - a_n)$ vs. $(621 - a_n)\, a_n$ is approximately a straight line through the origin, then the assumption is reasonable and the slope of the line is the value of b.

1. Rename a blank worksheet "**Bacteria**" and format it as in Figure 4.11. Enter the data from Table 4.2 in columns **A** and **B** and copy range **D2:E2** down to row 20.

	A	B	C	D	E
1	**n**	**a_n**	**Predicted**	**a_n(621 - a_n)**	**a_{n+1} - a_n**
2	0	10.3		=B2*(621-B2)	=B3-B2

FIGURE 4.11

2. Create a graph of the transformed data in columns **D** and **E** and fit a straight line through the origin as in Figure 4.12. We see that the line fits the data well, so our model appears to be reasonable.

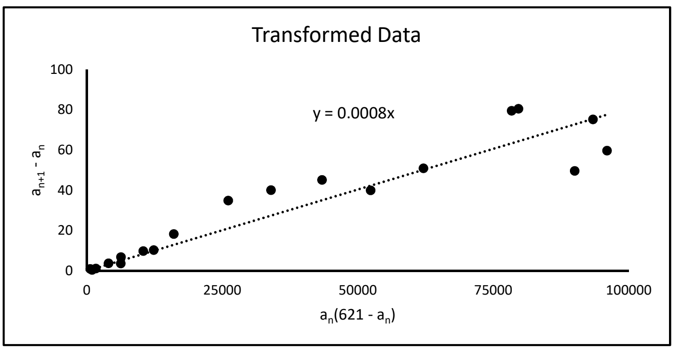


FIGURE 4.12

3. Using the slope of the line in Figure 4.12, our model is

$$a_{a+1} = a_n + 0.0008\,(621 - a_n)\, a_n$$

To test this model against the given data add the formula in Figure 4.13 and copy row 3 down to row 21.

	C
1	**Predicted**
2	10.3
3	=C2+0.0008*(621-C2)*C2

FIGURE 4.13

4. Use the data in columns **A**, **B**, and **C** to form a graph as in Figure 4.14. Notice that the "shape" of the predicted values is relatively close to the shape of the observed values, so the reasonableness of our model is verified. In Exercise 4.3.2 we will fit a logistic equation to this data in a different way which will yield a better fitting model.

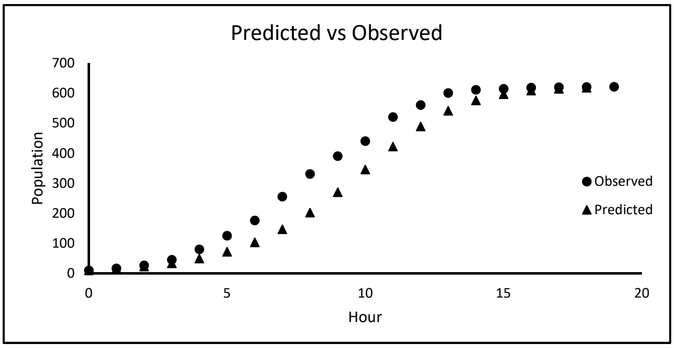


FIGURE 4.14

□

Example 4.3.2 (Sensitivity to the Intrinsic Growth Rate)
In Section 4.2 we looked at how the long–term behavior of the savings account changes as the initial deposit changes. Now we look at how the behavior of our constrained population model changes as the value of b changes. We take a strictly graphical approach.

1. Rename a blank worksheet "**Bacteria 2**" and format it as in Figure 4.15. Copy row 3 down to row 100.

	A	B	C
1	**n**	**a_n**	**b**
2	0	10.3	0.0008
3	=A2+1	=B2+$C$2*(621-B2)*B2	

FIGURE 4.15

2. Highlight columns **A** and **B** and form a graph as in Figure 4.16. Set the y-axis min and max to 0 and 800, respectively.

3. Next add a scroll bar, set the linked cell to **D2**, set the min and max to 0 and 900, and add the formula in Figure 4.17. This will allow us to vary the value of b between 0 and 0.0045.

4. Move the slider left and right and notice how the behavior of the system changes, especially when b increases above 0.0026. Several examples are shown in Figure 4.18.

Notice that the behavior of the system changes quite dramatically as b changes. For small values of b, the behavior is very regular. But for larger values, it is quite chaotic. This illustrates that we must be careful about using a model of this form if we are unsure of the value of b. The value of b dramatically affects the behavior of the system. If we use a wrong value of b, our analysis of the system could be very inaccurate. □

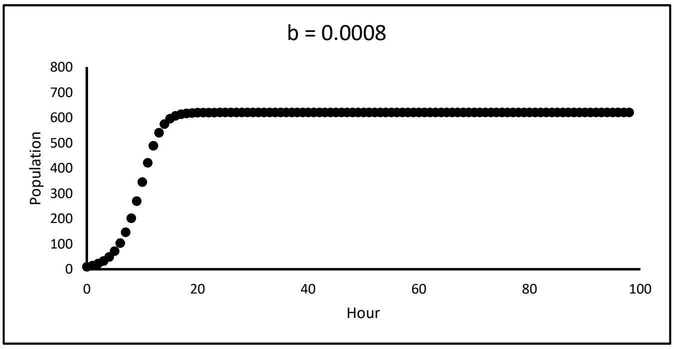


FIGURE 4.16

	C
1	**b**
2	=D2/200000

FIGURE 4.17

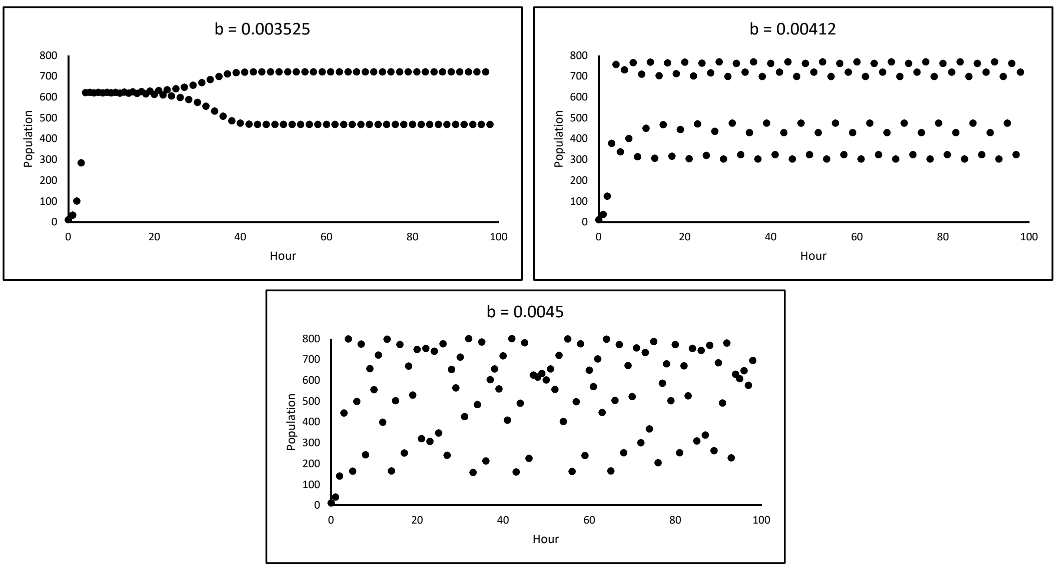


FIGURE 4.18

Exercises

4.3.1 The table below contains data on the population of foxes in a forest over a period of several years. Fit a discrete logistic equation to the data. How well does the model fit the data?

n	0	1	2	3	4	5	6	7	8	9	10
a_n	50	85	110	130	175	200	215	221	228	232	234

4.3.2 In this exercise we will fit a discrete logistic equation to the data in Table 4.2 using the least–squares criterion. This criterion says, informally, that when predicting the values of data $\{a_n \,|\, n = 1, \ldots, m\}$, the number

$$\sum_{i=1}^{m} (a_n - \text{Predicted})^2$$

(called the sum of squares) should be as small as possible. With a carrying capacity of 621, use a scroll bar to find a value of b in the discrete logistic equation that minimizes the sum of squares. Does this give the same model as the one found in Example 4.3.1? Does this new model fit the data any better than the model found in the example?

Here are some suggestions:

1. Modify the worksheet **Bacteria** by adding a column titled "$(a_n - \text{Predicted})^2$." Sum the values in this column to calculate the sum of squares.
2. Create a cell to hold the value of b. Reference this cell in your formula for the predicted values.
3. Add a scroll bar with a min and max of 0 and 1,000. Create a formula for the value of b equal to the scroll bar linked cell divided by 500,000.
4. Move the slider back and forth to find a value of b that minimizes the sum of squares.

4.3.3 Suppose we estimate that a forest can support a population of 10,000 deer and that the population of deer is described by the model $a_{n+1} = a_n + 0.00006(10000 - a_n)\, a_n$ where a_n is the population at the end of year n.

a. Suppose that we let hunters kill 700 deer at the end of each year. Write a model to describe this situation and analyze the long–term behavior of the population for different initial populations.

b. Suppose we start with a population of 9,000 deer and we allow hunters to kill m deer at the end of each year. Analyze the long–term behavior of the population for different values of m. What is the maximum value of m for which the population survives in the long–term?

c. Suppose we start with a population of 9,000 deer and we allow hunters to kill a certain proportion of the population (such as 0.25) at the end of each year. Write a model to describe this situation and analyze the long–term behavior of the population for different values of the proportion.

4.3.4 Consider a cup of coffee that is initially 100 °F, cools to 90 °F in 10 minutes, and sits in a room whose temperature is a constant 60 °F. A simple assumption for how the coffee cools is that the amount the temperature changes from one minute to the next is proportional to the difference between the coffee temperature and the room temperature. That is,

$$a_{n+1} - a_n = k\,(a_n - 60)$$

where a_n is the coffee temperature at minute n and k is some constant.

a. Implement this model to predict the coffee temperature for $0 \leq n \leq 150$. Use a scroll bar to find the value of k that yields $a_{10} \approx 90$.

b. Create a graph of a_n vs n.

c. Does this model seem reasonable? Briefly explain why or why not.

4.4 A Linear Predator–Prey Model

Consider a forest containing foxes and rabbits where the foxes eat the rabbits for food. We want to examine whether the two species can survive in the long–term. A forest is a very complex ecosystem. So to simplify the model, we will use the following assumptions:

1. The only source of food for the foxes is rabbits and the only predator of the rabbits is foxes.
2. Without rabbits present, foxes would die out.
3. Without foxes present, the population of rabbits would grow.
4. The presence of rabbits increases the rate at which the population of foxes grows.
5. The presence of foxes decreases the rate at which the population of rabbits grows.

We will model these populations using a discrete dynamical model. Each state of the system consists of the populations of foxes and rabbits at a point in time. Since this state consists of two components, this is a *two–dimensional discrete dynamical system*.

To create our model, we first need to define some variables. Let

$$F_n = \text{population of foxes at the end of month } n$$
$$R_n = \text{population of rabbits at the end of month } n$$

As in the bacteria model, the assumptions are stated in terms of rates of change, $\Delta F_n = F_{n+1} - F_n$ and $\Delta R_n = R_{n+1} - R_n$. There are many ways we could model these rates of change with the assumptions. In this section we will create a linear model. In the next section we will create a nonlinear model.

Assumptions 2 and 3 deal with the rates of change of each population in the absence of the other. A reasonable way to model these is to say that the rates are proportional to the populations. This yields the difference equations

$$\Delta F_n = F_{n+1} - F_n = -a F_n \tag{4.6}$$
$$\Delta R_n = R_{n+1} - R_n = d R_n \tag{4.7}$$

where both a and d are between 0 and 1. Note that the coefficient of proportionality in (4.6) is negative to reflect the fact that the foxes would die out (a negative rate of change) without rabbits. The coefficient in (4.7) is positive because the population of rabbits grows (a positive rate of change) without foxes.

Now, assumptions 4 and 5 say that these rates in Equations (4.6) and (4.7) either increase or decrease in the presence of the other species. So to incorporate these assumptions, we will simply add one term to each of Equations (4.6) and (4.7) yielding:

$$F_{n+1} - F_n = -a F_n + b R_n \tag{4.8}$$
$$R_{n+1} - R_n = -c F_n + d R_n \tag{4.9}$$

where b and c are non-negative. Note that the added term in (4.8) is positive to reflect the fact that the presence of rabbits increases the rate at which the population of foxes grows. The added term in (4.9) is negative since the presence of foxes decreases the rate at which rabbits grow.

Rewriting Equations (4.8) and (4.9) yields our model in the form of a system of linear equations

$$F_{n+1} = (1-a)F_n + bR_n \tag{4.10}$$
$$R_{n+1} = -cF_n + (1+d)R_n \tag{4.11}$$

Because our model has the form of a system of linear equations, it is called a *two-dimensional linear discrete dynamical system*.

The model could be written in matrix form as

$$\begin{bmatrix} F_{n+1} \\ R_{n+1} \end{bmatrix} = \begin{bmatrix} 1-a & b \\ -c & 1+d \end{bmatrix} \begin{bmatrix} F_n \\ R_n \end{bmatrix} \tag{4.12}$$

For a description of how to analyze this model using matrix and eigenvalue techniques, see, for instance, Lay, David C., *Linear Algebra and its Applications*, Third ed., Addison Wesley, 2003, pp. 342. We will take a strictly graphical approach to analyze the model.

The parameters $(1-a)$ and b are called the fox *death* and *birth* factors, respectively, while the parameters $-c$ and $(1+d)$ are called the rabbit death and birth factors, respectively.

1. Rename a blank worksheet "**Linear Predator–Prey**" and format is as Figure 4.19. The initial values of the parameters and populations are shown in the figure. Copy row 8 down to row 37 to model 30 months.

	A	B	C
1		**Factors**	
2		**Death**	**Birth**
3	**Foxes**	0.5	0.4
4	**Rabbits**	-0.17	1.1
5			
6	**Month**	**Foxes**	**Rabbits**
7	0	500	200
8	=A7+1	=$B$3*B7+C7*$C$3	=B7*$B$4+C7*$C$4

FIGURE 4.19

2. Next, add the graphs in Figure 4.20. The graphs of rabbits versus month and foxes versus month are called *time plots*. The curve in the graph of rabbits versus foxes is called a *trajectory* of the system. The plane on which a trajectory is drawn is called the *phase plane*. Notice that the trajectory tends toward the origin (0 foxes and 0 rabbits). This means that both species eventually die out. This is also shown in the time plots. If we change the initial populations, we note that the trajectories always tend toward the origin. This indicates that the populations always die out, regardless of the initial populations.

As in a one-dimensional discrete dyanmical system, two-dimensional systems can have an *equilibrium*.

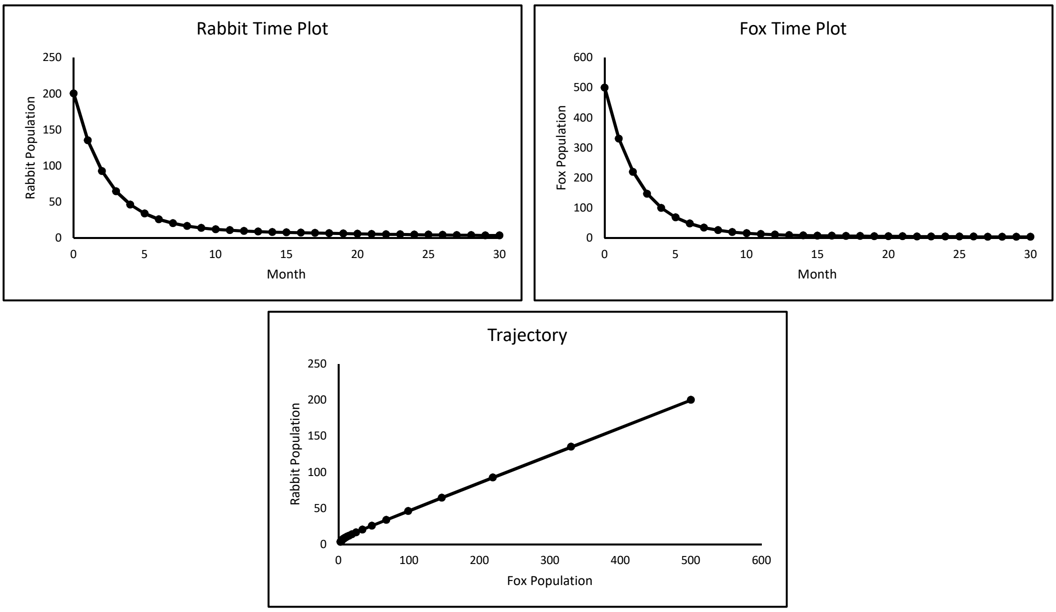


FIGURE 4.20

Definition 4.4.1. Let

$$a_{n+1} = f(a_n, b_n), \; b_{n+1} = g(a_n, b_n)$$

be a two-dimensional discrete dynamical system. A point (a, b) is an *equilibrium point* if $a_n = a$ and $b_n = b$ for all n whenever $a_0 = a$ and $b_0 = b$.

Stated another way, (a, b) is an equilibrium point if $f(a, b) = a$ and $g(a, b) = b$. For the system given by (4.10) and (4.11), the origin, $(0, 0)$ is an equilibrium point since

$$f(0, 0) = (1 - a)0 + b0 = 0 + 0 = 0$$
$$g(0, 0) = -c0 + (1 + d)0 = 0 + 0 = 0$$

Since the trajectories appear to be attracted to the origin, we have graphical evidence that the origin is an *attracting equilibrium*. If the trajectories tended to go away from the origin, we would say the origin is a *repelling equilibrium*.

Example 4.4.1 (Bifurcation)
Let's examine what happens if the rabbit death factor changes.

1. Add a scroll bar to the worksheet, set the min and max to 0 and 500, respectively, and the linked cell to **H1**.

2. Add the formula in Figure 4.21 to vary the rabbit death factor from 0 to -0.50 in increments of 0.001.

3. Move the slider on the scroll bar left and right and notice how the behavior of the system changes. For values of the death factor less than -0.123, the origin appears to be attracting and both populations die out, as in Figure 4.20. For values greater than -0.123, the origin appears to be repelling and both populations grow without bound, as in Figure 4.22. This change of behavior caused by a change in the value of a parameter is called a *bifurcation*.

	B
4	=-H1/1000

FIGURE 4.21

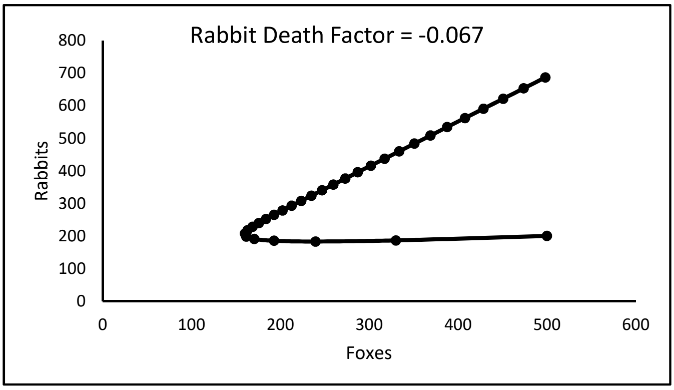


FIGURE 4.22

□

Notice that this model predicts that both populations either grow without bound or die off. These are two extremes. The model in Section 4.5 is a refinement of this one that allows for other possibilities.

Exercises

4.4.1 Suppose the rabbit death factor is -0.05 and the other factors are as shown in Figure 4.19. For initial populations of 500 foxes and 200 rabbits, we saw that the populations grow without bound in the long–run. Here we will investigate whether this is true for all initial populations.

a. Fix the initial rabbit population at 200. What happens if the initial fox population increases or decreases? How large can the initial fox population be for both species to survive?

b. Fix the initial fox population at 500. What happens if the initial rabbit population increases or decreases? How small can the initial rabbit population be for both species to survive?

4.4.2 Investigate sensitivity to the parameter rabbit birth factor using values of the other parameters and the initial populations shown in Figure 4.19 (i.e. increase and decrease the rabbit birth factor and observe what happens to the populations in the long–run).

a. For what values of the rabbit birth factor do the populations die out? For what values do the populations grow without bound? Are there any values for which the populations appear to reach an equilibrium?

b. A forest ranger knows that the populations will both die out under the parameters in Figure 4.19. To prevent this, she advocates building rabbit shelters in an attempt to increase the rabbit birth factor. Does this model predict that this would have any effect on the survival of the populations? Why or why not?

4.4.3 With the parameters and the initial populations shown in Figure 4.19 both populations die out. A hunting club claims that if they were allowed to kill (or harvest) a few foxes each month the populations would survive.

a. Modify the model to include this harvesting. Does the model support this claim? If it does, how many foxes could be harvested each month for both species to survive?

b. A conservation group claims that more foxes should be put into the forest each month for the populations to survive. Does your model support this claim?

4.4.4 A biologist models the populations of foxes and rabbits in another forest with the model

$$\begin{aligned} F_{n+1} &= 1.3F_n + 0.4R_n \\ R_{n+1} &= -0.6F_n + 1.05R_n \end{aligned}$$

What does this model suggest about the growth or death of foxes in the absence of rabbits? Is this model consistent with the assumptions used in building the model in this section? Why or why not?

4.4.5 Consider a two-dimensional system of the form

$$\begin{aligned} x_{n+1} &= ax_n + by_n \\ y_{n+1} &= cx_n + dy_n \end{aligned}$$

a. If $1 - d - \frac{cb}{1-a} \neq 0$, show that the only equilibrium point is $(0, 0)$.

b. If $1 - d - \frac{cb}{1-a} = 0$, show that any point of the form $(x, x(1-a)/b)$ is an equilibrium point.

4.4.6 Consider the linear two-dimensional discrete dynamical system

$$\begin{aligned} x_{n+1} &= 0.5x_n + 0.4y_n \\ y_{n+1} &= -0.104x_n + 1.1y_n. \end{aligned}$$

a. Numerically verify that the solution to the system is

$$\begin{aligned} x_n &= 10c_1(1.02^n) + 5c_2(0.58^n) \qquad \text{where} \qquad & c_1 &= \frac{y_0}{11} - \frac{x_0}{55} \\ y_n &= 13c_1(1.02^n) + c_2(0.58^n) & c_2 &= \frac{13x_0}{55} - \frac{2y_0}{11} \end{aligned}$$

for any initial conditions x_0 and y_0. (That is, calculate the state of the system for several values of n, say $n = 0$ to 30, as we did in Figure 4.19. Then calculate the state at each value of n using the claimed solution. Try different initial conditions. Verify that the solution always gives the correct value of the state.)

b. Using this solution, explain why the origin is a repelling equilibrium point.

4.5 A Nonlinear Predator–Prey Model

Let's consider a similar population of foxes and rabbits along with the same set of assumptions as in Section 4.4, but we will model the assumptions differently. We will start with modeling assumptions 2 and 3 the same way:

$$\Delta F_n = F_{n+1} - F_n = -a F_n \tag{4.13}$$

$$\Delta R_n = R_{n+1} - R_n = d R_n \tag{4.14}$$

where $0 < a \leq 1$ and $0 < d \leq 1$. In Section 4.4, the coefficients of F_n and R_n were kept constant. In this section we will model them as increasing or decreasing in the presence of the other population. Assumption 4 says that the presence of rabbits increases the rate of growth of foxes, so we write

$$F_{n+1} - F_n = (-a + bR_n)F_n \tag{4.15}$$

where $b \geq 0$. Likewise, assumption 5 says that the presence of foxes decreases the rate of growth of rabbits, so we have

$$R_{n+1} - R_n = (d - cF_n)R_n \tag{4.16}$$

where $c \geq 0$. Rewriting (4.15) and (4.16) we get our model:

$$F_{n+1} = (1 - a)F_n + bR_nF_n \tag{4.17}$$

$$R_{n+1} = -cR_nF_n + (1 + d)R_n \tag{4.18}$$

This type of model is called a *Lotka-Volterra* model, named after the researchers that first devised it in the 1920s and 1930s.

Note that both equations have a term involving R_nF_n; thus the model in nonlinear. This term can be interpreted as modeling the number of *interactions* of the two species. These interactions increase the number of foxes while decreasing the number of rabbits. Also note the similarities between this nonlinear model and the linear model in (4.10). We will refer to the parameters in this model using the same names as in the linear model.

This model can easily be implemented in Excel.

1. Rename a blank worksheet "**Nonlinear Predator–Prey**" and format it as in Figure 4.23. Copy row 8 down to row 507 to model 500 months. (Note that the parameters in this model do have *similar* meanings as in the linear model, but they do have different values. Also we have different initial populations.)

	A	B	C
1		**Factors**	
2		**Death**	**Birth**
3	**Foxes**	0.88	0.0001
4	**Rabbits**	-0.0003	1.039
5			
6	**Month**	**Foxes**	**Rabbits**
7	0	110	900
8	=A7+1	=$B$3*B7+$C$3*B7*C7	=$B$4*B7*C7+$C$4*C7

FIGURE 4.23

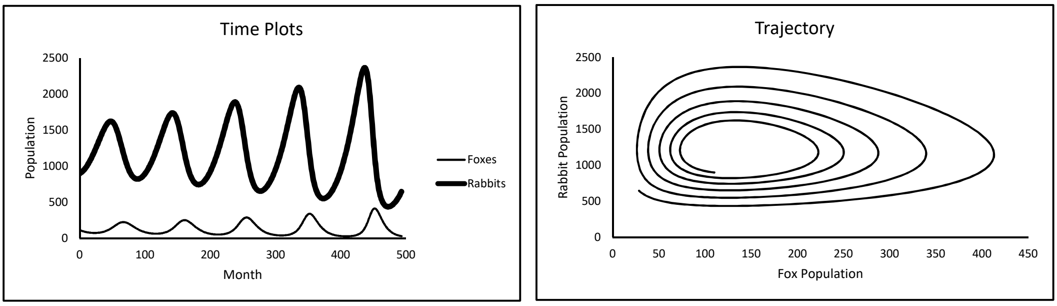


FIGURE 4.24

2. Create graphs similar to those in Figure 4.24.

This model predicts that the populations oscillate with the same period of oscillation, but with a phase shift, meaning they don't reach their peaks at the same time. These oscillations cause the spiraling nature of the trajectories in the graph of rabbits versus foxes. Oscillations such as this are actually observed in nature; thus this model appears to be more reasonable than the linear model.

Now let's calculate the equilibrium point of the system. Suppose (f, r) is an equilibrium point. By definition, this point must satisfy the system of equations

$$f = 0.88f + 0.0001fr$$
$$r = -0.0003fr + 1.039r$$

Assuming that $f \neq 0 \neq r$ yields the solution $f = 130$ and $r = 1,200$. Another equilibrium is (0, 0). Note that the point (130, 1200) is at the center of the spiral in the phase plane. If we change the starting populations in the worksheet to 130 foxes and 1200 rabbits we note that the populations do not change, as expected.

To determine if this equilibrium is attracting or repelling, we need to consider starting populations near the equilibrium. Changing the initial populations to 129 foxes and 1201 rabbits yields the trajectory shown in Figure 4.25. Notice that the trajectory moves away from the equilibrium. Trying other initial populations yields similar results. The fact that the trajectories move away from the equilibrium is evidence that the equilibrium is repelling.

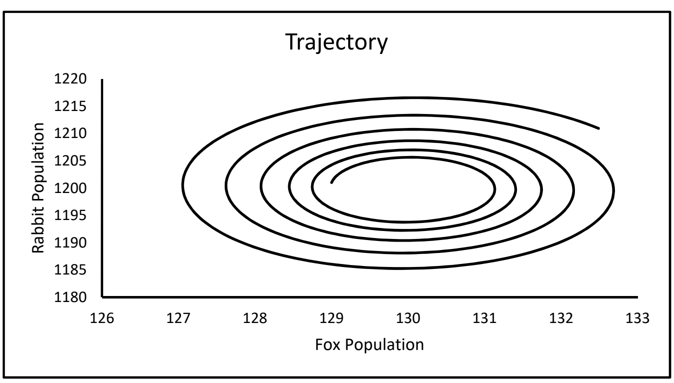


FIGURE 4.25

Example 4.5.1 (Hunting)
Suppose that a local hunting club wants to have a fox hunt at one of their next two meetings. The first meeting is 36 months from now and the next is 36 months after that. If they limit themselves to killing 50 foxes, how would the two options affect the long–term populations of the foxes and rabbits?

Starting with 110 foxes and 900 rabbits, the model predicts that at month 36, there will be approximately 88 foxes. Killing 50 foxes would reduce this number to 38. In the worksheet, changing the number of foxes in month 36 to 38 results in the graph shown in Figure 4.26. Notice that this causes a dramatic change in the behavior of the system. The populations fluctuate much more than they did in Figure 4.24. So, hunting 50 foxes in month 36 would have a great effect on the long–term populations.

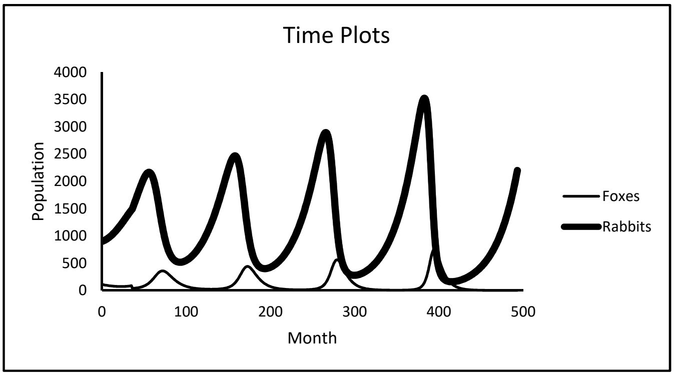


FIGURE 4.26

Now change the number of foxes in month 36 back to the original formula. In month 72, the model predicts a population of 212. Killing 50 foxes would leave 162. Changing the number of foxes in month 72 to 162 results in the graph shown in Figure 4.27. Notice that the fox population drops in month 72 which causes the rabbit population to initially grow. However, in the long–term the populations behave much as they did in Figure 4.24.

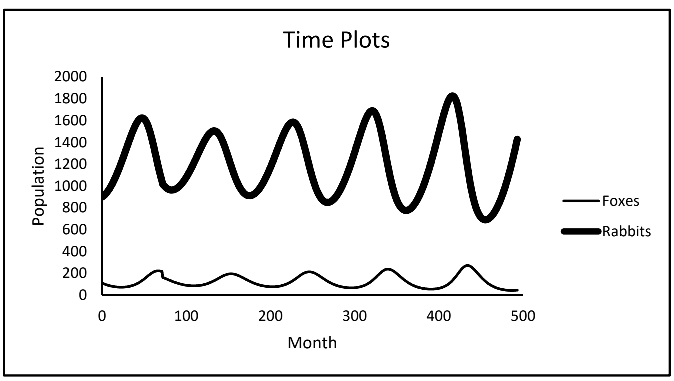


FIGURE 4.27

Why is there such a dramatic difference between these two cases? Note that originally near month 36, the fox population was near a local minimum. At month 72, it was near a local maximum. Killing 50 foxes near a time of a local minimum has a much greater effect than near a local maximum.

Thus the hunting club should not schedule the hunt at month 36; this would cause too great an effect on the populations. The hunt at month 72 would not cause a long–term effect.

□

Exercises

4.5.1 In the fox-rabbit model, we saw that the equilibrium point (130, 1200) is technically repelling.

a. Comment on how "repelling" this point is. That is, if the populations start near (130, 1200), do they very quickly move away from this point or not?

b. In Figure 4.23 we started with initial populations of (110, 900) and saw that the populations fluctuate quite a bit. Suppose a biologist advocates introducing *about* 20 new foxes and 300 new rabbits to the forest so that the initial populations are *near* (130, 1200) Would this help stabilize the populations, or would they fluctuate more? Explain.

4.5.2 Consider the parameter fox death factor.

a. Investigate the sensitivity of the system to the value of this parameter (use initial populations of 110 foxes and 900 rabbits). Specifically, comment on how the amplitudes of the oscillations change as this parameter changes.

b. Can this death factor ever be greater than or equal to 1? Why or why not? (**Hint**: See Equations (4.13) and (4.17).)

c. This parameter could be interpreted as the proportion of foxes that survive from one month to the next in the absence of rabbits. Suppose that a disease infects the fox population, decreasing the proportion that survives each month. What effect might this have on the populations?

4.5.3 Consider a model of the form

$$\begin{aligned} F_{n+1} &= aF_n + bF_nR_n \\ R_{n+1} &= cR_n + dF_nR_n \end{aligned}$$

where a, b, c, and d are non–zero constants. Find a formula for the equilibrium points of the system in terms of a, b, c, and d.

4.5.4 Suppose hunters are allowed to kill m rabbits at the end of each month.

a. Modify the model to take this into account (use the parameters and initial populations shown in Figure 4.23).

b. What effect will this have in the long–term? Would you say the system is sensitive to the parameter m?

c. How many rabbits could the hunters kill each month and still have the populations survive in the long–term?

4.5.5 A biologist is concerned by the large fluctuations in the fox and rabbit populations. To combat this, she advocates the building of rabbit shelters to help increase the survival of baby rabbits, thus increasing the rabbit birth factor. Does our model predict that this would have the intended effect (start with the parameters and initial conditions in Figure 4.23)? Explain your answer.

4.5.6 Consider a forest that contains Foxes and Wolves which compete for the same food resources. If F_n and W_n represent the populations of foxes and wolves, respectively, at the end of month n, a model for their populations is

$$F_{n+1} = 1.2\, F_n - 0.001\, F_n W_n$$
$$W_{n+1} = 1.3\, W_n - 0.002\, F_n W_n$$

a. Consider the parameters 1.2 and 1.3. Explain why both of them are greater than 1.

b. Why are the parameters -0.001 and -0.002 both negative?

c. Find the equilibria for this system and graphically determine if they are attracting or repelling (consider only values up to $n = 25$).

d. Is this model sensitive to the initial populations? Why or why not?

4.6 Epidemics

Consider a community of 1,000 people in which three members get sick with the flu. The following week, five new cases of the flu are reported. We are interested in modeling the spread of the disease through the community.

Consider the following assumptions:

1. Nobody enters or leaves the community and no one in the community has contact with anyone outside the community.

2. Each person is either Susceptible (able to get the flu), Infected (currently has the flu and able to spread it), or Removed (already had the flu and is not able to get it again). Initially each person is either susceptible or infected.

3. A susceptible person can get the flu only by contact with an infected person.

4. Once a person gets the flu, he/she cannot get it again.

5. The average duration of the flu is 2 weeks, during which time an infected person can spread the disease to a susceptible person.

The model we are going to build is called an *SIR* model (see, for instance, Allman, Elizabeth Spencer and John A. Rhodes, *Mathematical Models in Biology: An Introduction*, Cambridge University Press, 2004 for more information on this type of model). We begin by dividing the population into three categories: Susceptible, Infected, and Removed, as described in assumption 2. People move between these three categories as illustrated in Figure 4.28.

Let S_n, I_n, and R_n represent the numbers of people that are susceptible, infected, and removed, respectively, at the end of week n. As in previous models, we will begin by modeling the *change* of these variables in terms of difference equations.

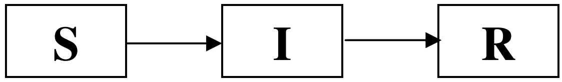


FIGURE 4.28

Let's begin by modeling R_n. Assumption 5 says that the average duration of the flu is 2 weeks. This means that, on average, about half the infected people will be healed (or removed) each week. Therefore, if $\gamma = 0.5$, a difference equation for R_n is

$$\Delta R_n = R_{n+1} - R_n = \gamma I_n \tag{4.19}$$

yielding the model

$$R_{n+1} = R_n + \gamma I_n \tag{4.20}$$

The parameter γ is called the *removal rate* and represents the proportion of the infected people that are removed each week. The quantity γI_n can be thought of as the number of "newly removed" people each week.

Now for I_n. This quantity will increase due to some newly infected people that come as a result of the interactions of the susceptible and infected people, and it will decrease due to the newly removed people. Therefore, a difference equation is

$$\Delta I_n = I_{n+1} - I_n = \alpha\, S_n I_n - \gamma I_n \tag{4.21}$$

The first term, $\alpha\, S_n I_n$, models the number of interactions of the susceptible and infected people, similar to the nonlinear predator–prey model. This quantity can be thought of as the number of "newly infected" people. Notice that the number of newly infected people in a week is a product of the number of infected and susceptible people in the previous *week*. The parameter α is called the "transmission coefficient." It is a measure of the likelihood that an interaction between an infected person and a susceptible person will result in an infection and is most likely a very small number. Equation (4.21) yields the model

$$I_{n+1} = I_n + \alpha\, S_n I_n - \gamma I_n \tag{4.22}$$

Lastly, to model S_n, note that this quantity is decreased only by the number of newly infected people and it does not increase. Thus a difference equation is:

$$\Delta S_n = S_{n+1} - S_n = -\alpha\, S_n I_n \tag{4.23}$$

Therefore the model is

$$S_{n+1} = S_n - \alpha\, S_n I_n \tag{4.24}$$

The overall model is given by the three equations

$$\begin{aligned} S_{n+1} &= S_n - \alpha\, S_n I_n \\ I_{n+1} &= I_n + \alpha\, S_n I_n - \gamma I_n \\ R_{n+1} &= R_n + \gamma I_n \end{aligned}$$

Now to determine the values of the parameters α and γ. We already noted that γ is the proportion of infected people removed each week. In this case, $\gamma = 0.5$. In general,

$$\gamma = \frac{1}{\text{average duration of the infectious period}}$$

To find the value of α, we need to use the fact that we started with 3 sick people and had 5 new cases the first week. In terms of the variables, this means that in week 0,

$$I_0 = 3 \text{ and } S_0 = 997$$

In week 1, the number of newly infected people is 5, so

$$5 = \alpha S_0 I_0 = \alpha(3 \cdot 997) \tag{4.25}$$

$$\Rightarrow \alpha = \frac{5}{3 \cdot 997} = 0.00167 \tag{4.26}$$

Now we can implement the model.

1. Rename a blank worksheet "**Epidemics**" and format it as in Figure 4.29. Copy row 8 down to row 32 to model 25 weeks.

	A	B	C	D	E	F
1		**Population**	1000			
2		**Transmission Coefficient**	0.00167			
3		**Removal Rate**	0.5			
4						
5					**Newly**	**Newly**
6	**Week**	**Susceptibles**	**Infectives**	**Removals**	**Infected**	**Removed**
7	0	=C1-C7	3	0	=$C$2*B7*C7	=$C$3*C7
8	=A7+1	=B7-E7	=C7+E7-F7	=D7+F7	=$C$2*B8*C8	=$C$3*C8

FIGURE 4.29

2. Next, use the data in columns Susceptibles, Infectives, and Removals to form a graph as in Figure 4.30.

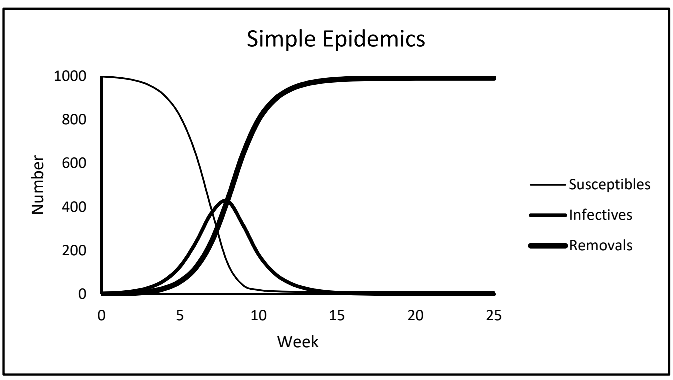


FIGURE 4.30

Notice that the graph shows that the worst of the epidemic will occur during week 8 when a total of approximately 427 people will have the flu. Also note that the numerical results show that approximately 9 people will never get the flu. In mathematical notation,

$$\lim_{n\to\infty} S_n \approx 9$$

The fact that this limit is not 0 is typical for an SIR model.

Now let's examine sensitivity to the initial number of cases. It is often the case that at the beginning of an epidemic, the number of initial cases (I_0) is underreported. Using our model we can easily assess this impact.

We will assume that the number of new cases in week 1, 5, is accurate. If the number of initial cases changes, the transmission coefficient will also change. Generalizing equation (4.26), we see that

$$\alpha = \frac{\text{number of new cases in week 1}}{(\text{initial number of cases})(\text{population} - \text{initial number of cases})} \tag{4.27}$$

We can use Formula (4.27) to easily calculate the value of α for different numbers of initial cases. To this end, modify the worksheet as in Figure 4.31.

	B	C
1	**Population**	1000
2	**Transmission Coefficient**	=C4/(C7*B7)
3	**Removal Rate**	0.5
4	**Number of New Cases in Week 1**	5

FIGURE 4.31

Change the number of infectives in week 0 to between 4 and 10 and note how the system changes. Three examples are given in Figure 4.32

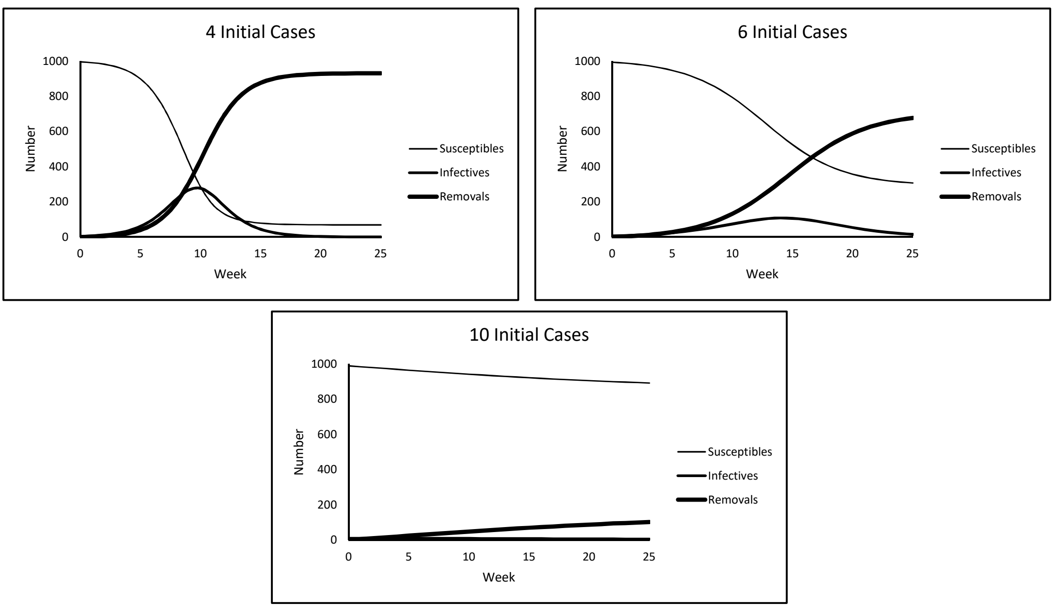


FIGURE 4.32

Notice that with 4 initial cases the value of the transmission coefficient decreases, the worst of the epidemic occurs about 2 weeks later, and the severity decreases. At the peak, only 277 people have the flu and in the long–run, about 69 people never get the flu.

As the number of initial cases increases, the severity decreases. Particularly note that with 10 or more initial cases, the number of infectives never increases. In this case we say that there is no epidemic. We see that if the initial number of infectives is higher than originally thought, the epidemic may not be as bad as originally thought.

Exercises

4.6.1 In this section we argued that the initial number of cases may be under-reported. If this were the case and the number of new cases in week 1 was constant, then the model predicts that the epidemic may not be as bad as originally thought. One could argue that the number of new cases in week 1 could also be under-reported. Use the model to investigate the severity of the epidemic if the number of new cases in week 1 is higher than 5 but the initial number of cases is 3. What happens if both numbers are higher than originally suspected?

4.6.2 In the original model we assumed that the population is constant. Now let's relax this assumption. Suppose that 25 people move into the community each week starting in week 1. Assume that each of these people is susceptible.

a. Modify the worksheet to incorporate this influx of people and describe what happens to the spread of the flu over a 100 week period (use $\alpha = 0.00167$, $\gamma = 0.5$ and $I_0 = 3$).

b. What happens if 100 new people move in each week?

4.6.3 Suppose the population is a constant 1,000, that initially 50 people have the flu ($I_0 = 50$), $\alpha = 0.00167$, and that $\gamma = 0.5$. To try to decrease the severity of the epidemic, the community quarantines some of those with the flu. One way to model quarantining is to simply modify the transmission coefficient. For instance, if 25% of those with the flu are quarantined, then only 75% of the interactions between those infected and those susceptible are capable of producing an infected person. Therefore, the number of newly infected people is given by

$$(0.75\,\alpha)\,S_n I_n$$

Thus the new "effective" transmission coefficient is $(0.75\,\alpha)$.

a. Add a cell to the worksheet "**Epidemics**" to hold the new parameter "Proportion Quarantined," set it equal to 0.25, and modify the model to incorporate this parameter. Describe what effect quarantining 25% has over not quarantining.

b. Add a scroll bar that allows the user to vary the proportion quarantined between 0 and 1. Describe what happens as the proportion quarantined changes. What happens if it equals 0? What if it equals 1?

c. Add a scroll bar that allows the user to vary the number of initial cases, I_0, between 0 and 1000.

d. Find a value of Proportion Quarantined that prevents an epidemic from occurring (i.e. the number of infectives never increases) regardless of the value of I_0. Estimate the smallest such value of Proportion Quarantined.

4.6.4 Consider a disease such as the common cold where a person is *not* immune once they are "healed." Once healed, a person becomes susceptible again. Such a disease could be modeled with an SIS model as illustrated in Figure 4.33.

a. Devise a set of equations for an SIS model. (**Hint:** There is no Removed category. Infected people are healed and are immediately added to the susceptible category.)

b. Implement your model in an Excel worksheet to describe the spread of the common cold through a population of 1,000 where initially 4 people have the cold and assuming that the cold lasts an average of 2 weeks (use $\alpha = 0.00167$). What do you observe?

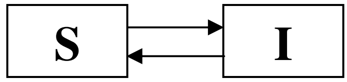


FIGURE 4.33

Project Ideas

1. Try to use a simple epidemic model to model the spread of diabetes in the US over time.

2. Use an epidemic model to model a hypothetical invasion of zombies.

3. Investigate different ways of modeling vaccinations in an epidemic model.

4. Model an epidemic involving both quarantining and vaccinations.

5. Use an epidemic model to model the infestation of lodgepole pine trees by the mountain pine beetle.

6. Create a worksheet to model the balance of a savings account over time. The worksheet should allow the user to input parameters such as interest rate, monthly income, percentage of monthly income for savings, etc.

7. Research and model the battle of Trafalgar.

8. Try to model the US population over time with a discrete logistic equation.

9. Research how eigenvalues and eigenvectors can be used to find analytic solutions of two-dimensional discrete dynamical systems.

10. Model the spread of the black plague in 14^{th} century London.

11. Suppose you just graduated college and have student loans to pay off. But you also think you need a new car. Should you A. hold on to your old junker car for as long as possible and pay its repair bills, or B. buy a new car now, pay fewer repair bills, and pay more toward your student loans? Create a worksheet to analyze this question.

12. Suppose you are a student just graduating from college, about to head into the work force, and you have a few student loans you will have to start thinking about paying off. Although you are just graduating, you are already thinking about saving for a house. Now you want to know, depending on the interest rates, would it be better to pay off all of your loans as fast as you can, or begin investing and saving for a house while making smaller payments on your student loans? Create a worksheet to analyze this question

For Further Reading

- For an extremely well–written treatise on discrete dynamical systems, see Sandefur, James T, *Discrete Dynamical Systems Theory and Applications*, Clarendon Press, 1990.

- For additional examples of discrete dynamical systems, see Meerschaert, Mark M., *Mathematical Modeling*, Second ed., Academic Press, 1999, pg. 141 – 152. Also see the bibliography given on page 152.
- For examples of discrete dyanmical systems applied to compound interest and mortgage payments, see Tung, Ka–Kit, *Topics in Mathematical Modeling*, Princeton University Press, 2007, pg. 54 – 67.
- For additional information on epidemic models, see Allman, Elizabeth Spencer and John A. Rhodes, *Mathematical Models in Biology: An Introduction*, Cambridge University Press, 2004.
- For information on analyzing discrete dynamical systems with eigenvalue and eigenvector methods, see Lay, David C., *Linear Algebra and its Applications*, Third ed., Addison Wesley, 2003, pg. 342 – 353.
- For a much different approach to modeling dynamical systems, see Hannon, Bruce and Matthias Ruth, *Dynamic Modeling*, Springer, 2001.

5

Differential Equations

Chapter Objectives

- Use Euler's Method to find numerical solutions to differential equations
- Numerically solve systems of differential equations
- Model mixing problems, population growth, and military combat with differential equations
- Introduce Runge-Kutta methods for numerically solving differential equations

5.1 Introduction

Often times it is very easy to describe how fast a quantity changes. For instance, in the bacteria population example in Chapter 4, we observed from the data that the rate at which the culture grows decreases as the population nears 621. This observation led to the difference equation

$$\Delta a_n = a_{n+1} - a_n = b(621 - a_n)\, a_n \tag{5.1}$$

Note in this example we worked with discrete increments of time. In reality, time is continuous so using discrete time units is a simplification. It is a convenient simplification because a difference equation such as (5.1) is very easy to solve for a_{n+1} in terms of a_n giving a recursive solution.

When measuring time continuously, we describe change with a *differential equation*. Differential equations are formed in the same basic way as difference equations, but finding their solutions can be much more complicated.

To illustrate how differential equations are formed, consider the following observation:

> When a hot cup of coffee is set on a desk, it initially cools very quickly. As the coffee gets closer to room temperature, it cools less quickly.

This simple observation is an example of Newton's Law of Cooling:

> The rate at which a hot object cools (or a cold object warms) is proportional to the difference between the temperature of the object and the temperature of its surrounding medium.

This law can be translated into the following differential equation:

$$\frac{dy}{dt} = k\,(y - T)$$

where

$$\begin{aligned} y(t) &= \text{ temperature of the object a time } t \\ T &= \text{ temperature of the medium (assumed to be constant)} \\ k &= \text{ constant of proportionality} \end{aligned}$$

This differential equation can be solved using basic techniques yielding the general solution:

$$y(t) = T + Ce^{kt}$$

where C is an arbitrary constant.

Example 5.1.1 (Newton's Law of Cooling)
Consider a cup of coffee that is initially 100 °F, cools to 90 °F in 10 minutes, and sits in a room whose temperature is a constant $T = 60\,^{\circ}\mathrm{F}$.

The general solution to Newton's Law of Cooling is $y(t) = T + Ce^{kt}$. To find the specific solution in this case we need to find the values of the constants C and k. The initial condition $y(0) = 100$ gives

$$100 = 60 + Ce^{k(0)} \quad \Rightarrow \quad C = 40$$

The condition $y(10) = 90$ gives

$$90 = 60 + 40e^{k(10)} \quad \Rightarrow \quad k \approx -0.02877$$

Thus the model is:

$$y(t) = 60 + 40e^{-0.02877t}. \tag{5.2}$$

A graph of this model is shown in Figure 5.1. This curve is called the *solution curve*.

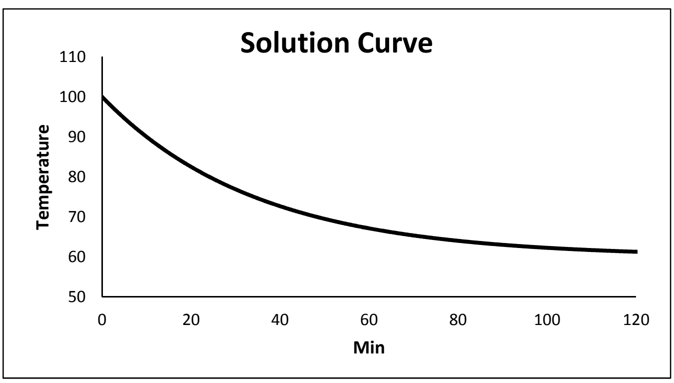


FIGURE 5.1

□

In this chapter, we do not analytically solve differential equations as done in Example 5.1.1. Instead we use a technique called *Euler's Method* to numerically approximate solution curves and then graphically analyze the results.

5.2 Euler's Method

Euler's method is a technique for approximating points on the solution curve of a differential equation. To illustrate the method, consider a differential equation of the form

$$\frac{dy}{dt} = F(y)$$

along with the initial condition $y(t_0) = y_0$ where t_0 are y_0 are some given values. As shown in Figure 5.2, the point (t_0, y_0) is a point on the solution curve. Now, let h be some small positive quantity and define time t_1 to be $t_1 = t_0 + h$. Our goal is to approximate the y-coordinate of the point $(t_1, y(t_1))$ on the solution curve.

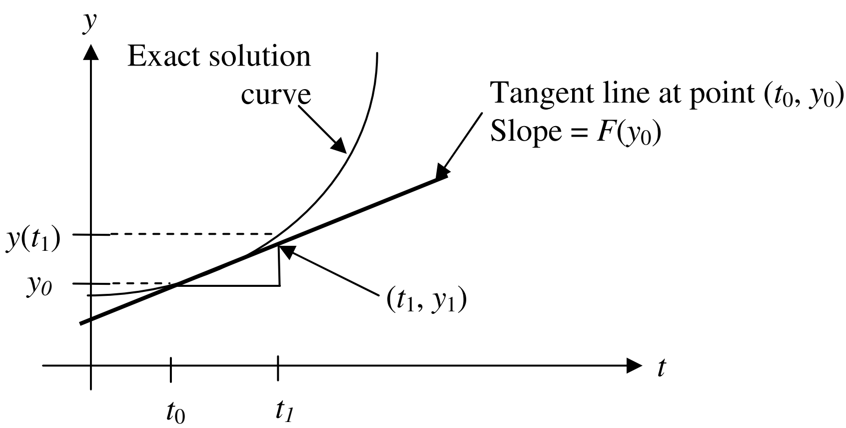


FIGURE 5.2

In the triangle in Figure 5.2, the base has length h and the hypotenuse is on a line with slope $F(y_0)$. Therefore, the height is

$$\text{height} = h\,F(y_0)$$

The y-coordinate of the base of the triangle is y_0. Thus the y-coordinate of the top of the triangle is

$$y_1 = y_0 + h\,F(y_0) \tag{5.3}$$

This y-coordinate is an approximation of $y(t_1)$. To approximate $y(t_2)$ where $t_2 = t_1 + h$, we can repeat this process, replacing y_0 with y_1. We continue to repeat this process as follows:

$$\begin{array}{cc}
t_1 = t_0 + h & y_1 = y_0 + h\,F(y_0) \\
t_2 = t_1 + h & y_2 = y_1 + h\,F(y_1) \\
\vdots & \vdots \\
t_{n+1} = t_n + h & y_{n+1} = y_n + h\,F(y_n)\,.
\end{array}$$

This algorithm is called *Euler's method*. The results from Euler's method can be interpreted in at least two ways:

a. **Numerically**: For each n, $y_n \approx y(t_n)$.

b. **Graphically**: Each point (t_n, y_n) is approximately a point on the solution curve.

Example 5.2.1 (Applying Euler's Method)
Euler's method is easy to implement in Excel. Here we apply it to the Newton's law of cooling problem in Example 5.1.1 and examine how the value of h affects the approximation. Rename a blank worksheet "**Euler**" and format it as in Figure 5.3. Copy row 5 down to row 1004 to calculate values at 1000 different time values.

	A	B
1	**h =**	0.5
2		
3	**Time**	**Approximate**
4	0	100
5	=A4+$B$1	=B4+$B$1*(-0.02877*(B4-60))

FIGURE 5.3

Numerically, these results tell us the approximate temperatures at different times. For instance, row 24 tells us that at time 10 min, the temperature is approximately 89.94 °F. Graphically, we can use the results to approximate the solution curve. Use columns **Time** and **Approximate** to create a graph as in Figure 5.4.

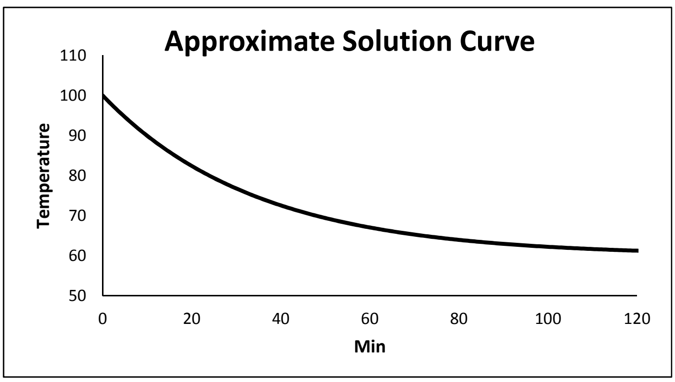


FIGURE 5.4

Observe that the approximate solution curve in Figure 5.4 looks very similar to the exact solution curve in Figure 5.2. To numerically compare these curves, add the formulas in Figure 5.5. The formulas in column **Exact** come from Equation (5.2) and yield the exact temperatures. The error is simply the difference between the approximate temperatures and the exact temperatures.

	C	D
3	**Exact**	**Error**
4	100	=C4-B4
5	=60+40*EXP(-0.02877*A5)	=C5-B5

FIGURE 5.5

To analyze the errors, use the columns **Time** and **Error** to create a graph as in Figure 5.6. Set the y-axis **min** and **max** to 0 and 0.25, respectively, and the x-axis **min** and **max** to 0 and 200, respectively.

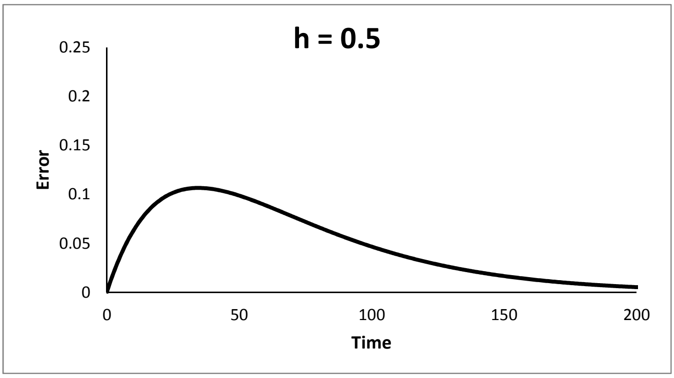


FIGURE 5.6

Figure 5.6 shows that errors are always less than about 0.12 and that as time increases, the errors get smaller. This shows that Euler's method gives very accurate results in this example.

To analyze how the value of h affects the errors, add a scroll bar with **min** and **max** of 0 and 1000, and set the **linked cell** to **E1**. Add the formula in Figure 5.7 to vary the value of h between 0 and 1.

	B
1	=E1/1000

FIGURE 5.7

Use the scroll bar to vary the value of h and observe what happens to the error. Two examples are shown in Figure 5.8. Note that as h gets smaller, the error decreases; as it gets larger, the error increases. However, note that the error is never greater than 0.25, which is very small relative to the exact values.

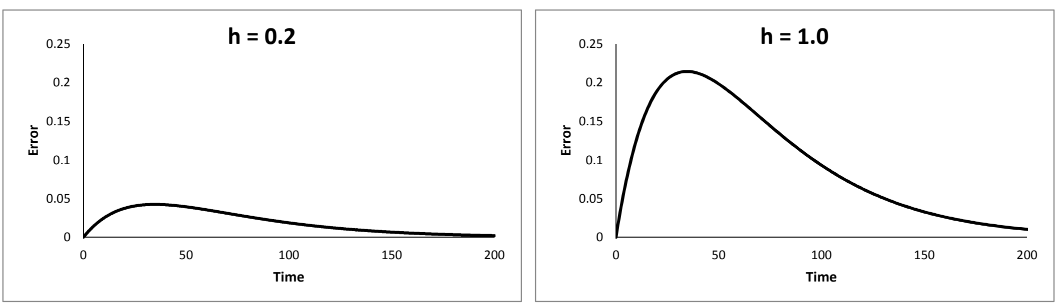


FIGURE 5.8

□

Example 5.2.1 illustrates that in general, the smaller h is, the better the approximation. But there is a trade-off. If we use a smaller value of h, we must do more iterations to graph the solution curve over the same interval of t-values. This begs two questions:

1. How do we choose the value of h?
2. How do we know if the approximate solution curve is accurate?

There is no perfect answer, but we do offer the following guideline:

> Cut the value of h in half. If the approximate solution curve does not change "significantly," then the approximation is probably accurate.

Example 5.2.2 (Logistic Equation)
Suppose that 25 panthers are released into a game preserve. Initially the population grows at a rate of approximately 25% per year, but because of limited food supplies, the preserve is believed to support only 200 panthers. We want to model the population over time.

Note that the information given deals with the rate of change. This suggests we create a differential equation to model the rate of change of the population. If $y(t)$ represents the population at year t, the initial rate of 25% suggests that we model

$$\frac{dy}{dt} = 0.25y$$

However, this model does not take into account the fact that the preserve can support only 200 panthers. It seems reasonable to assume that the rate of growth will decrease as y approaches 200. One way to model this is

$$\frac{dy}{dt} = 0.25\left(1 - \frac{y}{200}\right)y \tag{5.4}$$

Note that as $y \to 200$, $\left(1 - \frac{y}{200}\right) \to 0$ meaning that $\frac{dy}{dt} \to 0$. Equation (5.4) is called a *logistic differential equation*. Also note that this logistic differential equation is very similar to the logistic difference equation we derived for the bacteria population in Chapter 4. The general form of a logistic equation is

$$\frac{dy}{dt} = k\left(1 - \frac{y}{L}\right)y$$

The parameter L is called the *carrying capacity* and the parameter k is called the *unconstrained* (or *intrinsic*) *growth rate*.

To approximate the solution curve of Equation (5.4), rename a blank worksheet "**Logistic**" and format it as in Figure 5.9. Copy row 5 down to row 129 to model 25 years.

	A	B
1	**h =**	0.2
2		
3	**Year**	**Population**
4	0	25
5	=A4+$B$1	=B4+$B$1*(0.25*(1-B4/200)*B4)

FIGURE 5.9

Next, create a graph as in Figure 5.10.

Figure 5.10 shows that the rate of growth slows down as the population approaches 200, as expected. The population reaches the carrying capacity by year 25. Also note that this graph looks very similar to the graph of the bacteria population in Example 4.3.1.

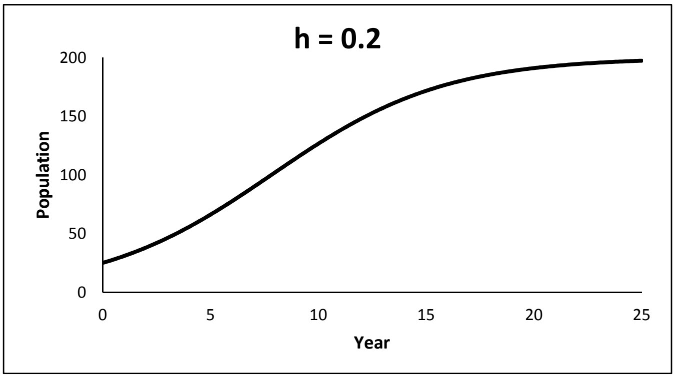


FIGURE 5.10

To get an idea of whether the approximate solution curve in Figure 5.10 is accurate or not, we follow the guideline given above. We change the value of h in cell **B2** to 0.1, copy row 5 down to row 254, and create a graph as in Figure 5.11.

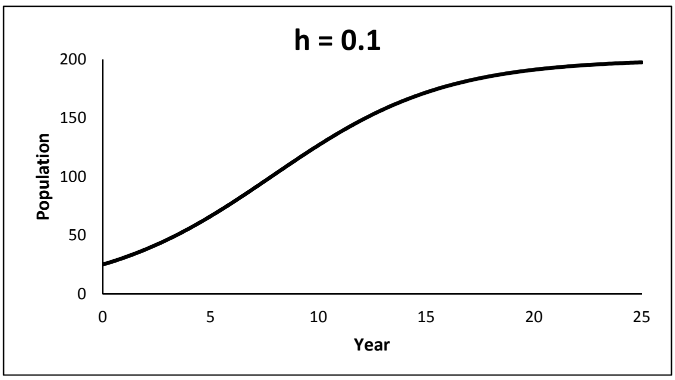


FIGURE 5.11

Observe that Figures 5.10 and 5.11 look almost identical. Thus $h = 0.2$ appears to give an accurate result. If the two figures were much different, then we would need to keep cutting h in half until the curves do not change much. □

Non-autonomous differential equations (meaning equations where the right-hand side explicitly depends on t) of the form

$$\frac{dy}{dt} = F(t, y)$$

arise frequently in applications. Euler's method can be easily adapted to these types of differential equations. The basic algorithm is given by

$$t_{n+1} = t_n + h, \quad y_{n+1} = y_n + hF(y_n, t_n).$$

The next example illustrates an application of a non-autonomous differential equation.

Example 5.2.3 (Bacteria Growth)
Let $y(t)$ denote the population of bacteria in a Petri dish t days after the bacteria begin growing. Suppose $y(t)$ is described by the differential equation

$$\frac{dy}{dt} = 150\sqrt{t}$$

for t between 0 and 10. If the initial population is 500, approximate the solution curve over the interval $0 \leq t \leq 10$ and approximate the population at time $t = 7$.

Rename a blank worksheet "**Bacteria**" and format it as in Figure 5.12. Copy row 5 down to row 104.

	A	B
1	**h =**	0.1
2		
3	**Day**	**Population**
4	0	500
5	=A4+$B$1	=B4+$B$1*150*SQRT(A4)

FIGURE 5.12

Create a graph of the solution curve as in Figure 5.13. Note that as time increases, the population grows faster.

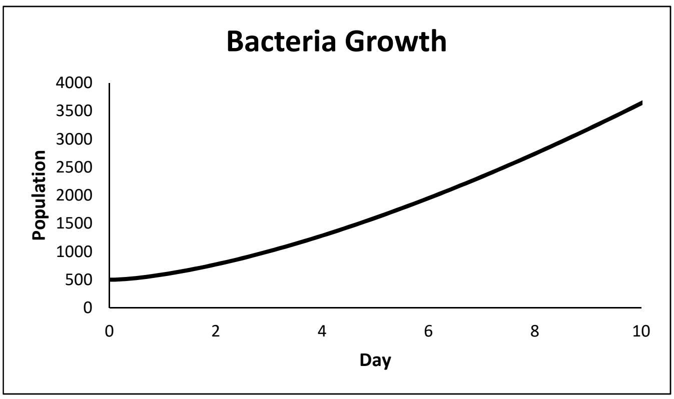


FIGURE 5.13

To determine if this approximate solution curve is accurate, we change the value of h in cell **B1** to 0.05, copy row 5 down to row 204, and graph the resulting approximate solution curve. Observe that this curve looks very similar to that in Figure 5.13. This indicates that $h = 0.1$ yields accurate results.

Now note that for $h = 0.1$, the calculations give $y(7) \approx 2331$. We interpret this result by saying that at the beginning of day 7, there will be approximately 2300 bacteria. $\square$

Exercises

Directions: Use Euler's method to approximate the solution curve of the differential equation in each problem. Choose a value of h and justify that this value gives accurate results using the guideline given in the text.

