Mathematical Modeling

Fourth Edition

Mark M. Meerschaert

Mathematical Modeling
Fourth Edition

Mathematical
Modeling
Fourth Edition

Mark M. Meerschaert

Department of Statistics and Probability
Michigan State University
A430 Wells Hall
East Lansing, MI 48824-1027 USA

For information on all Academic Press publications
visit our web site at *store.elsevier.com*

Preface

Mathematical modeling is the link between mathematics and the rest of the world. You ask a question. You think a bit, and then you refine the question, phrasing it in precise mathematical terms. Once the question becomes a mathematics question, you use mathematics to find an answer. Then finally (and this is the part that too many people forget), you have to reverse the process, translating the mathematical solution back into a comprehensible, no–nonsense answer to the original question. Some people are fluent in English, and some people are fluent in calculus. We have plenty of each. We need more people who are fluent in both languages and are willing and able to translate. These are the people who will be influential in solving the problems of the future.

This text, which is intended to serve as a general introduction to the area of mathematical modeling, is aimed at advanced undergraduate or beginning graduate students in mathematics and closely related fields. Formal prerequisites consist of the usual freshman–sophomore sequence in mathematics, including one–variable calculus, multivariable calculus, linear algebra, and differential equations. Prior exposure to computing and probability and statistics is useful, but is not required.

Unlike some textbooks that focus on one kind of mathematical model, this book covers the broad spectrum of modeling problems, from optimization to dynamical systems to stochastic processes. Unlike some other textbooks that assume knowledge of only a semester of calculus, this book challenges students to use *all* of the mathematics they know (because that is what it takes to solve real problems).

The overwhelming majority of mathematical models fall into one of three categories: optimization models; dynamic models; and probability models. The type of model used in a real application might be dictated by the problem at hand, but more often, it is a matter of choice. In many instances, more than one type of model will be used. For example, a large Monte Carlo simulation model may be used in conjunction with a smaller, more tractable deterministic dynamic model based on expected values.

This book is organized into three parts, corresponding to the three main categories of mathematical models. We begin with optimization models. A five-step method for mathematical modeling is introduced in Section 1 of Chapter 1, in the context of one–variable optimization problems. The remainder of the first chapter is an introduction to sensitivity analysis and robustness. These

fundamentals of mathematical modeling are used in a consistent way throughout the rest of the book. Exercises at the end of each chapter require students to master them as well. Chapter 2, on multivariable optimization, introduces decision variables, feasible and optimal solutions, and constraints. A review of the method of Lagrange multipliers is provided for the benefit of those students who were not exposed to this important technique in multivariable calculus. In the section on sensitivity analysis for problems with constraints, we learn that Lagrange multipliers represent shadow prices (some authors call them dual variables). This sets the stage for our discussion of linear programming later in Chapter 3. At the end of Chapter 3 is a section on discrete optimization that was added in the second edition. Here we give a practical introduction to integer programming using the branch–and–bound method. We also explore the connection between linear and integer programming problems, which allows an earlier introduction to the important issue of discrete versus continuous models. Chapter 3 covers some important computational techniques, including Newton's method in one and several variables, and linear and integer programming.

In the next part of the book, on dynamic models, students are introduced to the concepts of state and equilibrium. Later discussions of state space, state variables, and equilibrium for stochastic processes are intimately connected to what is done here. Nonlinear dynamical systems in both discrete and continuous time are covered. There is very little emphasis on exact analytical solutions in this part of the book, since most of these models admit no analytic solution. At the end of Chapter 6 is a section on chaos and fractals that was added in the second edition. We use both analytic and simulation methods to explore the behavior of discrete and continuous dynamic models, to understand how they can become chaotic under certain conditions. This section provides a practical and accessible introduction to the subject. Students gain experience with sensitive dependence to initial conditions, period doubling, and strange attractors that are fractal sets. Most important, these mathematical curiosities emerge from the study of real–world problems.

Finally, in the last part of the book, we introduce probability models. No prior knowledge of probability is assumed. Instead we build upon the material in the first two parts of the book, to introduce probability in a natural and intuitive way as it relates to real–world problems. Chapter 7 introduces the basic notions of random variables, probability distributions, the strong law of large numbers, and the central limit theorem. At the end of Chapter 7, Introduction to Probability Models, is a section on diffusion, which was added in the third edition. Here we give a gentle introduction to partial differential equations by focusing on the diffusion equation. We provide a simple derivation of the point source solution to this partial differential equation, using Fourier transforms, to arrive at the normal density. Then we connect the diffusion model to the central limit theorem introduced in the previous Section 7.3, Introduction to Statistics. This new section on diffusion grew out of a class taught at the University of Nevada for beginning graduate students in the earth sciences. The applications are to contaminant migration in the atmosphere and ground water. Chapter 8 covers the basic models of stochastic processes, including Markov chains, Markov pro-

cesses, and linear regression. At the end of Chapter 8, Stochastic Models, a new section on time series was added in the third edition. This section also serves as an introduction to multivariate regression models with more than one predictor. As a natural follow-up to the discussion in Section 8.3, Linear Regression, the new section on time series introduces the important idea of correlation. It also shows how to recognize correlated variables and include the dependence structure in a time series model. The discussion is focused on autoregressive models, since these are the most generally useful time series models. They are also the most convenient, in that they can be handled using widely available linear regression software. For the benefit of students with access to a statistical package, this section illustrates the proper application and interpretation of advanced methods including autocorrelation plots and sequential sums of squares. However, the entire section can also be covered using only a basic implementation of regression that allows multiple predictors and outputs the two basic measures: R^2 and the residual standard deviation s. This can all be done with a good spreadsheet or hand calculator. Chapter 9 treats simulation methods for stochastic models. The Monte Carlo method is introduced, and Markov property is applied to create efficient simulation algorithms. Analytic simulation methods are also explored, and compared to the Monte Carlo method. In the fourth edition, two new sections were added to the end of Chapter 9. The first new section covers particle tracking methods, for solving partial differential equations via Monte carlo simulation of the underlying stochastic process. The final section of the book introduces fractional calculus in the context of anomalous diffusion. The fractional diffusion equation is solved by particle tracking, and applied to a problem in ground water pollution. This section ties together the concepts of fractals, fractional derivatives, and probability distributions with heavy tails.

Each chapter in this book is followed by a set of challenging exercises. These exercises require significant effort, as well as a certain amount of creativity, on the part of the student. I did not invent the problems in this book. They are real problems. They were not designed to illustrate the use of any particular mathematical technique. Quite the opposite. We will occasionally go over some new mathematical techniques in this book *because the problem demands it.* I was determined that there would be no place in this book where a student could look up and ask, "What is all of this for?" Although typically oversimplified or grossly unrealistic, story problems embody the fundamental challenge in applying mathematics to solve real problems. For most students, story problems present plenty of challenge. This book teaches students how to solve story problems. There is a general method that can be applied successfully by any reasonably capable student to solve any story problem. It appears in Chapter 1, Section 1. This same general method is applied to problems of all kinds throughout the text.

Following the exercises in each chapter is a list of suggestions for further reading. This list includes references to a number of UMAP modules in applied mathematics that are relevant to the material in the chapter. UMAP modules can provide interesting supplements to the material in the text, or extra credit

projects. All of the UMAP modules are available at a nominal cost from the Consortium for Mathematics and Its Applications (`www.comap.com`).

One of the major themes of this book is the use of appropriate technology for solving mathematical problems. Computer algebra systems, graphics, and numerical methods all have their place in mathematics. Many students have not had an adequate introduction to these tools. In this course we introduce modern technology in context. Students are motivated to learn because the new technology provides a more convenient way to solve real–world problems. Computer algebra systems and 2–D graphics are useful throughout the course. Some 3–D graphics are used in Chapters 2 and 3 in the sections on multivariable optimization. Students who have already been introduced to 3–D graphics should be encouraged to use what they know. Numerical methods covered in the text include, among others, Newton's method, linear programming, the Euler method, linear regression, and Monte Carlo simulation.

The text contains numerous computer–generated graphs, along with instruction on the appropriate use of graphing utilities in mathematics. Computer algebra systems are used extensively in those chapters where significant algebraic calculation is required. The text includes computer output from the computer algebra systems Maple and Mathematica in Chapters 2, 4, 5, and 8. The chapters on computational techniques (Chapters 3, 6, and 9) discuss the appropriate use of numerical algorithms to solve problems that admit no analytic solution. Sections 3.3 and 3.4 on linear-integer programming include computer output from the popular linear programming package LINDO. Sections 8.3 and 8.4 on linear regression and time series include output from the commonly used statistical package Minitab.

Students need to be provided with access to appropriate technology in order to take full advantage of this textbook. We have tried to make it easy for instructors to use this textbook at their own institution, whatever their situation. Some will have the means to provide students with access to sophisticated computing facilities, while others will have to make do with less. The bare necessities include: (1) a software utility to draw 2–D graphs; and (2) a machine on which students can execute a few simple numerical algorithms. All of this can be done, for example, with a computer spreadsheet program or a programmable graphics calculator. The ideal situation would be to provide all students access to a good computer algebra system, a linear programming package, and a statistical computing package. The following is a partial list of appropriate software packages that can be used in conjunction with this textbook.

Computer Algebra Systems:

- Derive, Chartwell-Yorke Ltd., `www.chartwellyorke.com/derive`

- Maple, Waterloo Maple, Inc., `www.maplesoft.com`

- Mathcad, Parametric Technology Corp., `www.ptc.com/products/mathcad`

- Mathematica, Wolfram Research, Inc., `www.wolfram.com/mathematica`

- MATLAB, The MathWorks, Inc., `www.mathworks.com/products/matlab`

- Maxima, free download, `maxima.sourceforge.net`

Statistical Packages:

- Minitab, Minitab, Inc., `www.minitab.com`

- SAS, SAS Institute, Inc., `www.sas.com`

- SPSS, IBM Corp., `www.ibm.com/software/analytics/spss`

- S-PLUS, TIBCO Corp., `spotfire.tibco.com`

- R, R Foundation for Statistical Computing, free download, `www.r-project.org`

Linear Programming Packages:

- LINDO, LINDO Systems, Inc., `www.lindo.com`

- MPL, Maximal Software, Inc., `www.maximal-usa.com`

- AMPL, AMPL Optimization, LLC, `www.ampl.com`

- GAMS, GAMS Development Corp., `www.gams.com`

The numerical algorithms in the text are presented in the form of pseudocode. Some instructors will prefer to have students implement the algorithms on their own. On the other hand, if students are not going to be required to program, we want to make it easy for instructors to provide them with appropriate software. All of the algorithms in the text have been implemented on a variety of computer platforms that can be made available to users of this textbook at no additional cost. If you are interested in obtaining a copy, please contact the author, or go to `www.stt.msu.edu/users/mcubed/modeling.html` where you can download these implementations. Also, if you are willing to share your own implementation with other instructors and students, please send us a copy. With your permission, we will make copies available to others at no charge.

A complete and detailed solutions manual for instructors is available from the author or the publisher, for instructors who adopt the text for classroom use. Computer implementations of the algorithms used in the text can be downloaded for a variety of platforms, along with the computer files used to produce all of the graphics and computer outputs included in the text. These downloads are all available at `www.stt.msu.edu/users/mcubed/modeling.html`

The response to the first three editions of the text has been gratifying. The best part of this job is interacting with students and instructors who use this book. Please feel free to contact me with any comments or suggestions.

Mark M. Meerschaert
Department of Statistics and Probability
Michigan State University
C430 Wells Hall
East Lansing, MI 48824-1027 USA

Phone: (517) 353-8881
Fax: (517) 432-1405
Email: mcubed@stt.msu.edu
Web: www.stt.msu.edu/users/mcubed

Chapter 1

ONE VARIABLE OPTIMIZATION

Problems in optimization are the most common applications of mathematics. Whatever the activity in which we are engaged, we want to maximize the good that we do and minimize the unfortunate consequences or costs. Business managers attempt to control variables in order to maximize profit or to achieve a desired goal for production and delivery at a minimum cost. Managers of renewable resources such as fisheries and forests try to control harvest rates in order to maximize long–term yield. Government agencies set standards to minimize the environmental costs of producing consumer goods. Computer system managers try to maximize throughput and minimize delays. Farmers space their plantings to maximize yield. Physicians regulate medications to minimize harmful side effects. What all of these applications and many more have in common is a particular mathematical structure. One or more variables can be controlled to produce the best outcome in some other variable, subject in most cases to a variety of practical constraints on the control variables. Optimization models are designed to determine the values of the control variables which lead to the optimal outcome, given the constraints of the problem.

We begin our discussion of optimization models at a place where most students will already have some practical experience. One–variable optimization problems, sometimes called maximum–minimum problems, are typically discussed in first–semester calculus. A wide variety of practical applications can be handled using just these techniques. The purpose of this chapter, aside from a review of these basic techniques, is to introduce the fundamentals of mathematical modeling in a familiar setting.

1.1   The five-step Method

In this section we outline a general procedure that can be used to solve problems using mathematical modeling. We will illustrate this procedure, called the *five-*

step method, by using it to solve a one–variable maximum–minimum problem typical of those encountered by most students in the first semester of calculus.

Example 1.1. A pig weighing 200 pounds gains 5 pounds per day and costs 45 cents a day to keep. The market price for pigs is 65 cents per pound, but is falling 1 cent per day. When should the pig be sold?

The mathematical modeling approach to problem solving consists of five steps:

1. Ask the question.

2. Select the modeling approach.

3. Formulate the model.

4. Solve the model.

5. Answer the question.

The first step is to ask a question. The question must be phrased in mathematical terms, and it often requires a good deal of work to do this. In the process we are required to make a number of assumptions or suppositions about the way things really are. We should not be afraid to make a guess at this stage. We can always come back and make a better guess later on. Before we can ask a question in mathematical terms we need to define our terms. Go through the problem and make a list of variables. Include appropriate units. Next make a list of assumptions about these variables. Include any relations between variables (equations and inequalities) that are known or assumed. Having done all of this, we are ready to ask a question. Write down in explicit mathematical language the objective of this problem. Notice that the preliminary steps of listing variables, units, equations and inequalities, and other assumptions are really a part of the question. They frame the question.

In Example 1.1 the weight w of the pig (in lbs), the number of days t until we sell the pig, the cost C of keeping the pig t days (in dollars), the market price p for pigs ($/lb), the revenue R obtained when we sell the pig ($), and our resulting net profit P ($) are all variables. There are other numerical quantities involved in the problem, such as the initial weight of the pig (200 lbs). However, these are not variables. It is important at this stage to separate variables from those quantities that will remain constant.

Next we need to list our assumptions about the variables identified in the first stage of step 1. In the process we will take into account the effect of the constants in the problem. The weight of the pig starts at 200 lbs and goes up by 5 lbs/day so we have

$$(w \text{ lbs}) = (200 \text{ lbs}) + \left(\frac{5 \text{ lbs}}{\text{day}}\right)(t \text{ days}).$$

Variables: $t =$ time (days)
$w =$ weight of pig (lbs)
$p =$ price for pigs (\$/lb)
$C =$ cost of keeping pig t days (\$)
$R =$ revenue obtained by selling pig (\$)
$P =$ profit from sale of pig (\$)

Assumptions: $w = 200 + 5t$
$p = 0.65 - 0.01t$
$C = 0.45t$
$R = p \cdot w$
$P = R - C$
$t \geq 0$

Objective: Maximize P

Figure 1.1: Results of step 1 of the pig problem.

Notice that we have included units as a check that our equation makes sense. The other assumptions inherent in our problem are as follows:

$$\left(\frac{p \text{ dollars}}{\text{lb}}\right) = \left(\frac{0.65 \text{ dollars}}{\text{lb}}\right) - \left(\frac{0.01 \text{ dollars}}{\text{lb} \cdot \text{day}}\right)(t \text{ days})$$

$$(C \text{ dollars}) = \left(\frac{0.45 \text{ dollars}}{\text{day}}\right)(t \text{ days})$$

$$(R \text{ dollars}) = \left(\frac{p \text{ dollars}}{\text{lb}}\right)(w \text{ lbs})$$

$$(P \text{ dollars}) = (R \text{ dollars}) - (C \text{ dollars})$$

We are also assuming that $t \geq 0$. Our objective in this problem is to maximize our net profit, P dollars. Figure 1.1 summarizes the results of step 1, in a form convenient for later reference.

The three stages of step 1 (variables, assumptions, and objective) need not be completed in any particular order. For example, it is often useful to determine the objective early in step 1. In Example 1.1, it may not be readily apparent that R and C are variables until we have defined our objective, P, and we recall that $P = R - C$. One way to check that step 1 is complete is to see whether our objective P relates all the way back to the variable t. The best general advice about step 1 is to *do something*. Start by writing down whatever is immediately apparent (e.g., some of the variables can be found simply by reading over the problem and looking for nouns), and the rest of the pieces will probably fall into place.

Step 2 is to select the modeling approach. Now that we have a problem stated in mathematical language, we need to select a mathematical approach to use to get an answer. Many types of problems can be stated in a standard

form for which an effective general solution procedure exists. Most research in applied mathematics consists of identifying these general categories of problems and inventing efficient ways to solve them. There is a considerable body of literature in this area, and many new advances continue to be made. Few, if any, students in this course will have the experience and familiarity with the literature to make a good selection for the modeling approach. In this book, with rare exceptions, problems will specify the modeling approach to be used. Our example problem will be modeled as a one–variable optimization problem, or maximum–minimum problem.

We outline the modeling approach we have selected. For complete details we refer the reader to any introductory calculus textbook.

> We are given a real–valued function $y = f(x)$ defined on a subset S of the real line. There is a theorem that states that if f attains its maximum or minimum at an interior point $x \in S$, then $f'(x) = 0$, assuming that f is differentiable at x. This allows us to rule out any interior point $x \in S$ at which $f'(x) \neq 0$ as a candidate for max–min. This procedure works well as long as there are not too many exceptional points.

Step 3 is to formulate the model. We need to take the question exhibited in step 1 and reformulate it in the standard form selected in step 2, so that we can apply the standard general solution procedure. It is often convenient to change variable names if we will refer to a modeling approach that has been described using specific variable names, as is the case here. We write

$$P = R - C$$
$$= p \cdot w - 0.45t$$
$$= (0.65 - 0.01t)(200 + 5t) - 0.45t.$$

Let $y = P$ be the quantity we wish to maximize and $x = t$ the independent variable. Our problem now is to maximize

$$y = f(x)$$
$$= (0.65 - 0.01x)(200 + 5x) - 0.45x \qquad (1.1)$$

over the set $S = \{x : x \geq 0\}$.

Step 4 is to solve the model, using the standard solution procedure identified in step 2. In our example we want to find the maximum of the function $y = f(x)$ defined by Eq. (1.1) over the set $x \geq 0$. Figure 1.2 shows a graph of the function $f(x)$. Since f is quadratic in x, the graph is a parabola. We compute that

$$f'(x) = \frac{(8 - x)}{10},$$

so that $f'(x) = 0$ at the point $x = 8$. Since f is increasing on the interval $(-\infty, 8)$ and decreasing on $(8, \infty)$, the point $x = 8$ is the global maximum. At
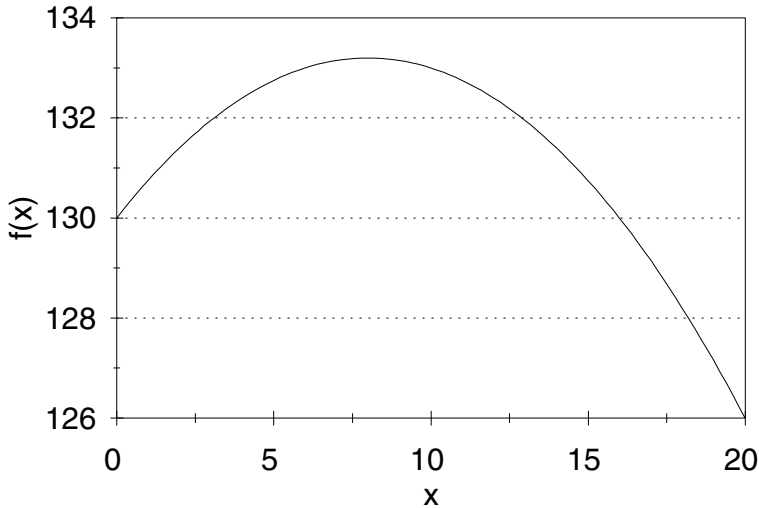
Figure 1.2: Graph of net profit $f(x) = (0.65 - 0.01x)(200 + 5x) - 0.45x$ versus time to sell x for the pig problem.

this point we have $y = f(8) = 133.20$. Since the point $(x, y) = (8, 133.20)$ is the global maximum of f over the entire real line, it is also the maximum over the set $x \geq 0$.

Step 5 is to answer the question posed originally in step 1; i.e., when to sell the pig in order to maximize profit. The answer obtained by our mathematical model is to sell the pig after eight days, thus obtaining a net profit of \$133.20. This answer is valid as long as the assumptions made in step 1 remain valid. Related questions and alternative assumptions can be addressed by changing what we did in step 1. Since we are dealing with a real problem (A farmer owns pigs. When should they be sold?), there is an element of risk involved in step 1. For that reason it is usually necessary to investigate several alternatives. This process, called *sensitivity analysis*, will be discussed in the next section.

The main purpose of this section was to introduce the five-step method for mathematical modeling. Figure 1.3 summarizes the method in a form convenient for later reference. In this book we will apply the five-step method to solve a wide variety of problems in mathematical modeling. Our discussion of step 2 will generally include a description of the modeling approach selected, along with an example or two. The reader who is already familiar with the modeling approach may choose to skip this part, or just skim to pick up the notation. Some of the other points summarized in Fig. 1.3, such as the use of "appropriate technology," will be expanded upon later in this book.

The exercises at the end of each chapter also require the application of the

Step 1. Ask the question.

- Make a list of all the variables in the problem, including appropriate units.

- Be careful not to confuse variables and constants.

- State any assumptions you are making about these variables, including equations and inequalities.

- Check units to make sure that your assumptions make sense.

- State the objective of the problem in precise mathematical terms.

Step 2. Select the modeling approach.

- Choose a general solution procedure to be followed in solving this problem.

- Generally speaking, success in this step requires experience, skill, and familiarity with the relevant literature.

- In this book we will usually specify the modeling approach to be used.

Step 3. Formulate the model.

- Restate the question posed in step 1 in the terms of the modeling approach specified in step 2.

- You may need to relabel some of the variables specified in step 1 in order to agree with the notation used in step 2.

- Note any additional assumptions made in order to fit the problem described in step 1 into the mathematical structure specified in step 2.

Step 4. Solve the model.

- Apply the general solution procedure specified in step 2 to the specific problem formulated in step 3.

- Be careful in your mathematics. Check your work for math errors. Does your answer make sense?

- Use appropriate technology. Computer algebra systems, graphics, and numerical software will increase the range of problems within your grasp, and they also help reduce math errors.

Step 5. Answer the question.

- Rephrase the results of step 4 in nontechnical terms.

- Avoid mathematical symbols and jargon.

- Anyone who can understand the statement of the question as it was presented to you should be able to understand your answer.

Figure 1.3: The five-step method.

five-step method. Getting in the habit of using the five-step method now will make it easier to succeed on the more difficult modeling problems to come. Be sure to pay particular attention to step 5. In the real world, it is not enough to be right. You also need the ability to communicate your findings to others, some of whom may not be as mathematically knowledgeable as you.

1.2   Sensitivity Analysis

The preceding section outlines the five-step approach to mathematical modeling. The process begins by making some assumptions about the problem. We are rarely certain enough about things to be able to expect all of these assumptions to be exactly valid. Therefore, we need to consider how sensitive our conclusions are to each of the assumptions we have made. This kind of sensitivity analysis is an important aspect of mathematical modeling. The details vary according to the modeling approach used, and so our discussion of sensitivity analysis will continue throughout this book. We will focus here on sensitivity analysis for simple one–variable optimization problems.

In the preceding section we used the pig problem (Example 1.1) to illustrate the five-step approach to mathematical modeling. Figure 1.1 summarizes the assumptions we made in solving that problem. In this instance the data and assumptions were mostly spelled out for us. Even so, we need to be critical. Data are obtained by measurement, observation, and sometimes sheer guess. We need to consider the possibility that the data are not precise.

Some data are naturally known with much more certainty than others. The current weight of the pig, the current price for pigs, and the cost per day of keeping the pig are easy to measure and are known to a great degree of certainty. The rate of growth of the pig is a bit less certain, and the rate at which the price is falling is even less certain. Let r denote the rate at which the price is falling. We assumed that $r = 0.01$ dollars per day, but let us now suppose that the actual value of r is different. By repeating the solution procedure for several different values of r, we can get an idea of the sensitivity of our answer to the value of r. Table 1.1 shows the results of solving our problem for a few selected values of r. Figure 1.4 contains the same sensitivity data in graphical form. We can see that the optimal time to sell is quite sensitive to the parameter r.

A more systematic method for measuring this sensitivity would be to treat r as an unknown parameter, following the same steps as before. Writing

$$p = 0.65 - rt,$$

we can proceed as before to obtain

$$y = f(x)$$
$$= (0.65 - rx)(200 + 5x) - 0.45x.$$

Then we can compute

$$f'(x) = \frac{-2(25rx + 500r - 7)}{5}$$

r	x
(\$/day)	(days)
0.008	15.0
0.009	11.1
0.010	8.0
0.011	5.5
0.012	3.3

Table 1.1: Sensitivity of best time to sell x to rate r at which price is falling for the pig problem.
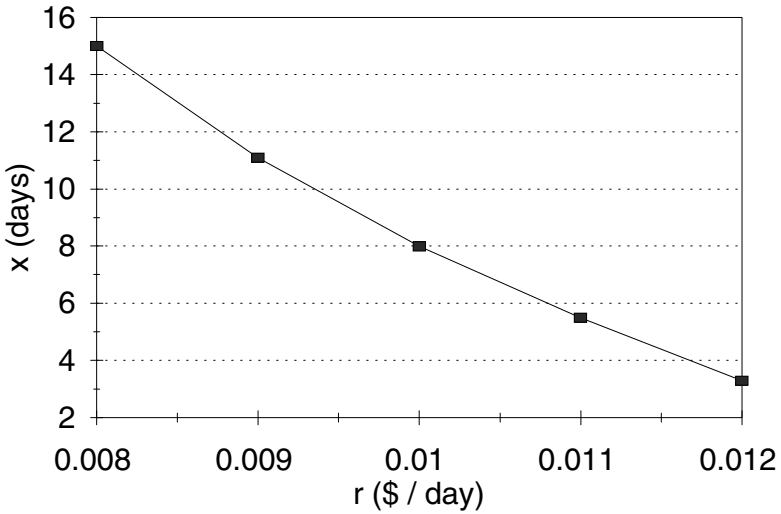

Figure 1.4: Graph of best time to sell x versus rate r at which price is falling for the pig problem.
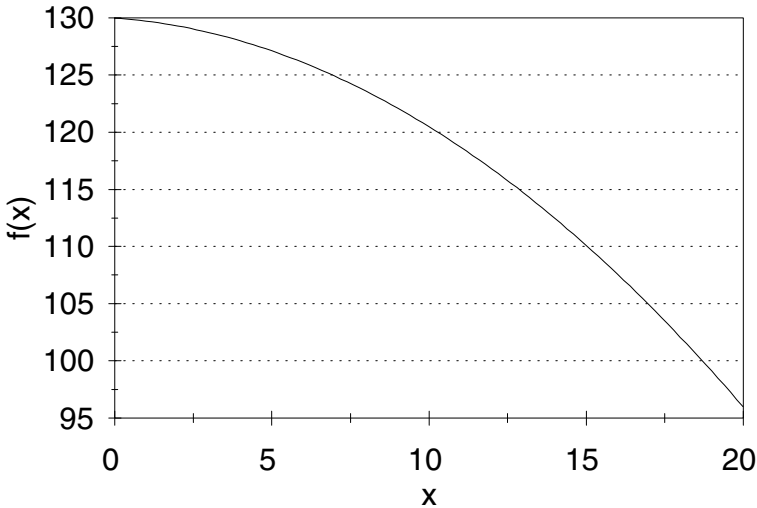
Figure 1.5: Graph of net profit $f(x) = (0.65 - 0.015x)(200 + 5x) - 0.45x$ versus time to sell x for the pig problem in the case $r = 0.015$.

so that $f'(x) = 0$ at the point

$$x = \frac{(7 - 500r)}{25r}. \tag{1.2}$$

The optimal time to sell is given by Eq. (1.2) as long as this expression is positive, i.e., as long as $0 < r \leq 0.014$. For $r > 0.014$, the vertex of the parabola $y = f(x)$ lies outside of the set $x \geq 0$ over which we are maximizing. In this case the optimal time to sell is at $x = 0$ since we have $f' < 0$ on the entire interval $[0, \infty)$. See Figure 1.5 for an illustration in the case $r = 0.015$.

We are also uncertain about the growth rate g of the pig. We have assumed that $g = 5$ lbs/day. More generally, we have that

$$w = 200 + gt,$$

which leads to the equation

$$f(x) = (0.65 - 0.01x)(200 + gx) - 0.45x, \tag{1.3}$$

so that

$$f'(x) = \frac{-(2gx + 5(49 - 13g))}{100}.$$

Now $f'(x) = 0$ at the point

$$x = \frac{5(13g - 49)}{2g}. \tag{1.4}$$
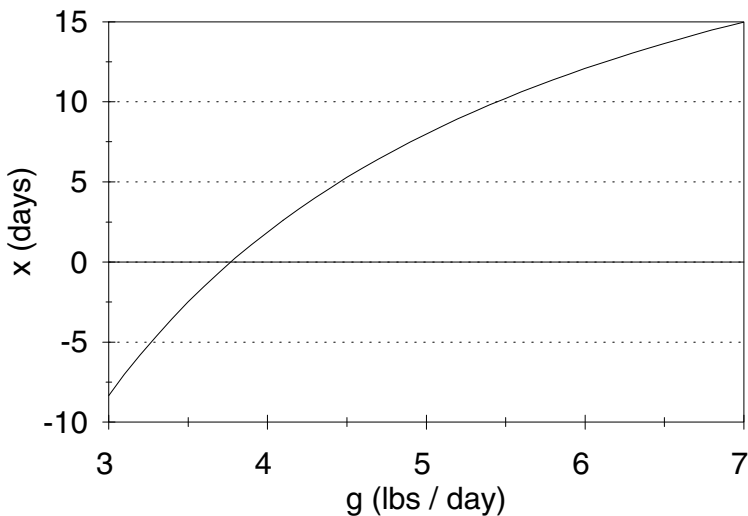
Figure 1.6: Graph of best time to sell x versus growth rate g for the pig problem.

The optimal time to sell is given by Eq. (1.4) so long as it represents a nonneg-
ative value of x. Figure 1.6 shows the relationship between the growth rate g
and the optimal time to sell.

It is most natural and most useful to interpret sensitivity data in terms of
relative change or percent change, rather than in absolute terms. For example, a
10% decrease in r leads to a 39% increase in x, while a 10% decrease in g leads
to a 34% decrease in x. If x changes by an amount Δx, the relative change
in x is given by $\Delta x/x$, and the percent change in x is $100\,\Delta x/x$. If r changes
by Δr, resulting in the change Δx in x, then the ratio between the relative
changes is $\Delta x/x$ divided by $\Delta r/r$. Letting $\Delta r \to 0$ and using the definition of
the derivative, we obtain

$$\frac{\Delta x/x}{\Delta r/r} \to \frac{dx}{dr} \cdot \frac{r}{x}.$$

We call this limiting quantity the *sensitivity* of x to r, and we will denote it by
$S(x, r)$. In the pig problem we have

$$\frac{dx}{dr} = \frac{-7}{25r^2}$$
$$= -2,800$$

at the point $r = 0.01$ and $x = 8$; thus

$$S(x, r) = \frac{dx}{dr} \cdot \frac{r}{x}$$

$$= (-2,800)\left(\frac{.01}{8}\right)$$

$$= \frac{-7}{2}.$$

If r goes up by 2%, then x goes down by 7%. Since

$$\frac{dx}{dg} = \frac{245}{2g^2}$$

$$= 4.9,$$

we have

$$S(x, g) = \frac{dx}{dg} \cdot \frac{g}{x}$$

$$= (4.9)\left(\frac{5}{8}\right)$$

$$= 3.0625,$$

so that a 1% increase in the growth rate of the pig would cause us to wait about 3% longer to sell the pig.

In order to compute the sensitivity $S(y, g)$, first substitute (1.4) into the objective function $y = f(x)$ from (1.3) to obtain

$$y = \left(0.65 - 0.01\left[\frac{5(13g - 49)}{2g}\right]\right)\left(200 + g\left[\frac{5(13g - 49)}{2g}\right]\right)$$

$$- 0.45\left[\frac{5(13g - 49)}{2g}\right]$$

$$= \frac{150.0625}{g} + 50.375 + 10.5625g.$$

Then compute the derivative

$$\frac{dy}{dg} = -\frac{150.0625}{g^2} + 10.5625,$$

and substitute $g = 5$ to get $dy/dg = 4.56$, which leads to

$$S(y, g) = \frac{dy}{dg} \cdot \frac{g}{y}$$

$$= (4.56)\left(\frac{5}{133.20}\right)$$

$$= 0.17.$$

If the pig grows 10% faster than expected, the expected net profit will be 1.7% larger. The computation of the derivative dy/dg in this case involves quite a bit of algebra. In Chapter 2, we will discuss how a *computer algebra system* can be used to perform the necessary algebraic computations.

The successful application of sensitivity analysis procedures requires good judgment. It is usually not possible to compute sensitivity coefficients for each parameter in the model, nor is this particularly desirable. We need to select those parameters about which there is the most uncertainty and perform sensitivity analysis on them. The interpretation of sensitivity coefficients also depends on the degree of uncertainty, the fundamental question being the extent to which our uncertainty about the data affects our confidence in the answer. In the pig problem, we are probably considerably more certain of the growth rate g than of the rate r at which prices fall. A 25% error in g would be quite surprising if we have observed the past history of growth in this pig or in similar animals. A 25% error in our estimate of r would not be at all surprising.

1.3   Sensitivity and Robustness

A mathematical model is robust if the conclusions it leads to remain true even though the model is not completely accurate. In real problems we will never have perfect information, and even if it were possible to construct a perfectly accurate model, we might be better off with a simpler and more tractable approximation. For this reason a consideration of robustness is a necessary ingredient in any mathematical modeling project.

In the preceding section we introduced the process of sensitivity analysis, which is a way to gauge the robustness of a model with respect to assumptions about the data. There are other assumptions made in step 1 of the mathematical modeling process which should also be examined. While it is often necessary to make assumptions for purposes of mathematical convenience and simplicity, it is the responsibility of the modeler to see to it that these assumptions are not so specialized as to invalidate the results of the modeling process.

Figure 1.1 contains a summary of the assumptions made in solving the pig problem. Aside from data values, the main assumptions are that both the weight of the pig and the selling price per pound are linear functions of time. These are obviously simplifying assumptions and cannot be expected to hold exactly. After all, according to these assumptions, a year from now the pig would weigh

$$w = 200 + 5(365)$$
$$= 2,025 \, \text{lbs}$$

and sell for

$$p = 0.65 - 0.01(365)$$
$$= -3.00 \, \text{dollars/lb.}$$

A more realistic model would take into account both the nonlinearity of these functions and the increasing uncertainty as time goes on.

How can a model give the right answer if the assumptions are wrong? While mathematical modeling strives for perfection, perfection can never be achieved. It would be more descriptive to say that mathematical modeling strives *toward* perfection. A well–constructed mathematical model will be robust, which is to say that while the answers it gives may not be perfectly correct, they will be close enough to be useful in a real–world context.

Let us examine the linearity assumptions made in the pig problem. Our basic equation is

$$P = pw - 0.45t,$$

where p is the selling price of the pig in dollars per pound, and w is the weight of the pig in pounds. If the original data and assumptions of the model are not too far off, then the best time to sell the pig is obtained by setting $P' = 0$. Calculate to find

$$p'w + pw' = 0.45$$

dollars per day. The term $p'w + pw'$ represents the rate of increase in the value of the pig. Our model tells us to keep the pig as long as the value of the pig is increasing faster than the cost of feeding it. Furthermore, the change in the pig's value has two components, $p'w$ and pw'. The first term, $p'w$, represents the loss in value due to a drop in price. The second term, pw', represents the gain in value due to the pig gaining weight. Consider the practical problems involved in the application of this more general model. The data required include a complete specification of both the future growth of the pig and the future changes in price as differentiable functions of time. There is no way to know these functions exactly. There is even some question as to whether they make sense. Can the pig be sold at 3 A.M. Sunday morning? Can price be an irrational number? Let us construct a realistic scenario. The farmer has a pig weighing approximately 200 lbs. The pig has been gaining about five lbs/day during the last week. Five days ago the pig could have been sold for 70 cents/lb but by now the price has dropped to 65 cents/lb. What should we do? The obvious approach is to project on the basis of this data ($w = 200$, $w' = 5$, $p = 0.65$, $p' = -0.01$) and decide when to sell. This is exactly what we did. We understand that p' and w' will not remain constant over the next few weeks, and that therefore p and w will not be linear functions of time. However, as long as p' and w' do not change too much over this period, the error involved in assuming they remain constant will not be too great.

We are now prepared to give a somewhat broader interpretation to the results of our sensitivity analysis from the preceding section. Recall that the sensitivity of the best time to sell (x) to changes in the growth rate w' was calculated to be 3. Suppose that in fact the growth rate over the next few weeks is somewhere between 4.5 and 5.5 lbs/day. This is within 10% of the assumed value. Then the best time to sell the pig will be within 30% of 8 days, or between 5 and 11 days. The amount of lost profit by selling at 8 days is less than 1 dollar.

With regard to price, suppose that we feel the value $p' = -.01$, or a 1 cent/day drop in price over the next few weeks, is a worst–case scenario. Prices are likely to drop more slowly in the future and may even level off ($p' = 0$).

All we can really say now is that we should wait at least 8 days to sell. For small values of p' (near zero), our model suggests waiting a very long time to sell. However, our model is not valid over long time intervals. The best course of action in this case is probably to keep the pig for a week, reestimate the parameter values p, p', w', and w, and start over.

1.4   Exercises

1. An automobile manufacturer makes a profit of \$1,500 on the sale of a certain model. It is estimated that for every \$100 of rebate, sales increase by 15%.

   (a) What amount of rebate will maximize profit? Use the five-step method, and model as a one–variable optimization problem.

   (b) Compute the sensitivity of your answer to the 15% assumption. Consider both the amount of rebate and the resulting profit.

   (c) Suppose that rebates actually generate only a 10% increase in sales per \$100. What is the effect? What if the response is somewhere between 10 and 15% per \$100 of rebate?

   (d) Under what circumstances would a rebate offer cause a reduction in profit?

2. In the pig problem, perform a sensitivity analysis based on the cost per day of keeping the pig. Consider both the effect on the best time to sell and on the resulting profit. If a new feed costing 60 cents/day would let the pig grow at a rate of 7 lbs/day, would it be worth switching feed? What is the minimum improvement in growth rate that would make this new feed worthwhile?

3. Reconsider the pig problem of Example 1.1, but now assume that the price for pigs is starting to level off. Let

$$p = 0.65 - 0.01t + 0.00004t^2 \tag{1.5}$$

   represent the price for pigs (cents/lb) after t days.

   (a) Graph Eq. (1.5) along with our original price equation. Explain why our original price equation could be considered as an approximation to Eq. (1.5) for values of t near zero.

   (b) Find the best time to sell the pig. Use the five-step method, and model as a one–variable optimization problem.

   (c) The parameter 0.00004 represents the rate at which price is leveling off. Conduct a sensitivity analysis on this parameter. Consider both the optimal time to sell and the resulting profit.

    (d) Compare the results of part (b) to the optimal solution contained in the text. Comment on the robustness of our assumptions about price.

4. An oil spill has fouled 200 miles of Pacific shoreline. The oil company responsible has been given 14 days to clean up the shoreline, after which a fine will be levied in the amount of $10,000/day. The local cleanup crew can scrub five miles of beach per week at a cost of $500/day. Additional crews can be brought in at a cost of $18,000 plus $800/day for each crew.

    (a) How many additional crews should be brought in to minimize the total cost to the company? Use the five-step method. How much will the clean–up cost?

    (b) Examine the sensitivity to the rate at which a crew can clean up the shoreline. Consider both the optimal number of crews and the total cost to the company.

    (c) Examine the sensitivity to the amount of the fine. Consider the number of days the company will take to clean up the spill and the total cost to the company.

    (d) The company has filed an appeal on the grounds that the amount of the fine is excessive. Assuming that the only purpose of the fine is to motivate the company to clean up the oil spill in a timely manner, is the fine excessive?

5. It is estimated that the growth rate of the fin whale population (per year) is $rx(1 - x/K)$, where $r = 0.08$ is the intrinsic growth rate, $K = 400,000$ is the maximum sustainable population, and x is the current population, now around 70,000. It is further estimated that the number of whales harvested per year is about $0.00001\ Ex$, where E is the level of fishing effort in boat–days. Given a fixed level of effort, population will eventually stabilize at the level where growth rate equals harvest rate.

    (a) What level of effort will maximize the sustained harvest rate? Model as a one–variable optimization problem using the five-step method.

    (b) Examine the sensitivity to the intrinsic growth rate. Consider both the optimum level of effort and the resulting population level.

    (c) Examine the sensitivity to the maximum sustainable population. Consider both the optimum level of effort and the resulting population level.

6. In Exercise 5, suppose that the cost of whaling is $500 per boat–day, and the price of a fin whale carcass is $6,000.

    (a) Find the level of effort that will maximize profit over the long term. Model as a one–variable optimization problem using the five-step method.

(b) Examine the sensitivity to the cost of whaling. Consider both the eventual profit in $/year and the level of effort.

(c) Examine the sensitivity to the price of a fin whale carcass. Consider both profit and level of effort.

(d) Over the past 30 years there have been several unsuccessful attempts to ban whaling worldwide. Examine the economic incentives for whalers to continue harvesting. In particular, determine the conditions (values of the two parameters: cost per boat–day and price per fin whale carcass) under which harvesting the fin whale produces a sustained profit over the long term.

7. Reconsider the pig problem of Example 1.1, but now suppose that our objective is to maximize our profit rate ($/day). Assume that we have already owned the pig for 90 days and have invested $100 in this pig to date.

(a) Find the best time to sell the pig. Use the five-step method, and model as a one–variable optimization problem.

(b) Examine the sensitivity to the growth rate of the pig. Consider both the best time to sell and the resulting profit rate.

(c) Examine the sensitivity to the rate at which the price for pigs is dropping. Consider both the best time to sell and the resulting profit rate.

8. Reconsider the pig problem of Example 1.1, but now take into account the fact that the growth rate of the pig decreases as the pig gets older. Assume that the pig will be fully grown in another five months.

(a) Find the best time to sell the pig in order to maximize profit. Use the five-step method, and model as a one–variable optimization problem.

(b) Examine the sensitivity to the time it will take until the pig is fully grown. Consider both the best time to sell and the resulting profit.

9. A local daily newspaper with a circulation of 80,000 subscribers is thinking of raising its subscription price. Currently the price is $1.50 per week, and it is estimated that the paper would lose 5,000 subscribers if the rate were to be raised by ten cents/week.

(a) Find the subscription price that maximizes profit. Use the five-step method, and model as a one–variable optimization problem.

(b) Examine the sensitivity of your answer in part (a) to the assumption of 5,000 lost subscribers. Calculate the optimal subscription rate assuming that this parameter is 3,000, 4,000, 5,000, 6,000, or 7,000.

(c) Let $n = 5,000$ denote the number of subscribers lost when the subscription price increases by ten cents. Calculate the optimal subscription price p as a function of n, and use this formula to determine the sensitivity $S(p, n)$.

(d) Should the paper change its subscription price? Justify your conclusions in plain English.

Further Reading

1. Cameron, D., Giordano, F. and Weir, M. *Modeling Using the Derivative: Numerical and Analytic Solutions.* UMAP module 625.

2. Cooper, L. and Sternberg, D. (1970) *Introduction to Methods of Optimization.* W. B. Saunders, Philadelphia.

3. Gill, P., Murray, W. and Wright, M. (1981) *Practical Optimization.* Academic Press, New York.

4. Meyer, W. (1984) *Concepts of Mathematical Modeling.* McGraw–Hill, New York.

5. Rudin W. (1976) *Principles of Mathematical Analysis.* 3rd Ed., McGraw–Hill, New York.

6. Whitley, W. *Five Applications of Max–Min Theory from Calculus.* UMAP module 341.

Chapter 2

MULTIVARIABLE OPTIMIZATION

Many optimization problems require the simultaneous consideration of a number of independent variables. In this chapter we consider the simplest category of multivariable optimization problems. The techniques should be familiar to most students from multivariable calculus. In this chapter we also introduce the use of computer algebra systems to handle some of the more complicated algebraic computations.

2.1 Unconstrained Optimization

The simplest type of multivariable optimization problems involves finding the maximum or minimum of a differentiable function of several variables over a nice set. Further complications arise, as we will see later, when the set over which we optimize takes a more complex form.

Example 2.1. A manufacturer of color TV sets is planning the introduction of two new products, a 19–inch LCD flat panel set with a manufacturer's suggested retail price (MSRP) of $339 and a 21–inch LCD flat panel set with an MSRP of $399. The cost to the company is $195 per 19–inch set and $225 per 21–inch set, plus an additional $400,000 in fixed costs. In the competitive market in which these sets will be sold, the number of sales per year will affect the average selling price. It is estimated that for each type of set, the average selling price drops by one cent for each additional unit sold. Furthermore, sales of the 19–inch set will affect sales of the 21–inch set, and vice–versa. It is estimated that the average selling price for the 19–inch set will be reduced by an additional 0.3 cents for each 21–inch set sold, and the price for the 21–inch set will decrease by 0.4 cents for each 19–inch set sold. How many units of each type of set should be manufactured?

Variables: s = number of 19–inch sets sold (per year)
t = number of 21–inch sets sold (per year)
p = selling price for a 19–inch set ($)
q = selling price for a 21–inch set ($)
C = cost of manufacturing sets ($/year)
R = revenue from the sale of sets ($/year)
P = profit from the sale of sets ($/year)

Assumptions: $p = 339 - 0.01s - 0.003t$
$q = 399 - 0.004s - 0.01t$
$R = ps + qt$
$C = 400,000 + 195s + 225t$
$P = R - C$
$s \geq 0$
$t \geq 0$

Objective: Maximize P

Figure 2.1: Results of step 1 of the color TV problem.

The five-step approach to mathematical modeling, introduced in the preceding chapter, will be used to solve this problem. Step 1 is to ask a question. We begin by making a list of variables. Next we write down the relations between variables and any other assumptions, such as nonnegativity. Finally, we formulate a question in mathematical terms, using the established notation. The results of step 1 are summarized in Figure 2.1.

Step 2 is to select the modeling approach. We will solve this problem as a multivariable unconstrained optimization problem. This type of problem is typically treated in introductory courses in multivariable calculus. We will outline the model and the general solution procedure here. We refer the reader to any introductory calculus textbook for details and mathematical proofs.

We are given a function $y = f(x_1, \ldots, x_n)$ on a subset S of the n–dimensional space $\mathbb{R}^n$. We wish to find the maximum and/or minimum values of f on the set S. There is a theorem that states that if f attains its maximum or minimum at an interior point $(x_1, \ldots, x_n)$ in S, then $\nabla f = 0$ at that point, assuming that f is differentiable at that point. In other words, at the extreme point

$$\frac{\partial f}{\partial x_1}(x_1, \ldots, x_n) = 0$$

$$\frac{\partial f}{\partial x_n}(x_1, \ldots, x_n) = 0. \tag{2.1}$$

The theorem allows us to rule out as a candidate for max–min any point in the interior of S for which any of the partial derivatives of f

do not equal zero. Thus, to find the max–min points we must solve simultaneously the n equations in n unknowns defined by Equation (2.1). Then we must also check any points on the boundary of S, as well as any points where one or more of the partial derivatives is undefined.

Step 3 is to formulate the model, using the standard form chosen in step 2. Let

$$\begin{aligned}
P &= R - C \\
&= ps + qt - (400,000 + 195s + 225t) \\
&= (339 - 0.01s - 0.003t)s \\
&\quad + (399 - 0.004s - 0.01t)t \\
&\quad - (400,000 + 195s + 225t).
\end{aligned}$$

Now let $y = P$ be the quantity we wish to maximize, and let $x_1 = s$, $x_2 = t$ be our decision variables. Our problem now is to maximize

$$\begin{aligned}
y &= f(x_1, x_2) \\
&= (339 - 0.01x_1 - 0.003x_2)x_1 \\
&\quad + (399 - 0.004x_1 - 0.01x_2)x_2 \\
&\quad - (400,000 + 195x_1 + 225x_2)
\end{aligned} \tag{2.2}$$

over the set

$$S = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}. \tag{2.3}$$

Step 4 is to solve the problem, using the standard solution methods outlined in step 2. The problem is to maximize the function f given by Eq. (2.2) over the set S defined in Eq. (2.3). Figure 2.2 shows a 3–D graph of the function f. This plot indicates that f attains its maximum in the interior of S. Figure 2.3 shows a plot of the level sets of f.

From this plot we can estimate that the maximum value of the function f occurs around $x_1 = 5,000$ and $x_2 = 7,000$. The function f is a paraboloid, and the vertex of the paraboloid is the unique solution to Eq. (2.1) obtained by setting $\nabla f = 0$. We compute that

$$\begin{aligned}
\frac{\partial f}{\partial x_1} &= 144 - 0.02x_1 - 0.007x_2 = 0 \\
\frac{\partial f}{\partial x_2} &= 174 - 0.007x_1 - 0.02x_2 = 0
\end{aligned} \tag{2.4}$$

at the point

$$\begin{aligned}
x_1 &= \frac{554,000}{117} \approx 4,735 \\
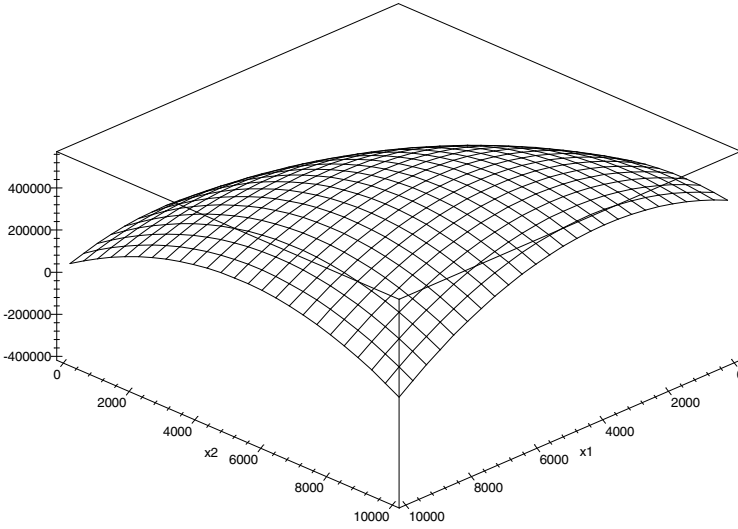x_2 &= \frac{824,000}{117} \approx 7,043.
\end{aligned} \tag{2.5}$$

Figure 2.2: 3–D graph of profit $y = f(x_1, x_2)$ from (2.2) versus production levels x_1 of 19–inch sets and x_2 of 21–inch sets for the color TV problem.
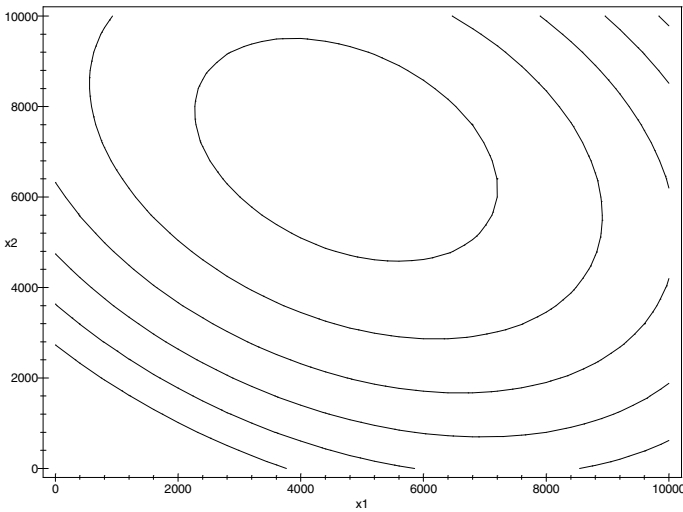

Figure 2.3: Contour plot showing level sets of profit $y = f(x_1, x_2)$ from (2.2) versus production levels x_1 of 19–inch sets and x_2 of 21–inch sets for the color TV problem.

The point (x_1, x_2) defined by Eq. (2.5) represents the global maximum of f over the entire real plane. It is therefore also the maximum of f over the set S defined in Eq. (2.3). The maximum value of f is obtained by substituting Eq. (2.5) back into Eq. (2.2), which yields

$$y = \frac{21,592,000}{39} \approx 553,641. \tag{2.6}$$

The calculations of step 4 in this problem are a bit cumbersome. In cases like this one, it is appropriate to use a *computer algebra system* to perform the necessary calculations. Computer algebra systems can differentiate, integrate, solve equations, and simplify algebraic expressions. Most packages can also perform matrix algebra, draw graphs, and solve some differential equations. Several good computer algebra systems (Maple, Mathematica, Derive, etc.) are available for both mainframe and personal computers, and many systems offer a student version at a substantially reduced price. The graphs in Figure 2.2 and Figure 2.3 were drawn using the computer algebra system Maple. Computer algebra systems are an example of the kind of "appropriate technology" we referred to in Fig. 1.3, in our summary of the five-step method.   Figure 2.4 shows the results of using the computer algebra system Mathematica to solve the current model.

There are several advantages to using a computer algebra system for a problem like this. It is more efficient and more accurate. You will be more productive if you learn to use this technology, and it will give you the freedom to concentrate more on the larger issues of problem solving instead of getting bogged down in the calculations. We will illustrate the use of computer algebra systems again in our sensitivity analysis calculations below, where the algebra is even more exacting.

The final step, step 5, is to answer the question in plain English. Simply stated, the company can maximize profits by manufacturing 4,735 of the 19–inch sets and 7,043 of the 21–inch sets, resulting in a net profit of $553,641 for the year. The average selling price for a 19–inch set is $270.52 and $309.63 for a 21–inch set. The projected revenue is $3,461,590, resulting in a profit margin (profit/revenue) of 16 percent. These figures indicate a profitable venture, so we would recommend that the company proceed with the introduction of these new products.

The conclusions of the preceding paragraph are based on the assumptions illustrated in Fig. 2.1. Before reporting our findings to the company, we should perform sensitivity analysis to insure that our conclusions are robust with respect to our assumptions about both the market for TV sets and the manufacturing process. Our main concern is the value of the decision variables x_1 and x_2, since the company must act on this information.

We illustrate the procedure for sensitivity analysis by examining the sensitivity to price elasticity for 19–inch sets, which we will denote by the variable a. Price elasticity is the term used by economists to describe the sensitivity of quantity sold to the asking price. In our model we assumed that $a = 0.01$

In[1]:= **y = (339 − x1 / 100 − 3 x2 / 1000) x1 +**
          (399 − 4 x1 / 1000 − x2 / 100) x2 −
          (400000 + 195 x1 + 225 x2)

Out[1]= $\left(-\dfrac{x1}{100} - \dfrac{3\,x2}{1000} + 339\right)x1 - 195\,x1 + \left(-\dfrac{x1}{250} - \dfrac{x2}{100} + 399\right)x2 - 225\,x2 - 400000$

In[2]:= **dydx1 = D[y, x1]**

Out[2]= $-\dfrac{x1}{50} - \dfrac{7\,x2}{1000} + 144$

In[3]:= **dydx2 = D[y, x2]**

Out[3]= $-\dfrac{7\,x1}{1000} - \dfrac{x2}{50} + 174$

In[4]:= **s = Solve[{dydx1 == 0, dydx2 == 0}, {x1, x2}]**

Out[4]= $\left\{\left\{x1 \to \dfrac{554000}{117},\ x2 \to \dfrac{824000}{117}\right\}\right\}$

In[5]:= **N[s]**

Out[5]= {{x1 → 4735.04, x2 → 7042.74}}

In[6]:= **y /. s**

Out[6]= $\left\{\dfrac{21592000}{39}\right\}$

In[7]:= **N[%]**

Out[7]= {553641.}

Figure 2.4: Optimal solution to the color TV problem using the computer algebra system Mathematica.

dollars per set. Substituting into our previous formula, we obtain

$$\begin{aligned}
y &= f(x_1,\, x_2) \\
&= (339 - ax_1 - 0.003x_2)x_1 \\
&\quad + (399 - 0.004x_1 - 0.01x_2)x_2 \\
&\quad - (400{,}000 + 195x_1 + 225x_2).
\end{aligned} \tag{2.7}$$

When we compute partial derivatives and set them equal to zero, we obtain

$$\begin{aligned}
\frac{\partial f}{\partial x_1} &= 144 - 2ax_1 - 0.007x_2 = 0 \\
\frac{\partial f}{\partial x_2} &= 174 - 0.007x_1 - 0.02x_2 = 0.
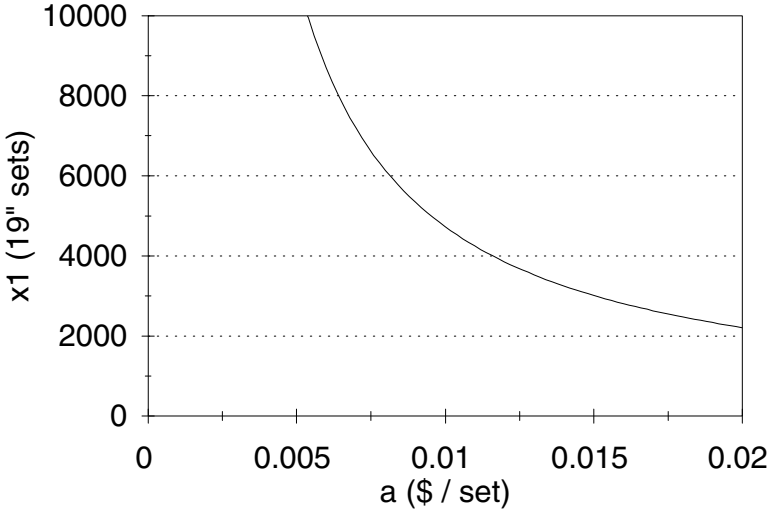\end{aligned} \tag{2.8}$$

Figure 2.5: Graph of optimum production level x_1 of 19–inch sets versus price elasticity a for the color TV problem.

Solving for x_1 and x_2 as before yields

$$x_1 = \frac{1,662,000}{40,000a - 49}$$

$$x_2 = 8,700 - \frac{581,700}{40,000a - 49}. \tag{2.9}$$

See Figs. 2.5 and 2.6 for the graphs of x_1 and x_2 versus a.

From these graphs it appears that a higher price elasticity a for 19–inch sets will reduce the optimal production level x_1 for 19–inch sets and increase the optimal production level x_2 for 21–inch sets. It also appears that x_1 is more sensitive to a than x_2, which seems to make sense. To get a numerical measure of these sensitivities we compute

$$\frac{dx_1}{da} = \frac{-66,480,000,000}{(40,000a - 49)^2}$$

$$= \frac{-22,160,000,000}{41,067}$$

at $a = 0.01$, so that

$$S(x_1, a) = \left(\frac{-22,160,000,000}{41,067} \right) \left(\frac{0.01}{554,000/117} \right)$$
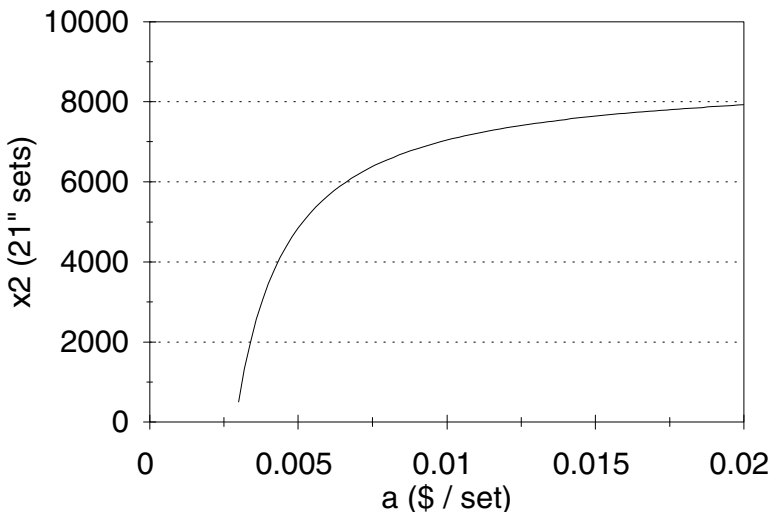
$$= -\frac{400}{351} \approx -1.1.$$

Figure 2.6: Graph of optimum production level x_2 of 21–inch sets versus price elasticity a for the color TV problem.

A similar calculation yields

$$S(x_2,\, a) = \frac{9,695}{36,153} \approx 0.27.$$

If the price elasticity for 19–inch sets were to increase by 10%, then we should make 11% fewer 19–inch sets and 2.7% more 21–inch sets.

Next, we consider the sensitivity of y to a. What effect will a change in the price elasticity for 19–inch sets have on our profits? To obtain a formula for y in terms of a, we substitute Eq. (2.9) back into Eq. (2.7) to get

$$
\begin{aligned}
y = &\left[339 - a\left(\frac{1,662,000}{40,000a - 49} \right) - 0.003\left(8,700 - \frac{581,700}{40,000a - 49} \right) \right] \\
&\times \left(\frac{1,662,000}{40,000a - 49} \right) \\
&+ \left[399 - 0.004\left(\frac{1,662,000}{40,000a - 49} \right) - 0.01\left(8,700 - \frac{581,700}{40,000a - 49} \right) \right] \quad (2.10) \\
&\times \left(8,700 - \frac{581,700}{40,000a - 49} \right) \\
&- \left[400,000 + 195\left(\frac{16,620,000}{40,000a - 49} \right) + 225\left(8,700 - \frac{581,700}{40,000a - 49} \right) \right].
\end{aligned}
$$

See Fig. 2.7 for a graph of y versus a. It appears that an increase in price elasticity for 19–inch sets will result in lower profits.
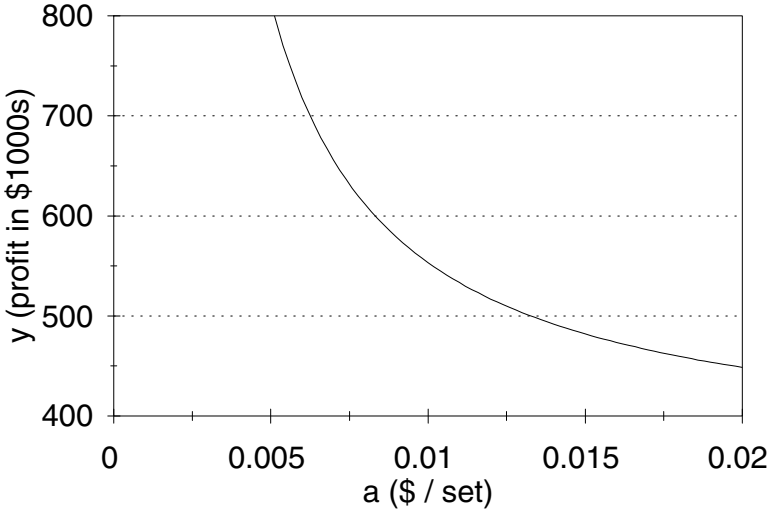
Figure 2.7: Graph of optimum profit y versus price elasticity a for the color TV problem.

To compute $S(y, a)$ we will need to obtain a formula for dy/da. One way to do this would be to employ one–variable methods directly on Eq. (2.10), perhaps with the aid of a computer algebra system. Another method, which is computationally more efficient, is to use the multivariable chain rule, which implies that

$$\frac{dy}{da} = \frac{\partial y}{\partial x_1}\frac{dx_1}{da} + \frac{\partial y}{\partial x_2}\frac{dx_2}{da} + \frac{\partial y}{\partial a}. \tag{2.11}$$

Since both $\partial y/\partial x_1$ and $\partial y/\partial x_2$ are zero at the optimum, we have

$$\frac{dy}{da} = \frac{\partial y}{\partial a} = -x_1^2$$

directly from Eq. (2.7), so

$$S(y, a) = -\left(\frac{554,000}{117}\right)^2 \frac{0.01}{(21,592,000/39)}$$

$$= -\frac{383,645}{947,349} \approx -0.40.$$

Thus, a 10% increase in price elasticity for 19–inch sets will result in a 4% drop in profit.

The fact that the term

$$\frac{\partial y}{\partial x_1}\frac{dx_1}{da} + \frac{\partial y}{\partial x_2}\frac{dx_2}{da} = 0$$

```
> y:=(339-a*x1-3*x2/1000)*x1
>     +(399-4*x1/1000-x2/100)*x2-(400000+195*x1+225*x2);
```

$$y := \left(339 - a\,x1 - \frac{3}{1000}\,x2 \right)\,x1 + \left(399 - \frac{1}{250}\,x1 - \frac{1}{100}\,x2 \right)\,x2 - 400000 - 195\,x1 - 225\,x2$$

```
> dydx1:=diff(y,x1);
```

$$dydx1 := -2\,a\,x1 + 144 - \frac{7}{1000}\,x2$$

```
> dydx2:=diff(y,x2);
```

$$dydx2 := -\frac{7}{1000}\,x1 - \frac{1}{50}\,x2 + 174$$

```
> s:=solve({dydx1=0,dydx2=0},{x1,x2});
```

$$s := \left\{ x2 = \frac{48000\,(-21 + 7250\,a)}{-49 + 40000\,a},\ x1 = \frac{1662000}{-49 + 40000\,a} \right\}$$

```
> assign(s);dx1da:=diff(x1,a);
```

$$dx1da := -\frac{66480000000}{(-49 + 40000\,a)^2}$$

```
> assign(a=1/100);x1;
```

$$\frac{554000}{117}$$

```
> sx1a:=dx1da*(a/x1);
```

$$sx1a := \frac{-400}{351}$$

```
> evalf(sx1a);
```

$$-1.139601140$$

```
> 
```

Figure 2.8: Calculation of the sensitivity $S(x_1, a)$ for the color TV problem using the computer algebra system Maple.

in Eq. (2.11) has its own real–world significance. This part of the derivative dy/da represents the effect on profits of the changing optimal production levels x_1 and x_2. The fact that it sums to zero means that small changes in production levels have (at least in the linear approximation) no effect on profits. Geometrically, since we are at the maximum point where the curve $y = f(x_1, x_2)$ is flat, small changes in x_1 and x_2 have little effect on y. Almost all of the drop in optimal profits caused by a 10% increase in price elasticity for 19–inch sets is due to the change in selling price. Therefore, the production levels given by our model will be very nearly optimal. For example, suppose that we have assumed $a = 0.01$, but that this price elasticity is in fact 10% higher. We will set our production levels using Eq. (2.5), which means that we will produce 11% too many 19–inch sets and around 3% too few 21–inch sets, compared to the optimal solution given by Eq. (2.9) with $a = 0.011$. Also, our profits will be 4% lower than expected. But what have we actually lost by applying the results of our model? Using Eq. (2.5) with $a = 0.011$, we will net a \$531,219 profit. The optimal profit would be \$533,514 (set $a = 0.011$ in Eq. (2.9) and substitute back

into Eq. (2.7)). Hence, we have lost only 0.43 percent of the potential maximum profit by applying the results of our model, even though our actual production levels were quite a way off from the optimum values. Our model appears to be extremely robust in this regard. Furthermore, a similar conclusion should hold for many similar problems, since it is basically due to the fact that $\nabla f = 0$ at a critical point.