5.2.1 Consider a cylindrical can with radius r filled with water which drains out through a small hole in the bottom of the can. Let $h(t)$ denote the height of the water in the can (in inches) at time t (in seconds). Using a principle called *Torricelli's law*, it can be shown that $h(t)$ is described by the differential equation

$$\pi r^2 \frac{dh}{dt} = -k\sqrt{h}$$

where $k > 0$ is a constant. If $r = 1$, $k = 0.372$, and the initial height is 8 inches, graph the approximate solution curve. At approximately what time is the can empty?

5.2.2 Suppose a certain population at time t, $y(t)$, is described by the differential equation

$$\frac{dy}{dt} = 0.00008(300 - y)(y - 200)^2.$$

Graph an approximate solution curve over the interval $0 \leq t \leq 20$, use a scroll bar to vary the initial population between 100 and 300, and describe the long-term behavior of the population for different initial populations.

5.2.3 Suppose the mass of a yeast culture is described by a logistic equation with a carrying capacity of 8 grams. At time $t = 0$, the culture weighs 0.5 grams. Two hours later, it weighs 2 grams.

a. Graph an approximate solution curve over the interval $0 \leq t \leq 10$. Use a scroll bar to vary the value of k.

b. Use the scroll bar to approximate the value of k so that the condition $y\,(2) = 2$ is satisfied.

c. At what time is the weight increasing most rapidly? Support your answer numerically.

5.2.4 Let $y(t)$ denote the population of rabbits (in thousands) in a certain forest at time t (in months). Suppose y is described by the differential equation

$$\frac{dy}{dt} = 1 + 3\cos\left(5\sqrt{t} - 9\right).$$

a. Graph an approximate solution curve over the interval $0 \leq t \leq 10$ if the initial population is 3000.

b. Describe, in words, the behavior of the population over this interval of time.

c. What is the approximate population at time $t = 5$?

5.2.5 In Example 5.2.1 we considered a Newton's law of cooling problem where the room temperature is a constant $T = 60\,^{\circ}\mathrm{F}$. Now suppose the room is warming up and that the room temperature is described by the function $T(t) = (60 + 0.05t)\,^{\circ}\mathrm{F}$. This means that the differential equation describing the temperature of the cup of coffee is

$$\frac{dy}{dt} = -0.02877\left[y - (60 + 0.05t)\right].$$

a. Graph an approximate solution curve over the interval $0 \leq t \leq 120$ if the initial temperature of the cup of coffee is $100\,^{\circ}\mathrm{F}$.

b. Graph the room temperature on the same plane as the solution curve and compare the two graphs. At approximately what time does the coffee temperature equal the room temperature?

5.2.6 Let $T(t)$ denote the temperature at time t hours from some starting point inside a building that is heated with a furnace. A simple model for T is

$$\frac{dT}{dt} = -k\left(T - M(t)\right) + k_U\left(T_D - T\right)$$

where

- $k > 0$ is a parameter measuring the level of insulation of the building (the more insulation, the lower the value of k),
- $M(t)$ is the temperature outside the building at time t,
- $k_U > 0$ is a parameter measuring the power of the furnace (the more powerful the furnace is, the higher the value of k_U), and
- T_D is the "desired" temperature inside the building.

a. According to this model, is the furnace heating the building faster when T is "close" to T_D or when T is "far" from T_D? Explain why.

b. Let $k = 0.25$, $M(t) = 8°\text{C}$, $k_U = 3$, and $T_D = 21$. If $T(0) = 12$, Use Euler's method to estimate the time at which the temperature reaches 18°C. Does the temperature ever reach the desired temperature of 21?

c. Consider the same building as in part a., but suppose the outside temperature is described by $M(t) = 12 + 4\cos(\pi t/12)$. Use Euler's method to estimate the inside temperature at time t. Graph both the inside and outside temperatures over the time interval [0, 72]. Do both temperatures reach their minimum and maximum values at the same time?

5.2.7 In Chapter 2, we modeled the velocity of a free-falling object with air resistance by assuming that the force due to air resistance is proportional to the velocity ($F_A = k\,v$). This yielded the differential equation

$$\frac{dv}{dt} + \frac{k}{m}v = g$$

where $g = 9.8\text{m/sec}^2$ and $m =$ mass of the object.

a. Suppose $m = 10\,\text{g}$, $v\,(0) = 0$, and $k = 3$. Use Euler's method to approximate $\lim_{t\to\infty} v\,(t)$.

b. Use a scroll bar to examine what happens to $\lim_{t\to\infty} v\,(t)$ as k changes. Do your results make sense? Why or why not?

c. Add a scroll bar to vary the value of m. What happens to $\lim_{t\to\infty} v\,(t)$ as m changes? Does this make sense?

5.2.8 Consider again the model of the velocity of a free-falling object with air resistance from Chapter 2. Now assume that the force due to air resistance is described by $F_A = (3/n)\,v^n$ where $0 < n < 5$. This assumption yields the differential equation

$$\frac{dv}{dt} + \frac{3}{m \cdot n}v^n = g$$

where $g = 9.8\text{m/sec}^2$ and $m =$ mass of the object.

a. Suppose $m = 10\,\text{g}$ and $v\,(0) = 0$. Use Euler's method to approximate $\lim_{t\to\infty} v\,(t)$.

b. Use a scroll bar to examine what happens to $\lim_{t\to\infty} v(t)$ for different values of n.

c. Compare the terminal velocity for $n > 1$ to that for $n = 1$. Also compare the amount of time needed to reach the terminal velocity. Repeat for $n < 1$.

d. If we wanted to model the free-fall of an object that is very aerodynamic, would we want to use a small value of n or a large value? Explain.

5.2.9 When two liquids of different temperatures are mixed together, the final temperature of the resulting mixture can be calculated with the following formula:

$$t_f = \frac{m_1 c_1 t_1 + m_2 c_2 t_2}{m_1 c_1 + m_2 c_2}$$

where m_1 and m_2 are the masses (in g), c_1 and c_2 are the specific heats (in J/g°C), and t_1 and t_2 are the initial temperatures (in °C) of the respective liquids. Suppose Sam pours some boiling water (100°C) over some instant coffee granules in a mug and then lets the coffee sit for 10 min to cool. Sam also wants to add some cold milk (5°C) to the coffee before he drinks it. He considers two options for when to add the milk:

i. At time $t = 0$ min, immediately after the water is poured into the mug.

ii. At time $t = 10$ min, immediately before he begins drinking the coffee.

Sam wants to know how each option would affect the final temperature of the coffee/milk mixture. Ignore the effects of the mug and the granules on the temperature. Assume the room temperature is 20°C, the constant of proportionality in Newton's law of cooling is $k = -0.06$, and the parameters are $m_1 = 240$, $m_2 = 5.15$, $c_1 = 4.2$, $c_2 = 3.77$ (liquid 1 is the coffee and liquid 2 is the milk).

a. Determine the temperature of the coffee/milk mixture at time $t = 10$ for each option.

b. Is there much of a difference between the two options?

c. Investigate the sensitivity of the system to the parameter m_2. Describe your observations.

5.3 Mixing Problems

A derivative dy/dt describes the rate of change of the quantity $y(t)$. In some situations a quantity is being increased by one factor and decreased by another. In such situations we can describe the overall rate of change using the following principle:

$$\text{Overall rate of change} = \text{Rate of increase} - \text{Rate of decrease}.$$

As an application of this principle, consider a tank that contains 50 gallons of a solution composed of 90% water and 10% alcohol. A second solution containing 50% alcohol is added to the tank at the rate of 2 gallons per minute. At the same time, solution is being drained from the tank at the rate of 3 gallons per minute. This situation is illustrated in Figure 5.14. Assuming the tank is continuously stirred, find the amount of pure alcohol in the tank from 0 to 50 min. Also calculate the concentration of alcohol in the tank at each time.

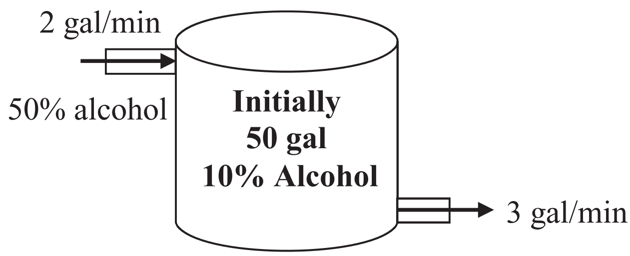


FIGURE 5.14

Let $y(t)$ denote the gallons of pure alcohol in the tank at time t min. The solution being added to the tank increases the value of y and the solution being drained out of the tank decreases the value of y. The principle from above suggests the following general differential equation model:

$$\begin{aligned}\frac{dy}{dt} &= \text{Rate of change of alcohol (in gal of alcohol/min)}\\ &= \text{Rate in} - \text{Rate out}\end{aligned}$$

Now we need to model the rate in and the rate out. Modeling the rate in is relatively easy. A 50% alcohol solution is being added to the tank at a constant 2 gallons per minute. Therefore,

$$\text{Rate in} = \left(\frac{0.5 \text{ gal alcohol}}{1 \text{ gal solution}}\right)\left(\frac{2 \text{ gal solution}}{\text{min}}\right) = \frac{1 \text{ gal alcohol}}{\text{min}}$$

Modeling the rate out is a little more complicated because the proportion of the solution in the tank that is alcohol is changing. Note that the volume of solution in the tank is decreasing at the rate of 1 gal/min. Therefore, the volume of solution in the tank at time t is simply $50 - t$. The amount of alcohol in the tank at time t is simply y. Therefore, similar to the model for rate in, we have

$$\text{Rate out} = \left(\frac{y \text{ gal alcohol}}{(50-t) \text{ gal solution}}\right)\left(\frac{3 \text{ gal solution}}{\text{min}}\right) = \frac{3y}{50-t}\,\frac{\text{gal alcohol}}{\text{min}}$$

Thus, the differential equation is

$$\frac{dy}{dt} = 1 - \frac{3y}{50-t}. \tag{5.5}$$

In general, the differential equation describing $y(t)$ has the form

$$\frac{dy}{dt} = \begin{pmatrix}\text{Rate of sol}\\ \text{flowing in}\end{pmatrix}\begin{pmatrix}\text{Conc of sol}\\ \text{flowing in}\end{pmatrix} - \begin{pmatrix}\text{Rate of sol}\\ \text{flowing out}\end{pmatrix}\left(\frac{y}{\text{Vol of sol in tank}}\right)$$

The concentration of the solution in the tank is then

$$\text{Concentration} = \frac{y}{\text{Vol of sol in tank}}.$$

Example 5.3.1 (Solving a Mixing Problem)
We can easily use ideas from Section 5.2 to approximate a solution curve of (5.5). Rename a blank worksheet "**Mixture**" and format it as in Figure 5.15. Copy row 5 down to row 504.

	A	B	C	D
1	**h =**	0.1		
2				
3	**Time**	**Vol Solution in Tank**	**Vol Alcohol**	**Concentration**
4	0	50	=5	=C4/B4
5	=A4+$B$1	=50-A5	=C4+$B$1*(1-3*C4/B4)	=C5/B5

FIGURE 5.15

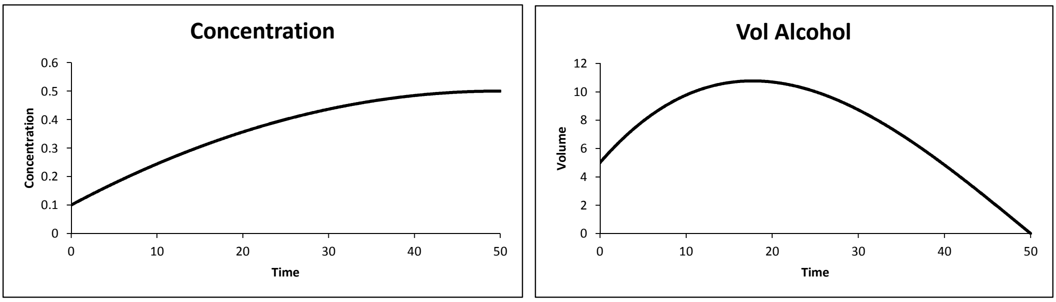


FIGURE 5.16

Create a graph as in Figure 5.16. To determine whether these approximate solution curves are accurate we change the value of h from 0.1 to 0.05 and observe that the curves do not change significantly. Thus $h = 0.1$ appears to give accurate results.

From Figure 5.16 we see that the amount of alcohol initially increases and then decreases to 0 as time approaches 50. The concentration increases from 0.1 and starts to level off near 0.5. We can also use the results to approximate, for instance, the time at which the concentration in the tank is 0.4. Examining the numerical results we see the concentration is approximately 0.4 at time 25 min. □

Example 5.3.2 (Tank with Valve)
Using Euler's method, we can easily solve mixing problems much more complicated than in Example 5.3.1. Consider the following scenario as illustrated in Figure 5.17: A solution of 50% alcohol and 50% water is poured at a constant rate of 1 L/min into a tank initially containing 5 L of pure water. The tank is well-stirred and the solution flowing out of the tank is controlled by a valve which is open for 1 min, closed for 1 min, open for 1 min, and so on. When the valve is open, solution flows out at a rate of 2 L/min. Approximate the time at which the concentration of alcohol in the tank is 0.4.

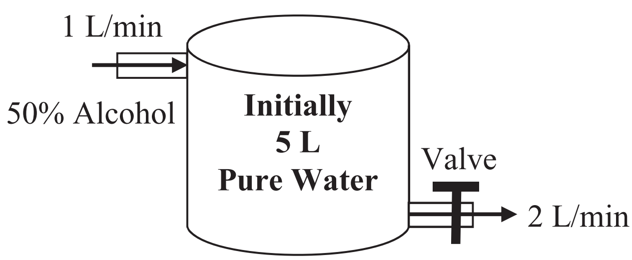


FIGURE 5.17

To model this scenario, note that the volume of solution in the tank is changing. When the valve is open, the volume decreases by 1 L/min. When the valve is closed, the volume increases by 1 L/min. Thus over a time of h min,

$$\text{Change in volume} = \begin{cases} 1 \cdot h \text{ L} & \text{if the valve is closed} \\ -1 \cdot h \text{ L} & \text{if the valve is open.} \end{cases}$$

Let $y(t)$ denote the volume of pure alcohol in the tank at time t. The differential equation describing y is

$$\frac{dy}{dt} = 1 \cdot 0.5 - \begin{pmatrix} 0 \text{ if the valve is closed} \\ 2 \text{ if the valve is open} \end{pmatrix} \left(\frac{y}{\text{Vol of sol in tank}} \right).$$

To approximate a solution curve to this differential equation, title a blank worksheet "**Tank with Valve**" and format it as in Figure 5.18. Copy row 5 down to row 254 and then create graphs as in Figure 5.19. To verify that these graphs are accurate, we change h from 0.2 to 0.1 and observe that the graphs do not change significantly.

	A	B	C	D	E
1	**h =**	0.2			
2					
3	**Time**	**Valve Open?**	**Vol Solution in Tank**	**Vol Alcohol**	**Concentration**
4	0	=IF(MOD(INT(A4),2)=0,1,0)	5	0	=D4/C4
5	=ROUND(A4+$B$1,1)	=IF(MOD(INT(A5),2)=0,1,0)	=IF(B4=1,C4-$B$1,C4+$B$1)	=D4+$B$1*(0.5-2*D4/C4*B4)	=D5/C5

FIGURE 5.18

The formulas in column **B** round the time down to the nearest whole number, calculate this rounded time modulus 2 and return a 1 if this result is 0, meaning the valve is open, and return 0 otherwise. The **ROUND** function in column **A** rounds off the time to one decimal point. This is necessary because when Excel adds a number repeatedly (as in adding h to the time in each step) it introduces a hidden "rounding" error. Thus when the time is supposed to be a whole number, Excel thinks it is not a whole number. The **ROUND** function corrects this error.

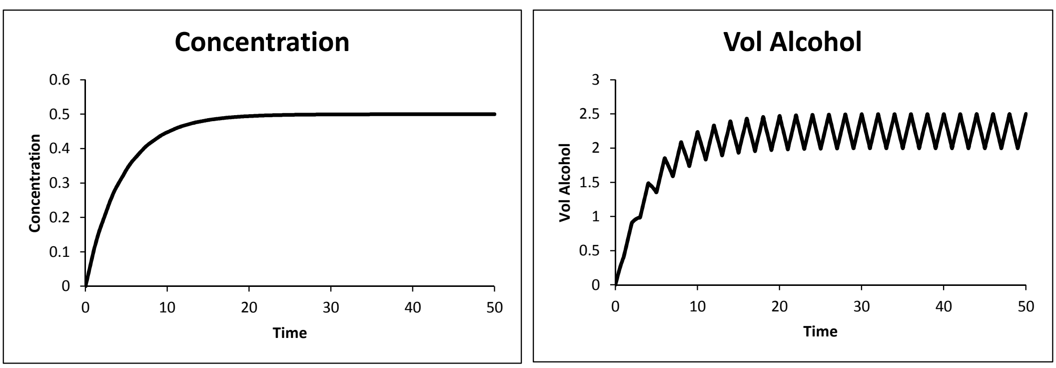


FIGURE 5.19

Notice that the volume of alcohol in the tank initially increases but then begins to fluctuate. The concentration initially increases and then levels out near 0.5. This is what we expect because the concentration of solution coming into the tank is 0.5. From the graph and the numerical results, we see that the concentration is 0.4 at approximately time 7 min. □

Exercises

Directions: Use Euler's method to solve the following problems. Choose a value of h and justify that this value gives accurate results using the guideline given in the text.

5.3.1 A 100-gal tank is full of a solution that contains 15 lb of salt. Starting at time $t = 0$, distilled water is poured into the tank at the rate of 5 gal per minute. The solution in the tank is continuously well-stirred and is drained out of the tank at the rate of 5 gal per minute.

a. Let $y(t)$ denote the weight of salt in the tank at time t. Form a differential equation to model dy/dt.

b. Approximate the amount of salt in the tank at time $t = 10$.

c. At approximately what time is the concentration of salt in the tank 0.1 lb/gal?

d. Approximate $\lim_{t\to\infty} y(t)$. Is this what you expect?

5.3.2 A solution containing 0.05 kg of salt per L flows at a constant rate of 6 L/min into a tank that initially holds 50 L of pure water. The tank is well-stirred and the solution flows out of the tank at a rate of 6 L/min.

a. Let $y(t)$ denote the mass of salt in the tank at time t. Form a differential equation to model dy/dt.

b. Approximate the time at which the concentration of salt in the tank is 0.035 kg/L.

c. Now suppose there was initially 0.5 kg of salt in the tank. Find the time at which the concentration is 0.03 kg/L.

5.3.3 A solution containing 0.07 kg of salt per L flows at a constant rate of 5 L/min into a tank that initially holds 15 L of water in which 0.5 kg of salt is dissolved. The tank is well-stirred and the solution flows out of the tank at a rate of 7 L/min.

a. Let $y(t)$ denote the mass of salt in the tank at time t. Form a differential equation to model dy/dt.

b. Approximate the time at which the amount of salt in the tank is maximized.

c. Approximate the time at which the concentration of salt in the tank is 0.06 kg/L.

5.3.4 A tank initially contains 50 L of pure water. A solution of 50% water and 50% alcohol is poured into the tank at a rate described by $r(t) = 1 + \cos(t)$ L/min. The tank is well-stirred and the solution flows out of the tank at a constant rate of 1 L/min.

a. Graph the approximate volume of alcohol in the tank over the interval $0 \leq t \leq 250$. (**Hint**: From time t_n to time $t_{n+1} = t_n + h$, the volume in the tank increases by $h \cdot r(t_n)$ L and decreases by $h \cdot 1$ L.)

b. Approximate the time at which there is 17.5 L of alcohol in the tank.

c. Graph the concentration of the alcohol in the tank over the interval $0 \leq t \leq 250$. What is the limit of the concentration as $t \to \infty$? Does this make sense? Explain why or why not.

5.3.5 A tank initially contains 50 L of a solution composed of 90% water and 10% alcohol. A second solution containing 25% alcohol is added to the tank at a constant rate of 1 L/min. The tank is well-stirred and the solution flowing out of the tank is controlled by a valve which is open for 2 min, closed for 2 min, open for 2 min, and so on. When the valve is open, solution flows out at a rate of 3 L/min.

a. Approximate the first time at which the volume of solution in the tank is 0 (**Hint**: The valve is open if the time modulus 4 is less than 2.)

b. Graph the volume of alcohol in the tank from time 0 to the time found in part a.

c. At approximately what time is the volume of alcohol in the tank maximized?

5.3.6 Let $y(t)$ denote the population (in thousands) of bacteria in a Petri dish at time t (in days). Suppose the population grows at the rate of $3y$. However, suppose that an antibiotic is continuously added to the Petri dish and kills off bacteria at a rate described by $f(t) = 100t$.

a. Form a differential equation to model dy/dt.

b. Graph an approximate solution curve if the initial population is 10,000.

c. At approximately what time does the population die out?

5.3.7 A radioactive isotope R_1 decays into another radioactive isotope R_2 which then decays into stable atoms. Let $y(t)$ denote the mass (in g) of R_2 at time t (in min). Suppose R_1 decays at a rate described by $50e^{-10t}$ g/min and that R_2 decays at a rate described by $2y$ g/min.

a. Form a differential equation to model dy/dt.

b. Graph an approximate solution curve if the initial mass of R_2 is 40 g.

c. At approximately what time is there 10 g of R_2 left?

5.3.8 Consider the heating and cooling of a building as described in Exercise 5.2.6 where $k = 0.25$, $k_U = 3$, $T_D = 21$, $T(0) = 12$, and $M(t) = 12 + 4\cos(\pi t/12)$.

a. Suppose the furnace is turned on between times 0 and 12, then off between 12 and 24, then on between 24 and 36, and so on. Graph an approximate solution curve over the interval $0 \leq t \leq 20$.

b. Now suppose the furnace is turned on all the time. Graph an approximate solution curve on the same plane as part a.

c. Suppose the building is a store that is open only when the furnace is turned on. Based on the results of parts a. and b., would you recommend leaving the furnace on all the time, or shutting it off when the store is closed? Explain.

5.3.9 Consider a reservoir with a volume of 8 billion cubic meters (m^3) and an initial pollution concentration of 0.0025 kg/m^3. There is a daily inflow of 500 million m^3 of water with a pollution concentration of 0.0005 kg/m^3 and an equal daily outflow of the well-mixed water in the reservoir. Approximately how long will it take to reduce the pollution concentration in the reservoir to 0.001 kg/m^3?

5.4 Systems of Differential Equations

A *system of differential equations* is a set of two or more related differential equations involving two or more unknown functions. In this section we restrict ourselves to a set of two equations with the general form

$$\frac{dx}{dt} = F(x, y), \quad \frac{dy}{dt} = G(x, y)$$

along with the initial conditions $x(t_0) = x_0$, $y(t_0) = y_0$. Euler's method for a system such as this is:

$$\begin{array}{rclrclrcl}
t_1 &=& t_0 + h & x_1 &=& x_0 + hF(x_0, y_0) & y_1 &=& y_0 + hG(x_0, y_0) \\
t_2 &=& t_1 + h & x_2 &=& x_1 + hF(x_1, y_1) & y_2 &=& y_1 + hG(x_1, y_1) \\
& \vdots & & & \vdots & & & \vdots & \\
t_{n+1} &=& t_n + h & x_{n+1} &=& x_n + hF(x_n, y_n) & y_{n+1} &=& y_n + hG(x_n, y_n).
\end{array}$$

Example 5.4.1 (Connected Tanks)
Consider the two connected tanks filled with salt water shown in Figure 5.20. Let $x(t)$ and $y(t)$ denote the masses of salt (in kg) in the tanks at time t where $x(0) = 4$ and $y(0) = 2$. We assume perfect mixing in both tanks. The goal of this example is to describe the long-term behavior of x and y.

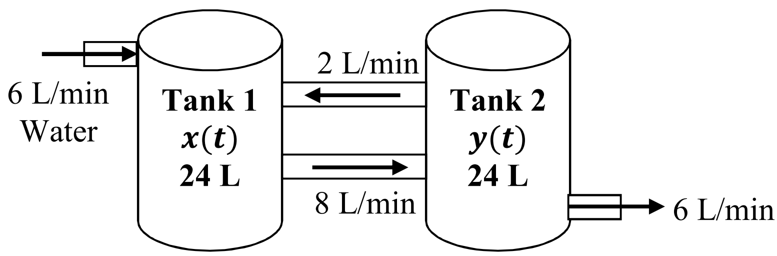


FIGURE 5.20

To set up the system of differential equations, we use the following principle from Section 5.3:

$$\text{Overall rate of change} = \text{Rate in} - \text{Rate out}.$$

First, observe that each tank is losing solution at the overall rate of 8 L/min and gaining solution at the rate of 8 L/min, so the volume of each tank is not changing. Now consider tank 1. This tank has pure water entering on the left at 6 L/min and solution from tank 2 entering on the right at 2 L/min. Therefore,

$$\text{Rate in} = \frac{0 \text{ kg}}{\text{L}} \times \frac{6 \text{ L}}{\text{min}} + \frac{y \text{ kg}}{24 \text{ L}} \times \frac{2 \text{ L}}{\text{min}} = \frac{y \text{ kg}}{12 \text{ min}}.$$

Likewise, tank 1 has solution leaving on the right at the rate of 8 L/min, so

$$\text{Rate out} = \frac{x \text{ kg}}{24 \text{ L}} \times \frac{8 \text{ L}}{\text{min}} = \frac{x \text{ kg}}{3 \text{ min}}.$$

Therefore, the differential equation for tank 1 is

$$\frac{dx}{dt} = \frac{y}{12} - \frac{x}{3}.$$

By a similar argument, the differential equation for tank 2 is

$$\frac{dy}{dt} = \frac{x}{3} - \frac{y}{3}.$$

To numerically solve this system using Euler's method with a step size of $h = 0.2$, rename a blank worksheet "**Connected Tanks**" and format it as in Figure 5.21. Copy row 5 down to row 204.

	A	B	C
1	**h =**	0.2	
2			
3	**t**	**x**	**y**
4	0	4	2
5	=A4+$B$1	=B4+$B$1*(C4/12-B4/3)	=C4+$B$1*(B4/3-C4/3)

FIGURE 5.21

To graphically analyze the results, create graphs of x vs. t and y vs. t as in Figure 5.22. These graphs are called *time plots*. In these graphs, we see that the mass of salt in tank 1 drops to 0 by about time 20 min. The mass of salt in tank 2 initially increases, but then drops to 0 by about time 30 min.

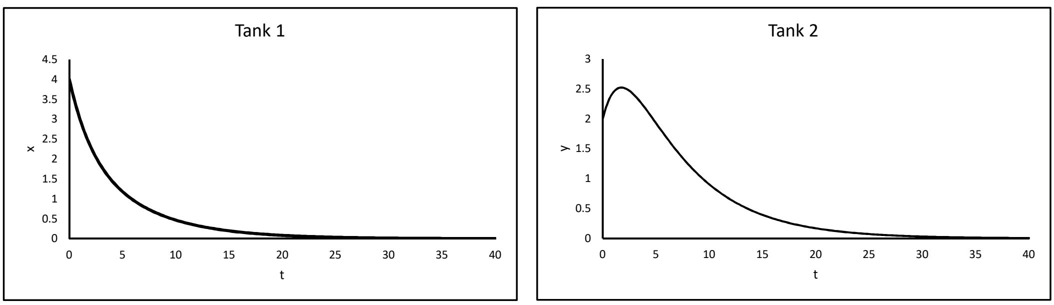


FIGURE 5.22

We can combine the two time plots into a single graph by graphing y vs. x as in Figure 5.23. The $x - y$ plane in this graph is called the *phase plane* and the curve is called a *trajectory*. The trajectory shows that the system starts at the point $(4, 2)$ (the initial condition). Moving along the trajectory to the left, we see that x decreases while y initially increases, but then begins to decrease. Both x and y eventually approach 0. This is exactly what we saw in the time plots.

□

The point $(0, 0)$ on the phase plane in Example 5.4.1 is called an *equilibrium point* of the system. Equilibrium points are central to the analysis of systems of differential equations.

Definition 5.4.1. An *equilibrium point* of a system of differential equations

$$\frac{dx}{dt} = F(x, y), \quad \frac{dy}{dt} = G(x, y)$$

is a point (x_0, y_0) such that $F(x_0, y_0) = 0 = G(x_0, y_0)$.

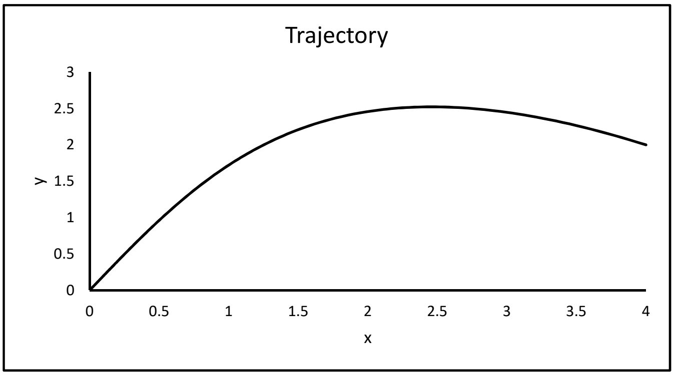


FIGURE 5.23

In simpler terms, an equilibrium point is a point on the phase plane where if we start there, we stay there forever. As with discrete dynamical systems, equilibrium points of systems of differential equations are points on the phase plane which typically attract or repel trajectories. Equilibrium points that attract trajectories are called *attracting*, *stable*, or *asymptotically stable*. Equilibrium points that repel trajectories are called *unstable* or *repelling*.

Example 5.4.2 (Analyzing an Equilibrium Point)
Consider again the system of connected tanks in Example 5.4.1. To confirm that (0, 0) really is an equilibrium point, change $x(0)$ and $y(0)$ to 0 in the worksheet **Connected Tanks** and observe that x and y stay at 0 forever. To determine whether (0, 0) is attracting or not, add the formulas in Figure 5.24 to randomly change the initial conditions to values between -5 and 5.

	B	C
3	**x**	**y**
4	=RANDBETWEEN(-5,5)	=RANDBETWEEN(-5,5)

FIGURE 5.24

On the graph of the trajectory, change the axes mins and maxes to -5 and 5 as in Figure 5.25. Press the **F9** key several times. Each time, a new set of initial conditions is generated. Observe that the trajectory always approaches the point (0, 0). This is graphical evidence that (0, 0) is an attracting equilibrium point.

□

Example 5.4.3 (Finding Equilibrium Points)
Graphically we have determined that (0, 0) is an equilibrium point of the system in Example 5.4.1, but are there possibly other equilibrium points? To algebraically find the equilibrium point(s) of a system of the form

$$\frac{dx}{dt} = F(x, y), \quad \frac{dy}{dt} = G(x, y),$$

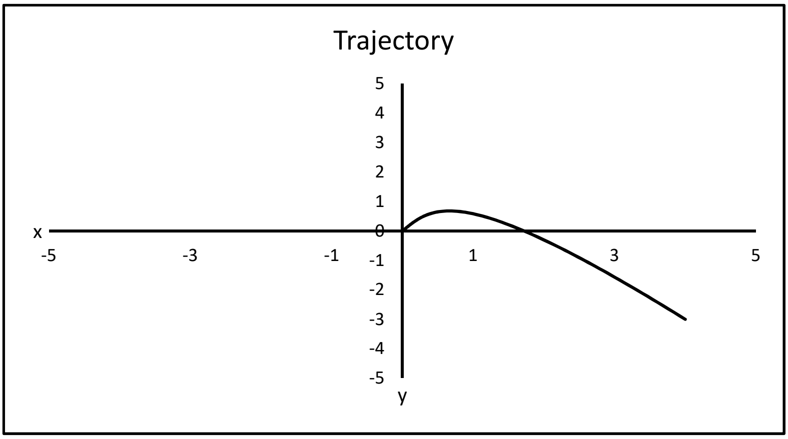


FIGURE 5.25

we need to set both $F(x, y)$ and $G(x, y)$ equal to 0 and solve for x and y. In Example 5.4.1, this yields the system of linear equation

$$\frac{y}{12} - \frac{x}{3} = 0$$
$$\frac{x}{3} - \frac{y}{3} = 0.$$

Solving this system using elementary linear algebra techniques (see Section 3.1) yields the only solution $x = y = 0$. Therefore, $(0, 0)$ is the *only* equilibrium point of the system. □

Exercises

5.4.1 Consider the connected tanks in Figure 5.26.

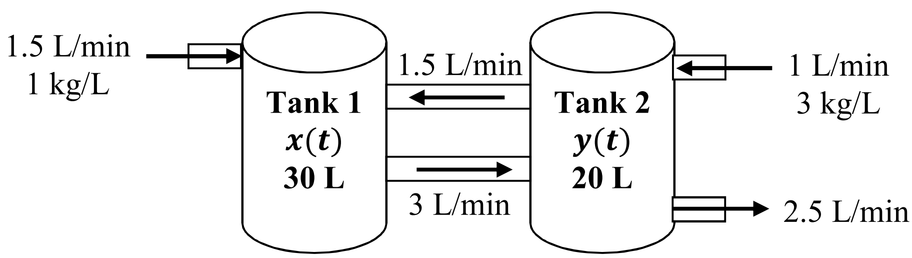


FIGURE 5.26

a. Find a system of differential equations for x and y.

b. Find the equilibrium point(s) of the system.

c. Graphically determine if the equilibrium point(s) are attracting or repelling. Use initial values between 0 and 50.

5.4.2 Consider a 60-L tank filled with water in which 2 kg of salt are dissolved. Fresh water is pumped into the tank at a rate of 3 L/min and the resulting mixture flows out at the rate of 3 L/min into another 60-L tank that initially is filled with pure water. From there, the mixture spills onto the ground at the rate of 3 L/min.

a. Assuming perfect mixing in both tanks, graph the approximate mass of salt in each tank over the interval $0 \leq t \leq 40$.

b. At what time is the mass of salt in the second tank at a maximum?

c. At what time do both tanks contain the same mass of salt?

5.4.3 The motions of a certain pendulum are described by the system of differential equations:

$$\frac{dx}{dt} = y, \ \frac{dy}{dt} = -5\sin x - \frac{9}{13}y$$

where $x = \theta$, the angle between the rod and the downward vertical direction, and $y = \frac{d\theta}{dt}$, the speed at which the angle changes. Find the equilibrium points for this system.

5.4.4 Consider a system of linear differential equations of the form

$$\frac{dx}{dt} = ax + by$$
$$\frac{dy}{dt} = cx + dy.$$

a. Show that the origin (0, 0) is an equilibrium point for any values of a, b, c, and d.

b. For each of the following values of the parameters (a, b, c, d), graphically determine if the origin is an attracting equilibrium point (meaning trajectories always move toward the origin) or repelling (meaning trajectories always move away from the origin).

i.$(-2, -5, 1, 4)$	ii.$(7, -1, 3, 3)$
iii$(-3, -2, -1, -1)$	iv.$(3, 1, -2, 1)$
v.$(-3, -9, 2, 3)$	

5.4.5 Insurgent forces have a strong foothold in the city of Urbania, a major metropolis in the center of the country of Ibestan. Intelligence estimates they currently have a force of 1570 fighters. The local police force has 2250 officers, many of which have had no formal training in law enforcement methods or modern tactics for addressing insurgent activity. Based on data collected over the past year, approximately 8% of insurgents switch sides and join the police each week whereas about 11% of police switch sides and join the insurgents. Intelligence also estimates that around 120 new insurgents arrive from the neighboring country of Moronka each week. Recruiting efforts in Ibestan yield about 85 new police recruits each week as well. In armed conflict with insurgent forces, the local police are able to capture or kill approximately 10% of the insurgent force each week on average while losing about 3% of their force. If $P(t)$ and $I(t)$ denote the number of police and insurgents, respectively, in week t, we can model this scenario with the system of differential equations

$$\frac{dP}{dt} = -0.03P(t) - 0.11P(t) + 0.08I(t) + 85$$
$$\frac{dI}{dt} = 0.11P(t) - 0.10I(t) - 0.08I(t) + 120$$

with the initial conditions $P(0) = 2250$ and $I(0) = 1570$.

a. Use Euler's method to graph the time plots over the interval $0 \leq t \leq 50$. Are the police able to maintain their numeric superiority over the insurgents?

b. Suppose the police were to increase their recruiting efforts. How many total new recruits do the police need each week to maintain their numeric superiority?

5.4.6 Consider a lake containing both bass and trout who compete for the same food source. Let $B(t)$ and $T(t)$ denote the populations, in thousands, of bass and trout, respectively, in month t. Suppose these functions are described by the system of differential equations

$$\begin{aligned} \frac{dB}{dt} &= (10 - B - T)B \\ \frac{dI}{dt} &= (15 - B - 3T)T. \end{aligned}$$

a. Find the equilibrium point(s) of the system.

b. Use Euler's method with $h = 0.1$ to graph both the time plots and trajectory over the interval $0 \leq t \leq 7$. Use random initial populations between 0 and 5.

c. Are the equilibrium point(s) from part a. attracting or repelling?

d. Does it appear the bass and trout can coexist? Briefly explain why or why not.

5.4.7 Consider two tanks of salt water separated by a thin permeable membrane through which salt can diffuse. Let $x_1(t)$ and $x_2(t)$ denote the masses of salt in the two tanks. If we assume that the rate at which salt diffuses though the membrane is proportional to the difference between the concentrations of the two solutions, we get the system of differential equations

$$\begin{aligned} \frac{dx_1}{dt} &= P\left(\frac{x_2}{V_2} - \frac{x_1}{V_1}\right) \\ \frac{dx_2}{dt} &= P\left(\frac{x_1}{V_1} - \frac{x_2}{V_2}\right) \end{aligned}$$

where P is a constant, called the *permeability* and V_1 and V_2 are the volumes of the respective tanks.

a. Suppose that $P = V_1 = V_2$. Plug these values into the system and simplify.

b. Show that any point where $x_1 = x_2$ is an equilibrium point of the system.

c. Use Euler's method with $h = 0.1$ to graph the time plots over the interval $0 \leq t \leq 3$. Use random initial values between 0 and 10.

d. What happens to the values of x_1 and x_2 in the long-term? Hypothesize how these long-term values are related to the initial values. Give numerical support for your hypothesis.

5.4.8 Consider the connected tanks in Figure 5.27 where each tank contains 100 L of salt water and solution flows through each pipe at the rate of 1 L/min.

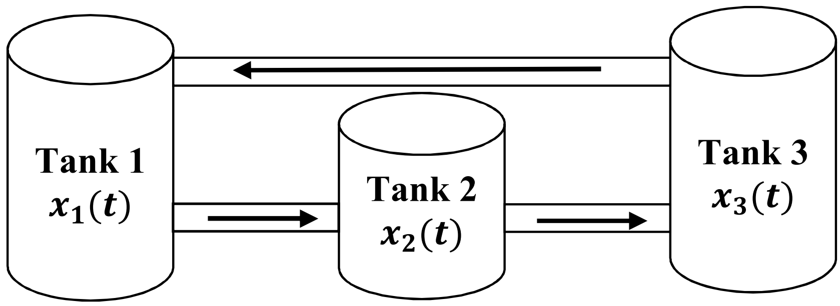


FIGURE 5.27

a. Find a system of differential equations for x_1, x_2, and x_3.

b. Suppose $x_1(0) = 5$, $x_2(0) = 12$, and $x_3(0) = 16$. Estimate $x_1(250)$, $x_2(250)$, and $x_3(250)$.

c. Hypothesize how the values of $x_1(250)$, $x_2(250)$, and $x_3(250)$ relate to the values of $x_1(0)$, $x_2(0)$, and $x_3(0)$. Explain how you came up with this hypothesis.

d. Support your hypothesis in part c. numerically and graphically by letting $x_1(0)$, $x_2(0)$, and $x_3(0)$ be random values between 0 and 20. Does your hypothesis always seem to be true?

5.4.9 Suppose the motion of a particle on the $x - y$ plane is described by the system of differential equations

$$\frac{dx_1}{dt} = x_2, \quad \frac{dx_2}{dt} = 2x_4 - 3x_1,$$
$$\frac{dx_3}{dt} = x_4, \quad \frac{dx_4}{dt} = -2x_2 - 3x_3,$$

where

- $x_1(t) = x(t)$ (the x-coordinate of the particle at time t),
- $x_2(t) = x'(t)$ (the rate at which the x-coordinate changes),
- $x_3(t) = y(t)$ (the y-coordinate of the particle at time t), and
- $x_4(t) = y'(t)$ (the rate at which the y-coordinate changes).

Consider the initial conditions $x_1(0) = 1$, $x_2(0) = x_3(0) = x_4(0) = 0$.

a. Use Euler's method with $h = \pi/500$ to approximate the solution to this system over the interval $0 \leq t \leq 2\pi$.

b. Graph y vs. x, that is, x_3 vs. x_1. This graph is an approximation of a curve called an *astroid*, also known as a *hypocycloid with four cusps*.

5.5 Quadratic Population Model

In this section we model the populations of two species with a system of two differential equations and graphically analyze the behavior of the system using Euler's method.

Example 5.5.1 (Competing Foxes and Wolves)
Consider a forest that contains foxes and wolves (we ignore all other animals in the forest). In the absence of any competition, foxes grow at the rate of 10% per year and wolves at a rate of 25% per year. The forest can support about 10,000 foxes or 6,000 wolves. The two species compete for the same resources, but the extent of this competition is not known. We want to know if both species can coexist or if one will dominate.

To model this scenario, define

$$\begin{aligned} F(t) &= \text{fox population at time } t \\ W(t) &= \text{wolf population at time } t \end{aligned}$$

We can use logistic equations as in Example 5.2.2 to model the two growth rates in terms of the carrying capacities:

$$\begin{aligned} \frac{dF}{dt} &= 0.10\left(1 - \frac{F}{10,000}\right)F \\ \frac{dW}{dt} &= 0.25\left(1 - \frac{W}{6,000}\right)W \end{aligned}$$

Now to model the effect of competition, we will assume that competition decreases the growth rates by an amount proportional to the product of the two populations (this product models the *interaction* of the two species). This yields the system

$$\begin{aligned} \frac{dF}{dt} &= 0.10\left(1 - \frac{F}{10,000}\right)F - c_1FW \\ \frac{dW}{dt} &= 0.25\left(1 - \frac{W}{6,000}\right)W - c_2FW \end{aligned}$$

where c_1 and c_2 are some unknown positive parameters. Rewriting this system, we get the model

$$\begin{aligned} \frac{dF}{dt} &= 0.10F - \frac{0.10}{10,000}F^2 - c_1FW \\ \frac{dW}{dt} &= 0.25W - \frac{0.25}{6,000}W^2 - c_2FW. \end{aligned}$$

The problem does not specify the values of c_1 and c_2, so let's think about possible values. The FW terms can be thought of as modeling competition *between* species while the square terms model competition *within* a species. Different species compete for similar, but different resources. Members within a species compete for the exact same resources. Therefore, it seems reasonable that competition within a species has a larger effect than competition between species. In other words, it seems reasonable that the coefficients of the FW terms are smaller in absolute value than the coefficients of the square terms. This suggests a model of the form

$$\begin{aligned}\frac{dF}{dt} &= 0.10F - \frac{0.10}{10,000}F^2 - \lambda\frac{0.10}{10,000}FW\\ \frac{dW}{dt} &= 0.25W - \frac{0.25}{6,000}W^2 - \lambda\frac{0.25}{6,000}FW,\end{aligned}$$

where λ is some parameter between 0 and 1. □

The model in Example 5.5.1 is an example of a *Quadratic Population Model* which has the general form

$$\begin{aligned}\frac{dx}{dt} &= a_1x + b_1x^2 + c_1xy\\ \frac{dy}{dt} &= a_2y + b_2y^2 + c_2xy.\end{aligned}$$

The parameter a describes how each population changes in the absence of any competition. The parameter b can be thought of as describing how competition *within* the species affects the rate of growth. The parameter c describes how competition *between* species affects the rate. The signs of these parameters tell us a lot about the system. For instance, in both the equations in Example 5.5.1, a is positive and b and c are negative. This tells us that both populations would grow in the absence of competition and that competition within each species and competition between species decrease both rates of growth.

The question in Example 5.5.1 is, "can the two species coexist?" To answer this question we need to analyze the long-term behavior of the populations and determine how this is affected by the value of λ.

Example 5.5.2 (Analyzing a Quadratic Population Model)
To graphically analyze the model in Example 5.5.1, follow these steps:

1. Rename a blank worksheet "**Quadratic**" and format it as in Figure 5.28. (Note that in this example, Species 1 is fox and Species 2 is wolf.)

	A	B	C	D
1	**Species 1**		**Species 2**	
2	**a₁ =**	0.1	**a₂ =**	0.25
3	**b₁ =**	=-0.1/10000	**b₂ =**	=-0.25/6000
4	**c₁ =**	=B5*B3	**c₂ =**	=B5*D3
5	**λ =**	0.5		

FIGURE 5.28

2. Add the formulas in Figure 5.29 and copy row 15 down to row 214.

	A	B
12	**h =**	1
13	**Time**	**Species 1**
14	0	1000
15	=A14+$B$12	=B14+$B$12*($B$2*B14+$B$3*B14^2+$B$4*B14*C14)

	C
13	**Species 2**
14	1000
15	=C14+$B$12*($D$2*C14+$D$3*C14^2+$D$4*B14*C14)

FIGURE 5.29

3. Create a graph as in Figure 5.30. Set the x-axis min and max to 0 and 10,000 and the y-axis min and max to 0 and 6000. The graph shows that the populations start at (1000, 1000) and stop changing at around (9333, 1333). Thus, with these parameters and initial populations, the foxes and wolves can indeed coexist.

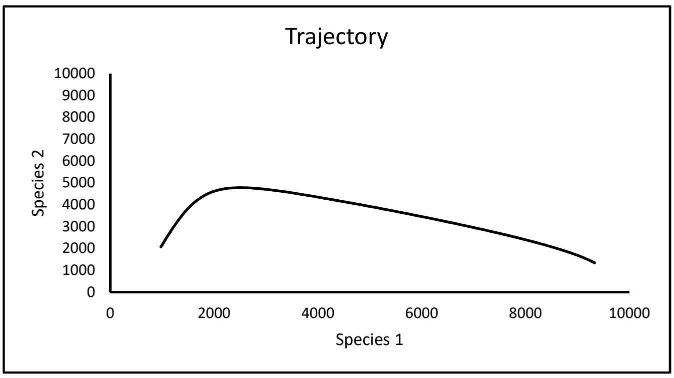


FIGURE 5.30

4. The initial populations of (1000, 1000) are arbitrary, so let's try different values and see what happens. Add the formulas in Figure 5.31 to generate random initial populations. (If these formulas do not work, select **Tools** → **Add-Ins...**, select **Analysis ToolPak** and press **OK**.)

	B	C
13	**Species 1**	**Species 2**
14	=RANDBETWEEN(0,10000)	=RANDBETWEEN(0,6000)

FIGURE 5.31

Press the **F9** key several times to try different initial populations. Note that for each one, the populations settle down at around (9333, 1333). Therefore, it appears that the populations can coexist regardless of the initial populations (at least with this value of λ).

□

It appears that the point (9333, 1333) is an equilibrium point of the system. Equilibrium points are important in population models because they can indicate what happens to the populations in the long–term.

Example 5.5.3 (Equilibrium Points of a Quadratic Population Model)
To find equilibrium points of a quadratic population model, we need to solve the algebraic system

$$0 = a_1x + b_1x^2 + c_1xy \tag{5.6}$$
$$0 = a_2y + b_2y^2 + c_2xy. \tag{5.7}$$

Obviously, (0, 0) is one equilibrium point. If $x \neq 0$ and $y = 0$, then (5.6) gives

$$\begin{aligned} & 0 = a_1x + b_1x^2 = x(a_1 + b_1x) \\ \Rightarrow \quad & 0 = a_1 + b_1x \\ \Rightarrow \quad & x = -\frac{a_1}{b_1} \end{aligned}$$

Thus

$$\left(-\frac{a_1}{b_1}, 0\right) \tag{5.8}$$

is another equilibrium point. If $x = 0$ and $y \neq 0$, similar calculations give

$$\left(0, -\frac{a_2}{b_2}\right) \tag{5.9}$$

as another. Assuming that $x \neq 0$ and $y \neq 0$ requires us to solve the system of equations

$$\begin{aligned} 0 &= a_1x + b_1x^2 + c_1xy \\ 0 &= a_2y + b_2y^2 + c_2xy \end{aligned}$$

for x and y. Dividing the first equation by x and the second by y and rewriting yields

$$\begin{aligned} b_1x + c_1y &= -a_1 \\ c_2x + b_2y &= -a_2 \end{aligned}$$

Writing this in matrix form gives

$$\begin{bmatrix} b_1 & c_1 \\ c_2 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -a_1 \\ -a_2 \end{bmatrix}$$

Now, Cramer's rule gives

$$x = \frac{\begin{vmatrix} -a_1 & c_1 \\ -a_2 & b_2 \end{vmatrix}}{\begin{vmatrix} b_1 & c_1 \\ c_2 & b_2 \end{vmatrix}} = \frac{-a_1b_2 + a_2c_1}{b_1b_2 - c_1c_2} \quad \text{and} \quad y = \frac{\begin{vmatrix} b_1 & -a_1 \\ c_2 & -a_2 \end{vmatrix}}{\begin{vmatrix} b_1 & c_1 \\ c_2 & b_2 \end{vmatrix}} = \frac{-a_2b_1 + a_1c_2}{b_1b_2 - c_1c_2}$$

Thus the fourth equilibrium point is

$$\left(\frac{-a_1b_2+a_2c_1}{b_1b_2-c_1c_2}, \frac{-a_2b_1+a_1c_2}{b_1b_2-c_1c_2}\right) \tag{5.10}$$

as long as $b_1b_2 - c_1c_2 \neq 0$. This one is particularly important because both x- and y-coordinates are (potentially) positive. At this equilibrium, the two species coexist and the populations do not change. □

Example 5.5.4 (Calculating Equilibrium Points)
We can easily implement Formulas (5.8), (5.9), and (5.10) in Excel and use the results to analyze the fox-wolf model in Example 5.5.1. In the worksheet **Quadratic**, add the formulas in Figure 5.32.

	A	B	C	D	E
8		**Equilibrium Points**			
9	**x =**	=(-B2*D3+D2*B4)/(B3*D3-B4*D4)	=-B2/B3	0	0
10	**y =**	=(-D2*B3+B2*D4)/(B3*D3-B4*D4)	0	=-D2/D3	0

FIGURE 5.32

We see that the system has equilibrium points of $(9333.\bar{3},\ 1333.\bar{3})$, $(10,000,\ 0)$, $(0,\ 6,000)$ and $(0,\ 0)$. Since all trajectories appear to be attracted to the point $(9333.\bar{3},\ 1333.\bar{3})$, we have graphical evidence that this equilibrium point is attracting. The other equilibrium points appear to be repelling.

Now let's see what happens to the equilibrium points as λ changes. Add a scrollbar, set the min and max to 0 and 1000, respectively, and the linked cell to **B6**. Add the formula in Figure 5.33.

	A	B
5	**λ =**	=B6/1000

FIGURE 5.33

Using the scroll bar, we see that for λ between 0 and 0.6, the first equilibrium point has positive x- and y-coordinates, and it appears to always be attracting. Thus the two species can coexist. For $\lambda > 0.6$, the y-coordinate of the equilibrium point is negative. This means that Species 2 (the wolves) will die out and the foxes will dominate.

So to answer the question "can the two species coexist?" we need to determine which is more likely: $\lambda < 0.6$ or $\lambda > 0.6$. It seems reasonable that competition within a species is much greater than competition between species. This means that c_i is much less than b_i, so λ must be very small (i.e. $\lambda < 0.6$). Therefore, we conclude that the two species can coexist. □

The general quadratic population model can be adapted to fit a variety of different scenarios, as the next example illustrates.

Example 5.5.5 (Predator–Prey System)
Consider the example of foxes and rabbits discussed in Chapter 4 where rabbits are the sole source of food for foxes. Let's suppose that without rabbits, foxes die at a rate of 8% per month, and without foxes, the rabbit population grows at a rate of 4% per month. The presence of rabbits increases the growth of the fox population and the presence of foxes decreases the growth of the rabbit population.

If $F(t)$ and $R(t)$ represent the populations of foxes and rabbits at time t, respectively, we can model this system using the differential equations (instead of difference equations as in Chapter 4):

$$\frac{dF}{dt} = -0.08F + c_1FR$$
$$\frac{dR}{dt} = 0.04R + c_2FR$$

In Chapter 4, we took $c_1 = 0.0001$ and $c_2 = -0.0003$. Note that this model is a quadratic population model without the square terms (i.e. $b_1 = b_2 = 0$). Entering these values of the parameters into the worksheet **Quadratic**, with a initial population of 100 foxes (Species 1) and 1000 rabbits (Species 2) and $h = 0.6$, yields a trajectory as shown in Figure 5.34 (you may need to change the scales on the axes to reproduce this graph).

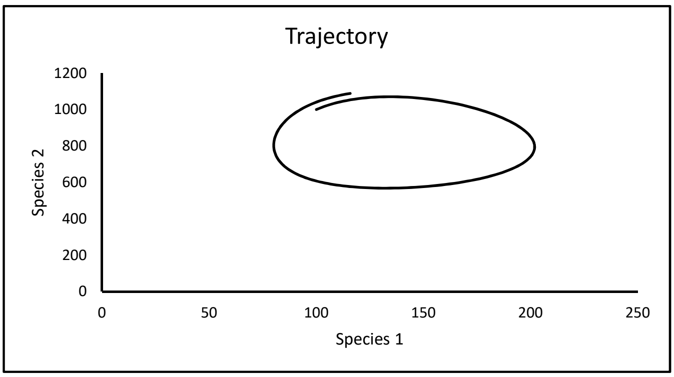


FIGURE 5.34

Note that the first equilibrium point is $(133.\bar{3}, 800)$ and the trajectory forms a loop around this point. (Note that in theory, the trajectory is a closed loop, but since our graph is only an approximation, the loop is open. Using a smaller value of h and more iterations of Euler's method would produce a more accurate graph.) Other initial populations give similar results. Initial populations further from the equilibrium produce larger loops. This means that the two populations vary more over time. Since the trajectories are not attracted to the equilibrium, we have evidence that it is repelling.

These loops mean that the populations oscillate over time. Some biologists argue that this type of model is not realistic because in nature, populations do not tend to oscillate. Rather, they tend to move toward an equilibrium point, as in the fox–wolf model.

□

Exercises

Directions for Exercises 5.5.1 - 5.5.4: For each given quadratic population model,

a. Modify the worksheet **Quadratic** to graph the trajectory. Set each initial population to a random number in the given interval. Use the given value of h.

b. Classify each of the four equilibrium points as attracting or repelling.

c. Will the populations be able to coexist? Why or why not?

5.5.1 Initial population interval: [0, 40], $h = 0.01$

$$\frac{dx}{dt} = 14x - 0.5x^2 - xy \qquad \frac{dy}{dt} = 16y - 0.5y^2 - xy$$

5.5.2 Initial population interval: [0, 10], $h = 0.01$

$$\frac{dx}{dt} = 14x - 2x^2 - xy \qquad \frac{dy}{dt} = 16y - 2y^2 - xy$$

5.5.3 Initial population interval: [0, 100], $h = 0.1$

$$\frac{dx}{dt} = 0.2x - 0.005xy \qquad \frac{dy}{dt} = -0.5y + 0.01xy$$

5.5.4 Initial population interval: [0, 10], $h = 0.01$

$$\frac{dx}{dt} = 5x - x^2 - xy \qquad \frac{dy}{dt} = -2y + 2xy$$

5.5.5 Investigate sensitivity of the fox–wolf model in Example 5.5.1 to the carrying capacities. That is, change the carrying capacities a small amount and analyze the resulting model. Does your final conclusion change?

5.5.6 Consider the predator–prey system model in Example 5.5.5. Use a scroll bar to investigate the sensitivity of the model to the parameter c_2.

a. What happens to the first equilibrium point (the point given by (5.10)) as c_2 gets closer to 0? What if it gets further from 0?

b. What happens to the variation within each population as c_2 gets closer to 0? What if it gets further from 0?

c. If shelters were built to protect the rabbits from the foxes, would c_2 get closer to 0 or further from 0? What does the model predict might happen to the populations? Would it increase the size of the rabbit population?

5.5.7 Suppose two populations, call them x and y, are described by a quadratic population model. Further suppose that in the absence of competition both populations grow. For each of the following scenarios, give possible values of the parameters a_1 through c_2.

a. For both populations, there is no competition within or between species.

b. For both populations, there is no competition within species and competition between species decreases the rate of growth of both species.

c. For both populations, competition within species decreases the rate of growth and competition between species increases the rate of growth.

d. For population x, competition within species increases the rate of growth and competition between species decreases the rate of growth. For population y, the opposite is true.

e. For both populations, competition decreases the rate of growth, but competition between species is much greater than competition within species.

5.5.8 Suppose the populations of two species are described by the following system of differential equations:

$$\begin{aligned}\frac{dx}{dt} &= -x + y\\ \frac{dy}{dt} &= -y + x.\end{aligned}$$

a. According to this model, can either species survive without the other? Explain why or why not.

b. Find the equilibrium points of this system.

c. Describe the long-term behavior of the system for different initial populations.

d. Would you describe this system as being sensitive to the initial populations? Explain why or why not.

5.5.9 A predator–prey model that takes into account harvesting (i.e., hunting) of the two species is

$$\begin{aligned}\frac{dx}{dt} &= a_1x - b_1xy - cx\\ \frac{dy}{dt} &= -a_2y + b_2xy - cy\end{aligned}$$

where $x(t)$ and $y(t)$ are the populations of the prey and predator species, respectively. All parameters are assumed to be positive. Assuming that $x \neq 0 \neq y$, find a formula for the equilibrium point in terms of the parameters.

5.5.10 Consider the predator-prey system in Example 5.5.5 with initial populations of 100 foxes and 1000 rabbits. In this model we used constant values of a_1 and a_2 (the numbers -0.08 and 0.04). In the terminology of Chapter 4, these are the fox death and rabbit birth factors, respectively. Now suppose these factors change throughout the year. Suppose that during the "summer", these factors are -0.08 and 0.04, respectively, but during the "winter" they are -0.1 and 0.01, respectively.

a. Use Euler's method to estimate the populations from month 0 to month 250 using $h = 1$. Create one graph of the fox population vs. month and another graph for the rabbit population. Let the winter occur from month 0 through 5, then 12 through 17, and so on.

b. Compare the populations when the factors are constant to the populations when the factors vary. Do the varying factors cause the ranges of the populations to increase or decrease?

5.5.11 Our original fox-wolf quadratic population model has the general form

$$\begin{aligned}\frac{dx}{dt} &= a_1x - \frac{a_1}{L_1}x^2 - \lambda\frac{a_1}{L_1}xy\\ \frac{dy}{dt} &= a_2y - \frac{a_2}{L_2}y^2 - \lambda\frac{a_2}{L_2}xy\end{aligned}$$

where a_1 and a_2 are the intrinsic growth rates, L_1 and L_2 are the carrying capacities of the respective species, and λ measures the effect of competition between species relative to competition between species ($\lambda < 1$ means competition between has less effect than within). The goal of this exercise is to explore the relationship between λ and the intrinsic growth rates.

a. Suppose $L_1 = L_2 = 2,000$ (i.e. the environment can support 2,000 of each species), $a_1 = 0.1$, $a_2 = 0.2$, $\lambda = 0.5$, and $x(0) = y(0) = 1,000$. Modify the worksheet **Quadratic** to calculate the equilibrium points and graph the trajectory.

b. Use a scroll bar to vary λ between 0 and 2. For what values of λ do the species coexist? What happens for other values of λ?

c. Now suppose $a_1 = 0.2$ and $a_2 = 0.1$. That is, suppose the intrinsic growth rates switch. Repeat part b.

d. Now suppose $a_1 = a_2 = 0.1$. Repeat part b.

e. Generalize your results. In the case with equal initial populations, which species has a better chance of surviving, the one with the higher intrinsic growth rate, or the one with the lower rate?

5.6 Volterra's Principle

The following scenario is described by Braun (Braun, Martin, "Why the Percentage of Sharks Caught in the Mediterranean Sea Rose Dramatically during World War I," in *Modules in Applied Mathematics* Volume 1 *Differential Equation Models*, ed. William F. Lucas, Springer-Verlag, 1983, p. 221, used by permission):

> In the mid-1920's the Italian biologist Umberto D'Ancona was studying variations in the population of various species of fish that interact with each other. In the course of his research, he came across data on percentages-of-total-catch of several species of fish that were brought into different Mediterranean ports in the years that spanned World War I. In particular, the data gave the percentage-of-total-catch of selachians (sharks, skates, rays, etc.) which are not very desirable as food fish. The data for the port of Fiume, Italy, during the years 1914 – 1923 is as follows:

1914	1915	1916	1918	1919	1920	1921	1922	1923
11.9%	21.4%	22.1%	36.4%	27.3%	16.0%	15.9%	14.8%	10.7%

> D'Ancona was puzzled by the very large increase in the percentage of selachians during the period of the war. Obviously, he reasoned, the increase in the percentage of selachians was due to the greatly reduced level of fishing during this period, but how does the intensity of fishing affect the fish populations. . . . It was also a concern to the fishing industry, since it would have obvious implications for the way fishing should be done.

D'Ancona took this problem to the famous Italian mathematician Vito Volterra. Volterra noted that the selachians are predators and the food fish are their prey. So he devised a simple predator–prey model:

$$\frac{dx}{dt} = a_1 x - b_1 xy$$
$$\frac{dy}{dt} = -a_2 y + b_2 xy$$

where $x(t)$ and $y(t)$ are the populations of the prey (the food fish) and predator (the selachians), respectively, and a_1, a_2, b_1, $b_2 > 0$. To model the impact of fishing, Volterra added an additional term to each equation:

$$\frac{dx}{dt} = a_1 x - b_1 xy - cx \tag{5.11}$$

$$\frac{dy}{dt} = -a_2 y + b_2 xy - cy \tag{5.12}$$

The parameter $c > 0$ could be thought of as the proportion of each population caught per unit of time. Rewriting (5.11) and (5.12) we get:

$$\frac{dx}{dt} = (a_1 - c)x - b_1 xy$$

$$\frac{dy}{dt} = (-a_2 - c)y + b_2 xy$$

We see that this model is really just a special case of the quadratic population model studied in Section 5.5. Using the formula derived in Exercise 5.5.9, we calculate that there is an equilibrium point at

$$\left(\frac{a_2 + c}{b_2}, \frac{a_1 - c}{b_1} \right)$$

Let's examine a system such as this graphically:

1. Rename a blank worksheet "**Volterra**" and format it as in Figure 5.35. Note that the value of these parameters do not come from the data. They are simply values we use to illustrate the point we want to make.

	A	B	C	D
1	**Prey**		**Predator**	
2	$\mathbf{a_1}$ **=**	0.04	$\mathbf{a_2}$ **=**	0.08
3	$\mathbf{b_1}$ **=**	0.0004	$\mathbf{b_2}$ **=**	0.0001
4	**c =**	0.03		
5				
6	**Equilibrium**			
7	**x =**	=(D2+B4)/D3		
8	**y =**	=(B2-B4)/B3		

FIGURE 5.35

2. Add the formulas in Figure 5.36 to compute numerical solutions using Euler's method. Copy row 13 down to row 1012. Use the calculations to create a graph as in Figure 5.37. As with the parameters, these initial populations are arbitrary.

   Notice that we get a loop as in the fox and rabbit predator–prey model. This means that the populations are periodic (like a sin or cos curve). Mathematically, this means there exists a time T such that

   $$x(T) = x(0) \quad \text{and} \quad y(T) = y(0)$$

3. To approximate the value of T for this model, we will use a scroll bar to graph the trajectory in Figure 5.37 over an interval of time so that it makes only one full "loop." Add a scroll bar, set the linked cell to **I1** and the min and max to 0 and 1000, respectively. Add the formulas in Figure 5.38.

	A	B
10	**h =**	0.25
11	**Time**	**Prey**
12	0	1000
13	=A12+$B$10	=B12+$B$10*($B$2*B12-$B$3*B12*C12-$B$4*B12)

	C
11	**Predator**
12	100
13	=C12+$B$10*(-$D$2*C12+$D$3*B12*C12-$B$4*C12)

FIGURE 5.36

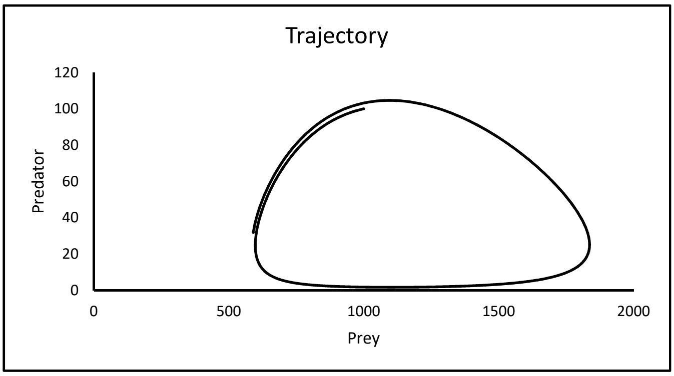


FIGURE 5.37

	F
1	**Time**
2	=I1*B10

FIGURE 5.38

4. Modify the formulas in Euler's method as in Figure 5.39. Copy this modified row 13 down to row 1012. These formulas will calculate the two populations *only* at each time less than the time in cell **F2**. For larger values of time, the formula will return **#N/A** which means "value not available" and is equivalent to a blank cell.

	B
13	=IF(A13<=$F$2,B12+$B$10*($B$2*B12-$B$3*B12*C12-$B$4*B12),NA())

	C
13	=IF(A13<=$F$2,C12+$B$10*(-$D$2*C12+$D$3*B12*C12-$B$4*C12),NA())

FIGURE 5.39

5. Use the scrollbar to find a time so that the graph looks similar to Figure 5.40. Notice that in Figure 5.40, the populations return to where they started, so they have each completed one full cycle. The corresponding time is approximately 220. This means $T \approx 220$.

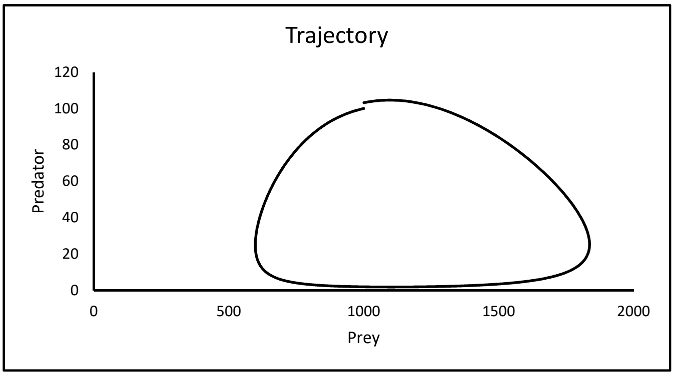


FIGURE 5.40

6. Next let's look at the average of each population over the period of time $[0, T]$. Add the formulas in Figure 5.41.

	E	F
7	**Prey Average =**	=SUMIF(B12:B1012,">0")/I1
8	**Predator Average =**	=SUMIF(C12:C1012,">0")/I1

FIGURE 5.41

The results of the calculations are shown in Figure 5.42. Comparing the averages to the equilibrium point we note that they're the same (at least approximately)!! If we change the initial populations and the parameters, and each time find T, we observe that the averages are always approximately equal to the equilibrium points. This illustrates theorem 5.6.1.

	A	B	C	D	E	F
6	**Equilibrium**					
7	**x =**	1100		**Prey Average =**		1104.35
8	**y =**	25		**Predator Average =**		25.33

FIGURE 5.42

Theorem 5.6.1. *Let $(x(t), y(t))$ be a periodic solution of the system described by (5.11) and (5.12) with period T. Define the average values of x and y as*

$$\bar{x} = \frac{1}{T}\int_0^T x(t)\,dt \quad \text{and} \quad \bar{y} = \frac{1}{T}\int_0^T y(t)\,dt$$

Then

$$\bar{x} = \frac{a_2 + c}{b_2} \quad \text{and} \quad \bar{y} = \frac{a_1 - c}{b_1}$$

□

Theorem 5.6.1 explains D'Ancona's observations. We see that a moderate amount of fishing ($c < a_1$) increases $\bar{x}$ (the average number of prey, or food fish) and decreases $\bar{y}$ (the average number of predators, or selachians). Conversely, decreased fishing (as happened during WWI) increases the number of selachians and decreases the number of food fish, on average. The fact that some fishing increases the number of food fish is known as *Volterra's Principle*.

Exercises

5.6.1 In the original "fishing" model, we assumed that both the predator and prey are caught at the same rate. Now consider a refined "insecticide" model where the predator and prey are killed at different rates:

$$\begin{aligned}\frac{dx}{dt} &= 0.05x - 0.0004xy - c_1 x \\ \frac{dy}{dt} &= -0.04y + 0.0001xy - c_2 y\end{aligned}$$

where $c_1 \neq c_2$ are positive numbers.

a. Assume that $c_2 = 0.01$. Modify the worksheet **Volterra** to graph trajectories for this refined model.

b. Assume that we start with 100 prey and 50 predator insects. For what values of c_1 is the prey species killed off before the predator species? For what values is the the predator species killed off first? For what values is neither species killed off? (**Suggestion**: We could consider the prey species be to be "killed off" first when its population drops below 1 before the population of the predator species does. Graphically, this is when the trajectory touches, or nearly touches, the predator–axis (the y-axis). Neither species is killed off when neither population drops below 1.)

5.6.2 Consider the competing foxes and wolves model from Example 5.5.1 with an additional term to model the hunting, or harvesting, of both species:

$$\begin{aligned}\frac{dF}{dt} &= 0.10F - \frac{0.10}{10,000}F^2 - 0.5\left(\frac{0.10}{10,000}\right)FW - h_0 F \\ \frac{dW}{dt} &= 0.25W - \frac{0.25}{6,000}W^2 - 0.5\left(\frac{0.25}{6,000}\right)FW - h_0 W\end{aligned}$$

where h_0 is a parameter measuring the amount of hunting.

a. Rewrite this model so it fits the form of a standard quadratic population model given in Section 5.5. Modify the worksheet **Quadratic** to implement this rewritten model. Create a cell for the value of h_0 (suppose we start with 1,500 foxes and 1,000 wolves).

b. Find the largest value of h_0 so that both coordinates of the equilibrium given by Formula (5.10) are positive. What does it mean if both of these coordinates are positive? What if one is negative?

c. Suppose that at time $t = 0$, when people started hunting the two species, the populations were 1,500 foxes and 1,000 wolves and that at time $t = 50$, the fox population is around 100. Find the value of h_0 necessary for this to happen.

d. What would happen to the two populations if the level of hunting found in part c. were to continue?

5.7 Lanchester Combat Models

In this section we present an application of systems of differential equations that is very much different than the population models presented earlier. However, we will see that the resulting models are not that much different.

During World War I, F. W. Lanchester devised several mathematical models of warfare. Since then, his models have been widely studied and adapted to a variety of scenarios ranging from “isolated battles to entire wars.” In this section we will graphically analyze one such model described by Courtney S. Coleman in the chapter “Combat Models” in the book *Modules in Applied Mathematics* Volume 1, ed. William F. Lucas, Springer-Verlag, 1983, pp. 109–131.

Let $A(t)$ represent the number of combatants in army A at time t. The rate at which $A(t)$ changes with respect to time, dA/dt, is affected by several factors including casualties caused by the opposing army, disease, desertions, and reinforcements. For simplicity, we will only consider the first factor.

The rate at which combatants are lost due to casualties caused by the opposing army is often referred to as the *combat loss rate* (CLR). In mathematical notation, this is described by the differential equation

$$\frac{dA}{dt} = -\text{CLR}$$

Armies are divided into two general categories: conventional and guerrilla. A conventional army operates in relative large units with an identifiable front line while a guerilla army operates in small units without a front line.

Consider a battle where a conventional army C goes against a smaller guerrilla army G. We can describe this scenario with a basic system of differential equations:

$$\begin{aligned} \frac{dC}{dt} &= -\text{CLR}_C \\ \frac{dG}{dt} &= -\text{CLR}_G \end{aligned}$$

Suppose that army C is out in the open in some formation and army G is hidden in the trees of a forest and that each combatant in each army is firing a gun.

Let’s consider CLR_C. This is the rate at which combatants are killed or wounded. Obviously the larger $G(t)$ is, the higher the rate. This suggests a proportionality relationship:

$$\frac{dC}{dt} = -g\,G(t) \tag{5.13}$$

The constant of proportionality g is called the *combat effectiveness coefficient* of army G and is defined as:

$$g = r_G p_G$$

where

r_G = Firing rate of army G (shots/day/combatant), and
p_G = Probability that a single shot from army G will hit an opponent.

Now consider CLR_G. It seems reasonable that this is proportional to $C(t)$. However, combatants in army C cannot see those in army G. So they are blindly firing into the forest (this has been called “spray and pray”). Thus, the larger $G(t)$ is, the larger the probability that a shot will hit an opponent. This suggests another proportionality relationship:

$$\frac{dG}{dt} = -c\,C(t)\;G(t) \tag{5.14}$$

The combat effectiveness coefficient c is defined in a similar fashion as g, $c = r_C p_C$. Since army C is blindly firing into the forest, the probability that a shot will hit an opponent can be described by

$$\begin{aligned} p_C &= \frac{\text{Area of the exposed part of the body of a single guerilla}}{\text{Total area occupied by the guerrillas}} \\ &= \frac{\text{Area of the exposed part of the body of a single guerilla}}{(\text{Area occupied by a single guerilla}) \cdot G_0} \end{aligned}$$

Where $G_0 = G\,(0)$ (the initial number of guerilla combatants). Putting (5.13) and (5.14) together, we have our model:

$$\frac{dC}{dt} = -g\,G \tag{5.15}$$

$$\frac{dG}{dt} = -c\,C\,G \tag{5.16}$$

(the time variable t has been dropped for simplicity). The type of battle modeled here was common in Vietnam where the conventional American and South Vietnamese army fought the guerilla North Vietnamese and Viet Cong army. For this reason, this model is called the "Vietnam" model.

Now we will use this model to analyze the question "what ratio of initial combatants ($C_0/G_0 = n$) is necessary for army C to win?" We say that army C "wins" when army G runs out of combatants first.

To answer this question, we will solve the system (5.15) and (5.16) for G in terms of C. Note that

$$\frac{dC}{dG} = \frac{-g\,G}{-c\,C\,G} = \frac{g}{c\,C}$$

Cross-multiplying, we get

$$g\,d\,G = c\,C\,d\,C \tag{5.17}$$

Now integrating both sides of (5.17) yields

$$g\,G = \frac{1}{2}c\,C^2 + M$$

where M is an arbitrary constant. Dividing by g gives

$$G = \frac{c}{2g}\,C^2 + \frac{M}{g} \tag{5.18}$$

Equation (5.18) does not give us $G\,(t)$ or $C\,(t)$ in terms of t, but it does give us a *relation* between $G\,(t)$ and $C\,(t)$. We can use this to answer the question graphically.

Now to find the value of M we need to use the conditions that $G\,(0) = G_0$ and $C\,(0) = C_0$. Evaluating (5.18) at $t = 0$ gives

$$\begin{aligned} G\,(0) &= \frac{c}{2g}\,C(0)^2 + \frac{M}{g} \\ \Rightarrow \quad G_0 &= \frac{c}{2g}\,C_0^2 + \frac{M}{g} \\ \Rightarrow \quad M &= g\,G_0 - \frac{c}{2}\,C_0^2 \end{aligned}$$

To implement this model and analyze it to find the value of n so that army C wins, follow these steps:

1. Rename a blank worksheet "**Vietnam**" and format it as in Figure 5.43. Copy row 11 down to row 309 to calculate 300 pairs of values of C and G. Note that the values of C in column **A** are in increments of 15 (i.e. they aren't calculated using any formula). The corresponding values of G in column **B** are calculated using Equation (5.18). Also note that the values of the parameters in Figure 5.43 are somewhat arbitrary, but reasonable.

	A	B
1	**Firing Rate =**	10
2	**Exposed Area =**	2
3	**Area/Guerrilla =**	1000
4	**p_G =**	0.1
5	**n =**	=D1/100
6	**g =**	=B1*B4
7	**c =**	=B1*B2/(B3*B10)
8	**M =**	=B6*B10-B7/2*A10^2
9	**Conventional**	**Guerrilla**
10	=B5*B10	150
11	=A10-15	=$B$7/(2*$B$6)*A11^2+$B$8/$B$6

FIGURE 5.43

2. Add a scroll bar, set the linked cell to **D1** and the min and max to 0 and 1500, respectively. Create a graph similar to those in Figure 5.44.

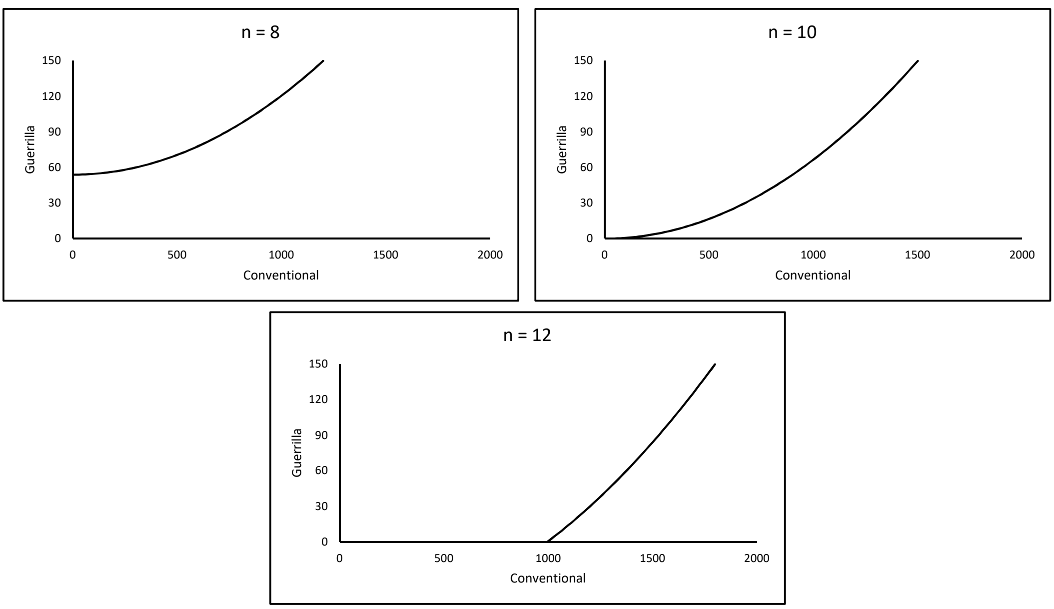


FIGURE 5.44

In 5.44 we see that for $n = 8$ (that is when army C is initially 8 times as large as army G), when $C(t) = 0$, $G(t) \approx 54$. This means that army G wins. We do not know the time t at which this happens, but this is not terribly important for our analysis.

For $n = 12$, when $G(t) = 0$, $C(t) \approx 1,000$. This means that army C wins. For $n = 10$, $C = 0 = G$ at approximately the same time (in other words they are both totally destroyed at the same time and nobody wins). Using the scroll bar to vary the value of n, we see that army C wins for $n > 10$ and loses for $n < 10$.

To perform a sensitivity analysis on the model, we change one initial condition or parameter at a time and find the approximate minimum value of n for army C to win. Changing G_0 or the Firing Rate we get the same results as above. Table 5.1 shows ranges of values of the other parameters and the associated ranges of the minimum values of n.

TABLE 5.1

	Exposed Area	**Area/Guerrilla**	p_G
Range of Values	1.5 – 4	500 – 1,750	0.05 – 0.15
Range of n	11.5 – 7	7 – 13.1	7 – 12.2

Summarizing our sensitivity analysis, we conclude that for army C to win, n must be *at least* 7, and probably more.

To verify our model we need some data. Table 5.2 (data adapted from a graph given by Coleman, p. 119) gives n for several guerrilla–conventional conflicts since WWII and the victors. In this data, n is computed using average force strengths over the period of time and does not take into account reinforcement rates or non–combat loss rates. Thus we should be careful about interpreting the data. Nevertheless, we see that the data do tend to support our conclusion.

TABLE 5.2

Conflict	n	**Victor**
Greece: 1946 – 49	9	
Malaya: 1945 – 54	18	Conventional
Kenya: 1953	10	
Philippines: 1948 – 52	4	
Indochina: 1945 – 54	2	
Indonesia: 1945 – 47	2	
Cuba: 1958 – 59	6	Guerrilla
Laos: 1959 – 62	3	
Algeria: 1956 – 62	10	
Vietnam: 1959	9	
Vietnam: 1968	6	
Vietnam: 1975	≈ 4	

Example 5.7.1 (Application to the Vietnam War)
In the spring of 1968 there were approximately 1,680,000 conventional forces lead by the U.S. and 280,000 guerrilla forces lead by the North Vietnamese and Viet Cong in Vietnam. This means the ratio of conventional forces to the guerrilla forces was approximately $\frac{1{,}680{,}000}{280{,}000} = 6$ (i.e. $n \approx 6$). Around this time, General Westmoreland, then commander of U.S. forces in South Vietnam, requested an additional 206,000 troops from President Johnson. Could this have actually helped?

With an additional 206,000 troops, the ratio of conventional forces to the guerrilla forces would have increased to $\frac{1{,}866{,}000}{280{,}000} \approx 6.7$ (assuming the size of the guerrilla force did not change). This ratio is too small for the conventional forces to win (at least as predicted by our model and supported by the data). Analysis such as this, and a myriad of other factors, caused President Johnson to reject the request for additional troops. □

Exercises

5.7.1 Suppose two conventional armies, x and y, are engaged in battle modeled by the system

$$\frac{dx}{dt} = -by$$
$$\frac{dy}{dt} = -cx$$

where $x(t)$ and $y(t)$ represent the number of combatants in armies x and y, respectively. Solving this system yields the relationship

$$y = \sqrt{\frac{c(x^2 - x_0^2)}{b} + y_0^2}$$

Parameters c and b are the combat effectiveness coefficients of armies x and y, respectively. They represent the strengths of the respective armies. Assume that army y is more powerful than army x (i.e. $b > c$).

a. Assume that $c = \lambda b$ for some $\lambda < 1$. Define $n = \frac{y_0}{x_0}$ (the ratio of initial forces). Create a spreadsheet to graph $y(t)$ vs. $x(t)$ for different values of λ and n. (**Suggestions**: Use $x_0 = 100$ with values of x in increments of 1. Also initially take $b = 1.5$. Try different values of b. Does the value of b really make a difference?)

b. Since army y is more powerful, for it to win it seems reasonable that y_0 can be less than x_0. How much less can y_0 be so that army y wins? To answer this question, define n_0 to be the value of n so that both armies are destroyed (i.e. the graph of $y(t)$ vs. $x(t)$ goes through the origin). Choose several values of λ and find the corresponding value of n_0 for each. Find a formula for n_0 in terms of λ. (**Hint**: It's a very simple formula.)

5.7.2 Suppose two conventional armies, x and y, are engaged in battle using weapons that can be aimed at specific targets (such as rifles) and weapons that can impact a large area (such as grenades). If $x(t)$ and $y(t)$ represent the number of combatants in army x and y, respectively at time t, a Lanchester model of this battle is:

$$\frac{dx}{dt} = -ay - bxy$$
$$\frac{dy}{dt} = -cx - dxy$$

The parameters a and b represent the effectiveness of army y's specific target weapons and their area weapons, respectively. The parameters c and d have the same meaning for army x.

Army x has a three-to-one numerical superiority at the beginning of the battle. However, army y is better trained, better equipped, and their weapons are more effective. This means that $a > c$ and $b > d$.

a. Assume that $c = 0.1$, $d = 0.001$, $x(0) = 3$, $y(0) = 1$, $a = \lambda c$, and $b = \lambda d$ for some $\lambda > 1$. Let λ_0 be the minimum value of λ necessary for army y to win the battle. Use a graphical approach and Euler's method to approximate the value of λ_0.

b. Repeat part 1, but now assume that army x has a 4:1 numerical superiority (i.e. $x(0) = 4$ and $y(0) = 1$). What about a 5:1 superiority?

c. Generalize your results in part 2. Suppose army x has an n:1 numerical superiority where $n \geq 1$. Conjecture a sufficient condition for λ guaranteeing army y wins. This sufficient condition should be in the form of a simple relationship between λ and n.

d. (**Extra Credit**) Prove your conjecture in part c. analytically (i.e. don't use graphs). **Hint**: Use the differential equations to find $\frac{dy}{dx}$ in terms of a, b, c, d, x, and y. Substitute in the relationships $a = \lambda c$, and $b = \lambda d$. Calculate this slope on the line $x = ny$. Show that your relationship in part c. guarantees this slope is less than $\frac{1}{n}$.

5.7.3 Consider the conventional vs. guerilla combat modeled in the worksheet **Vietnam**. We used the worksheet to graphically estimate the value of n so that the conventional army wins. In this exercise we find this value analytically.

a. Let G_1 denote the number of guerilla troops left when $C = 0$. Use Equation (5.18) to show that

$$G_1 = G_0 - \frac{c}{2g} C_0^2.$$

b. If $G_1 < 0$, then the conventional army wins. Use the fact that $n = C_0/G_0$ to show that the conventional army wins if

$$n > \frac{2g}{c\,C_0}.$$

c. Use the result from part b. and the facts that $c = r_C p_C$ and $p_C = A_1/(A_2 G_0)$, where A_1 = area of the exposed part of the body of a single guerilla and A_2 = area occupied by a single guerilla, to show that the conventional army wins if

$$n > \sqrt{\frac{2gA_2}{r_C A_1}}.$$

Graphically confirm this result with the spreadsheet.

5.8 Runge-Kutta Methods

Consider the problem of estimating the solution curve of a differential equation of the form

$$\frac{dy}{dt} = F(t, y)$$

along with the initial condition $y(t_0) = y_0$ where t_0 and y_0 are some given values. In Section 5.2 we did this using Euler's method. Euler's method is relatively simple to implement, but as we've seen, it can give inaccurate results. In this section we present two improved methods developed in the early 1900's by the German mathematicians Carl David Tolmè Runge and Martin Wilhelm Kutta.

Runge-Kutta 2$^{\text{nd}}$ Order Method

To motivate the first Runge-Kutta method, observe that as shown in Figure 5.2, Euler's method approximates $y'(t,\, y)$ over the interval $[t_n,\, t_{n+1}]$ with the constant value $F(t_n,\, y_n)$. Then the change in y over the interval is approximated by $h\,F(t_n,\, y_n)$. Stated another way, Euler's method approximates the change in y over the interval using the left end-point of the interval. One could argue that the midpoint of the interval would give a better approximation.

This idea leads to the following algorithm, called the *Runge-Kutta* 2^{nd} *order method*, or simply *RK2*:

$$\begin{aligned} k_1 &= h\,F(t_n, y_n) \\ k_2 &= h\,F\left(t_n + \frac{1}{2}h,\, y_n + \frac{1}{2}k_1\right) \\ x_{n+1} &= x_n + h \\ y_{n+1} &= y_i + k_2 \end{aligned}$$

Informally, k_1 is an approximate change in y over the interval $[t_n,\, t_{n+1}]$. The point $\left(t_n + \frac{1}{2}h,\, y_n + \frac{1}{2}k_1\right)$ is an approximate point on the solution curve at the midpoint of the interval. k_2 is an approximate change in y over the interval based on this point. Hopefully, k_2 is a better approximation than k_1.

Example 5.8.1 (Implementing RK2)
Let $y(t)$ represent the population (in thousands) of rabbits in a forest at year t. Suppose y is described by the differential equation

$$\frac{dy}{dt} = -t + 2y$$

with the initial condition $y(0) = 0.3$. To approximate the value of $y(3)$ using RK2 with $h = 0.1$, rename a blank workbook "**RK2**" and format it as in Figure 5.45. Copy row 6 down to row 35.

	A	B	C	D	E	F
1	**h =**	0.1				
2						
3						**RK2**
4	**Time**	**k$_1$**	**t + h/2**	**y + k$_1$/2**	**k$_2$**	**y**
5	0					0.3
6	=A5+$B$1	=$B$1*(-A5+2*F5)	=A5+$B$1/2	=F5+B6/2	=$B$1*(-C6+2*D6)	=F5+E6

FIGURE 5.45

It can be shown that the exact solution to this problem is given by $y(t) = 0.25(2t+1) + 0.05e^{2t}$. To compare RK2 to the exact solution and the results from Euler's method, add the formulas in Figure 5.46. Copy row 6 down to row 35.

Lastly, to compare the errors, add the graph in Figure 5.47.

We see that the exact value is $y(3) = 21.92143967$, RK2 gives $y(3) \approx 21.23789471$, and Euler's method gives $y(3) \approx 13.61881569$. RK2 yields smaller errors over the entire interval $[0,\, 3]$. Thus RK2 is more complicated than Euler's method, but it gives more accurate results.

□

	G	H	I	J
3	**Euler**	**Exact**	**RK2**	**Euler's**
4	**y**	**y**	**Error**	**Error**
5	0.3	0.3	=H5-F5	=H5-G5
6	=G5+$B$1*(-A5+2*G5)	=0.25*(2*A6+1)+0.05*EXP(2*A6)	=H6-F6	=H6-G6

FIGURE 5.46

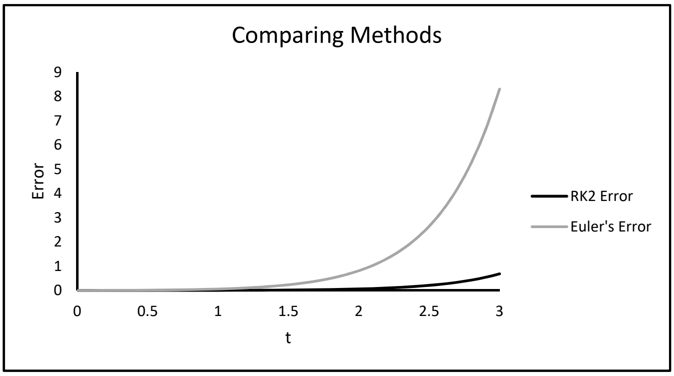


FIGURE 5.47

Runge-Kutta 4$^{\text{th}}$ Order Method

As mentioned above, k_2 in RK2 is an approximate change in y over the interval $[t_n,\, t_{n+1}]$ based on the midpoint of the interval. One could argue that we could get a better approximation by using the left endpoint, the right endpoint, and the midpoint, and then taking an average.

This idea leads to the following algorithm, called the *Runge-Kutta* 4^{th} *order method*, or simply *RK4*:

$$\begin{aligned}
k_1 &= h \cdot F\left(t_n,\, y_n\right) \\
k_2 &= h \cdot F\left(t_n + \frac{1}{2}h,\, y_n + \frac{1}{2}k_1\right) \\
k_3 &= h \cdot F\left(t_n + \frac{1}{2}h,\, y_n + \frac{1}{2}k_2\right) \\
k_4 &= h \cdot F\left(t_n + h,\, y_n + k_3\right) \\
t_{n+1} &= t_n + h \\
y_{n+1} &= y_n + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4
\end{aligned}$$

Informally, k_1 is an approximation of the change in y over the interval based on the left endpoint of the interval, k_2 and k_3 are approximations based on the midpoint, and k_4 is an approximation based on the right endpoint. The quantity $\frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4$ is a weighted average of these approximations.

Example 5.8.2 (Implementing RK4)
Consider Example 5.8.1. To approximate $y(3)$ using RK4 with $h = 0.1$, rename a blank worksheet "**RK4**," format it as in Figure 5.48 and copy row 6 down to row 35.

	A	B	C	D	E
1	**h =**	0.1			
2					
3					
4	**Time**	**k_1**	**t + h/2**	**y + k_1/2**	**k_2**
5	0				
6	=A5+$B$1	=$B$1*(-A5+2*J5)	=A5+$B$1/2	=J5+B6/2	=$B$1*(-C6+2*D6)

	F	G	H	I	J
3					**RK4**
4	**y + k_2/2**	**k_3**	**y + k_3**	**k_4**	**y**
5					0.3
6	=J5+E6/2	=$B$1*(-C6+2*F6)	=J5+G6	=$B$1*(-A6+2*H6)	=J5+B6/6+E6/3+G6/3+I6/6

FIGURE 5.48

We see that RK4 gives $y(3) \approx 21.92007319$, which is a better estimate of the exact value, $y(3) = 21.92143967$ than either Euler's method or RK2 give. This illustrates that even though RK4 is somewhat complicated, it gives much more accurate results.

□

Runge-Kutta 4$^{\text{th}}$ Order Method for Systems

Consider a system of two differential equations of the form

$$\frac{dx}{dt} = F(t, x, y)$$
$$\frac{dy}{dt} = G(t, x, y),$$

along with the initial conditions $x(t_0) = x_0$, $y(t_0) = y_0$. We can numerically approximate the solution to this system with the following algorithm, motivated by the RK4 algorithm from above:

$$k_1 = h \cdot F(t_n, x_n, y_n) \qquad l_1 = h \cdot G(t_n, x_n, y_n)$$
$$k_2 = h \cdot F\left(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_1, y_n + \frac{1}{2}l_1\right) \qquad l_2 = h \cdot G\left(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_1, y_n + \frac{1}{2}l_1\right)$$
$$k_3 = h \cdot F\left(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_2, y_n + \frac{1}{2}l_2\right) \qquad l_3 = h \cdot G\left(t_n + \frac{1}{2}h, x_n + \frac{1}{2}k_2, y_n + \frac{1}{2}l_2\right)$$
$$k_4 = h \cdot F(t_n + h, x_n + k_3, y_n + l_3) \qquad l_4 = h \cdot G(t_n + h, x_n + k_3, y_n + l_3)$$
$$x_{n+1} = x_n + \frac{1}{6}k_1 + \frac{1}{3}k_2 + \frac{1}{3}k_3 + \frac{1}{6}k_4 \qquad y_{n+1} = y_n + \frac{1}{6}l_1 + \frac{1}{3}l_2 + \frac{1}{3}l_3 + \frac{1}{6}l_4$$
$$t_{n+1} = t_n + h$$

Example 5.8.3 (Connected Tanks)
Consider the connected tanks in Example 5.4.1 described by the differential equations

$$\frac{dx}{dt} = \frac{1}{12}y - \frac{1}{3}x$$
$$\frac{dy}{dt} = \frac{1}{3}x - \frac{1}{3}y$$

with the initial conditions $x(0) = 4$ and $y(0) = 2$. To approximate $x(5)$ and $y(5)$ with RK4 using a step size of $h = 0.1$, rename a blank worksheet "**RK4 System**" and format it as in Figure 5.49. Copy row 6 down to row 55.

	A	B	C	D	E	F
1	**h =**	0.1				
2						
3						
4	**Time**	**k_1**	**l_1**	**t + h/2**	**x + k_1/2**	**y + l_1/2**
5	0					
6	=A5+$B$1	=$B$1*(R5/12-Q5/3)	=$B$1*(Q5/3-R5/3)	=A5+$B$1/2	=Q5+B6/2	=R5+C6/2

	G	H	I	J	K	L
4	**k_2**	**l_2**	**x + k_2/2**	**y + l_2/2**	**k_3**	**l_3**
5						
6	=$B$1*(F6/12-E6/3)	=$B$1*(E6/3-F6/3)	=Q5+G6/2	=R5+H6/2	=$B$1*(J6/12-I6/3)	=$B$1*(I6/3-J6/3)

	M	N	O	P	Q	R
4	**x + k_3**	**y + l_3**	**k_4**	**l_4**	**x**	**y**
5					4	2
6	=Q5+K6	=R5+L6	=$B$1*(N6/12-M6/3)	=$B$1*(M6/3-N6/3)	=Q5+B6/6+G6/3+K6/3+O6/6	=R5+C6/6+H6/3+L6/3+P6/6

FIGURE 5.49

From the worksheet, we see that $x(5) \approx 1.209623037$ and $y(5) \approx 1.926736014$ $\square$

Exercises

5.8.1 The reader might wonder what the "2" and the "4" stand for in RK2 and RK4. To answer this question, first note that as in Examples 5.8.1 and 5.8.1, RK2 and RK4 yield approximate values of the function. In these examples, we ended with approximations of $y(3)$. Now, define the *global truncation error* (GTE) as

$$\text{GTE} = |\text{exact} - \text{approximate}|\,.$$

The numbers in RK2 and RK4 stand for the *order of convergence*. Informally, a numeric approximation method is said to be k^{th} *order convergent* if dividing the step size h by some number M results in a GTE that is divided by approximately M^k, when h is "small enough." For instance with RK2, if we divide h by 2, the GTE should be divided by approximately 2^2. With RK4, the GTE should be divided by 2^4.

a. In Example 5.8.1 we approximated $y(3)$ with $h = 0.1$. Calculate the resulting GTE.

b. Redo Example 5.8.1 with $h = 0.05$, that is, divide h by 2. Calculate the resulting GTE.

c. Is the GTE from part b. approximately the GTE from part a. divided by 2^2?

d. Repeat parts a. and b. for RK4 in Example 5.8.2. Does dividing h by 2 result in a GTE divided by approximately 2^4?

e. Euler's method is 1^{st} order convergent. Repeat parts a. and b. for Euler's method in Example 5.8.1. Does dividing h by 2 result in a GTE divided by approximately 2^1?

5.8.2 In Example 2.3.3 we modeled a free-falling object with air resistance by assuming that the force due to air resistance is proportional to the velocity. This yielded the following differential equation for the velocity v:

$$m\frac{dv}{dt} = mg - kv.$$

The solution to this differential equation showed the terminal velocity is

$$\lim_{t\to\infty} v(t) = \frac{mg}{k}.$$

In certain cases, the force due to air resistance is proportional to v^r where $r > 1$ is some constant. This leads to the differential equation

$$m\frac{dv}{dt} = mg - kv^r,$$

a nonlinear equation, which is much more difficult to solve than a linear equation. The goal of this problem is to approximate the terminal velocity with RK4. Suppose $m = 1$, $g = 9.81$, $k = 2$, and $v(0) = 0$.

a. Use RK4 with a step size of $h = 0.2$ to approximate the terminal velocities for $r = 1.1$, 1.5, and 2.0.

b. If r increases, does the terminal velocity increase or decrease? Does this agree with your intuition? Why or why not?

5.8.3 Stefan's law of radiation states that the rate of change in temperature of a blackbody radiator at $T(t)$ degrees in a medium at $M(t)$ degrees is proportional to $M^4 - T^4$. That is,

$$\frac{dT}{dt} = k\left[(M(t))^4 - (T(t))^4\right]$$

where k is a constant. Suppose $k = 40^{-4}$, $M(t) = 70 + 15\cos(3.14t)$, and $T(0) = 100$. Use RK4 with a step size of $h = 0.1$ to graph the approximate solution curve over the interval $0 \le t \le 20$.

5.8.4 Consider a rocket blasting off straight upward from the surface of the earth with a constant thrust from the engines. As it accelerates upward, the rocket's velocity increases and its mass decreases from the burning of fuel. It can be shown that the rocket's velocity is described by the differential equation

$$m\frac{dv}{dt} - c\beta = -mg - kv$$

where

$\beta =$ burn rate of fuel	$m_0 =$ initial mass of rocket with fuel
$m(t) = m_0 - \beta t =$ mass of rocket at time t	$c =$ speed of exhaust gases relative to rocket
$k =$ constant measuring air resistance	$g =$ acceleration due to gravity (9.81 m/s^2)

The V-2 rocket that was used to attack London in WWII had an initial mass of 12,850 kg, a burn rate of 125.746 kg/sec, $c = 2{,}000$ m/s, $k = 1.45$ N per m/s, and $v(0) = 0$.

a. Plug these parameters into the differential equation and solve for dv/dt.

b. The engines burn for 70 seconds. Use RK4 with a step size of $h = 0.5$ to estimate the velocity at burnout.

5.8.5 This problem will illustrate that although RK4 is typically very accurate, it isn't perfect. Consider the differential equation

$$\frac{dy}{dt} = 5y - 6e^{-t}$$

with the initial condition $y(0) = 1$.

a. The exact solution to this differential equation is $y(t) = e^{-t}$. Graph this solution over the interval $0 \leq t \leq 5$.

b. Use RK4 to graph the approximate solution curve over the interval $0 \leq t \leq 5$ using step sizes of 0.1, 0.01, 0.005, and 0.001. Comment on the accuracy of these approximate curves.

Project Ideas

1. Model a guerrilla vs. guerrilla combat scenario.
2. Model the World War II battle of Iwo Jima with a Lanchester combat model.
3. Create a model of a traffic light to determine how long the light should remain yellow.
4. Create a worksheet to determine how long it will take an object in free-fall to hit the ground.
5. Create a worksheet to model the position of a parachutist. Consider different scenarios, one where the parachute opens instantly at some point during the fall, and another where the parachute takes a few seconds to fully open.
6. Use a quadratic population model to model the populations of wolves and elk in Yellowstone National Park.
7. Model a scenario involving three competing species.
8. Model a predator-prey-scavenger system.
9. Research the Solow economic growth model.
10. Model the velocity of a water rocket.
11. Model the trajectory of a baseball hit with a bat.
12. Model the motion of a damped harmonic oscillator.
13. Model a battle with reinforcements.
14. Model the blood alcohol content of a person over time while at a party.

15. Research numerical methods for approximating solutions to differential equations other than Euler's method, RK2, and RK4. Possible methods include Heun's method, middle point rule, Ralston's method, three eights method, RK5, and RK6.
16. Model pollution levels in the Great Lakes.
17. Devise a complicated system of connected tanks with valves and solve it.
18. Model the voltage across a capacitor in an RC circuit.
19. Analyze an SIR model where the susceptibles are broken into two categories: vaccinated and non-vaccinated. Assume the transmission coefficient for those vaccinated is lower than the non-vaccinated.
20. Model the flow of traffic through a stoplight.

For Further Reading

There are more books and articles written on differential equation models than any other type of model. Here are a few suggestions.

- For a good introduction to setting up and solving elementary differential equations, see Boyce, William E., and Richard C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, Seventh ed., John Wiley & Sons, 2001.
- For another good introduction, see Ledder, Glenn, *Differential Equations: A Modeling Approach*, McGraw–Hill, 2005.
- For applications of differential equations to a wide variety of scenarios, see *Modules in Applied Mathematics* Volume 1, *Differential Equation Models*, ed. William F. Lucas, Springer-Verlag, 1983.
- For several classic differential equation models, see Haberman, Richard, *Mathematical Models – Mechanical Vibrations, Population Dynamics, and Traffic Flow*, Society for Industrial and Applied Mathematics (SIAM), 1998.
- For more examples of differential equation models, see Dym, Clive L., *Principles of Mathematical Modeling*, Second edition, Elsevier, 2004.

6

Simulations

Chapter Objectives

- Define and motivate the idea of a simulation model
- Discuss ways of generating pseudorandom numbers
- Use density functions to model random events
- Model various scenarios with simulation models

6.1 Introduction

Mathematical modeling is all about describing systems (a system being a collection of components that operate together). Systems come in two general categories: *Deterministic* and *Probabilistic*. A deterministic system is one in which the behavior is determined once its parameters are set. A probabilistic system is one in which the behavior is determined, in part, by random events. Similarly, models can be put into these same two categories.

An example of a deterministic system is the area under a curve $y = f(x)$ over an interval $[a, b]$. The parameters of this system are the function $f(x)$ and the interval $[a, b]$. Once these parameters are set, the area is determined. Nothing else affects it. We can model this system with a deterministic model using elementary calculus:

$$\text{Area} = \int_a^b f(x)\, dx$$

Most real world systems are probabilistic. Inevitably a real world system involves some type of random event. Probabilistic systems are more difficult to model, so we typically treat them as if they were deterministic and create a deterministic model. Every model of a real world system we have created to this point in this book has been deterministic.

In this chapter we will introduce one very common type of probabilistic model, *simulation*. A simulation, in general, is any model that uses random numbers. Often simulations are used to imitate some type of real world behavior, but this does not have to be the case.

There are many reasons why one would construct a simulation model.

1. **System is far too complex to model analytically**. Consider a military air cargo transportation network. This system consists of many components including aircraft of different type, parking spots at airfields, fuel availability at airfields, different types of cargo, etc. It seems impossible to construct an analytical model that incorporates all of these components. Many simplifications would be needed, resulting in a very low-fidelity model.

2. **It may be difficult, costly, or dangerous to collect data for creating an empirical model**. Consider a hospital emergency room. If we were interested in modeling the waiting time of patients in terms of the number of doctors on staff, we could vary the number of doctors from week-to-week, collect data on waiting time, and construct an empirical model. This approach would certainly take much time and could result in dead patients. A much more practical approach is to simulate the behavior of the emergency room on a computer and vary the number of doctors in the simulation.

3. **The system may not exist yet**. Consider an aircraft on the drawing board. If we want to study the drag on the fuselage, we can't go fly it and collect data because it doesn't exist yet! We could simulate its behavior using a computer based on its design.

4. **System may contain random events that we do not want to over-simplify**. Consider a check-out line at a supermarket where customers arrive at an average of 2 per minute and the cashier can service an average of 3 customers per minute. If we wanted to model the waiting time of customers, we might be tempted to say that they won't have to wait at all because the service rate is greater than the arrival rate. This is a vast oversimplification. Instead, we will simulate the arrival and service of customers and analyze waiting time.

In this chapter, we focus on *Monte Carlo Simulations*. These simulations get their name from the fact that they are often used to study games of chance (such as those played in Monte Carlo). These simulations consist of three basic steps:

1. Construct a model that uses random numbers.

2. Evaluate, or "run," the model many times (possibly hundreds or thousands) using different random numbers each time.

3. Statistically analyze the results.

One advantage of using simulations to study real world behavior is that it allows the modeler to test "what–if" scenarios at very little cost. For example, in a simulation of a hospital emergency room, we could easily change the number of doctors or nurses on staff and observe the results. We could simulate many months or years of time in a matter of a few minutes at very little cost and no danger to anybody.

In this chapter we present a wide range of common types of simulation models and discuss some important topics related to the construction of simulation models.

6.2 Basic Examples

In this section we illustrate some of the basic concepts involved with a Monte Carlo simulation with three different examples.

Example 6.2.1 (Flipping a Coin)
Here we approximate the probability of getting at least 7 tails when a coin is flipped 10 times. In the terminology of elementary probability theory, flipping a coin 10 times is a *random experiment*, and getting at least 7 tails is an *event*. One way of estimating the probability of an event is to perform the random experiment many times, each time is called a *trial*, and count the number of times the event occurred, called the *number of successes*. We then calculate

$$P(\text{event}) \approx \frac{\text{Number of successes}}{\text{Total number of trials}}.$$

We apply this approach to estimate P(at least 7 tails).

Algorithm

a. Simulate 10 random flips of a coin and determine the number of tails obtained. This is one trial.

b. Repeat for 1000 trials.

c. Calculate the number of trials in which at least 7 tails were obtained. This is the number of successes.

d. Calculate $\text{P(least 7 tails)} \approx \dfrac{\text{Number of successes}}{1000}$.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet "**Coins**." Format the worksheet as shown in Figure 6.1. The formula in **B2** will select an integer between 0 and 1 with equal probability. An output of 1 indicates a Tail and a 0 indicates a Head. Copy the formulas in **B2** to the range **B2:K3** and then copy row 3 down to row 1001. (**Note:** If Excel returns the #NAME? error in cell **B2**, install and load the **Analysis ToolPak** add-in.)

	A	B	C	D	E	F	G	H	I	J	K	L
1	**Trial Number**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Total Tails**
2	1	=RANDBETWEEN(0,1)										=SUM(B2:K2)
3	=A2+1											=SUM(B3:K3)

FIGURE 6.1

2. Add the formula in Figure 6.2 to calculate the number of trials with at least 7 tails and the probability.

	N
1	**# successes**
2	=COUNTIF(L2:L1001,">=7")
3	**P(at least 7 tails)**
4	=N2/1000

FIGURE 6.2

Press the **F9** key several times to repeat these 1000 trials. Note that each time you press **F9**, you will probably get a different estimate of P(at least 7 tails). This illustrates that a simulation like this can only give an approximation of the true *theoretical* probability.

One benefit of a simulation like this is that we can easily modify it to consider different what-if scenarios. For instance, what if we use a *biased* coin where the probability of a tail on a single flip is different than 0.5? Add the cells in Figure 6.3 to store the probability of a tail on a single flip.

	N
6	**P(tail)**
7	0.4

FIGURE 6.3

Next, modify the formula in cell **B2** as shown in Figure 6.4, and copy **B2** to the range **B2:K1001**. Now we can easily set $P(\text{tail})$ to any value between 0 and 1 and estimate $P(\text{at least 7 tails})$.

	B
2	=IF(RAND()<$N$7,1,0)

FIGURE 6.4

Next we create a table that stores estimate of $P(\text{at least 7 tails})$ for different values of $P(\text{tail})$. Add the formulas in Figure 6.5 to begin setting up this table. Copy row 4 down to row 23 to calculate values of $P(\text{tail})$ between 0 and 1 in increments of 0.05.

	P	Q
1	**P(tail)**	**P(at least 7 tails)**
2		=N4
3	0	
4	=P3+0.05	

FIGURE 6.5

Next, highlight the range **P2:Q23**, and select **Data** → **What-If Analysis** → **Table...** Select **N7** as the **Column input cell:**, leave the **Row input cell:** blank, and press **OK**. Here's what this does:

1. The first number in the left column of the table (0 in this case) is "pasted" into the cell **N7**. This causes the random numbers to regenerate, which means another 1000 trials are run with $P(\text{tail}) = 0$. The results from the simulation are displayed in the cell **Q2** and then copied to the cell next to the 0 (**Q3**).

2. The next number in the first column of the table (0.05 in this case) is pasted into the cell **N7**. The simulation is run again and the results are copied to cell **Q4**.

3. This process is repeated until the bottom of the table is reached.

We can create a graph of the results as in Figure 6.6.

□

Example 6.2.2 (Area Under a Curve)
Here we construct a probabilistic model of a deterministic system. Specifically, we use a simulation to estimate the area under the curve $y = \sqrt{1 - x^2}$ over the interval $[-1, 1]$. This curve forms the top-half of a circle with radius 1, so the area under the curve is exactly $\pi/2$. This simulation could be seen as a way of estimating the value of π.

A graph of the curve is shown in Figure 6.7 along with a rectangle of height $h = 1$ and width $w = 2$ drawn around it.

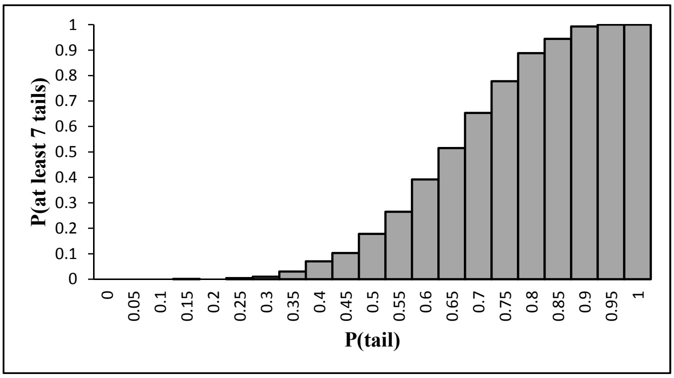


FIGURE 6.6

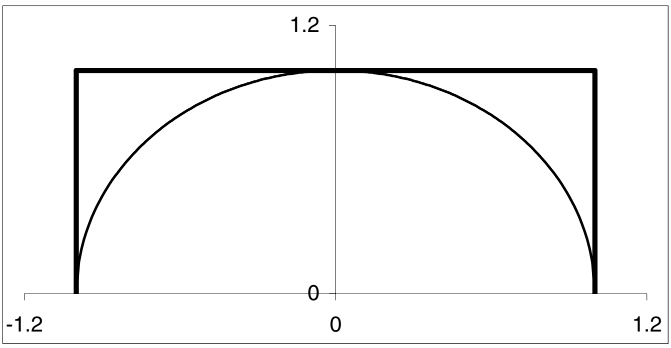


FIGURE 6.7

In the simulation, we randomly pick points inside the rectangle and determine if each one is above or below the curve. We then estimate the area under the curve using the relationship

$$\frac{\text{Area under the curve}}{\text{Area of the rectangle}} \approx \frac{\text{Number of points under the curve}}{\text{Total number of points}}$$

or equivalently,

$$\text{Area under the curve} \approx \frac{\text{Number of points under the curve}}{\text{Total number of points}}(\text{Area of the rectangle}) \quad (6.1)$$

Algorithm

a. Randomly pick 200 points inside the rectangle.

b. Determine if each point lies under the curve.

c. Count the number of points under the curve.

d. Use (6.1) to estimate the area under the curve. This is one trial.

e. Repeat for a total of 200 trials.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet "**Area**" and format it as in Figure 6.8. The parameters a and b are simply the lower and upper limits of the interval of interest. Copy row 6 down to row 205. The formula in cell **C6** tests whether the y–coordinate of the point lies under the curve $y = \sqrt{1 - x^2}$ which forms the top half of a circle. If it does, then the formula returns a 1. If not, it returns a 0.

	A	B	C
1	**a =**	-1	
2	**b =**	1	
3	**h =**	1	
4			
5	**x**	**y**	**Under Curve?**
6	=$B$1+RAND()*($B$2-$B$1)	=RAND()*$B$3	=IF(B6<SQRT(1-A6^2),1,0)

FIGURE 6.8

2. Add the formulas in Figure 6.9 to estimate the area. Press the **F9** key several times to repeat the simulation. Note that the estimated area fluctuates quite a bit from repetition to repetition, but the estimates are near the actual area of approximately 1.5708.

	E
1	**# Points Under the Curve**
2	=SUM(C6:C205)
3	**Area Under the Curve**
4	=E2/200*(B2-B1)*B3

FIGURE 6.9

3. To repeat the simulation 200 times, format the spreadsheet as in Figure 6.10 and copy row 4 down to row 202.

	G	H
1	**Trial #**	**Area**
2		=E4
3	1	
4	=G3+1	

FIGURE 6.10

4. Next, we calculate a data table in the range **G2:H202** as done in the flipping a coin simulation. Select **F1** as the column input cell and leave the row input cell blank. In this case, the trial number is not a meaningful parameter, so we simply paste it into a blank cell (we arbitrarily chose **F1**). Each time the trial number is pasted, the simulation is run again, and the result is stored in the table.

5. Add the formulas in Figure 6.11 to calculate the average of the 200 trials. Press **F9** several times and note the variation of the average. The average in general is closer to the true area than the single trial case. The average also has less variation.

	E
5	**Average**
6	=AVERAGE(H3:H202)

FIGURE 6.11

This example illustrates a very important point when analyzing results from simulations:

The more trials, the closer the average value is to the theoretical value.

The point is that if you want to estimate something with a simulation, you will get a more accurate estimate if you use many trials and take an average than if you use only a few trials. This concept is called the *Law of Large Numbers*.

□

Example 6.2.3 (Car Dealership Contest)
A car dealership is sponsoring a contest where the grand prize is a new car. Contests are to gather tickets which contain the letter "C," "A," or "R" from participating merchants. To win, one must obtain all three letters. 55% of the tickets contain a "C", 44% contain an "A", and 1% contain an "R." What is the expected number of tickets a contestant must gather to win the car?

Algorithm

a. Randomly generate 500 tickets.

b. Keep running totals of the number of each letter obtained.

c. Determine if at least one of each letter has been obtained.

d. Repeat for a total of 500 trials.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet "**Car**." Format the worksheet as shown in Figure 6.12. Copy row 6 down to row 504 to simulate 500 tickets. The letter on the ticket is determined by breaking up the interval (0, 1) into three sub-intervals of lengths 0.55, 0.44, and 0.01. Then we generate a random number between 0 and 1 with the RAND function. If this number is in the first sub-interval, the letter is an "A," etc.

	A	B	C
3		**Random**	
4	**Ticket**	**Number**	**Letter**
5	1	=RAND()	=IF(B5<0.55,"C",IF(B5<0.99,"A","R"))
6	=A5+1	=RAND()	=IF(B6<0.55,"C",IF(B6<0.99,"A","R"))

FIGURE 6.12

2. To keep running totals of the number of each letter obtained, add the formulas in Figure 6.13. Copy row 6 down to row 504.

	D	E	F
4	**Total C's**	**Total A's**	**Total R's**
5	=IF(C5="C",1,0)	=IF(C5="A",1,0)	=IF(C5="R",1,0)
6	=IF(C6="C",1,0)+D5	=IF(C6="A",1,0)+E5	=IF(C6="R",1,0)+F5

FIGURE 6.13

3. To determine when the contest is won and how many tickets were needed, add the formulas in Figure 6.14. Copy row 8 down to row 504. On the ticket the contest is won, "WIN" is displayed. If the contest should go on, a blank is displayed. If the contest is already over, "NA" is displayed. The total number of tickets needed is displayed in cell **G2**. This is the result of one trial.

	G
1	**Number of Tickets**
2	=COUNTBLANK(G5:G504)+1
3	
4	**Win?**
5	
6	
7	=IF(AND(D7>=1,E7>=1,F7>=1),"WIN","")
8	=IF(G7="",IF(AND(D8>=1,E8>=1,F8>=1),"WIN",""),"NA")

FIGURE 6.14

4. Lastly, to store the results of 500 trials and calculate the overall average, add the formulas in Figure 6.15. Copy row 5 down to row 503. Then create a data table in the range **I3:J503**, selecting any blank cell as the column input cell.

	H	I	J
1	**Average # Tickets**		**Number**
2	=AVERAGE(J4:J503)	**Play**	**of Tickets**
3			=G2
4		1	
5		=I4+1	

FIGURE 6.15

The simulation shows that the expected number of tickets needed to win the contest is just over 100.

□

Exercises

6.2.1 Modify the formulas in the worksheet **Coins** so there are a total of 5000 trials. Press **F9** several times. Compare the variability of the value of P(at least 7 tails) with 1000 trials as compared to 5000 trials.

6.2.2 Modify the worksheet **Coins** to approximate the probability that the number of Tails is between 4 and 7, inclusive.

6.2.3 Suppose you roll a fair four-sided die three times. Design a spreadsheet to estimate the probability that

a. the sum of the three rolls is at least 7,

b. at least one of the rolls is greater than 2, and

c. the first roll or the third roll is even.

(**Suggestion:** To determine whether a trial is a success for probabilities b. and c., use the **OR** function which works much like the **AND** function in Figure 6.14.)

6.2.4 Consider the area under a curve model in Example 6.2.2.

a. Increase the value of h in the simulation. How does this affect the quality of the estimation of the area?

b. Increase the number of points selected inside the rectangle. How does this affect the estimate?

c. Create a graph of the curve $y = \sqrt{1 - x^2}$ and the rectangle along with a dot for each randomly selected point, similar to Figure 6.16.

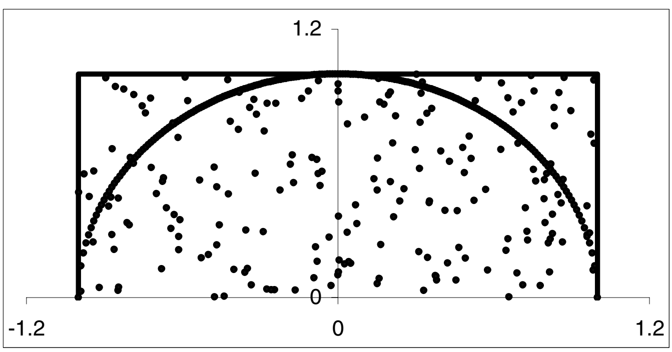


FIGURE 6.16

6.2.5 Design a simulation to estimate the value of $\int_0^{0.5} xe^x \sin(50x) + 1\, dx$.

6.2.6 A square is constructed such that the length of a side is randomly chosen between 2 and 3 inches (not necessarily an integer). Design a simulation to estimate the probability that the area of the square is between 5 and 6 in^2.

6.2.7 Consider the following game:

In each round, two numbers, call them N_1 and N_2, are chosen. In the first round, $N_1 = 1$, in the second, $N_1 = 2$, and so on until round 11 when $N_1 = 1$. This pattern repeats until the game is over. In each round, the number N_2 is a randomly chosen integer between 1 and 10, inclusive. The game is over when $N_1 = N_2$.

a. Simulate this game to estimate the expected number of rounds in the game.

b. Generalize the simulation by replacing the number 10 with any positive integer N. Allow the user to enter the value of N.

6.2.8 Consider the following game:

You choose an integer y between 1 and 20, inclusive. Then the dealer randomly chooses a number x, not necessarily an integer, between 0 and 10. The amount of money you "win" is given by $z = -x^2 + 8x - y^2 + 24y - 175$ (if $z > 0$, you get that many dollars, if $z < 0$, you pay that many dollars).

Simulate this game and determine the value of y you should choose to maximize the average winnings from many plays of the game. Consider using a data table which stores the average winnings for different values of y.

6.2.9 Consider the following coin-flipping game:

1. A single play of the game consists of repeatedly flipping a fair coin until the *difference* between the number of heads tossed and the number of tails is 4.

2. You are required to pay \$1 for each flip of the coin, and you may not quit during the play of the game.

3. You receive \$10 at the end of each play of the game.

The "winnings" from the game is defined as the \$10 received at the end minus the amount paid.

a. Simulate this game to estimate the expected winnings from many plays of the game.

b. Suppose we use a biased coin. Find value(s) of $P(\text{tail})$ that make the game fair, meaning the expected winnings is \$0.

6.2.10 Ally and Bernita are playing a simple game of tennis where the player who gets to 4 points first wins the game. Suppose Ally has a 60% probability to win any given point, and points cannot end in a tie.

a. Use a simulation to estimate Ally's probability of winning.

b. Now suppose the winner is the first player to score at least 4 points *and* score at least 2 more points that the opponent. Estimate Ally's probability of winning. Assume the game could go on indefinitely.

6.2.11 Consider the following game:

A card is selected from a well-shuffled standard deck of 52 cards. The player guesses the suite (4 possibilities) and the rank (13 possibilities) of the card. If only the suite is correct, the player receives \$2. If only the rank is correct, the player receives \$5. If both the suite and rank are correct, the player receives \$10. If neither is correct, the player receives nothing. Each play costs \$1. After each play the card is returned to the deck and the deck is shuffled.

Simulate this game to estimate the expected winnings.

6.2.12 Suppose a fair coin is flipped repeatedly.

a. Estimate the expected number of flips until the sequence HTH is observed. Repeat for HTT and HHH.

b. Now suppose the coin is flipped exactly 300 times. Estimate the expected number of times the sequence HTH is observed. Repeat for HTT and HHH.

6.3 Three Famous Problems

In this section we show how simulations can be used to approximate the solutions to three famous problems in elementary probability: the *Monty Hall Problem*, the *Birthday Problem*, and *Buffon's Needle Problem*.

Example 6.3.1 (The Monty Hall Problem)
In the famous game show *Let's Make a Deal*, hosted by Monty Hall, one of the games required a contestant to choose one of three doors. Behind one of the doors was a prize (like money or a car), and behind the other two doors were dummy prizes (like a donkey). Once the choice was made, Monty Hall would open up one of the un-chosen doors revealing one of the dummy prizes. The contestant was then given a choice to either switch to the remaining unopened door or keep the door that was already chosen. The contestant would get whatever "prize" was behind the door.

To illustrate this game, consider the scenario in Figure 6.17 where the real prize is behind door 2 (unbeknownst to the contestant) and the contestant chooses door 1. Monty would then open door 3 revealing a dummy prize. The contestant then had to decide whether to switch to door 2 (and consequently win the real prize) or don't switch (and consequently not win the real prize).

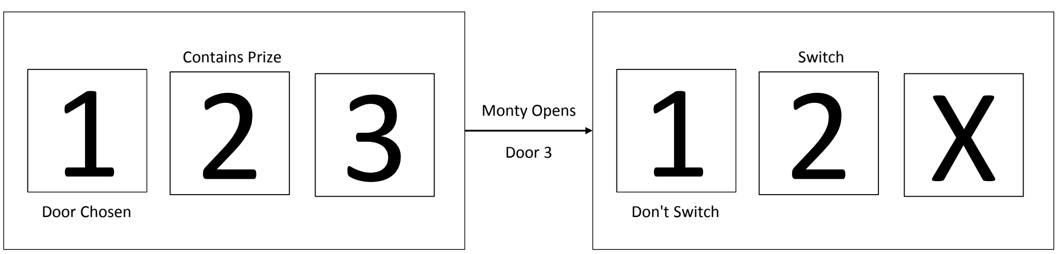


FIGURE 6.17

The dilemma facing the contestant is the decision of whether to switch or not. To help make this decision, we will use a simulation to estimate the following two probabilities:

1. Probability of winning if switching and

2. Probability of winning if not switching.

The key to estimating these probabilities is to realize that if not switching, the only way to win is to initially choose the door with the real prize. If not switching, the prize is won if the door with the real prize is not initially chosen.

Algorithm

a. Randomly designate the door with the real prize.

b. Randomly choose a door.

c. Determine if the real prize is won if switching.

d. Determine if the real prize is won if not switching.

e. Repeat for 1,000 trials.

f. Calculate $P(\text{win if switch}) \approx \dfrac{\text{number of times won when switching}}{\text{number of trials}}$.

g. Calculate $P(\text{win if don't switch}) \approx \dfrac{\text{number of times won when not switching}}{\text{number of trials}}$.

Rename a blank worksheet "**Monty**" and format it as in Figure 6.18. Copy row 8 down to row 1007 to perform 1000 trials.

	A	B	C	D	E
1		**Probabilities**			
2		**Win if**	**Win if don't**		
3		**switch**	**switch**		
4		=AVERAGE(D8:D1007)	=AVERAGE(E8:E1007)		
5					
6		**Door**	**Door**	**Win if**	**Win if don't**
7	**Trial**	**with prize**	**chosen**	**switch?**	**switch?**
8	1	=RANDBETWEEN(1,3)	=RANDBETWEEN(1,3)	=IF(B8<>C8,1,0)	=IF(B8=C8,1,0)

FIGURE 6.18

We see from the simulation that the probability of winning if switching is 2/3 and the probability of winning if not switching is 1/3. Thus the contestant should switch. This does not guarantee that the contestant will win, but the contestant is twice as likely to win if switching than if not switching. □

Example 6.3.2 (The Birthday Problem)
In a class of n students, what's the probability that at least two students will share a birthday (month and day)? This famous problem is known as the *birthday problem*. We assume that birthdays are uniformly distributed throughout the year (i.e. no day is more or less likely to be a birthday than any other day) and we ignore leap years.

Algorithm

a. Randomly generate an integer between 1 and 365 for each student in the class to represent birthdays ($1 =$ January 1, $2 =$ January 2, etc.).

b. For each day of the year, count the number of students in the class that have that day as their birthday.

c. Determine if some birthday is shared by at least two students. This is considered a success.

d. Repeat for 200 trials.

e. Determine the number of successes.

f. Calculate $P\,(\text{at least two people sharing a birthday}) \approx \dfrac{\text{number of successes}}{\text{number of trials}}$.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet "**Birthday**" and format it as in Figure 6.19. Copy row 3 down to row 102 to simulate a class of up to 100 students.

	A	B
1	**# Students =**	23
2	**Student**	**Birthday**
3	1	=IF(A3<=$B$1,RANDBETWEEN(1,365),0)

FIGURE 6.19

2. Add the formulas in Figure 6.20 and copy row 4 down to row 367. These formulas count the number of students who have a birthday on each day of the year and determine whether the trial is a success.

	D	E
1	**Success?**	=IF(COUNTIF(E3:E367,">=2")>=1,1,0)
2	**Day**	**Count**
3	1	=COUNTIF($B$3:$B$102,D3)
4	=D3+1	=COUNTIF($B$3:$B$102,D4)

FIGURE 6.20

3. Add the formulas in Figure 6.21 to set up a table to store the results of 200 trials and calculate the estimated probability. Copy row 6 down to row 204. Create a table in the range **G4:H204** to store the results from 200 trials. Select **F1** as the column input cell. Press **F9** to repeat the simulation several times. Note that for a class of 23 students the simulation gives a probability of approximately 0.50.

	G	H
1	**Probability**	=AVERAGE(H5:H204)
2		
3	**Trial**	**Success?**
4		=E1
5	1	
6	=G5+1	

FIGURE 6.21

One benefit of using a simulation is that we can easily modify it to estimate more complicated probabilities. For instance, if we wanted to estimate the probability that at least 3 students share a birthday in a class of 50 students, we could simply change the value of "# Students" to 50 and modify the formula to determine whether the trial is a success as in Figure 6.22.

□

	G	H
1	**Probability**	=AVERAGE(H5:H204)
2		
3	**Trial**	**Success?**
4		=E1
5	1	
6	=G5+1	

FIGURE 6.22

Example 6.3.3 (Buffon's Needle Problem)
A version of this problem was first solved by the French naturalist and mathematician, the Comte de Buffon (1707-1788). Suppose we randomly drop a needle of length $L \leq 1$ on a wood floor in which the joints between the planks are 1 unit apart. Find the probability that the needle "hits," or intersects, one of the joints.

Let x denote the distance of the midpoint of the needle to the nearest joint between the planks and θ denote the angle as illustrated in Figure 6.23. Note that $0 \leq x \leq 0.5$ and $0 \leq \theta \leq \pi/2$.

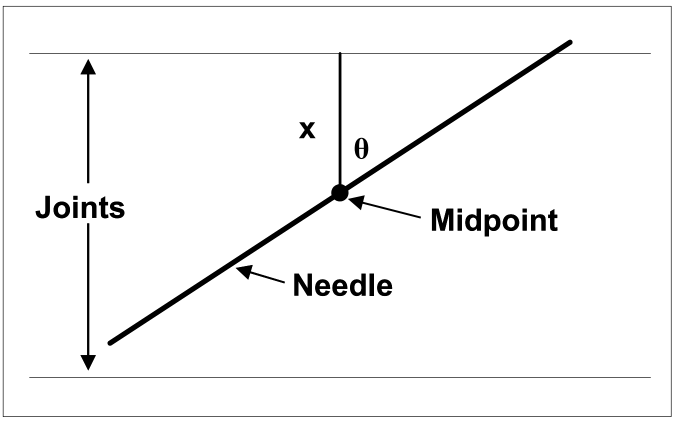


FIGURE 6.23

To determine whether the needle intersects a joint, observe that this can only occur if the hypotenuse of the right triangle in Figure 6.23 is less than $L/2$ (if the needle does not intersect a joint, then the needle must be "extended" to form the triangle, making the hypotenuse more than $L/2$). If h denotes the length of the hypotenuse, then using trigonometry we have

$$\cos\theta = \frac{x}{h} \quad \Rightarrow \quad h = \frac{x}{\cos\theta}.$$

Thus the needle intersects a joint if $x/\cos\theta < L/2$.

Algorithm

a. Randomly generate a number x between 0 and 0.5.

b. Randomly generate a number θ between 0 and $\pi/2$.

c. Calculate $h = x/\cos\theta$.

d. Determine if $h \leq L/2$. If this is true, then the trial is a success.

e. Repeat for 1000 trials.

f. Calculate $P\,(\text{intersecting a joint}) \approx \dfrac{\text{number of successes}}{\text{number of trials}}$.

Rename a blank worksheet "**Buffon**" and format it as in Figure 6.24. Adjust the value of L and observe that as L gets smaller, so does the probability.

	A	B	C	D	E
1	**L =**	1	**Probability =**	=AVERAGE(E4:E1003)	
2					
3	**Trial**	**x**	**θ**	**h**	**Hit Joint?**
4	1	=0.5*RAND()	=PI()/2*RAND()	=B4/COS(C4)	=IF(D4<=$B$1/2,1,0)

FIGURE 6.24

□

Exercises

6.3.1 Consider the Monty Hall problem.

a. In the simulation, the "door chosen" was random. What if the contestant always chooses door 1? Does this change either of the probabilities of winning? What if door 2 is always chosen? What about door 3?

b. Consider the following strategy:

> Always initially choose door 1. If door 2 is opened, then switch. Otherwise don't switch.

Create a simulation to estimate the probability of winning if this strategy were used every time. Assume that if the prize is behind door 1 then doors 2 and 3 are equally likely to be opened. Does this strategy do any better than switching every time?

6.3.2 Consider a generalization of the Monty Hall problem where there are N doors, exactly one of which contains the real prize. Assume the game is played the following way:

1. The contestant always initially chooses door 1.

2. If the door with the prize is door 1 or door N, then door 2 is opened. In this case, if the contestant switches, the door switched to is randomly selected between 3 and N, inclusive.

3. If the door with the prize is not door 1 or door N, then door N is opened. In this case, if the contestant switches, the door switched to is randomly selected between 2 and $N-1$, inclusive.

Create a simulation of this game where the user can input the value of N and estimate the probabilities of winning if switching and if not switching. (**Suggestion**: Create a column to determine the door to switch to. Use this to determine whether the game is won if a switch is made.)

a. Create a table which gives the probabilities for values of N between 3 and 25.

b. What happens to the difference between the probabilities of winning if switching and the probability if not switching as N gets larger? For "large" values of N, does it really matter if the contestant switches or not?

6.3.3 Consider the birthday problem and let P_n denote the probability that in a class of n students, at least 2 students share a birthday. It can be shown that P_n satisfies the following recursive relationship:

$$P_n = 1 - (1 - P_{n-1}) \cdot \frac{N - (n-1)}{N}$$

where N is the number of days in a year and $P_1 = 0$.

a. Create a spreadsheet to calculate the values of P_n for n between 1 and 100. Allow the user to enter the value of N. If $N = 365$, what is the value of P_{23}? Does this agree with the results of the simulation?

b. Graph P_n vs. n when $N = 365$. For what values of n is $P_n > 0.90$?

c. Use the relationship to find the probability that in a class of 6 students, at least 2 students share a birth *month*.

6.3.4 Modify the worksheet **Birthday** to estimate the solution to the following generalization of the birthday problem: If a teacher asks a class of n students to write down an integer between a and b, what's the probability that at least m of them will write down the same number.