All of the previous sensitivity analysis calculations could also be performed using a computer algebra system. In fact, this is the preferred method, assuming that one is available. Figure 2.8 illustrates how the computer algebra system Maple can be used to compute the sensitivity $S(x_1, a)$. The calculations of the other sensitivities are similar.

Sensitivity analysis for the other elasticities could be carried out in the same manner. While the particulars will differ, the form of the function f suggests that each affects y in essentially the same manner. In particular, we have a high degree of confidence that our model will lead to a good (nearly optimal) decision about production levels even in the presence of small errors in the estimation of price elasticities.

We will say just a few words on the more general subject of robustness. Our model is based on a linear price structure. Certainly, this is only an approximation. However, in practical applications we are likely to proceed as follows. We begin with an educated guess about the size of the market for our new products and with a reasonable average sale price. Then we estimate elasticities either on the basis of past experience with similar situations or on the basis of limited marketing studies. We should be able to get reasonable estimates for these elasticities over a certain range of sales levels. This range presumably includes the optimal levels. So in effect we are simply making a linear approximation of a nonlinear function over a fairly small region. This sort of approximation is well known to exhibit robustness. After all, this is the whole idea behind calculus.

2.2 Lagrange Multipliers

In this section we begin to consider optimization problems with a more sophisticated structure. As we noted at the beginning of the previous section, complications arise in the solution of multivariable optimization models when the set over which we optimize becomes more complex. In real problems we are led to consider these more complicated models by the existence of constraints on the independent variables.

Example 2.2. We reconsider the color TV problem (Example 2.1) introduced in the previous section. There we assumed that the company has the potential to produce any number of TV sets per year. Now we will introduce constraints based on the available production capacity. Consideration of these two new products came about because the company plans to discontinue manufacture of some older models, thus providing excess capacity at its assembly plant. This excess capacity could be used to increase production of other existing product lines, but the company feels that the new products will be more profitable. It

Variables: s = number of 19–inch sets sold (per year)
t = number of 21–inch sets sold (per year)
p = selling price for a 19–inch set (\$)
q = selling price for a 21–inch set (\$)
C = cost of manufacturing sets (\$/year)
R = revenue from the sale of sets (\$/year)
P = profit from the sale of sets (\$/year)

Assumptions: $p = 339 - 0.01s - 0.003t$
$q = 399 - 0.004s - 0.01t$
$R = ps + qt$
$C = 400,000 + 195s + 225t$
$P = R - C$
$s \leq 5000$
$t \leq 8000$
$s + t \leq 10,000$
$s \geq 0$
$t \geq 0$

Objective: Maximize P

Figure 2.9: Results of step 1 for the color TV problem with constraints.

is estimated that the available production capacity will be sufficient to produce 10,000 sets per year ($\approx$ 200 per week). The company has an ample supply of 19–inch and 21–inch LCD panels and other standard components; however, the circuit boards necessary for constructing the sets are currently in short supply. Also, the 19–inch TV requires a different board than the 21–inch model because of the internal configuration, which cannot be changed without a major redesign, which the company is not prepared to undertake at this time. The supplier is able to supply 8,000 boards per year for the 21–inch model and 5,000 boards per year for the 19–inch model. Taking this information into account, how should the company set production levels?

Once again we will employ the five-step method. The results of step 1 are shown in Figure 2.9. The only change is the addition of several constraints on the decision variables s and t. Step 2 is to select the modeling approach.

This problem will be modeled as a multivariable constrained optimization problem and solved using the method of Lagrange multipliers.

We are given a function $y = f(x_1, \ldots, x_n)$ and a set of constraints. For the moment we will assume that these constraints can

be expressed in the form of k functional equations

$$g_1(x_1, \ldots, x_n) = c_1$$
$$g_2(x_1, \ldots, x_n) = c_2$$
$$\vdots$$
$$g_k(x_1, \ldots, x_n) = c_k.$$

Later on we will explain how to handle inequality constraints. Our job is to optimize

$$y = f(x_1, \ldots, x_n)$$

over the set

$$S = \{(x_1, \ldots, x_n) : \ g_i(x_1, \ldots, x_n) = c_i \text{ for all } i = 1, \ldots, k\}.$$

There is a theorem that states that at an extreme point $x \in S$, we must have

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k.$$

We call $\lambda_1, \ldots, \lambda_k$ the *Lagrange multipliers*. This theorem assumes that $\nabla g_1, \ldots, \nabla g_k$ are linearly independent vectors (see Edwards (1973), p. 113). Then in order to locate the max–min points of f on the set S, we must solve the n Lagrange multiplier equations

$$\frac{\partial f}{\partial x_1} = \lambda_1 \frac{\partial g_1}{\partial x_1} + \cdots + \lambda_k \frac{\partial g_k}{\partial x_1}$$
$$\vdots$$
$$\frac{\partial f}{\partial x_n} = \lambda_1 \frac{\partial g_1}{\partial x_n} + \cdots + \lambda_k \frac{\partial g_k}{\partial x_n}$$

together with the k constraint equations

$$g_1(x_1, \ldots, x_n) = c_1$$
$$\vdots$$
$$g_k(x_1, \ldots, x_n) = c_k$$

for the variables $x_1, \ldots, x_n$ and $\lambda_1, \ldots, \lambda_k$. We must also check any exceptional points at which the gradient vectors $\nabla g_1, \ldots, \nabla g_k$ are not linearly independent.

The method of Lagrange multipliers is based on a geometrical interpretation of the gradient vector. Suppose for the moment that there is only one constraint equation,

$$g(x_1, \ldots, x_n) = c,$$

so that the Lagrange multiplier equation becomes

$$\nabla f = \lambda \nabla g.$$

The set $g = c$ is a curved surface of dimension $n - 1$ in $\mathbb{R}^n$, and for any point $x \in S$ the gradient vector $\nabla g(x)$ is perpendicular to S at that point. The gradient vector ∇f always points in the direction in which f increases the fastest. At a local max or min, the direction in which f increases fastest must also be perpendicular to S, so at that point we must have ∇f and ∇g pointing along the same line; i.e., $\nabla f = \lambda \nabla g$.

In the case of several constraints, the geometrical argument is similar. Now the set S represents the intersection of the k level surfaces $g_1 = c_1, \ldots, g_k = c_k$. Each one of these is an $(n - 1)$–dimensional subset of $\mathbb{R}^n$, so their intersection is an $(n - k)$–dimensional subset. At an extreme point, ∇f must be perpendicular to the set S. Therefore it must lie in the space spanned by the k vectors $\nabla g_1, \ldots, \nabla g_k$. The technical condition of linear independence ensures that the k vectors $\nabla g_1, \ldots, \nabla g_k$ actually do span a k–dimensional space. (In the case of a single constraint, linear independence simply means that $\nabla g \neq 0$.)

Example 2.3. Maximize $x + 2y + 3z$ over the set $x^2 + y^2 + z^2 = 3$.

This is a constrained multivariable optimization problem. Let

$$f(x, y, z) = x + 2y + 3z$$

denote the objective function, and let

$$g(x, y, z) = x^2 + y^2 + z^2$$

denote the constraint function. Compute

$$\nabla f = (1, 2, 3)$$
$$\nabla g = (2x, 2y, 2z).$$

At the maximum, $\nabla f = \lambda \nabla g$; in other words,

$$1 = 2x\lambda$$
$$2 = 2y\lambda$$
$$3 = 2z\lambda.$$

This gives three equations in four unknowns. Solving in terms of λ, we obtain

$$x = 1/2\lambda$$
$$y = 1/\lambda$$
$$z = 3/2\lambda.$$

Using the fact that
$$x^2 + y^2 + z^2 = 3,$$

we obtain a quadratic equation in λ, with two real roots. The root $\lambda = \sqrt{42}/6$ leads to

$$x = \frac{1}{2\lambda} = \frac{\sqrt{42}}{14}$$
$$y = \frac{1}{\lambda} = \frac{\sqrt{42}}{7}$$
$$z = \frac{3}{2\lambda} = \frac{3\sqrt{42}}{14},$$

so that the point

$$a = \left(\frac{\sqrt{42}}{14}, \frac{\sqrt{42}}{7}, \frac{3\sqrt{42}}{14} \right)$$

is one candidate for the maximum. The other root, $\lambda = -\sqrt{42}/6$, leads to another candidate, $b = -a$. Since $\nabla g \neq 0$ everywhere on the constraint set $g = 3$, a and b are the only two candidates for the maximum. Since f is a continuous function on the closed and bounded set $g = 3$, f must attain its maximum and minimum on this set. Then, since

$$f(a) = \sqrt{42}, \qquad \text{and} \qquad f(b) = -\sqrt{42},$$

the point a is the maximum and b is the minimum. Consider the geometry of this example. The constraint set S defined by the equation
$$x^2 + y^2 + z^2 = 3$$

is a sphere of radius $\sqrt{3}$ centered at the origin in $\mathbb{R}^3$. Level sets of the objective function

$$f(x, y, z) = x + 2y + 3z$$

are planes in R^3. The points a and b are the only two points on the sphere S at which one of these planes is tangent to the sphere. At the maximum point a, the gradient vectors ∇f and ∇g point in the same direction. At the minimum point b, ∇f and ∇g point in opposite directions.

Example 2.4. Maximize $x + 2y + 3z$ over the set $x^2 + y^2 + z^2 = 3$ and $x = 1$.

The objective function is

$$f(x, y, z) = x + 2y + 3z,$$

so
$$\nabla f = (1,\, 2,\, 3).$$

The constraint functions are
$$g_1(x,\, y,\, z) = x^2 + y^2 + z^2$$
$$g_2(x,\, y,\, z) = x.$$

Compute
$$\nabla g_1 = (2x,\, 2y,\, 2z)$$
$$\nabla g_2 = (1,\, 0,\, 0).$$

Then the Lagrange multiplier formula $\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$ yields
$$1 = 2x\lambda_1 + \lambda_2$$
$$2 = 2y\lambda_1$$
$$3 = 2z\lambda_1.$$

Solving for x, y, and z in terms of λ_1 and λ_2 gives
$$x = \frac{1 - \lambda_2}{2\lambda_1}$$
$$y = \frac{2}{2\lambda_1}$$
$$z = \frac{3}{2\lambda_1}.$$

Substituting into the constraint equation $x = 1$ gives $\lambda_2 = 1 - 2\lambda_1$. Substituting all of this into the remaining equation
$$x^2 + y^2 + z^2 = 3$$

yields a quadratic equation in λ_1, which gives $\lambda_1 = \pm\sqrt{26}/4$. Substituting back into the equations for x, y, and z yields the two following solutions:
$$c = \left(1,\, \frac{2\sqrt{26}}{13},\, \frac{3\sqrt{26}}{13}\right)$$
$$d = \left(1,\, \frac{-2\sqrt{26}}{13},\, \frac{-3\sqrt{26}}{13}\right).$$

Since the two gradient vectors ∇g_1 and ∇g_2 are linearly independent everywhere on the constraint set, the points c and d are the only candidates for a maximum. Since f must attain its maximum on this closed and bounded set, we need only evaluate f at each candidate point to find the maximum. The maximum is
$$f(c) = 1 + \sqrt{26},$$

and the point d is the location of the minimum. The constraint set S in this example is a circle in $\mathbb{R}^3$ formed by the intersection of the sphere

$$x^2 + y^2 + z^2 = 3$$

and the plane $x = 1$. As before, the level sets of the function f are planes in $\mathbb{R}^3$. At the points c and d these planes are tangent to the circle S.

Inequality constraints can be handled by a combination of the Lagrange multiplier technique and the techniques for unconstrained problems. Suppose that the problem in Example 2.4 is altered by replacing the $x = 1$ constraint with the inequality constraint $x \geq 1$. We can consider the set

$$S = \{(x,\, y,\, z) : x^2 + y^2 + z^2 = 3,\ x \geq 1\}$$

as the union of two components. The maximum over the first component

$$S_1 = \{(x,\, y,\, z) : x^2 + y^2 + z^2 = 3,\ x = 1\}$$

was found to occur at the point

$$c = \left(1,\ \sqrt{\frac{8}{13}},\ 1.5\sqrt{\frac{8}{13}}\right)$$

in our previous analysis, and we can calculate that

$$f(x,\, y,\, z) = 1 + 6.5\sqrt{\frac{8}{13}} = 6.01$$

at this point. To consider the remaining part

$$S_2 = \{(x,\, y,\, z) : x^2 + y^2 + z^2 = 3,\ x > 1\},$$

we return to our analysis from Example 2.3, noting that there is no local maximum of f anywhere on this set. Therefore, the maximum of f on S_1 must be the maximum of the function f on the set S. If we had considered the maximum of f over the set

$$S = \{(x, y, z) : x^2 + y^2 + z^2 = 3,\ x \leq 1\},$$

then the maximum would be at the point

$$a = \left(\frac{1}{2},\ \frac{2}{2},\ \frac{3}{2}\right) \cdot \sqrt{\frac{6}{7}}$$

found in our analysis of Example 2.3.

Returning now to the problem introduced at the beginning of this section, we are ready to continue the modeling process with step 3. We will formulate the revised color TV problem as a constrained multivariable optimization problem. We wish to maximize $y = P$ (profit) as a function of our two decision variables, $x_1 = s$ and $x_2 = t$. We have the same objective function

$$
\begin{aligned}
y &= f(x_1,\, x_2)\\
&= (339 - 0.01x_1 - 0.003x_2)x_1 + (399 - 0.004x_1 - 0.01x_2)x_2\\
&\quad - (400{,}000 + 195x_1 + 225x_2).
\end{aligned}
$$

We wish to maximize f over the set S consisting of all x_1 and x_2 satisfying the constraints

$$
\begin{aligned}
x_1 &\le 5{,}000\\
x_2 &\le 8{,}000\\
x_1 + x_2 &\le 10{,}000\\
x_1 &\ge 0\\
x_2 &\ge 0.
\end{aligned}
$$

The set S is called the *feasible region* because it represents the set of all feasible production levels. Figure 2.10 shows a graph of the feasible region for this problem.

We will apply Lagrange multiplier methods to find the maximum of $y = f(x_1,\, x_2)$ over the set S. Compute

$$
\nabla f = (144 - 0.02x_1 - 0.007x_2,\ 174 - 0.007x_1 - 0.02x_2).
$$

Since $\nabla f \ne 0$ in the interior of S, the maximum must occur on the boundary. Consider first the segment of the boundary on the constraint line

$$
g(x_1,\, x_2) = x_1 + x_2 = 10{,}000.
$$

Here $\nabla g = (1,\, 1)$, so the Lagrange multiplier equations are

$$
\begin{aligned}
144 - 0.02x_1 - 0.007x_2 &= \lambda\\
174 - 0.007x_1 - 0.02x_2 &= \lambda.
\end{aligned}
\tag{2.12}
$$

Solving these two equations together with the constraint equation

$$
x_1 + x_2 = 10{,}000
$$

yields

$$
\begin{aligned}
x_1 &= \frac{50{,}000}{13} \approx 3{,}846\\
x_2 &= \frac{80{,}000}{13} \approx 6{,}154\\
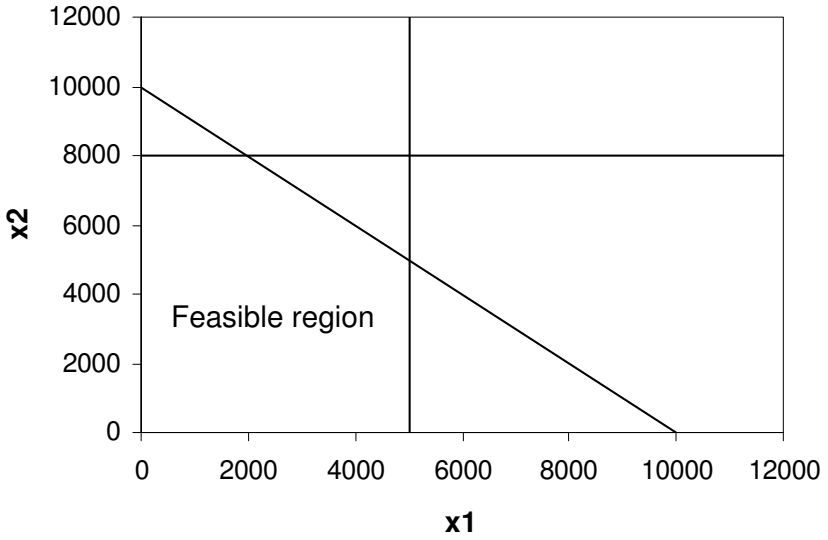\lambda &= 24.
\end{aligned}
$$

Figure 2.10: Graph showing the set of all feasible production levels x_1 of 19–inch sets and x_2 of 21–inch sets for the color TV problem with constraints.
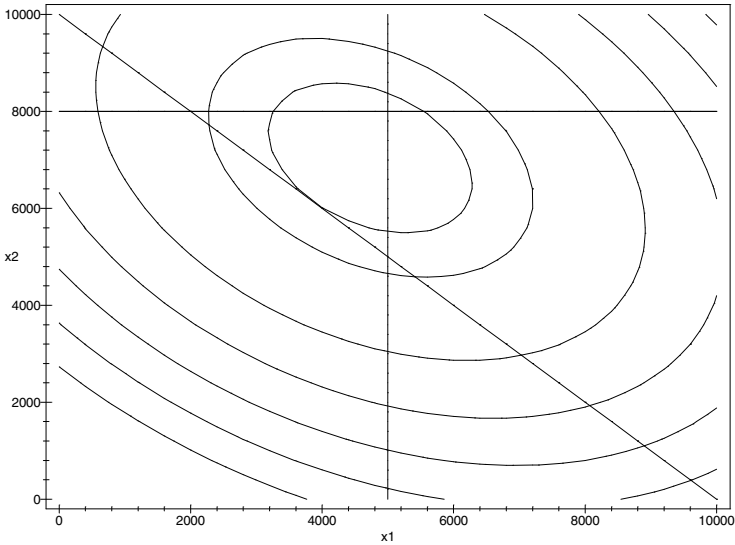

Figure 2.11: Graph showing level sets of profit $y = f(x_1, x_2)$ versus production levels x_1 of 19–inch sets and x_2 of 21–inch sets together with the set of all feasible production levels for the color TV problem with constraints.

Substituting back into Eq. (2.2), we obtain $y = 532,308$ at the maximum.

Figure 2.11 shows a Maple graph of the level sets of f together with the feasible region.

The level sets $f = C$ for $C = 0,\ 100,000,\ \ldots,\ 500,000$ form smaller and smaller concentric rings, all of which intersect the feasible region. The level set $f = 532,308$ forms the smallest ring. This set barely touches the feasible region S, and is tangent to the line $x_1 + x_2 = 10,000$ at the optimum point. This graphical evidence indicates that the critical point found by using Lagrange multipliers along the constraint line $x_1 + x_2 = 10,000$ is actually the maximum of the function f over the feasible region S.

In[1]:= **y = (339 − x1 / 100 − 3 x2 / 1000) x1 +**
             (399 − 4 x1 / 1000 − x2 / 100) x2 −
             (400000 + 195 x1 + 225 x2)

Out[1]= $\left(-\dfrac{x1}{100} - \dfrac{3\,x2}{1000} + 339\right)x1 - 195\,x1 + \left(-\dfrac{x1}{250} - \dfrac{x2}{100} + 399\right)x2 - 225\,x2 - 400000$

In[2]:= **dydx1 = D[y, x1]**

Out[2]= $-\dfrac{x1}{50} - \dfrac{7\,x2}{1000} + 144$

In[3]:= **dydx2 = D[y, x2]**

Out[3]= $-\dfrac{7\,x1}{1000} - \dfrac{x2}{50} + 174$

In[4]:= **s = Solve[{dydx1 == lambda, dydx2 == lambda, x1 + x2 == 10000}, {x1, x2, lambda}]**

Out[4]= $\left\{\left\{x1 \to \dfrac{50000}{13},\ x2 \to \dfrac{80000}{13},\ \text{lambda} \to 24\right\}\right\}$

In[5]:= **N[%]**

Out[5]= {{x1 → 3846.15, x2 → 6153.85, lambda → 24.}}

In[6]:= **y /. %**

Out[6]= {532308.}

Figure 2.12: Optimal solution to the color TV problem with constraints using the computer algebra system Mathematica.

An algebraic proof that this point is actually the maximum is a bit more complicated. By comparing values of f at this critical point with values at the endpoints $(5,000,\ 5,000)$ and $(2,000,\ 8,000)$, we can show that this critical point is the maximum over this line segment. Then we can optimize f over the remaining line segments and compare results. For example, the maximum of f over the line segment along the x_1 axis occurs at $x_1 = 5,000$. To see this,

apply Lagrange multipliers with $g(x_1, x_2) = x_2 = 0$. Here $\nabla g = (0, 1)$, so the Lagrange multiplier equations are

$$144 - 0.02x_1 - 0.007x_2 = 0$$
$$174 - 0.007x_1 - 0.02x_2 = \lambda.$$

Solving these two equations together with the constraint equation $x_2 = 0$ yields $x_1 = 7,200$, $x_2 = 0$, and $\lambda = 123.6$. This is outside the feasible region, so the max–min along this segment must occur at the endpoints $(0, 0)$ and $(5,000, 0)$. The first is the minimum and the second is the maximum, since the value of f at the second is greater. It is also possible to optimize along this line segment by substituting $x_2 = 0$ and using one variable methods. Since the largest value of f occurs on the slanted line segment, we have found the maximum over S. Some of the calculations in step 4 are rather involved. In such cases it is appropriate to use a computer algebra system to simplify the process of computing derivatives and solving equations. Figure 2.12 shows the results of using the computer algebra system Mathematica to perform the calculations of step 4 for the constraint line $x_1 + x_2 = 10,000$.

In plain English, the company can maximize profits by producing 3,846 of the 19–inch sets and 6,154 of the 21–inch sets for a total of 10,000 sets per year. This level of production uses all of the available excess production capacity. The resource constraints on the availability of TV circuit boards are not binding. This venture will produce an estimated profit of \$532,308 annually.

2.3   Sensitivity Analysis and Shadow Prices

In this section we discuss some of the specialized techniques for sensitivity analysis in Lagrange multiplier models. It turns out that the multipliers themselves have a real–world significance.

Before we report the results of our model analysis in Example 2.2, it is important to perform sensitivity analysis. At the end of Section 2.1 we investigated the sensitivity to price elasticity for a model without constraints. The procedure for our new model is not much different. We examine the sensitivity to a particular parameter value by generalizing the model slightly, replacing the assumed value with a variable. Suppose we want to look again at the price elasticity, a, for 19–inch sets. We rewrite the objective function as in Eq. (2.7) so that

$$\nabla f = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2} \right),$$

where $\partial f / \partial x_1$ and $\partial f / \partial x_2$ are given by Eq. (2.8). Now the Lagrange multiplier equations are

$$144 - 2ax_1 - 0.007x_2 = \lambda$$
$$174 - 0.007x_1 - 0.02x_2 = \lambda. \tag{2.13}$$

Solving together with the constraint equation

$$g(x_1, x_2) = x_1 + x_2 = 10,000$$

yields

$$x_1 = \frac{50,000}{1,000a + 3}$$

$$x_2 = 10,000 - \frac{50,000}{1,000a + 3} \tag{2.14}$$

$$\lambda = \frac{650}{1,000a + 3} - 26.$$

Then we have

$$\frac{dx_1}{da} = \frac{-50,000,000}{(1,000a + 3)^2}$$

$$\frac{dx_2}{da} = \frac{-dx_1}{da}, \tag{2.15}$$

so that at the point $x_1 = 3,846$, $x_2 = 6,154$, $a = 0.01$, we obtain

$$S(x_1, a) = \frac{dx_1}{da} \cdot \frac{a}{x_1} = -0.77$$

$$S(x_2, a) = \frac{dx_2}{da} \cdot \frac{a}{x_2} = 0.48.$$

See Figures 2.13 and 2.14 for the graphs of x_1 and x_2 versus a in this case. If the price elasticity of 19–inch sets increases, we will shift production from 19–inch to 21–inch sets. If it decreases, then we will produce more 19–inch sets and fewer 21–inch sets.

In any case, as long as the point (x_1, x_2) given by Eq. (2.14) lies between the other constraint lines ($.007 \leq a \leq .022$), we will always produce a total of 10,000 sets.

Now let us consider the sensitivity of our optimal profit y to the price elasticity a for 19–inch sets. To obtain a formula for y in terms of a, we substitute Eq. (2.14) back into Eq. (2.2) to get

$$y = \left[339 - a\left(\frac{50,000}{1,000a + 3}\right) - 0.003\left(10,000 - \frac{50,000}{1,000a + 3}\right)\right]\left(\frac{50,000}{1,000a + 3}\right)$$

$$+ \left[399 - 0.004\left(\frac{50,000}{1,000a + 3}\right) - 0.01\left(10,000 - \frac{50,000}{1,000a + 3}\right)\right]$$

$$\times \left(1,000 - \frac{50,000}{1,000a + 3}\right)$$

$$- \left[400,000 + 195\left(\frac{50,000}{1,000a + 3}\right) + 225\left(10,000 - \frac{50,000}{1,000a + 3}\right)\right].$$

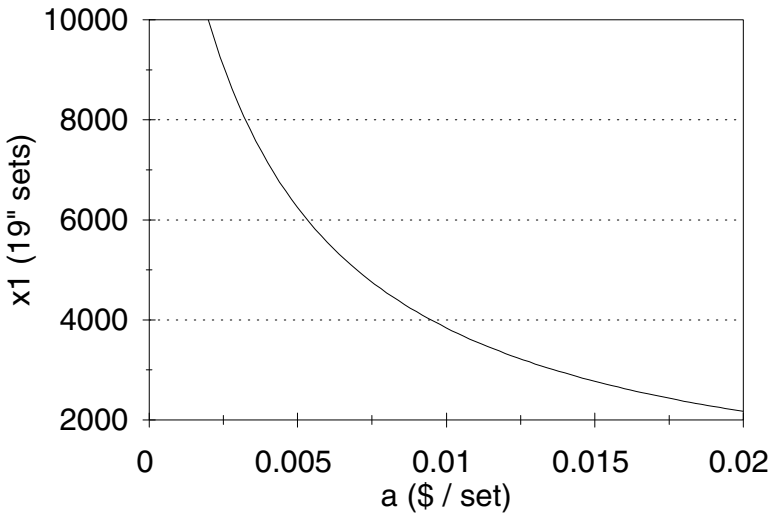See Figure 2.15 for the graph of y versus a.

Figure 2.13: Graph of optimal production level x_1 of 19–inch sets versus price elasticity a for the color TV problem with constraints.
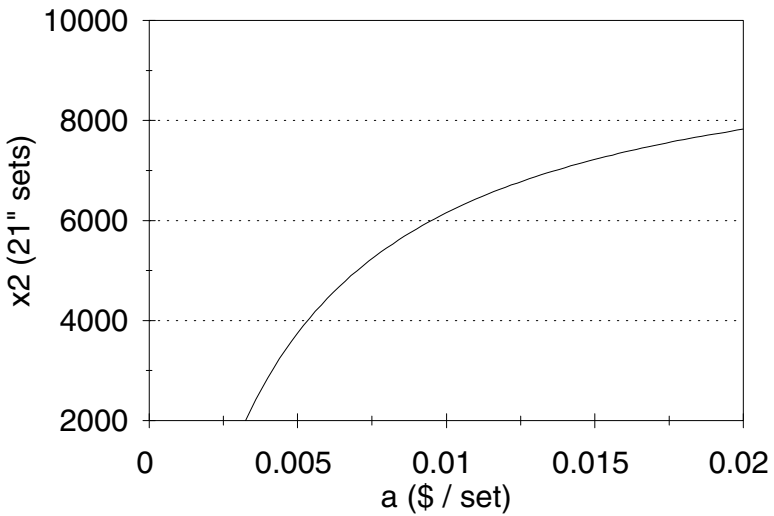

Figure 2.14: Graph of optimal production level x_2 of 21–inch sets versus price elasticity a for the color TV problem with constraints.
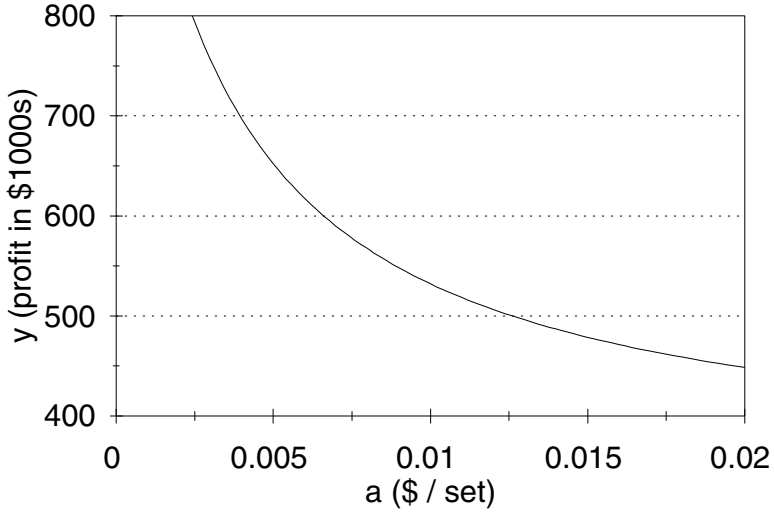
Figure 2.15: Graph of optimum profit y versus price elasticity a for the color TV problem with constraints.

In order to get a numerical measure of the sensitivity of y to a, we could apply one–variable techniques to the above expression, perhaps with the aid of a computer algebra system. Another method, which is much more efficient computationally, is to use the multivariable chain rule in Eq. (2.11). For any a, the gradient vector ∇f is perpendicular to the constraint line $g = 10,000$. Since

$$x(a) = (x_1(a),\, x_2(a))$$

is a point on the curve $g = 10,000$, the velocity vector

$$\frac{dx}{da} = \left(\frac{dx_1}{da},\, \frac{dx_2}{da} \right)$$

is tangent to the curve. But then ∇f is perpendicular to dx/da; i.e., the dot product

$$\nabla f \cdot \frac{dx}{da} = \left(\frac{\partial y}{\partial x_1},\, \frac{\partial y}{\partial x_2} \right) \cdot \left(\frac{dx_1}{da},\, \frac{dx_2}{da} \right)$$
$$= \frac{\partial y}{\partial x_1} \frac{dx_1}{da} + \frac{\partial y}{\partial x_2} \frac{dx_2}{da} = 0$$

in general. Therefore, we again obtain

$$\frac{dy}{da} = \frac{\partial y}{\partial a} = -x^2$$

as in Section 2.1. Now we can easily compute that

$$S(y,\, a) = \frac{dy}{da} \cdot \frac{a}{y}$$
$$= -(3,846)^2 \frac{0.01}{532,308}$$
$$= -0.28.$$

As in the unconstrained problem, an increase in price elasticity leads to lost profits. It is also true here that, as in the unconstrained problem, almost all of the lost profit is due to the fact that the selling price for 19–inch sets has decreased. If $a = 0.011$ and we use $x_1 = 3,846$, $x_2 = 6,154$ instead of the new optimum given by Eq. (2.13), we will not lose much potential profit. The gradient vector ∇f points in the direction of fastest increase of the objective function f, which represents our profits. We are not at the optimum, but the path between the optimum and the point (3,846, 6,154) is perpendicular to ∇f. Therefore, we may expect that the value of f at this point does not vary much from the optimum value. Hence our model leads to a nearly optimal decision even in the presence of small variations in a.

We remark that for this problem it would also make sense to use a computer algebra system to perform the necessary calculations. Figure 2.16 illustrates the use of the computer algebra system Maple to compute the sensitivity $S(x_2,\, a)$. The other sensitivities can be computed in a similar manner.

We will now consider the sensitivity of the optimal production levels x_1 and x_2 and the resulting profit y to the available manufacturing capacity of $c = 10,000$ sets per year. In order to do this we will return to our original problem, replacing the constraint $g = 10,000$ by the more general form $g = c$. The feasible region is similar to that pictured in Fig. 2.10, but now the slanted constraint line is moved a bit (remaining parallel to the line $x_1 + x_2 = 10,000$). For values of c near $10,000$, the maximum still occurs at the point on the constraint line

$$g(x_1,\, x_2) = x_1 + x_2 = c \qquad (2.16)$$

where $\nabla f = \lambda \nabla g$. Since both ∇f and ∇g are unchanged from our original problem, we have the same Lagrange multiplier equations from Eq. (2.12), which are to be solved along with the new constraint equation, Eq. (2.16). Solving, we obtain

$$x_1 = \frac{13c - 30,000}{26}$$
$$x_2 = \frac{13c + 30,000}{26} \qquad (2.17)$$
$$\lambda = \frac{3(106,000 - 9c)}{2,000}.$$

```
> y:=(339-a*x1-3*x2/1000)*x1
>    +(399-4*x1/1000-x2/100)*x2-(400000+195*x1+225*x2);
```

$$y := \left(339 - a\,x1 - \frac{3}{1000}\,x2\right)x1 + \left(399 - \frac{1}{250}\,x1 - \frac{1}{100}\,x2\right)x2 - 400000 - 195\,x1 - 225\,x2$$

```
> dydx1:=diff(y,x1);
```

$$dydx1 := -2\,a\,x1 + 144 - \frac{7}{1000}\,x2$$

```
> dydx2:=diff(y,x2);
```

$$dydx2 := -\frac{7}{1000}\,x1 - \frac{1}{50}\,x2 + 174$$

```
> s:=solve({dydx1=lambda,dydx2=lambda,x1+x2=10000},{x1,x2,lambda});
```

$$s := \{\, x1 = \frac{50000}{1000\,a + 3},\, x2 = 20000\,\frac{500\,a - 1}{1000\,a + 3},\, \lambda = -52\,\frac{500\,a - 11}{1000\,a + 3}\,\}$$

```
> assign(s);
> dx2da:=diff(x2,a);
```

$$dx2da := \frac{10000000}{1000\,a + 3} - 20000000\,\frac{500\,a - 1}{(1000\,a + 3)^2}$$

```
> assign(a=1/100);
> sx2a:=dx2da*(a/x2);
```

$$sx2a := \frac{25}{52}$$

```
> evalf(sx2a);
```

$$.4807692308$$

Figure 2.16: Calculation of the sensitivity $S(x_2,\, a)$ for the color TV problem with constraints using the computer algebra system Maple.

Now

$$\frac{dx_1}{dc} = \frac{1}{2}$$
$$\frac{dx_2}{dc} = \frac{1}{2}. \tag{2.18}$$

There is a simple geometric explanation for Eq. (2.18). Since ∇f points in the direction of the fastest increase of f, when we move the constraint line in Eq. (2.16), the new optimum (x_1, x_2) should be located at approximately the point where ∇f intersects the line in Eq. (2.16). Then

$$S(x_1,\, c) = \frac{1}{2} \cdot \frac{10,000}{3,846} \approx 1.3$$
$$S(x_2,\, c) = \frac{1}{2} \cdot \frac{10,000}{6,154} \approx 0.8.$$

To obtain the sensitivity of y to c, we compute

$$\frac{dy}{dc} = \frac{\partial y}{\partial x_1}\frac{dx_1}{dc} + \frac{\partial y}{\partial x_2}\frac{dx_2}{dc}$$
$$= (24)\left(\frac{1}{2}\right) + (24)\left(\frac{1}{2}\right)$$
$$= 24,$$

which is the value of the Lagrange multiplier λ. Now

$$S(y,\, c) = (24)\left(\frac{10,000}{532,308}\right) \approx 0.45.$$

The geometric explanation for $dy/dc = \lambda$ is as follows. We have $\nabla f = \lambda \nabla g$, and as we increase c, we move out in the direction ∇f. As we move in this direction, f increases λ times as fast as g.

The derivative $dy/dc = 24$ has an important real–world interpretation. The addition of 1 unit of production capacity $\Delta c = 1$ results in an increased profit $\Delta y = 24$ dollars. This is called a *shadow price*. It represents the value to the company of a certain resource (production capacity). If the company is interested in the possibility of increasing production capacity, which is, after all, the binding constraint, it should be willing to pay up to \$24 per unit of added capacity. Alternatively, it would be worthwhile to transfer production capacity from the manufacture of 19–inch and 21–inch LCD flat panel TV sets to an alternative product if and only if the new product would yield a profit greater than \$24 per unit.

The calculation of sensitivities in this problem can be simplified by using a computer algebra system. Figure 2.17 illustrates the use of the computer algebra system Maple to compute the sensitivity $S(y,\, c)$.

The other sensitivities can be computed in a similar manner. If you have the good fortune to have access to a computer algebra system, you should use it for your own work. Real–world problems often involve lengthy computations. Some facility in the use of a computer algebra system will make you more productive. It is also a lot more fun than calculation by hand.

Of course, the optimal level of profit y and the production levels x_1 and x_2 are totally insensitive to the values of the other constraint coefficients, since the other constraints, $x_1 \leq 5,000$ and $x_2 \leq 8,000$, are not binding. A small change in the upper bounds on x_1 or x_2 would change the feasible region, but the optimal solution would remain at (3,846, 6,154). Thus, the shadow prices for these resources are zero. The company would be unwilling to pay a premium to increase the available number of TV circuit boards, since they do not need them. This situation would not change unless the number of boards available was reduced to 3,846 or less for 19–inch sets, or to 6,154 or less for 21–inch sets. In the next example, we will consider what happens in this case.

Example 2.5. Suppose that in the constrained color TV problem of Example 2.2, the number of circuit boards available for 19–inch TVs is only 3,000 per year. What is the optimum production schedule?

```
> y:=(339-x1/100-3*x2/1000)*x1
>    +(399-4*x1/1000-x2/100)*x2-(400000+195*x1+225*x2);
```
$$y := \left(339 - \frac{1}{100}x1 - \frac{3}{1000}x2\right)x1 + \left(399 - \frac{1}{250}x1 - \frac{1}{100}x2\right)x2 - 400000 - 195\,x1 - 225\,x2$$

```
> dydx1:=diff(y,x1);
```
$$dydx1 := -\frac{1}{50}x1 + 144 - \frac{7}{1000}x2$$

```
> dydx2:=diff(y,x2);
```
$$dydx2 := -\frac{7}{1000}x1 - \frac{1}{50}x2 + 174$$

```
> s:=solve({dydx1=lambda,dydx2=lambda,x1+x2=c},{x1,x2,lambda});
```
$$s := \{\lambda = -\frac{27}{2000}c + 159,\, x1 = \frac{1}{2}c - \frac{15000}{13},\, x2 = \frac{1}{2}c + \frac{15000}{13}\}$$

```
> assign(s);
> dydc:=diff(y,c);
```
$$dydc := -\frac{27}{2000}c + 159$$

```
> assign(c=10000);
> dydc;
```
$$24$$

```
> syc:=dydc*(c/y);
```
$$syc := \frac{78}{173}$$

```
> evalf(syc);
```
$$.4508670520$$

Figure 2.17: Calculation of the sensitivity $S(y, c)$ for the color TV problem with constraints using the computer algebra system Maple.

In this case the point on the level curve $x_1 + x_2 = 10,000$ at which $f(x_1, x_2)$ is maximized occurs outside of the feasible region. The maximum of f on the feasible region occurs at the point (3,000, 7,000). This is the intersection of the constraint curves

$$g_1(x_1,\, x_2) = x_1 + x_2 = 10,000$$
$$g_2(x_1,\, x_2) = x_1 = 3,000.$$

At this point we have
$$\nabla f = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2.$$

In fact, we can easily compute

$$\nabla f = (35,\, 13)$$
$$\nabla g_1 = (1,\, 1)$$
$$\nabla g_2 = (1,\, 0)$$

at the point (3,000, 7,000); thus we have $\lambda_1 = 13$ and $\lambda_2 = 22$. Of course any vector in $\mathbb{R}^2$ can be written as a linear combination of $(1, 1)$ and $(1, 0)$. The point, however, of computing the Lagrange multipliers, is that even in the case of multiple constraints, they still represent the shadow prices for the binding constraints (production capacity and 19–inch boards). In other words, an additional unit of production capacity is worth \$13, and an additional circuit board is worth \$22.

For the convenience of the reader, we include here a proof of the fact that the Lagrange multipliers represent shadow prices. We are given a function $y = f(x_1, \ldots, x_n)$, which is to be optimized over the set defined by one or more constraint equations of the form

$$g_1(x_1, \ldots, x_n) = c_1$$
$$g_2(x_1, \ldots, x_n) = c_2$$
$$\vdots$$
$$g_k(x_1, \ldots, x_n) = c_k.$$

Suppose that the optimum occurs at a point x_0 at which the hypotheses of the Lagrange multiplier theorem are satisfied, so that at this point we have

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k. \tag{2.19}$$

Since the constraint equations can be written in any order, it will suffice to show that λ_1 is the shadow price corresponding to the first constraint. Let $x(t)$ denote the location of the optimum point over the set $g_1 = t$, $g_2 = c_2$, $\ldots$, $g_k = c_k$. Since $g_1(x(t)) = t$, we have $\nabla g(x(t)) \cdot x'(t) = 1$, and in particular $\nabla g(x_0) \cdot x'(c_1) = 1$. Since for $i = 2, \ldots, k$ we have $g_i(x(t)) = c_i$ constant for all t, we have $\nabla g_i(x(t)) \cdot x'(t) = 0$, and in particular $\nabla g_i(x_0) \cdot x'(c_1) = 0$. The shadow price is

$$\frac{d(f(x(t)))}{dt} = \nabla f(x(t)) \cdot x'(t)$$

evaluated at the point $t = c_1$. Then, from the fact that Eq. (2.19) holds at the point x_0, we obtain $\nabla f(x_0) \cdot x'(c_1) = \lambda_1 \nabla g_1(x_0) \cdot x'(c_1) = \lambda_1$ as desired.

2.4   Exercises

1. Ecologists use the following model to represent the growth process of two competing species, x and y:

$$\frac{dx}{dt} = r_1 x \left(1 - \frac{x}{K_1} \right) - \alpha_1 xy$$

$$\frac{dy}{dt} = r_2 y \left(1 - \frac{y}{K_2} \right) - \alpha_2 xy.$$

The variables x and y represent the number in each population; the parameters r_i represent the intrinsic growth rates of each species; K_i represents the maximum sustainable population in the absence of competition; and α_i represents the effects of competition.   Studies of the blue whale and fin whale populations have determined the following parameter values (t in years):

	Blue	Fin
r	0.05	0.08
K	150,000	400,000
α	10^{-8}	10^{-8}

   (a) Determine the population levels x and y that maximize the number of new whales born each year. Use the five-step method, and model as an unconstrained optimization problem.

   (b) Examine the sensitivity of the optimal population levels to the intrinsic growth rates r_1 and r_2.

   (c) Examine the sensitivity of the optimal population levels to the environmental carrying capacities K_1 and K_2.

   (d) Assuming that $\alpha_1 = \alpha_2 = \alpha$, is it ever optimal for one species to become extinct?

2. Reconsider the whale problem of Exercise 1, but now look at the total number of whales. We will say that the whale population levels x and y are feasible provided that both x and y are nonnegative. We will say that the population levels x and y are *sustainable* provided that both of the growth rates dx/dt and dy/dt are nonnegative.

   (a) Determine the population levels that are feasible, sustainable, and that maximize the total whale population $x + y$. Use the five-step method, and model as a constrained optimization problem.

   (b) Examine the sensitivity of the optimal population levels x and y to the intrinsic growth rates r_1 and r_2.

   (c) Examine the sensitivity of the optimal population levels x and y to the environmental carrying capacities K_1 and K_2.

(d) Assuming that $\alpha_1 = \alpha_2 = \alpha$, examine the sensitivity of the optimal population levels x and y to the strength of competition α. Is it ever optimal to drive one species to extinction?

3. Reconsider the whale problem of Exercise 1, but now consider the economic ramifications of harvesting.

   (a) A blue whale carcass is worth \$12,000, and a fin whale carcass is worth about half as much. Assuming that controlled harvesting can be used to maintain x and y at any desired level, what population levels will produce the maximum revenue? (Once population reaches the desired levels, the population levels will be kept constant by harvesting at a rate equal to the growth rate.) Use the five-step method. Model as an unconstrained optimization problem.

   (b) Examine the sensitivity of the optimal population levels x and y to the parameters r_1 and r_2.

   (c) Examine the sensitivity of revenue in \$/year to the parameters r_1 and r_2.

   (d) Assuming $\alpha_1 = \alpha_2 = \alpha$, study the sensitivity of x and y to α. At what point does it become economically optimal to drive a species to extinction?

4. In Exercise 1, suppose that the International Whaling Commission (IWC) has decreed that no population of whales may be sustained at a level less than half of the environmental carrying capacity K.

   (a) Find the population levels that maximize the sustained profit subject to these constraints. Use Lagrange multipliers.

   (b) Examine the sensitivity of the optimal population levels x and y and the sustained profit to the constraint coefficients.

   (c) The IWC feels that enforcement of the minimum population rules is most easily carried out in terms of quotas. Determine the quotas (maximum number of blue whales and fin whales harvested per year) that will have an equivalent effect to the $K/2$ rule.

   (d) The whalers, complaining that IWC quotas cost them a considerable amount of money, have petitioned for them to be relaxed. Analyze the potential effects of increased quotas on both the yearly revenue of the whalers and the population levels of the whales.

5. Consider the color TV problem without constraints (Example 2.1). Because the company's assembly plant is located overseas, the U.S. government has imposed a tariff of \$25 per unit.

   (a) Find the optimal production levels, taking the tariff into consideration. What does the tariff cost the company? How much of this cost is paid directly to the government, and how much represents lost sales?

(b) Would it be worthwhile for the company to relocate production facilities to the U.S. in order to avoid the tariff? Assume that the overseas facility can be leased to another manufacturer for $200,000 per year and that the cost of constructing and operating a new facility in the U.S. would amount to $550,000 annually. The construction costs have been amortized over the expected life of the new facility.

(c) The purpose of the tariff is to motivate manufacturing companies to operate plants in the U.S. What is the minimum tariff that would make it worthwhile for the company to relocate its facility?

(d) Given that the tariff is large enough to motivate the company to move its facility, how important is the actual tariff amount? Consider the sensitivity of both production levels and profit to the amount of the tariff.

6. A manufacturer of personal computers currently sells 10,000 units per month of a basic model. The cost of manufacture is $700/unit, and the wholesale price is $950. During the last quarter the manufacturer lowered the price $100 in a few test markets, and the result was a 50% increase in sales. The company has been advertising its product nationwide at a cost of $50,000 per month. The advertising agency claims that increasing the advertising budget by $10,000/month would result in a sales increase of 200 units/month. Management has agreed to consider an increase in the advertising budget to no more than $100,000/month.

(a) Determine the price and the advertising budget that will maximize profit. Use the five-step method. Model as a constrained optimization problem, and solve using the method of Lagrange multipliers.

(b) Determine the sensitivity of the decision variables (price and advertising) to price elasticity (the 50% number).

(c) Determine the sensitivity of the decision variables to the advertising agency's estimate of 200 new sales each time the advertising budget is increased by $10,000 per month.

(d) What is the value of the multiplier found in part (a)? What is the real–world significance of the multiplier? How could you use this information to convince top management to lift the ceiling on advertising expenditures?

7. A local daily newspaper has recently been acquired by a large media conglomerate. The paper currently sells for $1.50/week and has a circulation of 80,000 subscribers. Advertising sells for $250/page, and the paper currently sells 350 pages/week (50 pages/day). The new management is looking for ways to increase profits. It is estimated that an increase of ten cents/week in the subscription price will cause a drop in circulation of 5,000 subscribers. Increasing the price of advertising by $100/page will cause the paper to lose approximately 50 pages of advertising per week.

The loss of advertising will also affect circulation, since one of the reasons people buy the paper is for the advertisements. It is estimated that a loss of 50 pages of advertisements per week will reduce circulation by 1,000 subscriptions.

(a) Find the weekly subscription price and the advertising price that will maximize profit. Use the five-step method, and model as an unconstrained optimization problem.

(b) Examine the sensitivity of your conclusions in part (a) to the assumption of 5,000 lost sales when the price of the paper increases by ten cents.

(c) Examine the sensitivity of your conclusions in part (a) to the assumption of 50 pages/week of lost advertising sales when the price of advertising is increased by $100/page.

(d) Advertisers who currently place advertisements in the newspaper have the option of using direct mail to reach their customers. Direct mail would cost the equivalent of $500/page of newspaper advertising. How does this information alter your conclusions in part (a)?

8. Reconsider the newspaper problem of Exercise 7, but now suppose that advertisers have the option of using direct mail to reach their customers. Because of this, management has decided not to increase the price of advertising beyond $400/page.

(a) Find the weekly subscription price and the advertising price that will maximize profit. Use the five-step method, and model as a constrained optimization problem. Solve by the method of Lagrange multipliers.

(b) Determine the sensitivity of your decision variables (subscription price and advertising price) to the assumption of 5,000 lost sales when the price of the paper increases by ten cents.

(c) Determine the sensitivity of the two decision variables to the assumption of 50 pages of advertisements lost per week when the advertising price increases by $100/page.

(d) What is the value of the Lagrange multiplier found in part (a)? Interpret this number in terms of the sensitivity of profit to the $400/page assumption.

9. Reconsider the newspaper problem of Exercise 7, but now look at the newspaper's business expenses. The current weekly business expenses for the paper are as follows: $80,000 for the editorial department (news, features, editorials); $30,000 for the sales department (advertising); $30,000 for the circulation department; and $60,000 in fixed costs (mortgage, utilities, maintenance). The new management is considering cuts in the editorial

department. It is estimated that the paper can operate with a minimum of a $40,000/week editorial budget. Reducing the editorial budget will save money, but it will also affect the quality of the paper. Experience in other markets suggests that the paper will lose 2% of its subscribers and 1% of its advertisers for every 10% cut in the editorial budget. Management is also considering an increase in the sales budget. Recently the management of another paper in a similar market expanded its advertising sales budget by 20%. The result was a 15% increase in advertisements. The sales budget may be increased to as much as $50,000/week, but the overall budget for business expenses will not be increased beyond the current level of $200,000/week.

(a) Find the editorial and sales budget figures that maximize profit. Assume that the subscription price remains at $1.50/week, and the advertising price stays at $250/page. Use the five-step method, and model as a constrained optimization problem. Solve using the method of Lagrange multipliers.

(b) Calculate the shadow price for each constraint, and interpret their meaning.

(c) Draw a graph of the feasible region for this problem. Indicate the location of the optimal solution on this graph. Which of the constraints are binding at the optimal solution? How is this related to the shadow prices?

(d) Suppose that cuts in the editorial budget produce an unusually strong negative response in this market. Assume that a 10% cut in the editorial budget causes the paper to lose $p\%$ of its advertising and $2p\%$ of its subscribers. Determine the smallest value of p for which the paper would be better off not to cut the editorial budget.

10. A shipping company has the capacity to move 100 tons/day by air. The company charges $250/ton for air freight. Besides the weight constraint, the company can only move 50,000 ft^3 of cargo per day because of limited volume of aircraft storage compartments. The following amounts of cargo are available for shipping each day:

Cargo	Weight (tons)	Volume (ft^3/ton)
1	30	550
2	40	800
3	50	400

(a) Determine how many tons of each cargo should be shipped by air each day in order to maximize revenue. Use the five-step method, and model as a constrained optimization problem. Solve using Lagrange multipliers.

(b) Calculate the shadow prices for each constraint, and interpret their meaning.

(c) The company has the capability to reconfigure some of its older planes to increase the size of the cargo areas. The alterations would cost $200,000 per plane and would add 2,000 ft^3 per plane. The weight limits would stay the same. Assuming that the planes fly 250 days per year and that the remaining lifetime of the older planes is around five years, would it be worthwhile to make the alterations? To how many planes?

Further Reading

1. Beightler, C., Phillips, D. and Wilde, D. (1979) *Foundations of Optimization.* 2nd ed., Prentice–Hall, Englewood Cliffs, New Jersey.

2. Courant, R. (1937) *Differential and Integral Calculus.* vol. II, Wiley, New York.

3. Edwards, C. (1973) *Advanced Calculus of Several Variables.* Academic Press, New York.

4. Hundhausen, J., and Walsh, R. *The Gradient and Some of Its Applications.* UMAP module 431.

5. Hundhausen, J., and Walsh, R. *Unconstrained Optimization.* UMAP module 522.

6. Nievergelt, Y. *Price Elasticity of Demand: Gambling, Heroin, Marijuana, Whiskey, Prostitution, and Fish.* UMAP module 674.

7. Nevison, C. *Lagrange Multipliers: Applications to Economics.* UMAP module 270.

8. Peressini, A. *Lagrange Multipliers and the Design of Multistage Rockets.* UMAP module 517.

Chapter 3

COMPUTATIONAL METHODS FOR OPTIMIZATION

The preceding chapters have discussed some of the analytic techniques for solving optimization problems. These techniques form the basis for most optimization models. In this chapter we will study some of the computational problems that arise in real applications and discuss a few of the most popular methods for dealing with them.

3.1 One Variable Optimization

Even for simple one–variable optimization problems, the task of locating global extreme points can be exceedingly difficult. Real problems are usually messy. Even when the functions involved are differentiable everywhere, the computation of the derivative is often complicated. The worst part, however, is solving the equation $f'(x) = 0$. The plain and simple fact is that most equations cannot be solved analytically. The best we can do in most instances is to find an approximate solution by graphical or numerical techniques.

Example 3.1. Reconsider the pig problem of Example 1.1, but now take into account the fact that the growth rate of the pig is not constant. Assume that the pig is young, so that the growth rate is increasing. When should we sell the pig for maximum profit?

We will use the five-step method. Step 1 will be to modify the work we did in Section 1.1, as summarized in Fig. 1.1. Now we cannot simply assume that $w = 200 + 5t$. What would be a reasonable assumption to represent an increasing rate of growth? There are, of course, many possible answers to this question. Let us suppose for now that the growth rate of the pig is proportional

to its weight. In other words, let us assume that

$$\frac{dw}{dt} = cw. \tag{3.1}$$

From the fact that $dw/dt = 5$ lbs/day when $w = 200$ lbs, we conclude that $c = 0.025$. This leaves us with a simple differential equation to solve for w, namely

$$\frac{dw}{dt} = 0.025w, \ w(0) = 200. \tag{3.2}$$

We can solve Eq. (3.2) by separation of variables to obtain

$$w = 200e^{0.025t}. \tag{3.3}$$

Since all of our other assumptions are unchanged from what was presented in Fig. 1.1, this concludes step 1.

Step 2 is to select our modeling approach. We will model the problem as a one–variable optimization problem. The general solution procedure for one–variable optimization problems was outlined in Section 1.1. In this section we will explore some computational methods that can be used to implement this general solution procedure. Computational methods such as those we present here are often needed in real problems when calculations become either too hard or too tedious to perform by hand.

Step 3 is to formulate the model. The only difference between the present case and the problem formulation of Section 1.1 is that we have to replace the weight equation $w = 200 + 5t$ by Eq. (3.3). This leads to the new objective function

$$\begin{aligned} y &= f(x) \\ &= (0.65 - 0.01x)(200e^{0.025x}) - 0.45x, \end{aligned} \tag{3.4}$$

and our problem is to maximize the function in Eq. (3.4) over the set $S = \{x : x \geq 0\}$.

Step 4 is to solve the model. We will use the graphical method. Good graphing utilities for personal computers, and graphing calculators, are widely available. We start our graphical analysis of this problem by producing a graph of the function in Eq. (3.4) on the same scale as Fig. 1.2, our graph of the original objective function. In this case we are left with the feeling that there is more to see on the graph of this function. We would say that Fig. 3.1 is not a *complete graph* of this function over the set $S = [0, \infty)$.

Figure 3.2 *is* a complete graph. It shows all of the important features we need for the solution of our problem.

How do we know when we have a complete graph?   There is no simple answer to this question. Graphing is an exploratory technique. You need to experiment and to use good judgment. Of course, we do not need to look at negative values of x, but we also need not look beyond $x = 65$. After this point our formula says that the price for pigs is negative, which is clearly nonsense.
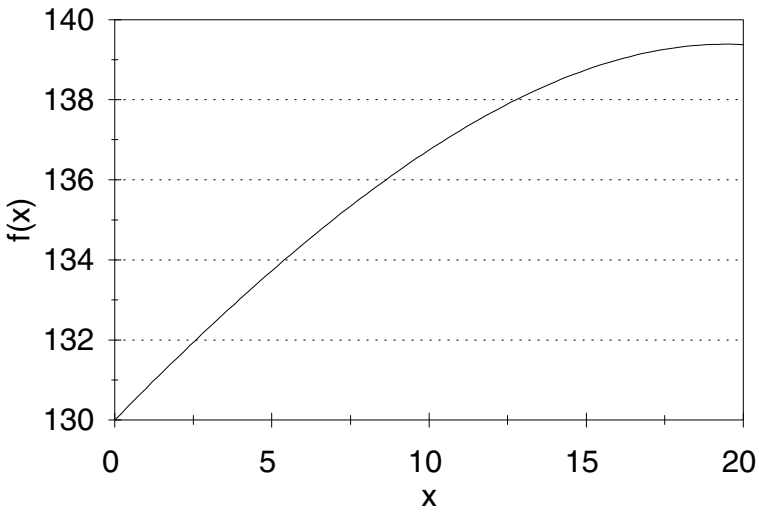
Figure 3.1: Graph of net profit $f(x)$ from (3.4) versus time to sell x for the pig problem with nonlinear weight model.
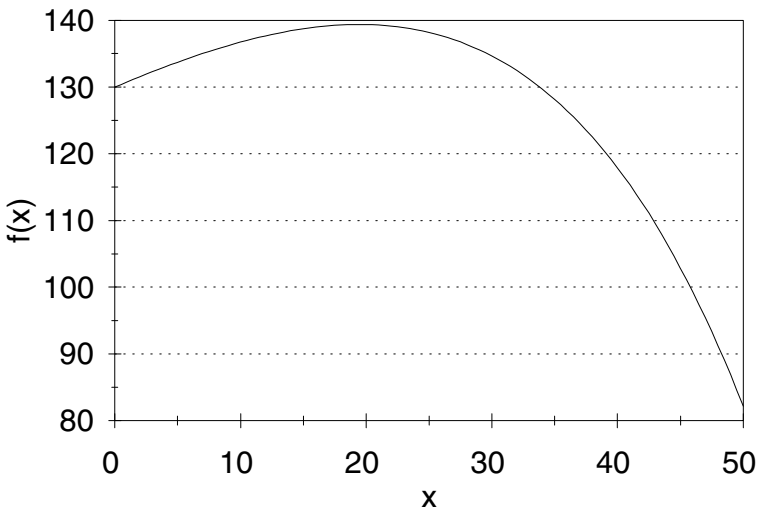

Figure 3.2: Complete graph of net profit $f(x)$ from (3.4) versus time to sell x for the pig problem with nonlinear weight model.
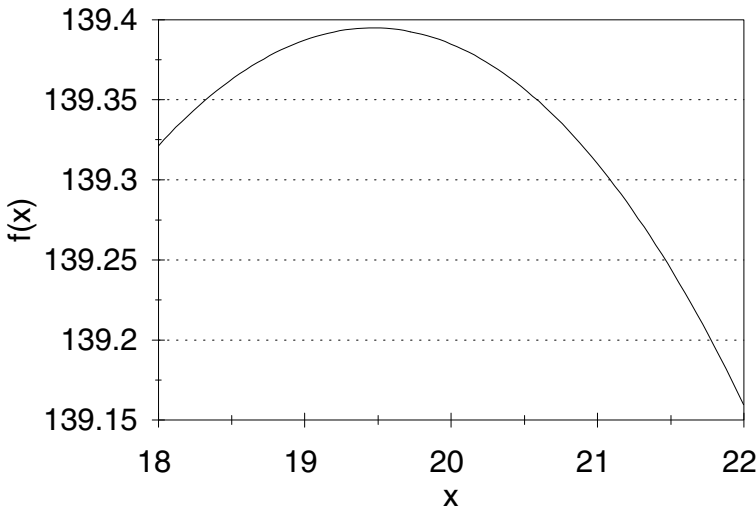
Figure 3.3: First zoom–in on the graph of net profit $f(x)$ from (3.4) versus time to sell x for the pig problem with nonlinear weight model.

From the graph in Fig. 3.2 we can conclude that the maximum occurs around $x = 20$ and $y = f(x) = 140$. To obtain a better estimate, we can zoom in on the maximum point of the graph. Figures 3.3 and 3.4 show the outcome of successive zooms.

On the basis of the graph in Fig. 3.4, we would estimate that the maximum occurs at

$$x = 19.5$$
$$y = f(x) = 139.395. \tag{3.5}$$

At this point we have found the location of the maximum to three significant digits, and the value of the maximum to six significant digits. Since $f'(x) = 0$ at the maximum, the function $f(x)$ is quite insensitive to changes in x near this point, so we are able to obtain more accuracy for $f(x)$ than for x.

Step 5 is to answer the question. After taking into account the fact that the growth rate of the young pig is still increasing, we now recommend waiting 19 or 20 days to sell. This should result in a net profit of approximately \$140.

The graphical method used in step 4 to locate the optimum point (see Eq. (3.5)) did not produce a high degree of accuracy. This is acceptable for the present problem because we do not need a high degree of accuracy. While it is true that the graphical method can be made to produce higher accuracy (by zooming in on the optimum point over and over again), there are more efficient computational methods that should be used in such cases. We will look at some of these next, in the course of our sensitivity analysis.
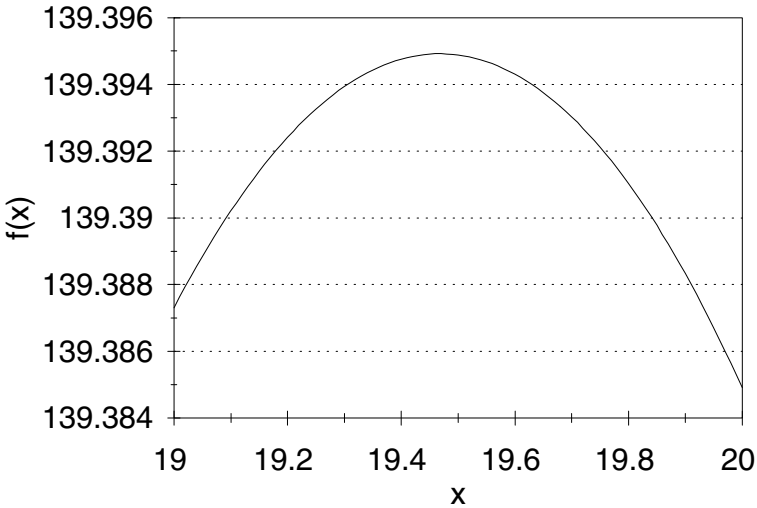
Figure 3.4: Second zoom–in on the graph of net profit $f(x)$ from (3.4) versus time to sell x for the pig problem with nonlinear weight model.

Let us examine the sensitivity of the optimum coordinates in Eq. (3.5) to the growth rate $c = 0.025$ for the young pig. One way to do this would be to repeat our graphical analysis for several different values of the parameter c. However, this would be tedious. We would prefer a more efficient method.

Let us begin by generalizing the model. Now we are assuming that

$$\frac{dw}{dt} = cw, \ \ w(0) = 200, \tag{3.6}$$

so that

$$w = 200e^{ct}. \tag{3.7}$$

This leads to the objective function

$$f(x) = (0.65 - 0.01x)(200e^{cx}) - 0.45x. \tag{3.8}$$

From our graphical analysis, we know that for $c = 0.025$, the optimum occurs at an interior critical point, at which point $f'(x) = 0$. Since f is a continuous function of c, it seems reasonable to conclude that the same holds for values of c near 0.025. In order to locate this interior critical point, we need to compute the derivative $f'(x)$ and solve the equation $f'(x) = 0$. The first part of this process (computing the derivative) is relatively easy. There is a standard method for computing derivatives that you learned in one–variable calculus. It can be applied to virtually any differentiable function. For complicated expressions the derivative can also be computed using a computer algebra system (Maple,

Mathematica, Derive, etc.) or a hand calculator (such as the HP–48). In our problem, it is not too difficult to compute by hand that

$$f'(x) = 200ce^{cx}(0.65 - 0.01x) - 2e^{cx} - 0.45. \qquad (3.9)$$

The second part of the process is to solve the equation

$$200ce^{cx}(0.65 - 0.01x) - 2e^{cx} - 0.45 = 0. \qquad (3.10)$$

You can try to solve this equation by hand, but you are unlikely to succeed in obtaining an algebraic solution. Some computer algebra systems can now solve (3.10) in terms of the W function, which was developed specifically to solve equations of this type. Ongoing research in mathematics and mathematical software will continue to expand our ability to solve equations, but the plain and simple fact is that most equations still cannot be solved by algebraic methods. While there does exist a general algebraic method for computing derivatives, there is no general algebraic method for solving equations. Even for polynomials, it is known that there can be no generally useful algebraic method of finding roots (i.e., there can be no analog of the quadratic formula) for degree five or above. This is why it is often necessary to resort to *numerical approximation* methods to solve algebraic equations.

We will use Newton's method to solve Eq. (3.10). You probably learned about Newton's method in one–variable calculus.

> We are given a differentiable function $F(x)$ and an approximation x_0 to the root $F(x) = 0$. Newton's method works by linear approximation. Near $x = x_0$ we have $F(x) \approx F(x_0) + F'(x_0)(x - x_0)$ by the usual tangent line approximation. To obtain a better estimate $x = x_1$ of the actual root to $F(x) = 0$, we set $F(x_0) + F'(x_0)(x - x_0) = 0$ and solve for $x = x_1$ to get $x_1 = x_0 - F(x_0)/F'(x_0)$. Geometrically, the tangent line to $y = F(x)$ at the point $x = x_0$ intersects the x–axis at the point $x = x_1$. As long as x_1 is not too far away from x_0, the tangent line approximation should be reasonably good, and so x_1 should be close to the actual root. Newton's method produces a sequence of increasingly accurate estimates $x_1, x_2, x_3, \ldots$, to the actual root by repeating this tangent line approximation. Once we are sufficiently close to the root, each successive approximation produced by Newton's method is accurate to about twice as many decimal places as the preceding estimate.