1. Your simulation should allow the user to input the values of n, a, b, and m, and automatically calculate the results.

2. Use a total of 500 trials.

3. Assume that $n \leq 200$, $0 \leq a < b \leq 365$, and that the students' choices are uniformly distributed.

4. If the value of m is in cell **E1**, then consider modifying the formula to determine a success as in Figure 6.25.

	D	E
2	**Success?**	=IF(COUNTIF(E4:E368,">="&E1)>=1,1,0)

FIGURE 6.25

6.3.5 For Buffon's needle problem, create a table which gives the probability of an intersection for different values of L between 0 and 1. Create a graph of the probability vs. L, fit a curve to the data, and use the curve to hypothesize the theoretical relationship between the probability and L. (**Hint**: It is a very simple relationship involving the number $2/\pi$.)

6.3.6 Two numbers x and y are randomly chosen from the interval (0, 1). Design a simulation to estimate the probability that the closest integer to y/x is even. (**Hint**: Calculate y/x and round it off to a whole number using the **ROUND** function. To determine if this number is even, use the **MOD** function.)

6.3.7 Suppose we randomly choose 50 pairs of numbers (x, y), each number from the interval $(0, 1)$. Let ϵ be a small positive number. Design a simulation to estimate the probability that for at least one pair, the ratio x/y or y/x is less than ϵ from the golden ratio (a number approximately equal to 1.618 and denoted ϕ). Allow the user to input the value of ϵ.

6.3.8 Consider a mouse in a maze consisting of three connected rooms. When in each room, the mouse performs the following actions:

Room 1

- With probability 1/3, wander around for 2 minutes and then exit the maze.
- With probability 2/3, wander around for 5 minutes and then move to room 2.

Room 2

- With probability 1/2, move to room 1.
- With probability 1/2, move to room 3.

Room 3

- With probability 1, wander around for 3 minutes and then move to room 2.

Suppose the mouse starts in room 2. Create a simulation to estimate the expected amount of time spent in the maze.

6.3.9 Another famous problem in elementary probability is the *drunkard's walk*: Suppose a drunk man stands one step from a cliff and then takes a sequence of steps. In each step, he moves one step closer to the cliff with probability p and one step farther from the cliff with probability $(1 - p)$.

a. Design a simulation to estimate the probability the man falls off the cliff. Allow the user to enter the value of p.

b. Create a graph $P(\text{falling})$ vs. p for various values of p between 0 and 1. What, if anything, do you note of interest?

6.3.10 Consider a circle of radius 1 centered at the origin. Suppose 3 points are randomly selected on the circle. What is the probability the resulting triangle contains the origin? Design a simulation to estimate this answer. Here are some suggestions:

a. Generate three random angles between 0 and 2π. These angles determine the polar coordinates of the three points.

b. Sort the angles from smallest to largest using the **SMALL** function. These sorted angles determine points A, B, and C, respectively, as in Figure 6.26. Calculate the x- and y-coordinates of each point.

c. Consider the lines AB, BC, and CA. As illustrated in Figure 6.26, the triangle contains the origin if one of these lines has an x-intercept in the interval $(-1, 0)$, and another has an x-intercept in $(0, 1)$. Find the x-intercept of each line.

d. Count the number of x-intercepts in the interval $(-1, 0)$, and the number in $(0, 1)$. Use this to determine if the triangle contains the origin.

e. Repeat this for many trials and summarize the results.

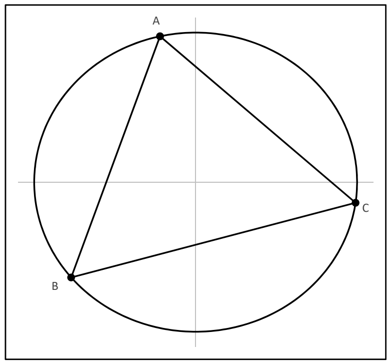


FIGURE 6.26

6.4 The Poker Problem

Each card in a standard deck of 52 cards is described by its rank (also called the *face value*), ace, 2, 3, . . ., 10, jack, queen, or king; and its suit, hearts, diamonds, clubs, or spades. Five cards are randomly dealt from the deck. A classic problem in elementary probability theory is calculating the probability of obtaining the different types of poker hands. Table 6.1 lists the ten different types of poker hands. In this section we build a simulation to estimate the probability of getting a full house. In the exercises you will estimate the probabilities of the other hands.

TABLE 6.1

	Name	**Description**
0.	No Value	None of the following types
1.	One Pair	Four distinct ranks, one rank occurs twice, each of the other three occurs once
2.	Two Pair	Three distinct ranks, two ranks occur twice each, one occurs once
3.	Three of a Kind	Three distinct ranks, one rank occurs three times, two ranks occur once each
4.	Straight	Five distinct ranks in sequential order, at least two distinct suits
5.	Flush	One suit, but the ranks are not arranged in sequential order
6.	Full House	Two distinct ranks, one rank occurs three times, one occurs twice
7.	Four of a Kind	Two distinct suits, one suit occurs four times, one occurs once
8.	Straight Flush	One suit, the ranks are arranged in sequential order, but not a royal flush
9.	Royal Flush	One suit, the ranks are in the sequential order 10, J, Q, K, A

Algorithm

a. Assign a number to each card in the deck.

b. Shuffle the deck.

c. Select the first five cards as our hand.

d. Determine the number of each suit in the hand.

e. Determine whether the hand is a full house.

f. Repeat for 5000 trials.

g. Calculate $P\,(\text{full house}) \approx \dfrac{\text{number of successes}}{\text{number of trials}}$.

To implement this algorithm, rename a blank workbook **Full House** and add the formulas in Figure 6.27 to assign a number to each card. Copy cell **A3** down to row 53 for a total of 52 cards. The ranks should go in order 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, and A. This sequence should restart with card 14. The first 13 cards are of suit Spades, the next 13 are Clubs, then Diamonds, and finally Hearts.

	A	B	C
1	**Number**	**Rank**	**Suit**
2	1	2	Spades
3	2	3	Spades

FIGURE 6.27

Next we need to shuffle the deck and select the first five cards as our hand. We might consider doing this by choosing random numbers between 1 and 52 with the **RANDBETWEEN** function, but then we might get the same card twice. (In the language of probability theory we want to select five cards *without replacement*.) Instead, we will choose 52 random numbers with the **RAND** function, rank them from 1 to 52, and let the sequence of ranks represent the shuffled deck. Add the formulas in Figure 6.28. Copy row 2 down to row 53.

	E	F
1	**Rand**	**Card**
2	=RAND()	=RANK(E2,$E$2:$E$53)

FIGURE 6.28

Next we need to convert the cards to rank and suit. Add the formulas in Figure 6.29. The **VLOOKUP** function in columns **G** and **H** will look at the left-most column in columns **A**–**C**, find the number that matches the card, and return the corresponding value in column **B** or **C**. Copy row 2 down to row 53.

	G	H
1	**Rank**	**Suit**
2	=VLOOKUP(F2,$A$2:$C$105,2)	=VLOOKUP(F2,$A$2:$C$105,3)

FIGURE 6.29

Next we need to count the number of each rank in the hand. Add the formulas in Figure 6.30. The ranks should go in order 2, 3, 4, 5, 6, 7, 8, 9, 10, J, Q, K, and A. Copy cell **K2** down to row 14.

	J	K
1	**Rank**	**Frequency**
2	**2**	=COUNTIF($G$2:$G$6,J2)

FIGURE 6.30

A full house could be described as getting a three-of-a-kind and a two-of-a-kind. Add the formulas in Figure 6.31 to determine if the hand is a full house.

	J	K
16	**3 of a Kind?**	=COUNTIF(K2:K14,3)
17	**# Pairs**	=COUNTIF(K2:K14,2)
18	**One Pair?**	=IF(K17=1,1,0)
19	**Full House?**	=IF(AND(K16=1,K18=1),1,0)

FIGURE 6.31

Lastly, add the formulas in Figure 6.32 to store the results of 5000 trials and calculate the estimated probability. Create a data table in the range **M2:N5002**, selecting any blank cell as the column input cell.

	M	N	O
1	**Trial**	**Result**	**P(Full House)**
2		=K19	=AVERAGE(N3:N5002)
3	1		

FIGURE 6.32

The theoretical value of $P(\text{full house})$ is 0.0014. Our estimate of the probability is in the neighborhood of 0.0014. Notice that each time the **F9** key is pressed, we get a different estimate of the probability. These estimates vary quite a bit in part because full houses occur so infrequently that even in 5000 trials, we don't observe a full house very many times.

Exercises

6.4.1 Estimate $P(\text{one pair})$, $P(\text{two pair})$, and $P(\text{three of a kind})$. (**Hint**: When determining if a hand is a one pair or a three of a kind, make sure it's not a full house.)

6.4.2 Estimate $P(\text{straight})$. Here are some suggestions:

a. A straight means, in part, that the hand contains 5 ranks in sequential order. Add a column to determine if the hand contains each individual rank.

b. For each rank, calculate the total number of that rank and the next 4 higher ranks in the hand. The hand contains 5 ranks in sequential order if one of these totals is 5.

c. An ace could be the high card or low card in the sequential order. Determine if this order has been obtained with ace as the low card.

d. A straight also means that the hand contains at least 2 distinct suits. Count the number of each suit in the hand. Determine if the hand contains at least two distinct suits.

6.4.3 Estimate $P(\text{flush})$. (**Hint**: A flush means the hand contains only one suit, but it's not a straight.)

6.4.4 Estimate $P(\text{four of a kind})$, $P(\text{straight flush})$, and $P(\text{royal flush})$.

6.4.5 Consider a simple variation of a poker game where just 4 cards are dealt from a well-shuffled deck instead of the usual 5. In this variation there are only two types of hands:

- **Two pair**: 2 cards of one rank and 2 cards of a different rank
- **All suits**: 4 cards of different suits.

Use a simulation to estimate the probability of getting a hand that is *both* two pair and all suits.

6.4.6 The procedure used to shuffle the deck can be used to simulate any scenario where selections are made without replacement. Consider a bag containing five red and three blue marbles. The bag is shaken and two marbles are chosen without replacement. Design a simulation to estimate the following probabilities.

a. $P(2 \text{ red})$

b. $P(2 \text{ blue})$

c. $P(1 \text{ red and } 1 \text{ blue})$

6.5 Random Number Generators

In this section we look at how computers generate "random" numbers. Random numbers are an essential part of all computer simulations. A simple definition of a *list of random numbers* is that it is a list of numbers in which there is *no* pattern and all possible numbers occur with equal frequency. The only way to get a truly random list of numbers is by mechanical means (i.e. numbered balls tumbling in a cage, rolling a die, etc.).

A computer generates a list of random numbers by using an iterative function where one output becomes the next input. The initial input, called the *seed*, is arbitrary (it is often chosen according to the clock time at which the algorithm begins) and each output becomes a number in the list.

Because the computer uses a deterministic algorithm, there *will* be a pattern to the list of numbers. Therefore, computers can never generate a true list of random numbers. The random numbers they generate are called *pseudorandom* numbers and the algorithm is called a *pseudorandom number generator*. A good list of pseudorandom numbers will, at the very least, have a pattern that is not at all obvious.

Excel has a built-in pseudorandom generator called **RAND** which generates numbers between 0 and 1. First we create a graph of numbers generated by this function to illustrate what we mean by "no pattern." Rename a blank worksheet "**Rand**," add the formulas in Figure 6.33, and copy row 2 down to row 501 to form a list of 500 pseudorandom numbers.

	A	B
1	**n**	**Random**
2	1	=RAND()

FIGURE 6.33

Create a graph of Random vs. n as in Figure 6.34. Push **F9** several times to create new lists of pseudorandom numbers. Note that the graph does not reveal any sort of relationship between the pseudorandom number and its position in the sequence and no number (or range of numbers) appears more frequently than any other (i.e. there is no pattern). There are many types of statistical tests that can be done to measure the randomness of a list of numbers. We will not discuss these here, but our simple graphical analysis indicates that **RAND** creates a good list of pseudorandom numbers.

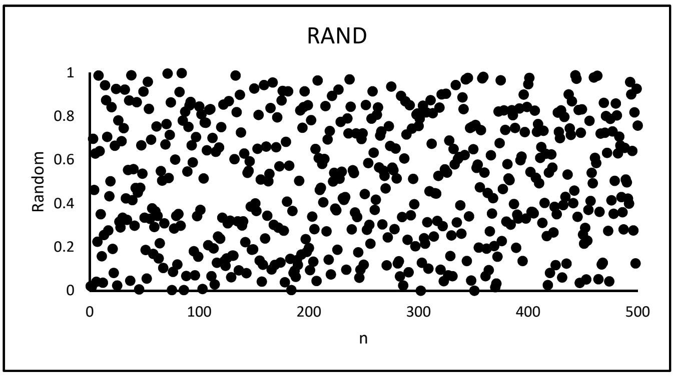


FIGURE 6.34

Because pseudorandom number generators form the heart of simulations and many security systems, much research has been done into developing and testing generators. Here we present some simple generator algorithms.

Example 6.5.1 (Linear Congruence)
This method uses modular arithmetic to generate pseudorandom integers. Three integers, a, b, and m, are chosen along with a seed x_0. Random integers are then generated using the function:

$$x_{n+1} = (a \cdot x_n + b) \bmod (m)$$

This method will produce a sequence of up to m integers between 0 and $m-1$, inclusively, before repeating, or *cycling*. For this reason, m is generally a very large integer such as 2^{32} or 2^{64}.

To implement this method using the values of $a = 1$, $b = 3$, $x_0 = 9$, and $m = 8$, rename a blank worksheet "**Linear**," and add the formulas in Figure 6.35. Copy row 6 down to row 25 to create a list of 20 pseudorandom numbers.

Figure 6.36 shows graphs of the lists of pseudorandom numbers for $a = 1$ and $a = 3$. Note how the lists start to repeat after $n = 8$ and $n = 4$, respectively. This is NOT what we want in a good list of pseudorandom numbers.

	A	B
1	**a =**	1
2	**b =**	3
3	**m =**	8
4	**n**	**x_n**
5	0	7
6	=A5+1	=MOD($B$1*B5+$B$2,$B$3)

FIGURE 6.35

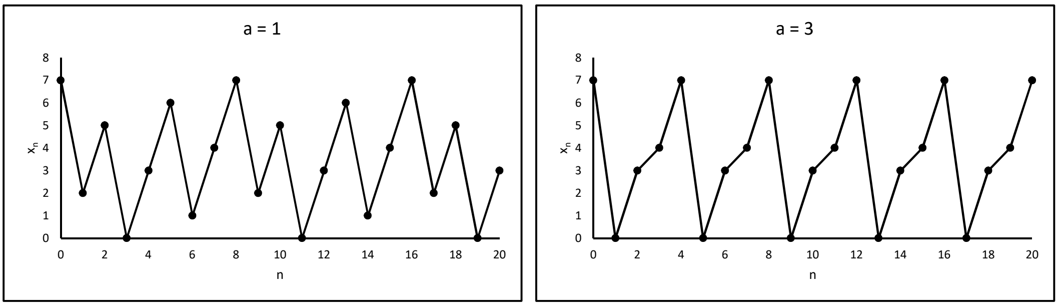


FIGURE 6.36

□

Example 6.5.2 (Old Excel Algorithm)
This algorithm was used by the **RAND** function in older versions of Excel and generates a list of up to 1 million different numbers [Microsoft Help and Support webpage, http://support.microsoft.com/kb/q86523/, June 2008.]:

1. x_0 = an arbitrary number between 0 and 1
2. x_{n+1} = fractional part of $(9821 * x_n + 0.211327)$

To implement this algorithm with $x_0 = 0.5$, rename a blank worksheet "**Old Rand**," add the formulas in Figure 6.37, and copy row 3 down to row 502.

	A	B
1	**n**	**x_n**
2	0	0.5
3	=A2+1	=MOD(9821*B2+0.211327,1)

FIGURE 6.37

A graph of x_n vs. n is shown in Figure 6.38. There does not appear to be any pattern. Other values of x_0 give similar results.

□

Example 6.5.3 (New Excel Algorithm)
The old Excel algorithm was sufficient for "casual" users (i.e. those who needed fewer than 1 million pseudorandom numbers). However, it did not pass a standard battery of tests for randomness named Diehard, so it was not sufficient for "power" users.

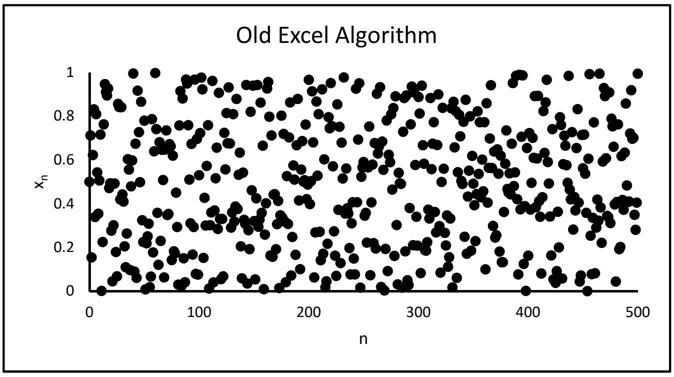


FIGURE 6.38

A new algorithm was developed that produces a list of up to 10^{13} numbers before cycling. It is a variation on the Linear Congruence method and is based on the idea "that if you take three random numbers on [0, 1] and sum them, the fractional part of the sum is itself a random number on [0,1]" [Microsoft Help and Support webpage http://support.microsoft.com/default.aspx?scid=kb;en-us;828795, June 2008].

The algorithm to generate a list of pseudorandom numbers between 0 and 1, $\{x_n : n = 0, 1, \ldots\}$, is given by:

a. Set a_0, b_0, and c_0 to integer values between 1 and 30,000

b. x_n = fractional part of $(a_n/30269 + b_n/30307 + c_n/30323)$

c. $a_{n+1} = (171 \times a_n) \bmod (30269)$

d. $b_{n+1} = (172 \times b_n) \bmod (30307)$

e. $c_{n+1} = (170 \times c_n) \bmod (30323)$

To implement this algorithm, rename a blank worksheet "**New Rand**," add the formulas in Figure 6.39, and copy row 3 down to row 502.

	A	B	C	D
1	**n**	**a_n**	**b_n**	**c_n**
2	0	15843	16235	9842
3	=A2+1	=MOD(171*B2,30269)	=MOD(172*C2,30307)	=MOD(170*D2,30323)

	E
1	**x_n**
2	=MOD(B2/30269+C2/30307+D2/30323,1)
3	=MOD(B3/30269+C3/30307+D3/30323,1)

FIGURE 6.39

A graph of x_n vs. n is shown in Figure 6.40. Again note that there does not appear to be any pattern.

□

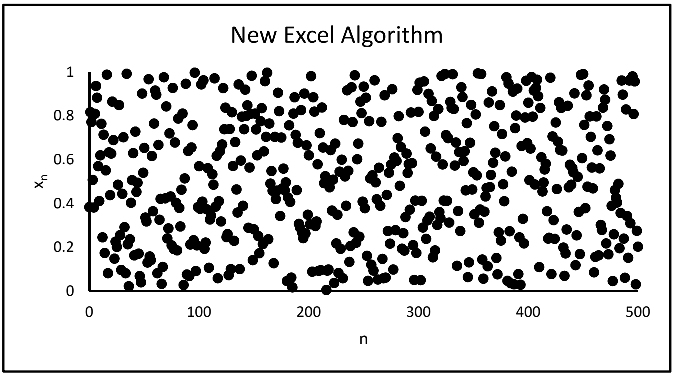


FIGURE 6.40

Exercises

6.5.1 Suppose we want to generate a list of pseudorandom numbers each of which has the value a, b, or c where a occurs $p_1 \times 100\%$ of the time, b occurs $p_2 \times 100\%$ of the time, and c occurs $p_3 \times 100\%$ of the time where $p_1 + p_2 + p_3 = 1$.

a. Design a spreadsheet which generates this list. You may use the **RAND** function. Make sure the user is able to input the values of a, b, c, p_1, p_2, and p_3.

b. Use the **COUNTIF** function to verify that the list contains the proper percentage of each number a, b, and c.

6.5.2 In each part below, use the **RANDBETWEEN** function to generate a list of equally likely pseudorandom numbers with the given specifications.

a. Even integers between 0 and 20, inclusive.

b. Numbers between 0 and 20, inclusive, with 2 decimal places.

c. Numbers from the set $\{-1, 1\}$. (**Hint**: Use **RANDBETWEEN** to choose between 2 numbers. Use the **IF** function to output a -1 for one of the numbers and a 1 for the other.)

d. Numbers from the set $[-10, -5] \cup [5, 10]$ with 2 decimal places.

6.5.3 Consider the following function for generating pseudorandom integers between the integers a and b, where $a < b$:

$$x = [\textbf{RAND} \cdot (b - a) + a] \text{ rounded to the nearest integer.}$$

a. Use this function to generate a list of 1000 integers. Allow the user to input the values of a and b.

b. Count the number of times each integer between a and b, inclusive, appears in the list. Do all the integers appear with nearly equal frequency?

c. Suggest a modification to this function so that it generates a list in which the integers between a and b, inclusive, appear with nearly equal frequency. Use your modification to generate another list of 1000 integers. Does the modification generate a list with nearly equal frequencies? Do not use the **RANDBETWEEN** function.

6.5.4 Another pseudorandom number generator is the *middle-square method*. The basic algorithm is as follows:

1. Start with an n-digit number x_0, called the *seed* (n is typically even).
2. Square x_0 to obtain an $2n$-digit number (add leading zero(s) if necessary).
3. Take the middle n digits as the next random number.

Implement this method to generate a list of 100 4-digit pseudorandom numbers. Use RANDBETWEEN to generate the seed. How well does this generator work? (**Suggestion**: After x is squared, first "chop" off the first two digits, then chop off the last two digits.)

6.5.5 Suppose a pseudorandom number generator gives integers between 0 and 9. One way to test if this generator gives integers with equal frequency is to apply a $\boldsymbol{\chi^2}$ *Goodness–of–fit test*. This can be done by generating a long list of integers (say 500) and then counting the number of times each integer appears in the list. These are called the *observed frequencies*, denoted by O. We then calculate the *expected frequencies*, denoted by E, which is the number of times we expect each integer to appear in the list if they do indeed occur with equal frequencies. If we have 10 different integers in a list of 500, we would expect each one to appear 50 times.

Then we calculate the test statistic χ^2 by

$$\chi^2 = \sum \frac{(O-E)^2}{E}$$

If this test statistic is "small" (for 10 different integers small is less than 16.9) then we can be 95% confident that the generator gives integers with equal frequencies. If it is large then we reject the claim that it gives integers with equal frequencies. (For a more detailed description of this test, see any introductory statistics textbook.)

a. The Excel function **RANDBETWEEN** gives pseudorandom integers between two specified values. Use it to generate a list of 500 integers between 0 and 9. Calculate the χ^2 test statistic as described above and press **F9** several times to get several different lists of integers. Does the **RANDBETWEEN** appear to give integers with equal frequencies?

b. Use the linear congruence algorithm with $a = 6$, $b = 9$, $m = 10$, and $x_0 = 2$ to generate a list of 500 integers between 0 and 9. Calculate the χ^2 test statistic. Does this linear congruence algorithm appear to give integers with equal frequencies? Try different values of a, b, and x_0. What do you observe?

6.6 Modeling Random Variables

Simulation is a useful tool for modeling the interaction of *random events*. A random event is an activity where we do not know the outcome until it occurs. Constructing a simulation involves modeling random events. One of the most important concepts used in modeling random events is the *random variable*.

Definition 6.6.1 (Random Variables)**.** A *random variable* is a rule for assigning real numbers to the observations of a random event. A *discrete random variable* can take only certain distinct values (such as integers). A *continuous random variable* can take any value within some interval.

In most cases, defining the random variable involved is rather obvious. For instance, suppose we roll a standard six–sided die, and we define the random variable X as the number on the top face of the die. This is an example of a discrete random variable. In another example, suppose we observe customers arriving at a check–out line at the grocery store. We define the random variable Y as the time between customer arrivals (called the inter-arrival times). This is an example of a continuous random variable.

In a simulation of a dice game, we need to generate values of the roll of the dice. In a simulation of a check–out line, we need to generate values of the inter–arrival times. In other words, simulations involve generating values of random variables. In this section we discuss how to do this using the **RAND** function.

We focus on continuous random variables. One of the most important tools used to model continuous random variables is the *density function*. Any function $f(x)$ is a density function if it satisfies the following two properties:

1. $f(x) \geq 0$ for all $x \in R$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1$

For a given random variable X, its density function $f(x)$ is used to calculate probabilities regarding the value of X by

$$P(X \leq a) = \int_{-\infty}^{a} f(t)\,dt \tag{6.2}$$

This means that the probabilities are related to the area under the graph of the density function. Related to the density function is the *cumulative distribution function* (*cdf*), $F(x)$, defined by

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

From Equation (6.2) we see that $F(x) = P(X \leq x)$. In terms of the graph of $f(x)$, $F(x)$ is the area under the curve $y = f(t)$ to the left of x.

Many functions could be density functions. We single out three important types of density functions that occur frequently in applications.

Example 6.6.1 (Uniform Density Function)
If a random variable X, $a \leq X \leq b$, has a uniform density function (we say X is *uniformly distributed*), the values are "spread out evenly" between a and b. The **RAND** function gives values of a pseudorandom variable that is uniformly distributed between 0 and 1. The density function of a uniformly distributed random variable is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise.} \end{cases}$$

A graph of this density function is shown in Figure 6.41.

The mean of a random variable X with a uniform distribution is $(b+a)/2$ and the standard deviation is $(b-a)/\sqrt{12}$.

□

Example 6.6.2 (Normal Density Function)
A normally distributed random variable X has a density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

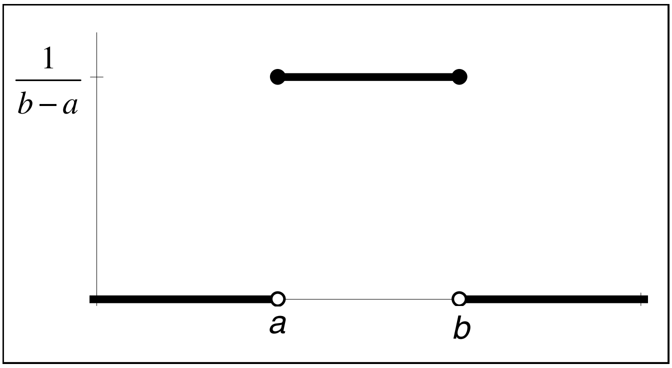


FIGURE 6.41

where μ is the mean and σ is the standard deviation. The graph of this density function is the familiar bell curve shown in Figure 6.42. In the graph we see that the variable X has a higher probability of taking values near the mean μ than farther away.

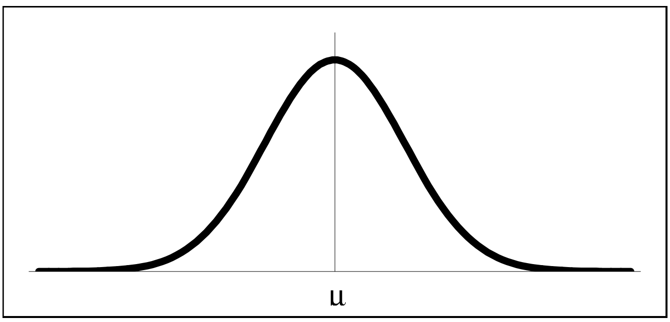


FIGURE 6.42

□

Example 6.6.3 (Exponential Density Function)
An exponential density function is often used to model waiting time between events. It has the form

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

where $\lambda > 0$ is a parameter. A graph of this density function is shown in Figure 6.43. In the graph we see that the variable X has a higher probability of taking values near 0 than those much larger.

The mean and standard deviation of a random variable with an exponential distribution are both equal to $1/\lambda$.

□

Once we know the density function, $f(x)$, for a random variable, X, we can generate values of it using the **RAND** function with the following general algorithm:

1. Find the cdf $y = F(x) = \int\limits_{-\infty}^{x} f(t)\,dt$
2. Find the inverse of the cdf, $x = F^{-1}(y)$

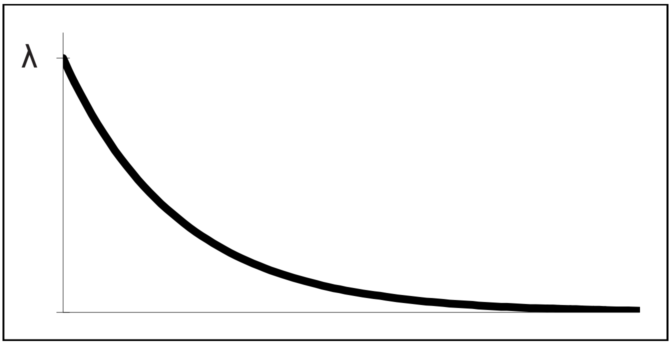


FIGURE 6.43

3. Use the **RAND** function to generate values of y and calculate a value of X by $x = F^{-1}(y)$.

Example 6.6.4 (Generating Values of an Exponential Distribution)
To illustrate this algorithm, consider generating values of an exponentially distributed random variable X.

Step 1: Find the cumulative distribution function

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} \lambda\, e^{-\lambda\, t} dt = 1 - e^{-\lambda\, x}$$

Step 2: Set $F(x) = y$ and solve for x.

$$y = 1 - e^{-\lambda x} \quad \Rightarrow \quad x = -\frac{1}{\lambda}\ln(1-y) \quad \Rightarrow \quad F^{-1}(y) = -\frac{1}{\lambda}\ln(1-y)$$

Step 3: Let $y = \textbf{RAND}$, so the formula is:

$$x = -\frac{1}{\lambda}\ln(1 - \textbf{RAND}) \tag{6.3}$$

Formula (7.5) can be implemented in Excel very easily. Rename a blank worksheet "**Exponential**" and format it as in Figure 6.44. Copy row 3 down to row 1002 to generate a list of 1000 values of this random variable.

	A	B
1	**λ =**	2
2	**x**	
3	=-1/$B$1*LN(1-RAND())	

FIGURE 6.44

□

In Example 6.6.4 we generated 1000 values of an exponentially distributed random variable. The calculations were rather easy to do, but they beg at least two questions:

1. What does it really mean to generate values of an exponentially distributed random variable?

2. How can we check if these calculations work the way we intended?

We can answer these questions by creating a very important type of graph called a *histogram*.

Example 6.6.5 (Generating a Histogram)
Generating histograms is rather easy to do in newer versions of Excel. Here we create and analyze a histogram of the 1000 values generated in Example 6.6.4. Follow these steps:

1. In the worksheet **Exponential**, select the range B3:B1003.

2. In the **Insert** ribbon, choose **Insert Statistic Chart** and select **Histogram** as shown in Figure 6.45.

FIGURE 6.45

3. Right-click on the horizontal axis of the resulting graph. Select **Format Axis**. Under **Axis Options**, change the **Number of bins** to 8. Under **Number**, select **Category** and **Number**, and set the number of decimal places to 2. The resulting graph should resemble Figure 6.46.

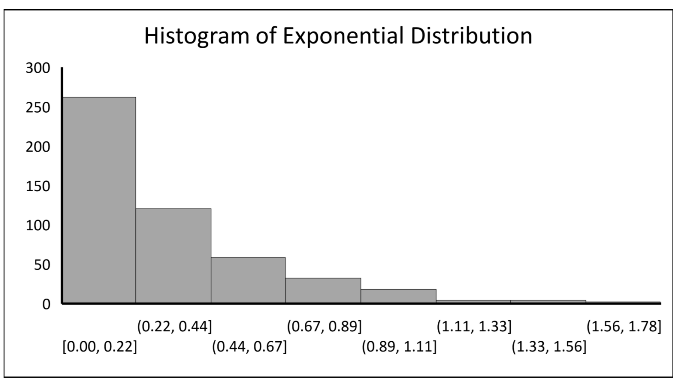


FIGURE 6.46

Excel generated this histogram by dividing the range of values into 8 *bins*, also called *subintervals* or *classes*, and then counting the number of values in each bin. The graph shows, for instance, that about 275 of the values are between 0 and 0.27. Larger values are less frequent. Most importantly, note that the "shape" of the histogram resembles the graph of the exponential density function in Figure 6.43.

The histogram answers our two questions:

1. Generating values of an exponentially distributed random variable means that we get many values close to 0. Larger values occur less frequently.

2. The calculations work if the shape of the histogram resembles the graph of the density function.

□

Example 6.6.6 (Generating Values of a Uniform Random Variable)
It can be shown that the inverse CDF of a uniformly distributed random variable is $F^{-1} = a + (b - a)y$ (see Exercise 6.6.1) is

$$F^{-1} = a + (b - a)y.$$

Rename a blank worksheet "**Uniform**" and format it as in Figure 6.33. Copy row 4 down to row 1003 to generate a list of 1000 values of this random variable.

	A	B
1	**a =**	5
2	**b =**	10
3	**x**	
4	=$B$1+($B$2-$B$1)*RAND()	

FIGURE 6.47

Generate a histogram of these values as in Figure 6.48. Note that all values occur with relatively equal frequency, and the shape of the histogram resembles the graph of the uniform density function in Figure 6.41.

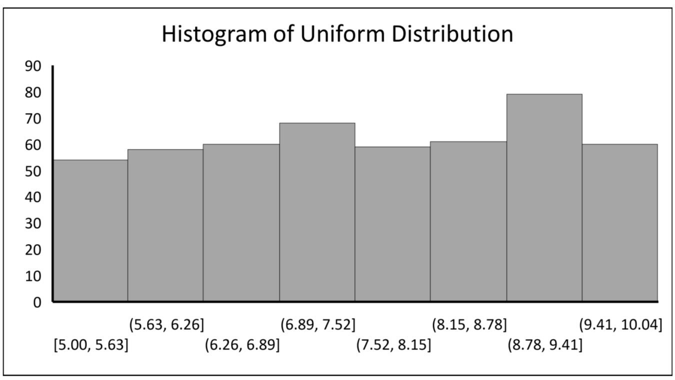


FIGURE 6.48

□

Example 6.6.7 (Generating Values of a Normal Random Variable)
We can derive a formula for generating values of a normally distributed random variable by calculating F and F^{-1}, but these calculations are very complicated. Fortunately, F^{-1} is built into Excel. Rename a blank worksheet "**Normal**" and format it as in Figure 6.49. Copy row 4 down to row 1003 to generate 1000 values of this random variable.

A histogram of these values is shown in Figure 6.50. The shape resembles the normal distribution bell curve, graphically confirming the formulas work correctly.

□

	A	B
1	μ =	0
2	σ =	1
3	x	
4	=NORM.INV(RAND(),$B$1,$B$2)	

FIGURE 6.49

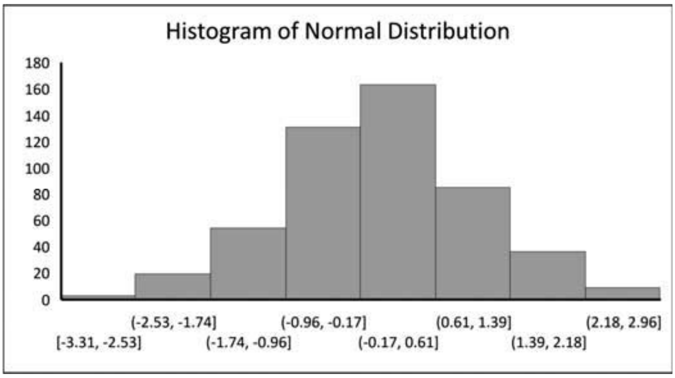


FIGURE 6.50

Example 6.6.8 (Application to Real Data)
Let's suppose we want to simulate the activity at a local grocery store check–out line. One random event we need to model is the arrival of customers. To do this, we observe 30 customers arriving at the check–out line and record the time between their arrivals (in minutes) as shown in Table 6.2.

TABLE 6.2

1.40	2.79	0.91	1.87	0.87	0.21	0.10	0.96	0.92	0.47
1.60	1.76	3.46	1.51	3.90	5.75	0.90	0.66	1.56	2.74
0.03	0.36	0.21	2.36	3.24	0.22	1.33	0.04	0.33	1.84

In our simulation, we need a density function to describe this random variable. Earlier we claimed that an exponential distribution is often used to describe the waiting time between events. We can easily test this claim by drawing a histogram of these data as shown in Figure 6.51. We see the histogram resembles the graph of the exponential density function, confirming the claim.

To generate values of this exponentially distributed random variable, we need to know the appropriate value of λ. Since the mean of a random variable with an exponential distribution is $1/\lambda$, we have $\lambda = 1/\text{mean}$. This is easily calculated using the formula in Figure 6.52 yielding $\lambda = 0.6772$.

□

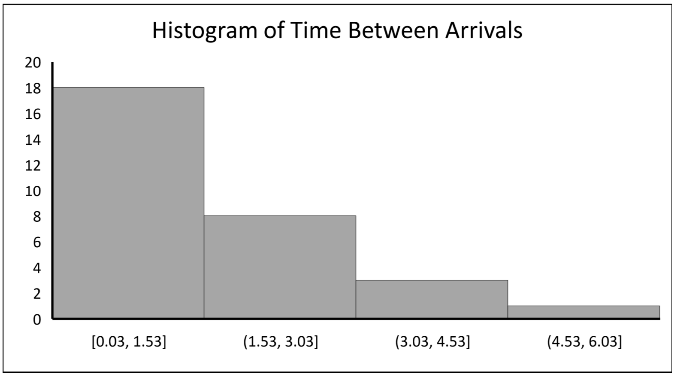


FIGURE 6.51

	C	D
2	**λ =**	=1/AVERAGE(A2:A31)

FIGURE 6.52

Exercises

6.6.1 Show that the inverse CDF of a uniformly distributed random variable is $F^{-1} = a + (b-a)y$.

6.6.2 A random variable X has the density function

$$f(x) = \begin{cases} 1/\left(2\sqrt{x}\right) & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}.$$

a. Find the cumulative distribution function, $F(x)$.

b. Find the inverse cumulative distribution function, $F^{-1}(y)$.

c. Use the inverse cumulative distribution function to generate 100 values of X.

d. Graph the density function over the interval $0 < x < 1$

e. Create a histogram of the 100 values generated in part c. Does the shape of the histogram resemble the graph of the density function?

6.6.3 Repeat Exercise 6.6.2 with the density function

$$f(x) = \begin{cases} x^3/4 & \text{for } 0 < x < 2 \\ 0 & \text{elsewhere} \end{cases},$$

but graph the density function over the interval $0 < x < 2$

6.6.4 Repeat Exercise 6.6.2 with the density function

$$f(x) = \begin{cases} 2(1-x) & \text{for } 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}.$$

(**Hint**: Use the quadratic formula to solve for x in terms of y.)

6.6.5 Create a graph of the exponential density function over the interval $0 \le x \le 10$. Use a scroll bar to vary the value of λ. Describe what happens to the shape of the graph as λ changes.

6.6.6 Create a graph of the normal density function over the interval $\mu - 3\sigma \le x \le \mu + 3\sigma$. Use scroll bars to vary the values of μ and λ. Describe what happens to the shape of the graph as μ and λ change.

6.6.7 Suppose we want to simulate the packaging of 16–ounce packages of carrots. One random variable involved is the actual weight of the packages. We measure and record the weight (in ounces) of 30 packages as shown in the table below.

15.99	16.18	16.16	16.38	16.10	16.38	16.11	16.43	16.21	16.36
16.28	16.34	16.20	16.21	16.27	16.08	16.15	16.31	16.06	16.06
16.26	16.10	15.94	16.52	16.16	16.01	16.26	16.19	16.30	16.08

Graphically determine if this random variable has a normal distribution. If so, estimate the mean μ and the standard deviation σ. (**Hint**: To estimate the standard deviation use the formula **STDEV**.)

6.6.8 Suppose we want to simulate the activity at a gas station. One random variable involved is the daily demand of gasoline. Given below are the daily demands (in thousands of gallons) on 30 randomly selected days at three different gas stations. For each gas station, determine which distribution, uniform, exponential, or normal, best models the random variable. For your choice of distribution, give appropriate value(s) of the parameter(s) (a and b for uniform, λ for exponential, and μ and σ for normal). Briefly explain your answers.

a.

3.44	2.88	0.25	6.04	0.25	3.07	1.88	5.62	3.74	5.41
1.15	7.82	1.06	1.85	0.30	1.40	9.93	0.17	1.31	2.49
2.78	7.10	3.19	4.28	1.51	1.11	1.06	1.27	1.69	1.46

b.

16.22	15.89	15.45	17.50	18.77	12.14	14.97	18.84	17.93	18.5
12.39	15.99	14.64	14.17	15.73	10.56	17.60	13.44	15.17	14.34
13.42	18.97	12.86	11.93	13.65	16.51	16.15	14.72	14.74	12.32

c.

19.87	22.89	22.10	20.49	21.19	18.27	20.94	16.25	16.29	14.17
22.60	24.13	19.79	16.65	23.53	21.51	24.03	15.17	14.29	17.27
18.64	24.87	21.62	15.69	21.34	14.81	21.03	22.18	20.99	20.78

6.6.9 Suppose the manager of a small board game store would like to simulate the activity in his store. One random variable involved is the daily sales. The manager looks over his records from the past 930 days and summarizes the daily sales data in the table below. (The first row, for instance, says that on 5 days the sales were between \$1000 and \$1099.)

Sales	Frequency
1000-1099	5
1100-1199	15
1200-1299	45
1300-1399	100
1400-1499	200
1500-1599	250
1600-1699	170
1700-1799	80
1800-1899	35
1900-1999	30

a. What type of distribution does this random variable appear to have? Briefly explain your reasoning.

b. The mean and standard deviation of summarized data such as these can be approximated using the formulas

$$\bar{x} \approx \frac{1}{n} \sum (f \cdot x) \quad \text{and} \quad s \approx \sqrt{\frac{n \sum (f \cdot x^2) - \left[\sum (f \cdot x)\right]^2}{n(n-1)}}$$

where n is the total number of data values (the sum of the frequencies), x is the midpoint of an interval, and f is the corresponding frequency. Use these formulas to approximate the mean and standard deviation.

6.6.10 In the worksheet **Exponential** we generated values of an exponential random variable with the formula $x = -\frac{1}{\lambda} \ln (1 - \textbf{RAND})$. Suppose we used the formula $x = -\frac{1}{\lambda} \ln (\textbf{RAND})$ instead. Try using this formula. Does it change the shape of the resulting histogram? Explain why or why not.

6.6.11 One of the most important theorems in probability theory is the *central limit theorem*. Informally, this theorem states that if we generate several independent values of a random variable and calculate the mean of these values, the mean is a value of a random variable with an approximately normal distribution regardless of the distribution of the original random variable. The purpose of this exercise is to graphically demonstrate this theorem is true.

a. Generate 30 values of a random variable with an exponential distribution (choose your own value of the parameter λ) and calculate the mean of these 30 values.

b. Repeat step a. 250 times.

c. Generate a histogram of the 250 values of the mean. Does this histogram resemble a bell curve?

d. Change the value of the parameter to several different values. What do you note about the shape of the histogram?

e. Repeat steps a. - d. for an original random variable with a normal distribution. Then repeat for a uniform distribution.

f. The central limit theorem also says that the mean of the means should approximately equal the mean of the original random variable. Also, the standard deviation of the means should approximately equal the standard deviation of the original random variable divided by the square root of the sample size (the sample size is 30 in this exercise). Check if these two points are true for each distribution.

g. What does all this say about the validity of the central limit theorem?

6.6.12 As mentioned in Section 6.5, Diehard is a battery of tests for random number generators. One of these tests is called the *3-D spheres test*. To illustrate the basic idea behind this test, consider the following simplified version:

1. Use the RANDBETWEEN function to generate 10 points in 3-D space inside a cube with sides of length 100. That is, generate 10 points of the form (x, y, z) with each coordinate is an integer between 0 and 100. (The real 3-D spheres test uses 4,000 points inside a sphere with sides of length 1,000.)
2. Calculate the Euclidean distance between each pair of points. This yields 45 distances.
3. Let $r =$ minimum of the distances in step 2. Calculate the volume of a sphere with radius r.
4. Repeat steps 1-3 a total of 500 times to collect 500 volumes. Calculate the average volume and generate a histogram of the volumes using 8 bins.

The average and distribution found in step 4 are the benchmarks by which a pseudo-random number generator is judged. If a generator does not yield the same average and distribution, it fails the test.

a. Perform steps 1-4. What is the average volume? What type of distribution does the volume appear to have?

b. Apply the test to the linear congruence generator described in Example 6.5.1 with parameters $a = 99$, $b = 3$, and $m = 101$. That is, perform steps 1-4, but generate the coordinates using the linear congruence generator. Use the RANDBETWEEN function to generate the seed x_0. Does this generator pass the test?

c. Now apply the test to the linear congruence generator with parameters $a = 54$, $b = 70$, and $m = 101$. Does this generator pass the test?

d. Try the parameters $a = 54$, $b = 68$, and $m = 101$. Does this generator pass the test?

e. Suppose we were to generate more points (like 4,000 in the real 3-D spheres test). Would this increase or decrease the average volume found in step 4? Explain.

6.7 A Theoretical Queuing Model

In simple terms, a *queue* is a waiting line. A prototypical example is people standing in line to buy movie tickets. Another example is an assembly line where bottles of soda wait to be filled with liquid. There are two important entities involved in a queue: *customers* and *servers*. Servers are whatever are used to process customers. In the movie ticket line, a customer is a person wanting to buy a ticket, and a server is a person selling tickets.

Suppose you arrive at a movie ticket booth to buy a ticket and see a long line of people waiting. Two questions might come to mind:

1. How long must I expect to wait?
2. What is the average number of people waiting in line?

In this section we build a simulation of a single-server queue to help answer these questions. Figure 6.53 illustrates the arrival of the first three customers. The arrival time of each customer is random and so is the service time. The wait time and service start time are calculated based on the completion of the previous customer.

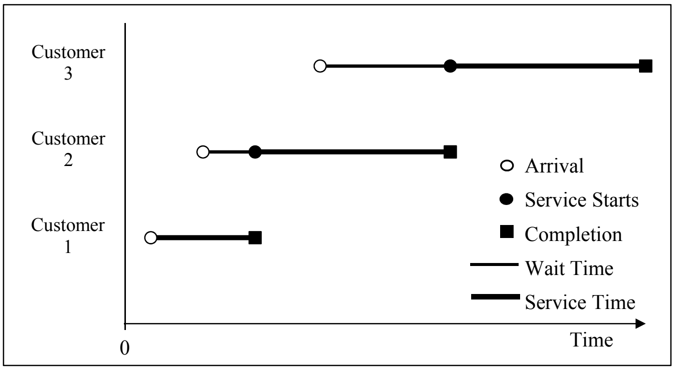


FIGURE 6.53

There are two random variables we need to model: (1) the time between arrivals and (2) the service time.

Time Between Arrivals: Theoretically it can be shown that the exponential distribution is a good model for time between arrivals. So we use this distribution with a parameter λ_1 to generate the time between arrivals. By properties of the exponential distribution,

- $1/\lambda_1$ = mean time between arrivals; and
- λ_1 = mean number of arrivals per time unit, called the *arrival rate*.

Service Time: This random variable could be modeled using any type of distribution, but often an exponential distribution is used. Call the associated parameter λ_2. By properties of the exponential distribution,

- $1/\lambda_2$ = mean service time; and
- λ_2 = mean number of customers serviced per time unit, called the *service rate*.

Example 6.7.1
Consider a single-server queue where the time between arrivals and the service time are described by exponential distributions (such a queue is called an *M/M/1 queue*) with an arrival rate of $\lambda_1 = 2$ customers per minute and a service rate of $\lambda_2 = 3$ customers per minute. Here we build a simulation to estimate the average wait time.

Algorithm

a. For each customer:

   (a) Generate a time between arrivals.
   (b) Calculate the arrival time.
   (c) Calculate the start time based on the finish time of the previous customer.
   (d) Generate the service time.
   (e) Calculate the completion time.

(f) Calculate the amount of time spent waiting in line.

(g) Calculate the cumulative wait time for all customers up to this point.

(h) Calculate the average wait time for all customers up to this point.

b. Repeat for 5000 customers.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet "**Queue**." Format the worksheet as shown in Figure 6.54.

	A	B	C	D	E	F
1	**Arrival Rate**	2	**Customer**	**Time Between**	**Arrival**	**Start**
2	**Service Rate**	3	**Number**	**Arrivals**	**Time**	**Time**
3						
4			1	=-1/$B$1*LN(1-RAND())	=D4	=E4
5			=C4+1	=-1/$B$1*LN(1-RAND())	=E4+D5	=MAX(E5,H4)

FIGURE 6.54

2. Add the formulas in Figure 6.55. Copy the formulas in the range **C5:K5** down to row 5003 to simulate 5000 customers arriving.

	G	H	I	J	K
1	**Service**	**Completion**	**Wait**	**Cumulative**	**Average**
2	**Time**	**Time**	**Time**	**Wait Time**	**Wait Time**
3					
4	=-1/$B$2*LN(1-RAND())	=F4+G4	=F4-E4	=I4	=J4/C4
5	=-1/$B$2*LN(1-RAND())	=F5+G5	=F5-E5	=J4+I5	=J5/C5

FIGURE 6.55

3. The overall average wait time is the average wait time of the last customer. To easily see this value, add the formula in Figure 6.56. We see that the overall average is about 0.666. This number can be verified theoretically (see Exercise 6.7.5).

	M
1	**Overall Average**
2	**Wait Time**
3	=K5003

FIGURE 6.56

4. Create a graph of **Average Wait Time** vs. **Customer Number** as in Figure 6.57 (fix the y–axis **min** and **max** to 0 and 1.2, respectively). The values on your graph may be different than in the figure due to the randomness of the simulation. Notice how much the average wait time varies as the number of customers increases. The average wait time doesn't settle down to close to the overall average wait time until about customer 1500.

In Exercise 6.7.2 we examine the number of people standing in line.

□

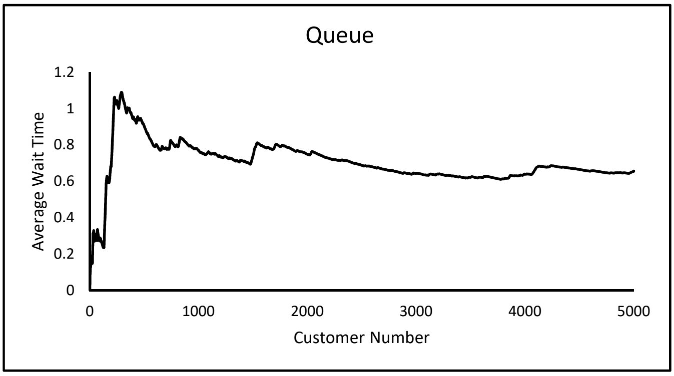


FIGURE 6.57

Exercises

6.7.1 In Example 6.7.1 we simulated a queue with a arrival rate of 2 and a service rate of 3. This means, theoretically, that an average of 2 customers will arrive per minute and an average of 3 customers will be serviced per minute. In the simulation, calculate the average number of customers that arrive per minute and the average number of customers that are serviced per minute. Do these averages agree with the theoretical values?

6.7.2 The *Queue Length* is the number of customers standing in line waiting to be serviced when a customer arrives. This does not include any customers being serviced at the time, nor does it include the arriving customer. Add a column to the worksheet **Queue** to calculate the Queue Length at the moment each customer arrives. Calculate the overall average queue length. (**Hint**: Stated another way, the queue length is the number of customers who have arrived but not yet started being serviced. Consider the scenario shown in Figure 6.58. When customer 81 arrives, customers 78, 79, and 80 have not yet started being serviced and are waiting in line, so the queue length is 3. Use the **COUNTIF** function to count the number waiting in line.)

Customer Number	Arrival Time	Start Time	Completion Time	Queue Length
76	30.58	30.72	31.20	0
77	30.86	31.20	31.86	0
78	31.16	31.86	32.19	1
79	31.43	32.19	32.40	1
80	31.56	32.40	32.68	2
81	31.73	32.68	32.71	3

FIGURE 6.58

6.7.3 Use the worksheet from Exercise 6.7.2 to estimate the probabilities in parts a. and b.

a. The probability that a single customer has a wait time of 0.

b. The probability that the queue length when a single customer arrives is 0.

c. Briefly explain why the answers to parts a. and b. are not the same

6.7.4 Consider a queue with an arrival rate of 1 and a service rate of 2.

a. Suppose both the time between arrivals and the service time are exponentially distributed, as in Example 6.7.1. Use a simulation to estimate the average wait time and average queue length.

b. Suppose the time between arrivals is exponentially distributed, but the service time is uniformly distributed between 0 and 1. Use a simulation to estimate the average wait time and average queue length. (**Hint**: This is very similar to part a, but we need to modify the formula for the service time. It's a very simple modification.)

c. Compare the average wait times and average queue lengths from parts a and b. Does changing the distribution of the service time significantly change the averages? Explain why or why not.

6.7.5 Using probability theory, it can be shown that in a single-server queue, if the time between arrivals has an exponential distribution and the service time has any distribution, then the expected number of customers waiting in line (called the expected queue length), L_q, and the expected waiting time in the queue, W_q, are given by

$$L_q = \frac{\lambda^2\sigma^2 + \rho^2}{2\,(1-\rho)} \quad \text{and} \quad W_q = \frac{L_q}{\lambda}$$

where

- $\lambda =$ mean number of arrivals per time unit (the arrival rate),
- $\mu =$ mean number of customers serviced per time unit (the service rate),
- $\sigma =$ standard deviation of the service time, and
- $\rho = \lambda/\mu$.

a. Calculate L_q and W_q for the queue in Example 6.7.1. Do these numbers agree with those found in the simulation?

b. Calculate L_q and W_q for a queue with an arrival rate of 5 and a service rate of 6, and verify these numbers with a simulation.

c. Calculate L_q and W_q for each of the two queues in Exercise 6.7.4. (**Hint**: See Example 6.6.1 to calculate the mean and standard deviation of the service time of the queue with the uniform service time.)

6.7.6 Suppose that customers arrive at the rate of 2 per minute into a single queue, but that there are two identical servers, each with a service rate of 3 per minute. Assume the time between arrivals and the service time both have exponential distributions.

a. Simulate this queue and estimate the expected wait time.

   Suggestions:

   - Start with the worksheet **Queue**.
   - Add two columns titled "Start Time Server 1" and "Start Time Server 2" and calculate the time that each customer would start if being serviced by the respective server.

- Add a column titled "Actual Start Time" and determine the time each customer will actually start being served.
- Add a column titled "Server Number" and determine which server each customer uses. Assume the default server is server 1. That is, if the customer could start with either server at the same time, server 1 is chosen.
- Add two columns titled "When is Server 1 Open?" and "When is Server 2 Open?" and calculate when each server would be available for another customer after the current customer is finished. For instance, if the current customer uses server 1, then server 1 will be open at the customer's actual start time plus the service time. This customer doesn't affect server 2, so server 2's open time will be the same as the previous customer.

b. Use the result given in Exercise 6.7.5 to calculate the expected waiting time if there were a single server with a service rate of 6. Does doubling the service rate yield the same expected waiting time as doubling the number of servers?

c. Suppose a theater has a single ticket-seller and wants to reduce customer waiting time. Assuming both of the following options cost the same, use the results of parts a. and b. to determine which would be the better option:

   i. Add equipment to double the service rate of the single ticket seller.

   ii. Add a second ticket seller.

6.8 A Scheduling Model

The Handyman Remodeling Company is considering placing a bid on a contract to remodel a living and dining room. The contract includes a large penalty if the job is not completed by the deadline of 3 weeks (21 days) from the start. The job would consist of 8 separate tasks with various precedence constraints as illustrated in Figure 6.59.

For each task, the project manager has estimated a *most likely* duration, m, an *optimistic* duration, a, and a *pessimistic* duration, b, in terms of the number of days. These numbers are given in Table 6.3.

TABLE 6.3

Task	a	m	b
Demolition	2	3	4
Electrical	2	3.5	4
Plumbing	1	2	2.5
Drywall	5	6	8
Painting	2	3	4.5
Lights	3	4	5
Carpet	1	1.25	1.5
Trim	2	3	4

Management wants to know the expected project duration and the probability of meeting the deadline so they can decide whether or not to place the bid. We will create a simulation to help answer this question.

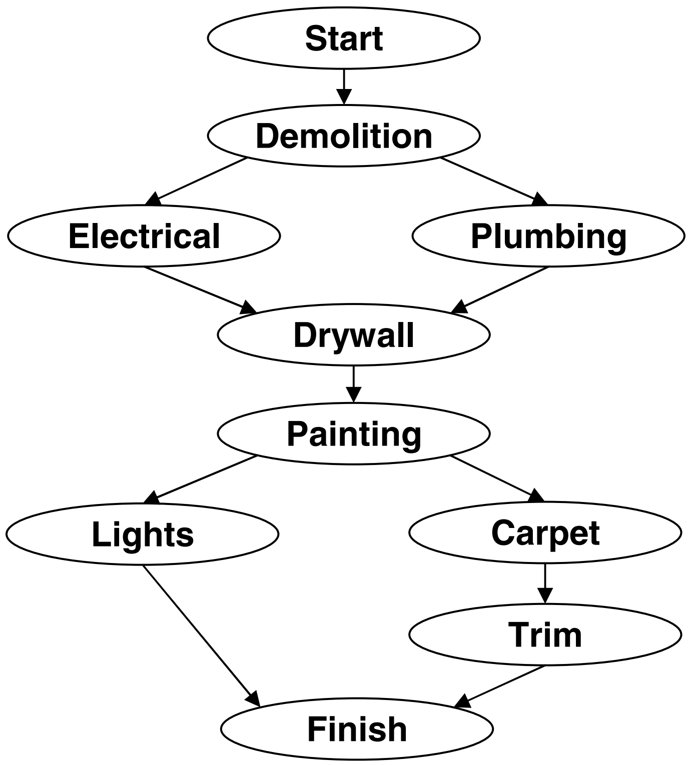


FIGURE 6.59

The duration of each task is a random variable with a distribution. The problem is that we do not know what the distribution is, and there is no reasonable way to determine what the density function is by collecting data. Therefore the best we can do is use a reasonable distribution to model these random variables. The distribution we will use is called a *triangular distribution*. Its density function is graphed in Figure 6.60.

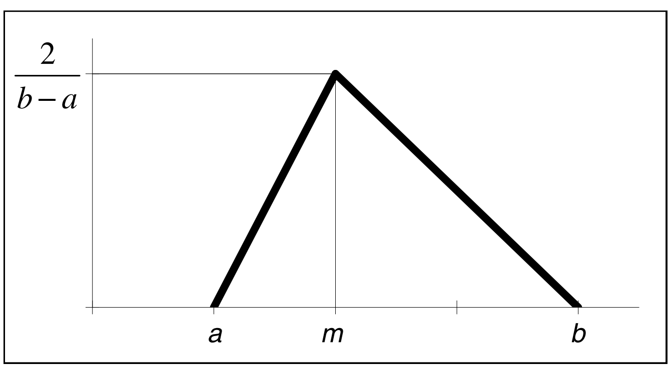


FIGURE 6.60

If we model the duration of a task with a triangular distribution, we see in Figure 6.60 that the probability the duration is near m is much higher than the probability it is near the optimistic or pessimistic estimate. This observation indicates that the Triangular Distribution is indeed a reasonable model.

It can be shown that the pdf for a Triangular Distribution is

$$f(x) = \begin{cases} \frac{2}{(b-a)(m-a)}x - \frac{2a}{(b-a)(m-a)}, & a \leq x \leq m \\ \frac{2}{(a-b)(b-m)}x - \frac{2b}{(a-b)(b-m)}, & m < x \leq b, \end{cases} \tag{6.4}$$

the cdf is

$$F(x) = \begin{cases} \frac{(x-a)^2}{(m-a)(b-a)}, & a \leq x \leq m \\ 1 - \frac{(b-x)^2}{(b-m)(b-a)}, & m < x \leq b, \end{cases} \tag{6.5}$$

and the inverse cdf is

$$F^{-1}(y) = \begin{cases} a + \sqrt{y(m-a)(b-a)}, & 0 \leq y \leq \frac{m-a}{b-a} \\ b - \sqrt{(1-y)(b-m)(b-a)}, & \frac{m-a}{b-a} < y \leq 1. \end{cases} \tag{6.6}$$

Algorithm

a. For each task:
   1. Determine the start time based on the finish times of the preceding activities.
   2. Generate the task duration.
   3. Calculate the finish time.

b. Determine the overall project duration.

c. Determine whether the deadline was met.

d. Repeat for 500 trials.

e. Calculate the average project duration and the percentage of the trials in which the deadline was met.

To implement this algorithm, follow these steps:

1. Rename a blank worksheet **Remodel** and format it as in Figure 6.61.

	A	B	C	D	E	F	G	H
1	**Task**	**a**	**m**	**b**	**Start Time**	**rand**	**Duration**	**Finish Time**
2	Demo	2	3	4	0	=RAND()		=E2+G2
3	Electrical	2	3.5	4	=H2	=RAND()		=E3+G3
4	Plumbing	1	2	2.5	=H2	=RAND()		=E4+G4
5	Drywall	5	6	8	=MAX(H3,H4)	=RAND()		=E5+G5
6	Painting	2	3	4.5	=H5	=RAND()		=E6+G6
7	Lights	3	4	5	=H6	=RAND()		=E7+G7
8	Carpet	1	1.25	1.5	=H6	=RAND()		=E8+G8
9	Trim	2	3	4	=H8	=RAND()		=E9+G9
10								
11							**Project Duration =**	=MAX(H7,H9)
12							**Deadline Met?**	=IF(H11<=21,1,0)

FIGURE 6.61

2. Add the formula for the inverse cdf as shown in Figure 6.62, and copy cell **G2** to the range **G3:G9**.

	G
2	=IF(F2<=(C2-B2)/(D2-B2),B2+SQRT(F2*(C2-B2)*(D2-B2)),D2-SQRT((1-F2)*(D2-C2)*(D2-B2)))

FIGURE 6.62

3. Start a table to store the results of 500 trials and find the average duration and percentage of successes as in Figure 6.63. Copy cell **J4** down to row 502. Create a table in the range **J2:L502** to store the results of 500 trials. Choose any blank cell for the **Column input cell**.

	I	J	K	L
1	**Average Duration**	**Trial**	**Project Duration**	**Success**
2	=AVERAGE(K3:K502)		=H11	=H12
3	**% Success**	1		
4	=AVERAGE(L3:L502)*100	=J3+1		

FIGURE 6.63

From the results we see that there is approximately an 81% chance of finishing on time and the average duration is just over 20 days. If this chance of finishing on time is too low, the project manager can experiment with different estimates to improve the chances.

Exercises

6.8.1 The quantity "Project Duration" is a random variable.

a. Approximate the mean and standard deviation of this random variable.

b. Create a histogram of the 500 values of this variable generated from the simulation, and determine which type of density function, uniform, exponential, or normal, best models this random variable.

6.8.2 Instead of modeling the duration of each task as a continuous random variable over the interval $[a, b]$, model it as a discrete variable which will take a value of a, m, or b each with a certain probability. For instance, suppose it equals a with probability 0.25, m with probability 0.5, and b with probability 0.25. Modify the worksheet from Exercise 6.8.1 to model the durations in this way. How does this change the mean, standard deviation, and distribution of the variable Project Duration? Try different values of the probabilities.

6.8.3 In this exercise we will graphically verify that our Excel formula for generating values of a random variable X described by a triangular distribution really does work.

a. Generate 500 values of a task duration with parameters $a = 3$, $m = 5.25$, and $b = 9$. Allow the user to change these parameters.

b. Create a histogram of the values in part a. Does the shape of the histogram resemble the graph of the triangular distribution density curve? Try some other values of the parameters.

6.8.4 Analytically show that the pdf, cdf, and inverse cdf of a triangular distribution are as given in Equations (6.4) - (6.6).

6.8.5 Martin has 4 days to study for his Mathematical Modeling final exam. At the end of day 0 he has an entire 100 pages of notes to read. He figures that if he spends h hours studying on any given day he will absorb $11.12\sqrt{h}$ pages of notes. He also figures that overnight he'll forget 10% of what he knew at the end of that day. So if we let h_i = the number of hours spent studying on day i and x_i = the number of pages absorbed by the end of day i, we have $x_0 = 0$ and

$$x_i = 0.9x_{i-1} + 11.12\sqrt{h_i}$$

He'd like to go into the exam having absorbed everything (i.e. $x_4 = 100$), and has planned to study a certain number of hours each day. However, he realizes that he may get lazy and might not be able to study as many hours as he planned. On the other hand, he may get very motivated and study more hours than anticipated. Therefore, for each day he has estimated a *most likely* study duration, an *optimistic* duration, and a *pessimistic* duration as shown in the table below. Design a simulation to estimate the probability that he will absorb everything and the expected number of pages he will absorb. Should he revise his time estimates? If so, how?

Day	Pessimistic	Most Likely	Optimistic
1	2	4	6
2	5	6	7.5
3	6	7	9
4	7.5	9	11

6.9 An Inventory Model

Consider the following scenario:

> The produce department at a neighborhood grocery store gets its bananas from a local supplier. To better schedule its deliveries, the supplier has asked the produce manager to place his order on a regular basis (i.e. every 5 days). The manager is trying to decide how often to place his order.
>
> The manager has room to store 50 boxes of bananas. Each time he places an order, he will order enough to completely replenish his stock. He will place his order at the end of the day and it will be delivered that evening. There is a $25 delivery fee for each delivery, regardless of the size of the order.
>
> The manager looks over his records from the previous month (30 days) for daily demand and notes that he sold between 1 and 10 boxes of bananas each day. The daily demand data is summarized in Table 6.4.
>
> Ideally, the manager wants to place his order before he runs out of bananas. For instance, if on the end of day 8 he has 2 boxes left and on day 9 he has demand for 5 boxes, he would have wished he had ordered 48 boxes on day 8. However, he doesn't want to place an order too often because of the delivery fee, so he's willing to occasionally run out of bananas.

TABLE 6.4

Demand	1	2	3	4	5	6	7	8	9	10
Number of Days	3	3	2	2	3	5	2	4	3	3

The above scenario is somewhat vague. So we first need to state a precise question to answer:

> How often should the manager place his order so that the department can meet demand at least 95% of the time?

Our simulation will consist of replicating the sale and delivery of bananas over a period of one year. We will vary the number of days between deliveries and keep track of the number of days demand was met.

The only random variable in this simulation is the daily demand. This is a discrete random variable and we will use the data in Table 6.4 to model the cdf. To do this, rename a blank worksheet **Bananas** and format it as in Figure 6.64. Enter the rest of the data from Table 6.4 and copy the range **O4:P4** down to row 12.

	M	N	O	P
1		**Num of**	**Relative**	**Cumulative**
2	**Demand**	**Days**	**Frequency**	**Frequency**
3	1	3	=N3/30	=O3
4	2	3	=N4/30	=O4+P3

FIGURE 6.64

The graph of the Cumulative Distribution Function, $F(x)$, is shown in Figure 6.65. (It is not necessary to create this graph.) To generate values of the demand based on this cdf we will first pick a uniformly distributed random number y using the **RAND** function. From that y-value on the graph, we will move horizontally until we hit a column. The corresponding x-value will be the value of $F^{-1}(y)$. For instance, $F^{-1}(0.065) = 1$ and $F^{-1}(0.75) = 8$.

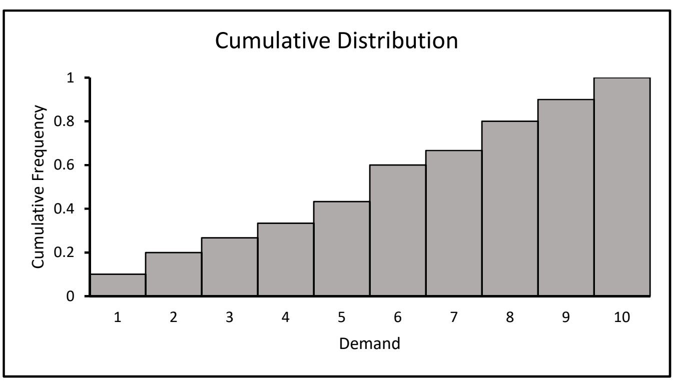


FIGURE 6.65

To calculate $F^{-1}(y)$, add the formulas in Figure 6.66 and copy row 4 down to row 12. The simulation will refer to the "lookup chart" to calculate the daily demand.

	R	S
1	**Lookup Chart**	
2	**Frequency**	**Demand**
3	0	1
4	=P3	2

FIGURE 6.66

Algorithm

a. Choose a value for Days Between Deliveries.

b. For each day:

   1. Generate a demand.
   2. Determine if the demand can be met.
   3. Determine if a delivery will be made.
   4. Calculate inventory at the end of the day.

c. Calculate the percentage of days the demand was met.

d. Repeat for 100 trials.

e. Try different values of Days Between Deliveries.

To implement this algorithm, follow these steps:

1. In the worksheet **Bananas**, add the formulas in Figure 6.67. Copy row 6 down to row 370. The **VLOOKUP** function in column **C** will look at the left-most column in the lookup chart, find the largest value less than or equal to the value of the random number, and return the corresponding value in the second column of the chart. This value will be the demand for that day.

	A	B	C	D	E	F
1			**Days Between Deliveries =**	5		
2						
3					**Demand**	**Inventory**
4	**Day**	**rand**	**Demand**	**Delivery?**	**Met?**	**at End of Day**
5	0					50
6	=A5+1	=RAND()	=VLOOKUP(B6,$R$3:$S$12,2)	=IF(MOD(A6,$D$1)=0,1,0)	=IF(F5>=C6,1,0)	=IF(D6=1,50,F5-C6)

FIGURE 6.67

2. Add the formulas in Figure 6.68 to calculate the percentage of days that demand was met, store the results from 100 trials, and calculate the overall results. Create a table in the range **I4:J104** to store the results from 100 trials. Choose any blank cell for the column input cell. Press **F9** several times to repeat the simulation. Note that with 5 days between deliveries, demand is met 100% of the time.

3. Change the value of **Days Between Deliveries** and rerun the simulation until you find the largest value that gives at least a 95% Overall Average. Note that for 8 days between deliveries demand is met about 95.5% of the time. For 9 days between deliveries demand is met only about 90% of the time. Thus 8 days between deliveries is the answer to the question. In the exercises we will consider a refinement to the model.

	I	J
1	**% Days Demand Met =**	=AVERAGE(E6:E370)*100
2	**Overall Average =**	=AVERAGE(J5:J104)
3	**Trial**	**%**
4		=J1
5	1	
6	=I5+1	

FIGURE 6.68

Exercises

6.9.1 In this exercise we will test whether our method for generating values of the daily demand works properly.

a. Create a spreadsheet that generates 2000 values of the daily demand.

b. The theoretical expected daily demand is $\sum_{x=1}^{10} x \cdot f(x)$ where $f(x)$ is the relative frequency of demand x. Use the data in Table 6.4 to calculate the expected daily demand.

c. Calculate the average of the 2000 values of the daily demand and compare this number to the theoretical number from part b.

d. Calculate the relative frequencies of the 2000 values of the daily demand and compare these to the relative frequencies from the data in Table 6.4. Does our method for generating values of the daily demand appear to work properly?

6.9.2 We modeled the demand random variable as a discrete variable whose distribution is motivated by Table 6.4. In this exercise we consider other distributions. Modify the simulation to model the demand with each given distribution and find the answer to the original question. Is this answer any different from the answer of 8 days found using the original distribution?

a. Discrete and uniform over the interval $[1, 10]$

b. Continuous and uniform over the interval $[1, 10]$

c. Normal with mean $\mu = 5.7$ and standard deviation $\sigma = 1.5$.

6.9.3 Suppose that the produce department makes \$4 profit for each box of bananas sold (if we don't take into account delivery costs).

a. Add a column of formulas to calculate daily profit. (**Note**: The daily profit is the number of boxes *sold* times \$4, not the number *demanded* times \$4.)

b. Add a formula to calculate the total delivery cost for the year (each delivery costs \$25).

c. Add a formula to calculate the total yearly profit from the sale of bananas.

d. Modify your table to store the profit and add a formula to calculate the average profit from all 100 trials.

e. Find the value of **Days Between Deliveries** that gives the maximum profit. How does this value compare to the solution to the original problem of 8 days between deliveries?

6.9.4 Consider a generalization of Exercise 6.5.1. Suppose we want to generate a list of pseudorandom numbers each of which has the value x_1, x_2, ..., or x_6 where x_1 occurs with probability p_1, x_2 occurs with probability p_2, etc. where $p_1 + \cdots + p_6 = 1$.

a. Design a spreadsheet which generates this list. Make sure the user is able to input the values of the x's and the p's. (**Hint**: Create a lookup table like we did in this section.)

b. Use the **COUNTIF** function to verify that the list contains the proper proportion of each number.

6.9.5 A local manufacturing plant produces a variety of products. The distribution of the total monthly demand is shown in the table below. Depending on conditions the average manufacturing cost per item is between \$60 and \$80 in integer values, and returns from distributors are between 120% and 130% of manufacturing costs. There is a fixed cost of \$2000/month for tooling, processing, etc. Build a simulation to estimate the average monthly profit. State all assumptions used.

Demand	**Probability**
300	0.05
320	0.10
340	0.20
360	0.30
380	0.25
400	0.10

6.9.6 The management at a bank is trying to improve customer satisfaction by offering better service. They want customers to wait on average less than 2 minutes and the average length of the queue to be 2 or fewer. The bank estimates it serves about 150 customers per day. Management observes that the service time and time between customer arrivals (in minutes) are typically whole numbers and that the distributions are as given in the tables below.

Time Between Arrivals	**Probability**	**Service Time**	**Probability**
0	0.10	1	0.25
1	0.15	2	0.20
2	0.10	3	0.40
3	0.35	4	0.15
4	0.25		
5	0.05		

a. Modify a worksheet from Section 6.7 to simulate customers arriving and being served at the bank. Use the simulation to estimate the average queue length and average wait time.

b. Are the bank's goals being met? If not, suggest improvements the bank can make.

6.9.7 In the Birthday Problem (Example 6.3.2) we simulated selecting a random class of n students and estimated the probability that at least two students share a birthday. Our model was based on the simplifying assumption that birthdays are uniformly distributed throughout the year and we ignored leap years. Table 6.9.7 shows the number of births (in tens of births) in the United States on each day of the year, including February 29, for the years 1994-2014 (data from http://thedailyviz.com/2016/09/17/how-common-is-your-birthday-dailyviz/ as accessed by Brennan DeForest, March 22, 2019). The upper left-hand corner corresponds to January 1, below that is January 2, etc. Redo the birthday problem simulation using these data to generate the birthdays. Estimate the smallest class size necessary so the probability that at least two students share a birthday is at least 0.50. Is this class size different than in the original simulation? (**Hint**: Create a lookup table like we did in this section.)

TABLE 6.5

15,584	22,298	22,022	21,634	21,594	17,592	22,982	24,144	22,092	21,528
18,614	22,126	21,546	21,754	21,564	20,808	23,216	24,296	22,552	21,710
21,626	21,786	22,274	21,728	21,802	22,974	23,498	24,110	22,366	22,502
22,038	22,030	21,908	21,690	21,438	24,216	22,936	24,458	21,856	22,364
21,906	22,030	21,828	21,992	22,328	23,888	23,384	24,214	22,064	22,284
21,822	21,796	22,006	21,764	22,690	23,538	23,842	23,626	22,204	21,962
21,850	21,208	22,362	21,328	22,512	23,476	23,576	23,840	22,024	22,264
21,220	23,272	21,934	21,606	22,442	23,588	23,096	23,948	21,630	21,916
21,248	22,376	21,478	21,470	22,328	23,130	23,362	23,890	19,956	21,482
22,046	21,896	21,842	21,462	22,480	22,362	23,274	23,732	22,700	21,786
21,950	21,708	21,948	22,004	22,320	23,360	23,542	23,986	22,162	21,698
21,868	21,880	21,776	22,226	22,050	23,508	23,286	23,722	22,260	21,902
21,244	21,346	21,790	21,806	22,166	23,536	23,650	23,108	22,258	21,766
21,952	21,772	22,090	21,434	22,444	23,436	23,310	23,144	22,382	22,880
21,092	22,016	21,746	22,146	22,320	23,544	22,904	22,978	22,162	21,710
21,246	22,222	21,428	21,898	22,392	23,090	23,152	23,440	22,616	21,904
21,802	21,854	21,558	21,890	22,082	22,856	23,240	23,144	22,360	22,382
21,766	21,808	20,600	21,910	22,576	23,328	23,474	23,348	21,854	22,704
21,382	21,948	22,008	22,080	22,156	23,372	23,710	22,980	22,078	22,962
21,650	21,454	21,798	22,142	22,530	23,398	23,848	22,544	22,282	23,350
21,648	21,716	22,438	21,488	22,506	23,214	23,600	22,670	22,154	23,870
21,346	22,106	21,800	22,032	22,678	23,536	23,110	22,648	21,484	24,018
21,730	5,231	21,278	21,394	22,352	23,162	21,860	22,618	22,480	23,360
22,098	22,258	21,718	22,140	23,004	22,820	22,000	22,274	22,458	22,776
21,902	21,604	21,780	22,314	22,596	23,228	22,238	23,112	22,044	20,676
21,686	22,148	21,660	22,566	22,260	23,186	22,432	22,536	22,250	16,138
21,646	21,978	21,652	22,244	22,488	23,198	22,862	22,028	22,346	13,148
21,670	21,958	22,118	21,798	22,656	23,032	22,586	21,536	22,510	19,086
21,134	21,842	21,906	21,998	22,812	23,550	22,796	22,298	22,884	23,330
21,504	22,174	20,778	22,386	22,748	23,160	23,984	22,522	23,134	23,710
21,766	21,952	21,624	22,508	23,180	22,664	24,602	22,230	21,328	23,912
21,858	21,530	21,766	22,576	23,114	23,138	24,286	22,592	19,766	23,778
21,898	21,880	21,818	23,050	22,702	23,220	23,006	22,298	20,030	20,788
21,686	21,862	21,794	22,734	23,094	23,172	24,448	21,700	19,908	
21,810	22,006	22,008	21,654	23,720	23,178	23,602	22,130	20,088	
21,370	21,308	21,782	20,802	23,656	23,902	23,764	22,114	19,436	
21,588	22,238	21,428	21,386	22,608	23,442	24,174	22,312	20,192	

Project Ideas

A popular general topic for a simulation project is to simulate a simple game of chance involving cards, dice, or spinners. The simulation can be used to answer questions such as

- What's the average length of the game?
- What's the average score?
- What's the distribution of the scores?
- What would happen if we used a biased die or deck of cards.

Many games of chance do involve some strategy. In the simulation, the strategy may have to be simplified, but the simulation can be used to compare different strategies. Here are a few specific simulations to consider:

1. Simulate the card game Blackjack (a.k.a. 21) to compare different strategies for when to hit (e.g. should you always hit at 16 or less, 17 or less, etc.).
2. Simulate the golf dice game GOLO to compare different strategies for keeping the dice in each round.
3. Simulate rolling dice from the game Dungeons and Dragons.
4. Simulate a fast food drive-thru. Compare different scenarios, one where the customers order and receive their food at the same window, and the other where they order their food at one window and pick up at a second window.
5. Simulate a variation of Buffon's needle problem where the needle has the shape of a circle, rectangle, or triangle.
6. Simulate a simple basketball game between two teams, each containing players of different abilities, and compare different substitution patterns.
7. Simulate a restaurant that seats 20 people. Compare the scenario where there are 5 tables that each seat 4 to another where there are 4 tables of 5. Which scenario allows the most number of customers to be served in a day?
8. Research the Die Hard battery of tests for randomness and try to implement one of the tests, or at least a simplified version of one.
9. Research ways of using simulations to estimate the value of π.
10. Research the Ant Colony Simulation optimization heuristic.
11. Simulate the board game Chutes and Ladders to estimate the average number of turns it takes for a player to win. How does this number change if there is more than one player?
12. Simulate the dice game Yahtzee.
13. Simulate the game Pokemon to estimate the probability of catching a wild Pidgey.
14. Simulate the casino game Roulette.

15. A bag contains several marbles of different colors. Suppose you randomly choose x marbles. Create a simulation to estimate the probably that exactly y of the chosen marbles are a certain color. Let the user input the values of x and y, the number of colors, and the number of marbles of each color.
16. Simulate the card game Up the River, Down the River.
17. Simulate the dice game Bunco.
18. Simulate the game Bingo.
19. Simulate the game Sorry.
20. Simulate the game Keno to test different strategies for picking the numbers.
21. Simulate the board game Battleship. Will a player find the battleship faster if they guess randomly or in a diagonal pattern?
22. Simulate the game of Nim.
23. Research an M/M/c queuing system where there are $c > 1$ servers.
24. Simulate spinning the Big Wheel on the game show The Price is Right!
25. Simulate the dice game Farkle.
26. Simulate the board game Axis & Allies to estimate the probability of winning with different combinations of attacker and defender forces.
27. Simulate rolling dice in the game of Risk.
28. Simulate the game Plinko from the game show The Price is Right to estimate the distribution of the outcomes.
29. Simulate the board game Hi Ho Cherry-O.
30. The game of Clue involves trying to solve a crime. Because the game ends when someone guesses the correct solution, players will try to guess the correct solution before the other players can, and may have to guess without fully eliminating every possible conclusion. Create a simulation to estimate the chances of winning a game of Clue if you were to guess knowing only a certain number of cards were eliminated.

For Further Reading

- For a classic reference on many of the concepts related to simulation, see Hillier, F. and G. Lieberman, *Introduction to Operations Research*, Seventh Edition, McGraw Hill, 2001, pg. 1084 – 1155.
- For a classic reference on everything related to simulation, see Law, Averill M. and W. David Kelton, *Simulation Modeling and Analysis* Second Edition, McGraw–Hill, Inc., 1991.
- For more examples of the concepts in this chapter, see Maki, Daniel and Maynard Thompson, *Mathematical Modeling and Computer Simulation*, Thomson Brooks/Cole, 2006.

7

Linear Optimization

Chapter Objectives

- Discuss the basic concepts of optimization problems
- Introduce linear programming
- Model transportation and assignment problems
- Discuss the basics of the Simplex Method

7.1 Introduction

In Calculus I, we solve problems such as

$$\textbf{Maximize } f(x) = -2x^2 + 3x + 2$$

by taking the derivative of the function f and setting it equal to 0. This is a very simple *Optimization Problem* (abbreviated OP). Practical situations often involve finding solutions to more complex optimization problems such as when a business is trying to decide how many units of a product to produce in order to maximize profit.

Every OP has two components: *decision variable(s)* and an *objective function*. The objective function is the function being maximized. The decision variable(s) are the variable(s) involved. The basic goal of an OP is to find values of the decision variable(s) that maximize the objective function.

Optimization problems are classified into two general categories, *constrained* and *unconstrained*. A constrained OP is one in which there are constraints on the values of the decision variable(s). An unconstrained OP has no such constraints. These constraints can be of many different forms, including:

1. Non-negativity (the decision variables must be non-negative)
2. Integrality (the decision variables must be integers)
3. Binary (the decision variables must be 0 or 1)
4. Equality (e.g. $x + y = 5$)
5. Inequality (e.g. $x + y \geq 6$)

Optimization problems are also classified into two other categories: *linear* or *nonlinear*. A linear OP is one in which the objective function and constraints are equations or inequalities of the form

$$2x_1 - 6x_2 + 8x_3 = 2$$

where x_1, x_2, and x_3 are decision variables (i.e. the objective function and constraints are equations or inequalities of the type studied in linear algebra). A *nonlinear* OP has an objective function or at least one constraint that is not of this type.

7.2 Linear Programming

The type of optimization problem we focus on in this chapter is called a *linear program* (LP), which is simply a linear constrained OP. More specifically, an optimization problem is called a *linear program* if it satisfies the following five properties:

1. There is a unique objective function.
2. Whenever a decision variable appears in either the objective function or a constraint function, it must appear with an exponent of 1, possibly multiplied by a constant.
3. No term contains products of decision variables.
4. All coefficients of decision variables are constants.
5. Decision variables are permitted to assume fractional as well as integer values.

An example of an LP is:

$$\begin{array}{rl} \textbf{Maximize} & Z = 25x_1 + 30x_2 \\ \textbf{Subject to} & \quad 20x_1 + 30x_2 \leq 600 \\ & \quad 5x_1 - \ 4x_2 \leq 4 \\ & \quad x_1 \geq 4,\ x_2 \geq 2. \end{array} \tag{7.1}$$

The objective function in this case is $25x_1 + 30x_2$, and its value is denoted Z. A *solution* to an LP (also called a *schedule* or a *feasible solution*) is any set of values of the decision variables that satisfies the constraints. *A* solution to the LP (7.1) is, for example, $x_1 = 5$, $x_2 = 6$. An *optimal solution* is a solution that gives the maximum (or minimum) value of the objective function over all possible solutions. Finding a solution is relatively easy. Finding an *optimal* solution is not as easy.

The algorithm used to find optimal solutions to linear programming problems is called the *Simplex method*. Excel uses this algorithm as part of its **Solver** tool. In Section 7.6 we will discuss the basics of how the Simplex method works. In this section we use Solver to solve some simple linear programs.

Example 7.2.1 (Making Fruit Baskets)
The manager of a produce department at a neighborhood supermarket is making fruit baskets for the busy holiday season. He sells two sizes of baskets: small and large. He has only 200 apples and 100 oranges remaining and is trying to decide how many of each size of basket he should make. Each small basket returns a profit of \$3 and requires 3 apples and 1 orange while each large basket returns a profit of \$4 and requires 2 apples and 2 oranges. Assuming he will sell all that he makes, how many of each size should he make to maximize his profit?

The first step in modeling a problem such as this is to organize all the information given into a *mixture chart* as shown in Table 7.1. The two sizes of baskets are generically called the *products* and the apples and oranges are called the *resources*.

TABLE 7.1

	Products		
Resources	**Small**	**Large**	**Amount Available**
Apples	3	2	200
Oranges	1	2	100
Profit	3	4	

The next step is to define variables and write a set of inequalities to model the situation. If s denotes the number of small baskets to produce and l denotes the number of large baskets to produce, the model is:

$$\begin{array}{rrl} \textbf{Maximize} & P = 3s + 4l & \\ \textbf{Subject to} & 3s + 2l & \leq 200 \\ & 1s + 2l & \leq 100 \\ & s,\ l & \geq 0 \end{array}$$

The objective function $3s + 4l$ gives the total profit. The first constraint says we can't use more than 200 apples while the second says we can't use more than 100 oranges. The last two constraints are non–negativity constraints which say we can't produce a negative number of baskets.

To solve this LP, rename a blank worksheet "**Fruit Baskets**" and format it as in Figure 7.1. The numbers in the range **B2:C2** are the values of s and l (these are not the optimal values, yet).

	A	B	C	D	E
1		**Small**	**Large**		
2	**Number**	1	1	**Amt Used**	**Amt Available**
3	**Apples**	3	2	=SUMPRODUCT($B$2:$C$2,B3:C3)	200
4	**Oranges**	1	2	=SUMPRODUCT($B$2:$C$2,B4:C4)	100
5				**Total Profit**	
6	**Profit**	3	4	=SUMPRODUCT($B$2:$C$2,B6:C6)	

FIGURE 7.1

Select **Data** → **Analysis** → Add-Ins... (If **Solver** is not available, select **File** → **Options** → **Add-Ins**. Then select **Excel Add-ins** next to **Manage:** and press the **Go**... button. Check the box next to **Solver Add-in** and press **OK**.) Format the Solver window as in Figure 7.2 and press the **Solve** button.

The drop-down box next to **Select a Solving Method** contains three different algorithms for solving optimization problems:

1. **GRG Nonlinear** - This is an algorithm for solving linear and nonlinear programs based on the *gradient method*. The basics of this method are discussed in Sections 8.4 and 8.5.

2. **Simplex LP** - This is an algorithm for solving linear programs that utilizes the simplex method. We choose this algorithm in this example because the program is linear. The basics of this method are discussed in Sections 7.5 and 7.6

3. **Evolutionary** - This algorithm falls into a general category called *evolutionary methods*. This method is used in Section 8.8.

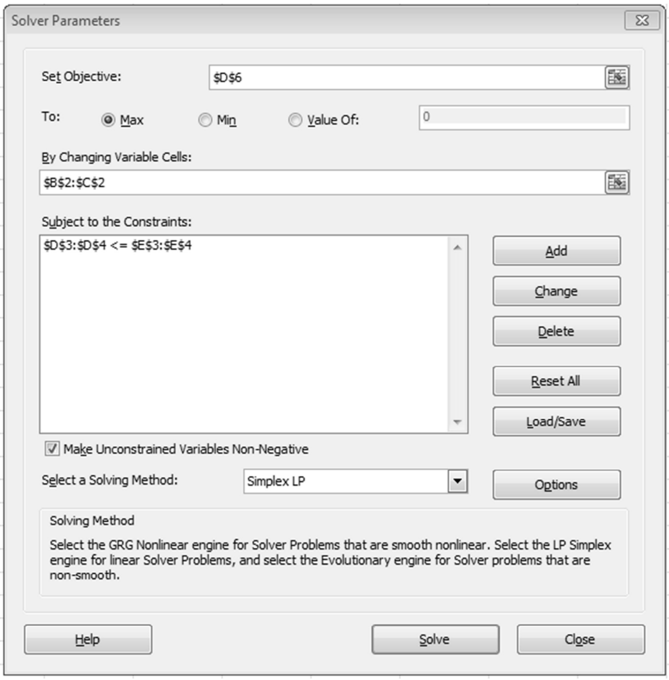


FIGURE 7.2

The worksheet should now look like Figure 7.3. From this we see that our optimal solution is to produce 50 small baskets and 25 large baskets (called the optimal *production schedule*), which will yield a profit of $250. Also note that we use all 200 apples and 100 oranges.

	A	B	C	D	E
1		**Small**	**Large**		
2	**Number**	50	25	**Amt Used**	**Amt Available**
3	**Apples**	3	2	200	200
4	**Oranges**	1	2	100	100
5				**Total Profit**	
6	**Profit**	3	4	250	

FIGURE 7.3

□

Example 7.2.2 (A Diet)
John has decided to go on a diet to lose weight and has limited himself to two types of food: protein shakes and pasta (plus vitamin supplements). He is concerned with getting enough protein and carbohydrates, but not too much fat. Table 7.2 lists the amounts of these nutrients provided by each type of food along with his daily requirements and cost information. How many servings of each food should he eat each day to meet the requirements and minimize the cost?

TABLE 7.2

	Grams per serving		
	Shake	Pasta	Daily Requirement
Carbs	1	35	≥ 80
Protein	7	6	≥ 75
Fat	5	1	≤ 50
Cost per serving	\$0.75	\$1.25	

If we let s denote the number of protein shake servings and p denote the number of pasta servings per day, our mathematical model is:

$$\begin{array}{rl} \textbf{Minimize} & C = 0.75s + 1.25p \\ \textbf{Subject to} & 1s + 35p \geq 80 \\ & 7s + 6p \geq 75 \\ & 5s + 1p \leq 50 \\ & s,\ p \geq 0 \end{array}$$

Note that in this case, the objective function gives the cost, not the profit. Solving this problem in a similar fashion to the previous example gives the results shown in Figure 7.4. Observe that the solution is not integral, which may or may not be an issue. It certainly is possible to make fractional servings, but we typically think of servings in terms of whole numbers. We could have added the constraint that the decision variables be integers, but then the problem would not be linear. A simpler way to interpret this mathematical solution in the real-world is to simply round-off the values to 9 shakes and 2 servings of pasta. This rounded-off solution does not meet the carbohydrate requirement (it provides only 79 grams of carbohydrates), but arguably this is close enough.

	A	B	C	D	E
1		**Shake**	**Pasta**		
2	**Number**	8.975	2.029	**Amt Consumed**	**Amt Needed**
3	**Carbs**	1	35	80	80
4	**Protein**	7	6	75	75
5	**Fat**	5	1	46.90376569	50
6				**Total Cost**	
7	**Cost**	0.75	1.25	9.267782427	

FIGURE 7.4

□

Exercises

Directions: Formulate each problem below as an LP and solve it with Solver. Make sure that all the constraints and the objective functions are linear.

7.2.1 A toy company manufactures plastic cars and trucks. Each car yields a profit of \$1.25 and requires 5 units of plastic and 15 minutes of labor to produce. Each truck yields a profit of \$0.95 and requires 2 units of plastic and 18 minutes of labor to produce. If the company has 60 units of plastic and 360 minutes of labor available, how many of each vehicle should they produce to maximize their profit?

7.2.2 The Sweetie Pie Baking Company operates three different plants. Two are used for mixing ingredients and baking, and the third is used strictly for packaging. Management is considering adding two new products to their line-up, a chocolate-chunk cookie and a raisin bread. The cookie will be baked in plant 1 while the bread will be baked in plant 2. They will both be packaged in plant 3. One batch of each product requires a certain amount of time in the different plants and each plant has a certain amount of production time available each week. The data is summarized in the table below. How many batches of each product should be produced each week to maximize total profit?

	Time needed		**Time available**
	Cookie	**Bread**	**per week**
Plant 1	3	0	6
Plant 2	0	2	10
Plant 3	1	2	10
Profit per batch	300	250	

7.2.3 An aid organization is sending boxes of food and clothing to assist hurricane disaster victims.

a. The army offers to transport the boxes, provided they fit in the available cargo space. Each 20-ft^3 box of food weighs 40 lbs, and each 30-ft^3 box of clothing weighs 20 lbs. The total weight cannot exceed 16,000 lbs, and the total volume must not exceed 18,000 ft^3. Each box of food will feed 10 people, while each box of clothing will help clothe 8 people. How many boxes of food and how many boxes of clothing should be sent in order to maximize the total number of people assisted? How many people will be assisted?

b. Suppose a trucking company offers to deliver 550 boxes of clothing and 165 boxes of food free of charge. They claim they can make the delivery more quickly than the army. Would you advise taking their offer? Briefly explain why or why not.

7.2.4 A hog farmer typically buys hog feed from a local farming cooperative which offers two brands of hog feed. Brand X costs \$25 a bag and contains 2 units of nutritional element A, 2 units of element B, and 2 units of element C. Brand Y costs \$20 a bag and contains 1 unit of nutritional element A, 9 units of element B, and 3 units of element C. The minimum requirements for nutrients A, B, and C are 12 units, 36 units, and 24 units, respectively. The farmer plans to buy several bags of each brand and mix them together.

a. Find the number of bags of each brand the farmer should purchase from the cooperative to meet the requirements at a minimum cost.

b. Suppose a competitor feed dealer offers to sell the farmer 4 bags of each brand for a total of \$160. Do you recommend the farmer accept the offer? Briefly explain why or why not.

7.2.5 An automobile repair company performs paint-less dent removal from hail damaged cars and trucks. Each vehicle must be processed in both the body assembly shop and the finishing shop. In the body shop it takes 0.5 man-hours to repair a car and 0.5 man-hours to repair a truck. There are 25 body shop man-hours available per day. In the finishing shop it takes 0.4 man-hours to finish a car and 0.6 man-hours to finish a truck. There are 24 finishing man-hours available per day. Each car contributes \$200 to overall profit, and each truck contributes \$225 to overall profit.

a. Find the number of cars and trucks the company can service a day to maximize overall profit.

b. Suppose the repair company is offered a contract to repair 30 cars and 30 trucks a day. Would you recommend the company accept the contract? What changes, if any, would the company need to make? Briefly explain your answer.

7.2.6 The Jones Furniture company produces tables and chairs. Each table requires 4 hours of carpentry labor, 3 hours of painting labor, and 100 board-ft. of wood. Each chair requires 4.5 hours of carpentry labor, 2 hours of painting labor, and 75 board-ft. of wood. Each table returns a profit of \$6, and each chair returns a profit of \$4.75. Each month, there are 2400 hours of carpentry labor, 1500 hours of painting labor, and 10,000 board-ft of wood available. Management requires at least 50 tables be produced each month.

a. Find the number of tables and chairs that should be produced each month to maximize overall profit.

b. Which resource, carpentry labor, painting labor, or wood, is the most limiting? If the company wanted to increase the amount of this resource, how much would you recommend they increase it so the company more efficiently utilizes the available resources?

7.2.7 The Nutty Goodness Company sells mixtures of peanuts, walnuts, and cashews. A new customer wants 100 lb of a mixture that is 45% peanuts, 30% walnuts, and 25% cashews. The company has run out of peanuts, so they are going to make the mixture by combining five different mixtures containing different percentages of peanuts, walnuts, and cashews as shown in the table below. Determine the amounts of each of the five different mixtures that should be blended to form 100 lb of the new mixture at a minimum cost.

	Mixture				
	1	**2**	**3**	**4**	**5**
Percentage of peanuts	45	20	55	50	55
Percentage of walnuts	23	20	42	20	35
Percentage of cashews	32	60	3	30	10
Cost (\$/lb)	4.80	5.20	4.90	4.60	4.30

7.2.8 The Nostalgic Transportation Company manufactures collectible models of old cars. Their Model T returns a profit of \$3 and requires 12 minutes of assembly time and their Model A returns a profit of \$5 and requires 25 minutes of assembly time. The plant manager estimates that due to maintenance and breakdowns, the machine used for assembly operates only 33 hours per week. Based on sales predictions, management requires that for every five Model T's produced, at least two Model A's must be produced.

a. How many of each model should be produced each week to maximize profit?

b. Suppose the plant manager believes that hiring an additional maintenance person could increase the machine's operational time to 39 hours per week. Find the new optimal production schedule in this case.

c. Based on your solution to part b., how much would hiring a new maintenance person increase weekly profit? Do you think it's economically beneficial for the company to hire this person? Briefly explain why or why not.

7.2.9 The Wetzel Woodworking Company produces tables and chairs. Each table returns a profit of \$30 and requires 5.5 hours of labor while each chair returns a profit of \$10 and requires 2.5 hours of labor. There are 40 hours of labor available per week. Customer demand requires that at least three times as many chairs be produced as tables. Tables and chairs produced in one week must be stored over the weekend and shipped out the following Monday. Tables require three times as much storage space as chairs. The warehouse has room to store the equivalent of at most six tables.

a. How many tables and chairs should be produced each week to maximize profit?

b. Suppose the company wants to increase storage capacity to eight tables. Find the new optimal production schedule. Is there really any economical benefit to adding the storage? Briefly explain why or why not.

7.2.10 The Cubicles Unlimited Company manufactures office desks and cabinets. Management is trying to decide how many desks and cabinets to manufacture in a certain week to fill an order that will be shipped out the following week. Here are the requirements:

- Each desk requires 45 minutes of assembly time and 30 minutes of painting time while each cabinet requires 30 minutes of assembly time and 21 minutes of painting time.
- The assembly department has 40.25 hours and the painting department has 28.5 hours of production time available during the week.
- At the beginning of the week there are 50 desks and 55 cabinets in stock.
- The order is for 75 desks and 95 cabinets.
- Company policy is to maximize the total inventory at the end of each week.

Find an optimal solution to this problem.

7.2.11 Linear programs can be written in matrix form. The *standard maximum problem* is defined as:

$$\begin{aligned} &\textbf{Maximize } \boldsymbol{c}^T\boldsymbol{x} \\ &\textbf{Subject to } A\boldsymbol{x} \leq \boldsymbol{b} \text{ and } \boldsymbol{x} \geq \boldsymbol{0}, \end{aligned} \tag{7.2}$$

where $\boldsymbol{c}$ is the vector of coefficients in the objective function, $\boldsymbol{x}$ is the vector of decision variables, A is the matrix of coefficients in the constraints, and $\boldsymbol{b}$ is the vector of constants on the right-hand side of the constraints. For instance, the LP in Example 7.2.1 written in this standard form is:

$$\begin{aligned} &\textbf{Maximize } \begin{bmatrix} 3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \\ &\textbf{Subject to } \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 200 \\ 100 \end{bmatrix} \text{ and } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned} \tag{7.3}$$

Every LP has an associated *dual linear program*. The dual of the standard maximum problem (7.2) is defined as

$$\begin{aligned} &\textbf{Minimize } \boldsymbol{y}^T\boldsymbol{b} \\ &\textbf{Subject to } \boldsymbol{y}^T A \geq \boldsymbol{c}^T \text{ and } \boldsymbol{y} \geq \boldsymbol{0}, \end{aligned}$$

where $\boldsymbol{y}$ is a vector of new decision variables. An LP and its dual are intimately connected. In this exercise we'll look at one of those connections. The study of the dual problem falls into a branch of mathematics called *duality theory*.

a. Show that the dual of the problem (7.3) is

$$\textbf{Minimize} \quad \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 200 \\ 100 \end{bmatrix}$$
$$\textbf{Subject to} \quad \begin{bmatrix} y_1 & y_2 \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} \geq \begin{bmatrix} 3 & 4 \end{bmatrix} \text{ and } \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

b. Solve the dual problem in part a. and compare this solution to the original problem found in Example 7.2.1. What similarities do you observe?

7.2.12 Find the dual of the following LP. Solve both the LP and its dual, and compare the solutions.

$$\begin{array}{lrl} \textbf{Maximize} & P = & x_1 + x_2 \\ \textbf{Subject to} & & x_1 + 2x_2 \leq 4 \\ & & 4x_1 + 2x_2 \leq 12 \\ & & -x_1 + x_2 \leq 1 \\ & & x_1, x_2 \geq 0 \end{array}$$

7.3 The Transportation Problem

Operations Research (OR) is a branch of applied mathematics that deals with researching the operations of organizations, such as businesses and industries, with the goal of helping them operate more efficiently. It is also known as Management Science and is closely related to the field of engineering called Industrial Engineering. Operations research began in earnest during World War II as an attempt to help the military allocate and transport resources more efficiently.

Operations research involves topics such as queuing theory, inventory theory, and simulations. Solving LP's is also a big part of OR. In OR, LP's are categorized according to the form of the model and the different categories are named according to the prototypical example of the form. In this and the next section we consider two very common types of LP's; the *transportation problem* and the *assignment problem*.

Example 7.3.1 (Delivering Bread)
The Better Bread Company has three bakeries located in the Midwest United States near wheat growing areas and four distribution warehouses scattered across the U.S. Management is studying ways to reduce shipping costs. Table 7.3 summarizes the weekly output of each bakery and the weekly allocation of each warehouse (in units of truckloads), along with the estimated shipping costs for each bakery-warehouse combination. Determine how many truckloads of bread should be assigned to each bakery-warehouse combination to minimize total cost.

TABLE 7.3

	Warehouse				
Bakery	**1**	**2**	**3**	**4**	**Output**
1	119	253	321	402	21
2	205	198	348	365	18
3	432	351	195	248	18
Allocation	15	10	12	20	

To form the model, let x_{ij} ($i = 1, 2, 3$; $j = 1, 2, 3, 4$) represent the number of truckloads to be shipped from bakery i to warehouse j. This may be formulated as a linear program as:

Minimize $C = 119x_{11} + 253x_{12} + \ldots + 248x_{34}$
Subject to

$$
\begin{array}{lllllllllllll}
x_{11} + x_{12} + x_{13} + x_{14} & & & & = 21 \\
& x_{21} + x_{22} + x_{23} + x_{24} & & & = 18 \\
& & x_{31} + x_{32} + x_{33} + x_{34} & & = 18 \\
x_{11} & + x_{21} & + x_{31} & & = 15 \\
\quad x_{12} & \quad + x_{22} & \quad + x_{32} & & = 10 \\
\qquad x_{13} & \qquad + x_{23} & \qquad + x_{33} & & = 12 \\
\qquad\quad x_{14} & \qquad\quad + x_{24} & \qquad\quad + x_{34} & & = 20
\end{array}
$$

$$x_{ij} \geq 0 \text{ for all } i = 1, 2, 3 \text{ and } j = 1, 2, 3, 4$$

The first constraint says that the total number of trucks coming out of bakery 1 must equal its total output. The next two constraints have similar meanings for bakeries 2 and 3. The fourth constraint says that the total number of trucks going to warehouse 1 must equal its total allocation. The next three constraints have similar meanings for warehouses 2, 3, and 4.

Notice the special pattern of coefficients in the constraints. It is this pattern that sets the transportation problem apart from others. Any problem that can be completely described by a parameter table like that in 7.3 will have this pattern and is thus called a "transportation problem" regardless of whether or not it has anything to do with transportation. Also notice that the total output from the bakeries equals the total allocation of the warehouses. In more general terminology we say the total *supply* from the *sources* equals the total *demand* from the *destinations*. This is a necessary condition for there to be a solution to this problem.

To solve this problem in Excel, rename a blank worksheet "**Bread**" and format it as in Figure 7.5. and format the Solver window as in Figure 7.7.

	B	C	D	E	F	G
2		**Warehouse**				
3	**Bakery**	**1**	**2**	**3**	**4**	**Output**
4	**1**	119	253	321	402	21
5	**2**	205	198	348	365	18
6	**3**	432	351	195	248	18
7	**Allocation**	15	10	12	20	

FIGURE 7.5

Next, format the worksheet as in Figure 7.6. The decision variables are held in the range **C12:F14**.

	B	C	D	E	F	G
10		**Warehouse**				
11	**Bakery**	**1**	**2**	**3**	**4**	**Total**
12	**1**					=SUM(C12:F12)
13	**2**					=SUM(C13:F13)
14	**3**					=SUM(C14:F14)
15	**Total**	=SUM(C12:C14)	=SUM(D12:D14)	=SUM(E12:E14)	=SUM(F12:F14)	
16						
17			**Total Cost**			
18			=SUMPRODUCT(C4:F6,C12:F14)			

FIGURE 7.6

Format the Solver window as in Figure 7.7.

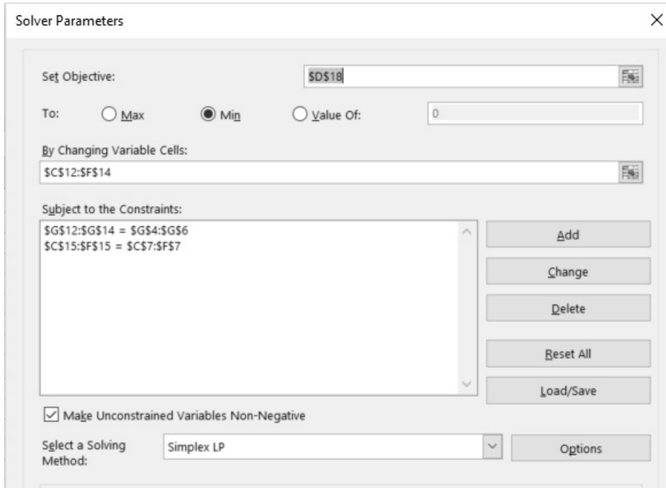


FIGURE 7.7

The solution is shown in Figure 7.8. It says that bakery 1 should ship 15 truckloads to warehouse 1 and 3 to warehouse 3. Bakery 2 should ship 10 truckloads to warehouse 2 and 8 to warehouse 4. Bakery 3 should ship 6 truckloads to warehouse 3 and 12 to warehouse 4.

	B	C	D	E	F	G
10		**Warehouse**				
11	**Bakery**	**1**	**2**	**3**	**4**	**Total**
12	**1**	15	0	6	0	21
13	**2**	0	10	0	8	18
14	**3**	0	0	6	12	18
15	**Total**	15	10	12	20	

FIGURE 7.8

□

Notice that in the optimal solution in Example 7.3.1, all the decision variables have integer values. This observation is generalized in the following theorem.

Theorem 7.3.1. *In a transportation problem, if all the supplies and demands have integer values, then all the decision variables in an optimal solution will have integer values.* □

Example 7.3.2 (Different Allocations)
Consider the same problem as in Example 7.3.1, but with slightly different warehouse allocation amounts as given in Table 7.4.

Notice that the total output is 57 and the total allocation is 54. Therefore, we cannot require all bakeries to produce all of their potential output since there isn't enough warehouse space to hold it all. We could model this problem as an LP using inequality constraints as:

TABLE 7.4

	Warehouse				
Bakery	**1**	**2**	**3**	**4**	**Output**
1	119	253	321	402	21
2	205	198	348	365	18
3	432	351	195	248	18
Allocation	13	10	12	19	

Minimize $C = 119x_{11} + 253x_{12} + \ldots + 248x_{34}$
Subject to

$$\begin{array}{llllr}
x_{11} + x_{12} + x_{13} + x_{14} & & & & \leq 21 \\
& x_{21} + x_{22} + x_{23} + x_{24} & & & \leq 18 \\
& & x_{31} + x_{32} + x_{33} + x_{34} & & \leq 18 \\
x_{11} & + x_{21} & + x_{31} & & = 13 \\
x_{12} & + x_{22} & + x_{32} & & = 10 \\
x_{13} & + x_{23} & + x_{33} & & = 12 \\
x_{14} & + x_{24} & + x_{34} & & = 19 \\
\end{array}$$

$$x_{ij} \geq 0 \text{ for all } i = 1, 2, 3 \text{ and } j = 1, 2, 3, 4$$

This LP could be solved with Solver using the Simplex method. However, the formulation above is not a transportation problem since there are some inequality constraints instead of all equality constraints. This may not seem like a big deal. However, there is a special version of the Simplex method, called the **Transportation Simplex method**, which is much more efficient at solving transportation problems than the generic Simplex method. For large problems with thousands of decision variables (as occur in real application), the Transportation Simplex method could save a great deal of computation time over the Simplex method. Therefore, if a problem can be modeled as a transportation problem, it is preferable to do so.

To model this problem as a transportation problem, we will introduce a "dummy" warehouse that will be allocated the excess 3 trucks of bread and given shipping costs of 0. The parameter table including this dummy warehouse is shown in Table 7.5.

TABLE 7.5

	Warehouse					
Bakery	**1**	**2**	**3**	**4**	**Dummy**	**Output**
1	119	253	321	402	0	21
2	205	198	348	365	0	18
3	432	351	195	248	0	18
Allocation	13	10	12	19	3	

Notice that in this version the total allocation equals the total output, so this version is a true transportation problem. Note that this problem has three additional decision variables. To solve this problem, we can simply modify the worksheet **Bread** to include this dummy warehouse (be sure to modify the formulas for the Totals and the Total Cost appropriately). The solution is shown in Figure 7.9.

	B	C	D	E	F	G	H
10		**Warehouse**					
11	**Bakery**	**1**	**2**	**3**	**4**	**Dummy**	**Total**
12	**1**	13	0	5	0	3	21
13	**2**	0	10	0	8	0	18
14	**3**	0	0	7	11	0	18
15	**Total**	13	10	12	19	3	

FIGURE 7.9

The solution says to ship 3 truckloads from bakery 1 to the dummy warehouse. This means in practice that bakery 1 should produce only 18 truckloads, 13 of which go to warehouse 1 and 5 of which go to warehouse 3. □

Example 7.3.3 (A New Requirement)
Consider the same scenario as in Example 7.3.2, except management is now requiring that bakery 1 produce exactly 21 truckloads. We could model this as before, except with the added constraint that the number of truckloads going from bakery 1 to the dummy warehouse is exactly 0, but then we wouldn't have a transportation problem.

Instead, we will use the *big-M method*. We will model the problem exactly as in Example 7.3.2, except we will assign a large cost, M, (in this case $M = 1000$) to a shipment from bakery 1 to the dummy warehouse. The parameter table is shown in Table 7.6.

TABLE 7.6

	Warehouse					
Bakery	**1**	**2**	**3**	**4**	**Dummy**	**Output**
1	119	253	321	402	1,000	21
2	205	198	348	365	0	18
3	432	351	195	248	0	18
Allocation	13	10	12	19	3	

Solving this requires us to change only this one cost value in the modified spreadsheet **Bread**. The solution is shown in Figure 7.10. Notice that no truckloads are shipped from bakery 1 to the dummy warehouse, as required. In this solution, bakery 2 should produce only 15 truckloads.

	B	C	D	E	F	G	H
10		**Warehouse**					
11	**Bakery**	**1**	**2**	**3**	**4**	**Dummy**	**Total**
12	**1**	13	0	8	0	0	21
13	**2**	0	10	0	5	3	18
14	**3**	0	0	4	14	0	18
15	**Total**	13	10	12	19	3	

FIGURE 7.10

□

Exercises

Directions: Formulate each problem below as a *transportation problem* by constructing an appropriate parameters table as in the "Data" section of Figure 7.5 and solve it with Solver. The parameter tables should meet the following requirements:

1. The sources are listed along the left-hand side.
2. The destinations are listed along the top.
3. The supply from each source is listed along the right-hand side.
4. The demand from each destination is listed along the bottom.
5. The total supply equals the total demand.

7.3.1 The manager of the produce department at a local supermarket buys his strawberries from two local suppliers, Sunnyside Farms and Green Valley Farms. The manager needs 3 cases of strawberries today and an additional 7 cases tomorrow. Sunnyside Farms can sell a maximum of 6 cases total at a price of \$7.25 per case today and \$6.35 per case tomorrow. Green Valley Farms can sell a maximum of 4 cases total at a price of \$5.75 per case today and \$5.25 per case tomorrow. How should the manager make his purchases to minimize the total cost while still meeting his daily requirements? (**Hint**: Let the sources be Sunnyside Farms and Green Valley Farms and the destinations be today and tomorrow. The costs would then be Sunnyside's price to sell today, Green Valley's price to sell tomorrow, etc.)

7.3.2 Supply and demand in the Armed Forces is extremely important. Suppose there are three warehouses that stock parts and fuel and five units which need supplies. The table below shows the transportation time (in hours) between each warehouse and unit, the demand (in tons) of each unit, and the supply (in tons) of each warehouse. Determine how to ship parts and fuel to minimize total transportation time.

	Unit					
Warehouse	**1**	**2**	**3**	**4**	**5**	**Supply**
1	5.5	6.5	8	7.5	8	5500
2	4.5	7.5	9	4.5	6	4000
3	10	6	7	8	7.5	5000
Demand	2500	3000	2500	3000	3500	

7.3.3 The Strikers Company manufactures bowling balls at its two plants in Orlando and Kansas City and ships them to distributors in six different states. The monthly output, demand, and unit shipping costs (in \$/bowling ball) are shown in the table below.

	Distributors						
Plants	**Ark.**	**Ala.**	**W. Vir.**	**Miss.**	**Ind.**	**Penn.**	**Output**
Orlando	0.5	0.35	0.6	0.45	0.8	0.75	3500
Kansas City	0.25	0.65	0.4	0.55	0.2	0.65	5000
Demand	1600	1800	1500	950	1250	1400	

a. Determine how the bowling balls should be shipped from the plants to the distributors to minimize the total shipping cost.

b. Suppose that management wants to find ways to reduce the unit shipping cost from Orlando to Indiana. How much does this cost need to be reduced before it has any impact on the minimum total shipping cost? Briefly explain your answer.

7.3.4 The Davenport Furniture Company ships student desks to schools from their three warehouses. They have received orders from four schools. The capacity of each warehouse, the quantities ordered, and unit shipping costs (in \$/desk) are shown in the table below.

	School				
Warehouse	**A**	**B**	**C**	**D**	**Capacity**
Dubuque	22	17	30	18	420
Des Moines	15	35	20	25	610
Omaha	28	21	16	14	340
Order	520	250	400	380	

a. Determine how the desks should be shipped from the warehouses to the schools to minimize the total shipping cost. Will the company be able to fulfill all the orders? Which school, if any, will not receive their full order?

b. To better meet demand, management is considering two options:

   1. Expand capacity at the Dubuque warehouse to 600 and increase the unit shipping costs from Dubuque to all the schools by \$5.
   2. Open a new warehouse in Peoria with a capacity of 300 and unit shipping costs of \$20, \$32, \$25, and \$26 to schools A, B, C, and D, respectively.

   Which of these two options would result in a lower total shipping cost?

7.3.5 The Rent-A-Jalopy company has six car rental locations in the tri-state area. At the end of a certain day, each location has several cars available and each needs a certain number for rental the next day. The number of cars available at each location, the number needed, and the drive times between each location (in minutes) are shown in the table below. Assuming the drive times are commutative, determine how cars should be moved between locations to minimize the total drive time.

	Location						
Location	**1**	**2**	**3**	**4**	**5**	**6**	**Available**
1	0	12	17	18	10	20	37
2		0	10	19	16	15	20
3			0	12	8	9	14
4				0	12	15	26
5					0	10	40
6						0	28
Needed	30	25	20	40	30	20	

7.3.6 The Great Openings Company, which manufactures doors and windows, is reassigning the production of three of its products to five of its plants. The costs to manufacture one unit of each product in each plant, the capacity of each plant, and the anticipated demand for each product are shown in the table below.

Plant	Product 1	2	3	Capacity
1	25	50	40	300
2	32	48	36	450
3	35	46	42	625
4	18	38	31	700
5	36	45	30	725
Demand	900	1100	600	

a. Determine how to assign the products to the plants to minimize total manufacturing cost.

b. Now suppose that management adds the requirements that plants 4 and 5 cannot manufacture product 2 and that plant 1 must manufacture exactly 300 units. Find the new optimal assignment.

7.3.7 The Better Bread Company also produces fruit cakes, most of which are sold during the holiday season. In September and October, the company ramps up their fruit cake production capacity, but then starts to cut back in November and December. The anticipated monthly demand, maximum production capacity, and production costs are shown in the table below (one unit is 1,000 fruit cakes and costs are in thousands of dollars). Fruit cakes can be produced in one month, stored, and then sold in another month, but there is a 0.0015 monthly cost to warehouse each unit. How should the monthly production be scheduled to minimize total cost while still meeting the anticipated demand? (**Hint:** Let the sources be the months produced and the destinations be the months sold. The cost c_{ij} would then be the total cost for a unit of fruit cakes that is produced in month i and sold in month j. For instance, $c_{11} = 1.18$ and $c_{14} = 1.1845$. Is it possible to produce a unit of fruit cakes in month 3 and sell it in month 1? What does this mean about c_{31}?)

Month	Anticipated Demand	Production Capacity	Unit Production Cost	Unit Storage Cost
Sept	5	35	1.18	0.0015
Oct	10	50	1.21	0.0015
Nov	40	40	1.19	0.0015
Dec	60	10	1.2	

7.3.8 In Theorem 7.3.1 we claimed that if the supplies and demands of a transportation problem are all integers, then the decision variables in an optimal solution will have integer values. To illustrate this idea, consider the following parameters table for the Better Bread Company where the supplies and demands are not integers. Find an optimal solution. Are the values of decision variables in the optimal solution all integers?

Bakery	Warehouse 1	2	3	4	Output
1	119	253	321	402	21.5
2	205	198	348	365	19.2
3	432	351	195	248	20.8
Allocation	14.8	11.7	12.6	22.4	

7.3.9 To further illustrate Theorem 7.3.1, consider the original parameters table for the Better Bread Company in Table 7.3. Replace some of the shipping costs with non-integer values of your choice and find an optimal solution. Are the values of decision variables in the optimal solution all integers?

7.3.10 In this exercise we will see that a transportation problem can have more than one optimal solution. Consider a bakery-warehouse scenario similar to Example 7.3.1, but with the unit costs, outputs, and allocations shown in the table below.

	Warehouse			
Bakery	**1**	**2**	**3**	**Output**
1	12	18	14	250
2	15	17	13	100
Allocation	150	80	120	

a. Use Solver to find an optimal solution to this problem.

b. Consider the solution shown in the table below. Verify that this solution is feasible. That is, verify the output and allocation constraints are met. Also, calculate the total cost of this solution and compare it to the optimal solution from part a. Is this solution optimal?

	Warehouse		
Bakery	**1**	**2**	**3**
1	150	0	100
2	0	80	20

7.3.11 In this section we solve transportation problems with the Simplex method using Solver. There are also iterative methods for solving these problems by hand. We won't discuss the full details of these methods, but we will discuss the first step which is to find an *initial feasible solution.* Consider Example 7.3.1. An initial feasible solution is an assignment of truckloads from the bakeries to the warehouses that meets the output and allocation requirements. The solution does not have to be optimal, just feasible. One method for finding an initial solution is called the *north-west corner method*. This method says we start in the upper left-hand corner of the parameters table (i.e. the north-west corner) and assign units as follows:

1. Assign all the supply of each row before moving down to the next row.

2. Meet the allocation requirement of each column before moving to the right to the next column.

In Example 7.3.1, this means we need to find an assignment of truckloads from the bakeries to the warehouses that meet the allocation and output constraints. To apply the north-west corner method, we first assign 15 truckloads from bakery 1 to warehouse 1. This satisfies the allocation constraint for warehouse 1. Next we assign 6 truckloads from bakery 1 to warehouse 2. This satisfies the output constraint for bakery 1. We assign 0 truckloads from bakery 1 to warehouses 3 and 4. We then move to bakery 2 and assign 0 truckloads to warehouse 1, 4 to warehouse 2, 12 to warehouse 3, and 2 to warehouse 4. We continue with bakery 3 in the same fashion. The resulting assignments are shown in the table below. We see that this solution is indeed feasible, but not optimal since the total cost is \$13,465 which is greater than the cost of the optimal solution found with Solver, \$12,757.

	Warehouse				
Bakery	**1**	**2**	**3**	**4**	**Total**
1	15	6	0	0	21
2	0	4	12	2	18
3	0	0	0	18	18
Total	15	10	12	20	

As an application of the north-west corner method, consider a different bakery-warehouse scenario whose unit costs, outputs, and allocations are shown in the table below.

	Warehouse			
Bakery	**1**	**2**	**3**	**Output**
1	2	6	1	30
2	2	4	1	60
3	6	9	5	20
Allocation	60	20	30	

a. Use the north-west corner method to find an initial feasible solution to this problem.

b. Use Solver to find an optimal solution to this problem. Compare this solution to that found in part a.

7.3.12 (This is not a transportation problem, so no parameters table is needed, but the decision variables do have similar meanings as in a transportation problem.) The Nutty Goodness Company is considering selling three new mixtures of peanuts, walnuts, and cashews. They have 1,000 lb of peanuts, 800 lb of walnuts, and 700 lb of cashews available for the first batch. Each mixture has a unique set of specifications as to the percentage of each type of nut as shown in the table below.

Mixture	Specifications	Selling price per pound
Walnut Lover's	At least 35% walnuts At most 25% cashews No restriction on peanuts	\$1.85
Cashew Lover's	At most 45% peanuts At least 45% cashews No restriction on walnuts	\$1.99
Premium	At most 10% peanuts Between 50% and 65% walnuts At least 15% cashews	\$2.75

If peanuts cost \$0.75 per lb, walnuts cost \$1.05 per lb, and cashews cost \$1.75 per lb, determine how much of each mixture they should make (and the amount of each type of nut in each mixture) to maximize total profit. (**Hint**: Let x_{ij} = amount of nut i in mixture j. In your spreadsheet, set up different cells to calculate the total amount of each mixture, the total amount of each nut in each mixture, and the percentage of each type of nut in each mixture. Also calculate total cost and total revenue. Remember, profit = revenue – cost.)

7.4 The Assignment Problem and Binary Constraints

The *Assignment Problem* (AP) is a special type of linear programming problem where we generically say that *workers* are being assigned to perform *jobs*. The simplest example is when employees are being assigned different types of jobs to perform. However, the workers may not always be people and the jobs may not always be literal jobs to perform. In Example 7.4.1 below, the workers are trucks and the jobs are packages that need to be delivered.

In an AP, the decisions to be made are which workers should perform which jobs in order to minimize some cost function. A problem is called an AP if and only if it satisfies the following assumptions (see Hillier, F. and G. Lieberman, *Introduction to Operations Research,* Seventh Edition, McGraw Hill, Boston, 2001, p. 382 for more information):

1. The number of workers, n, equals the number of jobs.
2. Each worker is to be assigned exactly one job.
3. Each job is to be performed by exactly one worker.
4. There is a cost c_{ij} associated with assigning worker i to job j.
5. The objective is to determine how all n assignments should be made to minimize the total cost.

The cost in the fourth assumption may be a literal dollar amount associated with assigning worker i to job j, or it may be some other value associated with that assignment (i.e. time), the total of which we want to minimize. The decision variables, x_{ij} for i, $j = 1, 2, \ldots, n$, are binary (they equal 0 or 1) with

$$x_{ij} = \begin{cases} 1 & \text{if worker } i \text{ performs job } j \\ 0 & \text{if not} \end{cases}$$

Assignment problems can be solved extremely efficiently by a special form of the transportation simplex method (which is in turn a special form of the simplex method).

Example 7.4.1 (On–Time Delivery Company)
A scheduler at the On–Time Delivery Company has three packages that need to be delivered and four available trucks of different types. Based on the locations of the packages and the locations of the trucks, the scheduler has determined a cost for each truck to deliver a package as shown in Table 7.7 (note truck 2 cannot deliver package 2). Such a table is called a *cost matrix*.

TABLE 7.7

	Package		
Truck	**1**	**2**	**3**
1	12	20	13
2	16	–	18
3	8	6	9
4	14	10	8

The goal is to determine how to assign trucks to packages to minimize the total delivery costs. Note that this is technically not an AP since there are more workers than jobs, so

TABLE 7.8

	Package				
Truck	**1**	**2**	**3**	**Dummy**	**Supply**
1	12	20	13	0	1
2	16	100	18	0	1
3	8	6	9	0	1
4	14	10	8	0	1
Demand	1	1	1	1	

we will add a dummy package. Each truck will supply one delivery and each package will demand one delivery. Also, we will use the big-M method to model the fact that truck 2 cannot deliver package 2. The resulting parameters table is shown in Table 7.8.

Notice that this is an appropriate parameter table for a transportation problem. So an AP is indeed a transportation problem. However, an AP has the additional feature that the supplies and the demands are all 1. This special feature can be exploited to find an algorithm that solves these problems extremely efficiently. Note that because the AP is a transportation problem, the supplies and demands are all integers, and the special structure of the problem, the optimal solution will be binary (i.e. we don't have to add the constraints that the decision variables are binary; we get this for free).

Solving this problem in the same way we did in Section 7.3 yields the solution shown in Figure 7.11. This shows that Truck 1 should deliver package 1, Truck 3 should deliver package 2, Truck 4 should deliver package 3, and Truck 2 should not deliver any package. Note that the big-M method prevented truck 2 from delivering package 2, as desired. However, the big-M method does not always prevent such unfeasible assignments (see Exercise 7.4.14).

	B	C	D	E	F	G
11				**Package**		
12	**Truck**	**1**	**2**	**3**	**Dummy**	**Sum**
13	**1**	1	0	0	0	1
14	**2**	0	0	0	1	1
15	**3**	0	1	0	0	1
16	**4**	0	0	1	0	1
17	**Sum**	1	1	1	1	
18		**Total Cost =**		26		

FIGURE 7.11

□

The next example illustrates a use of binary decision variables in a problem that involves assignments, but is not technically an assignment problem

Example 7.4.2 (Ace Manufacturing)
Ace Manufacturing Co. plans to start manufacturing four new products in three existing plants that have excess capacity. Table 7.9 shows the daily cost of producing one unit of each product in each plant (plant 2 cannot produce product 3), the available daily capacity of each plant, and the expected daily demand for each product.

TABLE 7.9

Plant	Product 1	2	3	4	Capacity
1	35	32	30	40	78
2	35	25	-	31	65
3	33	37	36	30	45
Demand	25	35	32	43	

Management wants each product to be produced in exactly one plant (e.g. we can't produce 12 units of product 1 in plant 1 and 13 units in plant 3). How should they assign products to plants to minimize the total production cost while still meeting demand?

Observe that plant 1 has enough capacity to produce any two products. Plant 2 could produce both products 1 and 2, but not both 3 and 4. Plant 3 has enough capacity to produce only one product. These restrictions make it complicated to model this problem as an AP, so instead we will model it as a linear program with binary constraints.

The first step in modeling this problem is to enter the unit costs, capacities, and demands. Rename a blank workbook **Ace** and format it as in Figure 7.12. Note that we use the big-M method to model the fact that plant 2 cannot produce product 3.

	B	C	D	E	F	G
1		**Product**				
2	**Plant**	**1**	**2**	**3**	**4**	**Capacity**
3	**1**	35	32	30	40	78
4	**2**	35	25	1000	31	65
5	**3**	33	37	36	30	45
6	**Demand**	25	35	32	43	

FIGURE 7.12

Now we need to calculate the total cost of assigning each product to each plant. This assignment cost is the demand times the unit cost. To make these calculations, add the formulas in Figure 7.13.

	B	C	D	E	F	G
8		**Product**				
9	**Plant**	**1**	**2**	**3**	**4**	**Capacity**
10	**1**	=C3*C$6	=D3*D$6	=E3*E$6	=F3*F$6	=G3
11	**2**	=C4*C$6	=D4*D$6	=E4*E$6	=F4*F$6	=G4
12	**3**	=C5*C$6	=D5*D$6	=E5*E$6	=F5*F$6	=G5
13	**Demand**	1	1	1	1	

FIGURE 7.13

To model the constraints and objective function, add the formulas in Figure 7.14. The binary decision variables are held in the range **C17:F19**. The formulas in column **Total** calculate the total number of units produced in each plant. These totals must be less than or equal to the capacities. The formulas in row **Sum** calculate the number of plants each product is assigned to. Each of these sums must equal 1.

Set up the Solver parameters as in Figure 7.15. Note that we need the binary constraint because this problem is not an AP. In Section 8.7 we will examine the basics of how Solver

	B	C	D	E	F	G
15				**Product**		
16	**Plant**	**1**	**2**	**3**	**4**	**Total**
17	**1**					=SUMPRODUCT(C17:F17,$C$6:$F$6)
18	**2**					=SUMPRODUCT(C18:F18,$C$6:$F$6)
19	**3**					=SUMPRODUCT(C19:F19,$C$6:$F$6)
20	**Sum**	=SUM(C17:C19)	=SUM(D17:D19)	=SUM(E17:E19)	=SUM(F17:F19)	
21			**Total Cost =**	=SUMPRODUCT(C10:F12,C17:F19)		

FIGURE 7.14

handles this binary constraint. (Before running Solver, click on **Options** and make sure that **Ignore Integer Constraints** in the **All Methods** tab is not checked.)

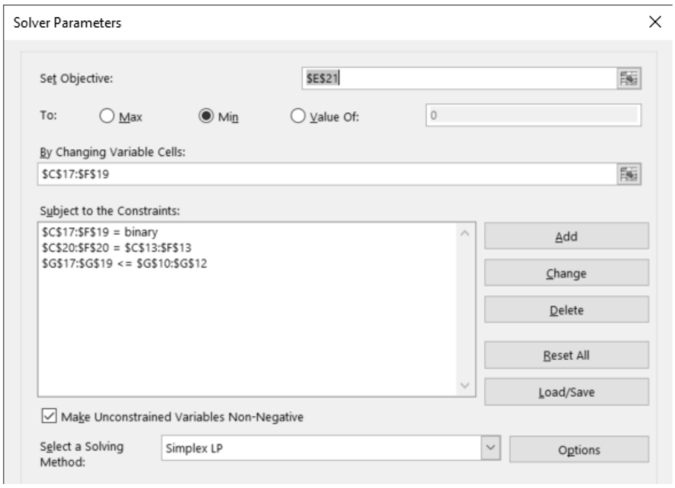


FIGURE 7.15

The solution is shown in Figure 7.16. It indicates that plant 1 should produce product 3, plant 2 should produce products 1 and 2, and plant 3 should produce product 4.

	B	C	D	E	F	G
15				**Product**		
16	**Plant**	**1**	**2**	**3**	**4**	**Total**
17	**1**	0	0	1	0	32
18	**2**	1	1	0	0	60
19	**3**	0	0	0	1	43
20	**Sum**	1	1	1	1	
21			**Total Cost =**	4000		

FIGURE 7.16

□

The last example illustrates a much different use of binary decision variables to model yes-no decisions.

Example 7.4.3 (Home Improvement Decisions)
Nathan and Laura are trying to sell their house which has two bedrooms and two bathrooms. To increase the house's value, they want to remodel one or more rooms. They have estimated

the costs of remodeling each room and their real estate agent has estimated the increase in the house's value if each room was remodeled as shown in Table 7.10 (where costs and increases in values are given in thousands of dollars). They have only \$10,000 to spend remodeling, and they have decided that they can't do both bathroom 2 and bedroom 2. They will only do bathroom 2 if they also do bathroom 1. Also, they will only do bedroom 2 if they also do bedroom 1. Which rooms should they remodel to maximize the total increase in their house's value?

TABLE 7.10

Room	**Decision Variable**	**Remodeling Cost**	**Increase in House Value**
Bathroom 1	x_1	6	9
Bedroom 1	x_2	3	5
Bathroom 2	x_3	5	6
Bedroom 2	x_4	2	4

There are four decisions to make in this problem: Do they remodel Bathroom 1?, Do they remodel Bedroom 1?, etc. So each one of the four decision variables will equal 1 if the associated decision is yes and 0 if the decision is no.

Our objective is then to maximize

$$Z = 9x_1 + 5x_2 + 6x_3 + 4x_4$$

The fact that they have only \$10,000 to spend means that

$$6x_1 + 3x_2 + 5x_3 + 2x_4 \leq 10$$

Since they can't do both bathroom 2 and bedroom 2, we can't have $x_3 = x_4 = 1$. So in terms of inequalities, we have

$$x_3 + x_4 \leq 1$$

Since they will only do bathroom 2 if they also do bathroom 1, we can only have $x_1 = 1$ and $x_3 = 1$, or $x_1 = 1$ and $x_3 = 0$, or $x_1 = 0$ and $x_3 = 0$. In terms of inequalities, we have

$$x_3 \leq x_1 \;\Rightarrow\; -x_1 + x_3 \leq 0$$

Likewise, since they will only do bedroom 2 if they also do bedroom 1, we have

$$x_4 \leq x_2 \;\Rightarrow\; -x_2 + x_4 \leq 0$$

So, putting it all together we get the program

$$\begin{array}{rrrrrrrrrl}
\textbf{Maximize} & Z = & 9x_1 & + & 5x_2 & + & 6x_3 & + & 4x_4 & \\
\textbf{Subject to} & & 6x_1 & + & 3x_2 & + & 5x_3 & + & 2x_4 & \leq 10 \\
 & & & & & & x_3 & + & x_4 & \leq 1 \\
 & & -x_1 & & & + & x_3 & & & \leq 0 \\
 & & & & -x_2 & & & + & x_4 & \leq 0 \\
 & & & & \multicolumn{6}{l}{x_1,\ x_2,\ x_3,\ x_4 \text{ are binary}}
\end{array}$$

Such a program is called a *binary integer program* (*BIP*). Notice that this program does not fit the form of a transportation problem, so we do need the additional binary constraints. To solve this in Excel, rename a blank worksheet **Improvement** and format it as in Figure 7.17, and format the Solver window as in Figure 7.18.