Figure 3.5 presents Newton's method in a form called *pseudocode*. This is a standard method for describing a numerical algorithm.

It is a fairly simple matter to transform pseudocode into a working computer program in any high–level computer language (BASIC, FORTRAN, C, PASCAL, etc.) or to implement on a spreadsheet. It is also possible to program in most computer algebra systems. For the numerical methods in this book, any of these options can be used. It is not recommended that these algorithms be implemented by hand.

Algorithm: NEWTON'S METHOD

Variables: $x(n)$ = approximate location of root after n iterations.
                N = number of iterations

Input: $x(0), N$

Process: Begin
                for $n = 1$ to N do
                    Begin
                    $x(n) \longleftarrow x(n-1) - F(x(n-1))/F'(x(n-1))$
                    End
                End

Output: $x(N)$

Figure 3.5: Pseudocode for Newton's Method in one variable.

c	x
0.022	11.626349
0.023	14.515929
0.024	17.116574
0.025	19.468159
0.026	21.603681
0.027	23.550685
0.028	25.332247

Table 3.1: Sensitivity of best time to sell x to growth rate parameter c for the pig problem with nonlinear weight model.

In our problem we want to use Newton's method to find a root of the equation

$$F(x) = 200ce^{cx}(0.65 - 0.01x) - 2e^{cx} - 0.45 = 0. \qquad (3.11)$$

For values of c near 0.025, we expect to find a root near the point $x = 19.5$. We used a computer implementation of Newton's method to produce the results in Table 3.1. For each value of c, we performed $N = 10$ iterations starting at the point $x(0) = 19.5$. An additional sensitivity run with $N = 15$ was used to verify the accuracy of our results.

Notice that our method of solution for Eq. (3.11) has two steps. First, we apply a *global method* (graphing) to locate an approximate solution. Then, we apply a fast *local method* to determine the exact solution to the desired accuracy. These are the two stages of numerical solution, and they are common to most generally useful solution methods. For one–variable optimization, graphing is the simplest and most useful global method. Newton's method is easy to program, and built–in numerical equation solvers are also available on most graph-

ing calculators, spreadsheets, and computer algebra systems. While the details vary, most of these solvers are based on some variation of Newton's method. These solvers can be used safely and effectively in the same manner as Newton's method. First, use a global method to approximate the root. Most solvers require either a point estimate or an interval estimate of the root. Then, exercise the solver and verify the result, either by performing sensitivity analysis on the available tolerance parameter(s), or by substituting the numerical solution back into the original equation. One word of caution: Casual and uncritical faith in the results of numerical solvers is dangerous. For many real problems, including some of the exercises in this book, inappropriate use of numerical solvers has been known to produce significant errors. Initial application of an appropriate global method, and subsequent verification of the root, are important parts of the numerical solution procedure. Some calculators, computer algebra systems, and spreadsheets also have numerical optimizers. Usually, these routines apply a variant of Newton's method based on a numerical approximation of the derivative. The same advice applies to these routines. Use a global method to approximate the optimum, exercise the numerical optimizer, and then perform sensitivity analysis on the tolerance parameters to ensure accuracy.

In order to relate our sensitivity results back to the original data in the problem, in Figure 3.6 we have plotted the root x, which represents the optimal time to sell, against the growth rate

$$g = 200c, \tag{3.12}$$

which was originally given as $g = 5$ lbs/day.

To obtain a numerical estimate of sensitivity, we solve Eq. (3.11) once more, setting $c = 0.02525$ (a 1% increase). The solution found was

$$x = 20.021136,$$

which represents a 2.84% increase in x. Thus, we estimate that $S(x, c) = 2.84$. Since

$$g = 200c,$$

we can easily show that

$$S(x, g) = S(x, c) = 2.84.$$

Also, if we let h denote the initial weight of the pig (we assumed $h = 200$ lbs), then since

$$h = 5/c,$$

we have

$$S(x, h) = \frac{dx}{dh} \cdot \frac{h}{x}$$
$$= \left(\frac{dx/dc}{dh/dc} \right) \left(\frac{5/c}{x} \right)$$
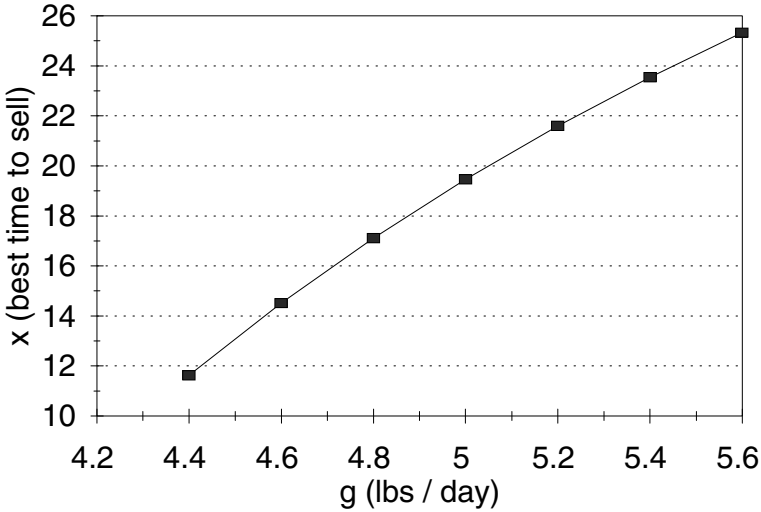$$= -S(x, c) = -2.84.$$

Figure 3.6: Graph of best time to sell x versus growth rate g for the pig problem with nonlinear weight model.

In fact, it is generally true that if y is proportional to z, then

$$S(x, y) = S(x, z), \tag{3.13}$$

and if y is inversely proportional to z, then

$$S(x, y) = -S(x, z). \tag{3.14}$$

Now we have computed the sensitivity of x to both the initial weight of the pig and the growth rate of the pig. The other sensitivities are unchanged from the original problem considered in Chapter 1, since the other parameters appear in the objective function in the same manner as before.

The fact that our optimal solution in the present case differs so markedly from what we found in Chapter 1 (19 or 20 days here versus 8 days there) raises some important questions about the robustness of our model. There is also serious concern about whether the assumption

$$p = 0.65 - 0.01t$$

would be valid over a three–week period. Alternative models for price could certainly be considered, along with more sophisticated models of the growth process for a representative pig. Some of these issues of robustness are addressed in the exercises at the end of this chapter. What we can say now is this: If the pig's rate of growth does not diminish, and if the price decline does not accelerate, then we should hold onto the pig for another week. At that time, we can reevaluate the situation on the basis of new data.

3	0	1	4	2	1
2	1	1	2	3	2
5	3	3	0	1	2
8	5	2	1	0	0
10	6	3	1	3	1
0	2	3	1	1	1

Figure 3.7: Map showing the number of emergency calls per year in each one square mile area of the city. Up is north and right is east.

3.2   Multivariable Optimization

The practical problems associated with locating the global optimum of a function of several variables are similar in many ways to those discussed in the preceding section. Additional complications arise because of the dimension of the problem. Graphical techniques are not available in dimensions $n > 3$, and solving $\nabla f = 0$ becomes more complicated as the number of independent variables increases. Constrained optimization is also more difficult because the geometry of the feasible region can be more complicated.

Example 3.2. A suburban community intends to replace its old fire station with a new facility. The old station was located at the historical city center. City planners intend to locate the new facility more scientifically. A statistical analysis of response–time data yielded an estimate of $3.2 + 1.7r^{0.91}$ minutes required to respond to a call r miles away from the station. (The derivation of this formula is the subject of Exercises 17 and 18 in Chapter 8.) Estimates of the frequency of calls from different areas of the city were obtained from the fire chief. They are presented in Figure 3.7. Each block represents one square mile, and the numbers inside each block represent the number of emergency calls per year for that block. Find the best location for the new facility.

   We will represent locations on the city map by coordinates (x, y), where x is the distance in miles to the west side of town, and y is the distance in miles to the south side. For example, $(0, 0)$ represents the lower left–hand corner of the map, and $(0, 6)$ represents the upper left–hand corner. For simplicity we will divide the city into nine 2×2–mile squares and assume that each emergency
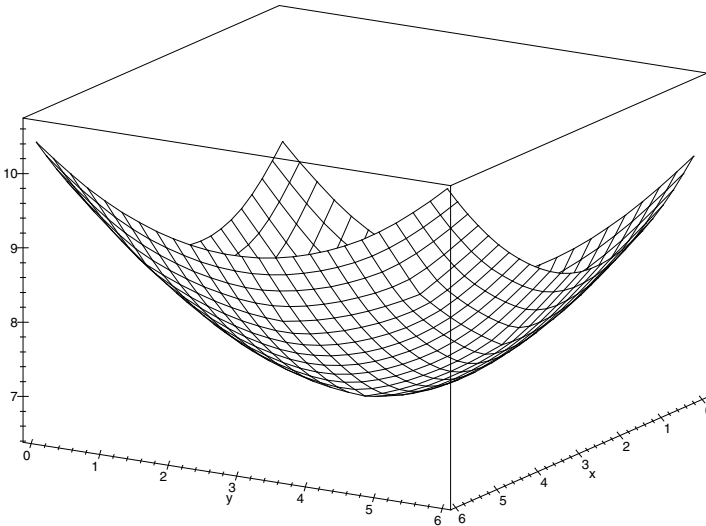
Figure 3.8: 3–D graph of average response time $z = f(x, y)$ from (3.15) versus map location (x, y) for the facility location problem.

is located at the center of the square. If (x, y) is the location of the new fire station, the average response time to a call is $z = f(x, y)$, where

$$z = 3.2 + 1.7 \left[6\sqrt{(x-1)^2 + (y-5)^2}^{0.91} \right.$$
$$+ 8\sqrt{(x-3)^2 + (y-5)^2}^{0.91} + 8\sqrt{(x-5)^2 + (y-5)^2}^{0.91}$$
$$+ 21\sqrt{(x-1)^2 + (y-3)^2}^{0.91} + 6\sqrt{(x-3)^2 + (y-3)^2}^{0.91} \quad (3.15)$$
$$+ 3\sqrt{(x-5)^2 + (y-3)^2}^{0.91} + 18\sqrt{(x-1)^2 + (y-1)^2}^{0.91}$$
$$\left. + 8\sqrt{(x-3)^2 + (y-1)^2}^{0.91} + 6\sqrt{(x-5)^2 + (y-1)^2}^{0.91} \right]/84.$$

The problem is to minimize $z = f(x, y)$ over the feasible region $0 \leq x \leq 6$, $0 \leq y \leq 6$.

Figure 3.8 shows a 3–D graph of the objective function f over the feasible region. It appears as though f attains its minimum at the unique interior point at which $\nabla f = 0$. Figure 3.9 shows a contour plot of the level sets of f, indicating that $\nabla f = 0$ near the point $x = 2$ and $y = 3$.

Now, it is certainly possible to compute ∇f in this problem, but it is not possible to solve $\nabla f = 0$ algebraically. Further graphical analysis is possible, but this is especially cumbersome for functions of more than one variable. What is needed here is a simple global method for estimating the minimum.
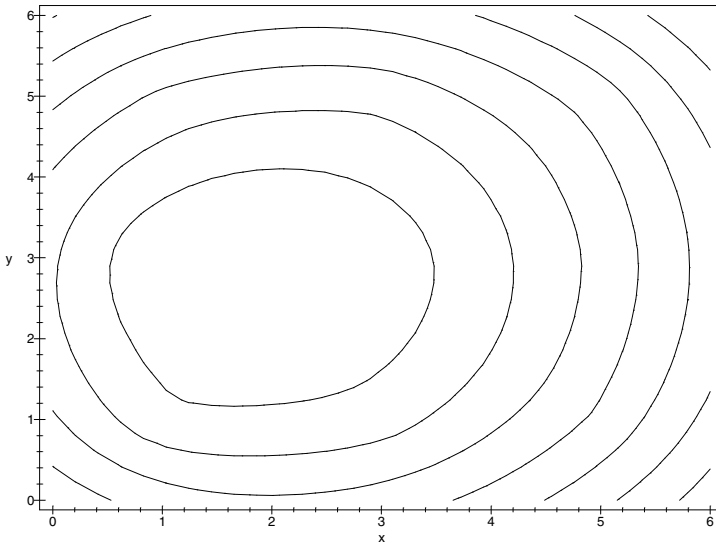
Figure 3.9: Contour plot of average response time $z = f(x, y)$ from (3.15) versus map location (x, y) for the facility location problem.

Figure 3.10 presents an algorithm for random search. This optimization method simply picks N feasible points at random and selects the one that yields the smallest value of the objective function. The notation Random $\{S\}$ denotes a randomly selected point from the set S. A computer implementation of random search was applied to the function in Eq. (3.15) with $a = 0$, $b = 6$, $c = 0$, $d = 6$, and $N = 1,000$. The resulting estimate of the minimum occurred at

$$x\,min = 1.66$$
$$y\,min = 2.73 \qquad\qquad (3.16)$$
$$z\,min = 6.46.$$

Since this algorithm involves random numbers, a repetition of the same computer implementation with the same inputs may produce slightly different outputs. The accuracy of random search is roughly the same as if the N points were to lie on an equally spaced grid over the entire feasible set. Such a grid would contain 32×32 points ($32^2 \approx 1,000$), so we would be accurate to within $6/32 \approx 0.2$ in both x and y. Since $\nabla f = 0$ at the minimum, we obtain much better accuracy in z. An alternative to random search would be a grid search (examine $z = F(x, y)$ at N equally spaced points). The performance of grid search is essentially the same as random search, but random search is more flexible and easier to implement.

The estimates in Eq. (3.16) of the optimum location (1.7, 2.7) and the re-

Algorithm: METHOD OF RANDOM SEARCH

Variables: a = lower limit on x
b = upper limit on x
c = lower limit on y
d = upper limit on y
N = number of iterations
$x\,min$ = approximate x coordinate of minimum
$y\,min$ = approximate y coordinate of minimum
$z\,min$ = approximate value of $F(x, y)$ at minimum

Input: a, b, c, d, N

Process: Begin
$x \longleftarrow$ Random $\{[a, b]\}$
$y \longleftarrow$ Random $\{[c, d]\}$
$z\,min \longleftarrow F(x, y)$
for $n = 1$ to N do
    Begin
    $x \longleftarrow$ Random $\{[a, b]\}$
    $y \longleftarrow$ Random $\{[c, d]\}$
    $z \longleftarrow F(x, y)$
    if $z < z\,min$ then
        $x\,min \longleftarrow x$
        $y\,min \longleftarrow y$
        $z\,min \longleftarrow z$
    End
End

Output: $x\,min$, $y\,min$, $z\,min$

Figure 3.10: Pseudocode for the method of random search.

sulting average response time of 6.46 minutes were obtained by evaluating the objective function at $N = 1,000$ random points in the feasible region. Better accuracy could be obtained by increasing N. However, the behavior of this simple, global method does not encourage such an approach. Each additional decimal place of accuracy requires increasing N by *a factor of one hundred.* Hence, this method is only suitable for obtaining a rough approximation of the optimum. For the present problem the answer we found is good enough. The simplifying assumptions we made earlier introduced errors on the order of one mile in emergency location, so there is no point in demanding more accuracy now. It is enough to state that the facility should be located around $(1.7, 2.7)$ on the map, for an average response time of around 6.5 minutes. The exact location will, in any case, depend on several factors not incorporated into our

model, such as the location of roads and the availability of land in the area of the optimal site. It is also reasonable to consider different definitions of an "optimal" location; see Exercise 3.6.

It is important to provide an estimate of the sensitivity of response time to the eventual facility location. Since $\nabla f = 0$ at the optimal location, we do not expect f to vary much near $(1.7, 2.7)$. To obtain a more concrete understanding of the sensitivity of f to (x, y) near the optimum, we reran the random search program, replacing f by $-f$ and bounding $1.5 \le x \le 2$, $2.5 \le y \le 3$. After $N = 100$ observations we found that the maximum of f over this region was approximately 6.49 minutes, or about 0.03 minutes longer than the observed optimum. It does not matter in any practical sense where in this half–mile square area we locate the station.

The random search methods employed in the preceding example are simple but slow. For some problems, where greater accuracy is required, such methods are unsuitable. If the functions involved are as complex as the objective in Example 3.2, then it will be very hard to obtain an accurate answer. More accurate and efficient methods for global optimization of functions of more than one variable are almost always based on the gradient. We now present an example in which these methods are more tractable, because the gradient is easier to compute.

Example 3.3. A manufacturer of lawn furniture makes two types of lawn chairs, one with a wood frame and one with a tubular aluminum frame. The wood–frame model costs \$18 per unit to manufacture, and the aluminum–frame model costs \$10 per unit. The company operates in a market where the number of units that can be sold depends on the price. It is estimated that in order to sell x units per day of the wood–frame model and y units per day of the aluminum–frame model, the selling price cannot exceed $10 + 31x^{-0.5} + 1.3y^{-0.2}$ \$/unit for wood–frame chairs, and $5 + 15y^{-0.4} + 0.8x^{-0.08}$ \$/unit for aluminum–frame chairs. Find the optimal production levels.

The objective is to maximize the profit function $z = f(x, y)$ \$/day over the feasible set of production levels $x \ge 0$, $y \ge 0$, where

$$z = x(10 + 31x^{-0.5} + 1.3y^{-0.2}) - 18x$$
$$+ y(5 + 15y^{-0.4} + 0.8x^{-0.08}) - 10y. \tag{3.17}$$

Figure 3.11 shows a graph of f. The graph indicates that f has a unique interior maximum at the point where $\nabla f = 0$. Figure 3.12 shows the level sets of f. This graph suggests that the maximum occurs around $x = 5$ and $y = 6$. We calculate the gradient $\nabla f(x, y) = (\partial z/\partial x, \partial z/\partial y)$ to get

$$\frac{\partial z}{\partial x} = 15.5x^{-0.5} - 8 + 1.3y^{-0.2} - 0.064yx^{-1.08}$$

$$\frac{\partial z}{\partial y} = 9y^{-0.4} - 5 + 0.8x^{-0.08} - 0.26xy^{-1.2}. \tag{3.18}$$
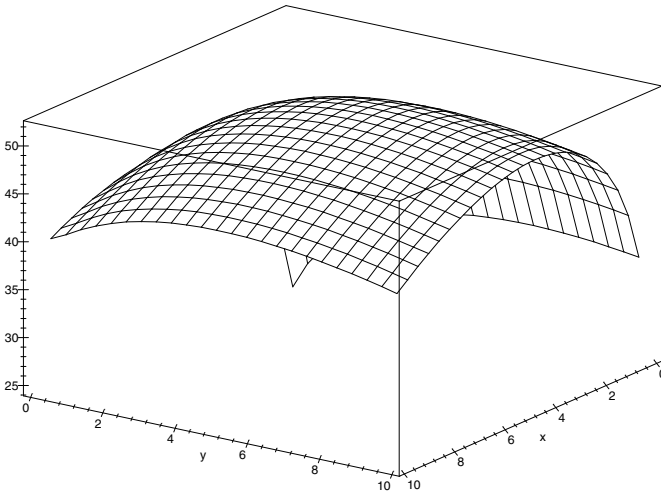
Figure 3.11: 3–D graph of profit $z = f(x, y)$ from (3.17) versus the number x of wood–frame chairs and the number y of aluminum–frame chairs produced per day in the lawn chair problem.
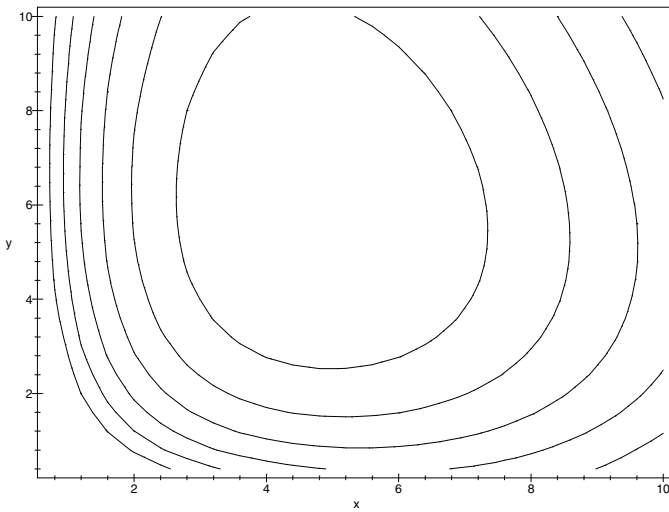

Figure 3.12: Contour graph of profit $z = f(x, y)$ from (3.17) versus the number x of wood–frame chairs and the number y of aluminum–frame chairs produced per day in the lawn chair problem.

A random search of $N = 1,000$ points over the region $0 \leq x \leq 10$, $0 \leq y \leq 10$ yields an observed maximum of \$52.06 per day at the production level $x = 4.8$ and $y = 5.9$ chairs per day. To obtain a more accurate numerical estimate of the optimum point, we will use a multivariable version of Newton's method to solve the gradient equation $\nabla f = 0$. Figure 3.13 gives the algorithm for Newton's method in two variables.

We are given a set of differentiable functions, $f_1, \ldots, f_n$, and an approximate root $(x_1(0), \ldots, x_n(0))$ to the system of equations

$$f_1(x_1, \ldots, x_n) = 0$$
$$\vdots \tag{3.19}$$
$$f_n(x_1, \ldots, x_n) = 0.$$

In order to simplify the explanation, and to clarify the connection between the one–variable and multivariable Newton's method, we will use vector notation. Let $x = (x_1, \ldots, x_n)$ and $F(x) = (f_1(x), \ldots, f_n(x))$ so that we can rewrite Eq. (3.19) in the form $F(x) = 0$ and $x(0)$ is the initial estimate of the root. Newton's method produces a sequence of increasingly accurate estimates $x(1)$, $x(2), x(3), \ldots$ to the actual root $F(x) = 0$ using a linear approximation. Near $x = x(0)$ we have $F(x) \approx F(x(0)) + A(x - x(0))$, where A is the matrix of partial derivatives

$$A = \begin{pmatrix} \partial f_1/\partial x_1 & \cdots & \partial f_1/\partial x_n \\ \vdots & & \vdots \\ \partial f_n/\partial x_1 & \cdots & \partial f_n/\partial x_n \end{pmatrix}$$

evaluated at $x = x(0)$. This is the multivariable version of the tangent line approximation. To obtain a better estimate $x = x(1)$ to the actual root $F(x) = 0$, we set $F(x(0)) + A(x - x(0)) = 0$ and solve for $x = x(1)$ to get $x(1) = x(0) - A^{-1}F(x(0))$. This is exactly the same formula as in the one–variable case, except that we cannot divide by the derivative matrix, so we take the inverse. Once we are sufficiently close to the root, each successive approximation is accurate to about twice as many decimal places as the preceding estimate. Just as in the one–variable case, the sequence of estimates produced by Newton's method converges rapidly to the actual root.

The pseudocode in Figure 3.13 implements Newton's method in two variables using the formula for the inverse of a 2×2 matrix

$$\begin{pmatrix} q & r \\ s & t \end{pmatrix}^{-1} = \frac{1}{qt - rs} \begin{pmatrix} t & -r \\ -s & q \end{pmatrix}.$$

In the general case, it is necessary to compute the inverse of the derivative matrix A by other methods. For more details on Newton's method in several variables, see Press, et al. (2002).

Algorithm: NEWTON'S METHOD IN TWO VARIABLES

Variables: $x(n)$ = approximate x coordinate of the root after n iterations.
$y(n)$ = approximate y coordinate of the root after n iterations.
N = number of iterations

Input: $x(0)$, $y(0)$, N

Process: Begin
   for $n = 1$ to N do
      Begin
      $q \longleftarrow \partial F/\partial x(x(n-1), y(n-1))$
      $r \longleftarrow \partial F/\partial y(x(n-1), y(n-1))$
      $s \longleftarrow \partial G/\partial x(x(n-1), y(n-1))$
      $t \longleftarrow \partial G/\partial y(x(n-1), y(n-1))$
      $u \longleftarrow -F(x(n-1), y(n-1))$
      $v \longleftarrow -G(x(n-1), y(n-1))$
      $D \longleftarrow qt - rs$
      $x(n) \longleftarrow x(n-1) + (ut - vr)/D$
      $y(n) \longleftarrow y(n-1) + (qv - su)/D$
      End
   End

Output: $x(N)$, $y(N)$

Figure 3.13: Pseudocode for Newton's method in two variables.

In our problem we have

$$F(x, y) = 15.5x^{-0.5} - 8 + 1.3y^{-0.2} - 0.064yx^{-1.08}$$
$$G(x, y) = 9y^{-0.4} - 5 + 0.8x^{-0.08} - 0.26xy^{-1.2}, \tag{3.20}$$

and so we can compute that

$$\frac{\partial F}{\partial x} = 0.06912yx^{-2.08} - 7.75x^{-1.5}$$

$$\frac{\partial F}{\partial y} = -0.064x^{-1.08} - 0.26y^{-1.2}$$

$$\frac{\partial G}{\partial x} = -0.064x^{-1.08} - 0.26y^{-1.2} \tag{3.21}$$

$$\frac{\partial G}{\partial y} = 0.312xy^{-2.2} - 3.6y^{-1.4}.$$

We used a computer implementation of Newton's method in two variables start-

ing at $x(0) = 5$, $y(0) = 5$. After $N = 10$ iterations we obtained

$$x = 4.68959$$
$$y = 5.85199$$

(3.22)

as our estimate of the root. A further sensitivity run with $N = 15$ was made to confirm these results. Substituting back into Eq. (3.17) yields $z = 52.0727$. Hence, the optimal solution to the lawn chair problem is to produce 4.69 wood–frame chairs and 5.85 aluminum–frame chairs per day, which should result in a net profit of \$52.07 per day.

As in the one–variable case, the accurate numerical solution of systems of equations is a two–step process. First, we must use a global method to estimate the location of the root. Graphical methods are useful if there are only two variables, but for most problems, a numerical method such as random search is necessary. A variety of more sophisticated global methods is also available, most of which are tailored to special classes of problems. Next, we use a fast local method to obtain an accurate solution, which must be verified either by sensitivity analysis of the parameters which control accuracy, or by substituting the numerical solution back into the original system of equations. Newton's method is a very fast local method that requires computing the partial derivatives. Various modifications exist which employ estimates of the partial derivatives; e.g., see Press, et al. (2002). Most spreadsheets and computer algebra systems have multivariable equation solvers, and solvers are also available in most numerical analysis software libraries and packages. To use one of these, begin with a global method to estimate the location of the root, then exercise the solver, and finally verify the solution as usual. Do not simply plug in the equations, use the solver, and accept the reported solution, as this can lead to serious errors. Many spreadsheets, computer algebra systems, and numerical analysis packages also have multivariable numerical optimizers. Usually these routines apply a variation of Newton's method based on numerical approximation of the derivatives. The same advice applies to these routines. Use a global method to approximate the optimum, exercise the numerical optimizer, and then perform sensitivity analysis on the tolerance parameters to ensure accuracy.

Multivariable optimization problems with constraints may be harder to solve. In some cases the optimum occurs in the interior of the feasible region, so that we may treat the problem as if there were no constraints. When the optimum occurs on the boundary, the situation is more complex. There are no simple, generally effective computational algorithms for solving such problems. The methods that do exist are tailored to the particular features of some special class of problems. In the next section we will discuss the most important of these.

3.3   Linear Programming

Multivariable optimization problems with constraints are almost always difficult to solve. A variety of computational techniques has been developed to handle

Requirements per acre	Corn	Wheat	Oats
Irrigation (acre–ft)	3.0	1.0	1.5
Labor (person–hrs/week)	0.8	0.2	0.3
Yield ($)	400	200	250

Table 3.2: Data for the farm problem of Example 3.4.

special types of multivariable optimization problems, but good general methods do not yet exist, even at the most sophisticated levels. The area of research that considers the development of new computational methods for such problems is called *nonlinear programming*, and it is very active.

The simplest type of multivariable constrained optimization problem is one where both the objective function and the constraint functions are linear. The study of computational methods for such problems is called *linear programming*. Software packages for linear programming are widely available and are in frequent use for problems in manufacturing, investment, farming, transportation, and government. Typical large–scale problems involve thousands of decision variables and thousands of constraints. There are many well–documented cases where an operations analysis based on a linear programming model has generated savings in the millions of dollars. Details can be found in the literature on operations research and management science.

Example 3.4. A family farm has 625 acres available for planting. The crops the family is considering are corn, wheat, and oats. It is anticipated that 1,000 acre–ft of water will be available for irrigation, and the farmers will be able to devote 300 hours of labor per week. Additional data are presented in Table 3.2. Find the amount of each crop that should be planted for maximum profit.

We will use the five-step method. The results of step 1 are shown in Figure 3.14. Step 2 is to select the modeling approach. We will model this problem as a linear programming problem.

The standard (inequality) form of a linear programming model is as follows: Maximize the objective function $y = f(x_1, \ldots, x_n) = c_1 x_1 + \cdots + c_n x_n$ over the feasible region defined by the constraints

$$a_{11} x_1 + \cdots + a_{1n} x_n \leq b_1$$
$$\vdots \qquad \qquad \vdots \quad \vdots \qquad \qquad (3.23)$$
$$a_{m1} x_1 + \cdots + a_{mn} x_n \leq b_m$$

and $x_1 \geq 0, \ldots, x_n \geq 0$. This is a special case of the multivariable

Variables: $x_1 =$ acres of corn planted
$x_2 =$ acres of wheat planted
$x_3 =$ acres of oats planted
$w =$ irrigation required (acre–ft)
$l =$ labor required (person–hrs/wk)
$t =$ total acreage planted
$y =$ total yield (\$)

Assumptions: $w = 3.0x_1 + 1.0x_2 + 1.5x_3$
$l = 0.8x_1 + 0.2x_2 + 0.3x_3$
$t = x_1 + x_2 + x_3$
$y = 400x_1 + 200x_2 + 250x_3$
$w \leq 1,000$
$l \leq 300$
$t \leq 625$
$x_1 \geq 0; \; x_2 \geq 0; \; x_3 \geq 0$

Objective: Maximize y

Figure 3.14: Results of step 1 of the farm problem.

constrained optimization problem discussed in Chapter 2. Let

$$g_1(x_1, \ldots, x_n) = a_{11}x_1 + \cdots + a_{1n}x_n$$

$$\vdots \tag{3.24}$$

$$g_m(x_1, \ldots, x_n) = a_{m1}x_1 + \cdots + a_{mn}x_n$$

and

$$g_{m+1}(x_1, \ldots, x_n) = x_1$$

$$\vdots$$

$$g_{m+n}(x_1, \ldots, x_n) = x_n.$$

The constraints can be written in the form $g_1 \leq b_1, \ldots, g_m \leq b_m$ and $g_{m+1} \geq 0, \ldots, g_{m+n} \geq 0$. The set of $(x_1, \ldots, x_n)$ that satisfies these constraints is called the *feasible region*. It represents all feasible values for the decision variables $x_1, \ldots, x_n$. Since $\nabla f = (c_1, \ldots, c_n)$ can never be zero, the function cannot attain its maximum in the interior of the feasible region. At a maximum on the boundary we must have

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_{m+n} \nabla g_{m+n} \tag{3.25}$$

with $\lambda_i \neq 0$ only if the ith constraint is binding. For $i = 1, \ldots, m$, the value of the Lagrange multiplier λ_i represents the potential increase in the maximum value of the objective function $f(x_1, \ldots, x_n)$

that would result from relaxing the ith constraint by one unit (changing it to $g_i(x_1, \ldots, x_n) \leq b_i + 1$). Computation of the optimal solution to a linear programming problem is usually obtained by computer, using a variation of the *simplex method*. This method is based on the fact (which you can easily verify) that the optimal solution must occur at one of the corner points of the feasible region. Rather than go into the details involved in the algebra of the simplex method, we will concentrate on what you need to know in order to use a linear programming package correctly.

In the simplex method, the coordinates of a corner point are computed using the alternative (equality) form of a linear programming problem: Maximize $y = c_1 x_1 + \cdots + c_n x_n$ over the set

$$
\begin{aligned}
a_{11}x_1 + \cdots + a_{1n}x_n + x_{n+1} &= b_1 \\
a_{21}x_1 + \cdots + a_{2n}x_n \quad\quad + x_{n+2} &= b_2 \\
\vdots \quad\quad\quad & \\
a_{m1}x_1 + \cdots + a_{mn}x_n \quad\quad\quad + x_{n+m} &= b_m
\end{aligned}
\tag{3.26}
$$

and $x_1 \geq 0, \ldots, x_{n+m} \geq 0$. The variable x_{n+i} is called a *slack variable* because it represents the amount of slack remaining in the ith constraint. The ith constraint is binding when the slack variable $x_{n+i} = 0$. The coordinates of a corner point can be obtained by setting n of the variables $x_1, \ldots, x_{m+n}$ equal to zero, and then solving the resulting m equations in m unknowns. The corner point is feasible if the remaining m variables (called *basic variables*) turn out to be nonnegative.

Suppose that we have a moderate size linear programming problem, with $n = 50$ variables and $m = 100$ constraints. The number of corner points is equal to the number of possible choices for which 50 of the 150 variables (50 decision variables plus 100 slack variables) we set equal to zero. The number of possible ways to choose 50 out of 150 is $(150!)/(50!)(100!)$, or around 2×10^{40}. A computer program that could check one corner point every nanosecond would take around 8×10^{30} years to solve this problem, which is representative of a typical linear programming application. The simplex method typically calculates only a selected subset of the corner points (a very small fraction of the total). A linear program of size $n = 50$, $m = 100$ can be solved quickly on a mainframe computer. As of this writing, a problem of this size is at the upper limits of what can be solved in a reasonable amount of time on a personal computer, but technological advances will certainly push these limits further in the near future. As a general rule, the execution time for a linear program using the simplex method is proportional to m^3, so that an order–of–magnitude improvement in processor speed more than doubles the size of the linear program that can be handled by a

given machine. For the problems in this book, any good computer implementation of the simplex method will be adequate. It is not recommended that these problems be solved by hand.

Step 3 is to formulate the linear programming model in the standard form. In our problem the decision variables are the number of acres of each crop $x_1, x_2,$ and x_3. We want to maximize the total yield, $y = 400x_1 + 200x_2 + 250x_3$, over the set

$$
\begin{aligned}
3.0x_1 + 1.0x_2 + 1.5x_3 &\leq 1,000 \\
0.8x_1 + 0.2x_2 + 0.3x_3 &\leq 300 \\
x_1 + x_2 + x_3 &\leq 625
\end{aligned}
\tag{3.27}
$$

and $x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$.

Step 4 is to solve the problem. We used a computer implementation of the simplex method called LINDO, written by Linus Schrage. The results of step 4 are shown in Figure 3.15. The optimal solution is $Z = 162,500$ at $x_1 = 187.5$, $x_2 = 437.5$, $x_3 = 0$. Since the slack variables for rows 2 and 4 are zero, the first and third constraints are binding. Since the slack variable for row 3 is equal to 62.5, the second constraint is not binding.

Step 5 is to answer the question. The question was how much of each crop to plant. The optimal solution is to plant 187.5 acres of corn, 437.5 acres of wheat, and no oats. This should yield $162,500. The optimal crop mixture we found uses all 625 acres and all 1,000 acre–ft of irrigation water, but only 237.5 of the available 300 person–hours of labor per week. Thus, there will be 62.5 person–hours per week that may be devoted to other profitable activities, or to leisure.

We will begin our sensitivity analysis by considering the amount of water available for irrigation. This amount will vary as a result of rainfall and temperature, which determine the status of the farm's irrigation pond. It would also be possible to purchase additional irrigation water from a nearby farm. Figure 3.16 illustrates the effect of one additional acre–ft of irrigation water on our optimal solution. Now we can plant an additional half–acre of corn (a more profitable crop), and in fact we save a little bit of labor (0.3 person–hours per week). The net result is an additional $100 in yield.

The $100 is the shadow price for this resource (irrigation water). The farm should be willing to purchase additional irrigation water for up to $100 per acre–ft. Alternatively, it should be unwilling to part with its own irrigation water for less than $100 per acre–ft. In Fig. 3.15, the shadow prices for the three resources (water, labor, and land) are called *dual prices*. They appear next to the corresponding slack variables. An additional acre of land would also be worth $100. Additional labor is worth $0, since there is an excess.

The dollar amount per acre that each crop will yield varies with the weather and the market. Figure 3.17 shows the effect of a slightly higher yield for corn. This would not affect our decision variables x_1, x_2, and x_3 (the amount of each crop we should plant). Our yield increases, of course, by $50x_1 = \$9,375$. It is

```
MAX      400 X1 + 200 X2 + 250 X3
SUBJECT TO
       2)   3 X1 + X2 + 1.5 X3 <=    1000
       3)   0.8 X1 + 0.2 X2 + 0.3 X3 <=   300
       4)   X1 + X2 + X3 <=   625
END


LP OPTIMUM FOUND AT STEP      2

        OBJECTIVE FUNCTION VALUE

      1)     162500.000

  VARIABLE          VALUE          REDUCED COST
       X1        187.500000            .000000
       X2        437.500000            .000000
       X3          .000000           -.000015


      ROW    SLACK OR SURPLUS      DUAL PRICES
       2)            .000000       100.000000
       3)          62.500000            .000000
       4)            .000000        99.999980

 NO. ITERATIONS=      2
```

Figure 3.15: Optimal solution to the farm problem using the linear programming package LINDO.

also interesting to note the change in the shadow prices. Water is at more of a premium when corn is more valuable. (Although both the water and the land constraints are binding, it is the water constraint that keeps us from planting more corn in place of wheat.)

Figure 3.18 shows what happens if oats yield a bit more than expected. A very small change in this parameter has a very significant effect on our optimal decision. Now we plant oats instead of wheat. We also plant considerably less corn than before. Apparently, our model is quite sensitive to this parameter. This being the case, it seems appropriate to consider the sensitivity to this parameter in more depth.

Let c denote the yield (\$/acre) for oats, so that our objective function $f(x) = 400x_1 + 200x_2 + cx_3$. Note that the value of c does not affect the shape of the feasible region S. Several additional model runs were made varying c. For $c \leq 250$ the optimal solution is at the corner point $(187.5, 437.5, 0)$, and for $c > 250$ the optimal solution is at the adjacent corner point $(41.6\bar{6}, 0, 583.3\bar{3})$.

```
MAX      400 X1 + 200 X2 + 250 X3
SUBJECT TO
       2)   3 X1 + X2 + 1.5 X3 <=    1001
       3)   0.8 X1 + 0.2 X2 + 0.3 X3 <=    300
       4)   X1 + X2 + X3 <=    625
END


LP OPTIMUM FOUND AT STEP       0

        OBJECTIVE FUNCTION VALUE

       1)    162600.000

 VARIABLE          VALUE           REDUCED COST
       X1        188.000000            .000000
       X2        437.000000            .000000
       X3           .000000           -.000015


       ROW    SLACK OR SURPLUS     DUAL PRICES
       2)            .000000      100.000000
       3)          62.200000           .000000
       4)            .000000       99.999980


NO. ITERATIONS=       0
```

Figure 3.16: Optimal solution to the farm problem with one extra acre–foot of water using the linear programming package LINDO.

Both points lie on the line formed by the intersection of the two planes

$$3.0x_1 + 1.0x_2 + 1.5x_3 = 1,000$$
$$x_1 + x_2 + x_3 = 625. \tag{3.28}$$

Consider the gradient vector $\nabla f = (400, 200, c)$ at any point along this line. For $c < 250$ the gradient vector points toward the corner point with $x_3 = 0$. For $c > 250$ the gradient vector points toward the corner point with $x_2 = 0$. As c increases, the gradient vector turns away from the former and toward the latter. At $c = 250$ the gradient vector ∇f is perpendicular to the line through these two points. For this value of c, any point along the line segment joining the two corner points will be optimal.

The practical ramification of our model's sensitivity to the parameter c is that we don't know whether we should plant oats or wheat. A small variation in yield would change our optimal decision. In light of the fact that the $/acre yield varies considerably with the weather and the market, it might be best to

```
MAX      450 X1 + 200 X2 + 250 X3
SUBJECT TO
       2)    3 X1 + X2 + 1.5 X3 <=    1000
       3)    0.8 X1 + 0.2 X2 + 0.3 X3 <=    300
       4)    X1 + X2 + X3 <=    625
END

LP OPTIMUM FOUND AT STEP      0

         OBJECTIVE FUNCTION VALUE

       1)    171875.000

   VARIABLE         VALUE         REDUCED COST
       X1       187.500000            .000000
       X2       437.500000            .000000
       X3          .000000          12.500000


     ROW    SLACK OR SURPLUS     DUAL PRICES
       2)          .000000       125.000000
       3)        62.500000          .000000
       4)          .000000        75.000000

NO. ITERATIONS=      0
```

Figure 3.17: Optimal solution to the farm problem with a higher yield for corn using the linear programming package LINDO.

present the farmer with more than one option. Any crop mixture of the form

$$\begin{aligned}
x_1 &= 187.5t + \ \ 41.6\overline{6}(1-t) \\
x_2 &= 437.5t + \ \ \ \ \ \ \ 0(1-t) \\
x_3 &= \ \ \ \ \ \ 0t + 583.3\overline{3}(1-t)
\end{aligned} \qquad (3.29)$$

for any $0 \leq t \leq 1$ will use all the available land and irrigation water. There is too much uncertainty in the figures for the \$/acre yield to tell which option would produce the most profit.

Sometimes sensitivity analysis is performed in response to client feedback after the results of an initial study have been presented. Suppose that after the farmer has seen the results of our analysis, a new seed catalog arrives with an advertisement for a new variety of corn. This variety of corn is more expensive, but is supposed to require less irrigation. Figure 3.19 shows the results of a sensitivity run in which we assumed that the corn planted requires only 2.5 acre–feet of irrigation water per acre (instead of 3.0). The new seed corn yields

```
MAX       400 X1 + 200 X2 + 260 X3
SUBJECT TO
        2)    3 X1 + X2 + 1.5 X3 <=    1000
        3)    0.8 X1 + 0.2 X2 + 0.3 X3 <=    300
        4)    X1 + X2 + X3 <=    625
END


LP OPTIMUM FOUND AT STEP      1

        OBJECTIVE FUNCTION VALUE

        1)     168333.300

  VARIABLE           VALUE          REDUCED COST
        X1        41.666670            .000000
        X2          .000000          13.333340
        X3       583.333300            .000000


        ROW    SLACK OR SURPLUS      DUAL PRICES
        2)            .000000         93.333340
        3)          91.666660           .000000
        4)            .000000        120.000000

 NO. ITERATIONS=        1
```

Figure 3.18: Optimal solution to the farm problem with a higher yield for oats using the linear programming package LINDO.

an additional \$12,500, and in this case we will, of course, plant more corn than before. It is also interesting to note that in this case the shadow price for water has increased 33%.

Finally, suppose that the farmer wishes to investigate the addition of a new crop, barley. An acre of barley requires 1.5 acre–ft of water, 0.25 person–hours of labor, and is expected to yield \$200. We represent the new crop in our model by adding a decision variable $x_4 =$ acres of barley. Figure 3.20 shows our model run. The results are essentially unchanged from our base case. A mixture of corn and wheat remains our optimal solution, and it is easy to see why. Although barley and wheat yield the same, barley requires more water and more labor.

Example 3.5. A large construction firm is currently excavating at three sites. Meanwhile, they are also building at four additional sites, where they require fill dirt. The excavations at sites 1, 2, and 3 produce 150, 400, and 325 cubic yards of dirt per day. The building sites A, B, C, and D require 175, 125, 225, and 450 cubic yards of dirt per day. Additional fill dirt can also be obtained from

```
MAX      400 X1 + 200 X2 + 250 X3
SUBJECT TO
        2)    2.5 X1 + X2 + 1.5 X3 <=    1000
        3)    0.8 X1 + 0.2 X2 + 0.3 X3 <=    300
        4)    X1 + X2 + X3 <=    625
END


LP OPTIMUM FOUND AT STEP       1

        OBJECTIVE FUNCTION VALUE

        1)      175000.000

    VARIABLE           VALUE          REDUCED COST
        X1         250.000000            .000000
        X2         375.000000            .000000
        X3            .000000          16.666680


        ROW    SLACK OR SURPLUS      DUAL PRICES
        2)            .000000         133.333300
        3)          25.000000            .000000
        4)            .000000          66.666660

    NO. ITERATIONS=       1
```

Figure 3.19: Optimal solution to the farm problem with a lower water requirement for corn using the linear programming package LINDO.

site 4 at a cost of 5 dollars per cubic yard. The cost of shipping fill dirt is about 20 dollars per mile for one truckload, and a truck carries 10 cubic yards of dirt. Table 3.3 gives the distance between sites in miles. Find the transportation plan that minimizes the cost to the company.

We will use the five-step method. The results of step 1 are shown in Figure 3.21. For example, we can ship no more than 150 cubic yards of dirt per day from site 1, and we must ship at least 175 cubic yards of dirt to site A. Since it costs $20 per mile to ship 10 cubic yards of dirt, the cost of shipping one cubic yard of dirt from site $i = 1, 2,$ or 3 to site $j = $ A, B, C, or D is $2 per mile. There is an additional cost of $5 per cubic yard if we ship from site $i = 4$. We assume that there is no limit to the amount of dirt we can obtain from site 4, if we are willing to pay the price.

Step 2 is to select the modeling approach. We will model this problem as a linear programming problem, and we will solve using a spreadsheet. Most spreadsheets include an equation solver or optimizer that implements the sim-

```
MAX      400 X1 + 200 X2 + 250 X3 + 200 X4
SUBJECT TO
        2)   3 X1 + X2 + 1.5 X3 + 1.5 X4 <=    1000
        3)   0.8 X1 + 0.2 X2 + 0.3 X3 + 0.25 X4 <=    300
        4)   X1 + X2 + X3 + X4 <=    625
END

LP OPTIMUM FOUND AT STEP      1

        OBJECTIVE FUNCTION VALUE

        1)    162500.000

  VARIABLE         VALUE          REDUCED COST
       X1       187.500000            .000000
       X2       437.500000            .000000
       X3          .000000           -.000015
       X4          .000000          49.999980


       ROW    SLACK OR SURPLUS      DUAL PRICES
       2)          .000000         100.000000
       3)        62.500000            .000000
       4)          .000000          99.999980

 NO. ITERATIONS=      1
```

Figure 3.20: Optimal solution to the farm problem with the addition of a new crop, barley, using the linear programming package LINDO.

plex method. If the linear programming problem is not in the standard form defined by Eq. (3.23) or (3.26), a simple transformation usually suffices. For example, to minimize the objective function $y = f(x_1, \ldots, x_n)$, we can maximize $-y$. Most implementations perform these transformations automatically, allowing for a more natural problem formulation. There is also a special streamlined version of the simplex method specially formulated for transportation problems. This transportation simplex method is much more efficient for large–scale problems. For the present problem, the ordinary simplex method will suffice.

Step 3 is to formulate the model. In this problem the decision variables are the number of cubic feet of dirt x_{ij} we want to ship from site i to site j, and

| | Site to receive fill dirt | | | |
Excavation site	A	B	C	D
1	5	2	6	10
2	4	5	7	5
3	7	6	4	4
4	9	10	6	2

Table 3.3: Mileage data for the dirt problem of Example 3.5: Miles between building sites.

Variables: x_{ij} = dirt shipped from site i to site j (cubic yards)
                      s_i = dirt shipped from site i (cubic yards)
                      r_j = dirt shipped to site j (cubic yards)
                      c_{ij} = cost of shipping dirt from site i to site j (\$/cubic yard)
                      d_{ij} = distance from site i to site j (miles)
                      C = total shipping cost (\$)

Assumptions: $s_1 = x_{1A} + x_{1B} + x_{1C} + x_{1D}$
                      $s_2 = x_{2A} + x_{2B} + x_{2C} + x_{2D}$
                      $s_3 = x_{3A} + x_{3B} + x_{3C} + x_{3D}$
                      $s_4 = x_{4A} + x_{4B} + x_{4C} + x_{4D}$
                      $r_A = x_{1A} + x_{2A} + x_{3A} + x_{4A}$
                      $r_B = x_{1B} + x_{2B} + x_{3B} + x_{4B}$
                      $r_C = x_{1C} + x_{2C} + x_{3C} + x_{4C}$
                      $r_D = x_{1D} + x_{2D} + x_{3D} + x_{4D}$
                      $s_1 \le 150, s_2 \le 400, s_3 \le 325$
                      $r_A \ge 175, r_B \ge 125, r_C \ge 225, r_D \ge 450$
                      $c_{ij} = 2d_{ij}$ if $i = 1, 2, 3$ or $(2d_{ij} + 5)$ if $i = 4$
                        where d_{ij} is given in Table 3.3
                      $C = c_{1A}x_{1A} + c_{1B}x_{1B} + c_{1C}x_{1C} + c_{1D}x_{1D}$
                        $+ c_{2A}x_{2A} + c_{2B}x_{2B} + c_{2C}x_{2C} + c_{2D}x_{2D}$
                        $+ c_{3A}x_{3A} + c_{3B}x_{3B} + c_{3C}x_{3C} + c_{3D}x_{3D}$
                        $+ c_{4A}x_{4A} + c_{4B}x_{4B} + c_{4C}x_{4C} + c_{4D}x_{4D}$
                      $x_{ij} \ge 0$ for $i = 1, 2, 3, 4$ and $j = $ A, B, C, D

Objective: Minimize C

Figure 3.21: Results of step 1 of the dirt problem.

A	A	B	C	D	E	F	G
1	costs						
2	site	A	B	C	D		
3	1	10	4	12	20		
4	2	8	10	14	10		
5	3	14	12	8	8		
6	4	23	25	17	9		
7	solution						
8	site	A	B	C	D	shipped	available
9	1	0	0	0	0	0	150
10	2	0	0	0	0	0	400
11	3	0	0	0	0	0	325
12	4	0	0	0	0	0	
13	received	0	0	0	0		
14	needed	175	125	225	450		
15	total cost	0					

Figure 3.22: Spreadsheet formulation for the dirt problem.

our objective is to minimize the total cost $y = C$ where

$$
\begin{aligned}
y = \ & 10x_{1A} + \ 4x_{1B} + 12x_{1C} + 20x_{1D} \\
+ \ & 8x_{2A} + 10x_{2B} + 14x_{2C} + 10x_{2D} \\
+ \ & 14x_{3A} + 12x_{3B} + \ 8x_{3C} + \ 8x_{3D} \\
+ \ & 23x_{4A} + 25x_{4B} + 17x_{4C} + \ 9x_{4D}
\end{aligned}
\tag{3.30}
$$

subject to the constraints

$$
\begin{aligned}
x_{1A} + x_{1B} + x_{1C} + x_{1D} &\le 150 \\
x_{2A} + x_{2B} + x_{2C} + x_{2D} &\le 400 \\
x_{3A} + x_{3B} + x_{3C} + x_{3D} &\le 325 \\
x_{1A} + x_{2A} + x_{3A} + x_{4A} &\ge 175 \\
x_{1B} + x_{2B} + x_{3B} + x_{4B} &\ge 125 \\
x_{1C} + x_{2C} + x_{3C} + x_{4C} &\ge 225 \\
x_{1D} + x_{2D} + x_{3D} + x_{4D} &\ge 450
\end{aligned}
\tag{3.31}
$$

and $x_{ij} \ge 0$ for $i = 1, 2, 3, 4$ and $j =$ A, B, C, D.

Figure 3.22 shows a spreadsheet set up for this problem. Most of the cells contain data. Cells F9 to F12, B13 to E13, and B15 contain formulas. For example, we input $F9 = B9 + C9 + D9 + E9$ and $B13 = B9 + B10 + B11 + B12$ while $B15 = B3 * B9 + C3 * C9 + \cdots + D6 * D12 + E6 * E12$. Then we must tell the spreadsheet which cell contains the objective, which cells contain decision variables, and we must specify the constraints. The details vary from one spreadsheet to another; consult the manual or on–line help facility for your spreadsheet if you are not sure how to proceed.

A	A	B	C	D	E	F	G
1	costs						
2	site	A	B	C	D		
3	1	10	4	12	20		
4	2	8	10	14	10		
5	3	14	12	8	8		
6	4	23	25	17	9		
7	solution						
8	site	A	B	C	D	shipped	available
9	1	0	125	0	0	125	150
10	2	175	0	0	0	175	400
11	3	0	0	225	100	325	325
12	4	0	0	0	350	350	
13	received	175	125	225	450		
14	needed	175	125	225	450		
15	total cost	7650					

Figure 3.23: Spreadsheet solution for the dirt problem showing the optimal shipping plan.

Step 4 is to solve the problem. Figure 3.23 shows the spreadsheet solution for this problem obtained using the Quattro Pro optimizer. Notice that all of the receiving constraints are binding, but not all the shipping constraints are binding. Step 5 is to answer the question. The optimal solution is to ship 125 cubic yards of dirt per day from site 1 to site B, and 175 cubic yards from site 2 to site A. Site 3 sends 225 cubic yards to site C and another 100 cubic yards to site D. The remaining 350 cubic yards of dirt needed at site D come from site 4. The added cost of purchasing dirt is offset by the fact that site 4 is close to site D. The total cost of this transportation plan is $7,650 per day. According to this shipping plan, we will not use all the dirt excavated from sites 1 and 2, so we will need to make other arrangements to dispose of the extra dirt.

We begin our sensitivity analysis by examining some of the reports that can be automatically generated by the spreadsheet optimizer. Figure 3.24 shows a sensitivity report on the constraints in our problem. Notice that the constraints on site 1 and site 2 are nonbinding, with a slack of 25 and 225 respectively. This means that there are 25 cubic yards of dirt being excavated from site 1, and 225 cubic yards from site 2, that are not being shipped under the optimal transportation plan. The dual values listed in Figure 3.24 are the shadow prices. The shadow prices for sites 1 and 2 are zero, since these constraints are nonbinding. The shadow price for site 3 is −1 dollar, meaning that the total cost would increase by −1 dollar if this constraint were increased by 1 cubic yard of dirt. One of the principal advantages of the spreadsheet implementation is the ease with which we can perform sensitivity analysis. We simply change the value in cell F12 from 325 to 326 and reoptimize. The resulting optimal solution (not shown) changes E11 to 101, E12 to 349 and B15 to 7649. All the other decision

Cell	Value	Constraint	Binding?	Slack	Dual Value	Increment	Decrement
site 1	125	<=150	No	25	0	Infinite	25
site 2	175	<=400	No	225	0	Infinite	225
site 3	325	<=325	Yes	0	-1	350	100
site A	175	>=175	Yes	0	8	225	175
site B	125	>=125	Yes	0	4	25	125
site C	225	>=225	Yes	0	9	100	225
site D	450	>=450	Yes	0	9	Infinite	350

Figure 3.24: Spreadsheet sensitivity report for the dirt problem showing sensitivity to changes in the amount of dirt available or the amount of dirt needed.

variables are unchanged. In other words, we ship one more cubic yard of dirt from site 3 to site D and we save \$1 per day.

The shadow prices for site A, B, C, D are all positive; e.g., the shadow price for site C is 9 dollars. If we require an additional 10 cubic yards at site C, then this will cost an additional \$90. To check this, we change cell D14 from 225 to 235 and reoptimize. The resulting optimal solution is shown in Figure 3.25. Note that the total cost increases by \$90 from \$7,650 to \$7,740. Now site C received 235 cubic yards from site 3, while site D receives 90 cubic yards from site 3 and 360 cubic yards from site 4. The other shipments remain the same.

The columns labelled "increment" and "decrement" show the amount of increase or decrease in the individual constraint amounts for which the shadow prices remain valid. If we increase the amount required at site C from 225 to 325 or less, then the total cost will increase by \$9 per cubic yard. If we go beyond this amount, then the nature of the solution will change. In this case it is easy to see why the nature of the solution changes if we need more than 325 cubic yards of dirt per day at site C. Right now we are getting all of the dirt for site C from site 3, but if we need more then 325 cubic yards at site C, then we must haul dirt from another site at a higher cost.

Figure 3.26 shows the spreadsheet report on sensitivity to objective function coefficients. The increment and decrement amounts indicate the amount of increase or decrease we can make in one of these cost–per–mile numbers without changing the optimal solution. For example, the current cost of shipping one cubic yard from site 1 to site B is 4 dollars. Lowering the cost by less than 4 dollars or raising the price by less than 6 dollars will not change the optimal shipping plan. The total cost will change, since we are shipping 125 cubic yards per day from site 1 to site B. Consider the geometry of the simplex method. When we change an objective function coefficient, we are not changing the feasible region at all. The optimal solution will only change if we alter the objective function enough so that the current corner point solution is no longer optimal. Then the optimal solution jumps to another corner point. We can

A	A	B	C	D	E	F	G
1	costs						
2	site	A	B	C	D		
3	1	10	4	12	20		
4	2	8	10	14	10		
5	3	14	12	8	8		
6	4	23	25	17	9		
7	solution						
8	site	A	B	C	D	shipped	available
9	1	0	125	0	0	125	150
10	2	175	0	0	0	175	400
11	3	0	0	235	90	325	325
12	4	0	0	0	360	360	
13	received	175	125	235	450		
14	needed	175	125	235	450		
15	total cost	7740					

Figure 3.25: Spreadsheet solution for the dirt problem with 10 additional cubic yards of dirt needed at site C.

Variable Cells	Starting	Final	Gradient	Increment	Decrement
x1A	0	0	10	Infinite	2
x2A	0	175	8	2	8
x3A	0	0	14	Infinite	7
x4A	0	0	23	Infinite	15
x1B	0	125	4	6	4
x2B	0	0	10	Infinite	6
x3B	0	0	12	Infinite	9
x4B	0	0	25	Infinite	21
x1C	0	0	12	Infinite	3
x2C	0	0	14	Infinite	5
x3C	0	225	8	3	9
x4C	0	0	17	Infinite	8
x1D	0	0	20	Infinite	11
x2D	0	0	10	Infinite	1
x3D	0	100	8	1	3
x4D	0	350	9	1	1

Figure 3.26: Spreadsheet sensitivity report for the dirt problem showing sensitivity to changes in the shipping costs.

A	A	B	C	D	E	F	G
1	costs						
2	site	A	B	C	D		
3	1	10	4	12	20		
4	2	8	10	14	10		
5	3	14	12	8	8		
6	4	23	25	17	9		
7	solution						
8	site	A	B	C	D	shipped	available
9	1	25	125	0	0	150	150
10	2	150	0	0	250	400	400
11	3	0	0	225	100	325	325
12	4	0	0	0	100	100	
13	received	175	125	225	450		
14	needed	175	125	225	450		
15	total cost	7950					

Figure 3.27: Spreadsheet solution for the dirt problem if we ship all of the excavated dirt.

verify this by changing the value in cell C3 and reoptimizing. If we insert any value between 0 and 10 into C3, we get the same shipping plan as in Figure 3.23. If we insert 11 into C3, then the shipping plan changes so that site B receives 125 cubic yards of dirt per day from site 2 and none from site 1.

Finally, we consider the robustness of our model. Our optimal solution in Figure 3.23 indicates that we should not ship all of the excavated dirt to another construction site. This leaves the company with a problem. Where do we put all the extra dirt? The company will need to put the extra dirt somewhere, and this will result in some additional cost. We do not have information about these costs, but we can explore some options. Suppose that we ship all of the excavated dirt to another construction site to be used as fill dirt. We know this is not the optimal solution, but how much more would it cost? We can find out by making a small change in our model. In step 1 we assumed that $s_1 \leq 150$, $s_2 \leq 400$, $s_3 \leq 325$. Assume now that $s_1 = 150$, $s_2 = 400$, $s_3 = 325$. The first three inequality constraints in Eq. (3.31) are replaced by

$$x_{1A} + x_{1B} + x_{1C} + x_{1D} = 150$$
$$x_{2A} + x_{2B} + x_{2C} + x_{2D} = 400 \qquad (3.32)$$
$$x_{3A} + x_{3B} + x_{3C} + x_{3D} = 325$$

and the rest of our linear programming problem remains the same. After changing these three constraints on our spreadsheet, we reoptimize to obtain the results shown in Figure 3.27. Sites B and C receive the same shipments as before, but now site A gets 25 cubic yards of fill dirt per day from site 1 and site D gets 250 cubic yards of fill dirt per day from site 2. According to this transportation plan, we only need to purchase 100 cubic yards from site 4. The new plan gets

rid of the extra 250 cubic yards of dirt and the additional cost is only $300 per day, a little more than one dollar per cubic yard. We do not really know which of these two plans is best. If the company needs to haul the excavated dirt away from the site, then they may favor the alternate plan. If they can use the dirt at or near the excavation site, they may prefer the original plan. We should present both plans to the company management and let them choose.

3.4 Discrete Optimization

All of the models used so far in this book have been based on continuous variables. In many real–world problems we must deal with variables that are discrete, like the integers. Discrete mathematics was once considered an arcane field with little or no practical applications. With the invention of digital computers, discrete mathematics has become extremely important. Discrete optimization is useful for problems in scheduling, inventory, investment, transportation, manufacturing, ecology, and computer science. Discrete models play an important role in the remainder of the text, and the relationship between continuous and discrete variables is one of the major issues in modeling.

In some cases a discrete optimization problem can be solved by simply enumerating the possibilities. In other cases we can employ continuous methods and then round off to the nearest integer. Nonlinear programming problems usually become much harder to solve when continuous decision variables are replaced by discrete variables. The feasible region becomes more complex when continuity is lost, and is often represented by a graph or tree structure. Effective solution algorithms have been developed for some classes of problems, and the improvement of these algorithms is a very active area of research. However, as in the continuous case, there are still no universally effective methods for solving discrete optimization problems.

In this section we will concentrate on one type of discrete optimization problem called *integer programming*. Integer programming is the discrete analog of the linear programming model considered in the preceding section. Aside from being the most commonly applied discrete optimization algorithm, the parallel with linear programmming will facilitate comparison between discrete and continuous models. An added bonus is that most linear programming implementations also solve integer programming problems, so that we can concentrate our attention on the model rather than learning a new software package.

Example 3.6. Reconsider the family farm problem of Example 3.4. The family has 625 acres available for planting. There are 5 plots of 120 acres each and another plot of 25 acres. The family wants to plant each plot with only one crop: corn, wheat, or oats. As before, 1,000 acre–ft of water will be available for irrigation, and the farmers will be able to devote 300 hours of labor per week. Additional data are presented in Table 3.2. Find the crop that should be planted in each plot for maximum profit.

We will use the five-step method. The results of step 1 are shown in Figure

3.28. Step 2 is to select the modeling approach. We will model this problem as
an integer programming problem.

Variables: $x_1 = $ number of 120 acre plots of corn planted
$x_2 = $ number of 120–acre plots of wheat planted
$x_3 = $ number of 120–acre plots of oats planted
$x_4 = $ number of 25–acre plots of corn planted
$x_5 = $ number of 25–acre plots of wheat planted
$x_6 = $ number of 25–acre plots of oats planted
$w = $ irrigation required (acre–ft)
$l = $ labor required (person–hrs/wk)
$t = $ total acreage planted
$y = $ total yield ($)

Assumptions: $w = 120(3.0x_1 + 1.0x_2 + 1.5x_3)$
$\quad + 25(3.0x_4 + 1.0x_5 + 1.5x_6)$
$l = 120(0.8x_1 + 0.2x_2 + 0.3x_3)$
$\quad + 25(0.8x_4 + 0.2x_5 + 0.3x_6)$
$t = 120(x_1 + x_2 + x_3) + 25(x_4 + x_5 + x_6)$
$y = 120(400x_1 + 200x_2 + 250x_3)$
$\quad + 25(400x_4 + 200x_5 + 250x_6)$
$w \le 1,000$
$l \le 300$
$t \le 625$
$x_1 + x_2 + x_3 \le 5$
$x_4 + x_5 + x_6 \le 1$
$x_1, \ldots, x_6$ are nonnegative integers

Objective: Maximize y

Figure 3.28: Results of step 1 of the modified farm problem.

An integer programming (IP) problem is a linear programming
(LP) problem in which the decision variables are further constrained
to take integer values. Both the objective function and the con-
straints must be linear. The most commonly used method for solving
an IP is the method of branch–and–bound. This method involves
branching through repeated solutions of LPs chosen to bound the
solution of the IP. If we remove the constraint that the decision
variables in a given IP take only integer values, we obtain what is
known as the LP relaxation. Since the feasible region of the LP
relaxation is larger than the feasible region of the corresponding IP,
any optimal solution of the LP relaxation for which all of the deci-
sion variables take integer values is also an optimal solution of the
IP. If some decision variable is not an integer, we branch to create

two additional LP relaxations. For example, if the optimal solution to the LP relaxation yields $x_1 = 11/3$, then we consider the original problem together with the additional constraint $x \leq 3$, and we also consider the original problem together with the additional constraint $x \geq 4$. Any integer solution must satisfy one of these additional constraints. By continuing to branch any time we find an optimal solution with a fractional decision variable, we obtain a binary tree of LP relaxations. If one of the new LPs has an optimal solution with all integer decision variables, this becomes a candidate for the solution to the original IP. Since the new LP has a smaller feasible region than the original, this integer solution also gives a useful lower bound on the optimal solution of the original IP. By systematically branching through the tree, and using these solution bounds, we will eventually solve the original IP. Because the branch–and–bound method requires many iterations of the simplex method, integer programming problems generally take much longer to solve than linear programming problems of the same size.

Step 3 is to formulate the problem. In our problem the decision variables are the number of 120–acre plots and the number of 25–acre plots to plant with corn, wheat, or oats. Notice that the variables x_4, x_5, and x_6 are binary decision variables. They can only take the values 0 or 1. Our integer programming problem in standard form is to maximize the total yield $y = 48000x_1 + 24000x_2 + 30000x_3 + 10000x_4 + 5000x_5 + 6250x_6$ over the set

$$
\begin{aligned}
375x_1 + 125x_2 + 187.5x_3 + 75x_4 + 25x_5 + 37.5x_6 &\leq 1000 \\
100x_1 + 25x_2 + 37.5x_3 + 20x_4 + 5x_5 + 7.5x_6 &\leq 300 \\
x_1 + x_2 + x_3 &\leq 5 \\
x_4 + x_5 + x_6 &\leq 1
\end{aligned}
\tag{3.33}
$$

where $x_1, \ldots, x_6$ are nonnegative integers.

Step 4 is to solve the problem. Figure 3.29 shows the solution of our integer programming problem using the popular linear programming package LINDO. The GIN 6 command specifies that the first six decision variable are nonnegative integers. Everything else in the problem specification is the same for the IP and the LP. The optimal solution is $y = 156,250$ which occurs when $x_3 = 5$, $x_6 = 1$, and the other decision variables are all zero. At the optimal solution the first two constraits are nonbinding, and the last two are binding.