The results are shown in Figure 7.19. They indicate that Nathan and Laura should remodel bedroom 1 and bathroom 1 which will increase their house's value by \$14,000.

□

	A	B	C	D	E	F	G	H	I
1		**Bath 1**	**Bed 1**	**Bath 2**	**Bed 2**				
2	**Variable**	**x_1**	**x_2**	**x_3**	**x_4**				
3	**Values**	**0**	**0**	**0**	**0**				
4	**Z =**	9	5	6	4	=	=SUMPRODUCT($B$3:$E$3,B4:E4)		**RHS**
5	**Constraint 1**	6	3	5	2	=	=SUMPRODUCT($B$3:$E$3,B5:E5)	≤	10
6	**Constraint 2**			1	1	=	=SUMPRODUCT($B$3:$E$3,B6:E6)	≤	1
7	**Constraint 3**	-1		1		=	=SUMPRODUCT($B$3:$E$3,B7:E7)	≤	0
8	**Constraint 4**		-1		1	=	=SUMPRODUCT($B$3:$E$3,B8:E8)	≤	0

FIGURE 7.17

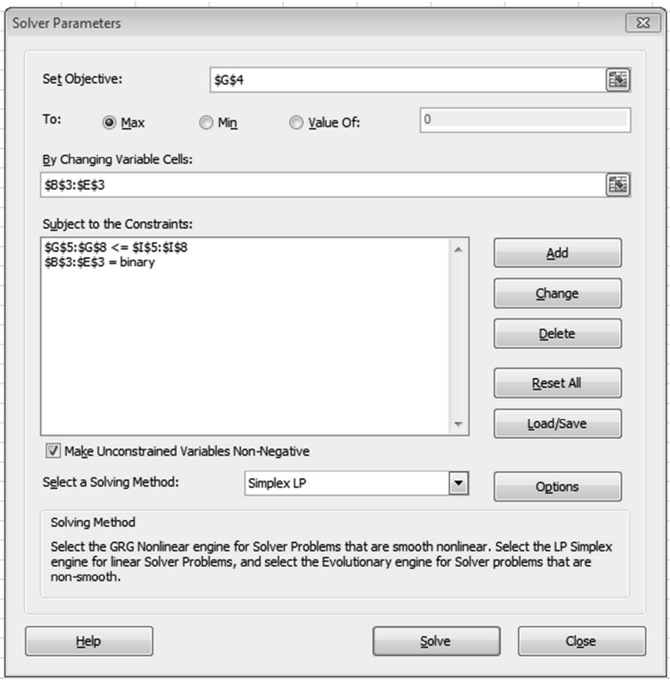


FIGURE 7.18

	A	B	C	D	E	F	G	H	I
1		**Bath 1**	**Bed 1**	**Bath 2**	**Bed 2**				
2	**Variable**	**x_1**	**x_2**	**x_3**	**x_4**				
3	**Values**	**1**	**1**	**0**	**0**				
4	**Z =**	9	5	6	4	=	14		**RHS**
5	**Constraint 1**	6	3	5	2	=	9	≤	10
6	**Constraint 2**			1	1	=	0	≤	1
7	**Constraint 3**	-1		1		=	-1	≤	0
8	**Constraint 4**		-1		1	=	-1	≤	0

FIGURE 7.19

Exercises

Directions: Model each problem below using binary decision variables and solve it with Solver.

7.4.1 The Keep-U-Clean janitorial service has four employees: Arlynn, Sharon, Jeff, and Nick. One of the buildings they clean requires four jobs: vacuuming, mopping, cleaning glass, and dusting. The time (in minutes) for each employee to perform each job is shown in the table below. Determine how to assign employees to jobs to minimize the total amount of time needed, assuming each employee does exactly one job.

Employee	Vacuum	Mop	Glass	Dust
Arlynn	13	4	7	6
Sharon	2	11	5	4
Jeff	6	7	2	8
Nick	2	4	5	9

7.4.2 The final round of a mathematics competition consists of four tests: arithmetic, algebra, calculus, and geometry. Each team is supposed to assign exactly 1 member to take each test. The team consisting of Brian, Ed, John, Larry, and Bruce has personal-best scores on each test as shown in the table below. How should they assign themselves to the tests so that no one takes more than one test and the sum of the corresponding scores is maximized?

Test	Brian	Ed	Larry	John	Bruce
Arithmetic	99.1	96.3	97.6	98.9	98.5
Algebra	99.3	98.9	98.2	99.2	98.8
Calculus	94.6	98.3	97.4	97.2	98.2
Geometry	95.3	98.5	98.6	98.5	94.2

7.4.3 The Buggy Bath Car Wash Company has five employees: Jared, Dustin, Matthew, Daniel, and Tim. One particular customer requires her car be washed, waxed, have the interior detailed, and have the wheels polished. The estimated time (in minutes) required for each employee to complete each task is shown in the table below where a - indicates the employee refuses to do that task. Determine how to assign employees to tasks to minimize the total time needed, assuming each employee does at most one task.

Employee	Wash	Wax	Interior	Wheels
Jared	23	19	31	19
Dustin	18	-	28	23
Matthew	27	21	29	29
Daniel	17	23	-	15
Tim	22	-	26	29

7.4.4 The coach of a high school swim team needs to form male and female medley relay teams. The medley consists of four different strokes, back, breast, butterfly, and freestyle. The members of the swim team, their genders, and their personal best times (in seconds) for each stroke are shown in the table below. Each team member may be assigned at most one stroke. Find the female team that minimizes the total relay swim times. Then do the same for the males.

	Gender	Back	Breast	Butterfly	Freestyle
Lilly	F	35.8	37.1	27.4	35.5
Kelly	F	36.3	32.2	28	25.6
Jenny	F	33.2	35.1	28	29.2
Angie	F	29.1	34.7	30.9	29.4
Colby	F	33.6	30.8	26.7	24.3
Elsie	F	27.2	32.8	28.7	27.2
James	M	25.3	29	25	23.1
Michael	M	27.8	25.8	28.2	25.7
Edward	M	28.8	32.5	26.8	26
Chris	M	26.8	28.7	22.2	21.5

7.4.5 A farmer, Stan, is going to harvest wheat on three different fields. He can haul the grain to two different elevators, the Farmers United elevator, which pays \$4.50 per bushel and can only accept 1,800 bushels of wheat; and the Sunflower elevator, which pay \$4.30 per bushel and can only accept 1,000 bushels. He predicts that fields 1, 2, and 3 will produce 1,000, 500, and 1,000 bushels, respectively. His grain trucks can haul 500 bushels each. Assume that only full trucks will be used to haul the wheat to the elevators. The price to haul one bushel from each field to each elevator is shown in the table below.

	Field		
Elevator	**1**	**2**	**3**
United	0.13	0.13	0.15
Sunflower	0.16	0.13	0.17

Stan needs to determine how much to haul from each field to each elevator to maximize the total profit. (**Hint**: The profit for a bushel is the price paid at the elevator minus the hauling cost. Consider having five jobs (or truckloads of wheat) and five workers. The cost for each assignment is the profit for each truckload.)

7.4.6 A manufacturing company needs to assign the production of four new products (products 1–4) to four existing plants (plants A–D). Production costs and sales revenue differ between plants. The costs and revenues (in thousands of dollars per product) are shown in the table below. Find an assignment of products to plants that maximizes total profit.

	Cost				Revenue			
Plant	**1**	**2**	**3**	**4**	**1**	**2**	**3**	**4**
A	49	65	40	48	59	69	50	57
B	53	55	38	72	56	70	44	81
C	48	56	42	53	59	73	47	75
D	61	61	54	71	68	63	62	87

7.4.7 Suppose Nathan and Laura are going to start their remodeling project. They have decided to do the drywall, painting, trim work, and finish plumbing themselves. They want to divide these four tasks between them so that each has exactly two tasks, but the total time they take is kept to a minimum. Each person has estimated the amount of time he/she will take for each task as shown in the table below.

	Drywall	Painting	Trim Work	Finish Plumbing
Nathan	10.0	8.5	9.5	3.0
Laura	9.5	6.0	9.0	4.0

Laura refuses to do both the painting and the drywall. Nathan won't do the finishing plumbing unless Laura does the trim work. How should they divide the tasks among themselves? Formulate this problem using binary decision variables and solve it with Solver. (**Hint**: Consider 8 decision variables, one for each person–task combination. You will need two constraints to make sure that each person does exactly two tasks. Four are needed to make sure that each task gets assigned. Two more are necessary for the additional requirements.)

7.4.8 The Janitor of the prestigious Lied Center finds himself at a dilemma. The president is going to show up the next day to watch the hit musical *Phantom of the Operations*. With a rampant case of bird flu causing all of the other janitors to call in sick, the janitor has the whole Lied Center to clean by himself. This involves the stage, balcony, bathrooms, lobby, light fixtures, and hallways. He only has 8 hours until the president arrives. As a last attempt for help, he calls his mom to help him clean. She agrees, but she can only clean for 6 hours, and she is a little slower in some rooms. Because of the layout of the rooms there are some weird requirements. They can only clean the bathrooms if they also clean the balcony. They cannot clean both the lobby and the halls. Finally, they will only clean the light fixtures if they clean the bathrooms and the stage, and only one person can clean a room. Each task gives them a nice boost in self esteem. The table below shows the time needed for each person to clean each room (in hours) and the boost in self esteem. Where should they clean to maximize the total boost in self esteem? Will every room get cleaned? (problem written by Paul Hammes, 2019)

	Janitor		**Mom**	
Room	**Time**	**Boost**	**Time**	**Boost**
Stage	2	2	3	3
Balcony	1	1	2	2
Bathrooms	3	3	1	1
Lobby	3	2	2	3
Lights	2	1	2	1
Hallway	3	3	4	2

7.4.9 A truck with a weight capacity of 16,000 lb and a volume capacity of 1,500 ft^3 is at a loading dock waiting to be loaded where six items are awaiting shipment. The dollar value, weight, and volume per pound of each item are shown in the table below.

Item	**Value ($)**	**Weight (lb)**	**Volume/lb (ft^3/lb)**
1	15000	5000	0.125
2	14000	4500	0.065
3	10400	3000	0.144
4	14250	3500	0.45
5	13000	4000	0.05
6	9750	3500	0.02

a. Determine which item(s) should be loaded on the truck to maximize the total value of the items loaded within the capacity of the truck.

b. Suppose there are two smaller trucks available, each with an 11,000 lb weight capacity and a 950 ft^3 volume capacity. Determine how items should be loaded onto these two trucks to maximize the total value of the items loaded within the capacity of each truck.

c. Consider the two trucks in part b., but add the requirement that there be the same total weight in each truck.

d. If it costs an additional \$5,000 to operate two trucks rather than one, is there any benefit to using two trucks in either part b. or part c. (i.e. is the increase in total value of items loaded greater than the additional expense)? Briefly explain your answer.

7.4.10 As part of its accreditation process, St. John School needs to form 7 committees to review different aspects of the school. Each committee consists of teachers, parents, and education board members. The 12 members of the education board have indicated preferences for their own membership on each committee in the form of rankings as shown in the table below (a ranking of 1 means that committee is most preferred, a 7 means least preferred). Determine how to best assign board members to committees so that each board member is on exactly one committee and each committee contains exactly one or two board members. (**Suggestion**: Consider each ranking a cost. Minimize the total cost.)

	Board Member											
Committee	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	3	6	5	6	1	5	2	1	1	2	2	1
2	5	4	3	2	2	3	3	7	7	5	7	4
3	6	5	2	5	6	6	5	2	5	7	3	3
4	2	1	1	4	5	1	4	3	6	6	5	7
5	1	2	7	7	3	7	7	4	3	1	1	5
6	4	7	6	3	4	2	1	5	2	3	6	6
7	7	3	4	1	7	4	6	6	4	4	4	2

7.4.11 The city of Colby has just consolidated all its public high schools into three large schools, and the school board is trying to decide how to assign students from the six different areas of the city to the schools in order to minimize busing costs. The number of students in each area, the busing costs, and the capacities of the schools are shown in the table below. (A cost of 0 indicates that students in that area can walk to school, so no busing is needed while a – indicates that assignment is unfeasible.)

	# of	**Busing Cost (\$/student)**		
Area	**Students**	**School 1**	**School 2**	**School 3**
1	450	300	0	700
2	600	–	400	500
3	550	600	300	200
4	350	200	500	–
5	500	0	–	400
6	450	500	300	0
	Capacity	900	1100	1000

Determine how to assign areas to schools to minimize the total busing cost for all students while meeting the capacity constraints of the schools. (Note that this problem does not need to be modeled as an assignment problem, but the decision variables should be binary.)

7.4.12 A professor needs to assign each student in a class of 14 a partner for the final project. She asks each student to give a compatibility rating between 0 and 10 for each of the other students in class. A higher rating indicates a higher level of compatibility. Each student gives themselves a rating of 0 since they can't be assigned to themselves. The results are shown in the rows of the table below (data collected by Joshua Hendrickson, 2019). Note that this table is not symmetric because students' feelings toward each other are not necessarily commutative. The total compatibility rating of an assignment is the sum of the ratings of the two students. For example, if student 1 were assigned to student 2, the total rating would be $8 + 7 = 15$. Determine how to assign partners to maximize the sum of the total compatibility ratings over all assignments.

	Rating of Other Students													
Student	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
1	0	8	4	5	3	4	2	2	7	6	5	3	2	6
2	7	0	10	5	2	4	3	4	2	1	1	1	9	4
3	3	1	0	1	7	7	7	3	5	8	2	7	2	5
4	4	8	5	0	2	6	4	8	6	1	8	5	4	9
5	10	10	3	8	0	8	9	2	4	9	7	10	9	4
6	1	4	10	4	7	0	5	7	4	2	8	10	3	7
7	5	6	8	10	2	2	0	8	7	4	1	10	6	2
8	1	8	2	9	9	3	6	0	4	8	7	8	1	7
9	5	3	6	9	10	6	10	3	0	4	7	4	5	6
10	8	5	7	3	4	4	8	10	7	0	2	7	9	10
11	6	4	6	4	8	1	9	1	7	5	0	3	4	4
12	1	5	4	6	10	5	7	6	8	8	1	0	1	1
13	4	5	5	7	7	3	9	7	1	2	10	9	0	4
14	7	8	4	8	1	1	3	2	3	6	9	1	7	0

7.4.13 There are 12 prison cells arranged along the outside walls of a prison block as illustrated in Figure 7.20. One of the cells contains exactly 1 prisoner, one contains exactly 2 prisoners, and so on. Figure out how to place the prisoners in the cells so that there are exactly 25 prisoners along each wall. (**Suggestion**: Treat this like an assignment problem where we assign the group of 1 prisoner to a cell, the group of 2 prisoners to a cell, and so on. Consider the cost of each possible assignment to be \$1. Add constraints to guarantee that there are 25 prisoners along each wall.)

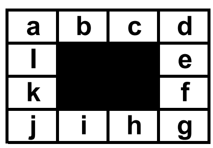


FIGURE 7.20

7.4.14 The cost matrix for an AP is shown below where a – indicates the assignment is unfeasible. Try using the big-M method to solve this AP. Can you find an M big enough to prevent all unfeasible assignments? What does this say about the existence of a solution to this AP?

	Job				
Worker	**1**	**2**	**3**	**4**	**5**
A	3	–	8	–	8
B	4	7	15	18	8
C	8	12	–	–	12
D	5	5	8	3	6
E	10	12	15	10	–

7.4.15 A prisoner of war camp consists of 20 prison cells arranged in 4 rows and 5 columns. The daily cost of keeping a prisoner in each cell is shown in the table below. Only one prisoner may be kept per cell.

	Column				
Row	**1**	**2**	**3**	**4**	**5**
1	2	4	3	5	1
2	6	2	3	1	5
3	3	5	6	4	2
4	1	2	6	4	5

Suppose that 9 prisoners are taken during a battle, 8 of which are normal prisoners and one is a really "bad guy." The bad guy is so bad that it is not safe to keep a prisoner in any cell adjacent to his. For instance, if the bad guy is in row 3, column 4, denoted (3, 4), then no prisoner may be kept in (3, 3), (2, 4), (3, 5) or (4, 4). Cells that share a corner with his, such as (2, 5) may hold a prisoner.

a. Suppose the bad guy is kept in cell (1, 1). Find an assignment of the other 8 prisoners to the remaining cells that minimizes the total cost.

b. Repeat part a. for every other possible location of the bad guy.

c. Which of the possibilities from part b. minimizes the total cost including the bad guy?

7.4.16 As mentioned in the text, an AP is a special type of transportation problem. One significance of this fact is that techniques used for solving transportation problems can also be used to solve AP's. For instance, we can use the north-west corner method as described in Exercise 7.3.11 to find an initial feasible solution to an AP. The only difference is that instead of assigning a certain number of units in each step, we simply make an assignment (or don't make an assignment). Apply the north-west corner method to find an initial feasible solution to the AP in Example 7.4.1 and compare it to the optimal solution. Make sure that you only make feasible assignments.

7.4.17 The following theorem about assignment problems is true:

> If a number (positive or negative) is added to all the entries of any row or column of the cost matrix of an assignment problem, then an optimal assignment for the resulting cost matrix is also an optimal assignment for the original cost matrix.

Demonstrate this theorem to be true by selecting a number to add to a row or column of the cost matrix in Example 7.4.1. Find an optimal solution for the resulting cost matrix using Solver and compare it it to the original optimal solution.

7.4.18 The Hungarian method is an algorithm for solving assignment problems that is based on the theorem in Exercise 7.4.17. Research this method and apply it to one or more assignment problems from this section.

7.5 Solving Linear Programs

In Section 7.2 we looked at the following linear problem:

> The manager of a produce department at a neighborhood supermarket is making fruit baskets for the busy holiday season. He sells two sizes of baskets: small and large. He has only 200 apples and 100 oranges remaining and is trying to decide how many of each size of basket he should make. Each small basket returns a profit of \$3 and requires 3 apples and 1 orange while each large basket returns a profit of \$4

and requires 2 apples and 2 oranges. Assuming he will sell all that he makes, how many of each size should he make to maximize his profit?

We modeled this problem with the linear program:

$$\begin{array}{rr} \textbf{Maximize} & P = 3s + 4l \\ \textbf{Subject to} & 3s + 2l \le 200 \\ & 1s + 2l \le 100 \\ & s,\ l \ge 0 \end{array} \tag{7.4}$$

where s = the number of small baskets to produce and l = the number of large baskets to produce. Solver gave an optimal solution of 50 small and 25 large baskets, which gives a maximum profit of \$250. In this section we will look at a graphical and an algebraic way to solve this program. In the next section we will examine how the Simplex method works.

Example 7.5.1 (Graphical Solution)
To examine a graphical solution to the model (7.4), open the worksheet "**Graphical Solution**" in the workbook **Linear Programming**, found on the website for the book, and follow these steps:

1. Enter the names of the decision variables, the objective functions, and the first two constraints as shown in Figure 7.21 (don't worry about the non-negativity constraints).

Decision Variables

s & l

Profit =	3	**s +**	4	**l**		
Constraint 1:	3	**s +**	2	**l**	**≤**	200
Constraint 2:	1	**s +**	2	**l**	**≤**	100

FIGURE 7.21

2. The resulting graph is shown in Figure 7.22 (s is on the horizontal axis and l is on the vertical axis). The graph contains the lines $3s + 2l = 200$ and $s + 2l = 100$, called the *constraint lines* (note that the non-negativity constraints mean that the s- and l-axis are also constraint lines). The area below the first line is the set of all points (i.e. combinations of numbers of small and large baskets) that satisfy the first constraint. Points below the second line satisfy the second constraint. Points below *both* lines satisfy *both* constraints and constitute what is called the *feasible region*. These points are all the feasible solutions.

3. Move the slider under **Desired Profit** to the right until the Desired Profit is 170.00. Move the slider under **Desired Schedule** to the right until $s = 22.0$ and $l = 26.0$. The resulting graph is shown in Figure 7.23. The thick black line labeled "Profit," called a *level curve*, is the set of all points which will give a profit of \$170.00. Notice that this line intersects the feasible region. This means that it *is* possible to produce a combination of small and large baskets that gives a profit of \$170.00. Specifically, the point $s = 22.0$ and $l = 26.0$ is on this level curve and in the feasible region. This means it is possible to produce 22 small baskets and 26 large baskets and profit \$170.00.

4. Continue to move the slider under **Desired Profit** to the right to increase the desired profit. Notice that for profits above \$250.00, the level curve does not intersect the

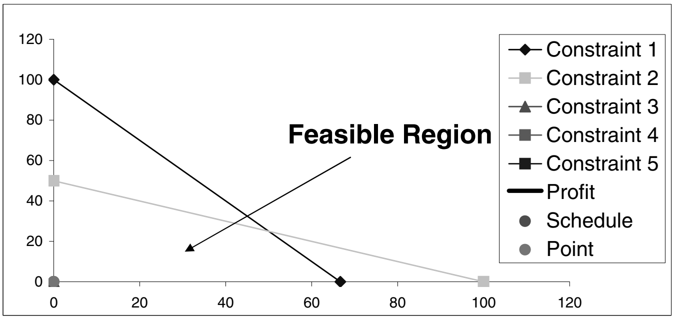


FIGURE 7.22

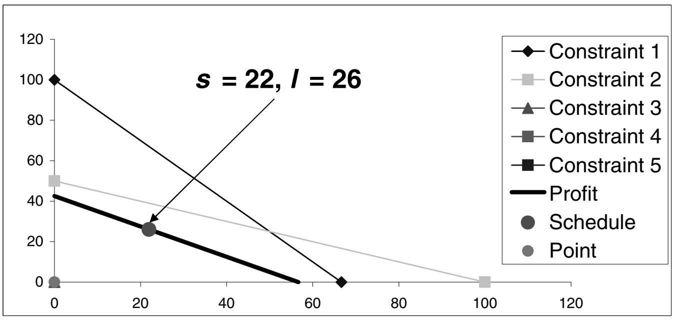


FIGURE 7.23

feasible region. This means that it is *not* possible to profit more than \$250.00. This is our maximum profit. Move the slider until the desired profit is exactly \$250.00 and note that the level curve intersects the feasible region at only one point. Using the slider under **Desired Schedule** we see that the coordinates of this point and $s = 50$ and $l = 25$ as shown in Figure 7.24. This is our optimal solution, which is exactly the same as that found by Solver.

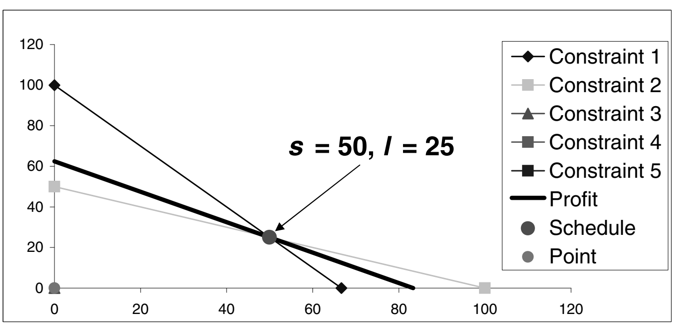


FIGURE 7.24

□

The feasible region in this, and all linear programs, forms what is called a *convex set* which means that if any two points in the set are joined by a line segment, the segment lies entirely within the set (i.e. it never "leaves" the set). A *corner-point* of the feasible region is a point of intersection of two or more constraint lines. Note that the optimal solution to this problem occurs at a corner point. For this reason, the solution is often called a *corner–point solution*. The feasible, but non-optimal, solution $s = 22.0$ and $l = 26.0$ is called an *interior–point solution*.

These observations are generalized in Theorem 7.5.1.

Theorem 7.5.1. *If the convex feasible region of a linear program is nonempty and bounded, then the maximum and minimum values of the objective function will occur at corner-points of the region. If the feasible region is unbounded, then the objective function may not attain maximum or minimum values.* □

Theorem 7.5.1 is important because it tells us that to find the optimal solution to a linear program, we only need to consider the corner-points of the feasible region. This observation forms the basis of the algebraic solution and the Simplex method.

Example 7.5.2 (Algebraic Solution)
The fact that the optimal solution to a linear programming problem occurs at a corner-point of the feasible region suggests that we simply need to find all the corner–points and pick the best one. To illustrate this idea, again consider the linear program (7.4):

$$\begin{array}{rl} \textbf{Maximize} & P = 3s + 4l \\ \textbf{Subject to} & \quad 3s + 2l \leq 200 \\ & \quad 1s + 2l \leq 100 \\ & \qquad s,\, l \geq 0 \end{array}$$

To find the corner–points, we will first find all the intersection points of the constraint lines. To do this, we will translate the first two inequality constraints into equality constraints by introducing the "slack" variables y_1 and y_2:

$$\begin{array}{rl} \textbf{Maximize} & P = 3s + 4l \\ \textbf{Subject to} & \quad 3s + 2l + y_1 \qquad = 200 \\ & \quad 1s + 2l \qquad + y_2 = 100 \\ & \qquad s,\, l,\, y_1,\, y_2 \geq 0 \end{array} \tag{7.5}$$

The variable y_1 is called a slack variable because it takes up the difference (i.e. the slack) between the quantity $3s + 2l$ and the number 200. Since $3s + 2l$ must be less than or equal to 200, y_1 must be non–negative. A similar explanation applies to y_2.

Now to calculate the points of intersection, consider the graph of the constraints shown in Figure 7.22. For any point on the constraint 1 line,

$$3s + 2l = 200$$

so $y_1 = 0$. For any point on the constraint 2 line,

$$1s + 2l = 100$$

so $y_2 = 0$. Therefore, any point at the intersection of these two constraint lines will be characterized by $y_1 = y_2 = 0$. So to find this point of intersection, we could set $y_1 = y_2 = 0$ in (7.5) and solve the resulting system of linear equations

$$\begin{aligned} 3s + 2l &= 200 \\ 1s + 2l &= 100 \end{aligned}$$

for s and l yielding $s = 50$, $l = 25$. This is one corner point.

TABLE 7.11

Point	s	l	y_1	y_2	Feasible?	Profit
1	0	0	200	100	Yes	0
2	0	100	0	−100	No	–
3	0	50	100	0	Yes	200
4	200/3	0	0	100/3	Yes	200
5	100	0	−100	0	No	–
6	50	25	0	0	Yes	250

The s axis is the constraint line for the constraint $l \geq 0$. Obviously for any point on this line, $l = 0$. Therefore, any point at the intersection of constraint 1 and the constraint $l \geq 0$ is characterized by $y_1 = l = 0$. So to find the point of intersection of these two constraints, we could set $y_1 = l = 0$ in (7.5) and solve the resulting system

$$\begin{aligned} 3s \qquad &= 200 \\ 1s + y_2 &= 100 \end{aligned}$$

for s and y_2. This yields $s = 200/3$, $y_2 = 100/3$. So another corner-point is $s = 200/3$, $l = 0$.

In general, to find the points of intersection of the constraint lines we need to set two variables in (7.5) equal to 0 and solve for the remaining variables. The number of ways we could choose which variables to set equal to 0 (i.e. the number of points of intersection) is given by

$$\binom{4}{2} = \frac{4!}{2!\,(4-2)!} = 6$$

□

All of these different combinations are shown in Table 7.11. Note that not all of these are feasible because some don't satisfy the non-negativity constraints in (7.5). We see that we have four feasible corner-point solutions, and the optimal one is $s = 50$, $l = 25$ (i.e. 50 small and 25 large baskets) with a profit of \$250. This is the exact same solution we got graphically and with Solver.

This basic idea seems simple enough, but there is a major problem with computational efficiency. Suppose we have a linear program with m decision variables and n inequality constraints. First we convert the constraints into equalities by introducing n slack variables, similar to above. This gives a total of $m + n$ variables. We find the points of intersection of the constraint lines by setting m variables equal to 0. This gives a system of n equations with n variables, which can be solved, in principle, by matrix techniques.

The number of ways we could choose these m variables to set equal to 0 (i.e. the number of points of intersection) is then given by:

$$\binom{m+n}{m} = \frac{(m+n)!}{m!\,((m+n)-m)!} = \frac{(m+n)!}{m!n!}$$

For a problem with 25 decision variables and 50 inequality constraints (which is relatively small for a real application), the number of points of intersection is

$$\binom{75}{25} = \frac{75!}{25!\,50!} = 5.25 \times 10^{19}$$

This is far too many points to check for any computer. So we want to find an algorithm that does not require finding and checking all points of intersection. One such algorithm is the Simplex method, which is discussed in the next section.

Exercises

7.5.1 Suppose the produce manager starts making 50 small and 25 large baskets but then discovers that two apples and five oranges are rotten. Is it still possible to make 50 small and 25 large baskets total? Why or why not? How many should he make now to maximize his profit?

7.5.2 Suppose the produce manager adds the constraint that he can't make more than 40 small baskets (assume he still has 200 apples and 100 oranges).

a. Graphically estimate the new optimal solution.

b. Suppose he limits himself to 30 small baskets. What is the new optimal solution?

c. As the number of allowed small baskets decreases, what happens to the optimal solution graphically?

d. Suppose we add the generic constraint $s \leq s_0$ where $0 \leq s_0 \leq 50$. Find a formula for the optimal solution (remember, the s and l must be integers).

7.5.3 Consider the linear program:

$$\begin{array}{rrl}
\textbf{Minimize} & C = & x + 2y \\
\textbf{Subject to} & & x + \ \ y \geq 6 \\
& & 3x + \ \ y \geq 9 \\
& & x,\, y \ \geq 0.
\end{array}$$

a. Graphically solve this program (note that the feasible region is *above* the constraint lines).

b. Find a value of the coefficient of x in the objective function so that the optimal solution is $x = 1.5$, $y = 4.5$ with $C = 15.00$.

7.5.4 Solve the linear program

$$\begin{array}{rrl}
\textbf{Maximize} & P = & 20x + 32y \\
\textbf{Subject to} & & 6x + \ \ y \leq 6 \\
& & 3x + 2y \leq 9 \\
& & x,\, y \ \geq 0
\end{array}$$

algebraically by enumerating all the possible corner-points, determining which are feasible, and choosing the best one as done in Table 7.11. Verify your solution using Solver.

7.6 The Simplex Method

Consider the linear program (7.4) (rewritten with decision variables x_1 and x_2 and objective function value z):

$$\begin{array}{rrl}
\textbf{Maximize} & z = & 3x_1 + 4x_2 \\
\textbf{Subject to} & & 3x_1 + 2x_2 \leq 200 \\
& & x_1 + 2x_2 \leq 100 \\
& & x_1,\, x_2 \ \geq 0
\end{array}$$

We solved this program with Solver, graphically, and algebraically and got an optimal solution of $x_1 = 50$, $x_2 = 25$ with a profit of 250.

In this section we illustrate some of the basic ideas behind the Simplex method by using it to solve this problem. Note that the version of the Simplex method we discuss here applies only to maximization linear programs with inequality constraints of the form $\leq$. To apply this version to a problem of another type would require reformatting the problem into this basic form.

Graphical Interpretation

Consider the graph of the feasible region of this problem in Figure 7.22. We know that the optimal solution will lie at one of the corner points of this feasible region. Graphically the Simplex method works by moving from one corner point to an adjacent corner point on the border of the feasible region (the border is called a **Simplex**).

1. Start at a corner point (typically we start at (0, 0)).

2. Move either up or to the right. This corresponds to increasing the value of x_1 or x_2.

3. To determine which direction to move, we look at the objective function $z = 3x_1 + 4x_2$. We want to maximize its value, so increasing x_2 will increase its value faster than increasing x_1. So move up.

4. To determine how far up we can go we look at the constraints

$$\begin{aligned} 3x_1 + 2x_2 &\leq 200 \\ 1x_1 + 2x_2 &\leq 100 \end{aligned}$$

   Since the decision variables are nonnegative, these constraints tell us that to remain feasible, at the very least we must have

$$\begin{gathered} 2x_2 \leq 200 \quad \Rightarrow \quad x_2 \leq 100 \\ \text{and} \\ 2x_2 \leq 100 \quad \Rightarrow \quad x_2 \leq 50. \end{gathered}$$

   The inequality $x_2 \leq 50$ is more restrictive than $x_2 \leq 100$, so we increase x_2 to 50. Graphically, this moves us to the corner point (0, 50).

These four steps move us from the first corner point to the next one. The basic idea behind the next move is the same, but performing the calculations requires us to rewrite the program. This is done most easily with matrices.

Tableau Form

To rewrite the program using matrices, we convert it into "tableau" form by adding slack variables, converting the constraints to equalities, and rewriting the objective function and adding it to the constraints:

$$\begin{array}{ll} \textbf{Maximize} & z \\ \textbf{Subject to} & 3x_1 + 2x_2 + y_1 \qquad\qquad = 200 \\ & x_1 + 2x_2 \qquad + y_2 \qquad = 100 \\ & -3x_1 - 4x_2 \qquad\qquad + z = 0 \\ & \qquad\qquad x_1,\, x_2,\, y_1,\, y_2 \geq 0 \end{array}$$

Now we can use this form of the problem to implement the Simplex Method. Use the worksheet "**2 Variable Simplex Method**" in the workbook **Linear Programming** to perform these steps:

Step 1: Form an initial "**Tableau**" as in Figure 7.25 (RHS stands for Right Hand Side). This tableau is really nothing more than a matrix of the coefficients in the tableau form of the program.

Tableau 0

Basic	x_1	x_2	y_1	y_2	z	RHS	Ratio
	3	2	1	0	0	200	
	1	2	0	1	0	100	
	-3	-4	0	0	1	0	

FIGURE 7.25

In the algebraic solution to this problem, we set two variables equal to 0 and solved for the others. Here, we take the same approach. The variables we set equal to 0 and those we solve for are given special names:

- **Non–basic variables** are those set equal to 0.
- **Basic variables** are those not necessarily equal to 0 (the ones we solve for).

Likewise, in the graphical interpretation of the Simplex method, we decided to increase x_2 first because it had the largest coefficient in the objective function. This variable and its corresponding column in the tableau are given special names:

- The **pivot column** is the column with the negative coefficient with the largest magnitude in the bottom row of the tableau.
- The **entering basic variable** is the variable corresponding to the pivot column.

In this case the pivot column is column 2 (corresponding to x_2).

Step 2: Identify the basic variable for each row of the tableau by entering it in the left-hand column of the tableau.

Identifying the basic and non-basic variables is relatively easy:

- In Tableau 0 (and only Tableau 0) the non–basic variables are the decision variables. The basic variables are the slack variables and z. The basic variable of each row is the basic variable with a coefficient of 1.
- When going from one tableau to the next, the basic variables remain the same with one exception. One of the basic variables, called the leaving basic variable (which is formally defined below) is replaced by the entering basic variable.

In the graphical interpretation of the simplex method, we decided how much we could increase x_2 by dividing the RHS of each constraint by the corresponding coefficient of x_2. We perform a similar step here.

Step 3: Compute the ratio of the RHS to the coefficient in the pivot column for the top two rows of the tableau. This can be done by entering the formulas in Figure 7.26.

Ratio
=I5/E5
=I6/E6

FIGURE 7.26

In the graphical interpretation of the simplex method, the smallest ratio that we just computed told us how much we could increase x_2. The row of the corresponding constraint is given a special name:

- The **pivot row** is the row with the smallest positive ratio of RHS to coefficient in the pivot column.
- The **leaving basic variable** is the variable corresponding to the pivot row.
- The **pivot** is the entry at the intersection of the pivot column and the pivot row.

In the next tableau, the leaving basic variable will be non–basic.

Step 4: Identify the Pivot Column, Pivot Row, and the Leaving Basic Variable by entering them to the right of the tableau.

At this point, Tableau 0 should look like Figure 7.27.

Tableau 0

Basic	x_1	x_2	y_1	y_2	z	RHS	Ratio
y1	3	2	1	0	0	200	100
y2	1	2	0	1	0	100	50
z	-3	-4	0	0	1	0	

Pivot Column =	2
Pivot Row =	2
Leaving Basic Variable =	y2

FIGURE 7.27

Step 5: Do elementary row operations to the matrix of coefficients so the pivot is 1 and all other entries in the pivot column are 0. This is called "clearing the pivot column."

Step 5 can be done by entering the formulas from Figure 7.28 into Tableau 1.

	D	E	F	G	H	I
11	x_1	x_2	y_1	y_2	z	RHS
13	=D5-2*D14	=E5-2*E14	=F5-2*F14	=G5-2*G14	=H5-2*H14	=I5-2*I14
14	=D6/2	=E6/2	=F6/2	=G6/2	=H6/2	=I6/2
15	=D7+4*D14	=E7+4*E14	=F7+4*F14	=G7+4*G14	=H7+4*H14	=I7+4*I14

FIGURE 7.28

Step 6: Repeat Steps 2 – 5 until there are no more negative entries in the bottom row.

Tableau 1

Basic	x_1	x_2	y_1	y_2	z	RHS	Ratio
y1	2	0	1	-1	0	100	50
x2	0.5	1	0	0.5	0	50	100
z	-1	0	0	2	1	200	

Pivot Column =	1
Pivot Row =	1
Leaving Basic Variable =	y1

FIGURE 7.29

Tableau 2

Basic	x_1	x_2	y_1	y_2	z	RHS	Ratio
x1	1	0	0.5	-0.5	0	50	
x2	0	1	-0.25	0.75	0	25	
z	0	0	0.5	1.5	1	250	

FIGURE 7.30

After repeating Steps 2 – 4, Tableau 1 should look like Figure 7.29. Note that in Tableau 0 the leaving basic variable is y_2. This means that in Tableau 1, y_2 is non-basic. The new basic variable is x_2, the variable corresponding to the pivot column.

Applying Step 5 to Tableau 1 results in Tableau 2 shown in Figure 7.30.

In Tableau 2 we see that there is no negative coefficient in the bottom row, so there is no pivot column. Therefore we are done. The non–basic variables (the ones set equal to 0) are y_1 and y_2, so the tableau gives the optimal solution shown in Figure 7.31. This is the exact same solution found with other methods. The numbers 0.5 and 1.5 in the bottom row of Tableau 2 have special meanings, as we will see in the next section.

x_1	x_2	y_1	y_2	z
50	25	0	0	250

FIGURE 7.31

Exercises

Directions: Solve each linear program below using the Simplex method in the worksheet **2 Variable Simplex Method** or **3 Variable Simplex Method**. Verify your solution with Solver.

7.6.1

$$\begin{array}{rl} \textbf{Maximize} & z = 14x_1 + 16x_2 \\ \textbf{Subject to} & x_1 + \quad x_2 \le 100 \\ & 20x_1 + 30x_2 \le 2400 \\ & x_1,\, x_2 \ge 0 \end{array}$$

7.6.2

$$\begin{array}{rl} \textbf{Maximize} & z = 1.25x_1 + 0.95x_2 \\ \textbf{Subject to} & 5x_1 + \quad 2x_2 \leq 60 \\ & 15x_1 + \quad 18x_2 \leq 360 \\ & x_1,\, x_2 \geq 0 \end{array}$$

7.6.3

$$\begin{array}{rl} \textbf{Maximize} & z = 4x_1 + 3x_2 + 6x_3 \\ \textbf{Subject to} & 3x_1 + \ x_2 + 3x_3 \leq 30 \\ & 2x_1 + 2x_2 + 3x_3 \leq 40 \\ & x_1,\, x_2,\, x_3 \geq 0 \end{array}$$

7.6.4

$$\begin{array}{rl} \textbf{Maximize} & z = x_1 + 2x_2 + 4x_3 \\ \textbf{Subject to} & 3x_1 + x_2 + 5x_3 \leq 10 \\ & x_1 + 4x_2 + x_3 \leq 8 \\ & 2x_1 + 2x_3 \leq 7 \\ & x_1,\, x_2,\, x_3 \geq 0 \end{array}$$

7.7 Sensitivity Analysis

Again, consider the fruit basket problem modeled by

$$\begin{array}{rl} \textbf{Maximize} & P = 3x + 4y \\ \textbf{Subject to} & 3x + 2y \leq 200 \\ & x + 2y \leq 100 \\ & x,\, x \geq 0 \end{array}$$

where $x =$ the number of small and $y =$ the number of large baskets to produce. This has an optimal solution of $x = 50$ and $y = 25$ with a maximum profit of \$250.

In doing sensitivity analysis we will analyze two questions:

1. How much can a unit profit change (i.e. a coefficient in the objective function) and the optimal solution still remain optimal?

2. How much will increasing the amount of a resource increase the maximum profit?

These two questions are important economically. If the profit for an item suddenly changes, management will want to adjust the production schedule to ensure they are still maximizing profit. They need to know when this is appropriate. Also, management may want to invest money in increasing some resource (e.g. hire more employees, modernize machinery, etc.). They need to know if this will indeed increase profit, and if so, how much. Another reason these questions are important is that the coefficients in the problem are often *estimates*. Answering these questions can help us understand how important the accuracy of the estimates are.

Changing Unit Profits

We will begin by graphically analyzing what happens to the optimal solution when a coefficient in the objective function changes. Enter the program in the worksheet **Graphical Solution**. Change the unit profit for a small basket (the coefficient of x in the objective function) to values above and below 3 and keep the profit for a large basket fixed at 4. For each one, find the optimal solution.

Figure 7.32 shows three such examples (where the dashed lines are the constraint lines and the solid lines are the level curves corresponding to the three different objective functions). Note that when the unit profit is 2.75 or 4, the optimal solution remains at (50, 25) (the maximum value of P is different in each case, though). When the unit profit is 9, the optimal solution changes to (66.67, 0). In general we see that for small changes in the unit profit, the solution remains the same. If the change is large enough, the solution changes. We make the same observations if we change the unit profit for a large basket while keeping the profit for a small basket fixed at 3.

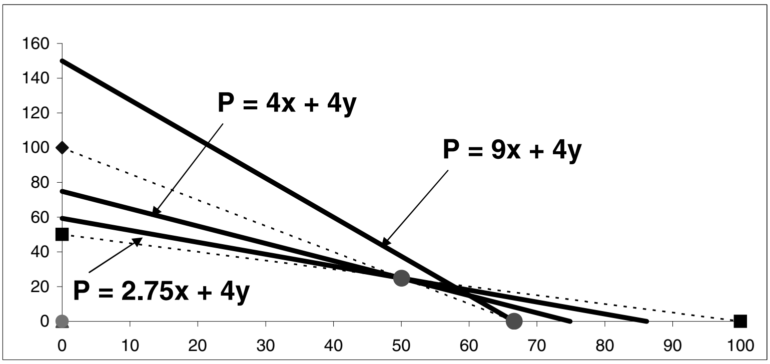


FIGURE 7.32

Graphically we observe that as long as the slope of the profit line is between the slopes of the constraint lines, the optimal solution will remain the same. To quantify this, we calculate the slopes of the constraint lines:

$$\begin{array}{llcl} \textbf{Constraint 1:} & 3x + 2y = 200 & \Rightarrow & y = -\frac{3}{2}x + 100 \\ \textbf{Constraint 2:} & x + 2y = 100 & \Rightarrow & y = -\frac{1}{2}x + \ \ 50 \end{array}$$

Therefore, as long as the slope of the profit line is between $-\frac{3}{2}$ and $-\frac{1}{2}$, the optimal solution will remain the same.

Now suppose the unit profit for small baskets is C. Then the profit equation is

$$P = Cx + 4y \;\Rightarrow\; y = -\frac{C}{4}x + \frac{P}{4}$$

Thus we see that the slope of the profit line is determined by only the value of C. The value of the profit P does not affect the slope, only the y-intercept. Therefore, if

$$-\frac{3}{2} \leq -\frac{C}{4} \leq -\frac{1}{2} \;\Rightarrow\; 2 \leq C \leq 6$$

then the optimal solution will remain at (50, 25). However, note that this is valid only if the unit profit for large baskets remains fixed at 4.

Now let K represent the unit profit for large baskets. The profit equation is

$$P = 3x + Ky \;\Rightarrow\; y = -\frac{3}{K}x + \frac{P}{K}$$

Again we see that the slope of the profit equation is determined by only the value of K. Therefore, if

$$-\frac{3}{2} \leq -\frac{3}{K} \leq -\frac{1}{2} \;\Rightarrow\; 2 \leq K \leq 6$$

then the optimal solution will remain at (50, 25). Similar to before, this is valid only if the unit profit for small baskets remains fixed at 3.

Thus the answer to the first question is this: The unit profit for a small basket can decrease as much as \$1 or increase as much as \$3. The unit profit for a large basket can change as much as \$2. In either case the solution (50, 25) will remain optimal. This is valid only if one unit profit changes. If they both change, then we must solve the problem again (or do more analysis).

Increasing Resources

We have already found the answer to the second question. It is given to us in the last row of the final tableau from the Simplex method shown in Figure 7.33.

Tableau 2

Basic	x_1	x_2	y_1	y_2	z	RHS	Ratio
x1	1	0	0.5	-0.5	0	50	
x2	0	1	-0.25	0.75	0	25	
z	0	0	0.5	1.5	1	250	

FIGURE 7.33

The numbers 0.5 and 1.5, the coefficients of the slack variables in the bottom row of the final tableau, are called the *shadow prices* of the apple and orange resources, respectively. These numbers tell us that increasing the number of apples by one unit will increase the maximum profit by \$0.5 and increasing the number of oranges by one unit will increase the maximum profit by \$1.50.

To understand why these shadow prices have these meanings, observe that the bottom row of tableau 2 is related to tableau 0 by the formula

$$\begin{aligned}(\text{Bottom row of tableau 2}) &= 0.5 \times (\text{row 1 of tableau 0})\\ &+ 1.5 \times (\text{row 2 of tableau 0})\\ &+ (\text{row 3 of tableau 0})\end{aligned}$$

Therefore, if the RHS of row 1 of tableau 0 is increased by 1 (i.e. the number of apples is increased by 1), then the RHS of the bottom row of tableau 2 is increased by 0.5 (i.e. the maximum profit is increased by 0.5). Likewise, if the RHS of row 2 of tableau 0 is increased by 1, then the RHS of the bottom row of tableau 2 is increased by 1.5.

In general this means that if the number of apples is changed by n units (positive or negative), then

$$\text{Change in maximum profit} = 0.5n \tag{7.6}$$

Note that this conclusion holds *only* if we change *only* the number of apples.

However, this conclusion is valid only if the number of apples changes within limits. To illustrate this, graph the feasible region of the program (with the original unit profits) in the worksheet **Graphical Solutions**. Change the number of apples (the RHS of constraint 1) to values above and below 200. Each time, find the optimal solution. Three such examples are given in Figure 7.34 where the solid line is the constraint 2 line, the dashed lines are the constraint 1 lines for three different numbers of apple, and the solid dots are the optimal solutions for the three cases.

Notice that for 230 apples, the optimal solution still lies at the intersection of the constraint 1 and constraint 2 lines (it is different than the original solution, but it still occurs at

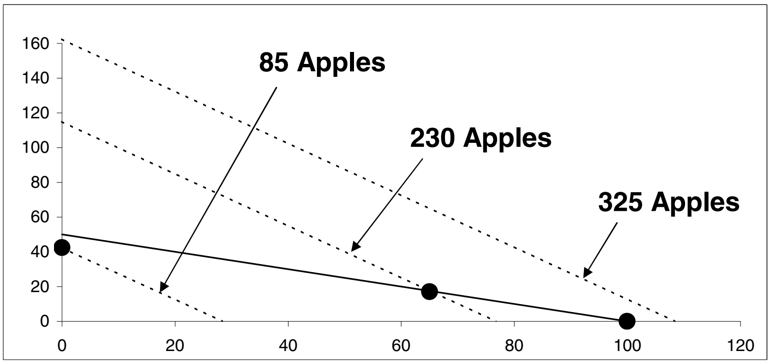


FIGURE 7.34

this intersection). For 325 apples, the number of apples is not a constraint and the optimal solution occurs at the x-intercept of the constraint 2 line. For 85 apples, the number of oranges is not a constraint and the optimal solution occurs at the y-intercept of the constraint 1 line. Thus we see that if the number of apples changes enough, the nature of our solution changes. This means that Equation (7.6) holds over only a limited domain.

Note graphically that for 325 apples the x- and y-intercepts of the constraint 1 line are both greater than the corresponding intercepts of the constraint 2 line. For 85 apples, the opposite is true. Based on this observation we can conclude that the optimal solution will occur at the intersection of the constraint 1 and constraint 2 lines only if this relation among the intercepts does not hold.

To calculate the domain over which Equation (7.6) holds, note that the constraint 2 line is given by $x + 2y = 100$, which has x- and y-intercepts of 100 and 50, respectively. Now suppose we have A apples available. Then the constraint 1 line is given by $3x + 2y = A$ which has x- and y-intercepts of $\frac{A}{3}$ and $\frac{A}{2}$, respectively.

The intercepts of the constraint lines are equal when:

$$\begin{array}{rcl} x\text{-intercepts: } 100 = \frac{A}{3} & \Rightarrow & A = 300 \\ y\text{-intercepts: } 50 = \frac{A}{2} & \Rightarrow & A = 100 \end{array}$$

Thus we conclude that Equation (7.6) holds when the number of apples is between 100 and 300 (or n is between -100 and $+100$). The economic interpretation of this is:

> If each apple costs less than \$0.50, the manager can increase the number of apples by up to 100 to increase the maximum profit. On the other hand, if each apple can be sold for *more* than \$0.50, up to 100 apples should be sold. This will decrease the profit from fruit baskets, but increase the total profit for the produce department.

Now let's fix the number of apples at 200 and change the number of oranges. The constraint 1 line is given by $3x + 2y = 200$ which has x- and y-intercepts of $\frac{200}{3}$ and 100, respectively. Suppose we have B oranges available. Then the constraint 2 line is given by $x + 2y = B$ which has x- and y-intercepts of B and $\frac{B}{2}$, respectively.

The intercepts of the constraint lines are equal when:

$$\begin{array}{rcl} x\text{-intercepts: } \frac{200}{3} = B & \Rightarrow & B \approx 66.7 \\ y\text{-intercepts: } 100 = \frac{B}{2} & \Rightarrow & B = 200 \end{array}$$

This means that up to 100 oranges may be added and the maximum profit will increase by \$1.50 (the shadow price for oranges) for each one. Up to 33 oranges may be removed and it will decrease by \$1.50 for each one.

Solver will generate a sensitivity analysis report after it solves a model. To examine this report for this program, return to the worksheet **Fruit Baskets** created in Section 7.2 (see Figure 7.1). Set the appropriate Solver parameters (be sure to select **Simplex LP** for the **Solving Method**) and press **Solve**. In the next window that appears, select **Sensitivity** under **Reports** and press **OK**. You should get the new worksheet containing the results shown in Figure 7.35.

	A	B	C	D	E	F	G	H
6	Adjustable Cells							
7				**Final**	**Reduced**	**Objective**	**Allowable**	**Allowable**
8		**Cell**	**Name**	**Value**	**Cost**	**Coefficient**	**Increase**	**Decrease**
9		\$B\$2	Number Small	50	0	3	3	1
10		\$C\$2	Number Large	25	0	4	2	2
11								
12	Constraints							
13				**Final**	**Shadow**	**Constraint**	**Allowable**	**Allowable**
14		**Cell**	**Name**	**Value**	**Price**	**R.H. Side**	**Increase**	**Decrease**
15		\$D\$3	Apples Amt Used	200	0.5	200	100	100
16		\$D\$4	Oranges Amt Used	100	1.5	100	100	33.33333333

FIGURE 7.35

Cell **F9** shows the unit profit for small baskets. Cells **G9** and **H9** tell us that this unit profit can increase by as much as 3 or decrease by as much as 1 (i.e. vary between 2 and 6) and the optimal solution will remain the same. Row 10 gives similar results for large baskets.

Cell **E15** gives the shadow price for apples. Cell **F15** shows the number of available apples. Cells **G15** and **H15** tell us that the interpretation of the shadow price is valid if the number of apples is increased by 100 or decreased by 100 (i.e. is between 100 and 300). Row 16 gives similar results for the oranges.

These are the exact same conclusions we reached earlier.

Exercises

7.7.1 Suppose the manager can buy a box of 50 oranges or a box of 50 apples for making fruit baskets for \$20 each, but he can purchase only one. Which one should he purchase? Why?

7.7.2 Suppose there are 350 apples and 100 oranges available for making fruit baskets. Generate a sensitivity report and explain why the shadow price for apples is 0, the allowable increase is ∞, and the allowable decrease is only 50.

7.7.3 A toy company manufactures cars and trucks. Each car uses 1 unit of plastic and 20 units of metal and yields a profit of \$14. Each truck uses 1 unit of plastic and 30 units of metal and yields a profit of \$16. The company has 100 units of plastic and 2400 units of metal available. To determine the number of each toy to manufacture to maximize profit, we need to solve the program

$$\begin{array}{rl} \textbf{Maximize} & P = 14x_1 + 16x_2 \\ \textbf{Subject to} & x_1 + \quad x_2 \leq 100 \\ & 20x_1 + 30x_2 \leq 2400 \\ & x_1,\, x_2 \geq 0 \end{array}$$

where x_1 = the number of cars and x_2 = the number of trucks to produce. This is the same program we solved in Exercise 7.6.2.

a. Use the final tableau in your solution to Exercise 7.6.2 to find the shadow prices of the cars and the trucks.

b. By hand, find the range over which the interpretation of the shadow price is valid for each resource.

c. By hand, find ranges for the unit profits over which the solution remains optimal.

d. Verify your calculations with Solver.

For Further Reading

- For a classic reference on everything related to operations research, see Hillier, F. and G. Lieberman, *Introduction to Operations Research*, Seventh Edition, McGraw Hill, 2001.
- For more information on the Simplex method, sensitivity analysis, nonlinear programming, and other topics from this chapter, see Winston, Wayne L. and Munirpallam Venkataramann, *Introduction to Mathematical Programming, Operations Research: Volume One*, Fourth Edition, Thomson Brooks/Cole, 2003.

8

Nonlinear Optimization

Chapter Objectives

- Use Newton's method to solve optimization problems
- Use the golden search method
- Introduce the one-dimensional gradient method
- Introduce the two-dimensional gradient method
- Use Lagrange multipliers
- Introduce branch and bound techniques
- Introduce the traveling salesman problem

8.1 Introduction

A *nonlinear program* is any program that does not fit the definition of a linear program. A linear program has a very special structure that can be utilized to construct an efficient solution method (i.e. the Simplex method). The drawback of this approach is that if a problem does not fit this special structure, the solution method cannot be applied.

As we will see, nonlinear programs are inherently more difficult to solve than linear programs. In fact, there is no efficient algorithm that guarantees an optimal solution to a general nonlinear program, like the Simplex method guarantees an optimal solution to a general linear program. However, certain types of nonlinear programs such as binary integer, convex, separable, and quadratic, have special structures that can be utilized to find efficient solution methods. In this chapter we present several numeric techniques for approximating the optimal solution to nonlinear programs.

We begin with a classic Calculus I application of a nonlinear program.

Example 8.1.1 (Calculus I Problem)
Engineers need to connect a wind turbine to a collector step-up transformer via buried cable. The turbine is located 2 km from a straight road and the transformer is located 6 km down the road, as illustrated in Figure 8.1. The soil between the turbine and the road is rocky and burying a cable costs \$2000/km. The soil along the road is easier to work with and burying a cable only costs \$1500/km. Find the path of the cable that minimizes the total cost.

The shortest path is a straight line from the turbine to the transformer, but this path is completely through the rocky soil so it would probably be expensive. Another option is to go straight down to the road and then along the entire length of the road. This would minimize the distance across rocky soil, but it would be the longest path. So this option

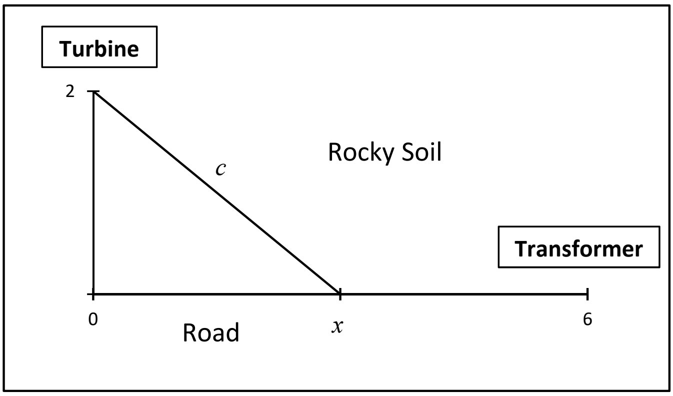


FIGURE 8.1

would probably be expensive as well. A third option is to run the cable to some point x along the road and then along the road, as illustrated in Figure 8.1. The question is, what's the value of x?

To answer this question, we need a function describing the total cost in terms of x. Using the Pythagorean theorem, the distance c from the turbine to the point x on the road is $\sqrt{2^2 + x^2}$, so the cost of this portion of the cable is $2000\sqrt{4 + x^2}$. The distance of the cable along the road is $6 - x$, so this cost is $1500(6 - x)$. Therefore, the total cost is $C = 2000\sqrt{4 + x^2} + 1500(6 - x)$.

We can write the program in the form

$$\begin{aligned} &\textbf{Minimize } C(x) = 2000\sqrt{4 + x^2} + 1500(6 - x) \\ &\textbf{Subject to } 0 \leq x \leq 6. \end{aligned}$$

To solve this program using Calculus I, we take the derivative of C with respect to x,

$$\frac{dC}{dx} = \frac{2000x}{\sqrt{4 + x^2}} - 1500,$$

set this derivative equal to 0, and solve,

$$\frac{2000x}{\sqrt{4 + x^2}} - 1500 \overset{\text{SET}}{=} 0 \quad \Rightarrow \quad x = \sqrt{\frac{36}{7}} \approx 2.267.$$

Values of x where the derivative is equal to 0 are called *critical points*. A critical point is a possible location of the optimal solution. There are two commonly used tests for determining whether a critical point is the location of the minimum or maximum value of the function, the *candidates test* and the *second derivative test*.

Candidates Test: This test says that all the possible locations of the minimum and maximum values of a continuous differentiable function on an interval are the critical points of the function and the endpoints of the interval (these points are called the *candidates*). We simply need to evaluate the function at each candidate, and then choose the best one.

For the function $C(x)$ in this example, the candidates are the critical point $x = 2.267$ and the end points of the constraint interval, $x = 0$ and $x = 6$. Then we evaluate $C(x)$ at these candidates:

$$C(0) = 2000\sqrt{4+0^2} + 1500(6-0) = 13000$$
$$C(2.267) = 2000\sqrt{4+2.267^2} + 1500(6-2.267) = 11645.75$$
$$C(6) = 2000\sqrt{4+6^2} + 1500(6-6) = 12649.11.$$

The smallest of these three costs is \$11,645.75, so we conclude that the optimal solution is $x = 2.267$. Therefore, the engineers should run the cable to a point 2.267 km down the road and then run the cable the remaining $6 - 2.267 = 3.733$ km along the road. The total cost is \$11,645.75.

Second Derivative Test: This test is a way of identifying whether a given critical point is the location of a minimum or maximum by evaluating the second derivative of the function at the critical point. If the second derivative is positive at the critical point, then the point is a location of a minimum. If negative, then the point is a location of a maximum.

For the function $C(x)$ in this example, the second derivative is

$$C''(x) = \frac{8000}{(4+x^2)^{3/2}},$$

and the second derivative evaluated at the critical point is

$$C''(2.267) = \frac{8000}{(4+2.267^2)^{3/2}} > 0.$$

Since the second derivative is positive, we conclude that the critical point $x = 2.267$ is the location of a minimum cost, the same conclusion reached with the candidates test. □

This example involved finding the derivative and setting it equal to 0 to find critical points. Critical points are possible values of x at which the function has a *local* (or *relative*) *maximum* or *minimum*. A local maximum occurs at x_0 if

$$f(x_0) \geq f(x) \text{ for all } x \text{ in some interval centered at } x_0.$$

A local minimum is defined similarly. On the other hand, a *global maximum* occurs at x_0 if

$$f(x_0) \geq f(x) \text{ for all } x \text{ in the domain of } f.$$

A global minimum is defined similarly. Generically, the local minimum and maximum are called *local extrema* and then global minimum and maximum are called *global extema*.

Graphically, local maximums occur at *high* points on the graph of a function. Global maximums occur at the *highest* point on the graph. Minimums are identified similarly.

Example 8.1.2 (Identifying Local and Global Extrema)
As an example of the definitions of local and global extrema, see the graph of $y = g(x)$ shown in Figure 8.2 whose domain is $0 \leq x \leq 4$. The function $g(x)$ has a local maximum near 0.6, a global maximum near 3.5, local minimums near 2 and at 4, and a global minimum at 0.

□

When solving a nonlinear program, we typically want to find the global extrema. As we'll see in the next section, sometimes it's hard to know if we've found a local or global extrema.

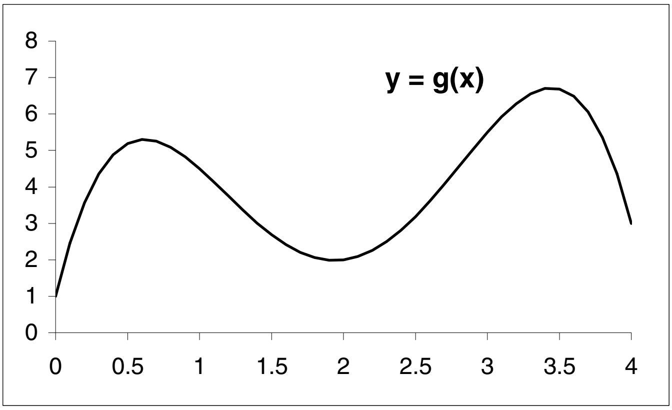


FIGURE 8.2

8.2 Newton's Method

Example 8.1.1 required us to find the derivative of the objective function, set it equal to 0, and then do some algebra to solve the equation. This algebra was arguably the most difficult part of the problem. In this section we present an algorithm for numerically approximating the solution to this equation called *Newton's method*. Like most numeric approximation techniques, it begins with a "guess," or approximation, of the solution, and then improves upon this approximation in several iterations.

Newton's method is an iterative method for approximating a solution to an equation of the form $f(x) = 0$ where $f(x)$ is a differentiable function. Such a solution is called a *root* or *zero* of the function. Graphically, this solution is an x-intercept of the function. The graph of a generic function $f(x)$ with root r is shown in Figure 8.3.

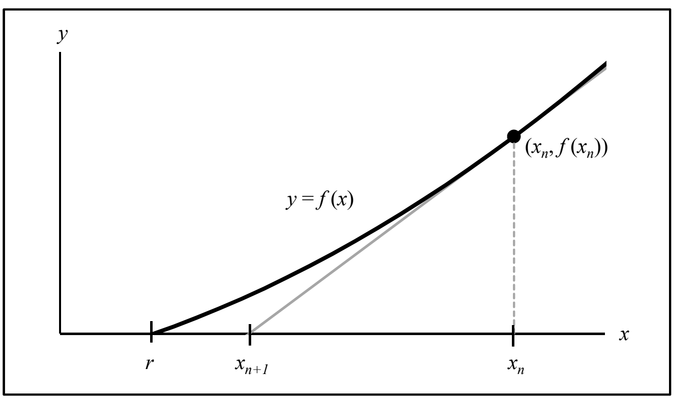


FIGURE 8.3

Figure 8.3 illustrates how Newton's method works. Let x_n be an approximation of the root r. Draw a tangent line at the point $(x_n, f(x_n))$. Let x_{n+1} be the x-intercept of this tangent line. Graphically, we see that x_{n+1} is a better approximation of r than x_n. To find x_{n+1}, consider the points $(x_{n+1}, 0)$ and $(x_n, f(x_n))$. We can calculate the slope between these two points in two ways:

1. They are both on the tangent line at the point $(x_n, f(x_n))$, so the slope is at the point $f'(x_n)$.

2. Using the definition of the slope between two points, the slope is

$$\frac{f(x_n) - 0}{x_n - x_{n+1}} = \frac{f(x_n)}{x_n - x_{n+1}}.$$

Equating these two calculations of the slope and solving for x_{n+1} yields

$$f'(x_n) = \frac{f(x_n)}{x_n - x_{n+1}} \quad \Rightarrow \quad x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

This idea leads to the following simple algorithm.

Newton's Method: To approximate a root r of a differentiable function $f(x)$

1. Let x_0 be an initial approximation of r

2. Calculate
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

3. Repeat step 2 for 10 iterations (10 is arbitrary).

Performing 10 iterations is completely arbitrary. In more sophisticated versions of the algorithm, step 2 is repeated until some "stopping criterion" is met (see Exercises 8.2.1 and 8.2.2).

Example 8.2.1 (Implementing Newton's Method)
Consider Example 8.1.1 where we found the minimum value of the function $C(x) = 2000\sqrt{4+x^2} + 1500(6-x)$ by taking the derivative and then setting it equal to 0. Let $f(x)$ denote the derivative of $C(x)$. That is, let

$$f(x) = C'(x) = \frac{2000x}{\sqrt{4+x^2}} - 1500,$$

and the derivative of $f(x)$ is

$$f'(x) = C''(x) = \frac{8000}{(4+x^2)^{3/2}}.$$

Instead of solving the equation $f(x) = 0$ algebraically, let's use Newton's method to approximate its solution using $x_0 = 3$ as an initial guess. Rename a blank worksheet "**Newton**" and format it as in Figure 8.4. Copy row 3 down to row 12.

	A	B	C	D
1	**n**	**x_n**	**$f(x_n)$**	**$f'(x_n)$**
2	0	3	=2000*B2/SQRT(4+B2^2)-1500	=2000/SQRT(4+B2^2)-2000*B2^2/(4+B2^2)^(3/2)
3	=A2+1	=B2-C2/D2	=2000*B3/SQRT(4+B3^2)-1500	=2000/SQRT(4+B3^2)-2000*B3^2/(4+B3^2)^(3/2)

FIGURE 8.4

The results are shown in Figure 8.5. The algorithm finds that there is a root at $r \approx 2.267787$, which agrees with the results of Example 8.1.1. Note that x_n practically stops changing after $n = 4$. Thus we could have stopped at $n = 4$.

	A	B	C	D
1	**n**	**x_n**	**$f(x_n)$**	**$f'(x_n)$**
12	10	2.267787	0	289.379

FIGURE 8.5

Now change the value of x_0 to various numbers between 0 and 6. Observe that sometimes the algorithm finds the root at $r \approx 2.267787$ and sometimes it doesn't. This illustrates that the algorithm does not work perfectly and that the value of x_0 can affect the results. □

Example 8.2.2 (Maximizing Revenue)
Suppose the revenue generated by a particular machine after x years is given by the function

$$h(x) = 1 - e^{-x} + (1 + x)^{-1}$$

for $0 \leq x \leq 5$. Find the maximum revenue.