Step 5 is to answer the question. If the family does not wish to split up individual plots (plan B), then the best plan is to plant oats in every plot. This results in an expected total yield of \$156,250 for the season. This is about 4% less than the projected total yield of \$162,500 if we allow more than one crop per plot (plan A, the optimal solution found in Example 3.4). Plan A uses all of the acreage available, all of the irrigation water available, and all but 62.5 of the 300 person–hours of labor available each week. Plan B uses all of the acreage available, but only 975 of the 1000 available acre–feet of irrigation water, and

```
MAX  48000 X1 + 24000 X2 + 30000 X3 + 10000 X4 + 5000 X5 + 6250 X6
SUBJECT TO
   2)    375 X1 + 125 X2 + 187.5 X3 + 75 X4 + 25 X5 + 37.5 X6 <= 1000
   3)    100 X1 + 25 X2 + 37.5 X3 + 20 X4 + 5 X5 + 7.5 X6 <= 300
   4)    X1 + X2 + X3 <= 5
   5)    X4 + X5 + X6 <= 1
END
GIN      6


        OBJECTIVE FUNCTION VALUE

        1)       156250.0

   VARIABLE          VALUE           REDUCED COST
        X1          0.000000       -48000.000000
        X2          0.000000       -24000.000000
        X3          5.000000       -30000.000000
        X4          0.000000       -10000.000000
        X5          0.000000        -5000.000000
        X6          1.000000        -6250.000000


        ROW    SLACK OR SURPLUS      DUAL PRICES
        2)          25.000000          0.000000
        3)         105.000000          0.000000
        4)           0.000000          0.000000
        5)           0.000000          0.000000


   NO. ITERATIONS=     101
```

Figure 3.29: Optimal solution to the modified farm problem using the linear programming package LINDO.

only 195 of the available 300 person–hours of labor available each week. We leave it to the family to decide which plan is best.

Sensitivity analysis can be very time–consuming for integer programming problems, because IPs take so much longer to solve than LPs. There are no shadow prices to guide us, because the value of the objective function at the optimal solution does not change smoothly as the constraints vary. Integer solutions will not usually lie exactly on the constraint boundaries, so the optimum may also be sensitive to small changes in nonbinding constraints. We begin by considering the amount of irrigation water available. Suppose that an additional 100 acre–feet of water is available. This changes our IP in Eq. (3.33) by substituting 1100 for 1000 in the first constraint. Figure 3.30 shows the LINDO solution to this IP. Now we plant one 120–acre plot of corn, one 120–acre plot

of wheat, and we plant oats everywhere else. Our optimal solution is quite sensitive to the amount of irrigation water available, even though this constraint was not binding in our original IP solution. The new plan yields an additional $12,000 in expected revenue.

```
MAX  48000 X1 + 24000 X2 + 30000 X3 + 10000 X4 + 5000 X5 + 6250 X6
SUBJECT TO
  2)    375 X1 + 125 X2 + 187.5 X3 + 75 X4 + 25 X5 + 37.5 X6 <= 1100
  3)    100 X1 + 25 X2 + 37.5 X3 + 20 X4 + 5 X5 + 7.5 X6 <= 300
  4)    X1 + X2 + X3 <= 5
  5)    X4 + X5 + X6 <= 1
END
GIN      6


        OBJECTIVE FUNCTION VALUE

        1)       168250.0

   VARIABLE           VALUE           REDUCED COST
        X1          1.000000         -48000.000000
        X2          1.000000         -24000.000000
        X3          3.000000         -30000.000000
        X4          0.000000         -10000.000000
        X5          0.000000          -5000.000000
        X6          1.000000          -6250.000000


        ROW    SLACK OR SURPLUS       DUAL PRICES
        2)          0.000000           0.000000
        3)         55.000000           0.000000
        4)          0.000000           0.000000
        5)          0.000000           0.000000


 NO. ITERATIONS=        62
```

Figure 3.30: Optimal solution to the modified farm problem with an additional 100 acre–feet of water available.

The optimal solution in Figure 3.29 shows that there are 25 acre–feet of unused irrigation water. If we lower the amount of water available, we will not change the optimal solution unless we go below 975 acre–feet. This is because the optimal solution will remain feasible, and since we are shrinking the feasible region, it must remain optimal as well. Figure 3.31 shows what happens if we have only 950 acre–feet of water available. The optimal IP solution is to plant one 120–acre plot of corn, four 120–acre plots of wheat, and to plant corn instead of oats in the smaller 25 acre plot. We use all the water and all the land, but we have 80 person–hours of labor per week to spare. The expected

```
 MAX   48000 X1 + 24000 X2 + 30000 X3 + 10000 X4 + 5000 X5 + 6250 X6
SUBJECT TO
  2)    375 X1 + 125 X2 + 187.5 X3 + 75 X4 + 25 X5 + 37.5 X6 <= 950
  3)    100 X1 + 25 X2 + 37.5 X3 + 20 X4 + 5 X5 + 7.5 X6 <= 300
  4)    X1 + X2 + X3 <= 5
  5)    X4 + X5 + X6 <= 1
END GIN      6


        OBJECTIVE FUNCTION VALUE

        1)        154000.0

   VARIABLE          VALUE         REDUCED COST
        X1         1.000000     -48000.000000
        X2         4.000000     -24000.000000
        X3         0.000000     -30000.000000
        X4         1.000000     -10000.000000
        X5         0.000000      -5000.000000
        X6         0.000000      -6250.000000


        ROW    SLACK OR SURPLUS      DUAL PRICES
        2)             0.000000         0.000000
        3)            80.000000         0.000000
        4)             0.000000         0.000000
        5)             0.000000         0.000000

  NO. ITERATIONS=       4
```

Figure 3.31: Optimal solution to the modified farm problem with 50 acre–feet less water available.

total yield is \$154,000, which is only \$2,250 less than before. This illustrates the unpredictable nature of IP solutions. For a 5% decrease in the amount of available irrigation water, instead of planting oats everywhere, we should not plant oats anywhere.

The main robustness issue is the relation between discrete and continuous optimization. We have already seen that devoting large plots to a single crop dramatically changes the optimal planting scheme. Now we will return to the original farm problem and examine the effect of changing the minimum plot size. Table 3.4 shows the results of several LINDO runs with different minimum plot sizes. For example, when the minimum plot size is 2 acres, we maximize

Minimum Plot Size (acres)	Corn (acres)	Wheat (acres)	Oats (acres)	Yield (dollars)
0	187.5	437.5	0	162,500
1	42	1	582	162,500
2	188	436	0	162,400
5	45	10	570	162,500
10	190	430	0	162,000
20	60	40	520	162,000
50	200	400	0	160,000
100	200	400	0	160,000
125	125	250	250	162,500
150	150	300	150	157,000
200	200	400	0	160,000
250	250	250	0	150,000
300	0	0	600	150,000
500	0	0	500	125,000

Table 3.4: Comparison of optimal planting schemes for the farm problem for different minimum plot sizes.

total yield $y = 800x_1 + 400x_2 + 500x_3$ over the set

$$
\begin{aligned}
6.0x_1 + 2.0x_2 + 3.0x_3 &\le 1000 \\
1.6x_1 + 0.4x_2 + 0.6x_3 &\le 300 \\
x_1 + x_2 + x_3 &\le 312
\end{aligned}
\tag{3.34}
$$

and x_1, x_2, x_3 are nonnnegative integers representing the number of 2–acre plots of corn, wheat, and oats to plant. Figure 3.32 shows the optimal solution for a minimum plot size of 2 acres using LINDO. The optimal solution is to plant 94 plots of corn (for a total of $2 \times 94 = 188$ acres), 218 plots of wheat (436 acres), and no oats. For this model excursion we do not plant in any leftover acreage, so that when the minimum plot size is 2 acres, we will leave 1 acre unplanted. The optimal solution in Table 3.4 changes significantly as the minimum plot size increases. Recall that, in the continuous optimization problem (LP), there are two optimal corner point solutions $(x_1, x_2, x_3) = (187.5, 437.5, 0)$ and $(x_1, x_2, x_3) = (41.6\bar{6}, 0, 583.3\bar{3})$ at which we achieve the optimal yield of $y = 162,500$. For small minimum plot sizes, the best planting scheme jumps between discrete approximations to these two solutions. For minimum plot sizes of 2, 10, and 50 acres, the solution is similar to the original corn–and–wheat planting scheme. For minimum plot sizes of 1, 5, and 20 acres, the solution is similar to the alternate corn–and–oats planting scheme. When the minimum plot size gets larger, both the optimal planting scheme and the expected total yield vary considerably. A larger minimum plot size may result in either a larger or smaller yield. For example, when we use a plot size of 125 acres, we obtain a scheme which matches the original $162,500 expected total yield. Note that

```
MAX      800 X1 + 400 X2 + 500 X3
SUBJECT TO
        2)    6 X1 + 2 X2 + 3 X3 <=    1000
        3)    1.6 X1 + 0.4 X2 + 0.6 X3 <=    300
        4)    X1 + X2 + X3 <=    312
END
GIN     3


       OBJECTIVE FUNCTION VALUE

     1)        162400.0

VARIABLE          VALUE           REDUCED COST
     X1         94.000000         -800.000000
     X2        218.000000         -400.000000
     X3          0.000000         -500.000000


     ROW    SLACK OR SURPLUS      DUAL PRICES
     2)          0.000000          0.000000
     3)         62.399998          0.000000
     4)          0.000000          0.000000


NO. ITERATIONS=      2
```

Figure 3.32: Optimal solution to the farm problem with a minimum plot size of 2 acres.

the 625 acres divide evenly into 125–acre plots, and that the optimal solution $(x_1, x_2, x_3) = (125, 250, 250)$ lies on the line segment of Eq. (3.29) giving the optimal solutions for the original LP. Most larger plot sizes give a considerably smaller expected yield, however, and the planting schemes vary widely.

Consider the geometry of the LP and its various IP approximations. When we choose a minimum plot size, we are restricting the feasible region to a lattice of integer points. If the plot size is small, then there are many such points, and they cover most of the feasible region, in the sense that there is a lattice point near every feasible point. Since the objective function is continuous, we can find a lattice point solution for the IP that is nearly optimal for the original LP. But when the lattice points are widely spaced, discretization will often change the optimal solution significantly, since there may be no lattice point near the LP optimum. Generally speaking, discretization makes little difference as long as the jumps between lattice points (in our example, the minimum plot size) represent a small percent change in the decision variables. Otherwise the IP solution may be very different than the solution to the corresponding LP relaxation.

Route	From	To	Miles	Cubic Yards
1	1	B	2	125
2	2	A	4	175
3	3	C	4	225
4	3	D	4	100
5	4	D	4	350

Table 3.5: Route data for the truck problem of Example 3.7.

Example 3.7. Reconsider the dirt problem of Example 3.5. The company has already determined the optimal shipping plan, assuming that dump trucks with a 10 cubic yard capacity are being used. The company also has 3 larger trucks available that can carry 20–cubic–yard loads. It may be possible to use these trucks to save some money on shipping. The 10–cubic–yard trucks take an average of 25 minutes to load, 5 minutes to unload, and travel at an average of 20 miles per hour. The cost of operating these trucks is $20 per mile per load. The larger 20–cubic–yard trucks take 35 minutes to load, 5 minutes to unload, and travel at an average of 20 miles per hour. The cost of operating these trucks is $30 per mile per load. How should these trucks be assigned in order to maximize our savings on shipping costs?

We will use the five-step method. Step 1 is to ask a question. The question is which trucks to assign to which routes. The optimal shipping routes found in Example 3.5 are shown in Table 3.5. There are 5 routes of varying length, and the amount of dirt shipped also varies. We will assume that either 10–cubic–yard trucks or 20–cubic–yard trucks, but not both, are used on each route. Since the larger trucks carry twice as much dirt at less than twice the cost, we want to assign these trucks to the routes where they will save us the most money. We need to compute the number of trucks needed and the total cost of shipping for each route and each type of truck, and then compute the potential savings for using larger trucks on each route. The optimal shipping plan calls for 125 cubic yards of dirt to be shipped from site 1 to site B, a distance of 2 miles. It takes the smaller trucks 20 minutes to load, 5 minutes to unload, and 6 minutes of travel time at 20 miles per hour, for a total of 31 minutes per load. We assume an 8–hour workday, so that each truck can be assigned no more than 480 minutes of work per day. It will take 13 loads to move the 125 cubic yards of dirt, for a total of $13 \times 31 = 403$ minutes of work, so one small truck is sufficient to service route 1. The cost of servicing route 1 with one small truck is

$$13 \text{ loads} \times \frac{2 \text{ miles}}{\text{load}} \times \frac{20 \text{ dollars}}{\text{mile}} = 520 \text{ dollars}.$$

It takes the larger trucks 30 minutes to load, 5 minutes to unload, and 6 minutes of travel time on route 1, for a total of 41 minutes per load. It will take 7 loads to move the 125 cubic yards of dirt, for a total of $7 \times 41 = 287$ minutes of work, so one large truck is enough to service route 1. The cost of servicing route 1

Variables: $x_i = 1$ if we use large trucks on route i
$x_i = 0$ if we use small trucks on route i
$T =$ total number of large trucks used
$y =$ total cost savings (\$)

Assumptions: $T = 1x_1 + 1x_2 + 2x_3 + 1x_4 + 2x_5$
$y = 100x_1 + 360x_2 + 400x_3 + 200x_4 + 640x_5$
$T \le 3$

Objective: Maximize y

Figure 3.33: Results of step 1 of the truck problem.

A	A	B	C	D	E	F
1		possible			actual	
2	route	trucks	savings	decision	trucks	savings
3	1	1	100	0	0	0
4	2	1	360	0	0	0
5	3	2	400	0	0	0
6	4	1	200	0	0	0
7	5	2	640	0	0	0
8	available		3		totals	0	0

Figure 3.34: Spreadsheet formulation for the truck problem.

with one large truck is

$$7 \text{ loads} \times \frac{2 \text{ miles}}{\text{load}} \times \frac{30 \text{ dollars}}{\text{mile}} = 420 \text{ dollars.}$$

The potential savings on route 1 is \$100. The savings and number of trucks required on the remaining routes is calculated similarly. The results of step 1 are summarized in Figure 3.33.

Step 2 is to select a modeling approach. We will model this problem as a binary integer programming (BIP) problem. A BIP is an IP with binary decision variables that can only take the integer values 0 or 1. It is common to use BIP to represent YES/NO decisions. Typical applications include assignment problems, scheduling, facility location, and investment portfolios. Specialized algorithms are available for BIPs that are much faster than the usual IP algorithms. For the small BIP problems in this book, any IP or BIP solver will suffice.

Step 3 is to formulate the problem. Figure 3.34 shows a spreadsheet formulation for this problem. Column D contains the binary decision variables x_i, 0 if small trucks are used and 1 if large trucks are used. Column E counts the number of large trucks used, and column F tallies our cost savings. For example,

A	A	B	C	D	E	F
1		possible			actual	
2	route	trucks	savings	decision	trucks	savings
3	1	1	100	0	0	0
4	2	1	360	1	1	360
5	3	2	400	0	0	0
6	4	1	200	0	0	0
7	5	2	640	1	2	640
8	available	3		totals	3	1000

Figure 3.35: Spreadsheet solution for the truck problem.

Trucks	Routes	Savings ($)	Marginal Savings ($)
1	2	360	360
2	5	640	280
3	2,5	1000	360
4	2,4,5	1200	200
5	2,3,5	1400	200
6	2,3,4,5	1600	200
7	all	1700	100

Table 3.6: Sensitivity to number of large trucks available for the truck problem.

E3=B3*D3 and F3=C3*D3, while E8=E3+$\cdots$+E7. We use the spreadsheet solver to maximize the total F8=F3+$\cdots$+F7 with constraints E8$\leq$B8, and D3 through D7 binary integers.

Step 4 is to solve the problem. Figure 3.35 shows the optimal assignment of trucks to routes, obtained from the spreadsheet optimization utility. The optimum value of $y = 1000$ is obtained at $x_2 = 1$, $x_5 = 1$, and $x_i = 0$ for $i = 1, 3, 4$. Step 5 is to answer the question. We wanted to know how to use the larger 20–cubic–yard trucks to save some money on our transportation costs. By using one of the large trucks on route 2 and the other two large trucks on route 5, we can save around $1,000 per day.

The first question for sensitivity analysis is how the potential cost savings depends on the number of large trucks available. The company may have other large trucks that can be reassigned, or they may be able to lease additional trucks. They may also have other projects in mind for their own trucks, so we will consider the effect of increasing or decreasing the number of large trucks available. Table 3.6 shows the results of our sensitivity runs. In each case we alter the constraint cell B8 and reoptimize.

We record the optimal decision (which routes to service with large trucks) as well as the expected cost savings. We also tabulate the marginal savings.

For example, the company can save \$1,200 by using 3 large trucks, or \$1,400 by using 4 large trucks, so the 4th truck saves the company \$200 per day. If the company can obtain the use of 1, 2, or 3 additional trucks for less than \$200 per day, it will save money. On the other hand, if the company has another activity where a large truck would save more than \$360 per day, then they would be better off using one less large truck for the transportation plan under consideration. The marginal savings are like shadow prices, since they tell us the potential impact of altering the constraint by one unit.

Next, we consider the impact of changes in the potential cost savings per route. Currently, we estimate that assigning 2 large trucks to route 5 will save the company \$640 per day. Another way to view this is that each truck on route 5 saves the company \$320 per day. Large trucks on routes 3 and 4 save \$200 per day, and a large truck on route 1 saves \$100 per day. It seems reasonable that small changes in the cost savings on route 5 will not alter our conclusion that assigning large trucks to routes 2 and 5 is the best plan. To verify this, we alter the value in cell C7 and reoptimize. Any value greater than 400 yields the same optimal decision. A value less than 400 leads us to assign large trucks to routes 2 and 3.

In the course of our robustness analysis, we will illustrate the way in which binary constraints can be used to restrict the possible decisions. Suppose management decides that large trucks should not be used on route 2, because the larger trucks will cause a public relations problem in that neighborhood. We could reformulate the problem to exclude this decision, but it is far easier to add the constraint $x_2 = 0$ to our current formulation. Now the optimal solution is to use large trucks on routes 4 and 5, with a total projected savings of \$840. This political concession costs the company \$160 per day. Now suppose that, for similar political reasons, management decides that if large trucks are used on route 4, then large trucks must be used on route 3 as well. This policy can be incorporated into our formulation by adding the linear constraint $x_3 \geq x_4$, or $x_3 - x_4 \geq 0$. If we add only this constraint, the optimal solution does not change, since the original optimal decision already satisfies this additional constraint (we don't use large trucks on either route). If we impose both constraints, then the optimal policy is to use large trucks on routes 3 and 4, with a total projected savings of \$600 per day.

3.5   Exercises

1. Reconsider the pig problem of Example 1.1, but now suppose that the price for pigs after t days is $p = 0.65e^{-(.01/.65)t}$ dollars/pound.

   (a) Show that the price for pigs is falling by one cent/day at time $t = 0$. What happens as t increases?

   (b) Find the optimal time to sell the pig. Use the five-step method, and model as a one–variable optimization problem.

(c) The parameter 0.01 represents the rate at which price is falling at time $t = 0$. Perform a sensitivity analysis on this parameter. Consider both the best time to sell and the resulting net profit.

(d) Compare the results of (b) to what was done in Section 1.1, and comment on the robustness of our model.

2. Reconsider the pig problem of Example 1.1, but now suppose that the weight of the pig after t days is $w = 800/(1 + 3e^{-t/30})$ lbs.

(a) Show that the pig is gaining about 5 lbs/day at $t = 0$. What happens as t increases?

(b) Find the optimal time to sell the pig. Use the five-step approach, and model as a one–variable optimization problem.

(c) The parameter 800 represents the eventual mature weight of the pig. Perform a sensitivity analysis for this parameter. Consider both the best time to sell and the profit obtained.

(d) Compare the results obtained in (b) to what was done in Sections 1.1 and 3.1. Comment on the robustness of our model. What general conclusions can we draw?

3. A statistical algorithm to determine the difference in effectiveness between two alternative treatments requires maximizing the quantity

$$\sum_{(k_1, k_2) \in E} \binom{n_1}{k_1} p_1^{k_1} (1 - p_1)^{n_1 - k_1} \binom{n_2}{k_2} p_2^{k_2} (1 - p_2)^{n_2 - k_2}$$

over the set

$$S = \{(p_1, p_2) : p_1 - p_2 = \Delta; \ p_1, p_2 \in [0, 1]\}.$$

Here E is a subset of the set

$$E_0 = \{(k_1, k_2) : k_1 = 0, 1, 2, \ldots, n_1; \ k_2 = 0, 1, 2, \ldots, n_2\}$$

and $\Delta \in [-1, 1]$. Find the maximum in the case $n_1 = n_2 = 4$, $\Delta = -0.1$,

$$E = \{(0, 4), (0, 3), (0, 2), (0, 1), (1, 4), (1, 3), (2, 4)\}$$

(Santner and Snell (1980)).

4. One method of evaluating the effectiveness of institutional trauma and burn medicine involves maximizing the function

$$f(p_1, \ldots, p_n) = \frac{\left(A - \sum_{i=1}^{n} p_i \right)}{\sqrt{B + \sum_{i=1}^{n} p_i (1 - p_i)}}$$

over the set

$$\{(p_i, \ldots, p_n) : a_i \leq p_i \leq b_i \text{ for all } i = 1, \ldots, n\}.$$

Maximize f in the case $n = 2$, $A = -5.92$, $B = 1.58$, $a_1 = 0.01$, $b_1 = 0.33$, $a_2 = 0.75$, and $b_2 = 0.85$ (Falk, J. et al., (1992)).

5. Reconsider the competing species model of Exercise 3 in Chapter 2. Assume that a level of effort E boat–days will result in the annual harvest of qEx blue whales and qEy fin whales, where the parameter q (catchability) is assumed to equal approximately 10^{-5}. Given a constant level of effort, assume that population levels will stabilize at the point where growth rate equals harvest rate.

   (a) Assuming that the cost of a whaling expedition is $250 per boat–day, find the level of effort that will maximize profit for the industry in the long run. Use the five-step method, and model as a one–variable optimization problem.

   (b) Examine the sensitivity to catchability q. Consider profit, level of effort, and the eventual stable population levels of the whales.

   (c) Increasing technology will certainly raise whale catchability. What will be the long–term effects on the whale populations and on the whaling industry?

6. Reconsider the facility location problem of Example 3.2, but now assume that the response time from point (x_0, y_0) to point (x_1, y_1) is proportional to the road travel distance $|x_1 - x_0| + |y_1 - y_0|$.

   (a) Find the location that minimizes average response time. Use the five-step method, and model as a multivariable unconstrained optimization problem.

   (b) Examine the sensitivity of the optimal location to the estimated number of emergencies in each 2×2–mile sector. Can you draw any general conclusions?

   (c) Comment on the robustness of this model. Compare the optimal location to that obtained in the analysis of Section 3.2. What do you think would happen if we assumed that response time was proportional to the straight–line distance $r = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2}$?

7. Reconsider the color TV problem of Example 2.1, but now use numerical methods instead of the analytic methods we employed in Chapter 2.

   (a) Determine the production levels x_1 and x_2 that maximize the objective function $y = f(x_1, x_2)$ in Eq. (2.2) of Chapter 2. Use the two–variable version of Newton's method.

(b) As in Section 2.1, let a denote the price elasticity for 19–inch sets. In part (a) we assumed $a = 0.01$. Now assume that a increases by 10% to $a = 0.011$ and repeat the optimization problem in part (a). Use your results to obtain a numerical estimate of the sensitivities $S(x_1, a)$, $S(x_2, a)$, and $S(y, a)$. Compare to the results obtained analytically in Section 2.1.

(c) Let b denote the price elasticity for 21–inch sets. Currently, $b = 0.01$. As in part (b), use numerical methods to estimate the sensitivities of x_1, x_2, and y to the parameter b.

(d) Compare the analytic methods of Section 2.1 to the numerical methods employed in this problem. Which do you prefer? Explain.

8. Reconsider Exercise 6 in Chapter 2, but now suppose that management has been persuaded to lift the ceiling on advertising expenditures. The assumption that sales vary as a linear function of the advertising budget is probably not reasonable over the wider range of advertising budget figures we now wish to consider. Suppose instead that sales are increased by 1,000 units each time the advertising budget is doubled.

(a) Find the price and the advertising budget that will maximize profit. Use the five-step method, and model as an unconstrained optimization problem.

(b) Determine the sensitivity of the decision variables (price and advertising budget) to price elasticity (the 50% number).

(c) Determine the sensitivity of the decision variables to the advertising agency's estimate of 1,000 new sales each time the advertising budget is doubled.

(d) What goes wrong in part (a) if we assume a linear relationship between advertising budget and sales? Why wasn't this an issue in Exercise 6 of Chapter 2?

9. (Continuation of Exercise 8) Repeat Exercise 8, but now assume an alternative model of the relationship between advertising expenditures and sales. Suppose that doubling the advertising budget results in 1,000 additional sales, but doubling it again results in only 500 additional sales, and so forth. Repeat parts (a) through (c) of Exercise 8. In part (c), determine the sensitivity to the assumption of 1,000 additional sales the first time the advertising budget is doubled. Compare your results to those obtained in Exercise 8, and comment on the robustness of the model.

10. Reconsider the newspaper problem of Exercise 7 in Chapter 2, but now suppose that we choose to maximize our profit margin (profit as a percentage of revenue). Assume that our business expenses remain fixed at $200,000 per week.

(a) Find the subscription price and the advertising price that maximize profit margin. Use the five-step method, and model as an unconstrained optimization problem. Find an approximate solution by the method of random search.

(b) Let $z = f(x, y)$ denote the objective function you obtained in part (a). Use a computer algebra system to determine $F = \partial f / \partial x$ and $G = \partial f / \partial y$. Then determine $\partial F / \partial x$, $\partial F / \partial y$, $\partial G / \partial x$, and $\partial G / \partial y$.

(c) Use Newton's method in two variables to obtain a precise answer to the question in part (a). Use the approximate solution from part (a) as your initial estimate. The required derivatives were calculated in part (b).

(d) If you have not previously solved part (a) of Exercise 7 in Chapter 2, then do so now. Use any method. Compare to the results of part (c). Does it matter whether we choose to maximize profit or profit margin? Explain.

11. Reconsider the lawn chair problem of Example 3.3. Notice that the objective function $f(x, y)$ tends to infinity as x or y approaches zero, and that $f(x, y)$ is undefined on the lines $x = 0$ and $y = 0$ that form the boundary of the feasible region. Presumably, the estimates of price elasticity are not accurate when extrapolated all the way to $x = 0$ or $y = 0$.

(a) Correct this model deficiency by altering the feasible region.

(b) Comment on the robustness of the decisions you made in part (a).

(c) Show that, for your corrected model, the optimal solution lies in the interior of the feasible region. Locate any local maximum of $f(x, y)$ on the boundary, and show that at every such point ∇f points into the interior.

12. Reconsider the newspaper problem of Exercise 9 in Chapter 2. Solve as a linear programming problem using a computer. Answer questions (a), (b), and (c) from the original problem.

13. Reconsider the cargo problem (Exercise 10 of Chapter 2). Solve as a linear programming problem using a computer. Answer questions (a), (b), and (c) from the original problem.

14. Reconsider the color TV problem of Example 2.2, but make the simplifying assumption that the company makes a profit of $80 per 19–inch set and $100 per 21–inch set.

(a) Find the optimal production levels. Use the five-step method, and solve as a linear programming problem using a computer.

(b) Determine the shadow prices for each constraint and explain what they mean. Which constraints are binding on the optimal solution?

		MI	NY	CA	GA	Output
	1	430	550	680	700	105
Textile Mill	2	510	590	890	685	160
	3	395	425	910	450	85
Allocation		70	100	105	75	

Table 3.7: Data for the transportation problem of Exercise 15. Shipping cost per truckload in dollars.

(c) Determine the sensitivity to the objective function coefficients (profit per set). Consider both profit and optimal production levels.

(d) Draw a graph of the feasible region (see Fig. 2.10) and include a picture of ∇f at the optimum. Describe geometrically what happens to the vector ∇f as we change one of the objective function coefficients. Use this geometric idea to determine how much each objective function coefficient can change before the current optimal solution is no longer optimal.

15. The Burningham textiles company has three textile mills located in the southern U.S. and four distribution centers located in Michigan, New York, California, and Georgia. The estimated yearly output from each mill, the allocation to each warehouse, and the shipping costs are tabulated in Table 3.7.

(a) Find the transportation plan that minimizes shipping costs. Use the five-step method, and solve as a linear programming problem using a computer.

(b) Determine the shadow prices for each of the output constraints. Would it be beneficial to shift production capacity from one mill to another? How much should the company be willing to spend to facilitate the shift?

16. A manufacturer of personal computers sells three desktop models. Model A costs $850 to manufacture and sells for $1,250; model B costs $950 to manufacture and sells for $1,400; and model C costs $1,500 to manufacture and sells for $2,500. The company purchases 10,000 desktop cases each month, and each computer requires one case. Models A and B use a 15–inch monitor, and the company can obtain 5,000 of these monitors per month. Model C uses a 17–inch monitor, and 7,500 of these are available per month. The remaining components are freely available. The company has 20,000 hours of production capacity available per month, and the production of each unit of model A, B, and C requires 1, 1.25, and 1.75 hours respectively.

(a) How many of each type of computer should the company produce? Use the five-step method and solve as a linear programming problem.

(b) Determine the shadow prices for each constraint. Explain what each shadow price means in the context of this problem.

(c) Next month, the company plans to put model C computers on sale for $2,199. How does this change the results of parts (a) and (b)?

(d) The company is considering a new desktop model D. This model will cost $1,250 to manufacture and will sell for $1,895. It requires a desktop case, a 17–inch monitor, and 1.5 hours of production. How will this change the results of parts (a) and (b)? Would you advise the company to go ahead with their plans to introduce this new model?

17. A retired engineer has $250,000 to invest and is willing to spend five hours per week managing her investments. Municipal bonds earn 6% per year and require no management. Real estate investments are expected to appreciate at 8% per year and require one hour of management per $100,000 invested. Blue chip stocks earn 10% per year and require 1.5 hours of management. Junk bonds earn 12% and require 2.5 hours, while grain futures earn 15% and require five hours per $100,000 invested.

(a) How should the retiree invest her money in order to maximize her expected earnings? Use the five-step method and solve as a linear programming problem.

(b) Determine the shadow prices for each constraint. Interpret each shadow price in the context of this problem.

(c) The retiree downloads software from the internet that allows her to effectively manage her grain futures in three hours per week per $100,000. How does this change the results in parts (a) and (b)?

(d) After a few disasters in the futures market, the engineer decides that risk is a significant factor in her investment strategy. An investment self–help book ranks municipal bonds, real estate, blue chip stocks, junk bonds, and grain futures as risk level 1, 4, 3, 6, and 10 respectively. The engineer decides that her investment portfolio should have an average risk level of no more than 4. How does this change the results in parts (a) and (b)?

18. The Green Supply Company manufactures plastic grocery bags and milk jugs. The company can obtain 5,000 lbs of used plastic bags, 18,000 lbs of used plastic milk jugs, and 40,000 of industrial plastic scraps per week at a cost of $18, $12, and $10 per 100 lbs respectively. The company has orders for 4,000 boxes of plastic bags and 80,000 milk jugs per week. One box of plastic bags requires 6 lbs of plastic, costs $5 to manufacture, and sells for $14. It costs $9 and requires 14 lbs of plastic to make 100 milk jugs, and the jugs retail for $20. The plastic bags must be at least 25% post–consumer recycled plastic (used milk jugs or bags) because of consumer preference, and the milk jugs must be at most 50% post–consumer recycled plastic for strength.

(a) Determine the optimal mix of plastics for each product. Use the five-step method and model as a linear programming problem.

(b) Determine the shadow prices for the each constraint, and interpret each shadow price in the context of this problem.

(c) A new supplier can provide industrial plastic scraps at $8 per 100 lbs. How does this change the results in parts (a) and (b)?

(d) A new customer has offered to buy 40,000 environmentally friendly milk jugs per week for $30 per 100 jugs. The jugs must contain at least 35% post–consumer recycled plastic. How does this change the results in parts (a) and (b)? Should the company accept this new customer?

19. Reconsider the dirt problem of Example 3.5, and assume that the company will only ship full truckloads of dirt.

(a) Suppose that the company uses dump trucks with a capacity of ten cubic yards of dirt, and find the optimal shipping plan. Use the five-step method and solve as an integer programming problem.

(b) Repeat part (a) assuming that the trucks have a capacity of five cubic yards.

(c) Repeat part (a) assuming that the trucks have a capacity of 20 cubic yards.

(d) Compare the results of parts (a), (b), and (c) to determine the robustness of the original linear programming model. Is the original shipping plan in Example 3.5 nearly optimal for any size truck?

20. (Variation of Exercise 14) Reconsider the color TV problem of Example 2.2, but make the simplifying assumption that the company makes a profit of $80 per 19–inch set and $100 per 21–inch set. In this problem we will explore the effect of discretization on the optimal solution.

(a) Find the optimal production levels. Use the five-step method, and solve as a linear programming problem using a computer.

(b) Because of mass production, the TV sets are actually made in batches of 30. Find the optimal number of batches to maximize profit. Solve as an integer programming problem using a computer.

(c) Repeat part (b) assuming batch sizes of 10, 20, 50, 100, 200, and 300. For each case, use integer programming to compute the optimal number of batches for each type of TV set.

(d) Compare the results in parts (a), (b), and (c) and comment on the robustness of the original linear programming solution found in part (a). How does the discretization of the feasible region affect the optimal solution? Consider both the optimal levels of production and the optimal profit.

21. (Variation of Exercise 15) The Burningham textiles company has three textile mills located in the southern U.S. and four distribution centers located in Michigan, New York, California, and Georgia. The estimated yearly output from each mill, the allocation to each warehouse, and the shipping costs are tabulated in Table 3.7.

   (a) Find the transportation plan that minimizes shipping costs. Use the five-step method, and solve as a linear programming problem using a computer.

   (b) The company has purchased three new trucks for shipping. The new trucks are more fuel–efficient, and are expected to reduce shipping costs by 50%. Assume that one truck can make one trip per week (52 per year), and that each route will use either all old or all new trucks. How should the company use the new trucks in order to maximize their cost savings? Use the five-step method, and solve as an integer programming problem using a computer.

   (c) Perform a sensitivity analysis on the number of new trucks. Determine the optimal shipping plan and the expected cost savings for $n = 4, 5, 6$, and 7 new trucks. If the cost of a new truck is $12,000 per year amortized over the expected life of the vehicle, how many additional trucks should the company purchase?

   (d) Because the State of California has instituted new pollution laws for trucks, the company must use the new trucks to ship to CA. How does this change the results of parts (b) and (c)? What do the new CA pollution laws cost the company?

22. Reconsider the textile shipping problem of Exercise 21, but now assume that both old and new trucks can be used on the same route. Answer the same questions as before. [Hint: Use one decision variable to indicate whether or not one new truck is used on route i, another decision variable to indicate whether a second truck is used, and so forth.]

23. A computer operating system stores files on a hard disk. Five large files of sizes 18, 23, 12, 125, and 45 MB are to be stored. Contiguous blocks of storage are available with size 25, 73, 38, and 156 MB, and each file must be stored in one contiguous block. In this problem we will explore an integer programming algorithm to assign files to storage blocks.

   (a) In order to reserve large contiguous blocks of storage for future use, we want to store each file in the smallest available block large enough to hold the file. Define the cost of storing file i in block j to be the size of block j, and determine the assignment of files to blocks that minimizes the total cost. Use the five-step method, and model as an integer programming problem.

   (b) Suppose that the 12 MB file expands to 19 MB. How does this effect the optimal solution found in part (a)? How much can this 12 MB file expand before the optimal solution changes?

(c) Suppose that the 18 MB file and the 23 MB file are to be stored in the same block, since they are used by the same program. How does this affect the optimal solution found in part (a)?

(d) One "greedy" algorithm for allocating blocks to files is to place each file in the first available block that is large enough to hold it. Apply this algorithm (by hand) and compare to the results of part (a). Is the IP solution found in part (a) significantly better than the results of the greedy algorithm?

(e) Why not just maximize the size of the largest remaining contiguous block of storage? Can this optimization problem be solved as an IP?

24. A technical manager is scheduling engineers to work on several upcoming projects. Projects A, B, and C will require 18, 12, and 30 person–months to complete. Engineers 1, 2, 3, and 4 are available to work on these projects. Their monthly salaries are \$3000, \$3500, \$3200, and \$3900 respectively.

(a) Determine the schedule (allocation of engineers to projects) that minimizes the total cost of finishing all the projects. Assume that engineers are assigned to only one project every six months, and that all projects are to be completed within 18 months. [Hint: Let the decision variable x_{ijk} indicate whether engineer i works on project j during period k.]

(b) Repeat part (a) assuming that engineer 1 is unavailable during period 2 because of a prior assignment. How does this affect the optimal solution? How much would it be worth to the technical manager to have engineer 1 reassigned to his team during period 2?

(c) Repeat part (a) assuming that engineers 2 and 3 cannot work together because of a personality conflict. How much does their personal animosity cost the company?

(d) The company is offered a \$10,000 performance bonus if they can complete project A within 6 months. How does this change the optimal schedule?

Further Reading

1. Beltrami, E. (1977) *Models for Public Systems Analysis*. Academic Press, New York.

2. Dantzig, G. (1963) *Linear Programming and Extensions*. Princeton University Press, Princeton, New Jersey.

3. Falk, J., Palocsay, S., Sacco, W., Copes, W. and Champion, H. (1992) Bounds on the Trauma Outcome Function via Optimization. *Operations Research* 40, Supp. No. 1, S86–S95.

4. Gearhart, W. and Pierce, J. *Fire Control and Land Management in the Chaparral.* UMAP module 687.

5. Hillier, F. and Lieberman, J. (1990) *Introduction to Operations Research.* McGraw–Hill, New York.

6. Maynard, J. *A Linear Programming Model for Scheduling Prison Guards.* UMAP module 272.

7. Press, W., Flannery, B., Teukolsky, S. and Vetterling, W. (2002) *Numerical Recipes in C++: The Art of Scientific Computing.* 2nd Ed., Cambridge University Press, New York. See also `www.numerical-recipes.com`

8. Polack, E. (1971) *Computational Methods in Optimization.* Academic Press, New York.

9. Santner, T. and Snell, M. (1980) Small–sample confidence intervals for $\rho_1 - \rho_2$ and ρ_1/ρ_2 in 2×2 contingency tables. *Journal of the American Statistical Association* 75, 386–394.

10. Straffin, P. *Newton's Method and Fractal Patterns.* UMAP module 716.

Chapter 4

INTRODUCTION TO DYNAMIC MODELS

Many problems of practical interest involve processes that evolve over time. Dynamic models are used to represent the changing behavior of these systems. Space flight, electrical circuits, chemical reactions, population growth, investments and annuities, military battles, the spread of disease, and pollution control are just a few of the many areas in which extensive use is made of dynamic models.

The five-step method and the fundamental principles of sensitivity analysis and robustness are as relevant and useful for dynamic models as they are for optimization models. We will continue to rely on them as we explore some of the most popular and generally applicable dynamic modeling techniques. In the course of this study, we will also introduce the important modeling concepts of state space, equilibrium, and stability. All of this will also be very useful in the last part of this book, where we explore stochastic models.

As a general rule, dynamic models are easy to formulate and hard to solve. Exact analytic solutions are available only for a few special cases, such as linear systems. Numerical methods usually do not provide a good qualitative understanding of system behavior. Therefore, the application of graphical techniques is usually employed as at least one part of the analysis of dynamic models. Because of the inherent simplicity of graphical techniques, along with their geometrical nature, this chapter also provides us with an ideal opportunity to introduce some of the deepest and most fundamental modeling concepts used for dynamic systems.

4.1   Steady State Analysis

In this section we will consider the simplest type of dynamic model. The mathematics required are elementary indeed. Even so, the practical applications for this model are numerous, and the absence of too much sophisticated technique

leaves us free to concentrate on some of the most fundamental ideas of dynamic modeling.

Example 4.1. In an unmanaged tract of forest area, hardwood and softwood trees compete for the available land and water. The more desirable hardwood trees grow more slowly, but are more durable and produce more valuable timber. Softwood trees compete with the hardwoods by growing rapidly and consuming the available water and soil nutrients. Hardwoods compete by growing taller than the softwoods can and shading new seedlings. They are also more resistant to disease. Can these two types of trees coexist on one tract of forest land indefinitely, or will one type of tree drive the other to extinction?

We will use the five-step method. Let H and S denote the populations of hardwood and softwood trees, respectively. A convenient unit often used by biologists is the biomass (tons per acre of living tree). We need to make some assumptions about the dynamics of these two populations. To begin with, we want to make assumptions that are as simple as possible without neglecting the most fundamental aspects of the problem. Later on, we can improve or enrich our model if necessary. It is reasonable to assume that in conditions of unrestricted growth (plenty of room, sunshine, water, and soil nutrients), the growth rate of a species is roughly proportional to the size of the species. Twice as many trees give rise to twice as many little trees. As population increases, members of the same species must compete for resources, and this inhibits growth. Thus, it is reasonable to assume that growth rate is roughly linear in population size for small populations and then falls off as population increases. The simplest growth rate function with these properties is

$$g(P) = rP - aP^2.$$

Here r is the intrinsic growth rate, and $a << r$ is a measure of the strength of resource limitations. If a is smaller, there is more room to grow.

The effect of competition is also due to resource limitations. The presence of hardwood trees limits the amount of sunlight, water, etc., available for the softwoods, and vice versa. The loss in growth rate due to competition depends on the size of both populations. A simple assumption is that this loss is proportional to the product of the two. Given these assumptions about growth and competition, we wish to know whether we can expect one species to die out over time. Figure 4.1 summarizes the results of step 1.

Step 2 is to select the modeling approach. We will model this problem as a dynamic model in steady state.

We are given functions

$$f_1(x_1, \ldots, x_n)$$

$$\vdots$$

$$f_n(x_1, \ldots, x_n)$$

Variables: H = hardwood population (tons/acre)
                   S = softwood population (tons/acre)
                   g_H = growth rate for hardwoods (tons/acre/year)
                   g_S = growth rate for softwoods (tons/acre/year)
                   c_H = loss due to competition for hardwoods (tons/acre/year)
                   c_S = loss due to competition for softwoods (tons/acre/year)

Assumptions: $g_H = r_1 H - a_1 H^2$
                   $g_S = r_2 S - a_2 S^2$
                   $c_H = b_1 SH$
                   $c_S = b_2 SH$
                   $H \geq 0, S \geq 0$
                   $r_1, r_2, a_1, a_2, b_1, b_2$ are positive reals

Objective: Determine whether $H \to 0$ or $S \to 0$

Figure 4.1: Results of step 1 for the tree problem.

defined on a subset S of $\mathbb{R}^n$. The functions $f_1, \ldots, f_n$ represent
the rate of change of each variable $x_1, \ldots, x_n$ respectively. A point
$(x_1, \ldots, x_n)$ in the set S is called an *equilibrium point* provided that

$$f_1(x_1, \ldots, x_n) = 0$$
$$\vdots \qquad\qquad\qquad (4.1)$$
$$f_n(x_1, \ldots, x_n) = 0$$

at this point. The rate of change of each of the variables $x_1, \ldots, x_n$
is then equal to zero, and so the system is at rest.

The variables $x_1, \ldots, x_n$ are called *state variables*, and S is called
the *state space*. Since the functions $f_1, \ldots, f_n$ depend only on the
current state $(x_1, \ldots, x_n)$ of the system, knowledge of the current
state suffices to determine the entire future of the system. What
happened in the past does not matter. We only need to know where
we are now, not how we got here. When we are at an equilibrium
point, defined by Eq. (4.1), we say that the system is in *steady state*.
At this point all of the rates of change are equal to zero. All of
the forces acting on the system are in balance. For this reason the
equations in (4.1) are sometimes referred to as the *balance equations*.
When a dynamic system is in steady state, it remains there forever.
Since all of the rates of change are equal to zero, any future time
will find us in exactly the same place we are right now.

In order to find the equilibrium states of a dynamic system, we
need to solve the n equations in n unknowns given by Eq. (4.1).
In very easy cases we can solve by hand. Sometimes we can solve
using a computer algebra system. All of the problems in this chap-

ter, including the exercises, can be solved using these techniques. Of course, many real problems give rise to systems of equations that cannot be solved analytically. We will treat such problems in Chapter 6, where we discuss computational methods for dynamical systems. (Alternatively, we could use the multivariable version of Newton's method introduced in Chapter 3.)

Step 3 is to formulate the model. Let $x_1 = H$ and $x_2 = S$ denote our two state variables, defined on the state space

$$\{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}.$$

The steady–state equations are

$$r_1 x_1 - a_1 x_1^2 - b_1 x_1 x_2 = 0$$
$$r_2 x_2 - a_2 x_2^2 - b_2 x_1 x_2 = 0. \tag{4.2}$$

We are interested in solutions of this system of equations that lie in the state space. These solutions represent the equilibrium points of our dynamic model.

Step 4 is to solve the model. Factoring out x_1 from the first equation and x_2 from the second, we find four solutions, three at the following coordinates:

$$(0, 0)$$
$$(0, r_2/a_2)$$
$$(r_1/a_1, 0)$$

and the fourth at the intersection of these two lines:

$$a_1 x_1 + b_1 x_2 = r_1$$
$$b_2 x_1 + a_2 x_2 = r_2.$$

See Fig. 4.2 for an illustration.

Solving by Cramer's rule yields

$$x_1 = \frac{r_1 a_2 - r_2 b_1}{a_1 a_2 - b_1 b_2}$$

$$x_2 = \frac{a_1 r_2 - b_2 r_1}{a_1 a_2 - b_1 b_2}.$$

If the two lines do not cross inside the state space, then there are only three equilibria. In this case the two species of trees cannot coexist in peaceful equilibrium.

We are interested to know the conditions under which $x_1 > 0$ and $x_2 > 0$. It is reasonable to assume that $a_i > b_i$, since the effect of competition between members of the same species should be stronger than the competition between species. The growth rate is

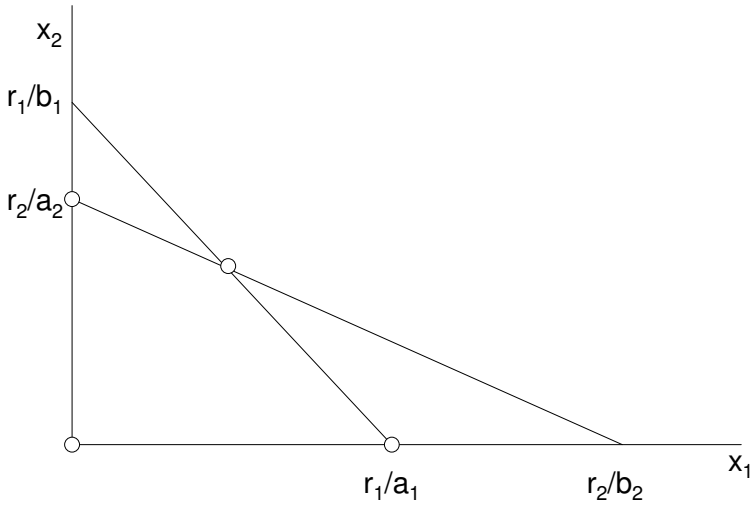$$r_i x_i - a_i x_i x_i - b_i x_i x_j,$$

Figure 4.2: Graph of softwoods x_2 versus hardwoods x_1 showing equilibria for the tree problem.

where the first term represents unrestricted growth, the second represents the effect of competition within a population, and the third represents competition between populations. Since the two types of trees do not occupy exactly the same ecological niche, we would suppose that for $x_i = x_j$ the effect of competition within a population would be the stronger. Hence $a_i > b_i$, so

$$a_1 a_2 - b_1 b_2 > 0.$$

The condition for coexistence is, therefore, that

$$r_1 a_2 - r_2 b_1 > 0$$
$$a_1 r_2 - b_2 r_1 > 0,$$

or, in other words,

$$\frac{r_2}{a_2} < \frac{r_1}{b_1} \text{ and } \frac{r_1}{a_1} < \frac{r_2}{b_2},$$

as shown in Fig. 4.2.

Step 5 is to report in plain English the results obtained from our model analysis. It is difficult to do so in this case, because our answer is qualified, and the qualifications involve unknown parameters. In order to communicate our results clearly, we would like to find a more tangible interpretation of our conditions for coexistence. Let us reexamine our model formulation to see if we can interpret the meaning of the ratios r_i/a_i and r_i/b_i in some straightforward way.

The parameters r_i measure growth tendency, and the parameters a_i and b_i measure the strength of competition within and between populations, respectively. Thus, the ratios r_i/a_i and r_i/b_i must measure the relative strength of growth versus competition. Let us try to go further. In the absence of competition between species, the growth rate is

$$r_i x_i - a_i x_i^2 = x_i(r_i - a_i x).$$

The ratio r_i/a_i represents the equilibrium population level in the absence of competition between species, or the level at which the population will stop growing of its own accord. Similarly, if we neglect the factor of competition within a population, the net growth rate is

$$r_i x_i - b_i x_i x_j = x_i(r_i - b_i x_j).$$

The ratio r_i/b_i thus represents the level of population j necessary to put an end to growth of population i. In light of this, we can now give our analysis results the following concrete interpretation.

For each type of tree (hardwood and softwood), there are two kinds of limits to growth. The first comes from competition with the other type of tree, and the second comes from competition between trees of the same type under crowded conditions. Thus, for each type of tree there is one point where growth will halt itself due to crowding, and another point where the growth of one type of tree will halt the growth of the other type due to competition. The condition for coexistence of both types is that each type reaches the point where it limits its own growth before it reaches the point where it limits the other's growth.

The steady–state analysis of this section leaves one important question unanswered. Given that a dynamic model has an equilibrium solution, will we ever get there? The answer depends on the dynamics of the model. An equilibrium point

$$x_0 = (x_1^0, \ldots, x_n^0)$$

is said to be *asymptotically stable* (or just *stable*) if whenever the state variables

$$(x_1(t), \ldots, x_n(t))$$

pass sufficiently close to x_0, they are drawn into the equilibrium. In other words,

$$(x_1(t), \ldots, x_n(t)) \to x_0.$$

Steady–state analysis cannot answer the question of stability, so we will have to defer further discussion of this topic to the next section.

4.2   Dynamical Systems

Dynamical system models are the most commonly used type of dynamic model. In a dynamical system model the forces of change are represented by differential equations. In this section we will focus on the graphical method for obtaining qualitative information about a dynamical system. The emphasis will be on questions of stability.

Variables: B = number of blue whales
                   F = number of fin whales
                   g_B = growth rate of blue whale population (per year)
                   g_F = growth rate of fin whale population (per year)
                   c_B = effect of competition on blue whales (whales per year)
                   c_F = effect of competition on fin whales (whales per year)

Assumptions: $g_B = 0.05B(1 - B/150,000)$
                   $g_F = 0.08F(1 - F/400,000)$
                   $c_B = c_F = \alpha BF$
                   $B \geq 0, F \geq 0$
                   α is a positive real constant

Objective: Determine whether dynamic system can reach stable equilibrium
                   starting from $B = 5,000$, $F = 70,000$

Figure 4.3: Results of step 1 for the whale problem.

Example 4.2. The blue whale and fin whale are two similar species that inhabit the same areas. Hence, they are thought to compete. The intrinsic growth rate of each species is estimated at 5% per year for the blue whale and 8% per year for the fin whale. The environmental carrying capacity (the maximum number of whales that the environment can support) is estimated at 150,000 blues and 400,000 fins. The extent to which the whales compete is unknown. In the last 100 years intense harvesting has reduced the whale population to around 5,000 blues and 70,000 fins. Will the blue whale become extinct?

We will use the five-step method. Notice that this problem is very similar to Example 4.1. Step 1 is to ask a question. We will use the number of blue and fin whales as state variables and make the simplest possible assumptions about growth and competition. The question we begin with is this: Can the two populations of whales grow to stable equilibrium starting from their current levels? The results of step 1 are summarized in Figure 4.3.

Step 2 is to select the modeling approach. We will model this problem as a dynamical system.

A *dynamical system* consists of n state variables $(x_1, \ldots, x_n)$ and a system of differential equations

$$\frac{dx_1}{dt} = f_1(x_1, \ldots, x_n)$$

$$\vdots \qquad \qquad \vdots \qquad\qquad\qquad (4.3)$$

$$\frac{dx_n}{dt} = f_n(x_1, \ldots, x_n)$$

defined on the state space $(x_1, \ldots, x_n) \in S$, where S is a subset of $\mathbb{R}^n$. The existence and uniqueness theorem of differential equations

states that if $f_1, \ldots, f_n$ have continuous first partial derivatives in a neighborhood of a point

$$x_0 = (x_1^0, \ldots, x_n^0),$$

then there exists a unique solution to this system of differential equations through this initial condition. See any introductory text on differential equations for details (e.g., Hirsch, et al. (1974) p. 162). Many other differential equation models can be reduced to the form (4.3). If the dynamics depend on time, we can introduce time as another state variable. If second derivatives are involved, we can include first derivatives as state variables, and so forth.

It is best to think of a solution to a dynamical system as a path through the state space. As long as differentiability assumptions are satisfied, there is a path through each point, and paths cannot cross except at an equilibrium. Let

$$x = (x_1, \ldots, x_n)$$
$$F(x) = (f_1(x), \ldots, f_n(x)).$$

Then the dynamical system equation is

$$\frac{dx}{dt} = F(x). \tag{4.4}$$

For a path $x(t)$, the derivative dx/dt represents the velocity vector. Hence, for every solution curve $x(t)$, we have that $F(x(t))$ is the velocity vector at each point. The vector field $F(x)$ tells us in what direction and how fast we are moving through the state space. Usually, a good idea of the qualitative behavior of a dynamical system in two variables can be obtained by drawing the vector field at selected points. The points where $F(x) = 0$ are the equilibria, and we will pay special attention to the vector field nearby these points.

Step 3 is to formulate the model. Let $x_1 = B$ and $x_2 = F$, and write

$$x_1' = f_1(x_1, x_2)$$
$$x_2' = f_2(x_1, x_2),$$

where

$$f_1(x_1, x_2) = 0.05x_1 \left(1 - \frac{x_1}{150,000}\right) - \alpha x_1 x_2$$
$$f_2(x_1, x_2) = 0.08x_2 \left(1 - \frac{x_2}{400,000}\right) - \alpha x_1 x_2. \tag{4.5}$$

The state space is

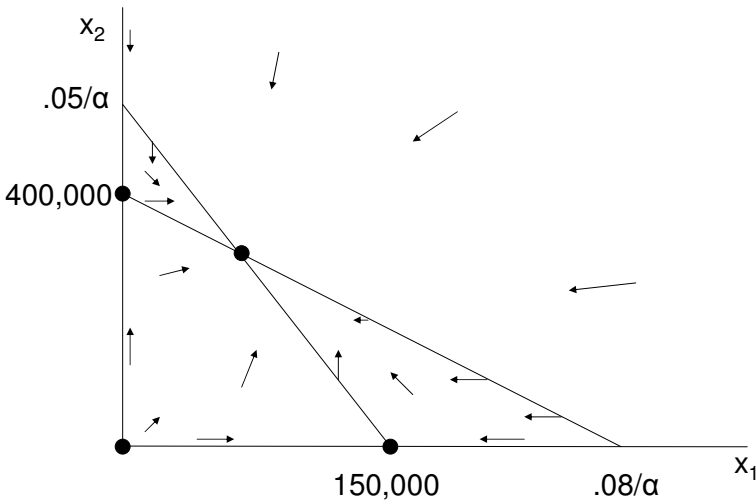$$S = \{(x_1, x_2) : x_1 \geq 0, x_2 \geq 0\}.$$

Figure 4.4: Graph of fin whales x_2 versus blue whales x_1 showing the vector field (4.5) for the whale problem.

Step 4 is to solve the model. We want to sketch a graph of the vector field for this problem. Start out by sketching the level sets $f_1 = 0$ and $f_2 = 0$. The equilibria will be at the intersection of these two. Furthermore, the velocity vectors will be vertical along $f_1 = 0$ ($x_1' = 0$) and horizontal along $f_2 = 0$ ($x_2' = 0$). Draw the velocity vectors along these two curves. Then fill in some of the velocity vectors in between. It helps to remember that (as long as $F(x)$ is continuous, which it usually is) both the length and direction of the vectors change continuously. In fact, for this kind of analysis, the length of the velocity vectors is not very important. See Figure 4.4 for the finished graph.

There are four equilibrium solutions, three at

$$(0, 0)$$
$$(150,000, 0) \qquad\qquad (4.6)$$
$$(0, 400,000)$$

and another at a point whose coordinates depend on α. Our graph assumes that

$$400,000 < (0.05/\alpha).$$

In this case it is easy to see that the equilibrium in the interior is the only stable one. In fact, any solution through a point in the interior of the state space will eventually converge upon this equilibrium. In particular, the solution with initial conditions $x_1(0) = 5,000$ and $x_2(0) = 70,000$ tends to this equilibrium as $t \to \infty$.

Step 5 is to summarize the results of our model analysis in nonmathematical terms. Based on our analysis, in the absence of further harvesting, the whale populations will grow back to their natural levels, and the ecological system will remain in stable equilibrium.

Of course, our conclusions are based on some rather broad assumptions. For example, we have assumed that the effect of competition is relatively small. If it were larger (i.e., if $(0.05/\alpha) < 400,000$), then the two species could not coexist. It does seem reasonable to make the assumption that α is small, since we know that the two species have coexisted for a long time before we began to harvest them. We have also made several simplifying assumptions about the growth process. The most critical of these is that for very small population levels, the population still tends to increase at the intrinsic growth rate. It is believed that some species have a minimum size (called the *minimum viable population level*) below which the growth rate is negative, ensuring the eventual extinction of the species. This assumption would, of course, change the behavior of our dynamical system. See Exercise 5 at the end of this chapter.

Finally, we address the questions of sensitivity analysis and robustness. First let us consider the sensitivity to the parameter α, for which we have very little information. For any value of $\alpha < 1.25 \times 10^{-7}$ there is a stable equilibrium $x_1 > 0,\ x_2 > 0$ at

$$x_1 = \frac{150,000(8,000,000\alpha - 1)}{D}$$
$$x_2 = \frac{400,000(1,875,000\alpha - 1)}{D}, \tag{4.7}$$

where

$$D = 15,000,000,000,000\alpha^2 - 1,$$

which we found by Cramer's Rule. For example, if $\alpha = 10^{-7}$, then

$$x_1 = \frac{600,000}{17} \approx 35,294$$
$$x_2 = \frac{6,500,000}{17} \approx 382,353. \tag{4.8}$$

The sensitivities at this point are

$$S(x_1, \alpha) = -\frac{21,882,352,927}{6,000,000,000} \approx -3.6$$

and

$$S(x_2, \alpha) = \frac{27}{221} \approx 0.122.$$

The above calculations could be performed either by hand or by using a computer algebra system. Figure 4.5 illustrates the computation of the sensitivity $S(x_1, \alpha)$ using the computer algebra system Maple.

```
> e1:=(5/100)*(1-x1/150000)-alpha*x2;
```
$$e1 := \frac{1}{20} - \frac{1}{3000000} x1 - \alpha \, x2$$

```
> e2:=(8/100)*(1-x2/400000)-alpha*x1;
```
$$e2 := \frac{2}{25} - \frac{1}{5000000} x2 - \alpha \, x1$$

```
> s:=solve({e1=0,e2=0},{x1,x2});
```
$$s := \left\{ x2 = \frac{400000\,(-1 + 1875000\,\alpha)}{-1 + 15000000000000\,\alpha^2}, \; x1 = \frac{150000\,(-1 + 8000000\,\alpha)}{-1 + 15000000000000\,\alpha^2} \right\}$$

```
> assign(s);
```
```
> dx1dalpha:=diff(x1,alpha);
```
$$dx1dalpha := \frac{1200000000000}{-1 + 15000000000000\,\alpha^2} - \frac{4500000000000000000\,(-1 + 8000000\,\alpha)\,\alpha}{(-1 + 15000000000000\,\alpha^2)^2}$$

```
> assign(alpha=10^(-7));
```
```
> sx1alpha:=dx1dalpha*(alpha/x1);
```
$$sx1alpha := \frac{-62}{17}$$

```
> evalf(sx1alpha);
```
$$-3.647058824$$

Figure 4.5: Calculation of the sensitivity $S(x_1, \alpha)$ for the whale problem using the computer algebra system Maple.

The blue whale population is most sensitive to α. If $\alpha = 10^{-8}$, then

$$x_1 = \frac{276,000,000}{1,997} \approx 138,207$$
$$x_2 = \frac{785,000,000}{1,997} \approx 393,090.$$

Of course we will always have $x_1 < 150,000$ and $x_2 < 400,000$, as is apparent from the graph in Fig. 4.4. But the most important features of this equilibrium are not its coordinates, but rather the fact that an equilibrium exists on $x_1 > 0$, $x_2 > 0$ and is stable. These conclusions remain valid over the entire range of $\alpha < 1.25 \times 10^{-7}$, which we believe to be plausible. Hence, we should say that our main conclusion is not at all sensitive to α. Similarly, our main conclusion is not at all sensitive to our data on intrinsic growth rate and carrying capacity, or even to the current whale populations.

Deeper questions of robustness revolve around the assumed form of the functions f_1 and f_2. We assumed that x_1'/x_1 and x_2'/x_2 are linear functions of x_1 and x_2, respectively. These lines represent the point where one species or the other stops growing. Suppose that we relax this linearity assumption. Let

$$x_1' = x_1 \, g_1(x_1, x_2)$$
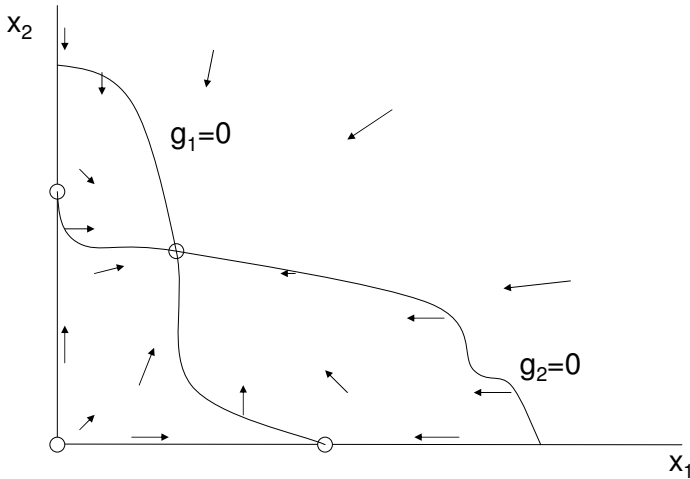$$x_2' = x_2 \, g_2(x_1, x_2).$$

Figure 4.6: Graph of fin whales x_2 versus blue whales x_1 showing vector field for the generalized whale problem.

All of our analysis results would certainly remain true if g_1 and g_2 were not linear, so long as the vector field had the same general features. See Figure 4.6 for an illustration.

4.3   Discrete Time Dynamical Systems

In some problems it is most natural to model the time variable as being discrete. When this happens, the usual differential equations are replaced by their discrete–time analog: difference equations. The relationship between discrete and continuous dynamics is the relationship between $\Delta x/\Delta t$ and dx/dt, so it is often assumed that the behavior of a dynamical system will be roughly the same whether we assume that time is continuous or discrete. However, this kind of logic overlooks one important point. There is a kind of time delay built into every discrete–time dynamical system, which is the length of the time step Δt. For systems in which the dynamic forces are very strong, this time delay can lead to unexpected results.

Example 4.3. Astronauts in training are required to practice a docking maneuver under manual control. As a part of this maneuver, it is required to bring an orbiting spacecraft to rest relative to another orbiting craft. The hand controls provide for variable acceleration and deceleration, and there is a device on board that measures the rate of closing between the two vehicles. The following

strategy has been proposed for bringing the craft to rest. First, look at the closing velocity. If it is zero, we are done. Otherwise, remember the closing velocity and look at the acceleration control. Move the acceleration control so that it is opposite to the closing velocity (i.e., if closing velocity is positive, we slow down, and we speed up if it is negative) and proportional in magnitude (i.e., we brake twice as hard if we find ourselves closing twice as fast). After a time, look at the closing velocity again and repeat the procedure. Under what circumstances will this strategy be effective?

We will use the five-step method. Let v_n denote the closing velocity observed at time t_n, the time of the nth observation. Let

$$\Delta v_n = v_{n+1} - v_n$$

denote the change in closing velocity as a result of our adjustments. We will denote the time between observations of the velocity indicator by

$$\Delta t_n = t_{n+1} - t_n.$$

This time interval naturally divides into two parts: the time it takes to adjust the velocity controls and the time between adjustment and the next observation of the velocity indicator. Write

$$\Delta t_n = c_n + w_n,$$

where c_n is the time to adjust the controls and w_n is the waiting time until the next observation. The parameter c_n is a function of astronaut response time, and we are free to choose w_n.

Let a_n denote the acceleration setting after the nth adjustment. Elementary physics yields

$$\Delta v_n = a_{n-1}c_n + a_n w_n.$$

The control law says to set acceleration proportional to $(-v_n)$; hence

$$a_n = -kv_n.$$

The results of step 1 are summarized in Figure 4.7.

Step 2 is to select the modeling approach. We will model this problem as a discrete–time dynamical system.

A discrete–time dynamical system consists of a number of state variables $(x_1, \ldots, x_n)$ defined on the state space $S \subseteq \mathbb{R}^n$ and a system of difference equations

$$\Delta x_1 = f_1(x_1, \ldots, x_n)$$

$$\vdots \qquad \vdots \qquad\qquad (4.9)$$

$$\Delta x_n = f_n(x_1, \ldots, x_n).$$

Variables: t_n = time of nth velocity observation (sec)
                     v_n = velocity at time t_n (m/sec)
                     c_n = time to make nth control adjustment (sec)
                     a_n = acceleration after nth adjustment (m/sec^2)
                     w_n = wait before $(n+1)$th observation (sec)

Assumptions: $t_{n+1} = t_n + c_n + w_n$
                     $v_{n+1} = v_n + a_{n-1}c_n + a_n w_n$
                     $a_n = -kv_n$
                     $c_n > 0$
                     $w_n \geq 0$

Objective: Determine whether $v_n \to 0$

Figure 4.7: Results of step 1 of the docking problem.

Here Δx_n represents the change in x_n over one time step. It is common to take time steps of length 1, which just amounts to selecting appropriate units. If time steps are of variable length, or if the dynamics of the system vary over time, then we include time as a state variable. If we let

$$x = (x_1, \ldots, x_n)$$
$$F = (f_1, \ldots, f_n),$$

then the equations of motion can be written in the form

$$\Delta x = F(x).$$

A solution to this difference equation model is a sequence of points

$$x(0), x(1), x(2), \ldots$$

in the state space with

$$\Delta x(n) = x(n+1) - x(n)$$
$$= F(x(n))$$

for all n. An equilibrium point x_0 is characterized by

$$F(x_0) = 0,$$

and the equilibrium is stable if

$$x(n) \to x_0$$

whenever $x(0)$ is sufficiently close to x_0. As in the continuous time case, many other difference equation models can be reduced to the form (4.9) by introducing additional state variables.
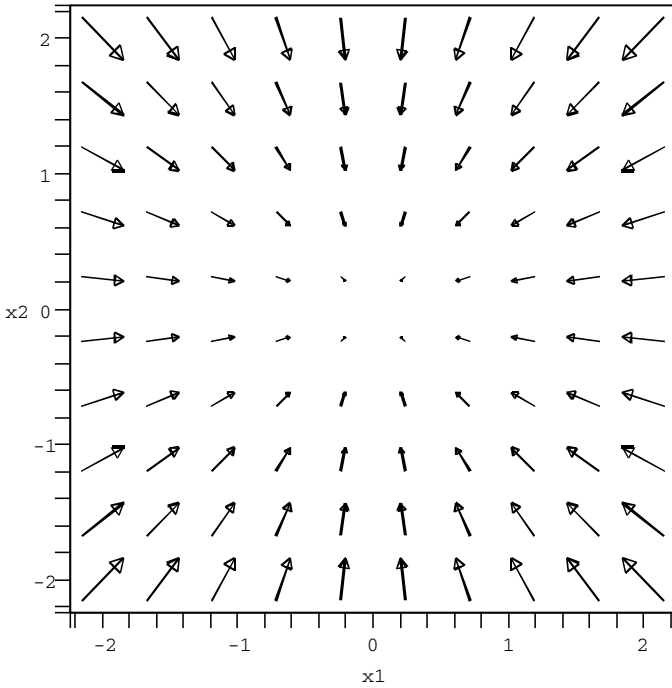
Figure 4.8: Vector field for Example 4.4.

Think of a solution as a sequence of points in the state space. The vector $F(x(n))$ connects the point $x(n)$ to the point $x(n+1)$. A graph of the vector field $F(x)$ can reveal much about the behavior of a discrete–time dynamical system.

Example 4.4. Let $x = (x_1, x_2)$, and consider the difference equation

$$\Delta x = -\lambda x, \tag{4.10}$$

where $\lambda > 0$. What is the behavior of solutions near the equilibrium point $x_0 = (0, 0)$?

Figure 4.8 shows a graph of the vector field $F(x) = -\lambda x$ in the case where $0 < \lambda < 1$. It is clear that $x_0 = (0, 0)$ is a stable equilibrium. Each step moves closer to x_0. Now let us consider what happens when λ becomes larger. Each of the vectors in Fig. 4.8 will stretch as λ increases. For $\lambda > 1$ the vectors are so long that they overshoot the equilibrium. For $\lambda > 2$ they are so long that the terminal point $x(n+1)$ is actually farther away from $(0, 0)$ than the starting point $x(n)$. In this case x_0 is an unstable equilibrium.

This simple example clearly illustrates the fact that discrete–time dynamical systems do not always behave like their continuous–time

analogs. Solutions to the differential equation

$$\frac{dx}{dt} = -\lambda x \qquad (4.11)$$

are all of the form

$$x(t) = x(0)e^{-\lambda t},$$

and the origin is a stable equilibrium regardless of $\lambda > 0$. The difference in behavior for the analogous difference equation in Eq. (4.10) is due to the inherent time delay. The approximation

$$\frac{dx}{dt} \approx \frac{\Delta x}{\Delta t}$$

is only valid for small Δt, where the term *small* depends on the sensitivity of x to t. It should only be relied upon in cases where Δx represents a small relative change in x. When this is not the case, the difference in behavior between discrete and continuous systems can be dramatic.

We return now to the docking problem of Example 4.3. Step 3 of the five-step method is to formulate the model. We are modeling the docking problem as a discrete–time dynamical system. From Fig. 4.7 we obtain

$$(v_{n+1} - v_n) = -k\,v_{n-1}\,c_n - k\,v_n w_n.$$

Hence, the change in velocity over the nth time step depends on both v_n and v_{n-1}. To simplify the analysis, let us assume that $c_n = c$ and $w_n = w$ for all n. Then the length of each time step is

$$\Delta t = c + w$$

seconds, and we do not need to include time as a state variable. We do, however, need to include both v_n and v_{n-1}. Let

$$x_1(n) = v_n$$
$$x_2(n) = v_{n-1}.$$

Compute

$$\Delta x_1 = -kwx_1 - kcx_2$$
$$\Delta x_2 = x_1 - x_2. \qquad (4.12)$$

The state space is $(x_1, x_2) \in \mathbb{R}^2$.

Step 4 is to solve the model. There is one equilibrium point $(0, 0)$ found at the intersection of the two lines

$$kwx_1 + kcx_2 = 0$$
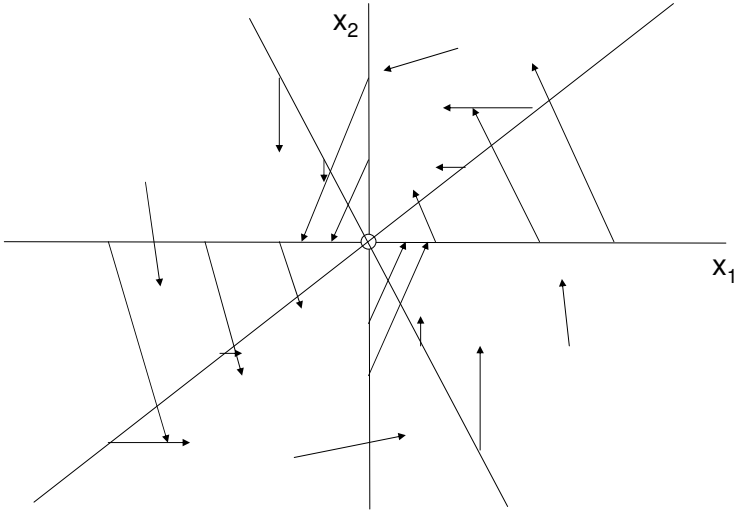$$x_1 - x_2 = 0.$$

Figure 4.9: Graph of previous velocity x_2 versus current velocity x_1 showing vector field for the docking problem.

These are the steady–state equations obtained by setting $\Delta x_1 = 0$ and $\Delta x_2 = 0$.

Figure 4.9 shows a graph of the vector field

$$F(x) = (-kwx_1 - kcx_2, \; x_1 - x_2).$$

It appears as though solutions will tend toward equilibrium, but it is hard to be sure. If k, c, and w are large, then the equilibrium is probably unstable, but once again it is difficult to tell.

In mathematics we often come across problems we cannot solve. Usually the best thing to do in such cases is to review our assumptions and consider whether we can reduce the problem to one that we can solve by making a further simplifying assumption. Of course this would be a meaningless and trivial exercise unless the simplified problem had some real significance.

In our docking problem we have expressed the change in velocity Δv_n as the sum of two components. One represents the change in velocity occurring between the time we read the velocity indicator and the time we adjust the acceleration controls. Suppose that we can do this very quickly. In particular, suppose that c is very much smaller than w. If v_n and v_{n-1} are not too different, the approximation

$$\Delta v_n \approx -kwv_n$$

should be reasonably accurate. The difference equation

$$\Delta x_1 = -kwx_1$$

is a familiar one from Example 4.4, and we know that we will get a stable equilibrium for any $kw < 2$. If $kw < 1$, we will approach the equilibrium asymptotically without overshooting.

Step 5 is to answer the analysis question in plain English. Maybe we should just say in plain English that we don't know the answer. However, we probably can do better than that. Let us report that a completely satisfactory solution is not obtainable by elementary graphical methods. In other words, it will take more work using more sophisticated methods to determine exactly the conditions under which the proposed control strategy will work. It does seem that the strategy will be effective in most cases as long as the time interval between control adjustments is not too long and the magnitude of those adjustments is not too large. The problem is complicated by the fact that there is a time delay between reading the velocity indicator and adjusting the controls. Since the actual closing velocity may change during this interval, we are acting on dated and inaccurate information. This adds an element of uncertainty to our calculations. If we ignore the effects of this time delay (which may be permissible if the delay is small), we can then draw some general conclusions, which are as follows.

The control strategy will work so long as the control adjustments are not too violent. Furthermore, the longer the interval between adjustments, the lighter those adjustments must be. In addition, the relationship is one of proportion. If we go twice as long between adjustments, we can only use half as much control. To be specific, if we adjust the controls once every 10 seconds, then we can only set the acceleration controls at 1/10 of the velocity setting to avoid overshooting the target velocity of zero. In order to allow for human and equipment error, we should actually set the controls somewhat lower, say 1/15 or 1/20 of velocity. More frequent adjustments require more frequent observations of the closing velocity indicator and more concentration on the part of the operator, but they do allow for the successful administration of more thrusting power under control. Presumably, this would be advantageous.

Normally, we would conclude our discussion of this problem with a fairly comprehensive sensitivity analysis. In view of the fact that we have not yet found a way to solve this problem, we will defer that discussion to a later chapter.

4.4   Exercises

1. Reconsider the tree problem of Example 4.1. Assume that

$$\frac{r_2}{a_2} < \frac{r_1}{b_1} \text{ and } \frac{r_1}{a_1} < \frac{r_2}{b_2}$$

so that the situation is as pictured in Fig. 4.2.

(a) Draw the vector field for this model.

(b) Classify each of the four equilibrium points as stable or unstable.

    (c) Can the two species of trees coexist in stable equilibrium?

    (d) Suppose that a logging operation removes all but a few of the valuable hardwood trees in this stand of forest. What does this model predict about the future of the two species of trees?

2. Reconsider the tree problem of Example 4.1, but now assume that

$$\frac{r_2}{a_2} < \frac{r_1}{b_1} \text{ and } \frac{r_1}{a_1} \geq \frac{r_2}{b_2}.$$

    (a) Locate each of the equilibrium points (x_1, x_2) in the state space $x_1 \geq 0$, $x_2 \geq 0$.

    (b) Draw the vector field for this case.

    (c) Classify each equilibrium as stable or unstable.

    (d) Suppose that we start out with an equal amount of hardwood and softwood trees. What does this model predict about the future of the two species?

3. Repeat Exercise 2, but now assume that

$$\frac{r_2}{a_2} \geq \frac{r_1}{b_1} \text{ and } \frac{r_1}{a_1} < \frac{r_2}{b_2}.$$

4. In the whale problem of Example 4.2 we used a logistic model of population growth, where the growth rate of population P in the absence of interspecies competition is

$$g(P) = rP \left(1 - \frac{P}{K} \right).$$

In this problem we will be using the simpler growth model

$$g(P) = rP.$$

    (a) Can both species of whales coexist? Use the five-step method, and model as a dynamical system in steady state.

    (b) Draw the vector field for this model. Indicate the location of each equilibrium point.

    (c) Classify each equilibrium point in the state space as stable or unstable.

    (d) Suppose that there are currently 5,000 blue whales and 70,000 fin whales. What does this model predict about the future of the two species?

5. In the whale problem of Example 4.2 we used a logistic model of population growth, where the growth rate of population P in the absence of interspecies competition is

$$g(P) = rP \left(1 - \frac{P}{K} \right).$$

In this problem we will be using a more complex model,

$$g(P) = rP \left(\frac{P - c}{P + c} \right) \left(1 - \frac{P}{K} \right),$$

in which the parameter c represents a minimum viable population level below which the growth rate is negative. Assume that $\alpha = 10^{-8}$ and that the minimum viable population level is 3,000 for blue whales and 15,000 for fin whales.

(a) Can the two species of whales coexist? Use the five-step method, and model as a dynamical system in steady state.

(b) Sketch the vector field for this model. Classify each equilibrium point as stable or unstable.

(c) Assuming that there are currently 5,000 blue whales and 70,000 fin whales, what does this model predict about the future of the two populations?

(d) Suppose that we have underestimated the minimum viable population for the blue whale, and that it is actually closer to 10,000. Now what happens to the two species?

6. Reconsider the whale problem of Example 4.2, and assume that $\alpha = 10^{-8}$. In this problem we will investigate the effects of harvesting on the two whale populations. Assume that a level of effort E boat–days will result in the annual harvest of qEx_1 blue whales and qEx_2 fin whales, where the parameter q (catchability) is assumed to equal approximately 10^{-5}.

(a) Under what conditions can both species continue to coexist in the presence of harvesting? Use the five-step method, and model as a dynamical system in steady state.

(b) Draw the vector field for this problem, assuming that the conditions identified in part (a) are satisfied.

(c) Find the minimum level of effort required to reduce the fin whale population to its current level of around 70,000 whales. Assume that we started out with 150,000 blue whales and 400,000 fin whales before mankind began to harvest them.

(d) Describe what would happen to the two populations if harvesting were allowed to continue at the level of effort identified in part (c). Draw the vector field in this case. This is the situation which led the IWC to call for an international ban on whaling.

7. One of the favorite foods of the blue whale is called krill. These tiny shrimp–like creatures are devoured in massive amounts to provide the principal food source for the huge whales. The maximum sustainable population for krill is 500 tons/acre. In the absence of predators, in uncrowded

conditions, the krill population grows at a rate of 25% per year. The presence of 500 tons/acre of krill increases blue whale population growth rate by 2% per year, and the presence of 150,000 blue whales decreases krill growth rate by 10% per year.

(a) Determine whether the whales and the krill can coexist in equilibrium. Use the five-step method, and model as a dynamical system in steady state.

(b) Draw the vector field for this problem. Classify each equilibrium point in the state space as stable or unstable.

(c) Describe what happens to the two populations over time. Assume that we start off with 5,000 blue whales and 750 tons/acre of krill.

(d) How sensitive are your conclusions in part (c) to the assumption of a 25% growth rate per year for krill?

8. Two armies are to engage in battle. The red army enjoys a three–to–one numerical superiority, but the blue army is better trained and better equipped. Let R and B denote the force levels of red and blue forces. The Lanchester model of combat states that

$$R' = -aB - bRB$$
$$B' = -cR - dRB,$$

where the first term accounts for direct fire (aimed at a specific target) and the second term accounts for attrition due to area fire (e.g., artillery). We are assuming that weapon effectiveness is higher for blue than for red; i.e., $a > c$ and $b > d$. But what kind of edge in weapon effectiveness would be necessary to counteract a $3 : 1$ numerical superiority?

(a) Assuming that $a = \lambda c$ and $b = \lambda d$ for some $\lambda > 1$, determine the approximate lower bound on λ necessary for blue to win the war. Use the five-step method, and model as a dynamical system.

(b) In part (a) you assumed red had an $n : 1$ numerical superiority. Discuss the sensitivity of your results in part (a) to the parameter $n \in (2, 5)$.

9. The following simple model is intended to represent the dynamics of supply and demand. Let P denote the selling price of a certain product and Q the quantity of this product being produced. The supply curve $Q = f(P)$ tells how much should be produced at a given price to maximize profit. The demand curve $P = g(Q)$ tells what price buyers should pay given a certain level of production in order to maximize their utility.

(a) Select a specific product and make an educated guess as to the form of the supply curve $Q = f(P)$ and the demand curve $P = g(Q)$.

(b) Use the results of part (a) to determine the equilibrium levels of P and Q.

(c) Formulate a dynamic model based on the assumption that P will tend toward the level dictated by the demand curve, while Q will tend to the level given by the supply curve.

(d) According to your model, is the (P, Q) equilibrium stable? Does it matter whether you assume a discrete–time or continuous–time model? (Economists usually assume a discrete–time model in order to represent the effect of a time delay.)

(e) Perform a sensitivity analysis for the assumptions you made in part (a). Consider the question of stability.

10. A population of 100,000 members is subject to a disease that is seldom fatal and leaves the victim immune to future infections by this disease. Infection can only occur when a susceptible person comes in direct contact with an infectious person. The infectious period lasts approximately three weeks. Last week, there were 18 new cases of the disease reported. This week, there were 40 new cases. It is estimated that 30% of the population is immune due to previous exposure.

   (a) What is the eventual number of people who will become infected? Use the five-step method, and model as a discrete–time dynamical system.

   (b) Estimate the maximum number of new cases in any one week.

   (c) Conduct a sensitivity analysis to investigate the effect of any assumptions you made in part (a) that were not supported by hard data.

   (d) Perform a sensitivity analysis for the number of cases (18) reported last week. It is thought by some that in early weeks the epidemic might be underreported.

11. Reconsider the docking problem of Example 4.3, and now assume that $c = 5$ sec, $w = 10$ sec, and $k = 0.02$.

   (a) Assuming an initial closing velocity of 50 m/sec, calculate the sequence of velocity observations v_0, v_1, v_2, ..., predicted by the model. Is the docking procedure successful?

   (b) An easier way to compute the solution in part (a) is to use the iteration function $G(x) = x + F(x)$, with the property that $x(n + 1) = G(x(n))$. Compute the iteration function for this problem, and use it to repeat the calculation in part (a).

   (c) Calculate the solution $x(1)$, $x(2)$, $x(3)$, ..., starting at $x(0) = (1, 0)$. Repeat, starting at $x(0) = (0, 1)$. What happens as $n \to \infty$? What does this imply about the stability of the equilibrium $(0, 0)$? [Hint: Every possible initial condition $x(0) = (a, b)$ can be written as a linear combination of the vectors $(1, 0)$ and $(0, 1)$, and $G(x)$ is a linear function of x.]

   (d) Are there any states x for which $G(x) = \lambda x$ for some real λ? If so, what happens to the system if we start with this initial condition?

Further Reading

1. Bailey, N. (1975) *The Mathematical Theory of Infectious Disease.* Hafner Press, New York.

2. Casstevens, T. *Population Dynamics of Governmental Bureaus.* UMAP module 494.

3. Clark, C. (1976) *Mathematical Bioeconomics: The Optimal Management of Renewable Resources.* Wiley, New York.

4. Giordano, F. and Leja, S. *Competitive Hunter Models.* UMAP module 628.

5. Greenwell, R. *Whales and Krill: A Mathematical Model.* UMAP module 610.

6. Greenwell, R. and Ng, H. *The Ricker Salmon Model.* UMAP module 653.

7. Hirsch, M. and Smale, S. (1974) *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York.

8. Horelick, B., Koont, S. and Gottleib, S. *Population Growth and the Logistic Curve.* UMAP module 68.

9. Horelick, B. and Koont, S. *Epidemics.* UMAP module 73.

10. Lanchester, F. (1956) Mathematics in Warfare. *The World of Mathematics.* Vol. 4, pp. 1240–1253.

11. May, R. (1976) *Theoretical Ecology: Principles and Applications.* Saunders, Philadelphia.

12. Morrow, J. *The Lotka–Volterra Predator–Prey Model.* UMAP module 675.

13. Sherbert, D. *Difference Equations with Applications.* UMAP module 322.

14. Tuchinsky, P. *Management of a Buffalo Herd.* UMAP module 207.

Chapter 5

ANALYSIS OF DYNAMIC MODELS

In this chapter we consider some of the most broadly applicable techniques for the analysis of discrete and continuous time dynamical systems. Except for a few special cases, these methods do not yield exact analytical solutions. Such exact analytical methods are more properly treated in a course on differential equations. In any case, most dynamical system models which arise in practice are not amenable to exact solution by any known technique. In this chapter, we will present techniques which can be applied to the analysis of almost any dynamical system model. These methods can provide important qualitative information about the behavior of dynamical systems, even when exact analytic solutions are not obtainable.

5.1   Eigenvalue Methods

When the equations of a dynamic model are linear, it is possible to obtain an exact analytical solution. While linear dynamics are rare in real life, the majority of dynamic systems can be approximated by linear systems, at least locally. Such linear approximations, especially in the neighborhood of an isolated equilibrium point, provide the basis for many of the most important analytical techniques available for dynamic modeling.

Example 5.1. Reconsider the tree problem of Example 4.1. Assume that hardwoods grow at a rate of 10% per year and softwoods at a rate of 25% per year. An acre of forest land can support about 10,000 tons of hardwoods or 6,000 tons of softwoods. The extent of competition has not been numerically determined. Can both types of trees coexist in stable equilibrium?

Step 1 of the five-step method was laid out in Figure 4.1. In this particular

case we have

$$r_1 = 0.10$$
$$r_2 = 0.25$$
$$a_1 = \frac{0.10}{10,000}$$
$$a_2 = \frac{0.25}{6,000}.$$

Step 2 is to select the modeling approach, including a method of analysis. We will analyze this nonlinear dynamical system by the eigenvalue method.

Suppose we are given a dynamical system $x' = F(x)$ where $x = (x_1, \ldots, x_n)$ is an element of the state space $S \subseteq \mathbb{R}^n$ and $F = (f_1, \ldots, f_n)$. A point $x_0 \in S$ is an equilibrium or steady state if and only if $F(x_0) = 0$. There is a theorem that states that an equilibrium point x_0 is asymptotically stable if the matrix

$$A = \begin{pmatrix} \partial f_1 / \partial x_1 \, (x_0) & \cdots & \partial f_1 / \partial x_n \, (x_0) \\ \vdots & & \vdots \\ \partial f_n / \partial x_1 \, (x_0) & \cdots & \partial f_n / \partial x_n \, (x_0) \end{pmatrix} \qquad (5.1)$$

has eigenvalues with all negative real parts. If any eigenvalue has a positive real part, then the equilibrium is unstable. In the remaining cases (pure imaginary eigenvalues) the test is inconclusive (Hirsch and Smale (1974) p. 187).

The eigenvalue method is based on a linear approximation. Even if $x' = F(x)$ is not linear, we will have

$$F(x) \approx A(x - x_0)$$

in the neighborhood of the equilibrium point. This is the same sort of linear approximation you saw in one–variable calculus, except that now the derivative of F is represented by a matrix. Some authors will call this matrix DF in analogy to the one–variable derivative. The linear approximation is good enough so that if the origin is a stable equilibrium of $x' = Ax$ (i.e., the point x_0 is a stable equilibrium of $x' = A(x - x_0)$), then x_0 is a stable equilibrium of $x' = F(x)$ as well. Therefore, it is enough to understand the eigenvalue test in the case of a linear system.

Undoubtedly, you solved some linear systems of differential equations in your introductory differential equations course, and you probably learned about the relation between solutions and eigenvalues. For example, if $Au = \lambda u$ (i.e., u is an eigenvector of A belonging to the eigenvalue λ), then $x(t) = ue^{\lambda t}$ is a solution to the initial value problem

$$x' = Ax, \ x(0) = u.$$

It is actually possible to write down the general solution to an $n \times n$ system of linear differential equations, although it is rather messy and requires a lot of linear algebra. One good thing that comes out of doing so, however, is a general description of solution behavior. This theorem says that for any solution $x(t)$ to the differential equation $x' = Ax$, where A is a matrix of constants, each coordinate is a linear combination of terms that look like one of

$$t^k e^{at} \cos(bt), t^k e^{at} \sin(bt)$$

where $a \pm ib$ is an eigenvalue of A (if the eigenvalue is real then $b = 0$), and k is a nonnegative integer less than n. From this general description it is easy to calculate that the origin is an asymptotically stable equilibrium of the system $x' = Ax$ if and only if every eigenvalue $a \pm ib$ has $a < 0$. (Hirsch and Smale (1974) p. 135)

Of course, a successful application of the eigenvalue method requires us to be able to compute the eigenvalues. In simple cases (e.g. on $\mathbb{R}^2$) it will be possible to compute eigenvalues by hand, or possibly with the aid of a computer algebra system. Otherwise, we will have to rely on approximate methods. Fortunately, there do exist numerical analysis software packages to compute the eigenvalues of an $n \times n$ matrix, and these are effective in most cases. (e.g., Press (1986))

Returning to Example 5.1, recall from Section 4.1 that there is an equilibrium at the point

$$x_1 = \frac{r_1 a_2 - r_2 b_1}{D}$$
$$x_2 = \frac{a_1 r_2 - b_2 r_1}{D}$$

where $D = a_1 a_2 - b_1 b_2$. We have now specified values for a_1, a_2, r_1, and r_2 but not for b_1 and b_2. We will, however, continue to assume that $b_i < a_i$. For the moment let us take $b_i = a_i/2$. Then the coordinates of the equilibrium point are $x_0 = (x_1^0, x_2^0)$, where

$$x_1^0 = \frac{28000}{3} \approx 9333$$
$$x_2^0 = \frac{4000}{3} \approx 1333$$

(5.2)

The dynamical system equations are $x' = F(x)$ where $F = (f_1, f_2)$ and

$$f_1(x_1, x_2) = 0.10 x_1 - \frac{0.10}{10000} x_1^2 - \frac{0.05}{10000} x_1 x_2$$
$$f_2(x_1, x_2) = 0.25 x_2 - \frac{0.25}{6000} x_2^2 - \frac{0.125}{6000} x_1 x_2$$

(5.3)

The partial derivatives are

$$
\begin{aligned}
\frac{\partial f_1}{\partial x_1} &= \frac{20000 - x_2}{200000} - \frac{x_1}{50000} \\
\frac{\partial f_1}{\partial x_2} &= \frac{-x_1}{200000} \\
\frac{\partial f_2}{\partial x_1} &= \frac{-x_2}{48000} \\
\frac{\partial f_2}{\partial x_2} &= \frac{-x_1}{48000} - \frac{x_2}{12000} + \frac{1}{4}
\end{aligned}
\tag{5.4}
$$

Evaluating the partial derivatives (5.4) at the equilibrium point (5.2) and substituting back into (5.1), we obtain

$$
A = \begin{pmatrix} -7/75 & -7/150 \\ -1/36 & -1/18 \end{pmatrix}.
\tag{5.5}
$$

The eigenvalues of this 2×2 matrix can be computed as the roots of the equation

$$
\begin{vmatrix} \lambda + 7/75 & 7/150 \\ 1/36 & \lambda + 1/18 \end{vmatrix} = 0.
$$

Evaluating the determinant, we obtain the equation

$$
\frac{1800\lambda^2 + 268\lambda + 7}{1800} = 0,
$$

and then we obtain

$$
\lambda = \frac{-67 \pm \sqrt{1339}}{900}.
$$

Since both eigenvalues have negative real parts, the equilibrium is stable.

The eigenvalue test for continuous time dynamical systems involves quite a bit of computation. This is an appropriate application for a computer algebra system. Figure 5.1 illustrates the use of the computer algebra system Mathematica to perform the computations in step 4 for the present problem.

Finally, we proceed to step 5. We have found that hardwoods and softwoods can coexist in stable equilibrium. There will be approximately 9,300 tons per acre of hardwoods and 1,300 tons per acre of softwoods in a mature, stable forest. These conclusions are based on certain plausible assumptions about the degree of competition between the two types of trees. A sensitivity analysis will be conducted to determine the effect of these assumptions on our broad conclusions.

For the sensitivity analysis, we will still assume that $b_i = t\, a_i$ but we will relax the assumption that $t = 1/2$. The conditions

$$
b_i < a_i
$$
$$
(r_i/a_i) < (r_j/b_j)
$$

In[1]:= **f1 = x1 / 10 − (x1 ^ 2 / 10) / 10000 − (5 x1 x2 / 100) / 10000**

Out[1]= $\dfrac{x1}{10} - \dfrac{x1^2}{100000} - \dfrac{x1\,x2}{200000}$

In[2]:= **f2 = 25 x2 / 100 − (25 x2 ^ 2 / 100) / 6000 − (125 x1 x2 / 1000) / 6000**

Out[2]= $\dfrac{x2}{4} - \dfrac{x1\,x2}{48000} - \dfrac{x2^2}{24000}$

In[3]:= **s = Solve[{f1 / x1 == 0, f2 / x2 == 0}, {x1, x2}]**

Out[3]= $\left\{\left\{x1 \to \dfrac{28000}{3},\ x2 \to \dfrac{4000}{3}\right\}\right\}$

In[4]:= **df = {{D[f1, x1], D[f1, x2]}, {D[f2, x1], D[f2, x2]}};**

In[6]:= **MatrixForm[df]**

Out[6]//MatrixForm=

$$\begin{pmatrix} \frac{1}{10} - \frac{x1}{50000} - \frac{x2}{200000} & -\frac{x1}{200000} \\ -\frac{x2}{48000} & \frac{1}{4} - \frac{x1}{48000} - \frac{x2}{12000} \end{pmatrix}$$

In[7]:= **A = df /. s**

Out[7]= $\left\{\left\{\left\{-\dfrac{7}{75},\ -\dfrac{7}{150}\right\},\ \left\{-\dfrac{1}{36},\ -\dfrac{1}{18}\right\}\right\}\right\}$

In[8]:= **Eigenvalues[A]**

Out[8]= $\left\{\dfrac{1}{900}\left(-67 - \sqrt{1339}\right),\ \dfrac{1}{900}\left(-67 + \sqrt{1339}\right)\right\}$

Figure 5.1: Calculations for step 4 of the tree problem using the computer algebra system Mathematica.

imply that $0 < t < 0.6$. The coordinates of the equilibrium point (x_1^0, x_2^0) are

$$\begin{aligned} x_1^0 &= \frac{10000 - 6000t}{1 - t^2} \\ x_2^0 &= \frac{6000 - 10000t}{1 - t^2}. \end{aligned} \tag{5.6}$$

The differential equations of this system are $x_i' = f_i(x_1, x_2)$ where

$$\begin{aligned} f_1(x_1, x_2) &= 0.10x_1 - \frac{0.10x_1^2}{10000} - \frac{0.10tx_1x_2}{10000} \\ f_2(x_1, x_2) &= 0.25x_2 - \frac{0.25x_2^2}{6000} - \frac{0.25tx_1x_2}{6000} \end{aligned} \tag{5.7}$$

and the partial derivatives are

$$\frac{\partial f_1}{\partial x_1} = \frac{10000 - tx_2}{100000} - \frac{x_1}{50000}$$

$$\frac{\partial f_1}{\partial x_2} = \frac{-tx_1}{100000}$$

$$\frac{\partial f_2}{\partial x_1} = \frac{-tx_2}{24000}$$

$$\frac{\partial f_2}{\partial x_2} = \frac{-tx_1}{24000} - \frac{x_2}{12000} + \frac{1}{4}. \tag{5.8}$$

Evaluating the partial derivatives (5.8) at the equilibrium point (5.6) and substituting back into (5.1) yields

$$A = \begin{pmatrix} \dfrac{5 - 3t}{50(t^2 - 1)} & \dfrac{t(5 - 3t)}{50(t^2 - 1)} \\ \dfrac{t(3 - 5t)}{12(t^2 - 1)} & \dfrac{3 - 5t}{12(t^2 - 1)} \end{pmatrix}. \tag{5.9}$$

The characteristic equation we must solve to find the eigenvalues is

$$\left[\lambda - \frac{5 - 3t}{50(t^2 - 1)}\right]\left[\lambda - \frac{3 - 5t}{12(t^2 - 1)}\right] - \left[\frac{t(3 - 5t)}{12(t^2 - 1)}\right]\left[\frac{t(5 - 3t)}{50(t^2 - 1)}\right] = 0. \tag{5.10}$$

Solving equation (5.10) for λ yields two roots:

$$\lambda_1 = \frac{143t - 105 + \sqrt{9000t^4 - 20400t^3 + 20449t^2 - 9630t + 2025}}{600(1 - t^2)}$$

$$\lambda_2 = \frac{143t - 105 - \sqrt{9000t^4 - 20400t^3 + 20449t^2 - 9630t + 2025}}{600(1 - t^2)}. \tag{5.11}$$

Figure 5.2 illustrates the use of the computer algebra system Maple to compute the eigenvalues for this problem. Computer algebra systems are especially useful for problems like this one, where the calculations become complicated, and there is an increased risk of error when doing all of the algebra by hand. Most computer algebra systems also include a graphing utility. The combination of graphics and algebra is important in problems such as the present one. Drawing a graph is often the easiest way to solve an inequality.

Figure 5.3 shows a graph of λ_1 and λ_2 versus t over the interval $0 < t < 0.6$. From this graph we can see that λ_1, λ_2 are always negative, so that the equilibrium is stable regardless of the strength of competition. (If you did Exercise 1 of Chapter 4, you probably drew the same conclusion from a graphical analysis.)

5.2   Eigenvalue Methods for Discrete Systems

The methods of the previous section apply only to continuous time dynamical systems. In this section we will present analogous methods for the stability

```
[ > with(linalg):
[ > f1:=x1/10-(x1^2/10)/10000-(t*x1*x2/10)/10000;
```

$$f1 := \frac{1}{10}x1 - \frac{1}{100000}x1^2 - \frac{1}{100000}t\, x1\, x2$$

```
[ > f2:=25*x2/100-(25*x2^2/100)/6000-(25*t*x1*x2/100)/6000;
```

$$f2 := \frac{1}{4}x2 - \frac{1}{24000}x2^2 - \frac{1}{24000}t\, x1\, x2$$

```
[ > df1dx1:=diff(f1,x1);
```

$$df1dx1 := \frac{1}{10} - \frac{1}{50000}x1 - \frac{1}{100000}t\, x2$$

```
[ > df1dx2:=diff(f1,x2);
```

$$df1dx2 := -\frac{1}{100000}t\, x1$$

```
[ > df2dx1:=diff(f2,x1);
```

$$df2dx1 := -\frac{1}{24000}t\, x2$$

```
[ > df2dx2:=diff(f2,x2);
```

$$df2dx2 := \frac{1}{4} - \frac{1}{12000}x2 - \frac{1}{24000}t\, x1$$

```
[ > s:=solve({f1/x1=0,f2/x2=0},{x1,x2});
```

$$s := \{ x2 = 2000\,\frac{-3+5\,t}{-1+t^2}, x1 = 2000\,\frac{-5+3\,t}{-1+t^2} \}$$

```
[ > assign(s);
[ > A:=array([[df1dx1,df1dx2],[df2dx1,df2dx2]]);
```

$$A := \begin{bmatrix} \dfrac{1}{10} - \dfrac{1}{25}\dfrac{-5+3\,t}{-1+t^2} - \dfrac{1}{50}\dfrac{t\,(-3+5\,t)}{-1+t^2} & -\dfrac{1}{50}\dfrac{t\,(-5+3\,t)}{-1+t^2} \\[2mm] -\dfrac{1}{12}\dfrac{t\,(-3+5\,t)}{-1+t^2} & \dfrac{1}{4} - \dfrac{1}{6}\dfrac{-3+5\,t}{-1+t^2} - \dfrac{1}{12}\dfrac{t\,(-5+3\,t)}{-1+t^2} \end{bmatrix}$$

```
[ > eigenvals(A);
```

$$\frac{1}{2}\frac{-286\,t+210+2\sqrt{20449\,t^2-9630\,t+2025-20400\,t^3+9000\,t^4}}{600\,t^2-600},$$

$$\frac{1}{2}\frac{-286\,t+210-2\sqrt{20449\,t^2-9630\,t+2025-20400\,t^3+9000\,t^4}}{600\,t^2-600}$$

Figure 5.2: Calculations for sensitivity analysis in the tree problem of Example 5.1 using the computer algebra system Maple.

analysis of discrete time dynamical systems. Once again, the basis for our analysis is a linear approximation, together with a calculation of eigenvalues.

Example 5.2. Reconsider the docking problem of Example 4.3, and assume now that it takes 5 seconds to make the control adjustments, and another 10 seconds until we can return from other tasks to observe the velocity indicator once again. Under these conditions, will our strategy for matching velocities be successful?

Step 1 of the five-step method was summarized in Figure 4.7. Now we will assume $c_n = 5, w_n = 10$. For the moment we will set $k = 0.02$, and then later we will perform a sensitivity analysis on k.

Step 2 is to select the modeling approach, including the method of solution. We will use an eigenvalue method.
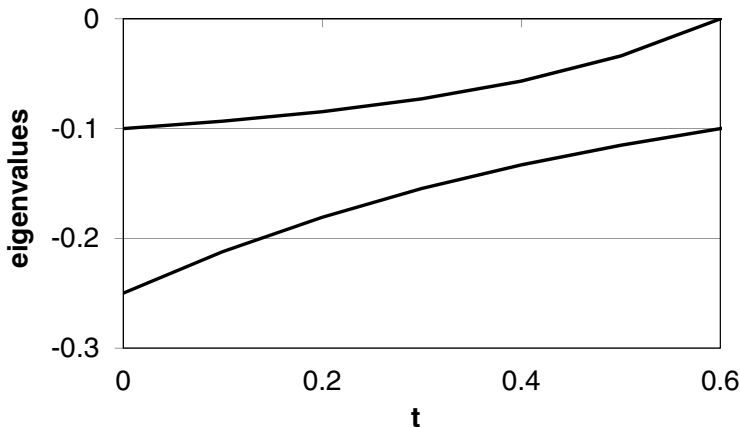
Figure 5.3: Graph of the eigenvalues λ_1 and λ_2 from (5.11) versus the parameter t in the tree problem.

Given a discrete time dynamical system

$$\Delta x = F(x)$$

where $x = (x_1, \ldots, x_n)$ and $F = (f_1, \ldots, f_n)$, let us define the iteration function

$$G(x) = x + F(x).$$

The sequence $x(0), x(1), x(2), \ldots$ is a solution to this system of difference equations if and only if

$$x(n+1) = G(x(n))$$

for all n. An equilibrium point x_0 is characterized by the fact that x_0 is a fixed point of the function $G(x)$; i.e., $G(x_0) = x_0$.

There is a theorem that says that an equilibrium point x_0 is (asymptotically) stable if every eigenvalue of the matrix of partial derivatives

$$A = \begin{pmatrix} \partial g_1/\partial x_1(x_0) & \cdots \partial g_1/\partial x_n(x_0) \\ \vdots & \vdots \\ \partial g_n/\partial x_1(x_0) & \cdots \partial g_n/\partial x_n(x_0) \end{pmatrix} \qquad (5.12)$$

has absolute value less than one. If the eigenvalue is complex $a \pm ib$ then by "absolute value" we mean the complex absolute value $\sqrt{a^2 + b^2}$. This simple test for stability is analogous to the eigenvalue test for continuous time dynamical systems presented in the preceding section (Hirsch and Smale (1974) p. 280).

As in the continuous case, the eigenvalue method for discrete time dynamical systems is based on a linear approximation. Even if the iteration function $G(x)$ is not linear, we will have

$$G(x) \approx A(x - x_0)$$

in the neighborhood of the equilibrium point x_0. In other words the behavior of the iteration function G in the neighborhood of the equilibrium point x_0 is approximately the same as the behavior of the linear function Ax near the origin. Therefore, the behavior of our original nonlinear system in the neighborhood of x_0 is approximately the same as the behavior of the linear discrete time dynamical system defined by the iteration function

$$x(n + 1) = Ax(n)$$

in the neighborhood of the origin. The linear approximation is good enough so that if the origin is a stable equilibrium of the linear system, then x_0 is a stable equilibrium of the original nonlinear system. So it only remains to discuss the conditions for stability of the linear system.

A matrix A is called a *linear contraction* if $A^n x \to 0$ for every x. There is a theorem that states that if every eigenvalue of a matrix A has absolute value less than one, then A is a linear contraction (Hirsch and Smale (1974) p. 279). It follows that the origin is a stable equilibrium of the discrete time dynamical system with iteration function A whenever the eigenvalues of A all have absolute value less than one. We illustrate the proof of this result in a simple case. Suppose that $Ax = \lambda x$ for *all* x. Then λ is an eigenvalue of A, and every nonzero vector x is an eigenvector belonging to λ. In this simple case we will always have

$$x(n + 1) = Ax(n) = \lambda x(n)$$

so that the origin is a stable equilibrium if and only if $|\lambda| < 1$.

Now we return to the docking problem. Step 3 is to formulate the model as necessary for the application of the techniques identified in step 2. In this case we already have a linear system with equilibrium $x_0 = (0, 0)$. The iteration function is

$$G(x_1, x_2) = x + F(x_1, x_2) = (g_1, g_2)$$

where

$$g_1(x_1, x_2) = 0.8x_1 - 0.1x_2$$
$$g_2(x_1, x_2) = x_1.$$

Moving on to step 4, we calculate

$$\begin{vmatrix} \lambda - 0.8 & 0.1 \\ -1 & \lambda - 0 \end{vmatrix} = 0$$

or $\lambda^2 - 0.8\lambda + 0.1 = 0$, from which we obtain

$$\lambda = \frac{4 \pm \sqrt{6}}{10}.$$

There are $n = 2$ distinct eigenvalues, and both are real and lie between -1 and $+1$. Hence the equilibrium $x_0 = 0$ is stable, and so we will get $x(t) \to (0,0)$ for any initial condition.

Step 5 is to state our results in plain English. We assumed 15 seconds between control adjustments; 5 seconds to make the adjustment and 10 seconds slack time. Using a correction factor of $1 : 50$, we can guarantee success for our proportional method of control. In practical terms, a correction factor of $1 : 50$ means that if the velocity indicator reads 50 m/sec we will set the acceleration controls for -1 m/sec^2; if the reading is 25 m/sec, we set the controls at -0.5 m/sec^2, and so on.

What follows is a sensitivity analysis for the parameter k. For a general k the iteration function is given by $G = (g_1, g_2)$ where

$$g_1(x_1, x_2) = (1 - 10k)x_1 - 5kx_2$$
$$g_2(x_1, x_2) = x_1$$

which leads to the characteristic equation

$$\lambda^2 - (1 - 10k)\lambda + 5k = 0.$$

The eigenvalues are

$$\lambda_1 = \frac{(1 - 10k) + \sqrt{(1 - 10k)^2 - 20k}}{2}$$
$$\lambda_2 = \frac{(1 - 10k) - \sqrt{(1 - 10k)^2 - 20k}}{2}. \tag{5.13}$$

The quantity under the radical in (5.13) becomes negative between

$$k_1 = \frac{4 - \sqrt{12}}{20} \approx 0.027$$

and

$$k_2 = \frac{4 + \sqrt{12}}{20} \approx 0.373.$$

Figure 5.4 shows a graph of λ_1 and λ_2 over the interval $0 < k \le k_1$. From the graph we can see that both eigenvalues have absolute value less than one, so that the equilibrium $(0,0)$ is stable over this entire range of k. For $k_1 < k < k_2$ both eigenvalues are complex, and the condition for stability is that

$$\left[\frac{(1 - 10k)}{2}\right]^2 + \left[\frac{\sqrt{20k - (1 - 10k)^2}}{2}\right]^2 < 1 \tag{5.14}$$

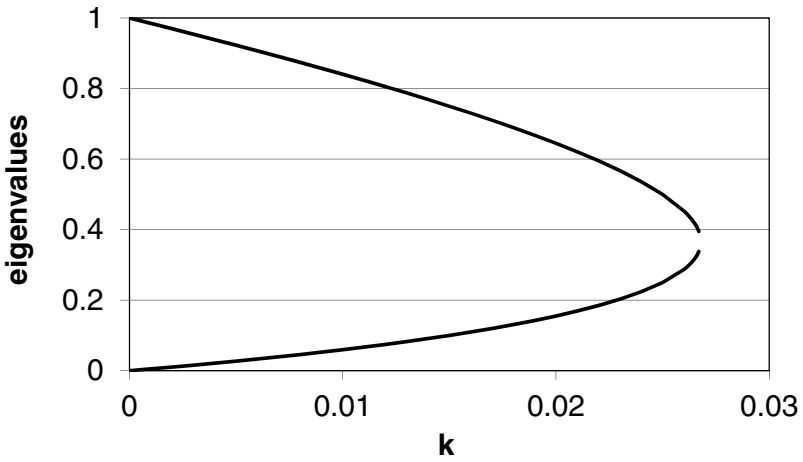Figure 5.4: Graph of the eigenvalues λ_1 and λ_2 from (5.13) versus control parameter k in the docking problem: case $0 < k \leq k_1$.
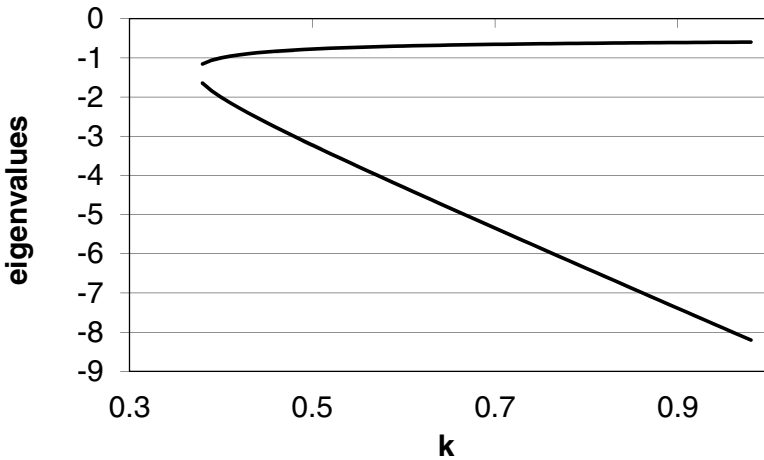

Figure 5.5: Graph of the eigenvalues λ_1 and λ_2 from (5.13) versus control parameter k in the docking problem: case $k \geq k_2$.
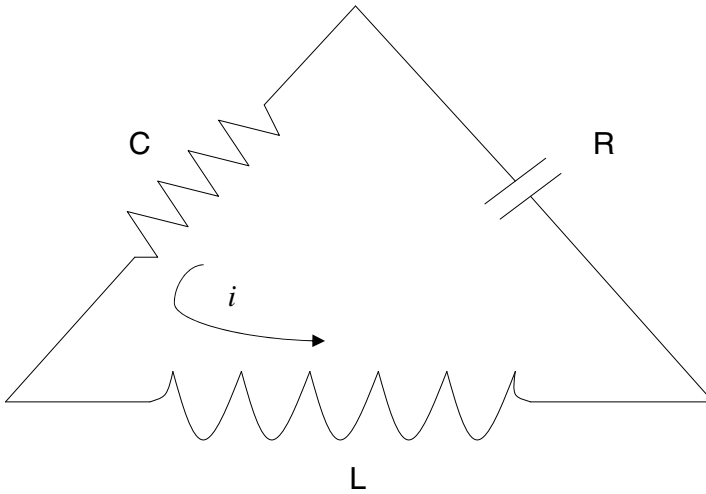
Figure 5.6: RLC circuit diagram for Example 5.3.

which reduces to $k < 1/5$. Figure 5.5 shows a graph of λ_1 and λ_2 for $k \geq k_2$. It is easy to see that the smaller eigenvalue λ_2 has absolute value greater than one for all such k. To summarize, the method will achieve matched velocities as long as $k < 0.2$, or at least a $1:5$ correction factor. Of course, it is of some interest to know which value of k is the most efficient. We will leave this problem for the exercises.

5.3   Phase Portraits

In Section 5.1 we introduced the eigenvalue test for stability in continuous time dynamical systems. This test is based on the idea of a linear approximation in the neighborhood of an isolated equilibrium point. In this section we will show how this simple idea can be used to obtain a graphical description of the behavior of a dynamical system near an equilibrium point. This information can then be used along with a sketch of the vector field to obtain a graphical description of the dynamics over the entire state space, called the *phase portrait*. Phase portraits are important in the analysis of nonlinear dynamical systems because, in most cases, it is not possible to obtain exact analytical solutions. At the end of this section we also include a brief discussion of some similar techniques for discrete time dynamical systems, again based on the idea of linear approximation.

Example 5.3. Consider the electrical circuit diagrammed in Figure 5.6. The circuit consists of a capacitor, a resistor, and an inductor in a simple closed

Variables: v_C = voltage across capacitor
                       i_C = current through capacitor
                       v_R = voltage across resistor
                       i_R = current through resistor
                       v_L = voltage across inductor
                       i_L = current through inductor

Assumptions: $C \, dv_C/dt = i_C$
                       $v_R = f(i_R)$
                       $L \, di_L/dt = v_L$
                       $i_C = i_R = i_L$
                       $v_C + v_R + v_L = 0$
                       $L = 1$
                       $C = 1/3$
                       $f(x) = x^3 + 4x$

Objective: Determine the behavior of all six variables over time

Figure 5.7: Step 1 of the RLC circuit problem.

loop. The effect of each component of the circuit is measured in terms of the re-
lationship between current and voltage on that branch of the loop. An idealized
physical model gives the relations

$$C\frac{dv_C}{dt} = i_C \text{ (capacitor)}$$

$$v_R = f(i_R) \text{ (resistor)}$$

$$L\frac{di_L}{dt} = v_L \text{ (inductor)}$$

where v_C represents the voltage across the capacitor, i_R represents the current
through the resistor, and so on. The function $f(x)$ is called the *v-i characteristic*
of the resistor. Usually $f(x)$ has the same sign as x. This is called a passive
resistor. Some control circuits use an active resistor, where $f(x)$ and x have
opposite sign for small x, see Example 5.4. In the classical linear model of
the RLC circuit, we assume that $f(x) = Rx$ where $R > 0$ is the resistance.
Kirchoff's current law states that the sum of the currents flowing into a node
equals the sum of the currents flowing out. Kirchoff's voltage law states that the
sum of the voltage drops along a closed loop must add up to zero. Determine
the behavior of this circuit over time in the case where $L = 1$, $C = 1/3$, and
$f(x) = x^3 + 4x$.

We will use the five-step method. The results of step 1 are summarized in
Figure 5.7. Step 2 is to select a modeling approach. We will model this problem
using a continuous time dynamical system, which we will analyze by sketching
the complete phase portrait.

Suppose that we are given a dynamical system $x' = F(x)$ where $x = (x_1 \ldots, x_n)$ and F has continuous first partial derivatives in the neighborhood of an equilibrium point x_0. Let A denote the matrix of first partial derivatives evaluated at the equilibrium point x_0 as defined by (5.1). We have stated previously that for x near x_0 the system $x' = F(x)$ behaves like the linear system $x' = A(x - x_0)$. Now we will be more specific.

The *phase portrait* of a continuous time dynamical system is simply a sketch of the state space showing a representative selection of solution curves. It is not hard to draw the phase portrait for a linear system (at least on $\mathbb{R}^2$) because we can always find an exact solution to a linear system of differential equations. Then we can just graph the solutions for a few initial conditions to get the phase portrait. We refer the reader to any textbook on differential equations for the details on how to solve linear systems of differential equations. For nonlinear systems, we can draw an approximate phase portrait in the neighborhood of each isolated equilibrium point by using the linear approximation.

A *homeomorphism* is a continuous function with a continuous inverse. The idea of a homeomorphism has to do with shapes and their generic properties. For example, consider a circle in the plane. The image of this circle under a homeomorphism

$$G : \mathbb{R}^2 \to \mathbb{R}^2$$

might be another circle, an ellipse, or even a square or a triangle. But it could not be a line segment. This would violate continuity. It also could not be a figure eight, because this would violate the property that G must have an inverse (so it must be one-to-one). There is a theorem that states that if the eigenvalues of A all have nonzero real parts, then there is a homeomorphism G that maps the phase portrait of the system $x' = Ax$ onto the phase portrait of $x' = F(x)$, with $G(0) = x_0$ (Hirsch and Smale (1974) p. 314). This theorem says that the phase portrait of $x' = F(x)$ around the point x_0 looks just like that of the linear system, except for some distortion. It would be as if we drew the phase portrait of the linear system on a sheet of rubber which we could stretch any way we like, but could not tear. This is a very powerful result. It means that we can get an actual picture (good enough for almost all practical purposes) of the behavior of a nonlinear dynamical system near each isolated equilibrium point just by analyzing its linear approximation. Then, to finish up the phase portrait on the rest of the state space, we combine what we have learned about the behavior of solutions near the equilibrium points with the information contained in a sketch of the vector field.

Step 3 is to formulate the model. We begin by considering the state space.

There are six state variables to begin with, but we can use Kirchoff's laws to reduce the number of degrees of freedom (the number of independent state variables) from six to two. Let $x_1 = i_R$ and notice that $x_1 = i_L = i_C$ as well. Let $x_2 = v_C$. Then we have

$$\frac{x_2'}{3} = x_1$$
$$v_R = x_1^3 + 4x_1$$
$$x_1' = v_L$$
$$x_2 + v_R + v_L = 0.$$

Substitute to obtain

$$\frac{x_2'}{3} = x_1$$
$$x_2 + x_1^3 + 4x_1 + x_1' = 0,$$

and then rearrange to get

$$x_1' = -x_1^3 - 4x_1 - x_2$$
$$x_2' = 3x_1. \tag{5.15}$$

Now if we let $x = (x_1, x_2)$, then Eq. (5.15) can be written in the form $x' = F(x)$ where $F = (f_1, f_2)$ and

$$f_1(x_1, x_2) = -x_1^3 - 4x_1 - x_2$$
$$f_2(x_1, x_2) = 3x_1. \tag{5.16}$$

This concludes step 3.

Step 4 is to solve the model. We will analyze the dynamical system (5.15) by sketching the complete phase portrait. Figure 5.8 shows a Maple graph of the vector field for this dynamical system. It is also a fairly simple matter to sketch the vector field by hand. Velocity vectors are horizontal on the curve $x_1 = 0$ where $x_2' = 0$, and vertical on the curve $x_2 = -x_1^3 - 4x_1$ where $x_1' = 0$. There is one equilibrium point $(0, 0)$ at the intersection of these two curves. From the vector field it is difficult to tell whether the equilibrium is stable or unstable. To obtain more information, we will analyze the linear system that approximates the behavior of (5.15) near the equilibrium $(0, 0)$.

Computing the partial derivatives from (5.16), we obtain

$$\frac{\partial f_1}{\partial x_1} = -3x_1^2 - 4$$
$$\frac{\partial f_1}{\partial x_2} = -1$$
$$\frac{\partial f_2}{\partial x_1} = 3 \tag{5.17}$$
$$\frac{\partial f_2}{\partial x_2} = 0.$$
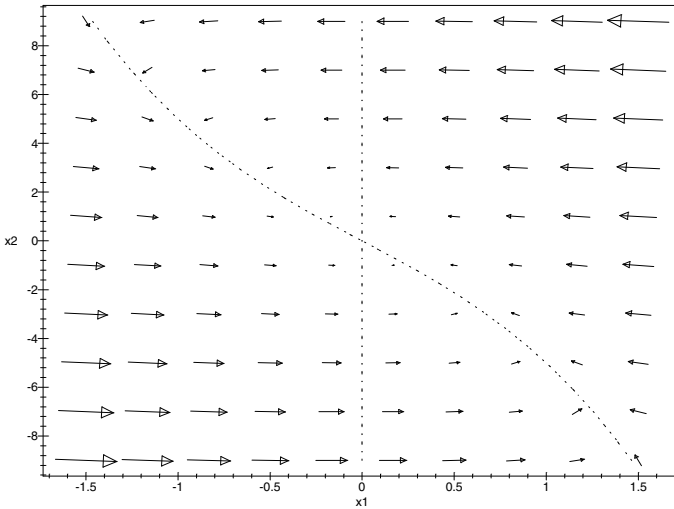
Figure 5.8: Graph of voltage x_2 versus current x_1 showing vector field (5.16) for the RLC circuit problem of Example 5.3.

Evaluating the partial derivatives (5.17) at the equilibrium point $(0,0)$ and substituting back into Eq. (5.1), we obtain

$$A = \begin{pmatrix} -4 & -1 \\ 3 & 0 \end{pmatrix}.$$

The eigenvalues of this 2×2 matrix can be computed as the roots of the equation

$$\begin{vmatrix} \lambda + 4 & 1 \\ -3 & \lambda \end{vmatrix} = 0.$$

Evaluating the determinant, we obtain the equation

$$\lambda^2 + 4\lambda + 3 = 0,$$

and then we obtain

$$\lambda = -3, -1.$$

Since both eigenvalues are negative, the equilibrium is stable.

To obtain additional information, we will solve the linear system $x' = Ax$. In this case we have

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} -4 & -1 \\ 3 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}. \tag{5.18}$$

We will solve the linear system (5.18) by the method of eigenvalues and eigenvectors. We have already calculated the eigenvalues $\lambda = -3, -1$. To compute the

eigenvector corresponding to the eigenvalue λ, we must find a nonzero solution to the equation

$$\begin{pmatrix} \lambda+4 & 1 \\ -3 & \lambda \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

For $\lambda = -3$ we have

$$\begin{pmatrix} 1 & 1 \\ -3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

from which we obtain

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

so that

$$\begin{pmatrix} -1 \\ 1 \end{pmatrix} e^{-3t}$$

is one solution to the linear system (5.18). For $\lambda = -1$ we have

$$\begin{pmatrix} 3 & 1 \\ -3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

from which we obtain

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix},$$

so that

$$\begin{pmatrix} -1 \\ 3 \end{pmatrix} e^{-t}$$

is another solution to the linear system (5.18). Then, the general solution to (5.18) can be written in the form

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = c_1 \begin{pmatrix} -1 \\ 1 \end{pmatrix} e^{-3t} + c_2 \begin{pmatrix} -1 \\ 3 \end{pmatrix} e^{-t} \tag{5.19}$$

where c_1, c_2 are arbitrary real constants.

Figure 5.9 shows the phase portrait for the linear system (5.18). This graph was obtained by plotting the solution curves (5.19) for a few select values of the constants c_1, c_2. For example, when $c_1 = 1$ and $c_2 = 1$, we plotted a parametric graph of

$$x_1(t) = -e^{-3t} - e^{-t}$$
$$x_2(t) = e^{-3t} + 3e^{-t}.$$

We superimposed a graph of the linear vector field in order to indicate the orientation of the solution curves. Whenever you plot a phase portrait, be sure to add arrows to indicate the direction of the flow.

Figure 5.10 shows the complete phase portrait for the original nonlinear dynamical system (5.15). This picture was obtained by combining the information in Figures 5.8 and 5.9 and using the fact that the phase portrait of the nonlinear
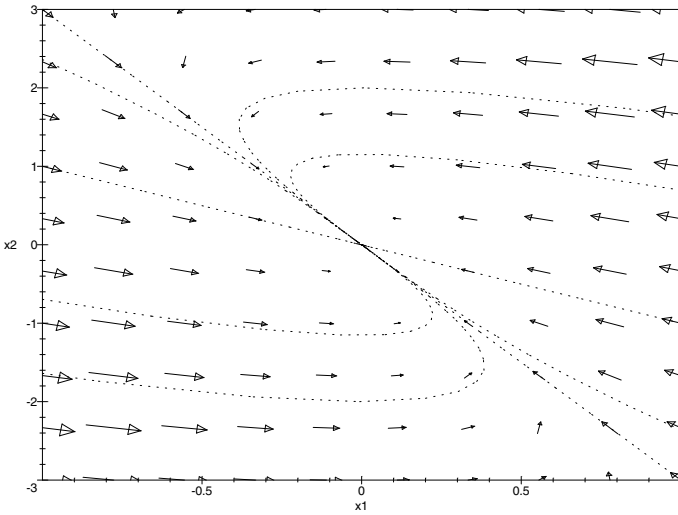
Figure 5.9: Graph of voltage x_2 versus current x_1 showing linear approximation to the phase portrait near $(0, 0)$ for the RLC circuit problem of Example 5.3.
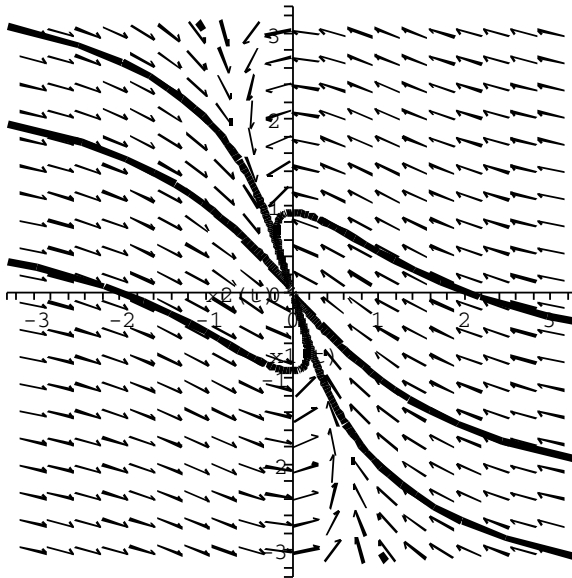

Figure 5.10: Graph of voltage x_2 versus current x_1 showing complete phase portrait for the RLC circuit problem of Example 5.3.

system (5.15) is homeomorphic to the phase portrait of the linear system (5.18). In this example there is not much qualitative difference between the behavior of the linear and the nonlinear systems.

Step 5 is to answer the question. The question was to describe the behavior of the RLC circuit. The overall behavior can be described in terms of two quantities, the current through the resistor and the voltage drop across the capacitor. Regardless of the initial state of the circuit, both quantities eventually tend to zero. Furthermore, it is eventually true that either voltage is positive and current is negative, or vice versa. For a complete graphical description of the way that current and voltage behave over time, see Figure 5.10, where x_1 represents current and x_2 represents voltage. The behavior of other quantities of interest can easily be described in terms of these two variables (see Figure 5.7 for details). For example, the variable x_1 actually represents the current through any branch of the circuit loop.

Next, we will perform a sensitivity analysis to determine the effect of small changes in our assumptions on our general conclusions. First let us consider the capacitance C. In our example we assumed that $C = 1/3$. Now we will generalize our model by letting C remain indeterminate. In this case we obtain the dynamical system

$$x_1' = -x_1^3 - 4x_1 - x_2$$
$$x_2' = \frac{x_1}{C} \tag{5.20}$$

in place of (5.15). Now we have

$$f_1(x_1, x_2) = -x_1^3 - 4x_1 - x_2$$
$$f_2(x_1, x_2) = \frac{x_1}{C}. \tag{5.21}$$

For values of C near $1/3$, the vector field for (5.21) is essentially the same as in Figure 5.8. Velocity vectors are still horizontal on the curve $x_1 = 0$ and vertical on the curve $x_2 = -x_1^3 - 4x_1$. There is still one equilibrium point $(0,0)$ at the intersection of these two curves.

Computing the partial derivatives from Eq. (5.21), we obtain

$$\frac{\partial f_1}{\partial x_1} = -3x_1^2 - 4$$
$$\frac{\partial f_1}{\partial x_2} = -1$$
$$\frac{\partial f_2}{\partial x_1} = \frac{1}{C} \tag{5.22}$$
$$\frac{\partial f_2}{\partial x_2} = 0.$$

Evaluating the partial derivatives (5.22) at the equilibrium point $(0,0)$ and

substituting back into Eq. (5.1), we obtain

$$A = \begin{pmatrix} -4 & -1 \\ 1/C & 0 \end{pmatrix}.$$

The eigenvalues of this matrix can be computed as the roots of the equation

$$\begin{vmatrix} \lambda+4 & 1 \\ -1/C & \lambda \end{vmatrix} = 0.$$

Evaluating the determinant, we obtain the equation

$$\lambda^2 + 4\lambda + \frac{1}{C} = 0.$$

The eigenvalues are

$$\lambda = -2 \pm \sqrt{4 - \frac{1}{C}}.$$

If $C > 1/4$, then we have two distinct real negative eigenvalues, and so the equilibrium is stable. In this case, the general solution to the linear system is

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = c_1 \begin{pmatrix} -1 \\ 2+\alpha \end{pmatrix} e^{(-2+\alpha)t} + c_2 \begin{pmatrix} -1 \\ 2-\alpha \end{pmatrix} e^{(-2-\alpha)t} \qquad (5.23)$$

where $\alpha^2 = 4 - 1/C$. The phase portrait of the linear system is about the same as Figure 5.9, except that the slope of the straight line solutions varies with C. Then for values of C greater than $1/4$, the phase portrait for the original nonlinear system is a lot like the one shown in Figure 5.10. We conclude that our general conclusions about this RLC circuit are not sensitive to the exact value of C as long as $C > 1/4$. A similar result may be expected for the inductance L. Generally speaking, the important characteristics of our solution (e.g., eigenvectors) depend continuously on these parameters.

Next, we consider the question of robustness. We assumed that the RLC circuit had v-i characteristic $f(x) = x^3 + 4x$. Suppose more generally that $f(0) = 0$ and that f is strictly increasing. Now the dynamical system equations are

$$\begin{aligned} x_1' &= -f(x_1) - x_2 \\ x_2' &= 3x_1. \end{aligned} \qquad (5.24)$$

Now we have

$$\begin{aligned} f_1(x_1, x_2) &= -f(x_1) - x_2 \\ f_2(x_1, x_2) &= 3x_1. \end{aligned} \qquad (5.25)$$

Let $R = f'(0)$. The linear approximation uses

$$A = \begin{pmatrix} -R & -1 \\ 3 & 0 \end{pmatrix}$$

and so the eigenvalues are the roots to the equation

$$\begin{vmatrix} \lambda + R & 1 \\ 3 & \lambda \end{vmatrix} = 0.$$

We compute that

$$\lambda = \frac{-R \pm \sqrt{R^2 - 12}}{2}.$$

As long as $R > \sqrt{12}$, we have two distinct real negative eigenvalues, and the behavior of the linear system is as depicted in Figure 5.9. Furthermore the behavior of the original nonlinear system cannot be too different from Figure 5.10. We conclude that our model of this RLC circuit is robust with regard to our assumptions about the form of the v-i characteristic.

Example 5.4. Consider the nonlinear RLC circuit with $L = 1, C = 1$, and v-i characteristic $f(x) = x^3 - x$. Determine the behavior of this circuit over time.

The modeling process is, of course, the same as for the previous example. Letting $x_1 = i_R$ and $x_2 = v_C$, we obtain the dynamical system

$$\begin{aligned} x_1' &= x_1 - x_1^3 - x_2 \\ x_2' &= x_1. \end{aligned} \tag{5.26}$$

See Figure 5.11 for a plot of the vector field. The velocity vectors are vertical on the curve $x_2 = x_1 - x_1^3$ and horizontal on the x_2 axis. The only equilibrium is the origin $(0,0)$. It is hard to tell from the vector field whether or not the origin is a stable equilibrium.

The matrix of partial derivatives is

$$A = \begin{pmatrix} 1 - 3x_1^2 & -1 \\ 1 & 0 \end{pmatrix}.$$

Evaluate at $x_1 = 0, x_2 = 0$ to obtain the linear system

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

which approximates the behavior of our nonlinear system near the origin. To obtain the eigenvalues we must solve

$$\begin{vmatrix} \lambda - 1 & 1 \\ -1 & \lambda - 0 \end{vmatrix} = 0$$

or $\lambda^2 - \lambda + 1 = 0$. The eigenvalues are

$$\lambda = 1/2 \pm i\sqrt{3}/2.$$

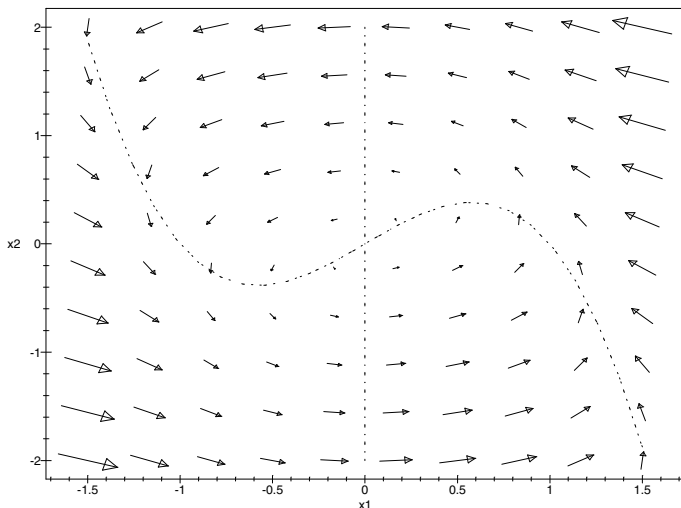Since the real part of every eigenvalue is positive, the origin is an unstable equilibrium.