Often the first step in an optimization problem is to graph the objective function as shown in Figure 8.6. We see that the revenue is maximized somewhere between $x = 2$ and $x = 3$. This gives us an idea for a starting value in Newton's method. It looks like $x_0 = 2.5$ might be a good choice.

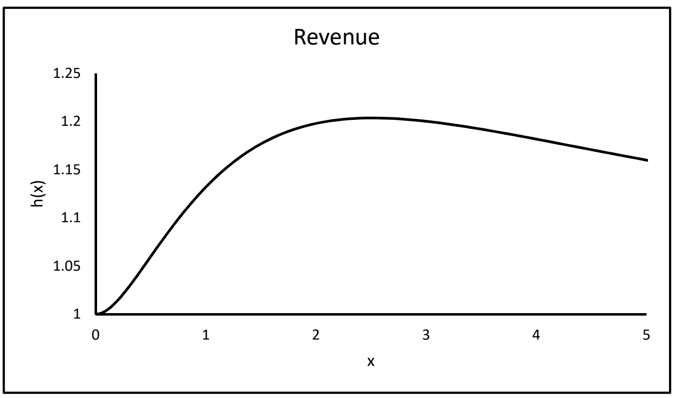


FIGURE 8.6

To use Newton's method to estimate the critical point, we first find $h'(x)$ and rename it $f(x)$:

$$f(x) = h'(x) = e^{-x} - (1 + x)^{-2}.$$

Now we need to find the root of $f(x)$. The derivative of $f(x)$ is

$$f'(x) = -e^{-x} + 2(1 + x)^{-3}.$$

To implement Newton's method to find the roots of $f(x)$, rename a blank worksheet "**Newton 2**" and format it as in Figure 8.7. Copy row 3 down to row 12.

The results are shown in Figure 8.8. They indicate that there is a root of $f(x)$ at $x \approx 2.512862$ and that

$$f'(2.512862) = g''(2.512862) = -0.0349.$$

	A	B	C	D
1	**n**	**x_n**	**$f(x_n)$**	**$f'(x_n)$**
2	0	2.5	=EXP(-B2)-1/(1+B2)^2	=-EXP(-B2)+2/(1+B2)^3
3	=A2+1	=B2-C2/D2	=EXP(-B3)-1/(1+B3)^2	=-EXP(-B3)+2/(1+B3)^3

FIGURE 8.7

Since $g''(2.512862) < 0$, the second derivative test tells us that the critical point 2.512862 is the location of a maximum, as the graph in Figure 8.6 suggested.

	A	B	C	D
1	**n**	**x_n**	**$f(x_n)$**	**$f'(x_n)$**
12	10	2.512862	0	-0.0349

FIGURE 8.8

Therefore, we conclude that the maximum revenue is

$$h(2.51862) = 1 - e^{-2.51862} + (1 + 2.51862)^{-1} \approx 1.203632.$$

In Exercise 8.2.6 we will examine how different values of x_0 affect the results. □

The last example illustrates a different, more academic, type of optimization problem and that we must be very careful about selecting x_0

Example 8.2.3 (Minimizing x)
Find the smallest positive value of x such that $x = \tan x$.

This is a different type of optimization problem than Example 8.1.1. We're not trying to find the optimal value of an objective function. Instead we're trying to find the smallest positive solution to the equation $x = \tan x$. This equation is not easy to solve algebraically, so we'll use Newton's method to approximate the solutions and then choose the smallest such solution.

First, we rewrite the equation,

$$x = \tan x \quad \Rightarrow \quad x - \tan x = 0.$$

If we let $f(x) = x - \tan x$, the problem can be stated as "find the smallest positive solution to $f(x) = 0$," a problem for which we can use Newton's method. A graph of $f(x)$ over the interval $0 \leq x \leq 8$ is shown in Figure 8.9. Note that there are three "pieces" of the graph and that the smallest positive x-intercept appears to be around 4.5.

The derivative of $f(x)$ is

$$f'(x) = 1 - \sec^2 x.$$

To implement Newton's method with an initial guess of $x_0 = 4.5$, rename a blank worksheet "**Newton 3**" and format it as in Figure 8.10.

The results give a root of $x \approx 4.493409$. Now suppose we use $x_0 = 4$. Simply changing x_0 to 4 in the worksheet yields the result $x_n \to \infty$. Why do we get such different results? To help explain why, consider the tangent line at the point $(4,\ f(4))$ as illustrated in Figure 8.11. We see that the tangent line has an x-intercept of about 6.1, meaning $x_1 \approx 6.1$. This drives the search to the third piece of the function, away from where we want.

□

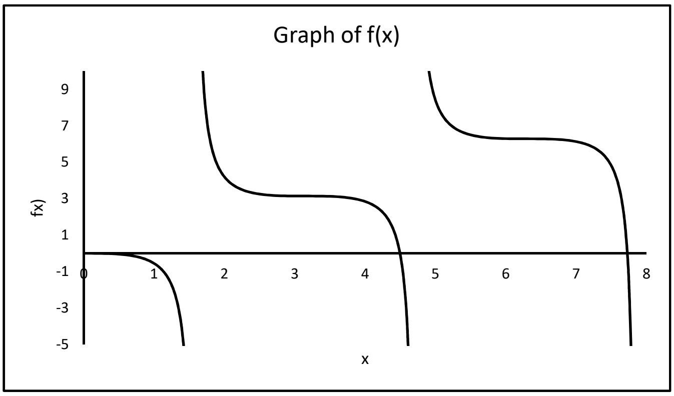


FIGURE 8.9

	A	B	C	D
1	**n**	$\mathbf{x_n}$	$\mathbf{f(x_n)}$	$\mathbf{f'(x_n)}$
2	0	4.5	=B2-TAN(B2)	=1-(SEC(B2))^2
3	=A2+1	=B2-C2/D2	=B3-TAN(B3)	=1-(SEC(B3))^2

FIGURE 8.10

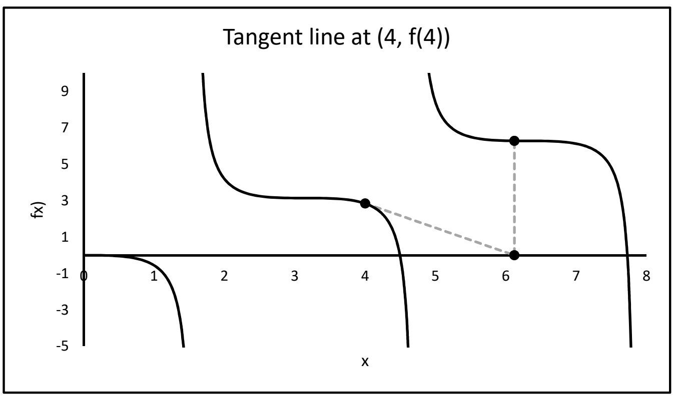


FIGURE 8.11

Exercises

8.2.1 In our simple Newton's method algorithm, we arbitrarily performed 10 iterations. In more sophisticated versions of the algorithm, step 2 is repeated until some "stopping criterion" is met. One very simple stopping criterion is that the algorithm terminates when x_n stops changing significantly. Specifically, the algorithm terminates when $|x_{n+1} - x_n| < \delta$ where $\delta > 0$ is some specified constant. Modify the worksheet **Newton** to incorporate this stopping criterion and output the number of iterations performed along with the approximation of the root. Allow the user to input the value of δ.

8.2.2 Another stopping criterion is based on the following rule-of-thumb: If x_{n+1} and x_n agree in the first m digits, then the approximation x_{n+1} is probably accurate to m digits.

Thus if we want an approximation with 9 decimal places of accuracy, for instance, we simply stop once x_{n+1} and x_n agree to 9 decimal places. As an application of this idea, consider the problem of solving the equation

$$x^2 - 2 = 0.$$

Elementary algebra shows the exact solution is $\sqrt{2}$. Use Newton's method to estimate this solution to 9 decimal places of accuracy using $x_0 = 1$. This result is an estimate of $\sqrt{2}$ to 9 decimal places. How many iterations are necessary?

8.2.3 Squares are cut out of the corners of an 11 in. $\times$ 8.5 in. sheet of paper. The resulting piece of paper is to be formed into an open-topped box. Find the dimensions of the squares so that the resulting box has as large a volume as possible.

8.2.4 Try using Newton's method to approximate the solution to $x^{1/3} = 0$. Try any value of x_0 different from 0. What do you notice? (**Hint**: When entering $x^{2/3}$ into Excel, enter it as $\left(x^2\right)^{1/3}$.)

8.2.5 A 12 m tall tree is cut to a height such that the cube root of the length of the cut off part equals the height of the part left standing. Use Newton's method to approximate the height of the part left standing.

8.2.6 Consider Example 8.2.2

a. Graph the function $f(x) = e^{-x} - (1+x)^{-2}$ over the interval $0 \le x \le 5$.

b. Suppose we start Newton's method with $x_0 = 0.4$. Does this find the root of $f(x)$ at $x \approx 2.512$? Use the graph of $f(x)$ to explain why not. Specifically, use the tangent line at the point $(0.4,\ f(0.4))$ to explain why x_1 is much different than x_0.

c. Now suppose we use $x_0 = 0.6$. Does the search do a better job at locating the root at $x \approx 2.512$? Explain why or why not.

d. Now suppose we use $x_0 = 4.5$. Explain what happens here.

8.2.7 The function $g(x) = \sin 3x - \cos x$ has numerous local minimum and maximum values over the interval $0 \le x \le 10$. The goal of this problem is to estimate the global minimum and maximum values over this interval.

a. Use Newton's method to estimate a critical point of $g(x)$ where x_0 is a random number between 0 and 10.

b. Use a data table to store the results of 100 trials of part a.

c. Calculate $g(x)$ for each result of part b. Find the maximum and minimum values of $g(x)$.

8.2.8 Use Newton's method to find at least two points on the graph of $y = \cos x$ where the tangent line goes through the origin.

8.2.9 Consider the problem of finding the maximum value of the function

$$g(x) = -16x^6 + 48x^5 - 61x^4 + 42x^3 - 16x^2 + 3.5x.$$

a. Use 100 iterations of Newton's method with $x_0 = 0.9$ to estimate the maximum value.

b. Now use $x_0 = 0.94$. Explain why it takes so many more iterations to get close to the same value as found in part a.

8.2.10 A 1-mile (5280 ft) long railroad track was constructed without expansion joints. Suppose the track heats up and the length increases by 1 ft causing the track to "bow up" to form a semicircle as in Figure 8.12 where θ is half the angle of the resulting circular sector and R is the radius of the circle. The goal of this problem, known as the *railroad track problem*, is to find the height x of the bowed rail above the ground.

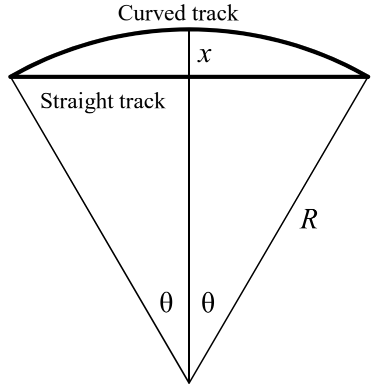


FIGURE 8.12

It can be shown that x satisfies the equation

$$\tan^{-1}\left(\frac{5280x}{2640^2 - x^2}\right) \cdot \left(\frac{2640^2 + x^2}{2x}\right) - \frac{5281}{2} = 0. \tag{8.1}$$

Solving this equation algebraically would be extremely difficult. Solving it with Newton's method would also be difficult because the derivative is complicated. One way to numerically approximate a derivative $f'(x_n)$ is by using secant lines. Specifically, we choose a small quantity d and then calculate

$$f'(x_n) \approx \frac{f(x_n + d) - f(x_n)}{d}.$$

a. Implement this approach in Newton's method to approximate the solution of equation (8.1). Use $d = 0.01$, but make d a parameter the user can change. (**Hint**: Use the formula **ATAN** for $\tan^{-1}$.)

b. **Extra Credit**: Derive equation (8.1).

8.3 The Golden Section Method

Newton's method relies on the derivative of a function. But not every function is differentiable. In this section we present a search method, called the *golden section method*, that does not rely on the derivative. This method begins with an interval of possible values at which the optimal solution occurs, and then narrows down the interval in each step. The method terminates when the interval is of a narrow width that is specified by the user.

The golden section method applies only to functions that are *unimodal* over the initial interval.

Definition 8.3.1. A function $f(x)$ is *unimodal* over an interval if there exists only one number m in the interval such that exactly one of the following is true:

- $f(x)$ is monotonically increasing for $x \leq m$ and monotonically decreasing for $x \geq m$
- $f(x)$ is monotonically decreasing for $x \leq m$ and monotonically increasing for $x \geq m$

Informally, a function is unimodal over an interval if it has only one local extrema over the interval. Graphically, a function is unimodal if it has only one high or low point on the interval. Unimodality is nice because there's only one local extrema to look for. When we find it, we've found the global extrema.

Example 8.3.1 (Identifying Unimodality)
Consider the function $f(x)$ graphed in Figure 8.13. Over the interval [0, 4] the function is not unimodal because it has two low points and one high point. However, $f(x)$ is unimodal over [0, 1] since it has only one low point over this interval. Similarly, the function is unimodal over [1, 3] and [3, 4].

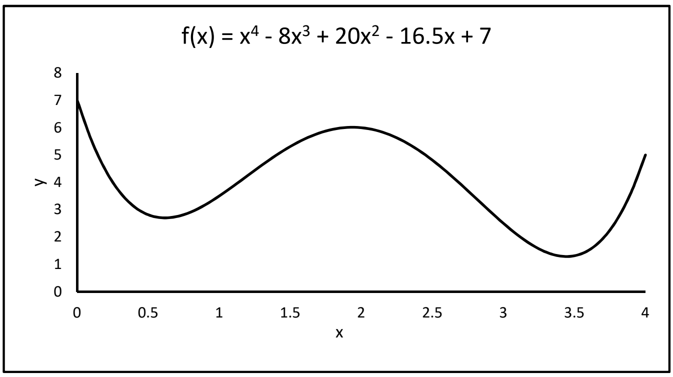


FIGURE 8.13

□

The golden section method relies on a number, called the *golden ratio*, that comes from a famous problem in Greek mathematics: Suppose we divide a line segment of length one into two pieces as illustrated in Figure 8.14 in such a way that the ratio of the longer piece to the shorter piece equals the ratio of the whole segment to the longer piece. Find the value of r.

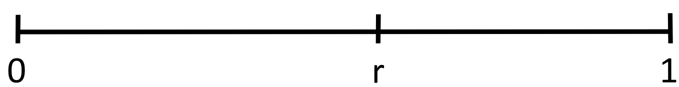


FIGURE 8.14

Algebraically, this problem is equivalent to solving the algebraic equation

$$\frac{r}{1-r} = \frac{1}{r}. \tag{8.2}$$

It can be shown that the solution is $r = \frac{\sqrt{5}-1}{2} \approx 0.618$ (see Exercise 8.3.1). The number $\frac{1}{r} \approx 1.618$ is called the *golden ratio*.

The next example helps motivate the golden section method.

Example 8.3.2 (Finding a Maximum)
Find the maximum value of the function $f(x) = 1 - e^{-x} + \frac{1}{1+x}$ on the interval [0, 20]. The graph of this function is shown in Figure 8.15. We see that the function is unimodal over [0, 20].

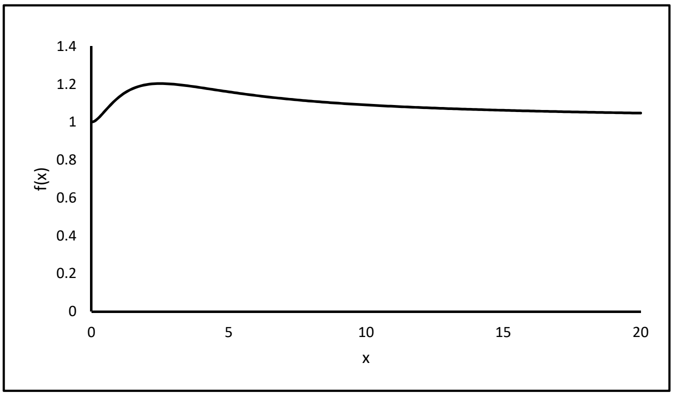


FIGURE 8.15

We begin the search by selecting two "test points" x_1 and x_2 in the interval [0, 20] in a way motivated by the Golden Ratio problem:

$$x_1 = 0 + (1 - 0.618)(20 - 0) = 7.64,$$
$$x_2 = 0 + 0.618(20 - 0) = 12.36.$$

Observe that x_1 defines an interval $[0,\, x_1]$ whose length is $(1 - r)$ times the length of [0, 20] while x_2 defines an interval $[0,\, x_2]$ whose length is r times the length of [0, 20]. Next we evaluate the function at these two test points:

$$f(x_1) = 1 - e^{-7.64} + \frac{1}{1 + 7.64} \approx 1.115,$$
$$f(x_2) = 1 - e^{-12.36} + \frac{1}{1 + 12.36} \approx 1.075.$$

These calculations show that $f(x)$ is decreasing on the interval $[x_1,\, x_2]$. Since $f(x)$ is unimodal, we can conclude that $f(x)$ is also decreasing on the interval $[x_2,\, 20]$. Therefore, we can eliminate the interval $[x_2,\, 20]$ as a location of the maximum value. This narrows the location of the maximum to the interval $[0,\, x_2]$. Then we can iterate this process. Each iteration yields a narrower interval for the location of the maximum. We can stop once the width of the interval becomes less than some specified value, called the *tolerance*. □

We generalize the above example in the following algorithm.

Golden Section Method: To find the maximum value of a unimodal function $f(x)$ on the interval $[a,\, b]$ with tolerance t:

1. Set $r = 0.618$ and define the test points

$$x_1 = a + (1 - r)(b - a),$$
$$x_2 = a + r(b - a).$$

2. Calculate $f(x_1)$ and $f(x_2)$.

3. Compare $f(x_1)$ and $f(x_2)$:

   a. If $f(x_1) \leq f(x_2)$, then the new interval is $[x_1, b]$ (i.e. a becomes x_1).

   b. If $f(x_1) > f(x_2)$, then the new interval is $[a, x_2]$ (i.e. b becomes x_2).

4. Calculate the width of the new interval. If this width is less than the tolerance t, then go to step 5. Otherwise, go back to step 1.

5. Estimate the location of the maximum value as the midpoint of the final interval. Call this number x^*. The approximate maximum value is $f(x^*)$.

We can modify this algorithm to find the minimum value of a unimodal function by switching the inequalities in step 3.

Example 8.3.3 (Implementing the Golden Section Method)
To find the maximum value of the function $f(x) = 1 - e^{-x} + \frac{1}{1+x}$ on the interval $[0, 20]$ with a tolerance of $t = 0.001$, rename a blank worksheet "**Golden Section**" and format it as in Figure 8.16. Copy row 10 down to row 59 to do up to 50 iterations.

	A	B	C	D	E
1	**Tolerance =**	0.001	**Num Iterations =**	=COUNTBLANK(I9:I59)	
2	**r =**	0.618			
3			**Final Interval**		
4			=OFFSET(B9,D1,0)	=OFFSET(C9,D1,0)	
5					
6		**x* =**	=(C4+D4)/2	**f(x*) =**	=1-EXP(-C6)+(1/(1+C6))
7					
8	**n**	**a**	**b**	**x_1**	**x_2**
9	0	0	20	=B9+(1-$B$2)*(C9-B9)	=B9+$B$2*(C9-B9)
10	=1+A9	=IF(F9>=G9,B9,D9)	=IF(F9>=G9,E9,C9)	=B10+(1-$B$2)*(C10-B10)	=B10+$B$2*(C10-B10)

	F	G	H	I
7			**Interval**	
8	**f(x_1)**	**f(x_2)**	**Width**	**Stop?**
9	=1-EXP(-D9)+(1/(1+D9))	=1-EXP(-E9)+(1/(1+E9))	=ABS(C9-B9)	=IF(H9<$B$1,"Stop", "")
10	=1-EXP(-D10)+(1/(1+D10))	=1-EXP(-E10)+(1/(1+E10))	=ABS(C10-B10)	=IF(I9="",IF(H10<$B$1,"Stop", ""),"NA")

FIGURE 8.16

The results show that the maximum value occurs somewhere in the interval $[2.5122, 2.5131]$, the midpoint of which is $x^* = 2.5127$, and the maximum value of $f(x)$ is approximately 1.2036. □

Exercises

8.3.1 Algebraically solve Equation 8.2 to show that $r = \frac{\sqrt{5}-1}{2}$.

8.3.2 Theoretically, it can be shown that the number of iterations needed in the golden section method is the smallest integer greater than or equal to k where

$$k = \frac{\ln\left(\frac{t}{b-a}\right)}{\ln(0.618)}$$

and $[a, b]$ is the initial interval. Demonstrate that this result is true by calculating k in the worksheet **Golden Section**, trying some different values of t, and observing that the number of iterations is indeed as claimed.

8.3.3 The reader might wonder if there is anything special about the number $r = 0.618$. To explore this question, change the value of r in the worksheet **Golden Section** to various values between 0.5 and 1. What happens to the number of iterations? Does 0.618 minimize the number of iterations?

8.3.4 The Golden Section method algorithm presented in this section is designed to find the maximum value of a unimodal function $f(x)$. To find the minimum value of a unimodal function $f(x)$, we have two options:

1. switch the inequalities in step 3, or
2. find the maximum value of $-f(x)$.

Use one of these options to resolve Example 8.1.1 with the Golden Section method and a tolerance of $t = 0.001$. How do your results compare to those found in the example?

8.3.5 Use the Golden Section method to solve each of the following problems.

a. Maximize $f(x) = -x^2 - 2x$ on $[-2, 1]$ with $t = 0.6$.

b. Maximize $f(x) = -x^2 - 3x$ on $[-3, 1]$ with $t = 0.6$.

c. Minimize $f(x) = x^2 + 2x$ on $[-3, 1]$ with $t = 0.01$.

d. Minimize $f(x) = x^3 - e^x$ on $[0, 2]$ with $t = 0.001$.

e. Maximize $f(x) = -x + e^x$ on $[-1, 3]$ with $t = 0.1$.

f. Maximize $f(x) = -|2 - x| - |5 - 4x| - |8 - 9x|$ on $[0, 3]$ with $t = 0.1$.

g. Maximize $f(x) = 91137 - 492.75x + 27550ln(x - 220)$ on $[220, 500]$ with $t = 0.01$.

8.3.6 Consider the problem of fitting a model of the form $W = kL^3$ to the data in the table below.

L	12.5	12.625	12.625	14.125	14.5	14.5	17.27	17.75
W	17	16	17	23	26	27	43	49

The goal of fitting a model of this form is to find the value of k so that the model "best" fits the data. In Section 2.4 we discussed several criteria for measuring how well a model fits a set of data. One of the criteria stated that the best fitting model should minimize the sum of the absolute values of the differences between the observed values and the predicted values. In this particular problem, this criterion says that the best value of k should minimize the function

$$f(k) = \left|17 - k \cdot 12.5^3\right| + \left|16 - k \cdot 12.625^3\right| + \cdots + \left|49 - k \cdot 17.75^3\right|.$$

Use the Golden Section method to find the best value of k with a tolerance of $t = 0.0001$. Assume that k is in the interval $[0, 1]$.

8.3.7 In the Golden Section algorithm presented in this section, x^*, the location of the maximum value of $f(x)$, is selected as the midpoint of the final interval. Another strategy for selecting x^* is to evaluate $f(x)$ at the right endpoint of the final interval, then at the right endpoint, and then at the midpoint. Then we select x^* as the one of these three points that yields the maximum value of $f(x)$. Modify the worksheet **Golden Section** to implement this strategy. Make sure your worksheet displays the value of x^*. Try some different values of the tolerance. Does this strategy yield different results than the original algorithm?

8.3.8 A variation on the golden section method is the *Fibonacci search method*, which relies on the Fibonacci sequence: $F_0 = 0$, $F_1 = 1$, and $F_n = F_{n-1} + F_{n-2}$ for $n \geq 2$. The steps in the Fibonacci search method are the same as the golden section method except for step 1:

1. Set N be the *smallest* value of n such that

$$F_n > \frac{b-a}{t}$$

and define the test points

$$x_1 = a + \left(\frac{F_{N-2}}{F_N}\right)(b-a),$$
$$x_2 = a + \left(\frac{F_{N-1}}{F_N}\right)(b-a).$$

Modify the worksheet **Golden Section** to implement the Fibonacci search method to find the maximum value of the function $f(x) = 1 - e^{-x} + \frac{1}{1+x}$ on the interval $[0, 20]$ with a tolerance of $t = 0.001$, as in Example 8.3.3. Compare the results from the Fibonacci and the golden section methods. (**Suggestions:** Use a "lookup chart" and the **VLOOKUP** function as in Section 6.9 to find N. If $\frac{b-a}{t} < 1$, then use $N = 2$.)

8.4 The One-Dimensional Gradient Method

Excel's Solver search engine contains an algorithm for solving nonlinear programs called *GRG Nonlinear*. In this section we discuss the basic ideas behind this algorithm for an unconstrained nonlinear program with one decision variable. This algorithm is based on the *gradient method* which relies on derivatives.

Example 8.4.1 (Using Solver)
Consider the following unconstrained nonlinear program with one decision variable:

$$\textbf{Minimize } f(x) = x^4 - 8x^3 + 20x^2 - 16.5x + 7$$

A graph of this function $f(x)$ is shown in Figure 8.17. We see that f has a local minimum near 0.6 and a global minimum near 3.4. In this problem, we are asking for the location of the global minimum.

To use Solver to solve this problem, rename a blank worksheet "**Solver**", and format it as in Figure 8.18.

Set the Solver parameters to minimize cell **B2** by changing cell **A2**. Do not add any constraints and do not make unconstrained variables non-negative. Select the solving method **GRG Nonlinear**. The solution it gives is shown in Figure 8.19. Notice that it found the local minimum. This is called *getting stuck at a local minimum*.

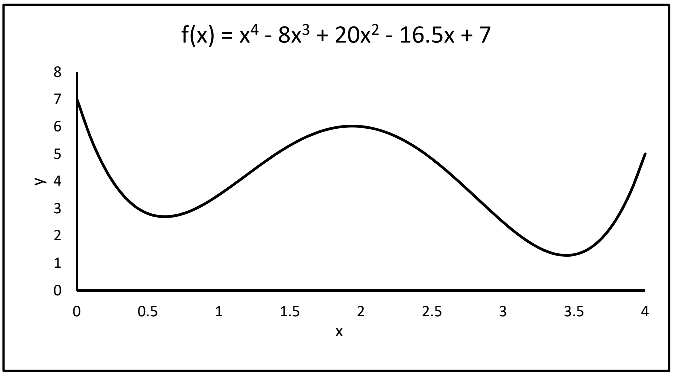


FIGURE 8.17

	A	B
1	**x**	**f(x)**
2	0	=A2^4-8*A2^3+20*A2^2-16.5*A2+7

FIGURE 8.18

	A	B
1	**x**	**f(x)**
2	0.618138	2.699114

FIGURE 8.19

	A	B
1	**x**	**f(x)**
2	3.444485	1.285246

FIGURE 8.20

Next set the value of x to 4 (i.e. enter 4 in cell **A2**) and resolve it. The results are shown in Figure 8.20. This time Solver found the global minimum.

□

To understand why Solver does not always find the global minimum, we need to examine how the gradient method works. The gradient method is an iterative approach for finding a critical point of the function (i.e. a value of x for which $f'(x) = 0$). A simplified version of the gradient method is given below.

One-dimensional Gradient Method: To minimize a differentiable function of one variable, $f(x)$

1. Choose an initial value, x_0.
2. Let $x_{k+1} = x_k - \lambda f'(x_k)$ where $\lambda > 0$ is some specified constant.
3. Repeat step 2 for 50 iterations (50 iterations is arbitrary).

The basic idea behind the gradient method is that if the derivative $f'(x_k)$ is positive, then the function is increasing, so we want to decrease x_k. If $f'(x_k)$ is negative, then the function is decreasing, so we want to increase x_k.

The constant λ affects how much x_k is changed in each step. In more sophisticated versions, the algorithm terminates when $|f'(x_k)|$ is sufficiently close to 0, rather than arbitrarily after 50 iterations (see Exercise 8.4.1). Ideally, when the algorithm terminates, x_k is near a critical point. We could modify this algorithm to maximize a function by changing the minus sign in step 2 to a plus sign.

Example 8.4.2 (Implementing the Gradient Method)
To implement the Gradient Method for the nonlinear program from Example 8.4.1, follow these steps:

1. First we create a graph of $f(x)$ over the interval $0 \leq x \leq 4$. Rename a blank workbook **1-Dim** and format it as in Figure 8.21. Copy row 3 down to row 42.

	A	B
1	**x**	**f(x)**
2	0	=A2^4-8*A2^3+20*A2^2-16.5*A2+7
3	=A2+0.1	=A3^4-8*A3^3+20*A3^2-16.5*A3+7

FIGURE 8.21

2. Create a graph similar to Figure 8.22.

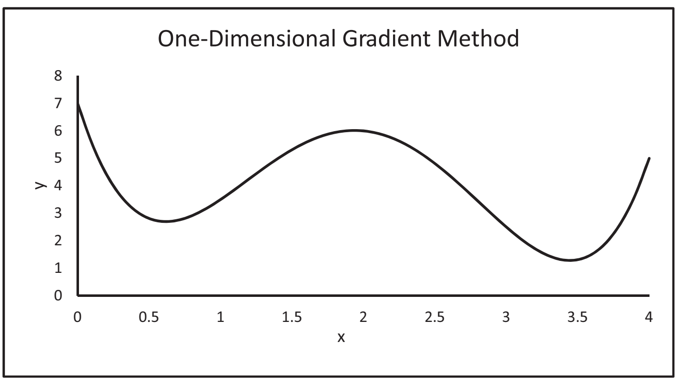


FIGURE 8.22

3. Add the formulas in Figure 8.23 and copy row 5 down to row 54.

	D	E	F	G
1	λ =	0.115		
2				
3	**k**	**x_k**	**f(x_k)**	**f'(x_k)**
4	0	0	=E4^4-8*E4^3+20*E4^2-16.5*E4+7	=4*E4^3-24*E4^2+40*E4-16.5
5	=D4+1	=E4-G4*$E$1	=E5^4-8*E5^3+20*E5^2-16.5*E5+7	=4*E5^3-24*E5^2+40*E5-16.5

FIGURE 8.23

4. To visualize how this algorithm works, add a scroll bar, link it to cell **J2** with a min and max of 0 and 50. Add the formulas in Figure 8.24.

	I	J	K	L
1			**x**	**f(x)**
2	**k =**	0	=OFFSET(E4,J2,0)	=OFFSET(F4,J2,0)

FIGURE 8.24

5. Add the cells **K2:L2** as a point on the graph, and format it so the graph resembles Figure 8.25.

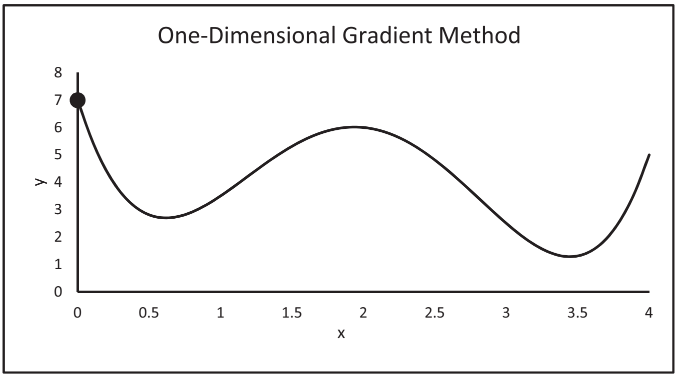


FIGURE 8.25

Move the slider back and forth and notice how the point on the graph eventually settles to one of the two local minima on the graph. Change the initial value of x and the value of λ and note how they affect the search and the solution. Note that if λ is large (e.g. $\lambda = 0.2$), the point jumps around and never settles at a local minimum. If λ is too small, the point doesn't move toward a local minimum very quickly. Also note that sometimes we get the local minimum at $x = 0.62$ and other times we get the global minimum at $x = 3.45$. $\square$

Exercises

8.4.1 In our simple gradient method algorithm, we arbitrarily performed 50 iterations. In more sophisticated versions of the algorithm, step 2 is repeated until some "stopping criterion" is met. One very simple stopping criterion is that the algorithm terminates when $|f'(x_k)| < \delta$ where $\delta > 0$ is some specified constant. Modify the worksheet **1-Dim** to incorporate this stopping criterion and output the number of iterations performed along with the minimum value of the function found. Allow for up to 200 iterations, and allow the user to input the value of δ.

8.4.2 Consider the problem of finding the maximum value of the following function:

$$f(x) = -x^6/6 + 3.5x^5 - 29x^4 + 127x^3 - 292x^2 + 331x.$$

a. Graph this function over the interval $0 \leq x \leq 9$.

b. Use 150 iterations of the gradient method to find the maximum value of this function. Use $\lambda = 0.001$.

c. Let x_0 be randomly chosen between 0 and 9. How well does the gradient method do at finding the global maximum of the function in 150 iterations?

d. Try using larger values of λ. How does this affect the performance of the algorithm?

8.4.3 As we saw in Example 8.4.2, the initial value of x, x_0 can affect the solution given by the gradient method. This suggests we try different values of x_0, and then select the best solution as the final solution.

a. Modify the worksheet **1-Dim** to let x_0 be a random number between 0 and 4.

b. Use a data table to store the results of the gradient method from 100 different values of x_0. Then display the best one.

c. How well does the data table do at finding the global minimum?

8.4.4 The gradient method requires knowledge of the derivative of the objective function. In Example 8.4.2, we used derivative rules to find $f'(x)$ and then used this to calculate $f'(x_k)$. This approach is not always completely practical because it requires the user to first find $f'(x)$, which may be very difficult. One way to numerically approximate $f'(x_k)$ is by using secant lines. Specifically, we choose a small quantity d and then calculate

$$f'(x_k) \approx \frac{f(x_k + d) - f(x_k)}{d}.$$

a. In the worksheet **1-Dim**, replace the formulas for $f'(x_k)$ with this approach for approximating $f'(x_k)$. Initially use $d = 0.01$, but let the user specify the value of d.

b. Try random values of x_0 between 0 and 4. Compare these results to the original.

c. Try different values of d. Does the value of d affect the results?

8.4.5 A chemical manufacturing company sells sulfuric acid at a price of \$100 per unit. The daily total production cost, in dollars, for x units is:

$$C(x) = 100000 + 50x + 0.0025x^2,$$

and the daily production is at most 7000 units.

a. How many units of sulfuric acid should the manufacturer produce to maximize daily profit, assuming the company sells all they produce? Solve this problem using Solver.

b. Now suppose the company doubles daily production capacity but the selling price and production cost remains the same. How many units of sulfuric acid should they produce now to maximize profit? Was all the extra production capacity really needed?

8.4.6 The Keyrific Computer Company manufactures keyboards in batches. They are trying to decide how many batches of keyboards to produce next year to minimize production and storage costs. Let x = the number of batches of keyboards produced per year. Define the parameters

- k = the annual storage cost of one keyboard,
- F = the fixed set up cost (includes insurance, machines, labor, etc.),
- v = the cost to produce one keyboard, and
- T = the total number of keyboards produced annually.

Suppose the costs are given by the functions

- Total production cost: $M(x) = x\left(F + \frac{vT}{x}\right)$,
- Average storage cost: $s(x) = \frac{kT}{2x}$, and
- Total cost function: $C(x) = x\left(F + \frac{vT}{x}\right) + \frac{kT}{2x}$.

Set up a spreadsheet that uses Solver to find the value of x that minimizes the total cost. Allow the user to enter the values of the parameters.

8.4.7 The Keyrific computer company also produces the SP6 computer. It costs $200 to produce each computer and there is a $5000 set up cost. The company is trying to decide how much to spend on advertising. They figure if they spend $$x$ on advertising, they will sell approximately $\sqrt{x}$ computers at $500 per computer. How much should they spend on advertising to maximize profit? Set up a spreadsheet that uses Solver to answer this question.

8.4.8 Each morning during rush hour, 10000 people travel from New Jersey to New York City. Each person either drives or takes the subway. If a person takes the subway, the trip lasts 40 minutes. If x is the number of people who drive, it takes $20 + 5(x/1000)$ minutes per person to make the trip.

a. How many people should drive to minimize the average travel time per person? Set up a spreadsheet that uses Solver to answer this question.

b. Suppose it costs $6 to take the subway, and the cost of traveling by car is $4 for gas. If time is money, and on average people work for $30/hour ($0.50 per minute), how many people should drive to minimize the average cost of commuting?

c. Now suppose the subway wants to expand its network and finance this by increasing its fare to $12. How many people should drive to minimize the average cost of commuting?

8.4.9 Dr. E. N. Throat has been taking x-rays of the trachea contracting during coughing. He has found that the trachea appears to contract by 33% (1/3) of its normal size. He has asked the department of mathematics to confirm or deny his claim. You perform some initial research and you find that under reasonable assumptions about the elasticity of the tracheal wall and about how air near the wall is slowed by friction, that the average velocity v of air moving through the trachea can be modeled by the equation

$$v = c\,(r_0 - r)\,r^2$$

where c is a positive constant, r_0 is the resting radius of the trachea in centimeters, and r is the radius of the trachea during coughing. Set up a spreadsheet that uses Solver to find the value of r that maximizes v. Do your results support or deny Dr. Throat's claim? Make sure you consider different values of the parameters c and r_0.

8.5 Two-Dimensional Gradient Method

The one-dimensional gradient method is an iterative method for approximating a local minimum or maximum value of a function of one variable $f(x)$. It is based on the idea that the derivative $f'(x)$ tells us whether a function is increasing or decreasing, thus giving us a direction in which to search. When we get close to a local minimum or maximum, $f'(x)$ gets close to 0.

In this section we apply a similar approach to a function of two variables, $f(x, y)$. This method involves the *gradient vector*.

Definition 8.5.1. Let $f(x, y)$ be a differentiable function of two variables, x and y. The *gradient vector* ∇f is defined as:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}.$$

It can be shown that the gradient vector ∇f at a point in the domain of f always points in the direction of the maximum rate of *increase* of the function, and that at a point of *local* minimum or maximum, (x_c, y_c) (called a *critical point*),

$$\nabla f(x_c, y_c) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Example 8.5.1 (A Function of Two Variables)
To better understand what the gradient tells us about a function, consider the function

$$f(x, y) = x^4 + y^4 - 4xy + 2.$$

First we examine some values of f for x and y between -2 and 2. Rename a blank worksheet "**Gradient**" and format it as in Figure 8.26. Copy row 4 down to row 11. Then copy column **C** to column **J**.

	A	B	C
1		**x**	
2	**y**	**-2**	**=B2+0.5**
3	**2**	=B\$2^4+\$A3^4-4*B\$2*\$A3+2	=C\$2^4+\$A3^4-4*C\$2*\$A3+2
4	**=A3-0.5**	=B\$2^4+\$A4^4-4*B\$2*\$A4+2	=C\$2^4+\$A4^4-4*C\$2*\$A4+2

FIGURE 8.26

The results are shown in Figure 8.27 (note that these results could be used to create a 3-D graph of the function in Excel called a *surface* or *contour* plot, but these graphs are somewhat crude, so use them at your discretion). There are three observations to highlight:

- The values of $f(x, y)$ range from 0 to 50.
- The global minimum value of 0 occurs twice, at $x = y = 1$ (the point $(1, 1)$) and at the point $(-1, -1)$.
- At the point $(0, 0)$, the value of f is 2 and the values "around" $(0, 0)$ are mostly larger than 2. This indicates that the point $(0, 0)$ is the location of a local, but not global, minimum value of f.

The partial derivatives of this function are

$$\frac{\partial f}{\partial x} = 4x^3 - 4y, \quad \text{and} \quad \frac{\partial f}{\partial y} = 4y^3 - 4x.$$

	A	B	C	D	E	F	G	H	I	J
1		**x**								
2	**y**	**-2**	**-1.5**	**-1**	**-0.5**	**0**	**0.5**	**1**	**1.5**	**2**
3	**2**	50.0	35.1	27.0	22.1	18.0	14.1	11.0	11.1	18.0
4	**1.5**	35.1	21.1	14.1	10.1	7.1	4.1	2.1	3.1	11.1
5	**1**	27.0	14.1	8.0	5.1	3.0	1.1	**0.0**	2.1	11.0
6	**0.5**	22.1	10.1	5.1	3.1	2.1	1.1	1.1	4.1	14.1
7	**0**	18.0	7.1	3.0	2.1	**2.0**	2.1	3.0	7.1	18.0
8	**-0.5**	14.1	4.1	1.1	1.1	2.1	3.1	5.1	10.1	22.1
9	**-1**	11.0	2.1	**0.0**	1.1	3.0	5.1	8.0	14.1	27.0
10	**-1.5**	11.1	3.1	2.1	4.1	7.1	10.1	14.1	21.1	35.1
11	**-2**	18.0	11.1	11.0	14.1	18.0	22.1	27.0	35.1	50.0

FIGURE 8.27

The gradient of f at (1, 1), for instance, is

$$\nabla f\,(1,\,1) = \begin{bmatrix} 4(1)^3 - 4(1) \\ 4(1)^3 - 4(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

This confirms that (1, 1) is the location of a local minimum or maximum. Similar calculations show that $\nabla f\,(-1,\,-1) = \nabla f\,(0,\,0) = (0,\,0)$, again confirming that $(-1,\,-1)$ and (0, 0) are the locations of local minimums or maximums.

Now consider the point $(1,\,-1)$. The gradient at this point is

$$\nabla f\,(1,\,-1) = \begin{bmatrix} 4(1)^3 - 4(-1) \\ 4(-1)^3 - 4(1) \end{bmatrix} = \begin{bmatrix} 8 \\ -8 \end{bmatrix}.$$

Informally, the positive x-coordinate of the gradient means that if we increase x "a little bit," the value of f should increase. Likewise, the negative y-coordinate means that decreasing y a little bit should increase the value of f. The equal magnitudes of the x- and y-coordinates mean that changing x and y the same small amount should increase the value of f the most. This is exactly what we see by examining Figure 8.27. □

The two-dimensional gradient method is an iterative approach for approximating a critical point of the function $f\,(x,\,y)$. A simplified version of the two-dimensional gradient method is given below.

Two-dimensional Gradient Method: To minimize a function of two variables, $f\,(x, y)$

1. Choose an initial point $(x_0,\,y_0)$.
2. Let $\begin{cases} x_{k+1} = x_k - \lambda \frac{\partial f}{\partial x}\,(x_k,\,y_k) \\ y_{k+1} = y_k - \lambda \frac{\partial f}{\partial y}\,(x_k,\,y_k) \end{cases}$ where $\lambda > 0$ is some specified constant.
3. Repeat step 2 for 50 iterations (50 is arbitrary).

This algorithm could be changed to maximize a function by changing the minus signs in step 2 to plus signs. In more sophisticated versions, the algorithm terminates when the norm of the gradient, $\|\nabla f\,(x_k,\,y_k)\|$, is sufficiently close to 0 (see Exercise 8.5.4).

Example 8.5.2 (Implementing the Two-dimensional Gradient Method)
Consider the problem of minimizing the function in Example 8.5.1,

$$f(x, y) = x^4 + y^4 - 4xy + 2,$$

on the domain $-2 \le x \le 2$, $-2 \le y \le 2$. To implement the two-dimensional gradient method with $\lambda = 0.02$ and the initial point $(0, 0.5)$, rename a blank worksheet **2-Dim** and format it as in Figure 8.28. Copy row 5 down to row 54.

	A	B	C	D	E	F
1		λ =	0.02			
2						
3	**k**	**x_k**	**y_k**	**$f(x_k,y_k)$**	**df/dx**	**df/dy**
4	0	0	0.5	=B4^4+C4^4-4*B4*C4+2	=4*B4^3-4*C4	=4*C4^3-4*B4
5	=A4+1	=B4-$C$1*E4	=C4-$C$1*F4	=B5^4+C5^4-4*B5*C5+2	=4*B5^3-4*C5	=4*C5^3-4*B5

FIGURE 8.28

The results are shown in Figure 8.29. The algorithm found the point (1, 1), approximately, where $f(1, 1) = 0$ and $f(1, 1) = (0, 0)$. Thus the algorithm found a locaion of the global minimum, as desired.

	A	B	C	D	E	F
3	**k**	**x_k**	**y_k**	**$f(x_k,y_k)$**	**df/dx**	**df/dy**
54	50	0.997875	0.997875	3.6E-05	-0.01695	-0.01694

FIGURE 8.29

In the one-dimensional gradient method we saw that the initial value can affect the solution. The same is true for the two-dimensional gradient method. Trying the different initial points shown in Table 8.1, we see that the algorithm doesn't always find the global minimum at $(-1, -1)$ or $(1, 1)$. It sometimes gets stuck at the local minimum at $(0, 0)$.

TABLE 8.1

Initial Point	$(-2, -2)$	$(-0.1, 0.1)$	$(0, 0)$	$(0.1, 0)$
Solution	$(-1, -1)$	$(0, 0)$	$(0, 0)$	$(1, 1)$

□

Example 8.5.3 (Using Solver)
Now let's examine how Solver handles this same minimization problem. Rename a blank worksheet "**Solver 2**" and format it as in Figure 8.30.

	A	B	C
1	**Solution**		
2	**x**	**y**	**f(x,y)**
3	0	0.5	=A3^4+B3^4-4*A3*B3+2

FIGURE 8.30

Format Solver to minimize cell **C3** by changing cells **A3:B3**. Add the constraints that both x and y need to be between -2 and 2. Select the solving method **GRG Nonlinear**. Notice that Solver found the solution (1, 1), which is indeed the location of the global minimum. However, if we enter the different initial points in Table 8.1 in the range **A3:B3** and rerun Solver, we find that Solver does not always find the global minimum. It can, and does, get stuck at a local minimum. This illustrates that we must be very careful about choosing the initial points. We might want to try several initial points (see Exercise 8.5.1. □

The above discussion illustrates that the gradient method is a rather simple algorithm for approximating critical points of a function. The drawback is that it cannot determine whether the solution it finds is a local or global minimum or maximum without some additional external knowledge of the function. Therefore, it, along with Solver, must be used with caution when solving nonlinear programs. To avoid potential problems with the gradient method, a problem should be modeled as linear rather than nonlinear whenever possible (see Exercise 8.5.11).

Exercises

8.5.1 Consider the problem of finding the minimum value of the function

$$f(x,\, y) = 2x^2 + 6xy + 6y^2 - 3x + 5y$$

on the domain $-10 \le x \le 10$, $-10 \le y \le 10$.

a. Modify the worksheet **2-Dim** to solve this problem. Use $\lambda = 0.1$ and let x_0 and y_0 be random numbers between -10 and 10.

b. Use a data table to store the results from 100 different initial points and display the best solution.

c. Now try using $\lambda = 0.2$. What happens in this case?

d. Now try using $\lambda = 0.0001$. What happens in this case?

8.5.2 Consider the problem of finding the *maximum* value of the function

$$f(x,\, y) = 90x - 0.1x^2 + 15y - 0.15y^2 - 0.05xy - 2,000.$$

a. Use 100 iterations of the two-dimensional gradient method with $\lambda = 0.5$ and the initial point (0, 0) to approximate the solution to this problem.

b. Now try using $\lambda = 0.1$. What happens in this case?

8.5.3 The distance the gradient method moves the point $(x_k,\, y_k)$ to $(x_{k+1},\, y_{k+1})$ is affected by two quantities: the length of the gradient, $\|\nabla f\,(x_k,\, y_k)\|$, and the value of λ. As the algorithm proceeds, $\|\nabla f\,(x_k,\, y_k)\|$ gets smaller and smaller, so the point does not move as far. As we have seen, if λ is too small, the algorithm doesn't find the optimal solution in a reasonable number of iterations.

One way to get around this problem is to increase the value of λ in each iteration. A simple algorithm for maximizing a function that incorporates this idea is given below:

1. Choose initial values $(x_0,\, y_0)$ and λ_0.

2. Let $\begin{cases} x_{k+1} = x_k + \lambda_k \frac{\partial f}{\partial x}\,(x_k,\, y_k) \\ y_{k+1} = y_k + \lambda_k \frac{\partial f}{\partial y}\,(x_k,\, y_k) \end{cases}$

3. Let $\lambda_{k+1} = \delta\lambda_k$ where $\delta > 1$ is some specified constant.

4. Repeat steps 2 and 3 for 50 iterations (50 is arbitrary).

a. Implement this algorithm to maximize the function in Exercise 8.5.2 using $\lambda_0 = 0.1$, $\delta = 1.1$, and the initial point (0, 0). How does the solution compare to that found in Exercise 8.5.2?

b. Now try using 100 iterations. What happens in this case?

How many iterations are really necessary for the gradient to get reasonably close to (0, 0)? What happens if we let the algorithm go for too many iterations?

8.5.4 In our simple gradient method algorithm, we arbitrarily performed 50 iterations. In more sophisticated versions of the algorithm, step 2 is repeated until some "stopping criterion" is met. One very simple stopping criterion is that the algorithm terminates when $\|\nabla f(x_k, y_k)\| < \delta$ where $\delta > 0$ is some specified constant. Modify the worksheet **2-Dim** to incorporate this stopping criterion and output the number of iterations performed along with the minimum value of the function found. Allow for up to 100 iterations, and allow the user to input the value of δ.

8.5.5 The gradient method requires knowledge of the derivative of the objective function. In Example 8.5.2, we used derivative rules to find $\nabla f(x, y)$ and then used this to calculate $\nabla f(x_k, y_k)$. This approach is not always completely practical because it requires the user to first find $\nabla f(x, y)$, which may be very difficult. One way to numerically approximate $\nabla f(x_k, y_k)$ is by choosing a small quantity d and then approximating the partial derivatives with

$$\frac{\partial f}{\partial x}(x_k, y_k) \approx \frac{f(x_k + d, y_k) - f(x_k, y_k)}{d}, \text{ and}$$
$$\frac{\partial f}{\partial y}(x_k, y_k) \approx \frac{f(x_k, y_k + d) - f(x_k, y_k)}{d}.$$

a. In the worksheet **2-Dim**, replace the formulas for $f'(x_k)$ with this approach for approximating the partial derivatives. Initially use the starting point (0, 0.5) and $d = 0.01$, but let the user specify the value of d.

b. Try the starting points in Table 8.1. Compare these results to the original.

c. Try different values of d. Does the value of d affect the results?

8.5.6 The least-square criterion for fitting a straight line $f(x) = mx + b$ to a set of data $(x_1, y_1), \ldots, (x_n, y_n)$, as discussed in Section 2.4, states that m and b should minimize the quantity

$$S = \sum_{i=1}^{n} (y_i - f(x_i))^2 = \sum_{i=1}^{n} (y_i - mx_i - b)^2.$$

The table below shows the shoe length and height of ten students (measured in inches). Use Solver to find the values of m and b that minimize S. Then plot the data and fit a linear trendline. Compare the results.

Shoe Length (x)	9	10	10.5	11	11.5	11.75	12	12.5	12.75	13
Height (y)	62	64	64.5	69	70	73	72	75	74	77

8.5.7 Suppose two populations a and b are modeled by the linear discrete dynamical system

$$\begin{aligned} a_{n+1} &= c_1 a_n + d_1 b_n \\ b_{n+1} &= c_2 a_n + d_2 b_n \end{aligned}$$

as discussed in Section 4.4 where a_n and b_n are the populations, in thousands, at the end of year n. The goal of this problem is to find appropriate values of the parameters c_1, c_2, d_1 and d_2.

a. Suppose that $a_0 = 19.28$, $b_0 = 33.44$, and $c_1 = c_2 = d_1 = d_2 = 1$. Use the model to predict the values of a_n and b_n for $n = 1, 2, 3,$ and 4.

b. The table below gives the observed values of a_n and b_n for $n = 1, 2, 3,$ and 4.

$\boldsymbol{n}$	1	2	3	4
$\boldsymbol{a_n}$	25.82	26.69	32.6	37.12
$\boldsymbol{b_n}$	21.47	35.11	35.57	43.96

Informally, the least-squares criterion says that the parameters c_1, c_2, d_1 and d_2 should be chosen to minimize the quantity

$$S = \sum_{i=1}^{n} (\text{Observed } a_n - \text{Predicted } a_n)^2 + \sum_{i=1}^{n} (\text{Observed } b_n - \text{Predicted } b_n)^2.$$

Use Solver to find the values of the parameters that minimize S.

8.5.8 Consider a triangle with vertices $A = (0, 0)$, $B = (3, 0)$, and $C = (1, 3)$. Let $P = (x, y)$ be a point on the $x - y$ plane and let AP, BP, and CP denote the distances from points A, B, and C to P, respectively.

a. Use Solver to find the point P that minimizes the sum of the distances $AP+BP+CP$. Such a point is called a *Fermat point*, a *Torricelli point*, or a *Fermat-Torricelli point*.

b. Now suppose vertex C changes to $C = (-2, 3)$. Repeat part a. What do you notice about this solution?

8.5.9 A small company is planning to install a central computer with cable links to five new departments. According to the floor plan, the peripheral computers for the five departments will be situated at the points in the table below. Use Solver to find the point at which they should locate the central computer to minimize the total amount of cable needed. (Assume cables can be run in straight lines from the central computer to the peripherals.)

	Department				
Point	**A**	**B**	**C**	**D**	**E**
$\boldsymbol{x}$	15	25	60	75	80
$\boldsymbol{y}$	60	90	75	60	25

8.5.10 The town of Schoolville is laid out on a 10×10 grid and is broken up into four subdivisions. The number of school-age students in each subdivision and the x- and y-coordinates of the center of each subdivision are shown in the table below. The school board plans to construct two new schools, A and B, with capacities of 500 and 800 students, respectively. They need to make two decisions:

1. the coordinates of where to build each school, and
2. how many students from each subdivision to send to each school.

The distance a student must travel to school is simply the Euclidean distance from the center of the student's subdivision to the school. Use Solver to find the solution that minimizes the total distance traveled by all the students.

Subdivision	Students	x	y
1	325	1	1
2	350	3	5
3	350	7	9
4	275	8	4

8.5.11 Consider Exercise 7.3.12.

a. Suppose we model the specifications in the following way:

Mixture	Specifications
Walnut Lover's	$0.35 \leq \dfrac{\text{weight of walnuts in W.L.}}{\text{total weight of W.L.}}$
	$0.25 \geq \dfrac{\text{weight of cashews in W.L.}}{\text{total weight of W.L.}}$
Cashew Lover's	$0.45 \geq \dfrac{\text{weight of peanuts in C.L.}}{\text{total weight of C.L.}}$
	$0.45 \leq \dfrac{\text{weight of cashews in C.L.}}{\text{total weight of C.L.}}$
Premium	$0.10 \geq \dfrac{\text{weight of peanuts in P.}}{\text{total weight of P.}}$
	$0.50 \leq \dfrac{\text{weight of walnuts in P.}}{\text{total weight of P.}} \leq 0.65$
	$0.15 \leq \dfrac{\text{weight of cashews in P.}}{\text{total weight of P.}}$

This model is nonlinear because the quantities in the numerators and the quantities in the denominators are functions of the decision variables. Therefore we have decision variables divided by decision variables, making the model nonlinear. Implement this nonlinear model in Excel (don't forget about the constraints on the total number of each type of nut available). Try to solve it with Solver using the GRG Nonlinear solving method. Start with all the decision variables equaling 0. What happens?

b. Now try to solve this nonlinear model starting with all the decision variables equaling 1.

c. We can turn the nonlinear model into a linear model by simply multiplying each inequality by the denominator:

Mixture	Specifications
Walnut Lover's	0.35(total weight of W.L.) $\leq$ weight of walnuts in W.L.
	0.25(total weight of W.L.) $\geq$ weight of cashews in W.L.
Cashew Lover's	0.45(total weight of C.L.) $\geq$ weight of peanuts in C.L.
	0.45(total weight of C.L.) $\leq$ weight of cashews in C.L.
Premium	0.10(total weight of P.) $\geq$ weight of peanuts in P.
	0.50(total weight of P.) $\leq$ weight of walnuts in P.
	0.65(total weight of P.) $\geq$ weight of walnuts in P.
	0.15(total weight of P.) $\leq$ weight of cashews in P.

This model is linear because we are not dividing any decision variables. Implement this model in Excel and solve it with the Simplex method. Start with all the decision variables equaling 0. Does this optimal solution equal that found in part b.?

8.6 Lagrange Multipliers

The optimization problems covered so far in this chapter have been relatively unconstrained, meaning there have been no constraints other than bounds on the values of individual variables. In this section we introduce a technique for solving constrained nonlinear optimization programs called the *method of Lagrange multipliers* which utilizes gradients. We limit our discussion to programs involving two decision variables and one constraint, though these ideas can be extended to any number of variables and constraints.

The first example is a geometric application of a nonlinear constrained optimization problem.

Example 8.6.1 (Minimizing Distance)
Find the points on the graph of the hyperbola $xy = 2$ that are closest to the origin.

A graph of the curve $xy = 2$ is shown in Figure 8.31. Graphically, we are looking for the points on this curve that are closest to the origin (indicated by the dots on the curve). The fact that the points must be on this curve can be written as the the constraint $xy - 2 = 0$.

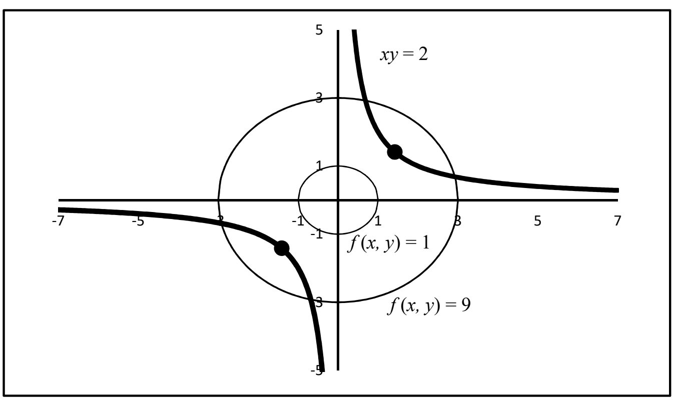


FIGURE 8.31

We want to minimize the distance to the origin. The distance from the origin to a generic point (x, y) is $\sqrt{x^2 + y^2}$. To simplify the algebra, we will minimize the square of the distance, $x^2 + y^2$. Thus we write our program as

$$\textbf{Minimize } f(x, y) = x^2 + y^2$$
$$\textbf{Subject to } xy - 2 = 0.$$

Figure 8.31 shows all the points such that $f(x, y) = 1$, called a *level curve* of $f(x, y)$. We see this curve does not intersect the curve $xy = 2$. Thus it is *not* possible for $f(x, y)$ to be a small as 1. Figure 8.31 also shows the level $f(x, y) = 9$. We see this curve does intersect the curve $xy = 2$. Thus it *is* possible for $f(x, y)$ to be as small as 9. However, it appears that there are smaller possible values of $f(x, y)$.

□

The following theorem, which we present without proof, is the key to solving a program such as in this example.

Theorem 8.6.1. *For the program*

$$\textbf{\textit{Minimize}} f(x, y)$$
$$\textbf{\textit{Subject to}} \; g(x, y) = 0,$$

if the optimal solution lies at the point (a, b) *where* $\nabla g(a, b) \neq \mathbf{0}$*, then*

$$\nabla f(a, b) = \lambda \nabla g(a, b) \tag{8.3}$$

for some constant λ *called a Lagrange multiplier.*

Theorem 8.6.1 gives a necessary condition for there to be an optimal solution to the program at the point (a, b). This theorem suggests that a solution to Equation (8.3) is a candidate for the location of an optimal solution. Like the gradient method, such a solution is not guaranteed to be a globally optimal solution.

Example 8.6.2 (Using Lagrange Multipliers)
Let's use Theorem 8.6.1 to solve Example 8.6.1. We have $f(x, y) = x^2 + y^2$ and $g(x, y) = xy - 2$. The gradients are

$$\nabla f(x, y) = \begin{bmatrix} 2x \\ 2y \end{bmatrix} \quad \text{and} \quad \nabla g(x, y) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

The requirement $\nabla f(x, y) = \lambda \nabla g(x, y)$ leads to the systems of equations

$$2x = \lambda y$$
$$2y = \lambda x.$$

Multiplying the first equation by x and the second by y yields

$$2x^2 = \lambda xy$$
$$2y^2 = \lambda xy.$$

Thus we see that

$$2x^2 = 2y^2 \quad \Rightarrow \quad x^2 = y^2.$$

Now, the constraint $xy - 2 = 0$ means that $xy = 2$ so that x and y must have the same sign. This, combined with the fact that $x^2 = y^2$ means that $x = y$. Combining this with the fact that $xy = 2$ yields

$$x^2 = 2 \quad \Rightarrow \quad x = y = \pm\sqrt{2}.$$

Thus the minimum distance occurs at the points $(-\sqrt{2}, -\sqrt{2})$ and $(\sqrt{2}, \sqrt{2})$. The equation $2x^2 = \lambda xy$ means that $\lambda = 2$. □

As we see, the algebra required for Lagrange multipliers can get rather involved. Luckily, Excel will do this automatically.

Example 8.6.3 (Using Excel)
To solve Example 8.6.1, rename a blank worksheet "**Min Dist**" and format it as in Figure 8.32.

	A	B	C	D
1	**x**	**y**	**f(x, y)**	**g(x, y)**
2	1	1	=A2^2+B2^2	=A2*B2-2

FIGURE 8.32

In **Solver**, set the objective to minimize **$C$2** by changing the variable cells **$A$2:$B$2**, add the constraint **$D$2=0**, and select the solving method **GRG Nonlinear**. After pressing **Solve**, select **Sensitivity** under **Reports** and press **OK**. The results are shown in Figure 8.33.

Variable Cells

Cell	**Name**	**Final Value**	**Reduced Gradient**
$A$2	x	1.414213534	0
$B$2	y	1.414213534	0

Constraints

Cell	**Name**	**Final Value**	**Lagrange Multiplier**
$D$2	g(x, y)	-7.96717E-08	2.000001686

FIGURE 8.33

Solver yields the same positive x and y solution we got algebraically (at least within rounding) and the same value of the Lagrange multiplier λ. Solver also displays a "reduced gradient," which has no real meaning for our purposes.

Now change the initial values of x and y to -1 and resolve. Solver yields the negative x and y solution we got algebraically. This illustrates that the initial values of the decision variables can affect the results, as we've seen before. □

The next example illustrates a practical application of the value of the Lagrange multiplier λ.

Example 8.6.4 (Producing E-phones)
A company is planning the production of a new brand of E-phones that are supposed to capture the market by storm. The two main input components of the new E-phone are the circuit boards and the relay switches that make the phone faster, smarter, and have more memory. The number of E-phones that can be produced is estimated to equal

$$E(a,\, b) = 200a^{1/2}b^{1/4}$$

where $E(a,\, b)$ is the number of phones produced, a is the number of hours of labor (in thousands) spent producing circuit boards, and b is the number of hours (in thousands) spent producing relays. Such a function is known to economists as a *Cobb-Douglas function*. Suppose that circuit board production costs \$5/hour and relay production costs \$10/hour and that the company has \$150 thousand to spend on labor. How many circuit boards and relays should be produced to maximize the number of E-phones produced?

The constraint is the money available for labor. This is easily modeled as

$$5a + 10b = 150 \quad \Rightarrow \quad 5a + 10b - 150 = 0.$$

Thus our program is

$$\textbf{Maximize } E(a,\, b) = 200a^{1/2}b^{1/4}$$
$$\textbf{Subject to } g(a,\, b) = 5a + 10b - 150 = 0.$$

We'll let Excel do all the work. Rename a blank worksheet "**E-phone**" and format it as in Figure 8.34.

	A	B	C	D
1	**a**	**b**	**E(a, b)**	**g(a, b)**
2	1	1	=200*A2^0.5*B2^0.25	=5*A2+10*B2-150

FIGURE 8.34

Set up **Solver** the same as Example 8.6.3, except choose to maximize cell **\$E\$2**. The sensitivity report is shown in Figure 8.35. These results show that we should spend 20 thousand hours producing circuit boards and 5 thousand hours producing relays and that the Lagrange multiplier is $\lambda = 6.687$. The worksheet **E-phones** shows that this solution results in the production of 1337.481 E-phones (in reality this number would be rounded off to a whole number).

To understand what the Lagrange multiplier $\lambda = 6.687$ means in a practical sense, let's suppose we increase the amount of money available for labor by \$1 thousand. This is easily done by changing the 150 in cell **\$D\$2** to 151. Resolving the program using the exact same **Solver** parameters yields the solution shown in Figure 8.36.

Note that the maximum value of $E(a,\, b)$ increased from 1337.481 to 1344.162, an increase of 6.687. □

This example illustrates the following interpretation of the Lagrange multiplier:

> If the Lagrange multiplier for a constraint resource is λ and the amount of that resource increases by a small amount Δ, then the maximum value of the objective function will increase by approximately $\lambda\Delta$.

This interpretation is similar to the *shadow price* discussed in Section 7.7.

One final note: These examples dealt with constrained optimization problems with only one constraint. The method of Lagrange multipliers can be applied to problems with more than one constraint. See Exercises 8.6.2 and 8.6.3 for examples.

Variable Cells

Cell	Name	Final Value	Reduced Gradient
$A$2	a	19.99999693	0
$B$2	b	5.000001535	0

Constraints

Cell	Name	Final Value	Lagrange Multiplier
$D$2	g(a, b)	9.9476E-13	6.687397907

FIGURE 8.35

	A	B	C	D
1	**a**	**b**	**E(a, b)**	**g(a, b)**
2	20.13333	5.033333	1344.162	2.5E-10

FIGURE 8.36

Exercises

8.6.1 Suppose x, y, and z are positive. Find the minimum value of $f(x, y, z) = x + y + z$ subject to the constraint $xyz = 1$.

8.6.2 Suppose we want to draw a right triangle with legs of length a and b and hypotenuse of length c that has a perimeter of 10. Find the dimensions that maximize the area $A = 0.5ab$. What type of triangle do we get? (**Hint**: Use the Pythagorean theorem for one of your constraints.)

8.6.3 Suppose we want to draw a triangle (not necessarily a right triangle) with sides of length a, b, and c that has a perimeter of 10. Let θ be the angle between sides a and b. Find the dimensions, and the angle θ, that maximize the area $A = 0.5ab\sin\theta$. What type of triangle do we get? (**Hint**: Use the law of cosines $c^2 = a^2 + b^2 - 2ab\cos\theta$ for one of your constraints.)

8.6.4 Consider the problem of finding the points(s) on the surface $z = xy + 5$ closest to the origin subject to the constraint that x, y, and z are all between -5 and 5.

a. Model this program in Excel and try solving it with Solver.

b. Try several different initial values of x, y, and z. Do you always get the same solution?

c. Try using the **Multistart** option in Solver. What solution do you get?