Figure 5.11: Graph of voltage x_2 versus current x_1 showing vector field from (5.26) for the RLC circuit problem of Example 5.4.

To obtain more information, we will solve the linear system. To find an eigenvector belonging to

$$\lambda = 1/2 + i\sqrt{3}/2,$$

we solve

$$\begin{pmatrix} -1/2 + i\sqrt{3}/2 & 1 \\ -1 & 1/2 + i\sqrt{3}/2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

and we obtain

$$x_1 = 2, \ x_2 = 1 - i\sqrt{3}.$$

Then we have the complex solution

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 - i\sqrt{3} \end{pmatrix} e^{(\frac{1}{2} + i\frac{\sqrt{3}}{2})t}.$$

Taking real and imaginary parts yields two linearly independent real solutions $u = (x_1, x_2)$, where

$$x_1(t) = 2e^{t/2} \cos(t\sqrt{3}/2)$$
$$x_2(t) = e^{t/2} \cos(t\sqrt{3}/2) + \sqrt{3}\, e^{t/2} \sin(t\sqrt{3}/2)$$

and $v = (x_1, x_2)$, where

$$x_1(t) = 2e^{t/2} \sin(t\sqrt{3}/2)$$
$$x_2(t) = e^{t/2} \sin(t\sqrt{3}/2) - \sqrt{3}\, e^{t/2} \cos(t\sqrt{3}/2).$$
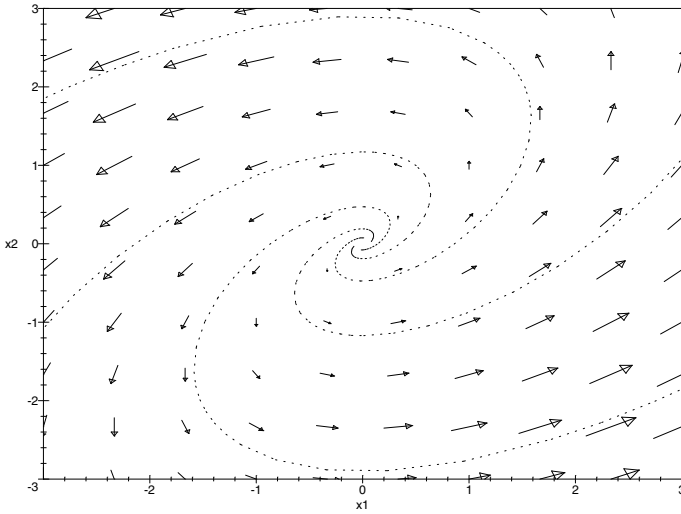
Figure 5.12: Graph of voltage x_2 versus current x_1 showing linear approximation to the phase portrait near $(0, 0)$ for the RLC circuit problem of Example 5.4.

The general solution is $c_1 u(t) + c_2 v(t)$. The phase portrait for this linear system is shown in Figure 5.12. This graph shows parametric plots of the solution for a few select values of c_1 and c_2. We superimpose a plot of the vector field in order to show the direction of the flow.

Note that, if we zoom in or zoom out on the origin in this linear phase portrait, it will look essentially the same. One of the defining characteristics of linear vector fields and linear phase portraits is that they look the same on every scale. The phase portrait for the nonlinear system in the neighborhood of the origin looks about the same, with some distortions. The solution curves near $(0,0)$ spiral outward, moving counterclockwise. If we continue to zoom in to the origin on a vector field or phase portrait for the nonlinear system, it will look more and more like the linear system. Further away from the origin, the behavior of the nonlinear system may vary significantly from that of the linear system.

In order to obtain the complete phase portrait of the nonlinear system, we need to combine the information from Figures 5.11 and 5.12. It is apparent from the vector field in Figure 5.11 that the behavior of solution curves changes dramatically farther away from the origin. There is still a general counterclockwise flow, but the solution curves do not spiral out to infinity as in the linear phase portrait. Solution curves that begin far away from the origin look like they are moving toward the origin as they continue their counterclockwise flow. Since the solution curves near the origin are spiraling outwards, and the solution

curves far away from the origin tend inward, and we know that solution curves do not cross, something interesting must be happening in the phase portrait. Whatever is happening, it is something that can never happen in a linear system. If a solution curve spirals outward in a linear phase portrait, then it must continue to spiral all the way out to infinity. In Section 6.3 we will explore the behavior of the dynamical system (5.26) using computational methods. We will wait until then to draw the complete phase portrait.

Before we leave the subject of linear approximation techniques, we should point out a few facts about discrete time dynamical systems. Suppose we have a discrete time dynamical system

$$\Delta x = F(x)$$

where $x = (x_1, \ldots, x_n)$, and let

$$G(x) = x + F(x)$$

denote the iteration function. At an equilibrium point x_0 we have $G(x_0) = x_0$. In Section 5.2 we used the approximation

$$G(x) \approx A(x - x_0)$$

for values of x near x_0, where A is the matrix of partial derivatives evaluated at $x = x_0$ as defined by (5.12).

One way to obtain a graphical picture of the iteration function $G(x)$ is to draw the image sets

$$G(S) = \{G(x) : x \in S\}$$

for various sets

$$S = \{x : |x - x_0| = r\}.$$

In dimension $n = 2$ the set S is a circle, and in dimension $n = 3$ it is a sphere. It is possible to show that, as long as the matrix A is nonsingular, there is a diffeomorphism $H(x)$ that maps the image sets $A(S)$ onto $G(S)$ in a neighborhood of the point x_0. If a point x lies inside of S, then $G(x)$ will be inside of $G(S)$. This allows a graphical interpretation of the dynamics.   Figures 5.13 through 5.15 illustrate the dynamics of the docking problem from Example 5.2. In this case, $G(S) = A(S)$, since G is linear. Starting at a state on (or inside) the set S, shown in Figure 5.13, the next state will be on (or inside, respectively) the set $A(S)$, shown in Figure 5.14, and then the next state will be on (or inside, respectively) the set $A^2(S) = A(A(S))$, shown in Figure 5.15. As $n \to \infty$, the set $A^n(S)$ gradually shrinks in toward the origin.
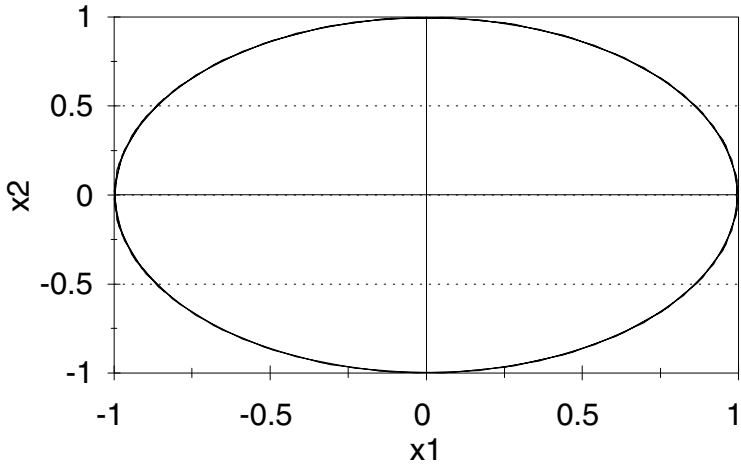
Figure 5.13: Dynamics of the docking problem showing the initial condition $S = \{(x_1, x_2) : x_1^2 + x_2^2 = 1\}$.
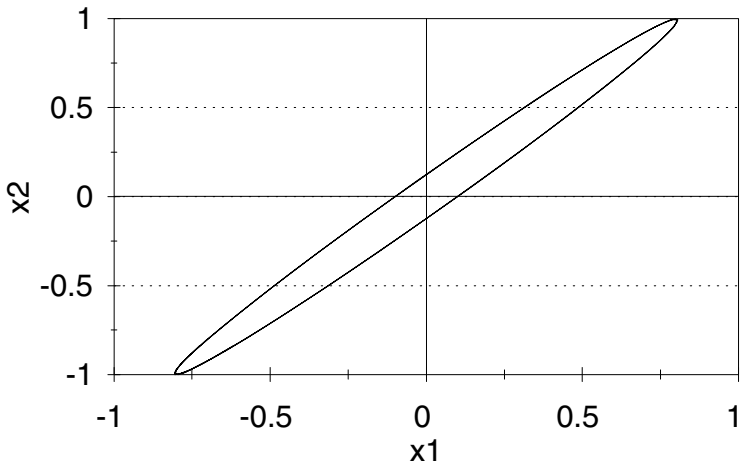

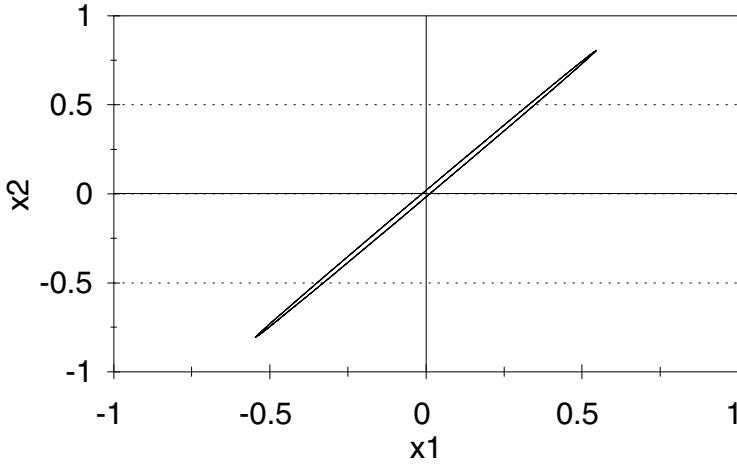Figure 5.14: Dynamics of the docking problem showing $A(S)$ after one iteration.

Figure 5.15: Dynamics of the docking problem showing $A^2(S)$ after two iterations.

5.4   Exercises

1. Reconsider Exercise 4 of Chapter 4.

   (a) Sketch the vector field for this model. Determine the location of each equilibrium in the state space. Can you tell from the vector field which of the equilibria are stable?

   (b) Use the eigenvalue method to test the stability of each equilibrium in the state space.

   (c) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

   (d) Sketch the complete phase portrait for this model, using the results of parts (a) and (c).

   (e) Given the current estimates of 5000 blue whales and 70,000 fin whales, what does this model predict about the future of the two species?

2. Reconsider Exercise 5 of Chapter 4.

   (a) Sketch the vector field for this model. Determine the location of each equilibrium in the state space.

(b) Use the eigenvalue method to test the stability of each equilibrium in the state space.

(c) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

(d) Sketch the complete phase portrait for this model, using the results of parts (a) and (c).

(e) Given the current estimates of 5000 blue whales and 70,000 fin whales, what does this model predict about the future of the two species?

3. Reconsider Exercise 6 of Chapter 4. Assume that the catchability coefficient $q = 10^{-5}$ and that the level of effort is $E = 3000$ boat-days per year.

(a) Sketch the vector field for this model. Determine the location of each equilibrium in the state space.

(b) Use the eigenvalue method to test the stability of each equilibrium in the state space.

(c) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

(d) Sketch the complete phase portrait for this model, using the results of parts (a) and (c).

(e) Given the current estimates of 5000 blue whales and 70,000 fin whales, what does this model predict about the future of the two species?

4. Repeat Exercise 3, but now assume a level of effort of $E = 6000$ boat-days per year.

5. Reconsider Exercise 7 of Chapter 4.

(a) Sketch the vector field for this model. Determine the location of each equilibrium in the state space.

(b) Use the eigenvalue method to test the stability of each equilibrium in the state space.

(c) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

(d) Sketch the complete phase portrait for this model, using the results of parts (a) and (c).

   (e) Suppose that an ecological disaster suddenly kills off 80% of the krill in the area, leaving 150,000 blue whales and only 100 tons per acre of krill. What does our model predict about the future of the whales and the krill?

6. Reconsider Exercise 9 of Chapter 4.

   (a) Use an eigenvalue method to determine the stability of the (P, Q) equilibrium, assuming a continuous time model.

   (b) Repeat part (a) assuming a discrete time model. To what do you attribute the difference in your results?

7. Reconsider the tree problem of Example 5.1. Assume $t = 1/2$.

   (a) Sketch the vector field for this model. Indicate the location of each equilibrium in the state space.

   (b) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

   (c) Sketch the complete phase portrait for this model, using the results of parts (a) and (b).

   (d) Suppose that a small number of hardwood trees is introduced into a mature stand of softwood trees. What does our model predict about the future of this forest?

8. Reconsider the tree problem of Example 5.1, but now suppose that the strength of the competition factor is too great to allow the coexistence of both hardwood and softwood trees. Assume $t = 3/4$.

   (a) Sketch the vector field for this model. Indicate the location of each equilibrium in the state space.

   (b) For each equilibrium point, determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the equilibrium point. Write the general solution to this linear system, and sketch the linear phase portrait.

   (c) Sketch the complete phase portrait for this model, using the results of parts (a) and (b).

   (d) Suppose that a small number of hardwood trees is introduced into a mature stand of softwood trees. What does our model predict about the future of this forest?

9. Reconsider the RLC circuit problem of Example 5.3 and perform a sensitivity analysis on the parameter L, which gives the inductance of the capacitor. Assume that $L > 0$.

   (a) Describe the vector field for the general case $L > 0$.

(b) Determine the range of L for which the equilibrium $(0,0)$ remains stable.

(c) Draw the phase portrait for the linear system in the case where there are two real eigenvalues.

(d) Use the results of parts (a) and (c) to draw the complete phase portrait. Comment on the sensitivity of our conclusions in Section 5.3 to the actual value of the parameter L.

10. Reconsider the RLC circuit problem of Example 5.3, but now suppose that the capacitance $C = 1/5$.

(a) Sketch the vector field for this model.

(b) Use the eigenvalue method to test the stability of the equilibrium at the origin.

(c) Determine the linear system that approximates the behavior of the original dynamical system in the neighborhood of the origin. Write the general solution to this linear system, and sketch the linear phase portrait.

(d) Sketch the complete phase portrait for this model, using the results of parts (a) and (c). How does the behavior of the RLC circuit change when the capacitance C is lowered?

11. (Continuation of Exercise 10) Reconsider the RLC circuit problem of Example 5.3, and now consider the effect of varying the capacitance C over the entire range $0 < C < \infty$.

(a) Sketch the phase portrait of the linear system that approximates the behavior of the RLC circuit in the neighborhood of the origin in the case $0 < C < 1/4$. Compare with the case $C > 1/4$, which was done in the text.

(b) Draw the complete phase portrait for the RLC circuit for the case $0 < C < 1/4$. Describe the changes that occur in the phase portrait as we transition between the two cases $0 < C < 1/4$ and $C > 1/4$.

(c) Draw the phase portrait of the linear system in the case $C = 1/4$ where there is only one eigenvalue. Sketch the phase portrait of the nonlinear system in this case. Explain how the phase portrait in this case represents an intermediate step between the case of two real distinct eigenvalues $(C > 1/4)$ and one pair of complex conjugate eigenvalues $(0 < C < 1/4)$.

(d) Reconsider the description of circuit behavior given in step 5 of Example 5.3 in the text. Describe in plain English the behavior of the RLC circuit in the more general case $C > 0$.

12. Reconsider the space docking problem of Example 5.2.

(a) Draw the vector field for this problem.

(b) Find the eigenvectors associated with the eigenvalues

$$\lambda = \frac{4 \pm \sqrt{6}}{10}$$

that were calculated in the text. Draw these eigenvectors on the picture from part (a). What do you notice about the vector field at these points?

(c) Calculate the rate at which closing velocity decreases (% per minute) if we start at an eigenvector.

(d) Generally speaking, what can you say about the rate of decrease in closing velocity for an arbitrary initial condition [Hint: Any (x_1, x_2) initial condition is a linear combination of the two eigenvectors found in part (b).]

13. (Continuation of Exercise 12) Reconsider the space docking problem of Example 5.2.

(a) As in Exercise 12, calculate the rate at which closing velocity decreases (% per minute) as a function of the control parameter k over the range of values $0 < k \leq 0.0268$.

(b) Find the value of k in part (a) that maximizes the rate of decrease in closing velocity.

(c) Explain the significance of your results in part (b) in terms of the efficiency of the control parameter k.

(d) Explain the problem with extending the approach used in this problem to find the most efficient value of k over the entire interval $0 < k < 0.2$ over which we have a stable control procedure.

14. Reconsider the space docking problem of Example 5.2, but now approximate the discrete time dynamical system

$$\Delta x_1 = -kw x_1 - kc x_2$$
$$\Delta x_2 = x_1 - x_2$$

by its continuous time analogue

$$\frac{dx_1}{dt} = -kw x_1 - kc x_2$$
$$\frac{dx_2}{dt} = x_1 - x_2$$

(a) Show that the continuous time model has a stable equilibrium at $(0, 0)$. Assume $w = 10, c = 5$, and $k = 0.02$.

(b) Solve the continuous model, using the method of eigenvalues and eigenvectors.

     (c) Draw the complete phase portrait for this model.

     (d) Comment on any differences between the behavior of the discrete and continuous models.

15. (Continuation of Exercise 14) Reconsider the docking problem of Example 5.2. As in Exercise 14, replace the discrete time model by its continuous time analogue.

     (a) Assume $w = 10$ and $c = 5$. For what values of k does the continuous model have a stable equilibrium at $(0, 0)$?

     (b) Solve the continuous model using the method of eigenvectors and eigenvalues.

     (c) Draw the complete phase portrait for this model. How does the phase portrait depend on k?

     (d) Comment on any differences between the continuous and discrete models.

16. Reconsider the tree problem of Example 5.1.

     (a) Can both types of trees coexist in equilibrium? Assume $b_i = a_i/2$. Use the five-step method, and model as a *discrete time* dynamical system with a time step of one year.

     (b) Use the eigenvalue test for discrete time dynamical systems to check the stability of the equilibrium you found in part (a).

     (c) Perform a sensitivity analysis on the parameter t, where $b_i = ta_i$. Determine the range of $0 < t < 0.6$ for which the equilibrium found in part (a) is stable.

     (d) Comment on any differences in results between the discrete time and continuous time models. As a practical matter, does it make any difference which we choose?

Further Reading

1. Beltrami, E. (1987) *Mathematics for Dynamic Modeling.* Academic Press, Orlando, Florida.

2. Frauenthal, J. (1979) *Introduction to Population Modeling.* UMAP Monograph.

3. Hirsch, M. and S. Smale (1974) *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York.

4. Keller, M., *Electrical Circuits and Applications of Matrix Methods: Analysis of Linear Circuits.* UMAP modules 108 and 112.

5. Rescigno, A. and I. Richardson, The Struggle for Life I, Two Species. *Bulletin of Mathematical Biophysics*, vol. 29, pp. 377-388.

6. Smale, S. (1972) On the Mathematical Foundations of Circuit Theory. *Journal of Differential Geometry*, vol. 7, pp. 193-210.

7. Wilde, C., *The Contraction Mapping Principle*. UMAP module 326.

Chapter 6

SIMULATION OF DYNAMIC MODELS

The technique of simulation has become the most important and popular method of analysis for dynamic models. Exact solution methods such as those taught in introductory courses in differential equations are of limited scope. The fact is that we do not know how to solve very many differential equations. The qualitative methods introduced in the two preceding chapters are more widely applicable, but for some problems we need a quantitative answer and a high degree of accuracy. Simulation methods provide both. Almost any dynamic model that will occur in practical applications can be simulated to a reasonable degree of accuracy. Furthermore, simulation techniques are very flexible. It is not hard to introduce more complex features such as time delays and stochastic elements, which are difficult to treat analytically.

The principal drawback of simulation comes in the area of sensitivity analysis. Without recourse to an analytic formula, the only way to test sensitivity to a particular parameter is to repeat the entire simulation for several values and then interpolate. This can be very expensive and time–consuming if there are several parameters to test. Even so, simulation is the method of choice for many problems. If we cannot solve analytically, and if we need quantitative solutions, then we really have no alternative but to simulate.

6.1   Introduction to Simulation

There are essentially two ways to approach the analysis of a dynamic system model. The analytic approach attempts to predict what will happen according to the model in a variety of circumstances. In the simulation approach, we build the model, turn it on, and find out.

Example 6.1. Two forces, which we will call red (R) and blue (B), are engaged in battle. In this conventional battle, attrition is due to direct fire (infantry) and

Variables: R = number of red divisions (units)
                      B = number of blue divisions (units)
                      D_R = red attrition rate due to direct fire (units/hour)
                      D_B = blue attrition rate due to direct fire (units/hour)
                      I_R = red attrition rate due to indirect fire (units/hour)
                      I_B = blue attrition rate due to indirect fire (units/hour)

Assumptions: $D_R = a_1 B$
                      $D_B = a_2 R$
                      $I_R = b_1 RB$
                      $I_B = b_2 RB$
                      $R \geq 0,\ B \geq 0$
                      $R(0) = 5,\ B(0) = 2$
                      a_1, a_2, b_1, b_2 are positive reals
                      $a_1 > a_2, b_1 > b_2$

Objective: Determine the conditions under which $R \to 0$
                      before $B \to 0$

Figure 6.1: Results of step 1 of the war problem.

area fire (artillery). The attrition rate due to direct fire is assumed proportional to the number of enemy infantry. The attrition rate due to artillery depends on both the amount of enemy artillery and the density of friendly troops. Red has amassed five divisions to attack a blue force of two divisions. Blue has the advantage of defense, and superior weapon effectiveness besides. How much more effective does blue have to be in order to prevail in battle?

We will use the five-step method. The results of step 1 are summarized in Figure 6.1. We have assumed that the attrition rate due to area fire is directly proportional to the product of enemy force level and friendly force level. It seems reasonable to assume at this stage that force level is proportional to force density. And since we have no information about the number of artillery versus infantry units, for this analysis we simply assume that artillery and infantry units are attrited in proportion to their numbers. Hence, the number of remaining artillery or infantry units on each side is assumed to remain in proportion to the total number of units.

Next is step 2. We will use a discrete–time dynamical system model, which we will solve by simulation. Figure 6.2 gives an algorithm for simulating a discrete–time dynamical system in two variables:

$$\Delta x_1 = f_1(x_1, x_2)$$
$$\Delta x_2 = f_2(x_1, x_2). \tag{6.1}$$

Step 3 is next. We will model the war problem as a discrete–time dynamical system with two state variables: $x_1 = R$, the number of red force units; and

Algorithm: DISCRETE–TIME SIMULATION

Variables: $x_1(n) =$ first state variable at time n
$x_2(n) =$ second state variable at time n
$N =$ number of time steps
Input: $x_1(0)$, $x_2(0)$, N

Process: Begin
for $n = 1$ to N do
   Begin
   $x_1(n) \leftarrow x_1(n-1) + f_1(x_1(n-1), x_2(n-1))$
   $x_2(n) \leftarrow x_2(n-1) + f_2(x_1(n-1), x_2(n-1))$
   End
End

Output: $x_1(1), \ldots, x_1(N)$
$x_2(1), \ldots, x_2(N)$

Figure 6.2: Pseudocode for discrete–time simulation.

$x_2 = B$, the number of blue force units. The difference equations are

$$\Delta x_1 = -a_1 x_2 - b_1 x_1 x_2$$
$$\Delta x_2 = -a_2 x_1 - b_2 x_1 x_2. \qquad (6.2)$$

We will start with $x_1(0) = 5$ and $x_2(0) = 2$ divisions of troops. We will use a time step of $\Delta t = 1$ hour. We will also need numerical values for a_i and b_i in order to run the simulation program. Unfortunately, we have been given no idea what they are supposed to be. We will have to make an educated guess. Suppose that a typical conventional battle lasts about 5 days and that engagement takes place about 12 hours per day. That means that one force is depleted in about 60 hours of battle. If a force were to be depleted by 5% per hour for 60 hours, the fraction remaining would be $(0.95)^{60} = 0.05$, which looks about right. We will assume that $a_2 = 0.05$. Since area fire is not generally as effective as direct fire in terms of attrition, we will assume $b_2 = 0.005$. (Recall that b_i is multiplied by both x_1 and x_2, which is why we made it so small.) Now blue is supposed to have greater weapon effectiveness than red, so we should have $a_1 > a_2$ and $b_1 > b_2$. Let us assume that $a_1 = \lambda a_2$ and $b_1 = \lambda b_2$ for some $\lambda > 1$. The analysis objective is to determine the smallest λ that will make $x_1 \to 0$ before $x_2 \to 0$. Now the difference equations are

$$\Delta x_1 = -\lambda(0.05)x_2 - \lambda(0.005)x_1 x_2$$
$$\Delta x_2 = -0.05 x_1 - 0.005 x_1 x_2. \qquad (6.3)$$

In step 4 we will solve the problem by running the simulation program for several values of λ. We begin by exercising the model for $\lambda = 1, 1.5, 2, 3,$ and 5.
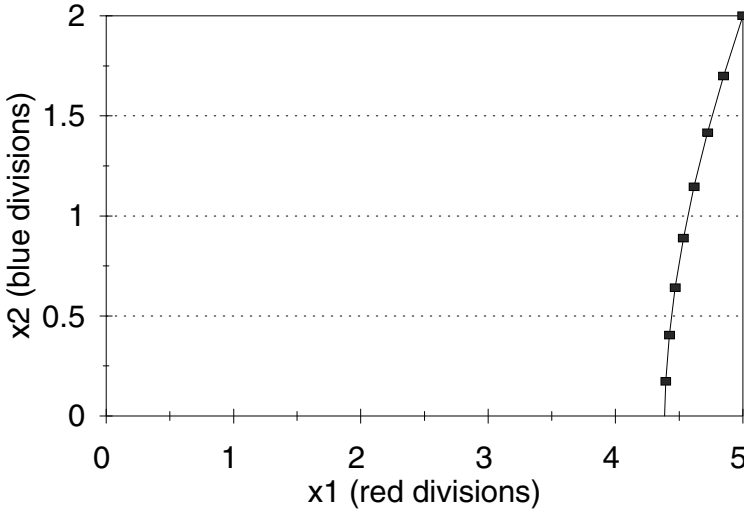
Figure 6.3: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 1.0$.

This should give us a good idea of how large λ needs to be, and it allows us to check our simulation against our intuitive grasp of the situation. For example, we should check that the larger λ is, the better blue does.

The results of this first batch of model runs are shown in Figures 6.3 through 6.7. A summary of our findings is contained in Table 6.1.

For each run we have recorded the value of λ, the duration of the battle, the identity of the winner, and the number of units remaining on the winning side. We decided to run the simulation for up to 14 days of combat (or $N = 168$ hours). The duration of battle is defined to be the number of hours of actual combat (there are 12 hours of fighting per day) until one of the variables x_1 or x_2 becomes zero or negative. If both sides survive 168 hours of combat, we call it a draw.

Advantage	Hours of	Winning	Remaining
(λ)	Combat	Side	Forces
1.0	8	red	4.4
1.5	9	red	4.1
2.0	9	red	3.7
3.0	10	red	3.0
5.0	17	red	1.0

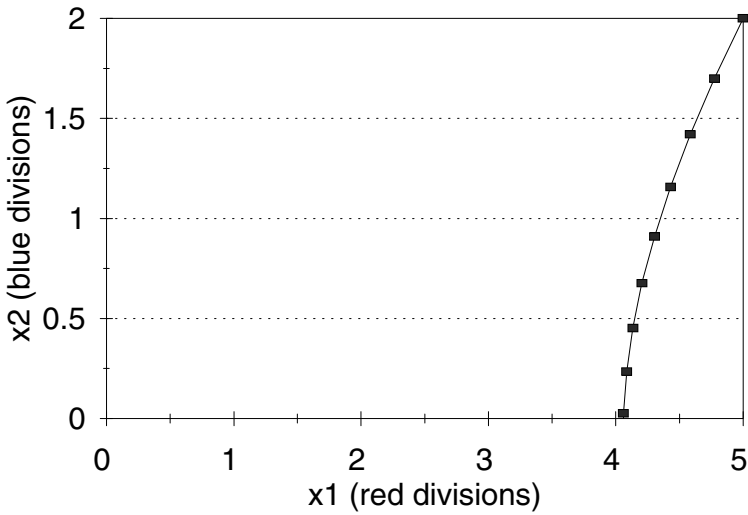Table 6.1: Summary of simulation results for the war problem.

Figure 6.4: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 1.5$.
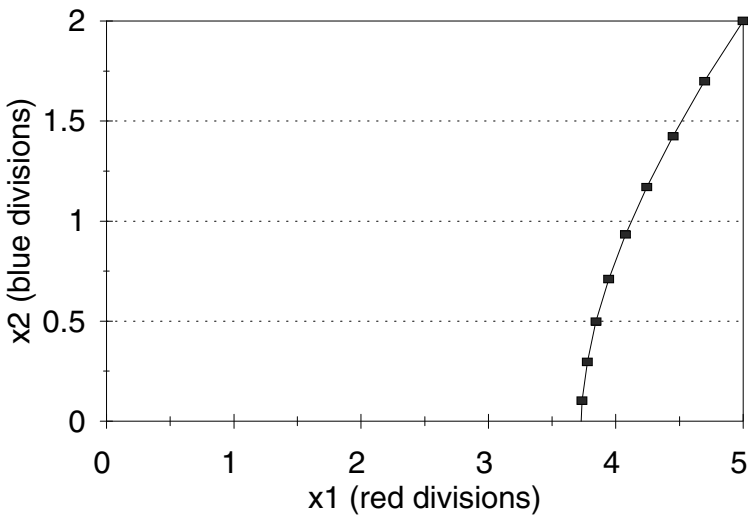

Figure 6.5: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 2.0$.
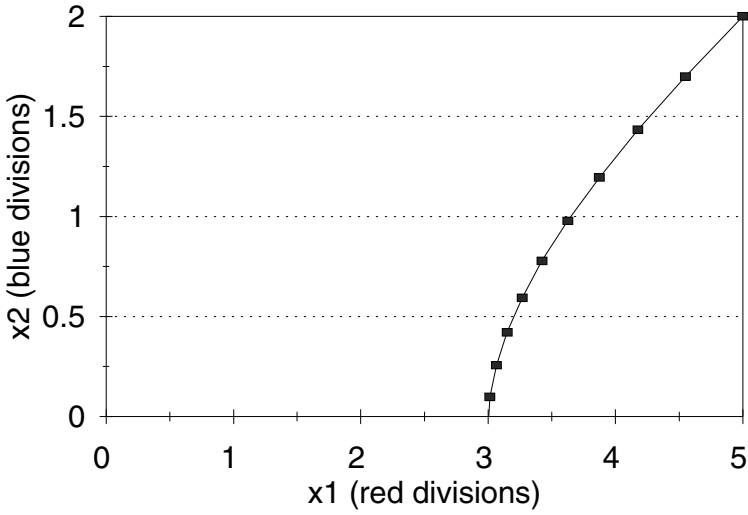
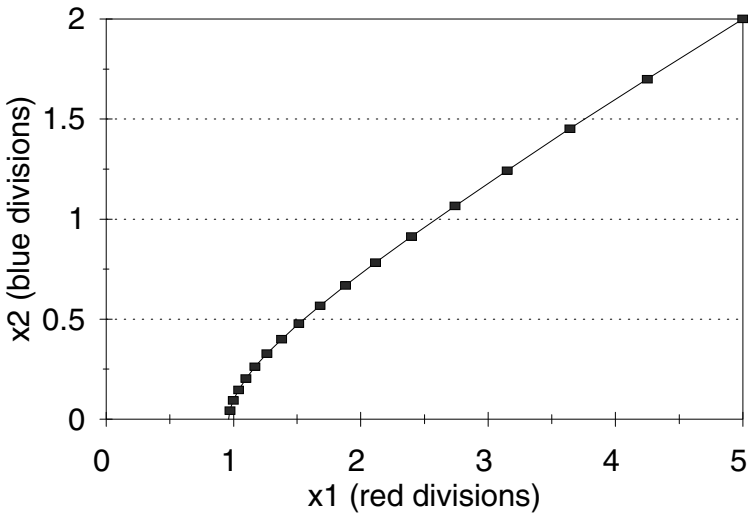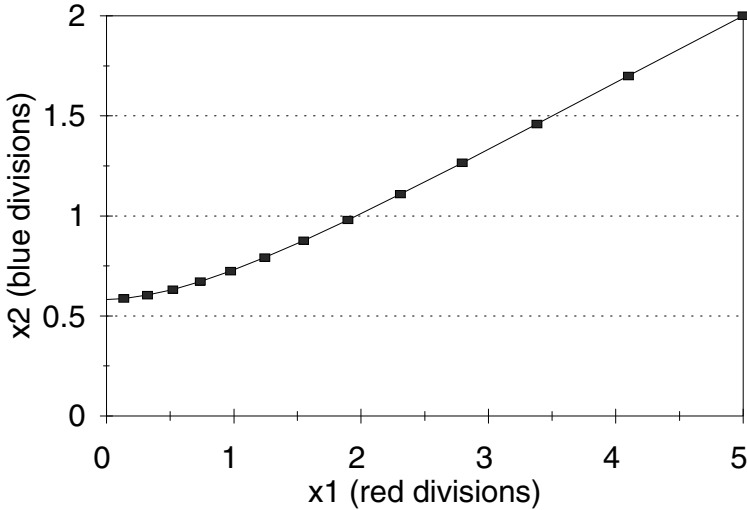Figure 6.6: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 3.0$.

Figure 6.7: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 5.0$.

Figure 6.8: Graph of blue divisions x_2 versus red divisions x_1 for the war problem: case $\lambda = 6.0$.

It does not look good for blue. Even with a $5 : 1$ edge in weapon effectiveness, blue will lose the battle. We decided to make a few more model runs to find out just how big λ would have to be for blue to win. At $\lambda = 6.0$, the blue side won after 13 hours of battle, with 0.6 units remaining (see Figure 6.8). A few more model runs, bisecting the interval $5.0 \leq \lambda \leq 6.0$, yielded a lower bound of $\lambda = 5.4$ for blue to win. At $\lambda = 5.3$, red was the winner.

Finally, we need to summarize our results. We simulated an engagement between an attacking red force of five divisions and a defending blue force of two divisions. We assumed that the two forces would engage and continue to fight until there emerged a clear victor. We wanted to investigate the extent to which a greater weapon effectiveness (kill rate) could offset a $5 : 2$ numerical disadvantage. We simulated a number of battles for different ratios of weapon effectiveness. We found that blue would need at least a $5.4 : 1$ advantage in weapon effectiveness to defend successfully against a numerically superior red force of 5 divisions.

Having finished the five-step method and answered the question stated in step 1, we need to perform a sensitivity analysis. This is particularly important in a problem such as this one, where most of our data came from sheer guesswork. We will begin by investigating the relationship between the magnitude of the attrition coefficients and the outcome of the battle. We had assumed that $a_2 = 0.05$, $b_2 = a_2/10$, $a_1 = \lambda a_2$, and $b_1 = \lambda b_2$. We will now vary a_2, keeping the same relative relationship between it and the other variables.

We will investigate the dependence of $\lambda_{\min}$ on a_2, where $\lambda_{\min}$ is defined to

Force ratio (red : blue)	Advantage required ($\lambda_{\min}$)
8 : 2	11.8
7 : 2	9.5
6 : 2	7.3
5 : 2	5.4
4 : 2	3.6
3 : 2	2.2
2 : 2	1.1

Table 6.2: Summary of simulation results showing the effect of force ratio for the war problem.

be the smallest value of λ for which blue wins. This requires making a number of model runs for each value of a_2. It turns out that there is no need to tabulate these results, because in every case we checked (from $a_2 = 0.01$ to $a_2 = 0.10$), we found $\lambda_{\min} = 5.4$, the same as our baseline case ($a_2 = 0.05$). There is apparently no sensitivity to the magnitude of the attrition coefficients.

Several more kinds of sensitivity analysis are possible, and the process probably should go on as long as time permits, curiosity persists, and the pressures of other obligations do not intrude. We were curious about the relationship between $\lambda_{\min}$ and the numerical superiority ratio of red versus blue, currently assumed to be 5 : 2. To study this, we returned to the baseline case, $a_2 = 0.05$, and made several model runs to determine $\lambda_{\min}$ for several values of the initial red force strength x_1, keeping $x_2 = 2$ fixed. The results of this model excursion are tabulated in Table 6.2. The case $x_1 = 2$ was run as a check. We found $\lambda_{\min} = 1.1$ in this case, because the case $\lambda = 1$ produced a draw.

6.2   Continuous–Time Models

In this section we discuss the fundamentals of simulating continuous–time dynamical systems. The methods presented here are simple and usually effective. The basic idea is to use the approximation

$$\frac{dx}{dt} \approx \frac{\Delta x}{\Delta t}$$

to replace our continuous–time model (differential equations) by a discrete–time model (difference equations). Then we can use the simulation methods we introduced in the preceding section.

Example 6.2. Reconsider the whale problem of Example 4.2. We know now that starting at the current population levels of $B = 5,000$, $F = 70,000$, and assuming a competition coefficient of $\alpha < 1.25 \times 10^{-7}$, both populations of whales will eventually grow back to their natural levels in the absence of any further harvesting. How long will this take?

We will use the five-step method. Step 1 is the same as before (see Fig. 4.3), except that now the objective is to determine how long it takes to get to the equilibrium starting from $B = 5,000$, $F = 70,000$.

Step 2 is to select the modeling approach. We have an analysis question that seems to require a quantitative method. The graphical methods of Chapter 4 tell us what will happen, but not how long it will take. The analytical methods reviewed in Chapter 5 are local in nature. We need a global method here. The best thing would be to solve the differential equations, but we don't know how. We will use a simulation; this seems to be the only choice we have.

There is some question as to whether we want to adopt a discrete–time or a continuous–time model. Let us consider, more generally, the case of a dynamic model in n variables, $x = (x_1, \ldots, x_n)$, where we are given the rates of change $F = (f_1, \ldots, f_n)$ for each of the variables $x_1, \ldots, x_n$, but we have not yet decided whether to model the system in discrete–time or continuous–time. The discrete–time model looks like

$$\Delta x_1 = f_1(x_1, \ldots, x_n)$$
$$\vdots \tag{6.4}$$
$$\Delta x_n = f_n(x_1, \ldots, x_n),$$

where Δx_i represents the change in x_i over 1 unit of time ($\Delta t = 1$). The units of time are already specified. The method for simulating such a system was discussed in the previous section.

If we decided on a continuous–time model instead, we would have

$$\frac{dx_1}{dt} = f_1(x_1, \ldots, x_n)$$
$$\vdots \tag{6.5}$$
$$\frac{dx_n}{dt} = f_n(x_1, \ldots, x_n),$$

which we would still need to figure out how to simulate. We certainly can't expect the computer to calculate $x(t)$ for every value of t. That would take an infinite amount of time to get nowhere. Instead we must calculate $x(t)$ at a finite number of points in time. In other words, we must replace the continuous–time model by a discrete–time model in order to simulate it. What would the discrete–time approximation to this continuous–time model look like? If we use a time step of $\Delta t = 1$ unit, it will be exactly the same as the discrete–time model we could have chosen in the first place. Hence, unless there is something wrong with choosing $\Delta t = 1$, we don't have to choose between discrete and continuous. Then we are done with step 2.

Step 3 is to formulate the model. As in Chapter 4, we let $x_1 = B$ and $x_2 = F$ represent the population levels of each species. The dynamical system equations

are

$$\frac{dx_1}{dt} = 0.05x_1\left(1 - \frac{x_1}{150,000}\right) - \alpha x_1 x_2$$
$$\frac{dx_2}{dt} = 0.08x_2\left(1 - \frac{x_2}{400,000}\right) - \alpha x_1 x_2$$

(6.6)

on the state space $x_1 \geq 0,\ x_2 \geq 0$. In order to simulate this model we will begin by transforming to a set of difference equations

$$\Delta x_1 = 0.05x_1\left(1 - \frac{x_1}{150,000}\right) - \alpha x_1 x_2$$
$$\Delta x_2 = 0.08x_2\left(1 - \frac{x_2}{400,000}\right) - \alpha x_1 x_2$$

(6.7)

over the same state space. Here, Δx_i represents the change in population x_i over a period of $\Delta t = 1$ year. We will have to supply a value for α in order to run the program. We will assume that $\alpha = 10^{-7}$ to start with. Later on, we will do a sensitivity analysis on α.

Step 4 is to solve the problem by simulating the system in Eq. (6.7) using a computer implementation of the algorithm in Fig. 6.2. We began by simulating $N = 20$ years, starting with

$$x_1(0) = 5,000$$
$$x_2(0) = 70,000.$$

Figures 6.9 and 6.10 show the results of our first model run. Both blue whale and fin whale populations grow steadily, but in 20 years they do not get close to the equilibrium values

$$x_1 = 35,294$$
$$x_2 = 382,352$$

predicted by our analysis back in Chapter 4.

Figures 6.11 and 6.12 show our simulation results when we input a value of N large enough to allow this discrete–time dynamical system to approach equilibrium.

Step 5 is to put our conclusions into plain English. It takes a long time for the whale populations to grow back: about 100 years for the fin whale, and several centuries for the more severely depleted blue whale.

We will now discuss the sensitivity of our results to the parameter α, which measures the intensity of competition between the two species. Figures 6.13 through 6.18 show the results of our simulation runs for several values of α. Of course, the equilibrium levels of both species change along with α.

However, the time it takes our model to converge to equilibrium changes very little. Our general conclusion is valid whatever the extent of competition: It will take centuries for the whales to grow back.
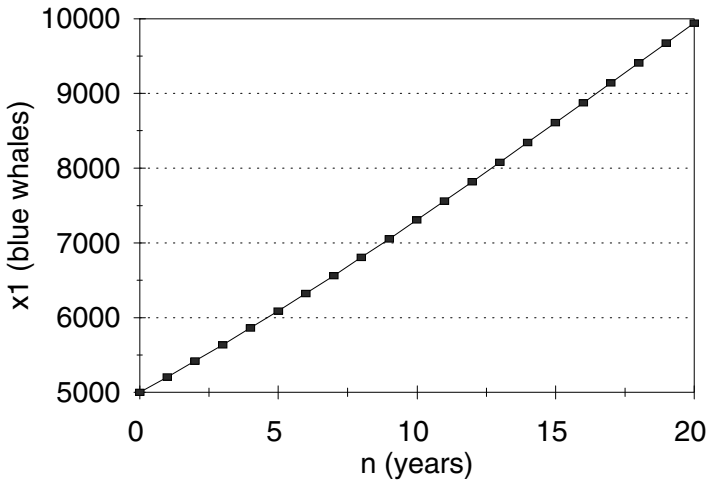
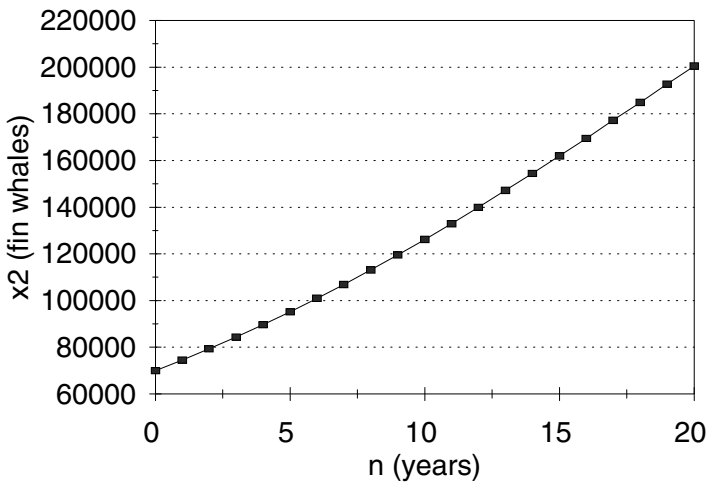Figure 6.9: Graph of blue whales x_1 versus time n for the whale problem: case $\alpha = 10^{-7}, N = 20$.

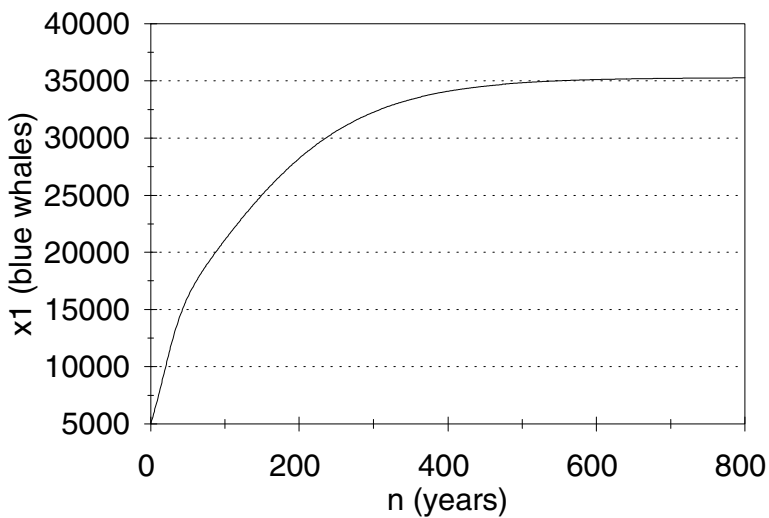Figure 6.10: Graph of fin whales x_2 versus time n for the whale problem: case $\alpha = 10^{-7}, N = 20$.

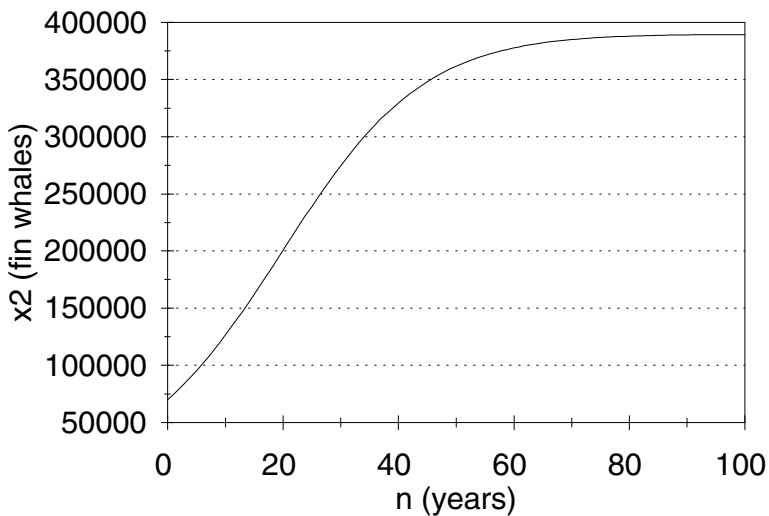Figure 6.11: Graph of blue whales x_1 versus time n for the whale problem: case $\alpha = 10^{-7}, N = 800$.

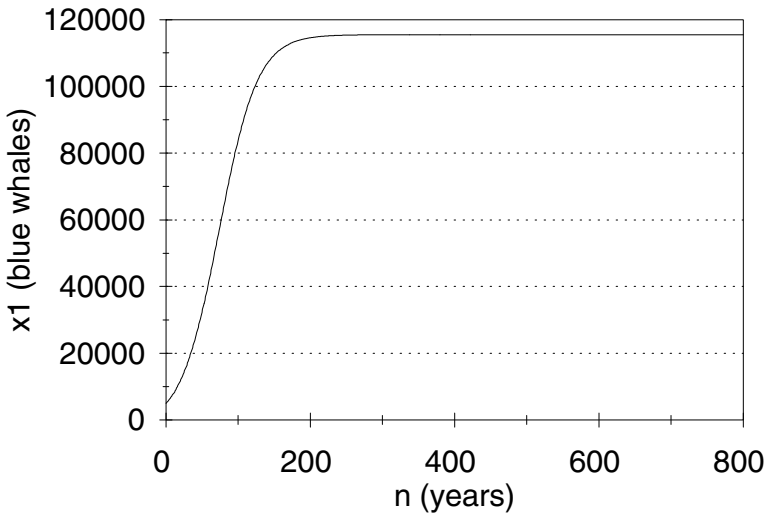Figure 6.12: Graph of fin whales x_2 versus time n for the whale problem: case $\alpha = 10^{-7}, N = 100$.

Figure 6.13: Graph of blue whales x_1 versus time n for the whale problem: case $\alpha = 3 \times 10^{-8}, N = 800$.
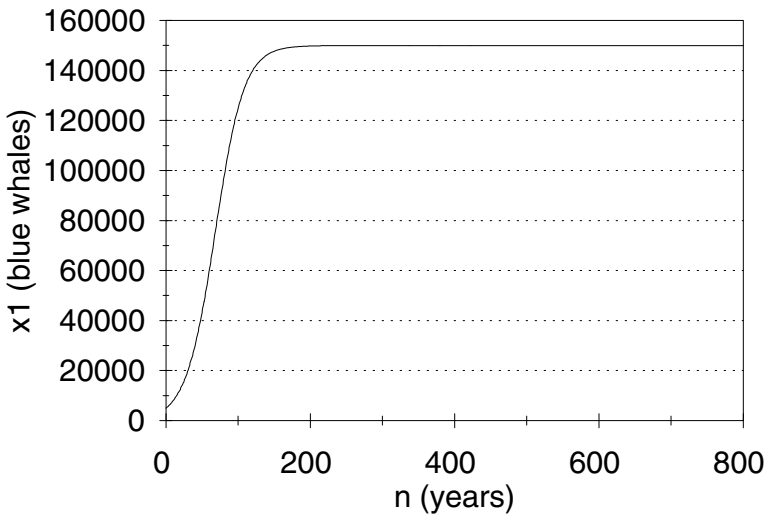

Figure 6.14: Graph of blue whales x_1 versus time n for the whale problem: case $\alpha = 10^{-8}, N = 800$.

Figure 6.15: Graph of blue whales x_1 versus time n for the whale problem: case $\alpha = 10^{-9}, N = 800$.

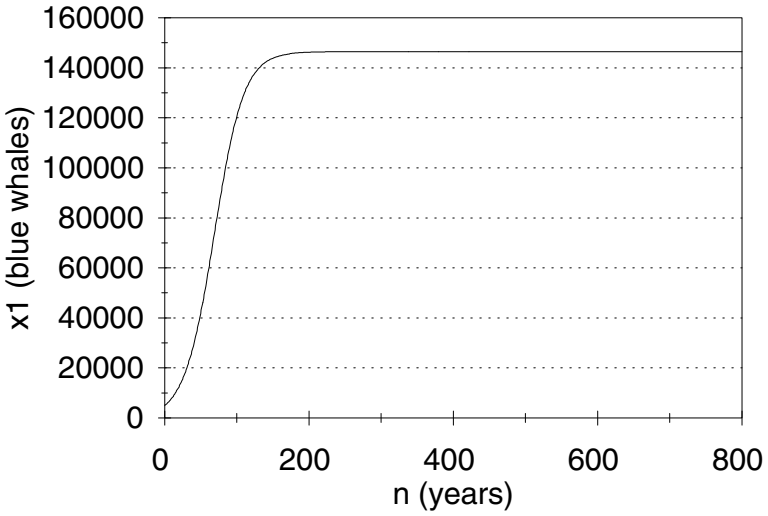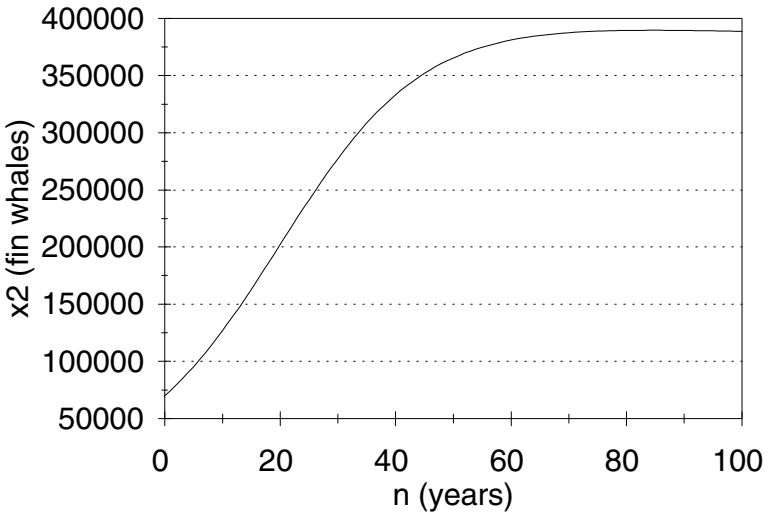Figure 6.16: Graph of fin whales x_2 versus time n for the whale problem: case $\alpha = 3 \times 10^{-8}, N = 100$.
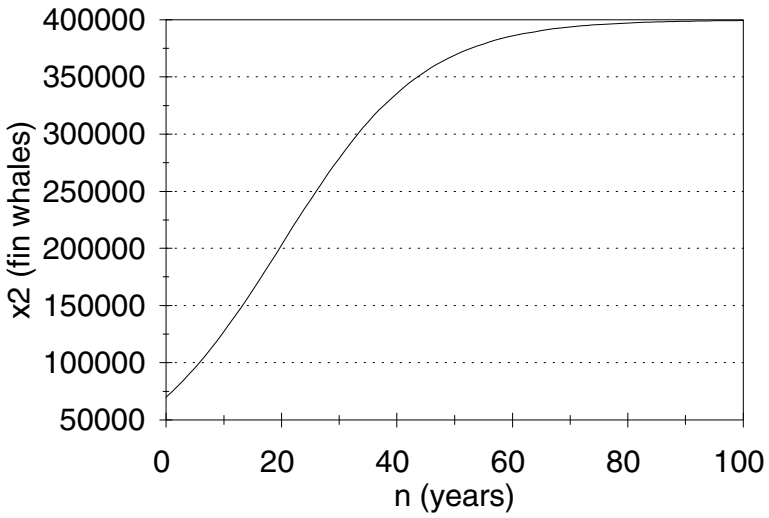
Figure 6.17: Graph of fin whales x_2 versus time n for the whale problem: case $\alpha = 10^{-8}, N = 100$.
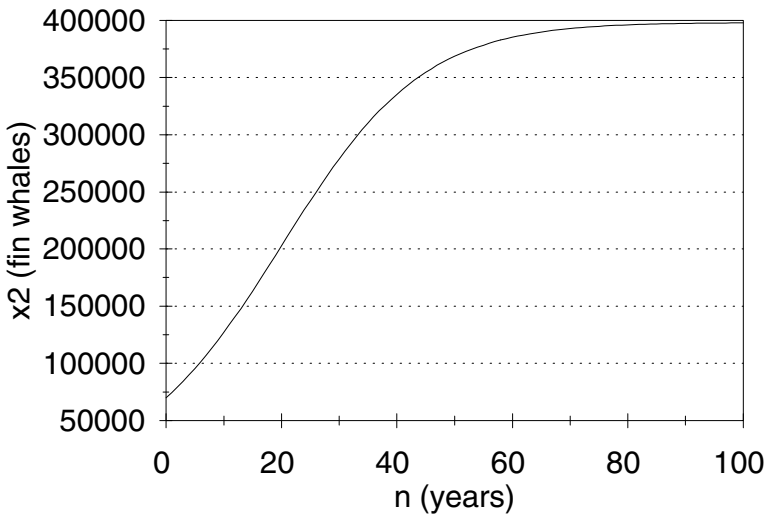

Figure 6.18: Graph of fin whales x_2 versus time n for the whale problem: case $\alpha = 10^{-9}, N = 100$.

6.3   The Euler Method

One of the reasons we simulate dynamic models is to obtain accurate quanti-
tative information about system behavior. For some applications the simple
simulation techniques of the previous section are too imprecise. More sophisti-
cated numerical analysis techniques are available, however, that can be used to
provide accurate solutions to initial value problems for almost any differential
equation model. In this section we present the simplest generally useful method
for solving systems of differential equations to any desired degree of accuracy.

Example 6.3. Reconsider the RLC circuit problem of Example 5.4 in the
previous chapter. Describe the behavior of this circuit.

Our analysis in Section 5.3 was successful only in determining the local
behavior of the dynamical system

$$
\begin{aligned}
x_1' &= x_1 - x_1^3 - x_2 \\
x_2' &= x_1
\end{aligned}
\tag{6.8}
$$

in the neighborhood of $(0, 0)$, which is the only equilibrium of this system. The
equilibrium is unstable, with nearby solution curves spiraling counterclockwise
and outward. A sketch of the vector field (see Fig. 5.11) reveals little new
information. There is a general counterclockwise rotation to the flow, but it is
hard to tell whether solution curves spiral inward, outward, or neither, in the
absence of additional information.

We will use the Euler method to simulate the dynamical system in Eq. (6.8).
Figure 6.19 gives an algorithm for the Euler method. Consider a continuous–
time dynamical system model

$$
x' = F(x)
$$

with $x = (x_1, \ldots, x_n)$ and $F = (f_1, \ldots, f_n)$, along with the initial condition
$x(t_0) = x_0$.

Starting from this initial condition, at each iteration the Euler method pro-
duces an estimate of $x(t + h)$ based on the current estimate of $x(t)$, using the
fact that

$$
x(t + h) - x(t) \approx h \, F(x(t)).
$$

The accuracy of the Euler method increases as the step size h becomes smaller;
i.e., as the number of steps N becomes larger. For small h the error in the
estimate $x(N)$ of the final value of the state variable x is roughly proportional
to h. In other words, using twice as many steps (i.e., reducing h by half)
produces results twice as accurate.

Figures 6.20 and 6.21 illustrate the results obtained by applying a computer
implementation of the Euler method to Eq. (6.8). Each graph in Figs. 6.20 and
6.21 is the result of several simulation runs. For each set of initial conditions,
we need to perform a sensitivity analysis on the input parameters T and N.

First, we enlarged T until any further enlargements produced essentially
the same picture (the solution just cycled around a few more times). Then

Algorithm: THE EULER METHOD

Variables: $t(n) =$ time after n steps
                  $x_1(n) =$ first state variable at time $t(n)$
                  $x_2(n) =$ second state variable at time $t(n)$
                  $N =$ number of steps
                  $T =$ time to end simulation

Input: $t(0)$, $x_1(0)$, $x_2(0)$, N, T

Process: Begin
                  $h \leftarrow (T - t(0))/N$
                  for $n = 0$ to $N - 1$ do
                  Begin
                      $x_1(n + 1) \leftarrow x_1(n) + h f_1(x_1(n), x_2(n))$
                      $x_2(n + 1) \leftarrow x_2(n) + h f_2(x_1(n), x_2(n))$
                      $t(n + 1) \leftarrow t(n) + h$
                  End
                  End

Output: $t(1)$, $\ldots$, $t(n)$
                  $x_1(1)$, $\ldots$, $x_1(N)$
                  $x_2(1)$, $\ldots$, $x_2(N)$

Figure 6.19: Pseudocode for the Euler method.

we enlarged N (i.e., decreased the step size) to check accuracy. If doubling N produced a graph that was indistinguishable from the one before, we judged that N was large enough for our purposes.

In Fig. 6.20 we started at $x_1(0) = -1$, $x_2(0) = -1.5$. The resulting solution curve spirals in toward the origin, with a counterclockwise rotation. However, before it gets too close to the origin, the solution settles into a more–or–less periodic behavior, cycling around the origin. When we start nearer the origin in Fig. 6.21, the same behavior occurs, except now the solution curve spirals outward. In both cases the solution approaches the same closed loop around the origin. This closed loop is called a *limit cycle*.

Figure 6.22 shows the complete phase portrait for this dynamical system. For any initial condition except $(x_1, x_2) = (0, 0)$, the solution curve tends to the same limit cycle. If we begin inside the loop, the curve spirals outward; if we begin outside the loop, the curve moves inward. The kind of behavior we see in Fig. 6.22 is a phenomenon that cannot occur in a linear dynamical system. If a solution to a linear dynamical system spirals in toward the origin, it must spiral all the way into the origin. If it spirals outward, then it spirals all the way out to infinity. This observation has modeling implications, of course. Any dynamical system exhibiting the kind of behavior shown in Fig. 6.22 cannot be
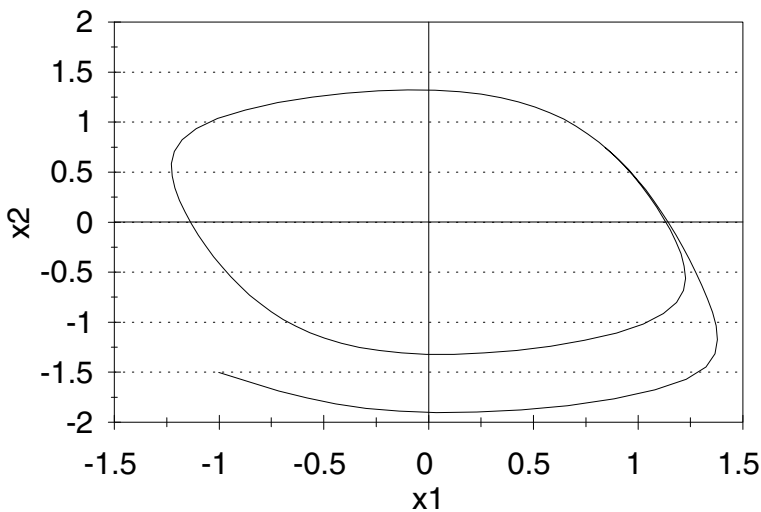
Figure 6.20: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = -1.0$, $x_2(0) = -1.5$.
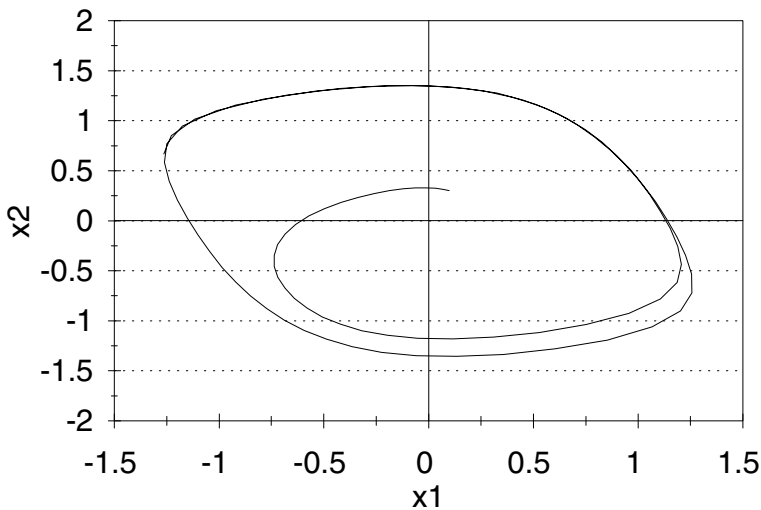

Figure 6.21: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = 0.1$, $x_2(0) = 0.3$.
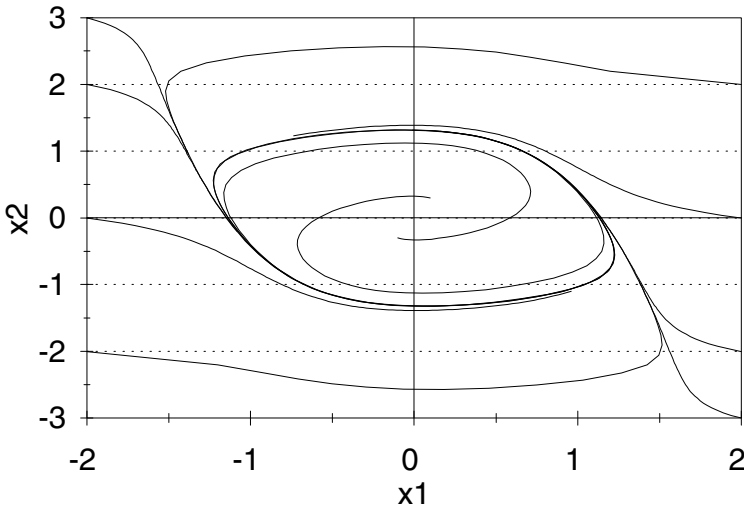
Figure 6.22: Graph of voltage x_2 versus current x_1 showing the complete phase portrait for the nonlinear RLC circuit problem of Example 6.3.

modeled adequately using linear differential equations.

The graphs in Figures 6.20–6.22 were produced using a spreadsheet implementation of the Euler method. The advantage of a spreadsheet implementation is that the computations and graphics are both performed on the same platform, and the results of changing initial conditions can be observed instantly. A simple computer program to implement this algorithm is effective, but the output is harder to interpret without graphics. Many graphing calculators and computer algebra systems also have built–in differential equation solvers, most of which are based on some variation of the Euler method. The Runge–Kutta method is one variation that uses a more sophisticated interpolation between $x(t)$ and $x(t + h)$; see Exercise 21 at the end of this chapter. No matter what kind of numerical method you use to solve differential equations, be sure to check your results by performing a sensitivity analysis on the parameters that control accuracy. Even the most sophisticated algorithms can produce serious errors unless they are used with care.

Next, we will perform a sensitivity analysis to determine the effect of small changes in our assumptions on our general conclusions. Here we will discuss the sensitivity to the capacitance C. Some additional questions of sensitivity and robustness are relegated to the exercises at the end of this chapter. In our example we assumed that $C = 1$. In the more general case we obtain the

dynamical system

$$x_1' = x_1 - x_1^3 - x_2$$
$$x_2' = \frac{x_1}{C}. \tag{6.9}$$

For any value of $C > 0$, the vector field is essentially the same as in Fig. 5.11. The velocity vectors are vertical on the curve $x_2 = x_1 - x_1^3$ and horizontal on the x_2 axis. The only equilibrium is the origin, $(0, 0)$.

The matrix of partial derivatives is

$$A = \begin{pmatrix} 1 - 3x_1^2 & -1 \\ 1/C & 0 \end{pmatrix}.$$

Evaluate at $x_1 = 0$, $x_2 = 0$ to obtain the linear system

$$\begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1/C & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{6.10}$$

which approximates the behavior of our nonlinear system near the origin. To obtain the eigenvalues, we must solve

$$\begin{vmatrix} \lambda - 1 & 1 \\ -1/C & \lambda - 0 \end{vmatrix} = 0,$$

or $\lambda^2 - \lambda + 1/C = 0$. The eigenvalues are

$$\lambda = \frac{1 \pm \sqrt{1 - \dfrac{4}{C}}}{2}. \tag{6.11}$$

As long as $0 < C < 4$, the quantity under the radical is negative, so we have two complex conjugate eigenvalues with positive real parts, making the origin an unstable equilibrium.

Next, we need to consider the phase portrait for the linear system. It is possible to solve the system in Eq. (6.10) in general by using the method of eigenvalues and eigenvectors, although it would be rather messy. Fortunately, in the present case it is not really necessary to determine a formula for the exact analytical solution to Eq. (6.10) in order to draw the phase portrait. We already know that the eigenvalues of this system are of the form $\lambda = a \pm ib$, where a is positive. As we mentioned previously (in Section 5.1, during the discussion of step 2 for Example 5.1), this implies that the coordinates of any solution curve must be linear combinations of the two terms $e^{at} \cos(bt)$ and $e^{at} \sin(bt)$. In other words, every solution curve spirals outward. A cursory examination of the vector field for Eq. (6.10) tells us that the spirals must rotate counterclockwise. We thus see that for any $0 < C < 4$, the phase portrait of the linear system in Eq. (6.10) looks much like the one in Fig. 5.10.

Our examination of the linear system in Eq. (6.10) shows that the behavior of the nonlinear system in the neighborhood of the origin must be essentially
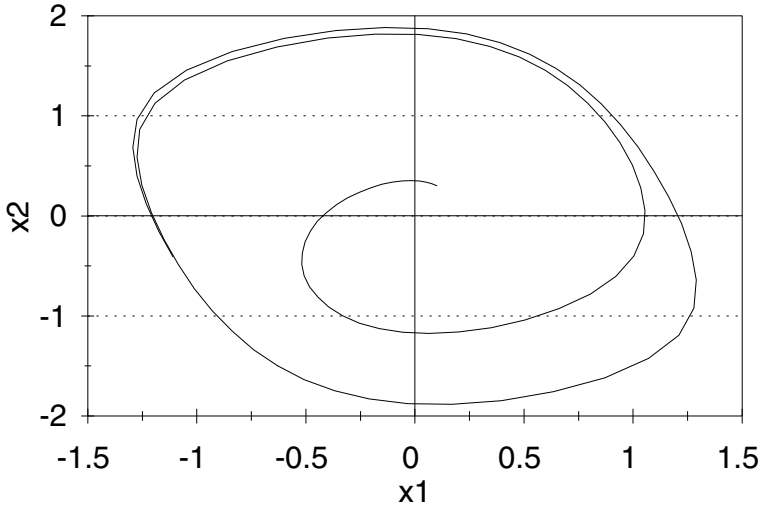
Figure 6.23: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = 0.1$, $x_2(0) = 0.3$, $C = 0.5$.

the same as in Fig. 6.22 for any value of C near the baseline case $C = 1$. To see what happens farther away from the origin, we need to simulate. Figures 6.23 through 6.26 show the results of simulating the dynamical system in Eq. (6.9) using the Euler method for several different values of C near 1. In each simulation run we started at the same initial condition as in Fig. 6.21.

In each case the solution curve spirals outward and is gradually attracted to a limit cycle. The limit cycle gets smaller as C increases. Several different initial conditions were used for each value of C tested (additional simulation runs are not shown). In each case, apparently, a single limit cycle attracts every solution curve away from the origin. We conclude that the RLC circuit of Example 6.3 has the behavior shown in Fig. 6.22 regardless of the exact value of the capacitance C, assuming that C is close to 1.

6.4 Chaos and Fractals

One of the most exciting mathematical discoveries of the twentieth century is the chaotic behavior of some dynamic models. *Chaos* is characterized by the apparently random behavior of solutions, with extreme sensitivity to initial conditions. Chaotic dynamical systems can give rise to exotic limit sets called *fractals*. Chaotic dynamical system models have been used to handle problems in turbulent fluid flow, ecosystems with aperiodic population fluctuations, cardiac arrythmia, occasional reversal of the earth's magnetic poles, complex

Figure 6.24: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = 0.1$, $x_2(0) = 0.3$, $C = 0.75$.

Figure 6.25: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = 0.1$, $x_2(0) = 0.3$, $C = 1.5$.
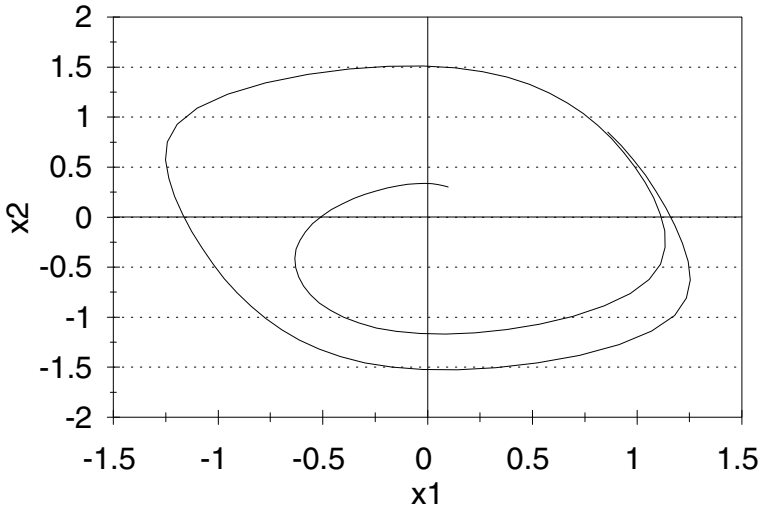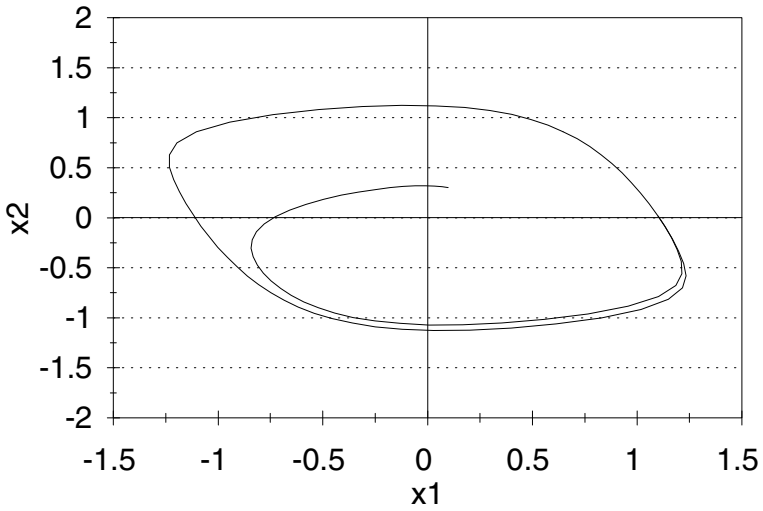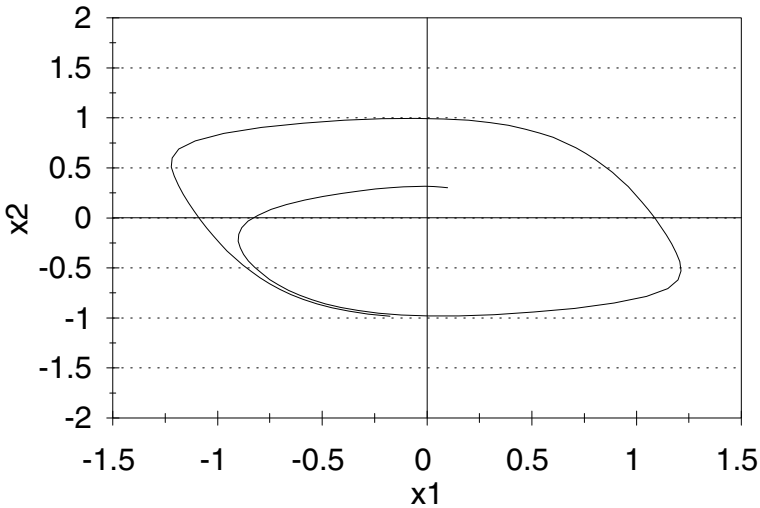
Figure 6.26: Graph of voltage x_2 versus current x_1 for the nonlinear RLC circuit problem: case $x_1(0) = 0.1$, $x_2(0) = 0.3$, $C = 2.0$.

chemical reactions, lasers, and the stock market. Many of these applications are controversial, and their implications are still being explored. One of the most surprising things about chaos is the way it can emerge from simple nonlinear dynamic models.

Example 6.4. Reconsider the whale problem of Example 4.2, but now suppose that we use a discrete model to project population growth with a time step of several years. We know that for a time step of one year, the discrete and continuous time models behave in essentially the same manner. How large a time step can we use and still retain the same qualitative behavior as the continuous time model? What happens to the model if we use too large a time step?

We will use the five-step method. The results of step 1 are the same as in Figure 4.3. In step 2 we specify a continuous–time dynamical system model, which we will solve by simulation using the Euler method.

Consider a continuous–time dynamical system model

$$\frac{dx}{dt} = F(x) \tag{6.12}$$

with $x = (x_1, \ldots, x_n)$ and $F = (f_1, \ldots, f_n)$, along with the initial condition $x(t_0) = x_0$. The Euler method uses a discrete time dynamical system

$$\frac{\Delta x}{\Delta t} = F(x) \tag{6.13}$$

to approximate the behavior of the continuous time system. One reason to use a large step size is to make long-range predictions. For example, if time t is measured in years, then $\Delta x = F(x)\Delta t$ is a simple projection of the change in the state variable x over the next Δt years, based on the current state information. If the step size Δt is chosen so that the relative change $\Delta x/x$ remains small, then the discrete time system in Eq. (6.13) will behave much like the original continuous time system in Eq. (6.12). If the step size is too big, then the discrete time system can exhibit very different behavior.

Example 6.5. Consider the simple linear differential equation

$$\frac{dx}{dt} = -x. \tag{6.14}$$

Compare the behavior of solutions to Eq. (6.14) to those of its discrete time analogue

$$\frac{\Delta x}{\Delta t} = -x. \tag{6.15}$$

Solutions to Eq. (6.14) are all of the form

$$x(t) = x(0)e^{-t}. \tag{6.16}$$

The origin is a stable equilibrium, and every solution curve decays exponentially fast to zero. For Eq. (6.15) the iteration function is

$$\begin{aligned} G(x) &= x + \Delta x \\ &= x - x\Delta t \\ &= (1 - \Delta t)x. \end{aligned}$$

Solutions to Eq. (6.15) are all of the form

$$x(n) = (1 - \Delta t)^n x(0).$$

If $0 < \Delta t < 1$, then $x(n) \to 0$ exponentially fast, and the behavior is much like that of the continuous time differential equation. If $1 < \Delta t < 2$, then we still have $x(n) \to 0$, but the sign of $x(n)$ oscillates between positive and negative. Finally, if $\Delta t > 2$, then $x(n)$ diverges to infinity as it oscillates in sign. In summary, the solutions of Eq. (6.15) behave much like those of Eq. (6.14) as long as the time step Δt is chosen so that the relative change $\Delta x/x$ remains small. If the time step Δt is too large, then Eq. (6.15) exhibits behavior entirely different from that of its continuous time analogue Eq. (6.14).

For linear dynamical systems, the time delays inherent in discrete approximations can lead to unexpected behavior. A stable equilibrium can become unstable, and new oscillations can occur.

For linear systems, this is about the only way in which the behavior of the discrete approximation can differ from that of the original continuous system. For nonlinear continuous time dynamical systems, however, discrete approximations can also exhibit chaotic behavior. In a chaotic dynamical system, there is an extreme sensitivity to initial conditions, along with apparently random behavior of individual solutions. Chaos is usually associated with systems in which nearby solutions tend to diverge, but overall remain bounded. This combination of factors can only occur in a nonlinear system.

The study of chaos in discrete time dynamical systems is an active area of research. Some iteration functions produce extremely complex sample paths, including *fractals*. A typical fractal is a set of points in the state space that is self–similar and whose *dimension* is not an integer. *Self–similar* means that the object contains smaller pieces that are exact scaled–down replicas of the whole. One simple way to measure dimension is to count the number of boxes needed to cover the object. For a one–dimensional object, it takes n times as many boxes if they are $1/n$ times as large; for a two–dimensional object, it takes n^2 times as many, and so on. For an object with fractal dimension d, the number of boxes it takes to cover the object increases like n^d as the box size $1/n$ tends to zero.

Step 3 is to formulate the model. We begin with the continuous time dynamical system model

$$
\begin{aligned}
\frac{dx_1}{dt} &= f_1(x_1, x_2) = 0.05x_1 \left(1 - \frac{x_1}{150,000}\right) - \alpha x_1 x_2 \\
\frac{dx_2}{dt} &= f_2(x_1, x_2) = 0.08x_2 \left(1 - \frac{x_2}{400,000}\right) - \alpha x_1 x_2
\end{aligned}
\tag{6.17}
$$

on the state space $x_1 \geq 0$, $x_2 \geq 0$, where x_1 denotes the population of blue whales and x_2 the population of fin whales. In order to simulate this model, we will transform to a set of difference equations

$$
\begin{aligned}
\Delta x_1 &= f_1(x_1, x_2)\Delta t \\
\Delta x_2 &= f_2(x_1, x_2)\Delta t
\end{aligned}
\tag{6.18}
$$

over the same state space. Then, for example, Δx_1 represents the change in the population of Blue whales over the next Δt years. We will assume that $\alpha = 10^{-8}$ to start with, and later on we will do a sensitivity analysis on α. Our objective is to determine the behavior of solutions to the discrete time dynamical system in Eq. (6.18) and compare to what we know about solutions to the continuous time model in Eq. (6.17).

In step 4 we solve the problem by simulating the system in Eq. (6.17) using a computer implementation of the Euler method for several different values of $h = \Delta t$. We assume that $x_1(0) = 5,000$ and $x_2(0) = 70,000$, as in Example 6.2. Figure 6.27 illustrates the results of our simulation with $N = 50$ iterations and
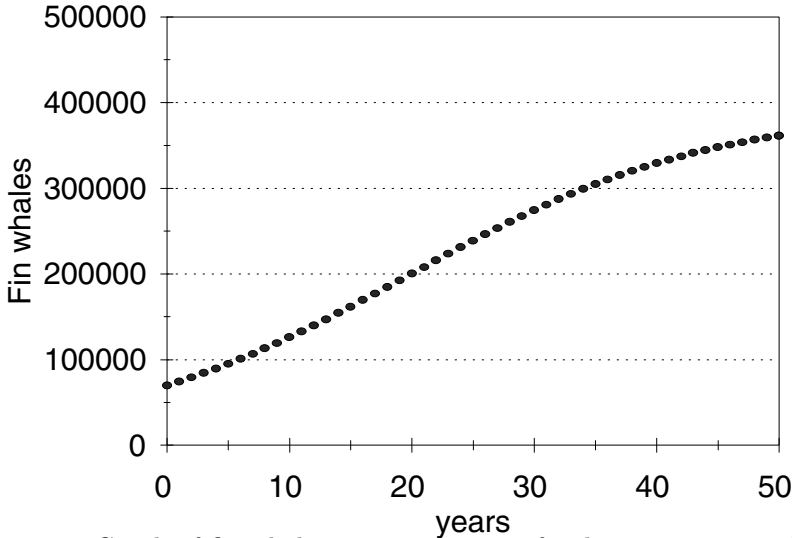
Figure 6.27: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 1$.

a time step of $h = 1$ years. In 50 years the fin whales grow back steadily but do not quite reach their eventual equilibrium level. In Figure 6.28 we increase the step size to $h = 2$ years. Now our simulation with $N = 50$ iterations shows the fin whale population approaching its equilibrium value. Using a larger time step h is an efficient way to project further into the future, but then something interesting happens.

Figure 6.29 shows the result of using a time step of $h = 24$ years. The solution still approaches its equilibrium, but now there is an oscillation. Figure 6.30 shows what happens when we use a time step of $h = 27$ years. Now the population actually diverges from the equilibrium and eventually settles into a discrete limit cycle of period two.

When $h = 32$, the solution settles down into a limit cycle of period four; see Figure 6.31. Figure 6.32 shows that when $h = 37$, the solution exhibits chaotic behavior. The effect is similar to that of a random number generator. When $h = 40$ (not shown), the solution quickly diverges to infinity. The behavior of the blue whale population is similar. Different initial conditions and different values of α produce similar results. In every case there is a transition from stability to instability as the step size h increases. As the equilibrium becomes unstable, first oscillations appear, then discrete limit cycles, and then chaos. Finally, if h is too large the solutions simply diverge.

Step 5 is to answer the question. The discrete approximation to the continuous time model is useful as long as the time step is small enough to keep the relative change in the state variables small at each time step. Using a larger

Figure 6.28: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 2$.

Figure 6.29: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 24$.

Figure 6.30: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 27$.
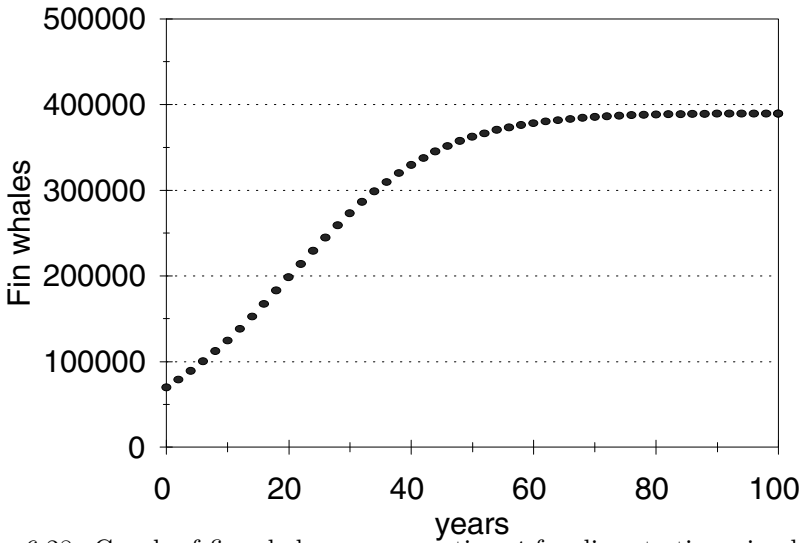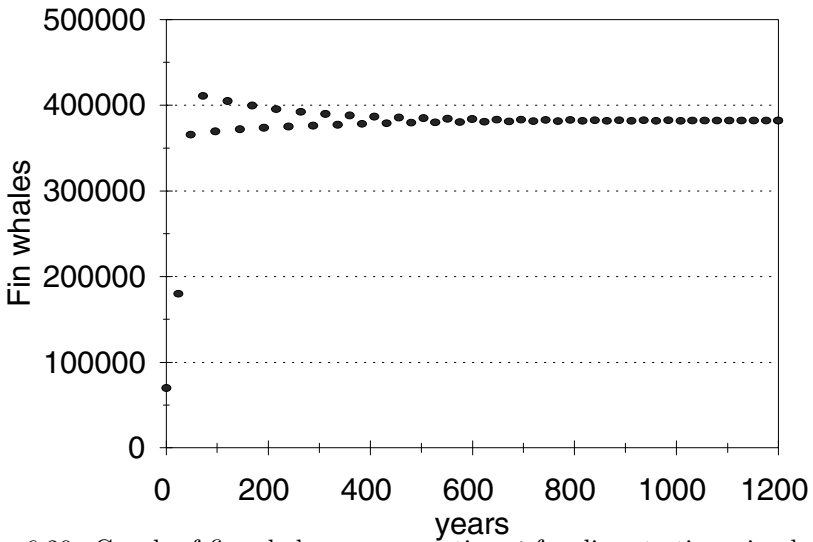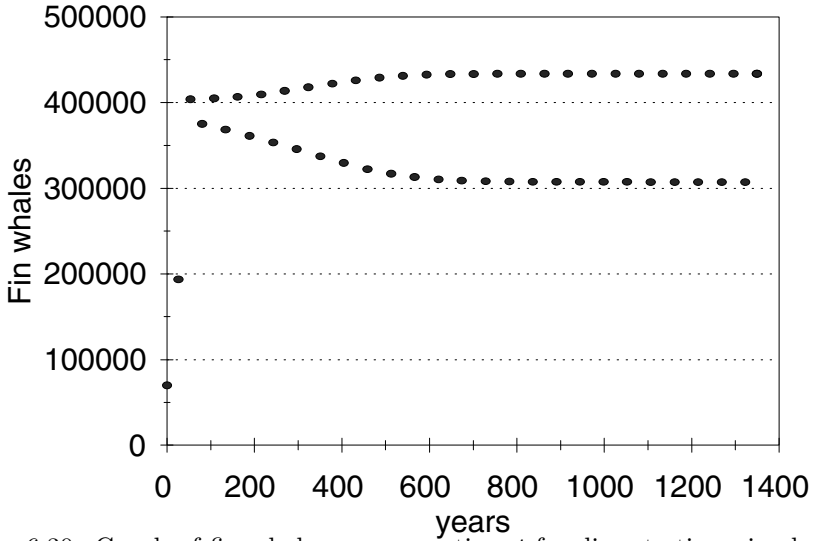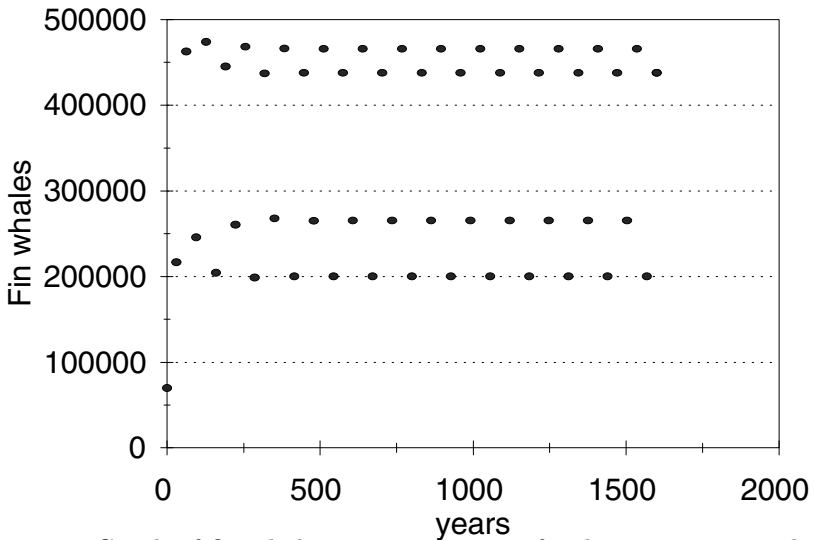

Figure 6.31: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 32$.
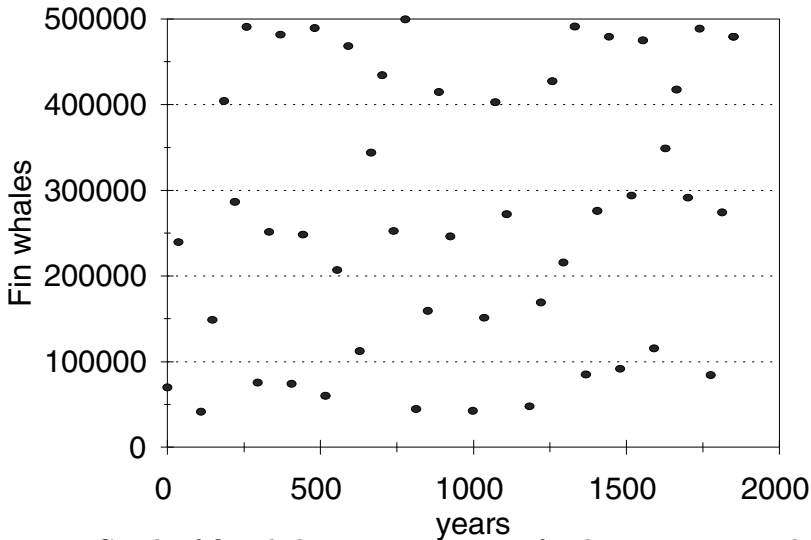
Figure 6.32: Graph of fin whales x_2 versus time t for discrete time simulation of the whale problem with time step $h = 37$.

time step allows us to project further into the future, but when the time step becomes too large, the behavior of the discrete time system no longer resembles that of the original continuous time model. It is interesting to observe the strange behavior of the discrete time system when a larger time step is employed; however, this behavior has no apparent connection to the real–world situation we are trying to model.

Many population models are based on some variation of the logistic model $x' = rx(1-x/K)$. Most of these models exhibit chaos in the discrete approximation. It is typical for the dynamics to transition from stable equilibrium to limit cycles to chaos (and then to unstable divergence) as the time step is increased. At the transition from limit cycles to chaos, the limit set typically becomes a fractal. See Exercise 25 at the end of this chapter for an illustration. There are many interesting books and articles on chaos and fractals. Strogatz (1994) is one good reference at an advanced undergraduate to beginning graduate level.

The emergence of chaos from the discrete approximation in the whale problem is interesting, but seems to have nothing to do with the real world. At best it is a mathematical curiosity, at worst a numerical headache. The next example, however, shows that chaos and fractals can also emerge naturally from realistic models of physical situations.

Example 6.6. Consider a layer of air that is heated from the bottom. In certain situations the warmer air rising up interacts with the colder air sinking down to form turbulent convection rolls. The complete derivation of the dynamics of

motion involves a system of partial differential equations, which can be solved by the method of Fourier transforms; see Lorentz (1963). A simplified representation involves three state variables. The variable x_1 represents the rate at which the convection rolls rotate, x_2 represents the temperature difference between the ascending and descending air currents, and x_3 represents the deviation from linearity of the vertical temperature profile, a positive value indicating that the temperature varies faster near the boundary. The equations of motion for this system are

$$\begin{aligned}
x_1' &= f_1(x_1, x_2, x_3) = -\sigma x_1 + \sigma x_2 \\
x_2' &= f_2(x_1, x_2, x_3) = -x_2 + r x_1 - x_1 x_3 \\
x_3' &= f_3(x_1, x_2, x_3) = -b x_3 + x_1 x_2,
\end{aligned} \tag{6.19}$$

and we will consider the realistic case where $\sigma = 10$ and $b = 8/3$. The remaining parameter r represents the temperature difference between the top and bottom of the air layer. Increasing r pumps more energy into the system, creating more vigorous dynamics. The dynamical system of Eq. (6.19) is called the *Lorentz equations*, after the meteorologist E. Lorentz who analyzed them.

To find the equilibrium points of Eq. (6.19), we solve the system of equations

$$\begin{aligned}
-\sigma x_1 + \sigma x_2 &= 0 \\
-x_2 + r x_1 - x_1 x_3 &= 0 \\
-b x_3 + x_1 x_2 &= 0
\end{aligned}$$

for the three state variables. Obviously, $(0, 0, 0)$ is one solution. The first equation implies $x_1 = x_2$. Substituting into the second equation, we obtain

$$\begin{aligned}
-x_1 + r x_1 - x_1 x_3 &= 0 \\
x_1(-1 + r - x_3) &= 0
\end{aligned}$$

so that if $x_1 \neq 0$, then $x_3 = r - 1$. Then, from the third equation we obtain $x_1^2 = b x_3 = b(r - 1)$. If $0 < r < 1$, there are no real roots to this equation, and so the origin is the only equilibrium point. If $r = 1$, then $x_3 = 0$, and again the origin is the only equilibrium point. If $r > 1$, then there are three equilibrium points

$$\begin{aligned}
E_0 &= (0, 0, 0) \\
E^+ &= (\sqrt{b(r-1)}, \sqrt{b(r-1)}, r - 1) \\
E^- &= (-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, r - 1).
\end{aligned}$$

Graphical analysis of the vector field is difficult, since we are now in three dimensions. Instead, we will perform an eigenvalue analysis to test the stability of these three equilibrium points. The matrix of partial derivatives is

$$DF = \begin{pmatrix} -\sigma & \sigma & 0 \\ r - x_3 & -1 & -x_1 \\ x_2 & x_1 & -b \end{pmatrix}.$$

At the equilibrium $E_0 = (0, 0, 0)$ with the parameter values $\sigma = 10$ and $b = 8/3$, this matrix becomes

$$A = \begin{pmatrix} -10 & 10 & 0 \\ r & -1 & 0 \\ 0 & 0 & -8/3 \end{pmatrix},$$

which has three real eigenvalues

$$\lambda_1 = \frac{-11 - \sqrt{81 + 40r}}{2}$$

$$\lambda_2 = \frac{-11 + \sqrt{81 + 40r}}{2}$$

$$\lambda_3 = \frac{-8}{3}$$

for any value of $r > 0$. If $0 < r < 1$, then all of these eigenvalues are negative, so that the origin is a stable equilibrium. If $r > 1$, then $\lambda_2 > 0$, so the origin is an unstable equilibrium.

The eigenvalue analysis for the remaining two equilibria is rather messy. Fortunately, the eigenvalues are the same at both E^+ and E^-. For $1 < r < r_1 \approx 1.35$, all three eigenvalues are real and negative. For any value of $r > r_1$, there is one eigenvalue $\lambda_1 < 0$ and one complex conjugate pair $\lambda_2 = \alpha + i\beta$, $\lambda_3 = \alpha - i\beta$. The real part α is negative for $r_1 < r < r_0$ and positive for $r > r_0$, where $r_0 \approx 24.8$. Thus, for $1 < r < r_0$, there are two stable equilibria at E^+ and E^-, while for $r > r_0$, every equilibrium is unstable. Solutions nearby these two equilibria will behave much like those of the linear system $x' = Ax$, where A is the matrix of partial derivatives DF evaluated at the equilibrium point. Every component of every linear solution can be written as a linear combination of terms of the form $e^{\lambda_1 t}$, $e^{\alpha t} \cos(\beta t)$, and $e^{\alpha t} \sin(\beta t)$. When $r_1 < r < r_0$, nearby solution curves spiral into the nonzero equilibrium points, and when $r > r_0$, they spiral outwards. It also turns out that solutions do not diverge to infinity when $r > r_0$. We have seen this kind of behavior before, in the nonlinear RLC circuit of Example 5.4. In that case a computer simulation showed that the solution curves settled into a limit cycle. We will now simulate the dynamical system of Eq. (6.19) to determine the long–term behavior of solutions.

The Euler method for three state variables uses exactly the same algorithm as in Figure 6.19 except for the addition of another state variable. A computer implementation of that algorithm was used to simulate the solution of Eq. (6.19) in the case $\sigma = 10$ and $b = 8/3$. Figure 6.33 shows the results of a simulation using $r = 8$ and initial conditions $(x_1, x_2, x_3) = (1, 1, 1)$. We graphed x_2 versus x_1 since these variables are the easiest to interpret. We used $N = 500$ and $T = 5$, so that the step size was $h = 0.01$. Additional sensitivity runs were made to ensure that increasing the simulation time T or decreasing the step size h led to essentially the same graph. As we expected from our earlier analysis, the solution curve spirals into the equilibrium point at $x_1 = x_2 = 4.32$ (and $x_3 = 7$). Recall that x_1 represents the rate at which the convection rolls rotate, and x_2 represents the temperature difference between the ascending and descending air
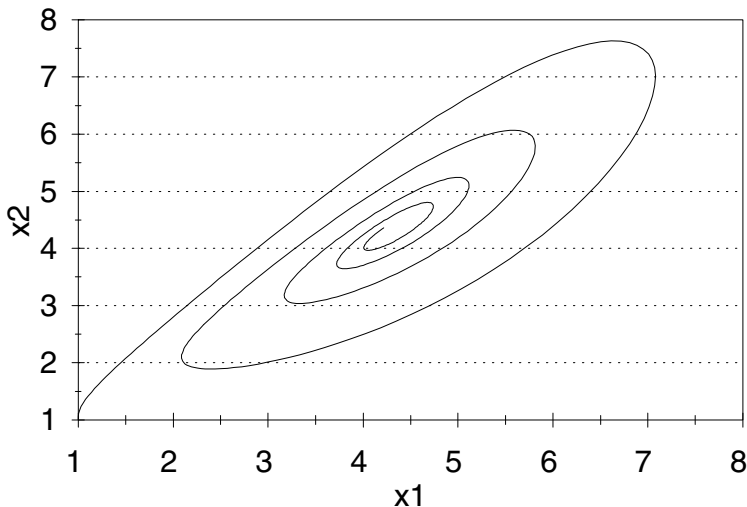
Figure 6.33: Graph of temperature differential x_2 versus convection rate x_1 for the weather problem with $r = 8$ and initial condition $(x_1, x_2, x_3) = (1, 1, 1)$.

currents. When $r = 8$, both of these quantities eventually settle down into a stable equilibrium. Figure 6.34 shows the simulation for the case $r = 8$ and initial condition $(x_1, x_2, x_3) = (7, 1, 2)$. Keep in mind that this graph is actually a projection of a three–dimensional picture. Of course, the real solution curve does not cross itself. That would violate the uniqueness of solutions. Once again, the solution spirals into the equilibrium.

As we increase the value of r, the simulation becomes extremely sensitive to discretization. Figure 6.35 shows the results of a simulation using $r = 18$ and initial conditions $(x_1, x_2, x_3) = (6.7, 6.7, 17)$. We used $N = 500$ and $T = 2.5$ for a step size of $h = 0.005$. The solution curve rotates rapidly about the equilibrium $E^+ = (6.733, 6.733, 17)$ while it spirals in very slowly. Figure 6.36 shows the results of the same simulation using $N = 500$ and $T = 5$ for a slightly larger step size of $h = 0.01$. In this case the solution spirals outward, away from the equilibrium. Of course, this is not really what is happening in the continuous time model, it is just an artifact of our simulation method. Because the system is very near the point between stability and instability, we must be careful to perform sensitivity analysis on the parameters N and T to ensure that the behavior of the discrete time system really reflects what is going on with the continuous time model.

Finally, we consider the case $r > r_0$, where we know that the equilibrium points are all unstable. Figure 6.37 shows the results of a simulation using $r = 28$ and initial conditions $(x_1, x_2, x_3) = (9, 8, 27)$. We used $N = 500$ and $T = 10$ for a step size of $h = 0.02$. A careful sensitivity analysis on N and T was performed

Figure 6.34: Graph of temperature differential x_2 versus convection rate x_1 for the weather problem with $r = 8$ and initial condition $(x_1, x_2, x_3) = (7, 1, 2)$.

Figure 6.35: Graph of temperature differential x_2 versus convection rate x_1 for the weather problem with $r = 18$ and initial condition $(x_1, x_2, x_3) = (6.7, 6.7, 17)$ using step size $h = 0.005$.

Figure 6.36: Graph of temperature differential x_2 versus convection rate x_1 for the weather problem with $r = 18$ and initial condition $(x_1, x_2, x_3) = (6.7, 6.7, 17)$ using step size $h = 0.01$.
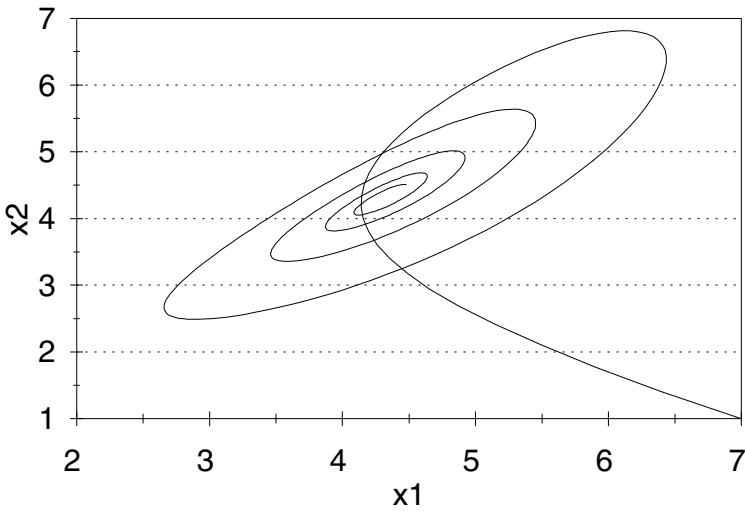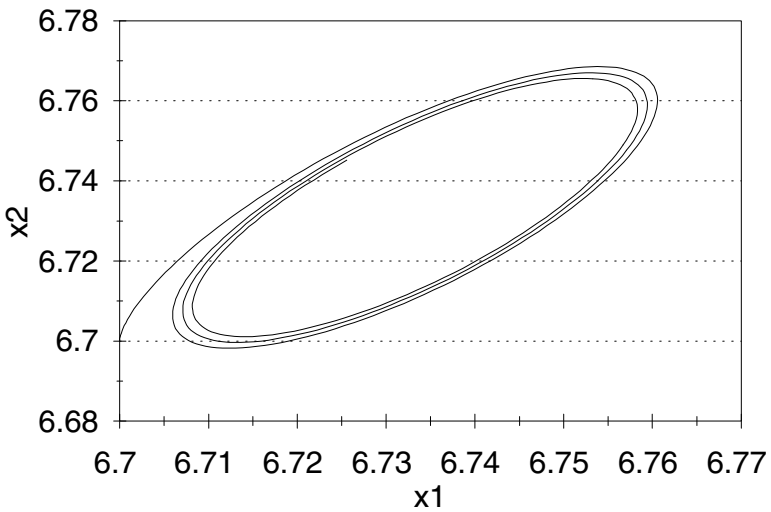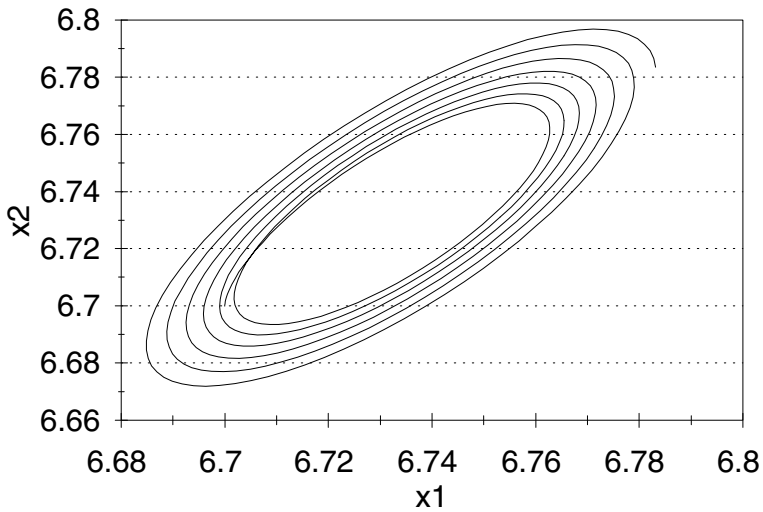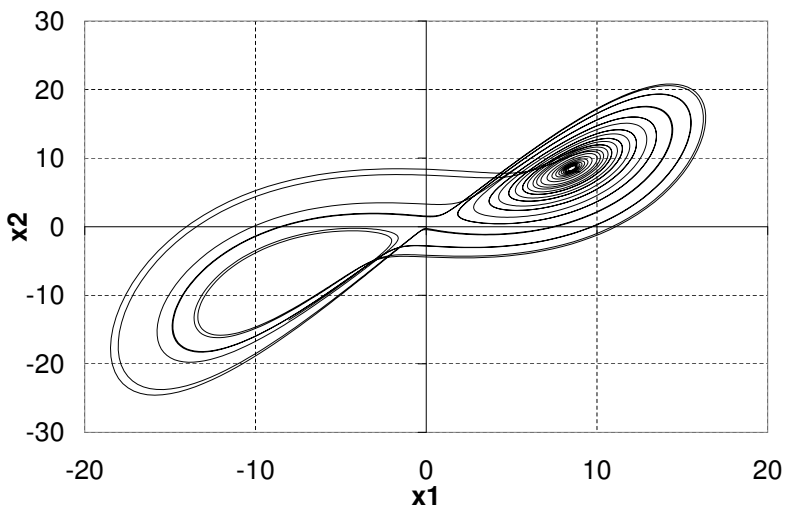

Figure 6.37: Graph of temperature differential x_2 versus convection rate x_1 for the weather problem with $r = 28$ and initial condition $(x_1, x_2, x_3) = (9, 8, 27)$.
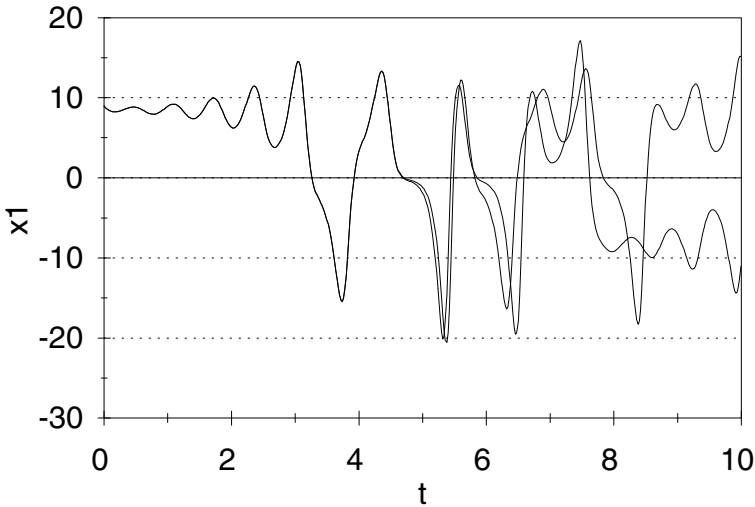
Figure 6.38: Graph of convection rate x_1 versus time t for the weather problem with $r = 28$ comparing two initial conditions $(x_1, x_2, x_3) = (9, 8, 27)$ and $(x_1, x_2, x_3) = (9.01, 8, 27)$.

to verify that this solution curve really represents the behavior of our continuous time model. At first the solution curve rotates rapidly about the equilibrium $E^+ = (8.485, 8.485, 27)$ while it spirals outward very slowly. Eventually, the solution curve heads out toward the equilibrium $E^- = (-8.485, -8.485, 27)$, where it spirals around for a while, and then eventually heads back towards E^+. Simulations using larger values of N and T indicate that the solution never repeats itself, yet remains bounded. It continues to cross over between the region around E^+ and the region around E^-.

Other solutions with nearby initial conditions show essentially the same behavior. There is also an extreme sensitivity to initial conditions. Solutions that begin very close together eventually move farther apart. The solution curve represented in Figure 6.37 started at the point $(x_1, x_2, x_3) = (9, 8, 27)$ at time $t = 0$, and by time $t = 10$, it was spiraling around E^-. A nearby solution curve, starting at the point $(x_1, x_2, x_3) = (9.01, 8, 27)$ at time $t = 0$, ends up spiraling around E^+ by time $t = 10$. Figure 6.38 compares the path of the x_1 coordinate for these two solution curves. By time $t = 10$ it would be impossible to guess that these two solutions began at almost the same initial condition. Yet if we let the simulation time T get larger and larger, we see that both solution curves trace out almost exactly the same shape in the state space. This limiting set is called a *strange attractor*. Although it is bounded, its length is infinite, and cross sections (for example, the collection of all points at which the curve intersects the plane $x_3 = 0$) are typically fractals. In fact, one reasonable analytic

approach for dealing with the strange attractor is to consider the mapping that sends a given intersection point to the next. This iteration function acts much like the one we analyzed in Example 6.4 for the logistic model of population growth. In this rather unexpected way, the analysis of discrete and continuous time dynamical systems finds a new connection.

6.5   Exercises

1. Reconsider the war problem of Example 6.1. In this problem we explore the effects of weather on combat. Bad weather and poor visibility decrease the effectiveness of direct fire weapons for both sides. The effectiveness of indirect fire weapons is relatively unaffected by the weather. We can represent the effects of bad weather in our model as follows. Let w denote the decrease in weapon effectiveness caused by bad weather conditions, and replace the dynamical system in Eq. (6.3) by

$$\Delta x_1 = -w\lambda(0.05)x_2 - \lambda(0.005)x_1x_2$$
$$\Delta x_2 = -w0.05x_1 - 0.005x_1x_2. \tag{6.20}$$

Here, the parameter $0 \le w \le 1$ represents a range of weather conditions, with $w = 1$ indicating the best weather and $w = 0$ indicating the worst weather.

(a) Use a computer implementation of the algorithm in Fig. 6.2 to simulate the discrete–time dynamical system in Eq. (6.20) in the case $\lambda = 3$. Assume that adverse weather conditions cause a 75% decrease in weapon effectiveness for both sides ($w = 0.25$). Who wins the battle, and how long does it take? How many divisions of troops remain on the winning side?

(b) Repeat your analysis for each of the cases $w = 0.1, 0.2, 0.5, 0.75$, and 0.9, and tabulate your results. Answer the same questions as in part (a).

(c) Which side benefits from fighting in adverse weather conditions? If you were the blue commander, would you expect red to attack on a sunny day or a rainy day?

(d) Examine the sensitivity of your results in parts (a), (b), and (c) to the degree of weapon superiority of blue over red. Repeat the simulations in parts (a) and (b) for $\lambda = 1.5, 2.0, 4.0$, and 5.0, and tabulate your results as before. Reconsider your conclusions in part (c). Are they still valid?

2. Reconsider the war problem of Example 6.1. In this problem we will consider the effect of tactics on the outcome of the battle. The red commander is considering the option of holding two of his five divisions in reserve until the second or third day of combat. You can simulate each possibility as

a deviation from the baseline case by running two separate simulations. First, simulate the first one or two days of battle, matching two blue divisions against three red divisions. Then, use the final outcome of that simulation as the initial conditions for the rest of the battle, except add two more divisions to the red side.

(a) Use a computer implementation of the algorithm in Fig. 6.2 to simulate the first phase of battle, in which two blue divisions fight three red divisions. Assume $\lambda = 2$, and tabulate the final force levels for the two cases (12 or 24 hours of battle).

(b) Use the results of part (a) to simulate the next phase of battle. Add two divisions to the final force levels for red, and continue the simulation. In each case, tell which side wins the battle, how many units remain on the winning side, and how long the battle lasts (total time for both phases of battle).

(c) The red commander may choose to commit all of his forces on the first day, or he may keep two divisions in reserve for one or two days. Which of the three strategies is better? Optimize on the basis of achieving victory at the minimum cost in terms of lost manpower.

(d) Perform a sensitivity analysis on the parameter λ, which describes the extent to which blue has weapon superiority. Repeat parts (a) and (b) for $\lambda = 1.0, 1.5, 3.0, 5.0$, and 6.0, and identify the optimal strategy for each value of λ. State your general conclusions concerning the optimal strategy for red.

3. Reconsider the war problem of Example 6.1. In this problem we will investigate the effect of tactical nuclear weapons on the battle. As a desperation move, the blue commander considers calling for a tactical nuclear strike. It is estimated that such a strike will kill or incapacitate 70% of the red force, and 35% of the blue force as well.

(a) Use a computer implementation of the algorithm in Fig. 6.2 to simulate the discrete–time dynamical system in Eq. (6.3), assuming that the blue commander calls for an immediate nuclear strike. Start with initial conditions $x_1 = (0.30)5.0, x_2 = (0.65)2.0$, and assume $\lambda = 3$. Who wins the battle, and how long does it take? How many divisions survive on the winning side? How does blue benefit from calling for a nuclear attack in this case?

(b) Simulate the case where the blue commander waits for six hours and then calls for a nuclear attack. Simulate six hours of battle, starting with $x_1 = 5$ divisions and $x_2 = 2$ divisions. Reduce the remaining number of troops for both sides to represent the results of a nuclear strike, and then continue the simulation. Answer the same questions as in part (a).

(c) Compare the results of parts (a) and (b) to the case of conventional combat summarized in the chapter. Discuss the benefits of a tactical

nuclear strike by blue. Can such a move be effective and, if so, when should the commander request the strike?

(d) Examine the sensitivity of your conclusions in part (c) to the extent λ of blue weapon superiority. Repeat the simulations of parts (a) and (b) for each of the cases $\lambda = 1.0$, 1.5, 2.0, 5.0, and 6.0, and answer the same questions as before.

4. Reconsider the space docking problem of Example 5.2. Suppose that our initial closing velocity is 50 m/sec under zero acceleration.

(a) Determine the time required for docking, assuming that the control factor $k = 0.02$. Use a computer implementation of the algorithm for discrete–time dynamical systems described in Fig. 6.2. Assume docking is complete when the closing velocity has been reduced to less than 0.1 m/sec in absolute value for all future time.

(b) Repeat the simulation of part (a) for each of the cases $k = 0.01$, 0.02, 0.03, ..., 0.20, and determine the time required for docking in each case. Which of these values of k results in the quickest time to dock?

(c) Repeat part (b), assuming an initial closing velocity of 25 m/sec.

(d) Repeat part (b), assuming an initial closing velocity of 100 m/sec. What conclusions can you draw about the optimal value of k for this docking procedure?

5. Reconsider the infectious disease problem introduced in Exercise 10 of Chapter 4. Use a computer implementation of the algorithm in Fig. 6.2 to simulate this discrete–time dynamical system model. Answer the questions in parts (a) through (d) from the original exercise.

6. In Exercise 4 of Chapter 4, we introduced a simplified model of population growth in the whale problem.

(a) Simulate this model, assuming that there are currently 5,000 blue whales and 70,000 fin whales. Use the simple simulation technique of Section 6.2, and assume $\alpha = 10^{-7}$. What happens to the two species of whales over the long term, according to this model?

(b) Examine the sensitivity of your conclusions in part (a) to the assumption that there are currently 5,000 blue whales. Repeat the simulation of part (a), assuming that there are originally 3,000, 4,000, 6,000, or 7,000 blue whales. How sensitive are your conclusions to the exact number of blue whales in the ocean at the present time?

(c) Examine the sensitivity of your conclusions in part (a) to the assumption that the intrinsic growth rate of the blue whale is 5% per year. Repeat the simulation of part (a), assuming that the actual rate is 3, 4, 6, or 7% per year. How sensitive are your conclusions to the actual intrinsic growth rate for the blue whales?

(d) Examine the sensitivity of your conclusions in part (a) to the competition coefficient α. Repeat the simulation of part (a) for each of the cases $\alpha = 10^{-9}$, 10^{-8}, 10^{-6}, and 10^{-5}, and tabulate your results. How sensitive are your general conclusions to the extent of competition between the two species?

7. In Exercise 5 of Chapter 4, we introduced a more sophisticated model of population growth in the whale problem.

   (a) Simulate this model, assuming that there are currently 5,000 blue whales and 70,000 fin whales. Use the simple simulation technique of Section 6.2, and assume that $\alpha = 10^{-8}$. What happens to the two species of whales over the long term, according to this model? Do both species of whales grow back, or will one or both species become extinct? How long does this take?

   (b) Examine the sensitivity of your conclusions in part (a) to the assumption that there are currently 5,000 blue whales. Repeat the simulation of part (a), assuming that there are originally 2,000, 3,000, 4,000, 6,000, or 8,000 blue whales. How sensitive are your conclusions to the exact number of blue whales in the ocean at the present time?

   (c) Examine the sensitivity of your conclusions in part (a) to the assumption that the intrinsic growth rate of the blue whale is 5% per year. Repeat the simulation of part (a), assuming that the actual rate is 2, 3, 4, 6, or 7% per year. How sensitive are your conclusions to the actual intrinsic growth rate for the blue whales?

   (d) Examine the sensitivity of your conclusions in part (a) to the assumption that the minimum viable population level of the blue whale is 3,000 whales. Repeat the simulation of part (a), assuming that the actual level is 1,000, 2,000, 4,000, 5,000, or 6,000 whales. How sensitive are your conclusions to the actual minimum viable population level for the blue whales?

8. Reconsider Exercise 6 in Chapter 4, and assume $\alpha = 10^{-8}$. Assume that the current population levels are $B = 5,000$ and $F = 70,000$.

   (a) Use a computer implementation of the simple algorithm used in Section 6.2 to determine the effect of harvesting. Assume $E = 3,000$ boat–days per year. What happens to the two species of whales over the long term, according to this model? Do both species of whales grow back, or will one or both species become extinct? How long does this take?

   (b) Repeat part (a), assuming that $E = 6,000$ boat–days per year.

   (c) For what range of E does the number of whales of both species approach a nonzero equilibrium?

(d) Repeat part (c) for each of the cases $\alpha = 10^{-9}$, 10^{-8}, 10^{-6}, and 10^{-5}, and tabulate your results. Discuss the sensitivity of your conclusions to the extent of interspecies competition.

9. Reconsider the whale harvesting problem of Exercise 6 in Chapter 4. In this problem we will explore the economic incentives for whalers to drive one species of whale to extinction. Assume that there are currently 5,000 blue whales and 70,000 fin whales.

   (a) Simulate this model, assuming $E = 3,000$ boat–days. Use the simple simulation technique of Section 6.2, and assume that $\alpha = 10^{-7}$. Determine the long–term harvest rate in blue whale units per year (2 fin whales = 1 blue whale unit).

   (b) Determine the level of effort that maximizes the long–term harvest rate in blue whale units. Simulate each of the cases $E = 500$, $1,000$, $1,500$, ..., $7,500$, boat–days per year. Which case results in the highest sustainable yield?

   (c) Assume that whalers harvest at the rate that maximizes their long–term sustainable yield. What happens to the two species of whales over the long term, according to this model? Do both species of whales grow back, or will one or both species become extinct? How long does this take?

   (d) Some economists argue that whalers will act to maximize the long–term sustainable yield for the entire industry. If so, would continued harvesting cause one or both species of whales to become extinct?

10. (Continuation of Exercise 9) Some economists argue that whalers will act in such a way as to maximize the total discounted revenue obtained by the entire whaling industry. Assume that harvesting produces revenue of $10,000 per blue whale unit, and assume a discount rate of 10%. If revenue R_i is obtained in year i, the total discounted revenue is defined as

$$R_0 + \lambda R_1 + \lambda^2 R_2 + \lambda^3 R_3 + \cdots,$$

where $1 - \lambda$ represents the discount rate ($\lambda = 0.9$ for this problem).

   (a) Simulate this model, assuming $E = 3,000$ boat–days. Use the simple simulation technique of Section 6.2, and assume that $\alpha = 10^{-7}$. Determine the total discounted revenue for this case.

   (b) Determine the level of effort that maximizes the total discounted revenue. Simulate each of the cases $E = 500$, $1,000$, $1,500$, ..., $7,500$, boat–days per year. Which case results in the highest yield?

   (c) Assume that whalers harvest at the rate that maximizes their total discounted revenue. What happens to the two species of whales over the long term, according to this model? Do both species of whales grow back, or will one or both species become extinct? How long does this take?

(d) Perform a sensitivity analysis on the parameter α, which measures the extent of interspecies competition. Consider each of the cases $\alpha = 10^{-9}$, 10^{-7}, 10^{-6}, and 10^{-5}, and tabulate your results. Discuss the sensitivity of your conclusions to the extent of interspecies competition.

11. Reconsider the predator–prey model of Exercise 7 in Chapter 4.

(a) Determine by simulation the equilibrium levels for whales and krill. Use the simple simulation technique discussed in Section 6.2. Begin at several different initial conditions and run the simulation until both population levels have settled down into steady state.

(b) Suppose that after both population levels have settled down into steady state, an ecological disaster kills off 20% of the whales and 80% of the krill. Describe what happens to the two species, and how long it takes.

(c) Suppose that harvesting has depleted the whales to 5% of their equilibrium population level, while krill remain at about the same level. Describe what happens once harvesting is stopped. How long does it take for the whales to grow back? What happens to the krill population?

(d) Examine the sensitivity of your results in part (c) to the assumption that 5% of the whales remain. Simulate each of the cases where 1, 3, 7, or 10% remain, and tabulate your results. How sensitive is the time it takes for the whales to grow back to the extent to which the population is depleted?

12. Reconsider the tree problem of Example 5.1.

(a) Determine how long it will take for both hardwoods and softwoods to grow to 90% of their stable equilibrium levels. Assume an initial population of 1,500 tons/acre of softwood trees and 100 tons/acre of hardwoods. This is the situation in which we are trying to introduce a new type of more valuable tree into an existing ecosystem. Assume $b_i = a_i/2$, and use the simple simulation technique introduced in Section 6.2.

(b) Determine the point at which the biomass of hardwood trees is increasing at the fastest rate.

(c) Assuming that hardwoods are worth four times as much as softwoods in $/ton, determine the point at which the value of the forest stand ($/acre) is increasing at the fastest rate.

13. (Continuation of Exercise 12) Clear–cutting is a technique in which all of the trees in the forest are harvested at one time and then replanted.

(a) Determine the optimal harvest policy for this forest; i.e., determine the number of years we should wait before cutting and replanting. Assume that replanting involves 100 tons/acre of hardwoods and 100 tons/acre of softwoods. Base your answer on the number of $/acre per year generated.

(b) Determine the optimal harvest policy, assuming that only hardwoods are replanted (200 tons/acre).

(c) Repeat part (b), but now assume that only softwoods are replanted (200 tons/acre).

(d) State the optimal clear–cutting policy for management of this tract of forest land. At what point would you consider selling the land rather than reforesting?

14. Reconsider the more sophisticated competing species model of Exercise 5 in Chapter 4. Assume $\alpha = 10^{-8}$.

(a) Use a computer implementation of the Euler method to simulate the behavior of this model, starting with the initial conditions $x_1 = 5,000$ blue whales and $x_2 = 70,000$ fin whales. Perform a sensitivity analysis on both T and N to ensure the validity of your results, as in the text. What happens to the two species of whales over the long term, according to this model? Do both species of whales grow back, or will one or both species become extinct? How long does this take?

(b) Repeat part (a) for a range of initial conditions for both blue and fin whales. Tabulate the results of your simulations, and answer the same questions as in part (a) for each case.

(c) Use the results of parts (a) and (b) in order to draw the complete phase portrait for this system.

(d) Indicate the region on the phase portrait where one or both species of whale are destined to become extinct.

15. Reconsider the RLC circuit problem of Example 6.3, and perform a sensitivity analysis on the parameter L, which represents inductance.

(a) Generalize the dynamical system model in Eq. (6.8) to represent the case $L > 0$. How does the vector field for this model vary with L?

(b) Determine the form of the linear system that approximates the behavior of the nonlinear RLC circuit model in the neighborhood of the origin. Calculate the eigenvalues for the linear system as a function of L. Determine the range of L over which both eigenvalues are complex with positive real part, as in the baseline case $L = 1$.

(c) Use a computer implementation of the Euler method to simulate the behavior of the RLC circuit for each of the cases $L = 0.5, 0.75, 1.5$, and 2.0. Use the same initial condition $x_1 = 0.1$, $x_2 = 0.3$ as in Fig. 6.21. For each case, perform a sensitivity analysis on both T and N to ensure the validity of your results, as in the text.

(d) Simulate several additional initial conditions for each value of L specified in part (c). Draw the complete phase portrait for each case. Describe how the phase portrait changes in response to changes in the inductance L.

16. Reconsider the RLC circuit problem of Example 6.3, and now consider what happens in the case of a large capacitance, $C > 4$.

(a) Solve the linear system in Eq. (6.10) by the method of eigenvalues and eigenvectors in the case $C > 4$.

(b) Draw the phase portrait for this linear system. How does the phase portrait change as a function of C?

(c) Use a computer implementation of the Euler method to simulate the RLC circuit for each of the cases $C = 5$, 6, 8, and 10. Use the same initial condition $x_1 = 0.1$, $x_2 = 0.3$ as in Fig. 6.21. For each case, perform a sensitivity analysis on both T and N to ensure the validity of your results, as in the text.

(d) Simulate several additional initial conditions for each value of C specified in part (c). Draw the complete phase portrait for each case. Contrast with the case $0 < C < 4$ discussed in the text. What changes occur in the phase portrait as we transition between the two cases?

17. Reconsider the RLC circuit problem of Example 6.3, and now consider the robustness of our general conclusions with respect to the assumption that the resistor in this RLC circuit has v–i characteristic $f(x) = x^3 - x$. In this problem we will assume that $f(x) = x^3 - ax$, where the parameter a may represent any positive real number. (The case $a = -4$ was the subject of Example 5.3.)

(a) Generalize the dynamical system model in Eq. (6.8) to represent the general case $a > 0$. How does the vector field for this model vary with a?

(b) Determine the form of the linear system that approximates the behavior of the nonlinear RLC circuit model in the neighborhood of the origin. Calculate the eigenvalues for the linear system as a function of a.

(c) Draw the phase portrait for this linear system. How does the phase portrait change as a function of a?

(d) Use a computer implementation of the Euler method to simulate the RLC circuit for each of the cases $a = 0.5$, 0.75, 1.5, and 2.0, and draw the complete phase portrait for each case. What changes occur in the phase portrait as we change a? What do you conclude about the robustness of this model?

18. Reconsider the RLC circuit problem of Example 6.3, and now consider the robustness of our general conclusions with respect to the assumption that the resistor in this RLC circuit has v–i characteristic $f(x) = x^3 - x$. In this problem we will assume that $f(x) = x|x|^{1+b} - x$, where $b > 0$.

   (a) Generalize the dynamical system model in Eq. (6.8) to represent the general case $b > 0$. How does the vector field for this model vary with b?

   (b) Determine the form of the linear system that approximates the behavior of the nonlinear RLC circuit model in the neighborhood of the origin. Calculate the eigenvalues for the linear system as a function of b.

   (c) Draw the phase portrait for this linear system. How does the phase portrait change as a function of b?

   (d) Use a computer implementation of the Euler method to simulate the RLC circuit for each of the cases $b = 0.5$, 0.75, 1.25, and 1.5. Draw the complete phase portrait for each case. What changes occur in the phase portrait as we change b? What do you conclude about the robustness of this model?

19. A pendulum consists of a 100 g weight at the end of a lightweight rod 120 cm in length. The other end of the rod is fixed, but can rotate freely. The frictional forces acting on the moving pendulum are thought to be roughly proportional to its angular velocity.

   (a) The pendulum is lifted manually until the rod makes a 45^0 angle with the vertical. Then the pendulum is released. Determine the subsequent motion of the pendulum. Use the five-step method, and model as a continuous–time dynamical system. Simulate using the Euler method. Assume that the force due to friction is of magnitude $k\theta'$, where θ' is the angular velocity in radians per second and the coefficient of friction is $k = 0.05$ g/sec.

   (b) Use a linear approximation to determine the approximate behavior of the system near equilibrium. Assume that the magnitude of the frictional force is $k\theta'$. How does the local behavior depend on k?

   (c) Determine the period of the pendulum. How does period vary with k?

   (d) This size pendulum will be used as part of the mechanism for a grandfather clock. In order to maintain a certain amplitude of oscillation, a force is to be applied periodically. How much force should be applied, and how often, to produce an amplitude of $\pm 30^0$? How does the answer depend on the amplitude desired? [Hint: Simulate one period of the pendulum oscillation. Vary the initial angular velocity $\theta'(0)$ to obtain periodic behavior.]

20. (Chaos) This problem illustrates the striking difference between the behavior of continuous–time and discrete–time dynamical systems that can occur even in simple models.

   (a) Show that the continuous–time dynamical system

   $$x_1' = (a - 1)x_1 - ax_1^2$$
   $$x_2' = x_1 - x_2$$

   has a stable equilibrium at $x_1 = x_2 = (a - 1)/a$ for any $a > 1$.

   (b) Show that the analogous discrete–time dynamical system

   $$\Delta x_1 = (a - 1)x_1 - ax_1^2$$
   $$\Delta x_2 = x_1 - x_2$$

   also has an equilibrium at $x_1 = x_2 = (a - 1)/a$ for any $a > 1$.

   (c) Use a simulation to explore the stability of the equilibrium $x_1 = x_2 = (a - 1)/a$ and the behavior of nearby solutions for the discrete–time dynamical system. For each of the cases $a = 1.5, 2.0, 2.5, 3.0, 3.5$, and 4.0, try several different initial conditions near the equilibrium point and report what you see. (The case $a = 4.0$ represents a simple model of chaos, the apparently random behavior of a deterministic dynamical system.)

21. (Programming exercise) An alternative method that can be used to simulate dynamical systems is the Runge–Kutta method.

   Figure 6.39 gives an algorithm for the Runge–Kutta method to simulate a dynamical system in two variables,

   $$\frac{dx_1}{dt} = f_1(x_1, x_2)$$
   $$\frac{dx_2}{dt} = f_1(x_1, x_2).$$

   For a fairly small step size h, Runge–Kutta has the property that doubling the number of steps (halving h) produces results approximately 16 times more accurate.

   (a) Implement the Runge–Kutta method on a computer.

   (b) Verify your computer implementation by using it to solve the linear system given by Eq. (5.18) in Chapter 5. Compare your results to the analytic solution in Eq. (5.19) for the case $c_1 = 1$, $c_2 = 0$.

   (c) Verify the results obtained in Figs. 6.20 and 6.21 for the RLC circuit problem of Example 6.3.

22. Reconsider the whale problem of Example 6.4. In this problem we will explore the behavior of the blue whale population. Assume that $\alpha = 10^{-8}$ and begin with $x_1 = 5,000$ blue whales and $x_2 = 70,000$ fin whales.

Algorithm: RUNGE–KUTTA METHOD

Variables: $t(n) =$ time after n steps
$x_1(n) =$ first state variable at time $t(n)$
$x_2(n) =$ second state variable at time $t(n)$
$N =$ number of steps
$T =$ time to end simulation

Input: $t(0), x_1(0), x_2(0), N, T$

Process: Begin
$h \leftarrow (T - t(0))/N$
for $n = 0$ to $N - 1$ do
Begin
$r_1 \leftarrow f_1(x_1(n), x_2(n))$
$s_1 \leftarrow f_2(x_1(n), x_2(n))$
$r_2 \leftarrow f_1(x_1(n) + (h/2)r_1, x_2(n) + (h/2)s_1)$
$s_2 \leftarrow f_2(x_1(n) + (h/2)r_1, x_2(n) + (h/2)s_1)$
$r_3 \leftarrow f_1(x_1(n) + (h/2)r_2, x_2(n) + (h/2)s_2)$
$s_3 \leftarrow f_2(x_1(n) + (h/2)r_2, x_2(n) + (h/2)s_2)$
$r_4 \leftarrow f_1(x_1(n) + hr_3, x_2(n) + hs_3)$
$s_4 \leftarrow f_2(x_1(n) + hr_3, x_2(n) + hs_3)$
$x_1(n + 1) \leftarrow x_1(n) + (h/6)(r_1 + 2r_2 + 2r_3 + r_4)$
$x_2(n + 1) \leftarrow x_2(n) + (h/6)(s_1 + 2s_2 + 2s_3 + s_4)$
$t(n + 1) \leftarrow t(n) + h$
End
End

Output: $t(1), \ldots, t(N); x_1(1), \ldots, x_1(N); x_2(1), \ldots, x_2(N)$

Figure 6.39: Pseudocode for the Runge–Kutta method.

(a) Use the Euler method with a time step of $h = \Delta t = 1$ year. Simulate $N = 50$ time steps and describe the behavior of the blue whale population over time.

(b) Repeat the simulation of part (a) with $N = 50$ time steps for each of the cases $h = 5, 10, 20, 30, 35$, and 40. How does the behavior of the blue whale population change as the time step h increases?