8.6.5 A castle is surrounded by a 3-m wide moat and the moat is surrounded by a 2-m tall outer wall as in Figure 8.37. Soldiers attacking the castle plan to lay a ladder on top of the outer wall to reach the castle wall, as indicated in the figure. Find the minimum length of ladder needed. (**Hint**: Let x and y be as indicated in the figure. Find the length of the

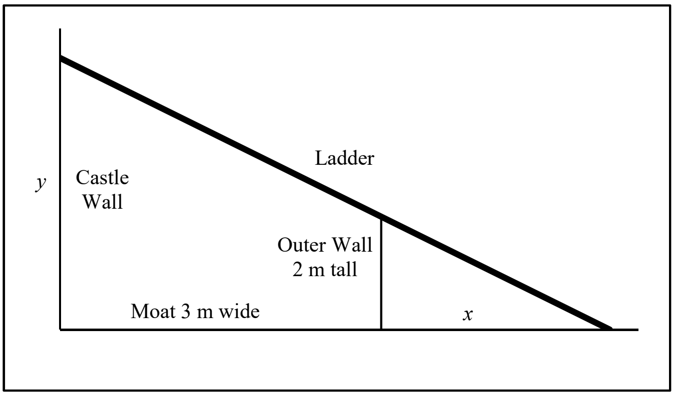


FIGURE 8.37

ladder in terms of x and y. Use similar triangles to set up a relationship between x and y. Use this relationship as a constraint.)

8.6.6 Let I_1 be the line of intersection of the two planes $2x+y+2z = 15$ and $x+2y+3z = 30$ and I_2 be the line of intersection of the two planes $x - y - 2z = 15$ and $3x - 2y - 3z = 20$. Find the points on these two lines that are closest together. (**Hint**: Let (x_1, y_1, z_1) be a generic point on line I_1 and (x_2, y_2, z_2) be a generic point on line I_2. Find the distance between these two points in terms of the x's, y's, and z's. We want to minimize this distance. Use the equations for the planes to set up the constraints.)

8.6.7 The graph of the curve $x^2+6y^2+3xy = 50$ is an ellipse. Find the point of this ellipse with the largest x-coordinate.

8.6.8 Consider a rectangular box with dimensions x, y, and z and fixed volume V. Find the dimensions that minimize the surface area of the box. Try various values of V. What type of box do you get?

8.6.9 Suppose the number of units produced E, in thousands, of a certain product is described by the Cobb-Douglas function

$$E(a, b) = 1.2a^{0.3}b^{0.6}$$

where a is the amount of capital and b is the amount of labor used. Further suppose that the total production cost is described by the function

$$T = 10000a + 7000b.$$

a. If there is a budget of \$55,060 to spend on production costs, find the maximum number of units that can be produced. Also find the Lagrange multiplier.

b. Now suppose the budget increases by \$10,000. Resolve the problem.

c. Is the result of part b. consistent with the practical interpretation of the Lagrange multiplier? Briefly explain why or why not.

8.6.10 A company manufactures perfumes and can purchase up to 1925 oz of the main chemical ingredient for \$10 per oz. At a cost of \$3 per oz the chemical can be manufactured into an ounce of perfume #11, and at a cost of \$5 per oz the chemical can be manufactured into an ounce of the higher priced perfume #2. An advertising firm estimates that if x ounces of perfume #1 are manufactured, it will sell for $\$(30 - 0.01x)$ per ounce. If y ounces of perfume #2 are produced, it can sell for $\$(50 - 0.02y)$ per ounce.

a. Find the values of x and y that maximize the company's profit.

b. What is the value of the Lagrange multiplier for the constraint that the company can purchase up to 1925 oz? Interpret the meaning of this value in the context of this problem.

8.6.11 A company manufactures flat screen TVs and wants to introduce two new TVs. Both TVs contain similar internal parts but one has a 55-inch screen and the other a 60-inch screen. In addition to \$40,000 in fixed costs it costs the company \$450 to produce the 55-inch TV and \$550 to produce the 60-inch screen. The manufacturer's suggested retail price is \$590 for the 55-inch and \$610 for the 60-inch. In a competitive market, sales help to reduce prices. For each size of TV, each additional TV sold the price drops by \$0.10. Additionally the price of the 55-inch TV is reduced by \$0.03 for every 60-inch TV sold and the price of a 60-inch TV is reduced by \$0.04 for every 55-inch TV sold. It takes 4 hours to build a 55-inch TV and 7 hours to assemble a 60-inch. There are 3,000 hours available for labor. Assuming the company sells all the TVs it produces, determine how many of each type of TV the company should manufacture to maximize profits.

8.6.12 The value of a Lagrange multiplier for a resource constraint tells us approximately how much the objective function will increase if the amount of the corresponding resource is increased by a small amount. In a problem with two or more resource constraints, each constraint has a Lagrange multiplier which tells us approximately how much the objective function will increase if the corresponding resource is increased a small amount while the other resource(s) are held constant. As an application of this idea, consider the following problem: The manager of a produce department is trying to decide how many of various types of fruit baskets to make to maximize profit. She has only 3 apples and 7 oranges left and models the problem as

$$\begin{aligned} \textbf{Maximize } & \text{Profit} = x^2 + y^2 + 3z \\ \textbf{Subject to } & x + y = 3 \text{ (apple constraint)} \\ & x + 3y + 2z = 7 \text{ (orange constraint)} \end{aligned}$$

where x, y, and z are the number of each type of basket produced.

a. Solve this problem and find the value of the Lagrange multiplier for each constraint.

b. If the manager has some extra money to spend on apples or oranges (but not both), should she purchase more apples or oranges? Briefly explain your reasoning.

8.6.13 Suppose a newspaper publisher must purchase three types of paper stock. They must meet their demand while minimizing the total cost using an *Economic Order Quantity* (EOQ) model. Let Q_i be the number of rolls of type i purchased per week. The cost of purchasing and storing type i per week is given by

$$C(Q_i) = \frac{a_i d_i}{Q_i} + h_i \frac{Q_i}{2}$$

where

- a_i is the order cost (in dollars),
- d_i is the order rate (in rolls/week),
- h_i is the storage cost (in dollars/week/roll), and
- $Q_i/2$ is the average amount on hand.

The total cost per week is the sum of the three costs,

$$C(Q_1, Q_2, Q_3) = C(Q_1) + C(Q_2) + C(Q_3).$$

The only constraint is the amount of available storage area. The amount of storage area required for type i is

$$S(Q_i) = s_i Q_i$$

where s_i is the number of ft^2 required per roll. The total required storage area is the sum of the three areas,

$$S(Q_1, Q_2, Q_3) = S(Q_1) + S(Q_2) + S(Q_3).$$

The rolls cannot be stacked; they must be be laid side-by-side on the floor. The values of the parameters are given in the table below

Parameter	Type I	Type II	Type III
a	25	18	20
d	32	24	20
h	1	1.5	2
s	4	3	2

a. Find the optimal solution if there is 200 ft^2 of storage space available. Make sure your solution is given in whole numbers.

b. Now suppose the publisher wants to increase their storage capability. Will increasing storage actually decrease total cost? If so, how much additional storage would you recommend? Briefly explain your reasoning.

8.7 Branch and Bound

Branch and bound is a technique often used for solving *integer programs*, which are programs where the decision variables are required to be integers. In this section we focus on a specific type of integer program called a *binary integer program* (*BIP*) where the decision variables are required to be either 0 or 1 (i.e. are binary).

To illustrate how branch and bound works, consider the following scenario from Example 7.4.3:

> Nathan and Laura are trying to sell their house which has two bedrooms and two bathrooms. To increase the house's value, they want to remodel one or more rooms. They have estimated the costs of remodeling each room and their real estate agent has estimated the increase in the house's value if each room was remodeled as shown in Table 8.2 (where costs and increases in values are given in thousands of dollars). They have only \$10,000 to spend remodeling, and they have decided that they can't do both bathroom 2 and bedroom 2. They will only do bathroom 2 if they also do bathroom 1. Also, they will only do bedroom 2 if they also do bedroom 1. Which rooms should they remodel to maximize the total increase in their house's value?

TABLE 8.2

Room	Decision Variable	Remodeling Cost	Increase in House Value
Bathroom 1	x_1	6	9
Bedroom 1	x_2	3	5
Bathroom 2	x_3	5	6
Bedroom 2	x_4	2	4

We modeled this scenario with the BIP

$$\begin{array}{rrcl} \textbf{Maximize} & Z = & 9x_1 + 5x_2 + 6x_3 + 4x_4 & \\ \textbf{Subject to} & & 6x_1 + 3x_2 + 5x_3 + 2x_4 & \le 10 \\ & & x_3 + x_4 & \le 1 \\ & & -x_1 + x_3 & \le 0 \\ & & -x_2 + x_4 & \le 0 \\ & & x_1,\, x_2,\, x_3,\, x_4 \text{ are binary} & , \end{array} \tag{8.4}$$

and solved it using Solver with a binary constraint yielding the optimal solution $x_1 = x_2 = 1$, $x_3 = x_4 = 0$ (denoted (1, 1, 0, 0)) with an objective function value of $Z = 14$. Observe that the objective function and all the constraints, except the binary constraint, in Program (8.4) are linear. For this reason, this program is called a *linear binary integer program.* Branch and bound is one technique for handling the binary constraint.

In an integer program such as this, the number of possible solutions is finite and the list of possible solutions is, in principle, easy to enumerate. In this example, there are four decision variables and each one has two possible values, so there are a total of $2^4 = 16$ possible binary solutions. A very crude method to solve this problem is to:

1. List all possible solutions.

2. Test each one for feasibility.

3. Evaluate the objective function at each feasible solution.

4. Choose the feasible solution that gives the largest value of the objective function.

This method certainly would work in principle. However, for large problems with many thousands of decision variables (as occur in real applications) the number of possible solutions can be extremely large making this method computationally unfeasible. Branch and bound is a more intelligent, systematic way of testing different possibilities and determining which ones do not need to be tested.

In branch and bound we set one or more variables equal to certain values, solve for the remaining variables, and compare the solution to what has already been found. We can organize our work in a *solution tree* as shown in Figure 8.38 which can be found in the workbook **Branch and Bound** available on the website for this book.

To solve Program (8.4), follow these steps in the workbook **Branch and Bound**:

Step 1: Solve a relaxation

A *relaxation* is a version of the problem of interest in which one or more of the constraints has been changed slightly (or "relaxed"), resulting in a simpler problem that can be solved with existing techniques, like the Simplex method. Without the binary constraint, this program is linear which can be solved with the Simplex method. So we relax the binary

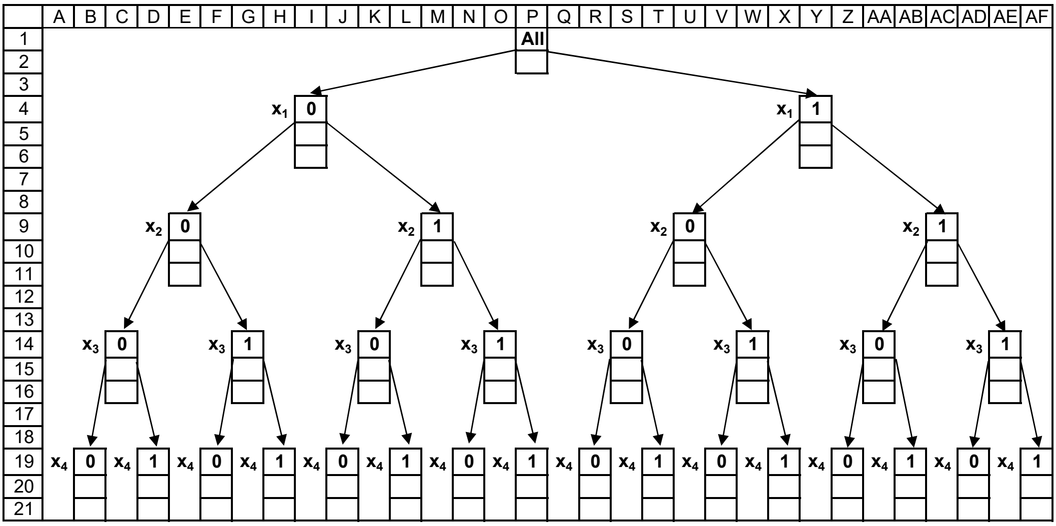


FIGURE 8.38

constraint by only requiring that the decision variables be between 0 and 1. This yields the following relaxation:

$$
\begin{array}{rrl}
\textbf{Maximize} & Z = & 9x_1 + 5x_2 + 6x_3 + 4x_4 \\
\textbf{Subject to} & & 6x_1 + 3x_2 + 5x_3 + 2x_4 \leq 10 \\
& & x_3 + x_4 \leq 1 \\
& & -x_1 + x_3 \leq 0 \\
& & -x_2 + x_4 \leq 0 \\
& & 0 \leq x_1, x_2, x_3, x_4 \leq 1
\end{array}
$$

This relaxation is programmed into the workbook **Branch and Bound** as shown in Figure 8.39. The appropriate constraints are already entered into Solver (the reader should verify this). The upper and lower bounds in rows 24 and 26 are bounds for each variable as given in the last constraint of the relaxation. In this case the bounds are all 0 and 1. In upcoming steps, the bounds may be different for each variable.

	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
23																**x_1**	**x_2**	**x_3**	**x_4**
24					**Coefficients**										**Upper Bound**	1	1	1	1
25	**Objective**			9	5	6	4	=	0						**Solution =**	0	0	0	0
26	**Constraint 1**			6	3	5	2	=	0	≤	10				**Lower Bound**	0	0	0	0
27	**Constraint 2**			0	0	1	1	=	0	≤	1								
28	**Constraint 3**			-1	0	1	0	=	0	≤	0								
29	**Constraint 4**			0	-1	0	1	=	0	≤	0								

FIGURE 8.39

Running Solver by simply opening the Solver window and pressing the **Solve** button yields an optimal solution of (0.83, 1, 0, 1) with $Z = 16.5$. This solution is obviously not feasible since it is not binary. If this solution were binary, then we would be done since this solution would satisfy all the constraints in the original program, including the binary constraints. However, it does tell us that the best binary solution will have $Z \leq 16.5$. This gives us an upper-bound on the optimal value of the objective function.

Step 2: Branch and create sub-problems

Branching is the process of setting the values of one or more decision variables to either 0 or 1 which creates new relaxations, called *sub-problems*. The first step in branching is shown in Figure 8.40. The number 16.5 entered under "**All**" reminds us that 16.5 is an upper-bound on the best binary solution.

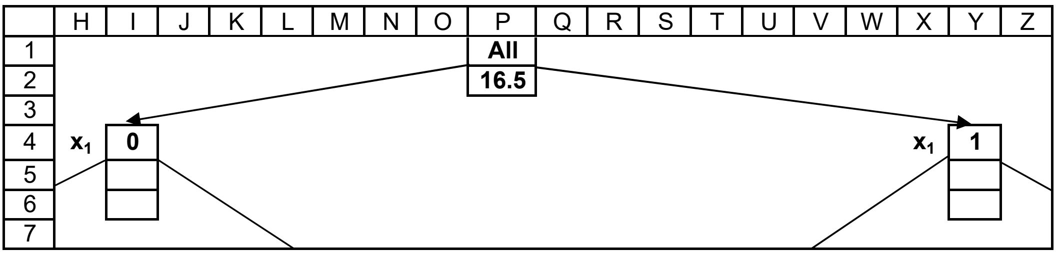


FIGURE 8.40

Sub-problem 1: First we will branch to the left and set $x_1 = 0$. We modify our relaxation by adjusting the upper bound of x_1 to 0 as in Figure 8.41.

	T	U	V	W	X	Y	Z
23				x_1	x_2	x_3	x_4
24	**Upper Bound**			0	1	1	1
25	**Solution =**			0.8	1	0	1
26	**Lower Bound**			0	0	0	0

FIGURE 8.41

Running Solver yields the optimal solution (0, 1, 0, 1) with $Z = 9$. This means that *for all* solutions with $x_1 = 0$, the best is (0, 1, 0, 1) with $Z = 9$. This solution and its associated value of Z are given special names:

- The *incumbent* is the best binary solution found so far.
- Z^* is the value of the objective function at the incumbent solution.

Thus at this point, the incumbent solution is (0, 1, 0, 1) and $Z^* = 9$. We record these as in Figure 8.42.

	C	D	E	F	G
24	**Incumbent =**			**(0,1,0,1)**	
25		**Z*=**		**9**	

FIGURE 8.42

Sub-problem 2: Next we branch to the right and set $x_1 = 1$. We modify the relaxation by adjusting the upper and lower bounds of x_1 to 1 as in Figure 8.43.

	T	U	V	W	X	Y	Z
23				x_1	x_2	x_3	x_4
24	Upper Bound			1	1	1	1
25	Solution =			0	1	0	1
26	Lower Bound			1	0	0	0

FIGURE 8.43

Running Solver yields the optimal solution (1, 0.8, 0, 0.8) with $Z = 16.2$. This means that *for all* solutions with $x_1 = 1$, the best is (1, 0.8, 0, 0.8) with $Z = 16.2$ (i.e. $Z = 16$ is an upper bound on all binary solutions with $x_1 = 1$).

Step 3: Fathom sub-problems

In this step we determine which of the sub-problems in step 2 are worth further consideration. Those that are not deemed worthy are *fathomed*.

- **Sub-problem 1**: Here we set $x_1 = 0$ and got the solution, (0, 1, 0, 1), which is binary. Therefore, there is no reason to examine any other possible solution with $x_1 = 0$ since we have found the best one. Thus we fathom it.
- **Sub-problem 2**: Here we set $x_1 = 1$ and got a non-binary solution, but $Z = 16.2$ which is greater than $Z^* = 9$. This means that it *may* be possible to find a feasible binary solution with $x_1 = 1$ that is better than the incumbent solution. So this sub-problem is worth further consideration and we *branch* on this sub-problem.

In general, there are three reasons why we fathom a sub-problem:

1. Its value of Z is less than or equal to Z^*.
2. It has no feasible solution.
3. Its solution is binary.

We record the results of each sub-problem as in Figure 8.44. In the first box under each sub-problem we record the associated value of Z. Below that we record the results from step 3. "F3" means we fathom the sub-problem because of reason 3. "B" means we branch on the sub-problem.

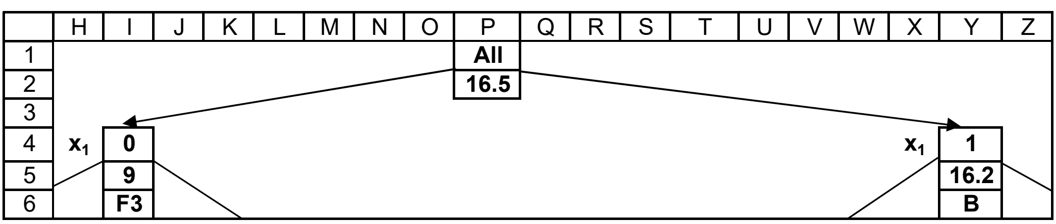


FIGURE 8.44

Step 4: Branch on newest remaining sub-problem

From the sub-problem with $x_1 = 1$, we branch to form two new sub-problems, one in which $x_2 = 0$ and one in which $x_2 = 1$. This is illustrated in Figure 8.45.

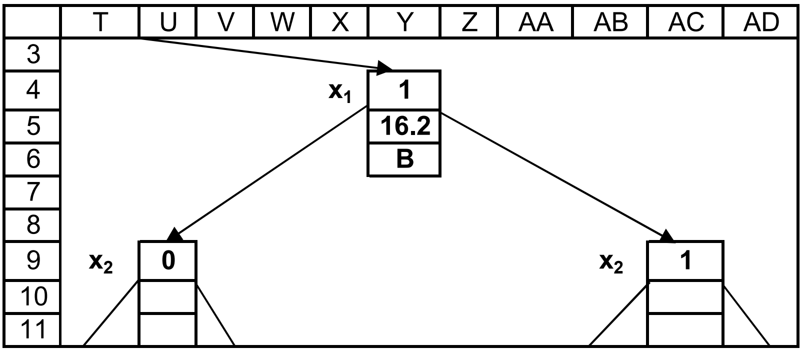


FIGURE 8.45

	T	U	V	W	X	Y	Z
23				x_1	x_2	x_3	x_4
24	**Upper Bound**			1	0	1	1
25	**Solution =**			1	0.8	0	0.8
26	**Lower Bound**			1	0	0	0

FIGURE 8.46

- **Sub-problem 3**: We branch left and set $x_2 = 0$ along with $x_1 = 1$. We modify the relaxation as in Figure 8.46. Solving this yields a solution of (1, 0, 0.8, 0) with $Z = 13.8$.

- **Sub-problem 4**: We branch right and set $x_2 = 1$ along with $x_1 = 1$. We modify the relaxation appropriately, solve, and obtain a solution of (1, 1, 0, 0.5) with $Z = 16$.

Step 5: Repeat steps 3 – 4 until there are no remaining sub-problems

We apply step 3 to determine which of sub-problems 3 or 4 should be fathomed. Neither sub-problem meets any of the three fathoming criteria. So we do not fathom either sub-problem.

We decide to branch on sub-problem 4 since it has a larger value of Z. We record the results as in Figure 8.47.

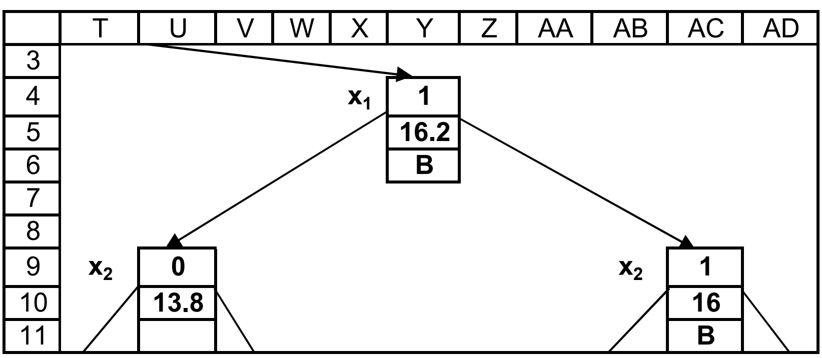


FIGURE 8.47

So we branch on $x_2 = 1$. Setting $x_3 = 0$ yields a non-binary solution with $Z = 16$. Setting $x_3 = 1$ is unfeasible, so we fathom it by criteria 2. Branching on $x_3 = 0$ and setting $x_4 = 0$ yields the solution (1, 1, 0, 0) with $Z = 14$. This is better than the incumbent solution with $Z^* = 9$. So (1, 1, 0, 0) becomes our new incumbent solution and $Z^* = 14$. We also fathom it by criteria 3. Setting $x_4 = 1$ is unfeasible so we fathom it by criteria 2. We record our results as in Figure 8.48.

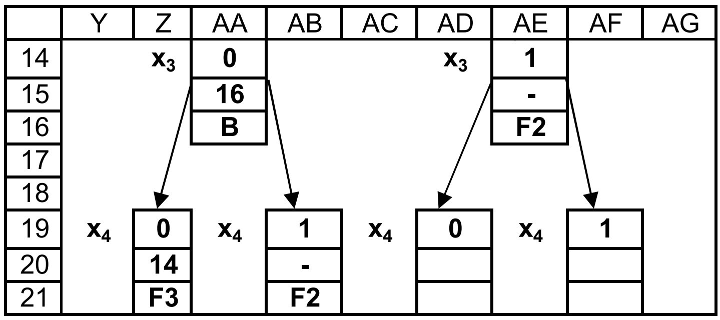


FIGURE 8.48

We still have one remaining sub-problem to consider where $x_1 = 1$ and $x_2 = 0$, which yielded $Z = 13.8$. This value of Z is less than our value of $Z^* = 14$, so we fathom it by criteria 1. We record this result as in Figure 8.49.

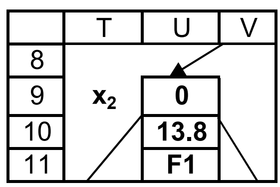


FIGURE 8.49

Thus there are no sub-problems left, so we are done. The optimal solution is (1, 1, 0, 0) with $Z^* = 14$. This means that Nathan and Laura should remodel bathroom 1 and bedroom 1. It will increase their house's value by \$14,000. This agrees with the solution Solver found with the binary constraint.

This branch and bound technique can be extended to general integer problems where each variable can take on several integer values. The basic idea is the same, except that on each branch, we have more than two options.

Exercises

8.7.1 Use the branch and bound technique to solve the following BIP:

$$\begin{array}{rrl}
\textbf{Maximize} & Z = & 9x_1 + 5x_2 + 6x_3 + 4x_4 \\
\textbf{Subject to} & & x_1 + 7x_2 - 4x_3 + 3x_4 \leq 8 \\
& & x_1 + x_3 + x_4 \leq 1 \\
& & -x_1 - x_2 + x_3 + x_4 \leq 0 \\
& & 2x_1 - x_2 + 2x_3 - x_4 \leq 0 \\
& & x_1, x_2, x_3, x_4 \text{ are binary}
\end{array}$$

8.7.2 Use the branch and bound technique to solve the following BIP:

$$\begin{array}{rrl} \textbf{Maximize} & Z = & 15x_1 + 3x_2 - 2x_3 + 4x_4 \\ \textbf{Subject to} & & 4x_1 - 2x_2 + 3x_3 + x_4 \le 8 \\ & & 2x_1 + x_3 - 3x_4 \le 1 \\ & & x_1 + 3x_2 + x_3 + x_4 \le 4 \\ & & 5x_1 - x_2 + 3x_3 + x_4 \le 10 \\ & & x_1,\ x_2,\ x_3,\ x_4 \text{ are binary} \end{array}$$

8.7.3 Consider the parameter table below for a transportation problem involving shipping truckloads of bread from bakeries to warehouses. Suppose that there is a price break in the shipping costs if the number of truckloads between a bakery and warehouse is large enough. If the number of truckloads is equal to or greater than the number in the "Breaks" chart, then the shipping costs are as shown in the "Reduced Costs" chart.

	Warehouse			Breaks Warehouse		Reduced Costs Warehouse	
Bakery	**1**	**2**	**Output**	**1**	**2**	**1**	**2**
1	764	375	50	20	25	375	350
2	390	416	95	70	30	350	375
Allocation	80	65					

Design a branch and bound technique to solve this problem. Explain how your technique works and show the results. (**Suggestion**: Use the workbook **Branch and Bound**. Each time you branch, force one of the variables to be above or below the break point and use the resulting cost in the objective function. Solve a relaxation to find a lower bound on the total cost down that branch.)

8.8 The Traveling Salesman Problem

The *traveling salesman problem* (TSP) is relatively simple to state:

> A traveling salesman needs to visit n cities. What route should he take to visit all the cities, return to where he started, and minimize the total distance traveled?

At first glance a TSP may look like a transportation problem from Section 7.3, but it turns out to be much more difficult to model and solve. There is currently no algorithm guaranteed to find a globally optimal solution for any value of n. However, there are many heuristic algorithms which can give "good" solutions, at least. In this section we use the *evolutionary algorithm* in Solver to solve a relatively simple TSP.

Example 8.8.1 (A Traveling Salesman Problem)
Consider a salesman who needs to visit 10 cities whose coordinates on an $x-y$ plane are shown in Table 8.3.

Figure 8.50 shows a map of these cities. We assume there are roads between each pair of distinct cities, the distances between cities are their Euclidean distances, and the distances are commutative (i.e. the distance from city 1 to 2 is the same as from 2 to 1). The cities along with the collection of roads is called a *graph*. The cities are called *nodes* and the roads are called *edges*. Because of the fact that there are edges between each pair of nodes, the graph is called *complete*.

TABLE 8.3

City	1	2	3	4	5	6	7	8	9	10
$\boldsymbol{x}$	22	28	22	34	10	16	4	34	23	1
$\boldsymbol{y}$	15	34	21	3	8	18	13	29	1	34

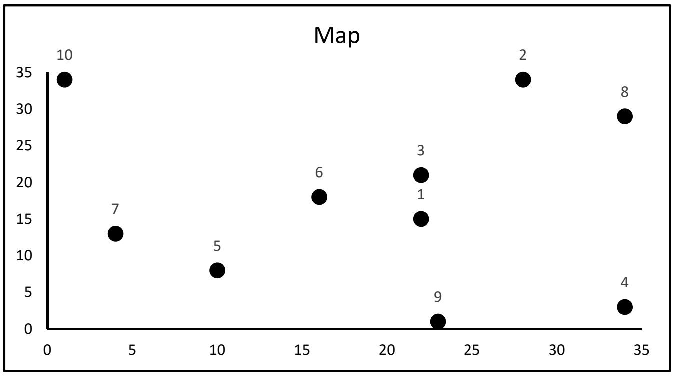


FIGURE 8.50

A solution to this TSP is a sequence of cities so that the salesman visits every city once and he ends where he began. Such a sequence is called a *tour*. One such tour is $1-2-3-4-5-6-7-8-9-10$ (where the salesman returns to city 1 after city 10). Note that all the numbers in this sequence are different (this point will become important when we use Solver). This tour is illustrated in Figure 8.51. Based on the illustration, this route does not seem very logical, so it's obviously not an optimal solution to the TSP.

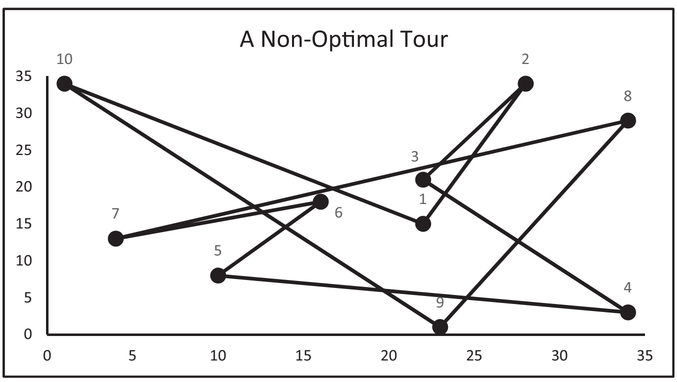


FIGURE 8.51

□

One way to model Example 8.8.1 is as a binary integer program with the decision variables

$$x_{ij} = \begin{cases} 1, & \text{if the salesman travels from city } i \text{ to city } j \\ 0, & \text{otherwise} \end{cases}$$

for $1 \le i, j \le 10$. Let d_{ij} denote the distance from city i to city j, N denote the set $\{1, \ldots, 10\}$, S denote a proper subset of N containing at least two elements, and $|S|$ denote the number of elements in S. With this notation we can model the TSP as

$$\begin{aligned}
\textbf{Minimize} \quad & \sum_{i=1}^{10}\sum_{j=1}^{10} d_{ij}x_{ij} && (8.5)\\
\textbf{Subject to} \quad & \sum_{i=1}^{10} x_{ij} = 1, \ \forall\, j \in N \\
& \sum_{j=1}^{10} x_{ij} = 1, \ \forall\, i \in N \\
& \sum_{i\in S}^{10}\sum_{j\in S}^{10} x_{ij} \le |S| - 1, \ \forall\, S \\
& x_{ij} \text{ are binary.}
\end{aligned}$$

The objective function in this model is simply the total distance traveled. Each constraint is really a set of 10 or more constraints. Here are the practical meanings of these sets of constraints:

1. Each city is traveled *to*.
2. Each city is traveled *from*.
3. Sub-tours are prohibited. For instance, the sub-tour $1 - 2 - 3$ (and returning to city 1) is not allowed to be part of the solution (see Exercise 8.8.10). These constraints are called *sub-tour eliminating constraints*.

The first two sets of constraints are fairly straightforward. The third set is quite complicated. It specifies one constraint for each proper subset of N containing at least two elements. There are $2^{10} - 2 - 10 = 1012$ such subsets. This is far too many constraints to be practical.

Instead of using this binary integer formulation of the TSP, we'll use the evolutionary search algorithm in Solver which contains a very simple tool for modeling a TSP. The evolutionary search algorithm is very complex and we will not go into the details here. Informally, the algorithm is based on ideas from evolutionary biology. The algorithm begins with a solution and then intelligently changes it a bit (or mutates it) to hopefully find a better solution. This process repeats itself until no better solution is found.

Example 8.8.2 (Solving a TSP)
To solve the TSP described in Example 8.8.1 with the evolutionary algorithm in Solver, follow these steps:

1. Rename a blank worksheet "**TSP**" and format it as in Figure 8.52. Add the rest of the coordinates shown above in Table 8.3.

	B	C	D
2	**City**	**x**	**y**
3	1	22	15

FIGURE 8.52

2. Create the graph in Figure 8.50. To add the city numbers to the graph, right-click on one of the cities and select **Add Data Labels**. Right-click a city again and select **Format Data Labels**. Check the box next to **Value From Cells** and **Select Range B3:B12**. Un-check the boxes next to **Y Value** and **Show Leader Lines**.

3. Next we need to calculate the distances between each pair of cities. Set up a table as in Figure 8.53. Copy row 4 down to row 12.

	F	G	H	I	J	K	L	M	N	O	P
1		**Distances**									
2		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
3	**1**										
4	**2**										

FIGURE 8.53

4. Add the formula in Figure 8.54 to calculate the distances. Copy this formula to the range **G3:P12**.

	G
3	=SQRT((OFFSET($B$2,$F3,1)-OFFSET($B$2,G$2,1))^2+(OFFSET($B$2,$F3,2)-OFFSET($B$2,G$2,2))^2)

FIGURE 8.54

5. To find the total length of a tour, add the formulas in Figure 8.55 (note the difference in the pattern in cell **C25**).

	B	C
14		**Distance to**
15	**Tour**	**Next City**
16	1	=OFFSET($F$2,B16,B17)
17	2	=OFFSET($F$2,B17,B18)
18	3	=OFFSET($F$2,B18,B19)
19	4	=OFFSET($F$2,B19,B20)
20	5	=OFFSET($F$2,B20,B21)
21	6	=OFFSET($F$2,B21,B22)
22	7	=OFFSET($F$2,B22,B23)
23	8	=OFFSET($F$2,B23,B24)
24	9	=OFFSET($F$2,B24,B25)
25	10	=OFFSET($F$2,B25,B16)
26	**Total =**	=SUM(C16:C25)

FIGURE 8.55

6. To visualize a tour, add the formulas in Figure 8.56. Add the range **E16:F26** to the chart in Figure 8.50 and format this new data series so the chart resembles Figure 8.51.

7. Lastly, format the **Solver** window as in Figure 8.57 and press **Solve**. It will take Solver several seconds to return a solution.

	E	F
14	**Coordinates**	
15	**x**	**y**
16	=OFFSET($C$2,B16,0)	=OFFSET($D$2,B16,0)
17	=OFFSET($C$2,B17,0)	=OFFSET($D$2,B17,0)
18	=OFFSET($C$2,B18,0)	=OFFSET($D$2,B18,0)
19	=OFFSET($C$2,B19,0)	=OFFSET($D$2,B19,0)
20	=OFFSET($C$2,B20,0)	=OFFSET($D$2,B20,0)
21	=OFFSET($C$2,B21,0)	=OFFSET($D$2,B21,0)
22	=OFFSET($C$2,B22,0)	=OFFSET($D$2,B22,0)
23	=OFFSET($C$2,B23,0)	=OFFSET($D$2,B23,0)
24	=OFFSET($C$2,B24,0)	=OFFSET($D$2,B24,0)
25	=OFFSET($C$2,B25,0)	=OFFSET($D$2,B25,0)
26	=E16	=F16

FIGURE 8.56

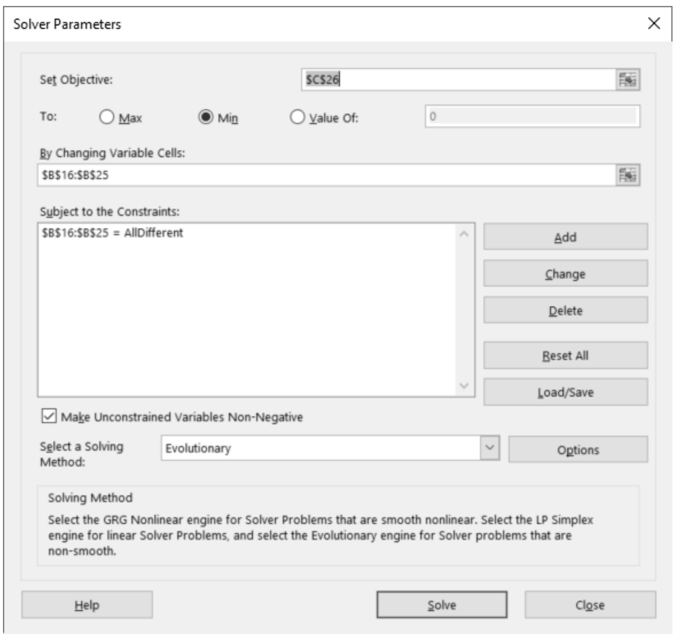


FIGURE 8.57

Solver finds the tour $8-3-6-1-4-9-5-7-10-2$ with a total distance of 134.588, which is illustrated in Figure 8.58. We do not claim that this is a globally optimal solution, but it seems reasonable. Note that we did not specify where the salesman starts. The total distance would be the same if he started in any city and followed this tour. For instance, the tour $4-9-5-7-10-2-8-3-6-1$ has the same total distance.

□

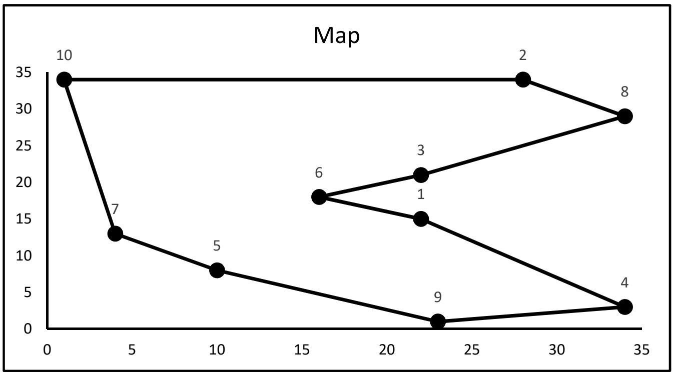


FIGURE 8.58

Exercises

8.8.1 One simple heuristic for solving a TSP is the *nearest neighbor algorithm* which says that at each step of the tour, the salesman should visit the closest city that has not already been visited. Apply this algorithm to find a solution to Example 8.8.1 starting at city 1. Compare the resulting tour to the one found by Solver.

8.8.2 One fact about an optimal solution to a TSP is that the tour should not cross itself. As an example of why this fact is true, consider the tour 1−3−6−5−7−10−2−9−8−4 from Example 8.8.1. This tour is illustrated in Figure 8.59. Show that this tour is not optimal by switching around just 3 cities to find a shorter tour.

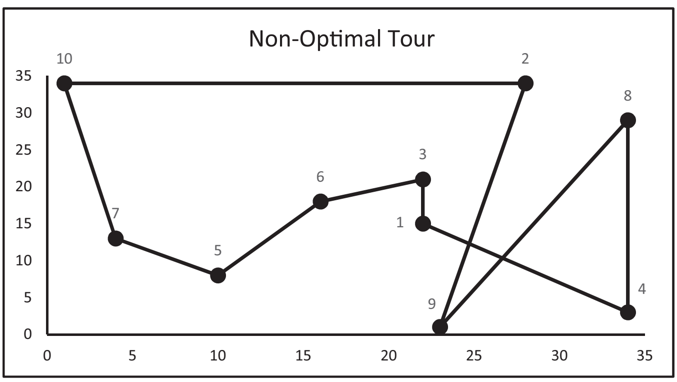


FIGURE 8.59

8.8.3 In the classical TSP, we require the salesman to end the tour where he started. Such a tour is called a *closed tour*. In an *open tour* we do not require the salesman to end where he started, simply that he visit each city. Modify the worksheet **TSP** to solve the open tour version of Example 8.8.1 where the salesman starts at city 10 (it does not matter exactly where he ends). (**Hint**: When modifying the Solver parameters, you may need to first press **Reset All** and then enter all the desired parameters.)

8.8.4 In Example 8.8.1, we assumed the graph was complete (i.e. there was a road between each pair of distinct cities). Consider the non-complete graph shown in Figure 8.60 consisting of 10 cities and 13 roads. The numbers next to each road are the times (in minutes) to travel between the respective cities. Modify the worksheet **TSP** to find a tour that starts at city 10, visits all cities, and minimizes the total time needed (the salesman does not need to end at city 10). (**Suggestion**: Use the big-M method to model the fact that there are not roads between every pair of cities.)

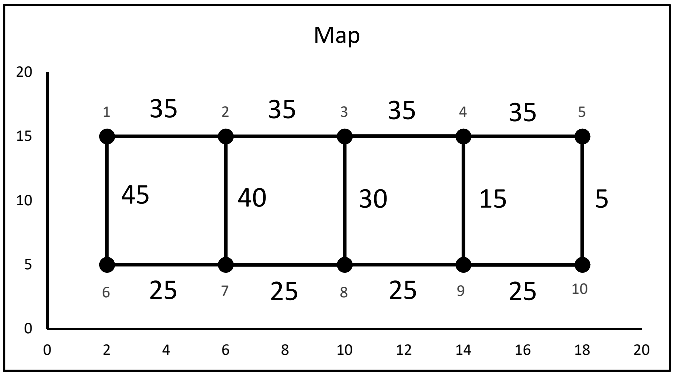


FIGURE 8.60

8.8.5 One approach to finding a globally optimal solution to a TSP, called the *brute-force algorithm*, is to simply list all possible tours, calculate the total distance of each, and then choose the one with the shortest total. In this exercise we'll think about how many tours really need to be considered when solving a TSP with 10 cities.

a. A tour is really a permutation of the set $\{1, \ldots, 10\}$. Find the number of such permutations.

b. Now consider the permutations $1-2-3-4-5-6-7-8-9-10$ and $2-3-4-5-6-7-8-9-10-1$. What is the difference between these tours? Will they have the same length? How many tours are there with these characteristics?

c. Consider the permutations $1-2-3-4-5-6-7-8-9-10$ and $1-10-9-8-7-6-5-4-3-2$. What is the difference between these tours? Will they have the same length?

d. Put parts a., b., and c. together to find the number of tours with (possibly) distinct lengths that we need to consider in the brute-force algorithm.

e. Does the brute-force algorithm seem like a reasonable solution method? Briefly explain your answer.

8.8.6 Consider a TSP with 5 cities. Follow these steps to use a version of the brute-force algorithm as described in Exercise 8.8.5 to find a globally optimal solution:

a. Design a spreadsheet that allows the user to input the coordinates of the 5 cities.

b. Create a table that contains the $4! = 24$ different tours that start at city 5.

c. Calculate the total distance of each tour.

d. Find the tour with the smallest total distance.

8.8.7 Consider the following algorithm for solving Example 8.8.1 with a simulation:

1. Assume the tour always starts at city 10.
2. Use an approach similar to that for generating a shuffled deck of cards as described in Section 6.4 to randomly choose the other 9 cities in the tour.
3. Use a table to store the results of 500 random tours.
4. Find the tour in the table with the shortest total distance.

Implement this algorithm. How well does it do compared to Solver?

8.8.8 The table below contains the approximate $x-y$ coordinates of all 15 National League MLB stadiums (data collected by Sam Otte, 2019). Find a tour of minimal distance that visits all 15 stadiums.

Stadium	**Team**	**City**	x	y
1	Giants	San Francisco, CA	2	21
2	Dodgers	Los Angeles, CA	2.5	16
3	Padres	San Diego, CA	3	15
4	Diamondbacks	Phoenix, AZ	7	12
5	Rockies	Denver, CO	12	18
6	Cardinals	St. Louis, MO	20.5	18
7	Cubs	Chicago, IL	22.5	22
8	Brewers	Milwaukee, WI	22	25
9	Reds	Cincinnati, OH	25	20
10	Pirates	Pittsburgh, PA	27.5	22
11	Braves	Atlanta, GA	25.5	12
12	Marlins	Miami, FL	28.5	5
13	Mets	New York, NY	30.5	24.5
14	Nationals	Washington, DC	29.5	19
15	Phillies	Philadelphia, PA	30	23

8.8.9 Find the driving distances between each school in your school's athletic conference. Then find a tour of minimal distance that visits each school.

8.8.10 Consider the binary integer program formulation of the TSP in Equations (8.5) and the sub-tour $1-2-3$. If this sub-tour were allowed, we would have $x_{12} = x_{23} = x_{31} = 1$ (with possibly other decision variables equaling 1 as well). Show that this would violate the third constraint with $S = \{1, 2, 3\}$.

For Further Reading

Here are several good references for more information on nonlinear programming:

- Bazarra, M. and C. Shetty, H.D. Scherali. *Nonlinear Programming: Theory and Applications*. New York: Wiley. 1993.
- Fox, William. P. "Teaching Nonlinear Programming with Minitab." COED Journal. Vol. II No. 1, January-March 1992, pages 80-84. 1992.

- Fox, William P. "Using Microcomputers in Undergraduate Nonlinear Optimization." Collegiate Microcomputer. VOL XI(3). Pages 214-218. 1993.

- Fox, William P. and Margie Witherspoon, "Single Variable Optimization When Calculus Fails: Golden Section Search Methods in Nonlinear Optimization Using MAPLE," COED. VOL XI (2), 2001, pages 50-56.

- Giordano, F. W. Fox, & S. Horton, 2013. *A First Course in Mathematical Modeling*. 5th Edition. Boston, MA.: Cengage.

- Phillips, D.T., A. Ravindran, and J. Solberg, 1976. *Operations Research*. New York, New York: John Wiley and Sons.

- Rao, S.S., 1979. *Optimization: Theory and Applications*. New Delhi, India: Wiley Eastern Limited.

- Winston, Wayne, 2002. *Introduction to Mathematical Programming: Applications and Algorithms*. 4th Edition. Belmont. CA: Duxbury Press, ITP.

A

Spreadsheet Basics

Here we will explain some of the basic terminology and tools used to build the models in this book. These explanations apply directly to Office Excel 2016, although most of them apply to other versions of Excel. Other free spreadsheet programs such as OpenOffice.org Calc (which is available for download as part of the OpenOffice.org suite at http://www.openoffice.org/) and Google Sheets can be used to build almost all of the models in this book. Many of the commands used in these free programs are similar to those in Excel, but some are significantly different.

A.1 Basic Terminology

When you first open Excel, you should get a window that looks similar to Figure A.1. (Your window will probably not look exactly like that in Figure A.1 due to the placement of toolbars and icons.)

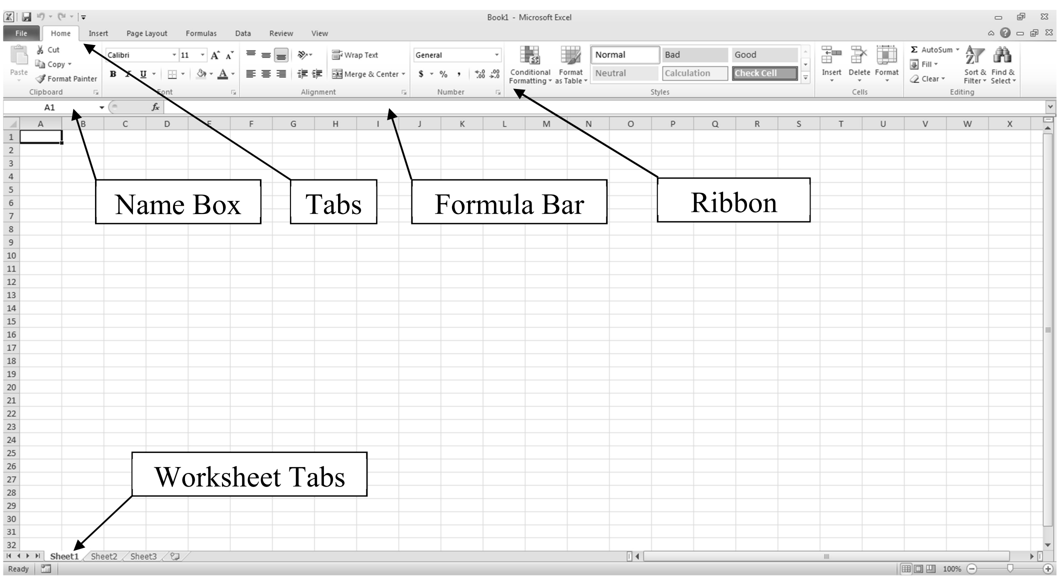


FIGURE A.1

Each rectangle, called a *cell*, is a place where data, text, or formulas can be entered. A collection of cells is called a *worksheet*. The name of the worksheet is given in the *worksheet tab* near the bottom of the worksheet. The worksheet name can be changed by right–clicking on the worksheet tab and selecting **Rename**. A collection of worksheets is called a *workbook*.

Worksheets can be added or deleted from a workbook by right–clicking on the worksheet tab and selecting either **Insert...** or **delete**. Worksheets can also be added by clicking on the plus sign next to the worksheet tabs. A worksheet tab can be moved to the left or right by left–clicking, holding, and dragging.

The name of each *column* is listed along the top of the worksheet while the number of each *row* is listed along the left–hand side of the worksheet. The width or height of a column or row can be changed by left–clicking and holding on the right or bottom and then dragging to the desired width or height. A cell is named according to its column and row position. The *selected cell* has a thicker border around it and its name is shown in the *name box*. The selected cell can be changed using the arrow keys or by clicking on another cell.

A two–dimensional *range* of cells can be selected by left–clicking on the cell in the upper left–hand corner of the range, holding, and dragging to the cell in the lower right–hand corner of the range. This highlights these cells indicating they are all selected. Ranges are referred to by the cells in the upper left–hand and lower right–hand corners in the form **(Upper Left):(Lower Right)**. The *formula bar* displays the contents of the selected cell.

A.2 Entering Text, Data, and Formulas

Text, data, and formulas are easily entered by selecting the desired cell, typing the desired contents, and pressing **Enter**. To practice doing this, format a blank worksheet as in Figure A.2. This worksheet contains two columns of data named "**x**" and "**y**" and a third column named "**z**" which we will define later.

	A	B	C
1	**x**	**y**	**z**
2	1	5	
3	2	8	

FIGURE A.2

Notice that when you press enter, the selected cell moves to the cell directly under the previous one. By default text is left–justified. The text in row 1 can be changed to bold and centered by selecting the range **A1:A3**, and then clicking on the bold font icon and then the center icon located in the toolbar.

Now suppose we want to define the quantity z to be $z + y$. We can easily do this by entering the formula in Figure A.3. Every formula begins with an equal sign. This formula can be entered by typing it as in the figure and then pressing **Enter**, or you can type $=$, click on cell **A2**, type $+$, click on cell **B2**, and then press **Enter**.

	C
2	=A2+B2

FIGURE A.3

Once the formula is entered, select cell **C2** and click in the formula bar. Notice how different colored boxes are put around cells **A2** and **B2** and that the **A2** and **B2** in the formula are changed to the corresponding colors. This feature simplifies the process of debugging formulas.

To calculate the second value of z, we could type the formula **=A3+B3** in cell **C3**, but there is an easier way. Select cell **C2**, left–click and hold on the dark square in the lower right–hand corner of the cell. Then drag the box down one row and release. The results are shown in Figure A.4. This is exactly what we want.

	C
1	**z**
2	=A2+B2
3	=A3+B3

FIGURE A.4

A.2.1 Understanding Cell References

To understand why the formula in cell **C2** copied down to cell **C3** in this way, we need to understand what we mean when we reference cells in formulas. The formula in cell **C2** should not be interpreted as "add cell **A2** to cell **B2**." Rather, it should be interpreted as "add the cell two columns to the left and in the same row to the cell one column to the left and in the same row." In other words, these cell references are *relative*. When this formula is copied down one row, the cell "two columns to the left and in the same row" is now **A3** and the cell "one column to the left and in the same row" is now **B3**.

Now, change the formula in cell **C2** to that shown in Figure A.5. The **$**'s can be entered manually or you can delete the contents of **C2**, then type =, click on cell **A2**, press the **F4** key, type +, click on cell **B2**, press the **F4** key, and then press **Enter**.

	C
2	=$A$2+$B$2

FIGURE A.5

Copy the formula in cell **C2** down to **C3**. The results are shown in Figure A.6. Notice that the formula did not change. This is because the **$**'s "fix" the row and column reference. So the formula in **C2** really does mean "add cell **A2** to cell **B2**." When we copy it down, the meaning does not change.

	C
1	**z**
2	=$A$2+$B$2
3	=$A$2+$B$2

FIGURE A.6

Now, change the formula in cell **C2** to that shown in Figure A.7. The **$**'s can be manually entered or they can entered by selecting the cells and pressing **F4** two or three times, similar to above.

Copy the formula in cell **C2** down one row and to the right one column. The results are shown in Figure A.8.

We get these results because the **$** in **A$2** "fixes" the row at 2, but the column is still relative. When we copy down, this row does not change, but when we copy to the right, the

	C
1	**z**
2	=A$2+$B2

FIGURE A.7

	C	D
1	**z**	
2	=A$2+$B2	=B$2+$B2
3	=A$2+$B3	

FIGURE A.8

column changes to **B**. Likewise, the **$** in **$B2** fixes the column at **B**, but the row is still relative. When we copy down, this row changes, but when we copy to the right, the column does not change.

A.2.2 Formatting Cells

The formats of a cell or range can be easily changed by first selecting the cell or range and then right–clicking within the cell or range. Selecting **Format Cells...** yields the window shown in Figure A.9.

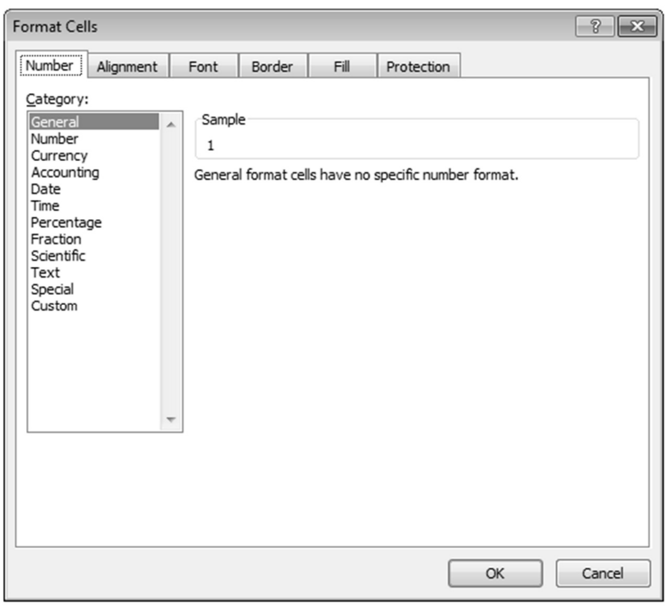


FIGURE A.9

Several of these tabs are useful for building the models in this book:

- **Number** – The number tab allows you to change the way numbers are displayed. For instance, selecting **Number** under **Category:** allows you to, among other things, set the number of displayed decimal places.

- **Font** – The font tab allows you to change the font, font style, and size of text. It also allows you to add effects such as superscript or subscript.
- **Border** – The border tab allows you to change the border around and between cells.
- **Patterns** – The patterns tab allows you to change the background color and pattern of cells.

A.3 Creating Charts and Graphs

To illustrate the process of creating charts and graphs, format a blank worksheet as in Figure A.10.

	A	B
1	**x**	**y**
2	1	2
3	2	5
4	3	9
5	4	12
6	5	13

FIGURE A.10

To create a simple plot of **y** vs. **x**, follow these steps:

1. Select the range **A1:B6** and click on the **Insert** tab. In the **Charts** section of the ribbon, select a **Scatter** plot as shown in Figure A.11. The resulting chart is shown in Figure A.12.

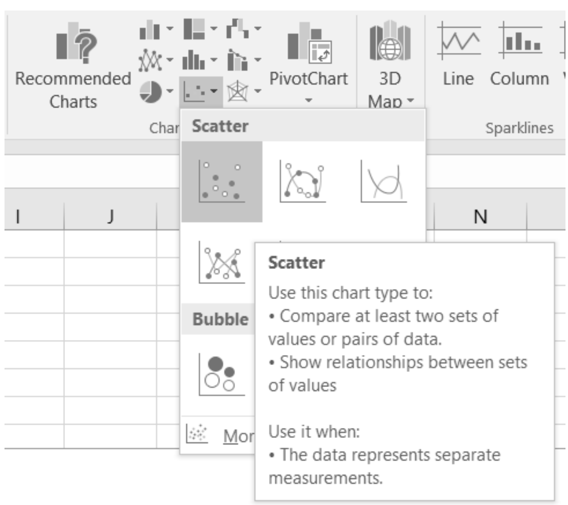


FIGURE A.11

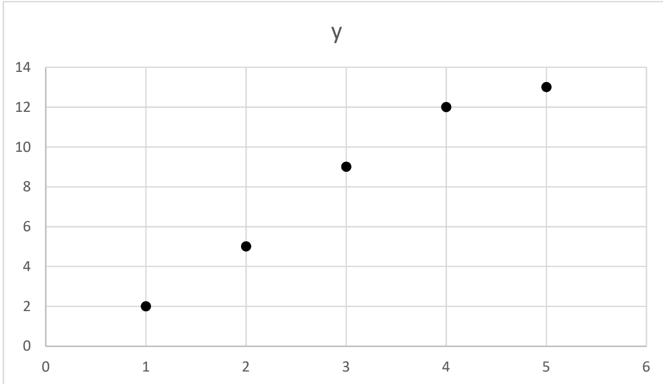


FIGURE A.12

2. Left-clicking anywhere on the chart causes a large plus sign to appear near the upper right-hand corner of the chart. Clicking on this plus sign allows you to select and deselect several chart elements. For instance, selecting **Axis Titles** and **Chart Title** causes textboxes to appear on the chart where you can enter desired titles. Deselecting **Gridlines** gets rid of the horizontal and vertical gridlines. After entering some titles, your graph can resemble Figure A.13. It's a good habit to always give charts a meaningful title. Other aspects of the chart, such as color and line thickness, can be modified by double-clicking anywhere on the chart.

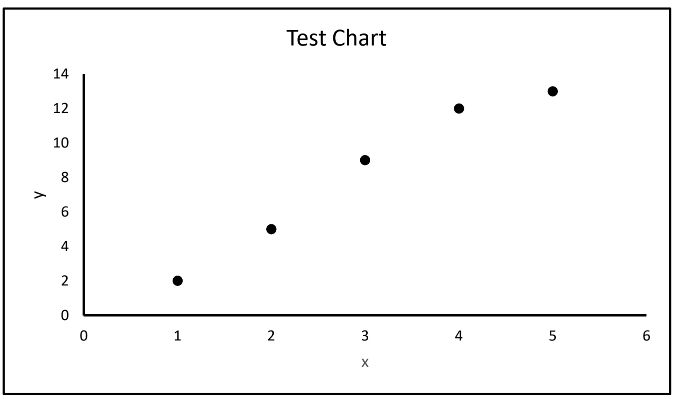


FIGURE A.13

3. The format of the x– and y–axes can be changed by double–clicking on an axis.

A.3.1 Adding Data to a Chart

Suppose we wanted to graph the data in Figure A.14 on the same $x - y$ plane in Figure A.13. Follow these steps:

1. Add the data in Figure A.14 to the worksheet.

2. Right-click anywhere on the chart and select **Select Data...**

	D	E
1	**x**	**y**
2	2	3
3	4	6
4	6	9

FIGURE A.14

3. Press the **Add** button. Format the resulting window as in Figure A.15. Press **OK** twice.

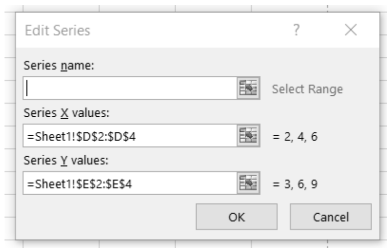


FIGURE A.15

4. The resulting graph is shown in Figure A.16. You can change options of each series, such as the color and shape of the marker, by double-clicking on one of the markers and using the window that appears on the right-hand side of the screen.

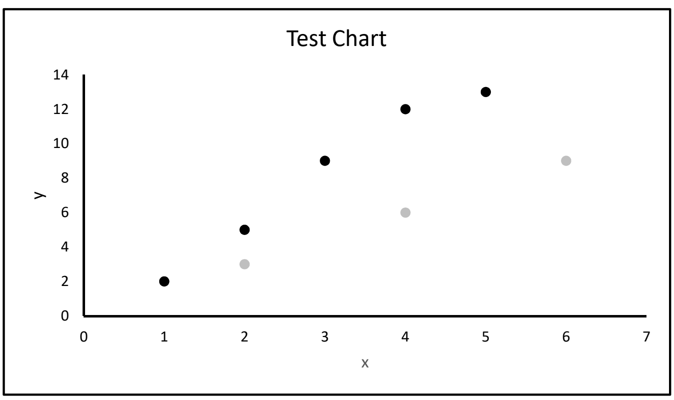


FIGURE A.16

5. When graphing multiple sets of data, we typically want a legend to identify each set. We can easily add a legend by clicking anywhere on the chart, clicking on the plus sign that appears near the upper right-hand corner of the chart, and selecting **legend**. The result is shown in Figure A.17. The name of a series can be changed by left-clicking on the chart, selecting **Select Data...**, selecting the desired series, clicking **Edit**, entering the desired name under **Series name**, and clicking **OK**.

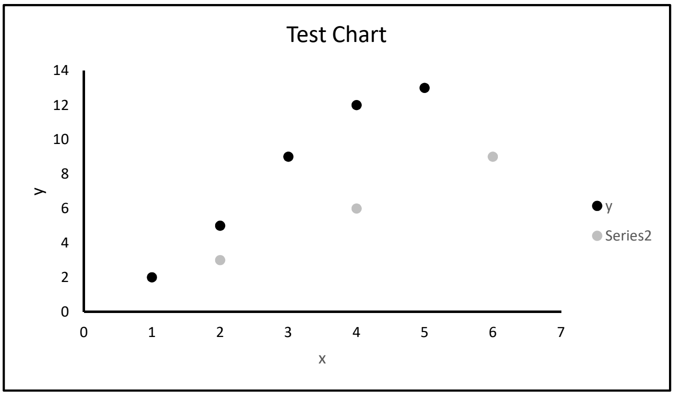


FIGURE A.17

A.3.2 Graphing Functions

Excel does not have a built–in tool for graphing functions, but we can easily create an $x - y$ table and then plot the points. For example, to graph the function $f(x) = x^2$ over the interval $[-2, 2]$, format a blank worksheet as in Figure A.18. Copy row 3 down to row 42.

	A	B
1	**x**	**y**
2	-2	=A2^2
3	=A2+0.1	=A3^2

FIGURE A.18

Select columns **A** and **B** by left–clicking and holding on the column **A** header and then dragging to column **B**. Insert a chart of the type **Scatter with Smooth Lines**. Once the chart is created, deselect the gridlines.

Now we'll adjust the x and y axes. Double-click on the x-axis. Under **Axis Options** set the **Minimum** and **Maximum** to -2 and $+2$. Do the same for the y-axis, setting the min and max to 0 and 4. After changing the chart title, the result should look like Figure A.19.

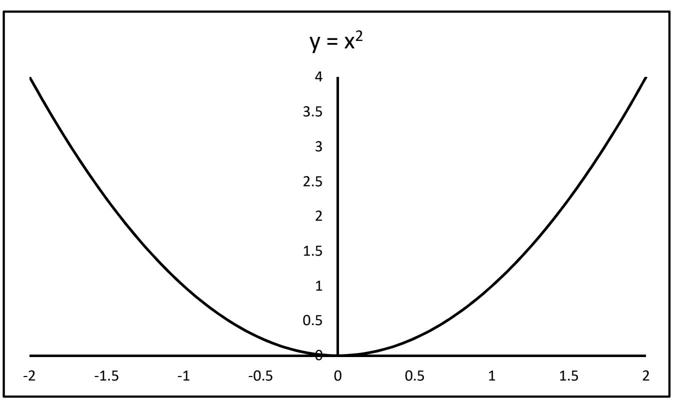


FIGURE A.19

A.4 Scroll Bars

Scroll bars allow us to change the value of a cell with a graphical interface. This allows us to dynamically change the values of parameters within a model and analyze the results. In a blank worksheet, select **View** → **Toolbars** → **Control Toolbox**. A window similar to that in Figure A.20 should appear.

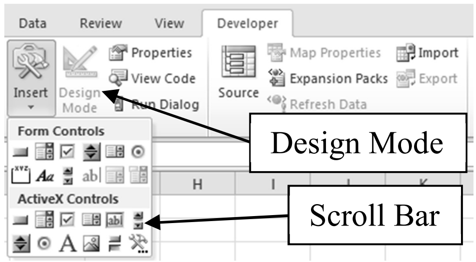


FIGURE A.20

For our purposes, there are only two important buttons in this window: **Design Mode** and **Scroll Bar**. When you press the scroll bar button, the cursor changes to a small cross. Use this to draw a long, skinny rectangle. Right–click on the resulting scroll bar and select **Properties**. A window similar to that in Figure A.21 should appear.

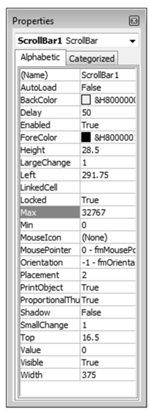


FIGURE A.21

There are three important properties we need to change. The **LinkedCell** is the cell whose value we want to change. Set this to **A1** by typing **A1** in the box next to it. The **Min** and **Max** are the minimum and maximum values of the cell. Set these to 0 and 1,000, respectively. Close the properties window and click on the **Design Mode** button. The scroll bar is now ready to use. Move the slider on the scroll bar back and forth and note that the number in cell **A1** changes between 0 and 1,000 in increments of 1. The scroll bar properties can be changed by clicking on the **Design Mode** button and right–clicking on the scroll bar.

In most instances, we may want the value of a cell to change in increments other than 1. This can be accomplished using a formula that references the linked cell. For instance, enter the formula shown in Figure A.22. Move the slider back and forth and note that the number in cell **A2** changes between 0 and 100 in increments of 0.1.

	A
2	=A1/10

FIGURE A.22

(**Note**: There is a somewhat easier way to create scroll bars using the **Forms** toolbar. However, these scroll bars do not work as well with graphs. With the method described above, if the scroll bar changes a value on a graph, the graph changes in a continuous manner as the slider is moved back and forth. If the **Forms** toolbar is used, the graph will not change until you release the mouse button.)

A.5 Array Formulas

Excel can perform simple matrix operations such as addition, multiplication, and finding inverses. For example, if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$, to compute $C = A + B$, format a blank worksheet as in Figure A.23. To center the text "A" between cells **A1** and **B1**, select the range **A1:B1** and then press the **Merge and Center** icon in the toolbar. (If this icon is not available, select **Format** → **Cells...** → **Alignment**, check the box next to **Merge cells**, and press **OK**.)

	A	B	C	D	E	F	G	H
1	**A**			**B**			**C**	
2	1	2		5	6			
3	3	4		7	8			

FIGURE A.23

Next, select the range **G2:H3**, type **=A2:B3+D2:E3**, and press the combination of keys **Ctrl-Shift-Enter** (this combination tells Excel to compute an array formula). The results are shown in Figure A.24. Notice that when you select any cell in the range **G2:H3**, the formula is in curly brackets, {...}. This indicates that an array formula has been entered.

	G	H
1	**C**	
2	6	8
3	10	12

FIGURE A.24

Index