(c) Repeat part (b) for $\alpha = 10^{-7}$ and $\alpha = 10^{-9}$. How sensitive are your conclusions in part (b) to the assumption that $\alpha = 10^{-8}$?

(d) Repeat part (b) starting from the initial condition $x_1 = 150,000$ blue whales and $x_2 = 400,000$ fin whales, and assume that $\alpha = 10^{-8}$. How sensitive are your conclusions in part (b) to the assumption that we start with $x_1 = 5,000$ blue whales and $x_2 = 70,000$ fin whales?

23. Reconsider the whale problem of Example 6.4. In this problem we will explore the sensitivity to initial conditions for a chaotic dynamical system. Assume that $\alpha = 10^{-8}$ and that we begin with $x_2(0) = 70,000$ fin whales.

    (a) Use the Euler method with a time step of $h = \Delta t = 35$ years. Simulate to determine the number of blue whales $x_1(T)$ that remain after $T = 1750$ years ($N = 50$ time steps), using the initial condition $x_1(0) = 5,000$ blue whales.

    (b) Repeat the simulation of part (a) using the initial condition $x_1(0) = 5050$ and determine the resulting blue whale population $x_1(T)$ after $T = 1750$ years. Compare to the results of part (a) and compute the sensitivity of the eventual population level to the initial condition. Note that the relative change in the initial condition is $\Delta x_1(0)/x_1(0) = 0.01$ and the relative change in the eventual population level is $\Delta x_1(T)/x_1(T)$.

    (c) Repeat part (b) for each of the initial conditions $x_1 = 5005, 5000.5,$ 5000.05, 5000.005 and comment on the relation between the sensitivity and the difference $\Delta x_1(0)$ in the initial condition.

    (d) How sensitive is this chaotic dynamical system to small changes in the initial condition? If we estimate the current state of such a system, can we reliably predict its future?

24. Reconsider the whale problem of Example 6.4. In this problem we will explore the transition from stability to instability in the discrete approximation of Eq. (6.18) as the step size $h = \Delta t$ increases. Assume that $\alpha = 10^{-8}$.

    (a) Compute the coordinates of the equilibrium in the positive first quadrant for the continuous time dynamical system of Eq. (6.17). Use the eigenvalue test for continuous time dynamical systems to show that this equilibrium is stable.

    (b) Explain why the iteration function for the discrete approximation is given by $G(x) = x + hF(x)$ where $h = \Delta t$. Write down the iteration function for the discrete time dynamical system of Eq. (6.18).

    (c) Write down the matrix of partial derivatives $A = DG$ evaluated at the equilibrium point found in part (a), and compute the eigenvalues of this matrix as a function of the step size h.

    (d) Use the eigenvalue test for discrete time dynamical systems to determine the largest step size h for which the equilibrium found in part (a) remains stable in the discrete approximation. Compare with the results in the text.

25. Reconsider the whale problem of Example 6.4. In this problem we will use simulation to explore the fractal limit sets in the discrete approximation of Eq. (6.18) for different step sizes $h = \Delta t$.

(a) Use a computer implementation of the Euler method to reproduce the results shown in Figure 6.31 in the text. Assume that $\alpha = 10^{-8}$ and use a step size of $h = \Delta t = 32$ years with initial conditions $x_1(0) = 5,000$ blue whales and $x_2(0) = 70,000$ fin whales.

(b) Plot fin whales $x_2(n)$ versus blue whales $x_1(n)$ for $n = 100, \ldots, 1000$. Your graph should show the limit set consisting of four points.

(c) Repeat the simulation of part (a) for step size $h = 33, 34, \ldots, 37$. For each case, plot the limit set as in part (b). How does the limit set change as the step size increases?

(d) Repeat part (c) for initial condition $x_1(0) = 150,000$ blue whales and $x_2(0) = 400,000$ fin whales. Does the limit set depend on the initial condition?

(e) Repeat part (c) for $\alpha = 3 \times 10^{-8}$. Does the limit set depend on the competition parameter α?

26. Reconsider the weather problem of Example 6.6.

(a) Use a computer implementation of the Euler method to reproduce the results of Figure 6.33 in the text. Assume $\sigma = 10$, $b = 8/3$, $r = 8$, and use the initial condition $(x_1, x_2, x_3) = (1, 1, 1)$.

(b) Use the results of part (a) to plot the deviation x_3 from linearity of the temperature profile versus the rate x_1 at which the convection rolls rotate. Perform a sensitivity analysis on the step size h to ensure that your plot represents the true behavior of the continuous time dynamical system.

(c) Repeat part (b) for the initial condition $(x_1, x_2, x_3) = (7, 1, 2)$. Does the solution curve approach the equilibrium found in the text?

(d) Repeat part (b) for $r = 18$ and $r = 28$. How does the solution behavior change as r increases?

27. Reconsider the weather problem of Example 6.6.

(a) Use a computer implementation of the Euler method with $N = 500$ and $T = 2.5$ (step size $h = 0.005$) to reproduce the results of Figure 6.35 in the text. Assume $\sigma = 10$, $b = 8/3$, $r = 18$, and use the initial condition $(x_1, x_2, x_3) = (6.7, 6.7, 17)$.

(b) Repeat part (a) for larger step sizes $h = 0.01, 0.015, \ldots, 0.03$. You can keep $N = 500$ and increase $T = 5, 7.5, 10, 12.5$, and 15. How does the simulated solution curve change as the step size increases?

(c) Find the coordinates of the equilibrium point E^+ when $\sigma = 10$, $b = 8/3$, and $r = 24$. Verify by simulation. Begin the simulation at the initial condition E^+ and check that the solution remains at this point.

(d) Is the equilibrium point E^+ found in part (c) stable? Verify by simulation. Begin the simulation at the initial condition $E^+ + (0.1, 0.1, 0)$ and determine whether the solution tends toward the equilibrium point E^+. How small a step size h must be used to ensure that the simulation results represent the true behavior of the continuous time dynamical system?

Further Reading

1. Acton, F. (1970) *Numerical Methods That Work*. Harper and Row, New York.

2. Brams, S., Davis, M. and Straffin, P. *The Geometry of the Arms Race*. UMAP module 311.

3. Dahlquist, G. and Bjorck, A. *Numerical Methods*. Prentice-Hall, Englewood Cliffs, New Jersey.

4. Gearhart, W. and Martelli, M. *A Blood Cell Population Model, Dynamical Diseases, and Chaos*. UMAP module 709.

5. Gleick, J. (1987) *Chaos: Making a New Science*. R. R. Donnelley, Harrisonburg, Virginia.

6. Press, W., Flannery, B., Teukolsky, S. and Vetterling, W. (1987). *Numerical Recipies*. Cambridge University Press, New York.

7. Smith, H. *Nuclear Deterrence*. UMAP module 327.

8. Strogatz, S. (1994) *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering*. Addison Wesley, Reading, Massachusetts.

9. Zinnes, D., Gillespie, J. and Tahim, G. *The Richardson Arms Race Model*. UMAP module 308.

Chapter 7

INTRODUCTION TO PROBABILITY MODELS

Most real–life problems contain elements of uncertainty. In some models we may introduce random elements to account for uncertainties in human behavior. In other models we may be unsure of the exact physical parameters of a system, or we may be unsure of the exact physical laws that govern its dynamics. It has even been suggested in some cases that physical parameters and physical laws are essentially random; for example, in quantum mechanics. Sometimes probabilities are introduced into a model as a matter of convenience, sometimes as a matter of necessity. In either case, it is here in the realm of probability that mathematical modeling becomes most interesting and useful.

Probability is a familiar and intuitive idea. In this chapter we begin our treatment of probability models. We do not assume any prior background in formal probability theory. We will introduce the basic concepts of probability here in a natural way as they emerge in the study of real problems.

7.1 Discrete Probability Models

The most simple and intuitive probability models are those involving a discrete set of possible outcomes and no time dynamic elements. Such models are frequently encountered in the real world.

Example 7.1. An electronics manufacturer produces a variety of diodes. Quality control engineers attempt to insure that faulty diodes will be detected in the factory before they are shipped. It is estimated that 0.3% of the diodes produced will be faulty. It is possible to test each diode individually. It is also possible to place a number of diodes in series and test the entire group. If this test fails, it means that one or more of the diodes in that group are faulty. The estimated testing cost is 5 cents for a single diode, and $4+n$ cents for a group of $n > 1$ diodes. If a group test fails, then each diode in the group must be retested

Variables: n = number of diodes per test group
                   C = testing cost for one group (cents)
                   A = average testing cost (cents/diode)

Assumptions: If $n = 1$, then $A = 5$ cents
                   Otherwise ($n > 1$), we have $C = 4 + n$
                   if the group test indicates that all
                   diodes are good, and $C = (4 + n) + 5n$
                   if the group test indicates a failure.
                   $A = $ (Average value of C)/n

Objective: Find the value of n that minimizes A

Figure 7.1: Results of step 1 for the diode problem.

individually to find the bad one(s). Find the most cost–effective quality control procedure for detecting bad diodes.

We will use the five-step method. The results of step 1 are summarized in Figure 7.1. The variable n is a decision variable, and we are free to choose any $n = 1, 2, 3, \ldots$, but the variable C is the random outcome of the quality control procedure we select. We say C is a random variable. The quantity A, however, is not random. It represents the average or expected value of the random variable C/n.

Step 2 is to select the modeling approach. We will use a discrete probability model.

Consider a random variable X, which can take any of a discrete set of values

$$X \in \{x_1, x_2, x_3, \ldots\},$$

and suppose that $X = x_i$ occurs with probability p_i. We will write

$$\Pr\{X = x_i\} = p_i.$$

Of course, we must have

$$\Sigma p_i = 1.$$

Since X takes the value x_i with probability p_i, the average or expected value of X should be a weighted average of the possible values x_i, weighted according to their relative likelihoods p_i. We will write

$$EX = \Sigma x_i p_i. \tag{7.1}$$

The probabilities p_i represent what we will call the probability distribution of X.

Example 7.2. In a simple game of chance, two dice are rolled and the bank pays the player the number of dollars shown on the dice. How much would you pay to play this game?
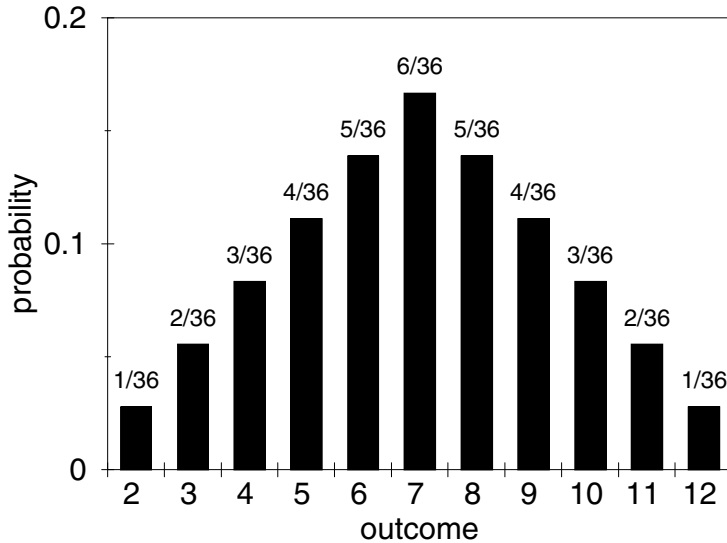
Figure 7.2: Histogram of probability versus outcome showing the distribution of the sum of two dice.

Let X denote the number shown on the dice. There are $6 \times 6 = 36$ possible outcomes, and each is equally likely. There is only one way to roll a 2, so

$$\Pr\{X = 2\} = 1/36.$$

There are two ways to roll a 3 (1 and 2, or 2 and 1), so

$$\Pr\{X = 3\} = 2/36.$$

The complete probability distribution of X is illustrated in Figure 7.2. The expected value of X is

$$EX = 2(1/36) + 3(2/36) + \cdots + 12(1/36),$$

or $EX = 7$. After many repetitions of this game, you would expect to win about seven dollars per roll. Therefore, it would be worthwhile to play the game if it cost no more than seven dollars to play.

To be more specific, suppose that you play the game over and over. Let X_n denote the amount you win on the nth roll of the dice. Each X_n has the same distribution, and the different X_n are independent. The amount won on one roll does not depend on the amount won in the previous roll. There is a theorem called "the strong law of large numbers," which says that for any sequence of

independent, identically distributed random variables X, X_1, X_2, $X_3, \ldots$ with EX finite, we will have

$$\frac{X_1 + \cdots + X_n}{n} \to EX \qquad (7.2)$$

as $n \to \infty$ with probability 1. In other words, if you play the game for a long time, you are virtually certain to win about \$7 per roll (Ross (1985) p. 70).

The formal definition of *independence* is as follows. Let Y and Z denote two random variables with

$$Y \in \{y_1, y_2, y_3, \ldots\}$$

and

$$Z \in \{z_1, z_2, z_3, \ldots\}.$$

We say that Y and Z are independent if it is generally true that

$$\Pr\{Y = y_i \text{ and } Z = z_j\} = \Pr\{Y = y_i\} \Pr\{Z = z_j\}. \qquad (7.3)$$

For example, let Y and Z denote the number on the first and second dice, respectively. Then

$$\Pr\{Y = 2, Z = 1\} = \Pr\{Y = 2\} \Pr\{Z = 1\} = (1/6)(1/6) = (1/36),$$

and likewise for each possible outcome. The random variables Y and Z are independent. The number that comes up on the second die has nothing to do with what happened on the first.

Returning to the diode problem of Example 7.1, we see that the random variable C takes on one of two possible values for any fixed $n > 1$. If all the diodes are good, then

$$C = 4 + n.$$

Otherwise,

$$C = (4 + n) + 5n,$$

since we have to retest each diode. Letting p denote the probability that all the diodes are good, the remaining possibility (one or more bad diodes) must have probability $1 - p$. Then the average or expected value of C is

$$EC = (4 + n)p + [(4 + n) + 5n](1 - p). \qquad (7.4)$$

Now for step 4. There are n diodes, and the probability that one individual diode is bad is 0.003. In other words, the probability that one individual diode is good is 0.997. Assuming independence, it follows that the probability that all n diodes in one test group are good is $p = 0.997^n$.

The expected value of the random variable C is

$$EC = (4 + n)0.997^n + [(4 + n) + 5n](1 - 0.997^n)$$
$$= (4 + n) + 5n(1 - 0.997^n)$$
$$= 4 + 6n - 5n(0.997)^n.$$

Hence, the average testing cost per diode is

$$A = \frac{4}{n} + 6 - 5(0.997)^n. \tag{7.5}$$

The strong law of large numbers tells us that this formula represents the long–run average cost we will experience if we use test groups of size n. Now all we need is to minimize A as a function of n. We leave the details to the reader (see Exercise 1). The minimum of $A = 1.48$ cents/diode occurs at $n = 17$.

We conclude with step 5. Quality control procedures for detecting faulty diodes can be made considerably more economical by group testing methods. Individual testing costs approximately 5 cents/unit. Bad diodes occur only rarely, at a rate of 3 per 1,000. By testing groups of 17 diodes each, in series, we can reduce testing costs by a factor of three (to 1.5 cents/diode) without sacrificing quality.

Sensitivity analysis is critical in this type of problem. The implementation of a quality control procedure will depend on several factors outside the scope of our model. It may be easier to test diodes in batches of 10 or 20, or perhaps n should be a multiple of 4 or 5, depending on the details of our manufacturing process. Fortunately, the average cost A does not vary significantly between $n = 10$ and $n = 35$. Again, we leave the details to the reader. The parameter $q = 0.003$, which represents the failure rate in the manufacturing process, must also be considered. For example, this value may vary with the environmental conditions inside the plant. Generalizing on our previous model, we have

$$A = \frac{4}{n} + 6 - 5(1 - q)^n. \tag{7.6}$$

At $n = 17$ we have

$$S(A, q) = \frac{dA}{dq} \cdot \frac{q}{A} = 0.16,$$

so small variations in q are not likely to affect our cost very much.

A more general robustness analysis would consider the assumption of independence. We have assumed that there is no correlation between the times of successive failures in the manufacturing process. It may be, in fact, that bad diodes tend to be produced in batches, perhaps due to a passing anomaly in the manufacturing environment, such as a vibration or a power surge. The mathematical analysis of dependent random variable models cannot be dealt with in its entirety here. The stochastic process models introduced in the next chapter are capable of representing some kinds of dependence, while some other types of dependence admit no tractable analytic formulations. Problems in robustness

Variables: $\lambda =$ decay rate (per second)
                        $T_n =$ time of nth observed decay

Assumptions: Radioactive decays occur at random
                        with rate λ. $T_{n+1} - T_n \geq 3 \times 10^{-9}$ for all n

Objective: Find λ on the basis of a finite number
                        of observations $T_1, \ldots, T_n$

Figure 7.3: Results of step 1 of the radioactive decay problem.

are very much an active and intriguing branch of current research in probability theory. In practice, simulation results tend to indicate that expected value models based on independent random variables are quite robust. More importantly, it has been found through experience that such models provide useful, accurate approximations of real–life behavior in most cases.

7.2   Continuous Probability Models

In this section we consider probability models based on random variables that take values over a continuum. Such models are particularly convenient for representing random times. The mathematical theory required is completely analogous to the discrete case, except that now integrals replace sums.

Example 7.3. A "type I counter" is used to measure the radioactive decay in a sample of fissionable material. Decays occur at random, at an unknown rate, and the purpose of the counter is to measure the decay rate. Each radioactive decay locks the counter for a period of 3×10^{-9} seconds, during which time any decays that occur are not counted. How should the data received from the counter be adjusted to account for the lost information?

We will use the five-step method. The results of step 1 are summarized in Figure 7.3. Step 2 is to select the modeling approach. We will use a continuous probability model.

> Suppose that X is a random variable that takes values on the real line. A convenient way to describe the probability structure of X is to specify the function
> $$F(x) = \Pr\{X \leq x\},$$
> called the *distribution function* of X. If $F(x)$ is differentiable, we call the function
> $$f(x) = F'(x)$$
> the *density function* of X. Then for any real numbers a and b we have
> $$\Pr\{a < X \leq b\} = F(b) - F(a) = \int_a^b f(x)\,dx. \qquad (7.7)$$

In other words, the area under the density curve represents probability. The mean or expected value of X is defined by

$$EX = \int_{-\infty}^{\infty} x f(x)\, dx, \tag{7.8}$$

which is directly analogous to the discrete case, as can be seen by considering the Riemann sum for the integral. (See Exercise 13 for details.) It is also worthwhile pointing out that this notation and terminology were originally adapted from a problem in physics, namely the center of mass problem. If a wire or rigid rod is laid out along the x axis, and $f(x)$ represents the density (gms/cm) at point x, then the integral of $f(x)$ represents mass, and the integral of $xf(x)$ represents the center of mass (assuming, as we do in probability, that the total mass is equal to 1).

The special case of random arrivals occurs frequently in applications. Suppose that arrivals (e.g., customers, phone calls, radioactive decays) occur at random with rate λ, and let X denote the random time between two successive arrivals. It is common to assume that X has the distribution function

$$F(t) = 1 - e^{-\lambda t}, \tag{7.9}$$

so that the density function of X is

$$f(t) = \lambda e^{-\lambda t}. \tag{7.10}$$

This distribution is called the *exponential distribution* with rate parameter λ.

One very important property of the exponential distribution is its "lack of memory". For any $t > 0$ and $s > 0$, we have

$$\begin{aligned}
\Pr\{X > s + t \mid X > s\} &= \frac{\Pr\{X > s + t\}}{\Pr\{X > s\}} \\
&= \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} \qquad (7.11) \\
&= \Pr\{X > t\}.
\end{aligned}$$

In other words, the fact that we have already waited s units of time for the next arrival does not affect the (conditional) distribution of the time until the next arrival. The exponential distribution "forgets" that we have already waited this long. The probability in Eq. (7.11) is called a *conditional probability*. Formally, the probability of event A occurring, given the event B occurs, is

$$\Pr\{A \mid B\} = \frac{\Pr\{A \text{ and } B\}}{\Pr\{B\}}. \tag{7.12}$$

In other words, $\Pr\{A \mid B\}$ is the relative likelihood of A among all possible events once B has occurred.

Now we proceed to step 3, the model formulation. We are assuming that radioactive decays occur at random at an unknown rate λ. We will model this process by assuming that the times between successive radioactive decays are independent and identically distributed with an exponential distribution with rate parameter λ. Let

$$X_n = T_n - T_{n-1}$$

denote the times between successive observations of a radioactive decay. Of course, X_n does not have the same distribution as the time between successive decays, because of the lock time. In fact, $X_n \geq 3 \times 10^{-9}$ with probability 1, which is certainly not true for the exponential distribution.

The random time X_n consists of two parts. First, we must wait $a = 3 \times 10^{-9}$ seconds while the counter is locked, and then we must wait an additional Y_n seconds until the next decay. Now Y_n is not simply the time between two decays, because it begins at the end of the lock time, not at a decay time. However, the memoryless property of the exponential distribution guarantees that Y_n is still exponential with rate parameter λ.

Step 4 is to solve the model. Since $X_n = a + Y_n$, we have $EX_n = a + EY_n$, where

$$EY_n = \int_0^\infty t\lambda e^{-\lambda t}\, dt.$$

Integrate by parts to find $EY_n = 1/\lambda$. Thus, $EX_n = a + 1/\lambda$. The strong law of large numbers says that

$$\lim_{n\to\infty} \frac{X_1 + \cdots + X_n}{n} = a + \frac{1}{\lambda}$$

with probability 1. In other words, $(T_n/n) \to a + 1/\lambda$. For large n it will be approximately true that

$$\frac{T_n}{n} = a + \frac{1}{\lambda}. \tag{7.13}$$

Solving for λ, we obtain

$$\lambda = \frac{n}{T_n - na}. \tag{7.14}$$

Finally, we conclude with step 5. We have obtained a formula for decay rate that corrects for the decays missed while the counter is locked. All that is required is to record the length of observation and the number of decays recorded. The distribution of those decays in the observation interval is not required to determine λ.

Sensitivity analysis should consider the lock time a, which must be determined empirically. The accuracy to which we can determine a will affect the accuracy of λ. From Eq. (7.14) we calculate that

$$\frac{d\lambda}{da} = \lambda^2.$$

The sensitivity of λ to a is then

$$S(\lambda, a) = \lambda^2(a/\lambda) = \lambda a.$$

This is also the expected number of decays during the lock time. We can therefore get a better estimate of λ (in relative terms) for a less intensely radioactive source. One simple way to achieve this is to use fewer grams of radioactive material in our sample. Another important source of potential error comes from the assumption that

$$(X_1 + \cdots + X_n)/n = a + 1/\lambda.$$

Of course, this is not exactly true. Random fluctuations will cause the empirical rate to vary from the mean, although we do have convergence as $n \to \infty$. The study of such random fluctuations is the subject of the next section.

Finally, there is the matter of robustness. We have made an assumption about the decay process that appears to be very special. We have assumed that times between decays are independent and that they have a particular distribution (exponential with rate parameter λ). Such an arrival process is called a *Poisson process*. The Poisson process is commonly used to represent random arrivals. Its use can be justified in part by the fact that many real–world arrival processes have interarrival times that are at least approximately exponential. This can be verified by collecting data on arrival times. But this does not answer the question of *why* the exponential distribution occurs.

It turns out that there is a mathematical reason for expecting an arrival process to look Poisson. Consider a large number of arrival processes that are independent of one another. We make no assumption about the interarrival time distribution of an arrival process, only that the interarrival times are independent and identically distributed. There is a theorem that states, under fairly general conditions, that the arrival process obtained by merging all of these independent processes has to look Poisson. (The merged process tends to Poisson as the number of merged processes tends to infinity.) This is why the Poisson process, based on the exponential distribution, is such a robust model (see Feller (1971) p. 370).

7.3 Introduction to Statistics

In any modeling situation it is desirable to get quantitative measures of performance. For probability models an additional complication is involved in deriving such parameters of system behavior. We must have a way to deal with the random fluctuations in system behavior that are characteristic of probability models. Statistics is the study of measurement in the presence of random fluctuations. The appropriate use of statistical methods must be a part of the analysis of any probability model.

Example 7.4. An emergency 911 service in a local community received an average of 171 calls per month for house fires over the past year. On the basis of this data, the rate of house fire emergencies was estimated at 171 per month. The next month only 153 calls were received. Does this indicate an actual reduction in the rate of house fires, or is it simply a random fluctuation?

Variables: $\lambda =$ Rate of house fire reports (per month)

$X_n =$ time between $(n-1)$st and nth fire (months)

Assumptions: House fires occur at random with rate λ;

i.e., X_1, X_2,... are independent, and each X_n has

an exponential distribution with rate parameter λ

Objective: Determine the probability that as few as 153 calls

would be received in one month, given $\lambda = 171$

Figure 7.4: Results of step 1 of the house fire problem.

We will use the five-step method. The results of step 1 are summarized in Figure 7.4. We are assuming exponential interarrival times for the emergency calls. Step 2 is to determine the modeling approach. We will model this as a statistical inference problem.

Suppose that X, X_1, X_2, X_3,..., are independent random variables, all with the same distribution. Recall that if X is discrete, the average or expected value is

$$EX = \Sigma\, x_k \Pr\{X = x_k\},$$

and if X is continuous with density $f(x)$, then

$$EX = \int xf(x)\,dx.$$

Another distributional parameter, called the *variance*, measures the extent to which X tends to deviate from the mean EX. In general we define

$$VX = E(X - EX)^2. \tag{7.15}$$

If X is discrete, we have

$$VX = \Sigma(x_k - EX)^2 \Pr\{X = x_k\}, \tag{7.16}$$

and if X is continuous with density $f(x)$, we have

$$VX = \int (x - EX)^2 f(x)\,dx. \tag{7.17}$$

There is a result called the *central limit theorem*, that states that as $n \to \infty$, the distribution of the sum $X_1 + \cdots + X_n$ gets closer and closer to a certain type of distribution called a *normal distribution*. Specifically, if we let $\mu = EX$ and $\sigma^2 = VX$, then for all t real we have

$$\lim_{n\to\infty} \Pr\left\{\frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le t\right\} \to \Phi(t) \tag{7.18}$$
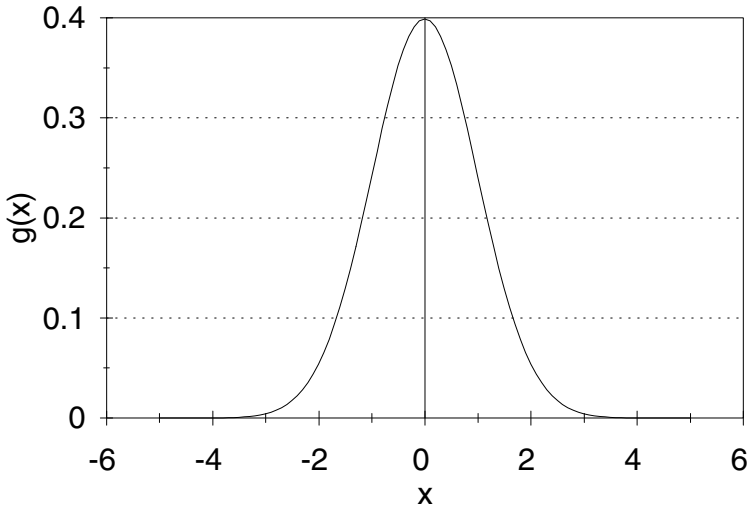
Figure 7.5: Graph of the standard normal density function (7.19).

where $\Phi(t)$ is a special distribution function called the *standard normal distribution*. The density function for the standard normal distribution is defined for all x by

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \tag{7.19}$$

so that, for all t, we have

$$\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx. \tag{7.20}$$

Figure 7.5 shows a graph of the standard normal density. Numerical integration shows that the area between $-1 \le x \le 1$ is approximately 0.68, and the area between $-2 \le x \le 2$ is approximately 0.95. Thus, for all n sufficiently large, we will have

$$-1 \le \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le 1$$

about 68% of the time, and

$$-2 \le \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \le 2 \tag{7.21}$$

about 95% of the time. In other words, we are 68% sure that

$$n\mu - \sigma\sqrt{n} \le X_1 + \cdots + X_n \le n\mu + \sigma\sqrt{n},$$

and 95% sure that

$$n\mu - 2\sigma\sqrt{n} \le X_1 + \cdots + X_n \le n\mu + 2\sigma\sqrt{n}. \qquad (7.22)$$

It is common in practice to accept the 95% interval from Eq. (7.22) as the range of normal variation in a random sample. In cases where the sum $X_1 + \cdots + X_n$ does not lie in the interval from Eq. (7.22), we say that the deviation is *statistically significant* at the 95% level.

Now we move on to step 3, the model formulation. We are assuming that the times between calls, X_n, are exponentially distributed, with density function

$$f(x) = \lambda e^{-\lambda x}$$

on $x \ge 0$. We have previously calculated that

$$\mu = EX_n = 1/\lambda.$$

The variance

$$\sigma^2 = VX_n$$

is given by

$$\sigma^2 = \int_0^\infty (x - 1/\lambda)^2 \lambda e^{-\lambda x}\, dx,$$

and the central limit theorem gives probability estimates of the extent to which

$$(X_1 + \cdots + X_n)$$

can vary from its mean, n/λ. In particular, we know that Eq. (7.22) holds with probability 0.95.

On to step 4. We calculate that

$$\sigma^2 = 1/\lambda^2,$$

using integration by parts. Substituting $\mu = 1/\lambda$ and $\sigma = 1/\lambda$ into Eq. (7.22), we find that the relation

$$\frac{n}{\lambda} - \frac{2\sqrt{n}}{\lambda} \le X_1 + \cdots + X_n \le \frac{n}{\lambda} + \frac{2\sqrt{n}}{\lambda} \qquad (7.23)$$

must hold with probability 0.95. Substituting $\lambda = 171$ and $n = 153$ into Eq. (7.23), we are 95% sure that

$$\frac{153}{171} - \frac{2\sqrt{153}}{171} \le X_1 + \cdots + X_{153} \le \frac{153}{171} + \frac{2\sqrt{153}}{171},$$

or in other words,

$$0.75 \le X_1 + \cdots + X_{153} \le 1.04.$$

Therefore, our observation that

$$X_1 + \cdots + X_{153} \approx 1$$

is within the range of normal variation.

Finally, step 5. There is insufficient evidence to conclude that the rate of house fire emergency calls has declined. The variation in the observed number of calls may well be the result of normal random variation. Of course, if the observed number of calls per month continues to be this low, then we would reassess the situation.

A few items should be included in our sensitivity analysis. First of all, we have concluded that 153 calls in one month is within the range of normal variation. More generally, suppose that n calls are received in one month. Substituting $\lambda = 171$ into Eq. (7.23), we conclude that

$$\frac{n}{171} - \frac{2\sqrt{n}}{171} \leq X_1 + \cdots + X_n \leq \frac{n}{171} + \frac{2\sqrt{n}}{171} \tag{7.24}$$

with probability 0.95. Since the interval

$$\frac{n}{171} \pm \frac{2\sqrt{n}}{171}$$

contains 1 for any value of $n \in [147, 199]$, we would conclude more generally that 95% of the time there will be between 147 and 198 calls in a month. In other words, the range of normal variation for this community is 147 to 198 emergency calls per month.

Now let us consider the sensitivity of our conclusions to the assumption that the actual expected number of emergency calls is 171 per month. Less specifically, assume that there is an average of λ emergency calls per month. We have observed $n = 153$ calls in a one–month period. Substituting into Eq. (7.23), we conclude that

$$\frac{153}{\lambda} - \frac{2\sqrt{153}}{\lambda} \leq X_1 + \cdots + X_{153} \leq \frac{153}{\lambda} + \frac{2\sqrt{153}}{\lambda} \tag{7.25}$$

with probability 0.95. Since the interval

$$\frac{153}{\lambda} \pm \frac{2\sqrt{153}}{\lambda}$$

contains 1 for any value of λ between 128 and 178, we conclude that a month with 153 emergency calls is within the range of normal variation for any community in which the average number of emergency calls is between 128 and 178 per month.

There is a final matter of robustness that requires comment. We have assumed that the times between calls, X_n, are exponential. However, the central limit theorem remains true for any distribution as long as μ and σ are finite. Hence, our conclusion is really not sensitive to the assumption of an exponential distribution. It only requires that σ is not too much smaller than μ (for an exponential distribution, $\mu = \sigma$). As we remarked at the end of Section 7.2, there is a good reason to expect this to be the case. Of course, we can always check by estimating μ and σ from the data.

7.4   Diffusion

The normal density function introduced in Section 7.3 is also important in another way. Brownian motion is the accumulation of small, independent random movements of particles. The sum of those independent random movements represents the location of the particle. The central limit theorem says that the distribution of this sum can be approximated by a normal probability density. Hence, the normal density is a model for the diffusion of small particles. In this section, we introduce the diffusion equation, a partial differential equation for particle spreading, in a way that highlights the close connection between the deterministic model and its probabilistic counterpart.

Example 7.5. An accident at an industrial plant ten kilometers upwind of a small town releases an airborne pollutant. One hour after release, a toxic cloud 2000 meters long is headed toward the town at a wind speed of 3 kilometers per hour. The maximum concentration of pollutant in the cloud is 20 times the safe level. What is the maximum concentration expected in town, when will it occur, and how long until the concentration of pollutant falls back below a safe level?

   We will use the five-step method. The first step is to ask a question. We want to know the concentration of pollutant in town and how it varies over time. Let us assume that the cloud of pollutant moves at a constant velocity in the direction of the town. Because of diffusion, the pollutant will also spread as it moves, lowering the peak concentration. Hence, we expect the concentration to diminish with time. We will assume that the wind speed is constant at three kilometers per hour, so that the distance between the town and the plume center is decreasing at three kilometers per hour. We can use the fact that the cloud has spread to a length of 2000 meters in the first hour to estimate the rate of spreading. Then our objective is to predict how the concentration at the town location will vary over time as the cloud of pollutant passes through. The results of step 1 are summarized in Figure 7.6.
   Step 2 is to select the modeling approach. We will use a diffusion model.

> Diffusion is the spreading of particles due to small random movements. The relative concentration of contaminant $C(x, t)$ at location x at time t is described by a partial differential equation called the *diffusion equation*. Here the term *relative concentration* indicates that the concentration function has been normalized so that the total mass of particles $\int C(x, t)\, dx = 1$. This is useful to emphasize the connection to probability theory. The diffusion equation results from a combination of two elements. First of all, the law of conservation of mass states that
> $$\frac{\partial C}{\partial t} = -\frac{\partial q}{\partial x} \qquad (7.26)$$
> where $q(x, t)$ is the *particle flux*, the number of particles passing through the point x per unit time. The change in concentration

Variables: t = Time since release of pollutant (hrs)
                    μ = Distance travelled by plume center (km)
                    x = Distance between plume center and town (km)
                    s = Plume spread at time t (km)
                    P = Pollution concentration in town (times safe level)

Assumptions: $\mu = 3t$
                    $x = 10 - \mu$
                    Peak concentration $P = 20$ when $t = 1$ hour
                    Plume spread is $s = 2000$ meters at $t = 1$ hour

Objective: Determine the maximum pollution level in town, and
                    time until pollution falls back down to a safe level

Figure 7.6: Results of Step 1 of the pollution problem.

inside a small box of width Δx at location x is the result of the difference between the particle flux $q(x,t)$ into the box on the left side, and the flux $q(x+\Delta x, t) \approx q(x,t) + \Delta x\, \partial q/\partial x(x,t)$ out of the box at the right. If $\partial q/\partial x > 0$, then the mass is exiting on the right faster than it enters on the left, so the concentration will decrease. The net loss of mass over a time interval Δt is $\Delta M \approx -\Delta x\, \partial q/\partial x\, \Delta t$, and hence, the concentration $C = M/\Delta x$ changes by an amount $\Delta C = \Delta M/\Delta x \approx -\partial q/\partial x\, \Delta t$, which leads to $\Delta C/\Delta t \approx -\partial q/\partial x$. Taking limits as $\Delta x \to 0$ yields the conservation of mass equation (7.26). The second element is *Fick's Law*, the empirical observation that diffusive particle flux is proportional to the concentration gradient (particles tend to diffuse from areas of high concentration to areas of low concentration), or in other words,

$$q = -\frac{D}{2}\frac{\partial C}{\partial x} \tag{7.27}$$

for some constant $D > 0$, called the *diffusivity*. Combining (7.26) and (7.27) yields the diffusion equation

$$\frac{\partial C}{\partial t} = \frac{D}{2}\frac{\partial^2 C}{\partial x^2}, \tag{7.28}$$

which can be solved to predict the spread of contaminants.

The simplest way to solve the diffusion equation is to use *Fourier transforms*. The Fourier transform of a function $f(x)$ is given by the integral formula

$$\hat{f}(k) = \int_{-\infty}^{\infty} e^{-ikx} f(x)dx. \tag{7.29}$$

Using integration by parts, it is easy to show that the derivative $f'(x)$ has Fourier transform $(ik)\hat{f}(k)$. Tables of Fourier transforms

are widely available, and computer algebra systems like Maple and Mathematica also compute Fourier transforms and their inverses (i.e., compute $f(x)$ for a given $\hat{f}(k)$ and vice versa). One useful formula that can be verified from tables or a computer algebra system is that the standard normal density (7.19) has Fourier transform

$$\hat{g}(k) = \int_{-\infty}^{\infty} e^{-ikx} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx = e^{-k^2/2}. \qquad (7.30)$$

The standard normal distribution is connected with the diffusion model because sums of independent particle movements are asymptotically normal. Since we are interested in the way this distribution of particles spreads over time, we need to understand the scaling properties of the normal density. If Z is a standard normal random variable with mean zero and variance one, then $X = \sigma Z$ has mean zero and its variance is $E(X^2) = E(\sigma^2 Z^2) = \sigma^2$. The distribution function of X is given by

$$F(x) = P(X \leq x) = P(\sigma Z \leq x) = P(Z \leq x/\sigma) = \int_{-\infty}^{x/\sigma} g(t)dt.$$

Substituting $u = \sigma t$ gives

$$F(x) = \int_{-\infty}^{x} g(\sigma^{-1}u)\sigma^{-1}du,$$

which shows that the random variable $X = \sigma Z$ has density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/(2\sigma^2)}.$$

Then a change of variables $t = \sigma x$ in (7.30) shows that the Fourier transform of this density function is

$$\hat{f}(k) = e^{-(\sigma^2 k^2)/2}.$$

In the limiting case $\sigma \downarrow 0$ the Fourier transform is identically equal to one. This corresponds to a random variable concentrated at the origin, so that its spread is zero.

Now let us see how we can use Fourier transforms to solve the diffusion equation. Take Fourier transforms

$$\hat{C}(k, t) = \int_{-\infty}^{\infty} e^{-ikx} C(x, t)dx$$

in the diffusion equation (7.28) to obtain

$$\frac{d\hat{C}}{dt} = \frac{D}{2}(ik)^2\hat{C} = -\frac{D}{2}k^2\hat{C}, \qquad (7.31)$$

which is a very simple ordinary differential equation ($u' = au$) for the Fourier transformed concentration. The solution to this differential equation with initial condition $\hat{C}(k,0) = 1$ for all k is $\hat{C}(k,t) = e^{-Dtk^2/2}$. This initial condition implies that at time $t = 0$ the contaminant plume is concentrated at the point $x = 0$ (the limiting case $\sigma \downarrow 0$). Now invert the Fourier transform to get

$$C(x,t) = \frac{1}{\sqrt{2\pi Dt}} \, e^{-x^2/(2Dt)}, \qquad (7.32)$$

the point source solution to the diffusion equation (7.28).

The connection between the diffusion equation (7.28) and the central limit theorem from Section 7.3 can now be made explicit. Suppose that over a small time interval Δt a contaminant particle makes a small random movement X_i and that all of these movements are independent. Then the particle location at time $n\Delta t$ is given by the sum $X_1 + \cdots + X_n$ of these movements. Assume that the mean jump $E(X_i) = 0$ and the variance of a jump is $\sigma^2 = E(X_i^2)$. Then the central limit theorem (7.18) implies

$$\frac{X_1 + \cdots + X_n}{\sigma\sqrt{n}} \approx Z$$

standard normal, so that $X_1 + \cdots + X_n \approx \sigma\sqrt{n}Z$ in the sense of having approximately the same distribution. Hence, the probability distribution of particle location at time $t = n\Delta t$ is approximately Gaussian with variance $n\sigma^2 = (t/\Delta t)\sigma^2$. Taking $D = \sigma^2/\Delta t$, so that $n\sigma^2 = Dt$, we can relate the limiting random particle location $\sqrt{Dt}Z$ at time t with its probability density $C(x,t)$ in (7.32). For example, this relation shows that the contaminants spread at a rate proportional to $\sqrt{t}$. The normal density solution to the diffusion equation represents the relative concentration of contaminants, since it integrates to one. To represent concentration, it is only necessary to multiply by the total mass of contaminants.

Returning to the pollution problem, we continue with Step 3. The concentration P of pollutant is 20 times the safe level at time $t = 1$ hour and has spread to a width of $s = 2000$ meters. We assume a coordinate system where the pollutant release occurred at location 0 and the town is at location 10 kilometers. We also assume that the center of the plume has moved to location $\mu = 3t$ kilometers after t hours. Using the diffusion model, we can represent the contaminant plume at time $t = 1$ as a normal density with center or mean $\mu = 3$ kilometers and standard deviation $\sigma = 0.500$ kilometers, so that the interval $\mu \pm 2\sigma$ contains the bulk (about 95%) of the pollutant. In other words, we assume that $s = 4\sigma$. Then we can model the relative concentration of pollutant x kilometers away from the plume center of mass at time t by equation (7.32),

where $D = \sigma^2 = 0.25$. This leads to the equation

$$C(x,t) = \frac{1}{\sqrt{0.5\pi t}} \, e^{-x^2/(0.5t)} \tag{7.33}$$

for the relative concentration x kilometers away from the plume center of mass at time t. Next, we note that the pollution concentration P is proportional to the relative concentration C. Using the units of % of safe level, we wish to set $P = 20$ when $t = 1$ and $x = 0$. Hence, we solve

$$P = 20 = P_0 \frac{1}{\sqrt{0.5\pi}} \, e^{-0^2/0.5}$$

to get $P = P_0 C$, where $P_0 = 20\sqrt{0.5\pi}$. Putting this all together, we get that the pollution level in the town, at time t and distance $x = 10 - \mu = 10 - 3t$ kilometers from the plume center of mass, is given by

$$P = \frac{20}{\sqrt{t}} \, e^{-(10-3t)^2/(0.5t)}, \tag{7.34}$$

and we want to answer the following questions: What is the maximum value of P and when does it occur? When will P fall below the safe level of $P = 1$?

Now we move on to Step 4. We want to maximize equation (7.34) over the set $t > 0$, and we want to solve the equation $P = 1$ for the largest positive root $t > 0$. We leave the details to the reader (see Exercise 15). The maximum of $P = 10.97$ occurs at the point $t_0 \approx 3.3$. There are two roots to the equation $P = 1$ over the set $t > 0$ that occur at (approximately) $t = 2.7$ and at $t = 4.1$.

Finally, for Step 5. The cloud of pollutant will expose the town to unsafe levels of contamination, almost 11 times the safe level. This peak level of risk will occur approximately 3 hours and 20 minutes after the pollution event. After around 4 hours the contamination level will fall back below safe levels. We also note that the level of contamination in the town will rise to an unsafe level for the first time at approximately 2 hours and 40 minutes after the accident.

Sensitivity analysis should focus on the parameters in our model that contain a significant amount of uncertainty. The biggest uncertainty is probably in the wind speed v, which could change over time. In our model we assumed $v = 3.0$ kilometers per hour. Generalizing the model, we now have that

$$P = \frac{20}{\sqrt{t}} \, e^{-(10-vt)^2/(0.5t)}, \tag{7.35}$$

and we repeat the analysis for different values of v. Table 7.1 shows the results of this exercise. It is clear that the wind speed has a very significant effect on the risk to the town. A higher wind speed produces a shorter exposure, but at a higher level of contamination. A lower wind speed does the opposite; it leads to a much longer exposure but at lower levels. Solving the problem again with a velocity of $v = 3.03$ (a 1% increase) yields a 0.5% increase in the maximum concentration M, so we conclude that $S(M, v) = 0.5$. The time T

Wind Speed (km/hr)	Highest Concentration	Time of Max Level (hrs)	Time Until Safe Level (hrs)
1.0	6.3	9.9	13.4
2.0	9.0	5.0	6.3
3.0	11.0	3.3	4.1
4.0	12.7	2.5	3.0
5.0	14.2	2.0	2.3

Table 7.1: Sensitivity analysis results for the pollution problem.

at which the maximum concentration M occurs goes down by -1.0 percent, so we conclude that $S(T, v) = -1$. The time L until the concentration falls to a safe level also decreases by 1% when the velocity increases by 1%, so we also have $S(L, v) = -1$. Of course, we could also compute these sensitivities from Table 7.1, with nearly the same results. We can also use the results of our sensitivity analysis to bound the parameters of risk in the event of variations in wind speed. For example, if the wind speed varies between 3 and 4 kilometers per hour over the next few hours, then we conservatively estimate that the maximum contamination in the town will be less than 12.7 times the safe level, and that contamination will drop to safe levels within 4.1 hours.

   Finally, we come to the matter of robustness. Since the normal solution to the diffusion equation is connected to the central limit theorem, we expect a great deal of robustness. After all, independent random particle jumps have to converge to a normal distribution, regardless of the underlying distribution of the individual jumps. Another limitation of our model is that we assume the wind speed is the same at every point in space. It is true that the wind speed may vary, but we can easily account for this in our model by letting the velocity $v(t)$ vary over time. Then we can make the same kind of calculations if we are given accurate data or predictions of wind speed. In either case, the diffusion model predicts that a cloud of particles will spread away from its center of mass at a rate proportional to the square root of time. This is due to the fact that the particle density describes a normal random variable $\sqrt{Dt}Z$ with standard deviation (spread) $\sqrt{Dt}$. In many applications, it has been found that clouds of diffusing particles spread at a different rate than the classical model predicts. This is called *anomalous diffusion*, and it is a very active research area. See Exercise 18 for an illustration.

7.5   Exercises

1. Consider the diode problem of Example 7.1. Let

$$A(x) = \frac{4}{x} + 6 - 5(0.997)^x$$

denote the average cost function.

(a) Show that on the interval $x > 0$, $A(x)$ has a unique minimum value at the point where $A'(x) = 0$.

(b) Use a numerical method to estimate the minimum to within 0.1.

(c) Find the minimum value of $A(x)$ over the set $x = 1, 2, 3, \ldots$

(d) Find the maximum of $A(x)$ over the set $10 \le x \le 35$.

2. Reconsider the diode problem of Example 7.1. In this exercise we will investigate the problem of estimating the failure rate q. In Example 7.1 we assumed that $q = 0.003$; in other words, that 3 out of 1,000 diodes are flawed.

(a) Suppose that we test a batch of 1,000 diodes and we find that 3 are flawed. On this basis we estimate that $q = 3/1,000$. How accurate is this estimate? Use the five-step method, and model as a statistical inference problem. [Suggestion: Define a random variable X_n to represent the status of the nth diode. Let $X_n = 0$ if the nth diode is good, or $X_n = 1$ if the nth diode has a flaw. Then the sum

$$\frac{X_1 + \cdots + X_n}{n}$$

represents the fraction of bad diodes, and the central limit theorem can be used to estimate the likely variation of this sum from the true mean q.]

(b) Repeat part (a), assuming that 30 bad diodes were found in a batch of 10,000.

(c) How many diodes need to be tested in order to be 95% sure that we have determined the failure rate to within 10% of its true value? Assume that the true value is close to our original estimate of $q = 0.003$.

(d) How many diodes need to be tested in order to be 95% sure that we have determined the failure rate to within 1% of its true value? Assume that the true value is close to our original estimate of $q = 0.003$.

3. Consider the radioactive decay problem of Example 7.3, and suppose that our counter registers an average of 10^7 decays/sec over a period of 30 seconds.

(a) Use Eq. (7.14) to estimate the actual decay rate λ.

(b) Generalize Eq. (7.13), using the central limit theorem. Calculate the range of normal variation (at the 95% level) for the observed decay rate T_n/n for an arbitrary value of the true decay rate λ.

(c) Determine the range of λ for which the observed decay rate $T_n/n = 10^7$ is within the range of normal variation at the 95% level. How accurate is our estimate in part (a)?

(d) How long would we have to sample this radioactive material to be 95% sure we have determined the true decay rate λ to six significant digits (i.e., to within $0.5 \times 10^{-6} \lambda$)?

4. It can be shown that the area beneath the graph of the standard normal density function (see Eq. (7.19)) between $-3 \leq x \leq 3$ is approximately 0.997. In other words, for large n we are about 99.7% sure that

$$n\mu - 3\sigma\sqrt{n} \leq X_1 + \cdots + X_n \leq n\mu + 3\sigma\sqrt{n}.$$

Use this fact to repeat the calculations in Exercise 3, but use the 99.7% confidence level. Comment on the sensitivity of your answer in part (d) to the confidence level.

5. Reconsider the house fire problem of Example 7.4. In this exercise we will investigate the problem of estimating the rate λ at which emergency calls occur.

   (a) Suppose that 2,050 emergency calls are received in a one–year period. Estimate the rate λ of house fires per month.

   (b) Assuming that the true value of λ is 171 calls per month, calculate the range of normal variation for the number of emergency calls received in one year.

   (c) Calculate the range of λ for which 2,050 calls in one year is within the range of normal variation. How accurate is our estimate of the true rate λ at which house fires occur?

   (d) How many years of data would be required to obtain an estimate of λ accurate to the nearest integer (an error of ± 0.5)?

6. Reconsider the house fire problem of Example 7.4. The underlying random process is called a Poisson process because it can be shown that the number of arrivals (calls) N_t during a time interval of length t has a Poisson distribution. Specifically,

$$\Pr\{N_t = n\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

for all $n = 0, 1, 2, \ldots$

   (a) Show that
$$EN_t = \lambda t$$
   and
$$VN_t = \lambda t.$$

   (b) Use the Poisson distribution to calculate the probability that the number of calls received in a given month deviates from the mean of 171 by as much as 18 calls.

(c) Generalize the calculation of part (b) to determine the exact range of normal variation (at the 95% level) for the number of calls in a one–month period.

(d) Compare the exact method used in part (c) with the approximate calculation of the range of normal variation that is included in the discussion of sensitivity analysis for Example 7.4 in the text. Which method would be more appropriate for determining the range of normal variation in the number of calls received in a single day? In a year?

7. The Michigan state lottery runs a game in which you pay $1 to buy a ticket containing a three–digit number of your choice. If your number is drawn at the end of the day, you win $500.

(a) Suppose you were to buy one ticket per week for a year. What are your chances of coming out a winner for the year? [Hint: It is easy to compute the probability of coming out a loser!]

(b) Can you improve your chances of coming out a winner this year by purchasing more than one ticket per week? Calculate the probability of coming out a winner if you buy n tickets a week, for $n = 1, 2, 3, \ldots, 9$.

(c) Suppose that the state lottery sells 1,000,000 tickets this week. What is the range of likely variation in the amount of money the state will make this week? How likely is it that the state will lose money this week? Use the central limit theorem.

(d) What is wrong with using the central limit theorem to answer the question in part (a) or (b)?

8. (Murphy's Law, part I) You are staying at a downtown hotel. In front of the hotel there is a taxicab stand. Taxis arrive at random, at a rate of about one every five minutes.

(a) How long do you expect to wait for a taxicab, assuming that there is none at the hotel when you exit?

(b) The time until the arrival of the next taxicab is called a *forward recurrence time*. The time since the most recent arrival is called a *backward recurrence time*. For a Poisson process it can be shown that backward and forward recurrence times have the same distribution. (The probabilistic behavior of the process is the same if we let time run in reverse.) Using this fact, how long on average has it been since the last taxi arrived at the time you exit the hotel?

(c) On average, the time between taxicab arrivals is five minutes. How long on average is the length of time between the arrival of the cab you just missed and the one you have to wait for?

9. (Murphy's Law, part II) You are at the supermarket checkout stand. After waiting for what seems like an unusually long time to check out, you decide to conduct a scientific experiment. One by one, you measure the length of time each customer has to wait. You continue until you find one who has to wait longer than you did.

   (a) Let X denote the time you had to wait, and let X_n denote the time the nth customer had to wait. Let N denote the number of the first customer n for which $X_n \geq X$. To be fair, assume that $X, X_1, X_2, \ldots$ are identically distributed. Explain why the probability that $N \geq n$ (i.e., that out of the group consisting of you and the first $n - 1$ customers, you waited the longest) must equal $1/n$.

   (b) Calculate the probability distribution of the random variable N.

   (c) Calculate the expected value EN, which represents the number of customers you must observe, on average, until you find one who waited longer than you did.

10. (Murphy's Law, part III) A doctor of internal medicine with a busy practice expects to be called into the hospital to respond to a serious heart attack on average about once every two weeks. Assume that heart attacks in this physician's patient population occur at random with this rate. One such emergency call is a challenge. Two such calls in a single day is a disaster.

   (a) How many heart attacks should the physician expect to respond to in a single year?

   (b) Explain why the probability that n heart attacks during the year all occur on different days is

   $$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - n + 1}{365}.$$

   (c) What is the probability that the doctor has to respond to two or more heart attacks on the same day sometime this year?

11. A squadron of 16 bombers needs to penetrate air defenses to reach its target. They can either fly low and expose themselves to the air defense guns, or fly high and expose themselves to surface–to–air missiles. In either case, the air defense firing sequence proceeds in three stages. First, they must detect the target, then they must acquire the target (lock on target), and finally they must hit the target. Each of these stages may or may not succeed. The probabilities are as follows:

AD Type	P_{detect}	P_{acquire}	P_{hit}
Low	0.90	0.80	0.05
High	0.75	0.95	0.70

The guns can fire 20 shells per minute, and the missile installation can fire three per minute. The proposed flight path will expose the planes for one minute if they fly low, and five minutes if they fly high.

(a) Determine the optimal flight path (low or high). The objective is to maximize the number of bombers that survive to strike the target.

(b) Each individual bomber has a 70% chance to destroy the target. Use the results of part (a) to determine the chances of success (target destroyed) for this mission.

(c) Determine the minimum number of bombers necessary to guarantee a 95% chance of mission success.

(d) Perform a sensitivity analysis with respect to the probability $p = 0.7$ that an individual bomber can destroy the target. Consider the number of bombers that must be sent to guarantee a 95% chance of mission success.

(e) Bad weather reduces both P_{detect} and p, the probability that a bomber can destroy the target. If all of these probabilities are reduced in the same proportion, which side gains an advantage in bad weather?

12. A scanning radio communications sensor attempts to detect radio emissions and pinpoint their locations. The sensor scans 4,096 frequency bands. It takes the sensor 0.1 seconds to detect a signal. If no signal is detected, it moves to the next frequency. If a signal is detected, it takes an additional five seconds to get a location fix. There is no signal except on about 100 of the frequency bands, but the sensor does not know which ones are being used, so it must scan them all. On the busy frequencies, the percent utilization (i.e., the fraction of time that the signal is on) varies from 30% to 70%. An additional complication is that emissions on the same frequency can come from several different sources, so the sensor must continue to scan all frequencies even after a source is located.

(a) Determine the approximate detection rate for this system. Assume that all frequency bands are scanned sequentially.

(b) Suppose the sensor has the capability of remembering 25 high–priority frequency bands, which are scanned ten times as often as the others. Assuming that the sensor is eventually able to identify 25 busy frequencies and gives them high priority, how does the detection rate change?

(c) Suppose that, in order to obtain useful information, we must be able to detect emissions on a particular frequency at a minimum rate of once every three minutes. Determine the optimal number of high–priority channels.

(d) Perform a sensitivity analysis on the answer to part (c) with respect to the average utilization rate for busy frequencies.

13. In this problem we will explore the parallel between discrete and continuous random variables. Suppose that X is a continuous random variable with distribution function $F(x)$ and density function $f(x) = F'(x)$. For each n we define a discrete random variable X_n with approximately the same distribution as X. Partition the real line into intervals of length $\Delta x = n^{-1}$, and let I_i denote the ith interval in the partition. Then, for each i, select a point x_i in the ith interval, and define

$$p_i = \Pr\{X_n = x_i\} = f(x_i)\Delta x.$$

 (a) Explain why we can always choose the points x_i so that we will have

$$p_i = \Pr\{X \in I_i\}$$

 for all i. This ensures that $\sum p_i = 1$.

 (b) Derive a formula that represents the probability that $a < X_n \leq b$ in terms of the density function f for any two real numbers a and b.

 (c) Derive a formula that represents the mean EX_n in terms of the density function f.

 (d) Use the results of part (b) to show that as $n \to \infty$ (or, equivalently, $\Delta x \to 0$), we have

$$\Pr\{a < X_n \leq b\} \to \Pr\{a < X \leq b\}$$

 for any two real numbers a and b. We say that X_n *converges in distribution* to X.

 (e) Use the results of part (c) to show that as $n \to \infty$ (or, equivalently, $\Delta x \to 0$), we have

$$EX_n \to EX.$$

 We say that X_n *converges in mean* to X.

14. (Geometric distribution) In this problem we investigate the discrete version of the exponential distribution. Suppose that arrivals occur at random at times $i = 1, 2, 3, \ldots$ Then the time X between two successive arrivals has the geometric distribution

$$\Pr\{X = i\} = p(1 - p)^{i-1},$$

where p is the probability of an arrival occurring at time i.

 (a) Show that $\Pr\{X > i\} = (1 - p)^i$. [Hint: Use the geometric series $1 + x + x^2 + x^3 + \cdots = (1 - x)^{-1}$.]

 (b) Use part (a) to show that X has the memoryless property $\Pr\{X > i + j | X > j\} = \Pr\{X > i\}$.

 (c) Compute $EX = 1/p$. [Hint: Differentiate the geometric series to obtain $1 + 2x + 3x^2 + \cdots = (1 - x)^{-2}$.] Explain why p is the arrival rate for the discrete process.

(d) Customers arrive at a public telephone at random at the rate of one every ten minutes. If Y is the time of first afternoon arrival, use the exponential model to compute $\Pr\{Y > 5\}$. If X is the number of minutes until the next arrival (take the clock time Y and just record the minutes X), use the geometric model to compute $\Pr\{X > 5\}$. Compare.

15. Consider the pollution problem of Example 7.5. Let

$$P(t) = \frac{20}{\sqrt{t}} e^{-(10-3t)^2/(0.5t)}$$

denote the pollution concentration in town at time t.

(a) Graph the function $P(t)$ and comment on its important features.

(b) Show that on the interval $t > 0$, $P(t)$ has a unique maximum at the point where $P'(t) = 0$.

(c) Use a numerical method to estimate the maximum to within 0.1.

(d) Show that the equation $P(t) = 1$ has two positive roots.

(e) Use a numerical method to estimate each positive root of the equation $P(t) = 1$ to within 0.1.

16. A chemical spill has contaminated the ground water near a municipal well. One year after the spill, the contaminant plume is 500 meters downstream of the spill, with a width of 200 meters. The concentration at the center of the plume is 3.6 parts per million.

(a) How long will it take until the maximum concentration reaches the municipal well located 1,800 meters downstream? What will the concentration be? Assume a one–dimensional diffusion model with constant velocity, as in Example 7.5.

(b) When will the concentration at the municipal well fall below a safe level of 0.001 parts per million?

(c) Compute the sensitivity of the answers in parts (a) and (b) to the groundwater velocity.

(d) Compute the sensitivity of the answers in parts (a) and (b) to the measured width of the contaminant plume.

17. (2–D pollution problem) In this problem we consider a two–dimensional diffusion model for the pollution problem of Example 7.5. The relative concentration $C(x, y, t)$ at location (x, y) at time t follows a bivariate normal distribution

$$C(x, y, t) = \frac{1}{\sqrt{2\pi Dt}} e^{-x^2/(2Dt)} \cdot \frac{1}{\sqrt{2\pi Dt}} e^{-y^2/(2Dt)},$$

where the center of the plume is assumed to be at location $x = 0$, $y = 0$. This bivariate density function solves the 2-D diffusion equation

$$\frac{\partial C}{\partial t} = \frac{D}{2} \frac{\partial^2 C}{\partial x^2} + \frac{D}{2} \frac{\partial^2 C}{\partial y^2}$$

and the connection to the probability model is the same as before, except that we also consider the spreading effect of small random jumps in the y direction. As in Example 7.5, an accident at an industrial plant ten kilometers upwind of a small town releases an airborne pollutant. One hour after release, a toxic cloud 2,000 meters wide is headed toward the town at a wind speed of three kilometers per hour, and the peak concentration is 20 times the safe level. Assume that the wind is blowing in the positive x direction, so that the plume center of mass is at location $(3, 0)$ at time $t = 1$, and the town is at location $(10, 0)$.

(a) Make a 3–D plot of the concentration plume at time $t = 1$. Comment on its most important features.

(b) How long will it take until the maximum concentration reaches the town located ten kilometers downwind? What will the concentration be?

(c) When will the concentration in the town fall below a safe level of 0.001 parts per million?

(d) Repeat parts (b) and (c), assuming that the wind is not blowing directly into town. Assume that at time $t = 1$ hour the plume center of mass is at location $(2.95, 0.5)$. How much difference does the wind direction make?

(e) Compare the results of parts (b) and (c) to what was found in the text. Does it make a significant difference whether we use a 1-D or a 2-D diffusion model?

18. (Anomalous diffusion) In this exercise we will investigate a model for anomalous super-diffusion, where a plume spreads faster than the classical diffusion equation predicts. Reconsider the pollution problem of Example 7.5, but now assume that the dispersivity $D(t)$ grows with time, so that $D(t) = 0.25t^{0.4}$ in Eq. (7.32).

(a) Repeat the calculations of Example 7.5. How long will it take until the maximum concentration reaches the town? What will the concentration be?

(b) When will the concentration in town fall to a safe level?

(c) Examine the sensitivity of the answers in parts (a) and (b) to the scaling parameter $p = 0.4$. Repeat parts (a) and (b) for $p = 0.2$, 0.3, 0.5, and 0.6, and discuss.

(d) Compare the results of parts (b) and (c) to what was found in the text. How does the possibility of anomalous super-diffusion affect the conclusions in the text?

Further Reading

1. Barnier, W. *Expected Loss in Keno.* UMAP module 574.

2. Berresford, G. *Random Walks and Fluctuations.* UMAP module 538.

3. Billingsley, P. (1979) *Probability and Measure.* Wiley, New York.

4. Carlson, R. *Conditional Probability and Ambiguous Information.* UMAP module 391.

5. Feller, W. (1971) *An Introduction to Probability Theory and Its Applications.* Vol. 2, 2nd ed., Wiley, New York.

6. Moore, P. and McCabe, G. (1989) *Introduction to the Practice of Statistics.* W.H. Freeman, New York.

7. Ross, S. (1985) *Introduction to Probability Models.* 3rd ed., Academic Press, New York.

8. Watkins, S. *Expected Value at Jai Alai and Pari-Mutuel Gambling.* UMAP module 631.

9. Wheatcraft, S. and Tyler, S. (1988) An explanation of scale-dependent dispersivity in heterogeneous aquifers using concepts of fractal geometry. *Water Resources Research* **24**, 566–578.

10. Wilde, C. *The Poisson Random Process.* UMAP module 340.

Chapter 8

STOCHASTIC MODELS

The deterministic dynamic models of Part II of this book do not allow for the explicit representation of uncertainty. When random effects are taken into account, the resulting model is called a *stochastic model*. Several kinds of general stochastic models are in wide use today. In this chapter we will introduce the most important and commonly used stochastic models.

8.1 Markov Chains

A Markov chain is a discrete–time stochastic model. It is a generalization of the discrete–time dynamical system model introduced in Section 4.3. Although the model is simple, the number and diversity of applications are surprisingly large. In this section we will introduce the general Markov chain model. We will also introduce a concept of steady state appropriate for a stochastic model.

Example 8.1. A pet store sells a limited number of 20–gallon aquariums. At the end of each week, the store manager takes inventory and places orders. Store policy is to order three new 20–gallon aquariums at the end of the week if all of the current inventory has been sold. If even one of the 20–gallon aquarium remains in stock, no new units are ordered. This policy is based on the observation that the store only sells an average of one of the 20–gallon aquariums per week. Is this policy adequate to guard against potential lost sales of 20–gallon aquariums due to a customer requesting one when they are out of stock?

We will use the five-step method. Step 1 is to ask a question. The store begins each sales week with an inventory of between one and three of the 20–gallon aquariums. The number of sales in one week depends on both the supply and the demand. The demand averages one per week but is subject to random fluctuations. It is possible that on some weeks demand will exceed supply, even if we start the week with the maximum inventory of three units. We would like to calculate the probability that demand exceeds supply. In order to get a specific answer, we need to make an assumption about the probabilistic nature

251

Variables: $S_n =$ supply of aquariums at the
                              beginning of week n
                        $D_n =$ demand for aquariums during
                              week n

Assumptions: If $D_{n-1} < S_{n-1}$, then $S_n = S_{n-1} - D_{n-1}$
                   If $D_{n-1} \geq S_{n-1}$, then $S_n = 3$
                   $\Pr\{D_n = k\} = e^{-1}/k!$

Objective: Calculate $P_r\{D_n > S_n\}$

Figure 8.1: Results of step 1 of the inventory problem.

of the demand. It seems reasonable to assume that potential buyers arrive at random at a rate of one per week. Hence, the number of potential buyers in one week will have a Poisson distribution with mean one. (The Poisson distribution was introduced in Exercise 6 of Chapter 7.) Figure 8.1 summarizes the results of step 1. Step 2 is to select the modeling approach. We will use a Markov chain model.

A *Markov chain* can best be described as a sequence of random jumps. For the purposes of this book, we will assume that these jumps can only involve a finite discrete set of locations or states. Suppose that the random variables X_n take values in a finite discrete set. There is no harm in assuming that

$$X_n \in \{1,\, 2,\, 3,\, \ldots,\, m\}.$$

We say that the sequence $\{X_n\}$ is a Markov chain provided that the probability that $X_{n+1} = j$ depends only on X_n. If we define

$$p_{ij} = \Pr\{X_{n+1} = j | X_n = i\}, \tag{8.1}$$

then the entire future history of the process $\{X_n\}$ is determined by the p_{ij} and the probability distribution of the initial X_0. Of course, when we say "determined", we mean that the probabilities $\Pr\{X_n = i\}$ are determined. The actual value of X_n depends on random factors.

Example 8.2. Describe the behavior of the following Markov chain. The state variable

$$X_n \in \{1,\, 2,\, 3\}.$$

If $X_n = 1$, then $X_{n+1} = 1$, 2, or 3 with equal probability. If $X_n = 2$, then $X_{n+1} = 1$ with probability 0.7, and $X_{n+1} = 2$ with probability 0.3. If $X_n = 3$, then $X_{n+1} = 1$ with probability 1.
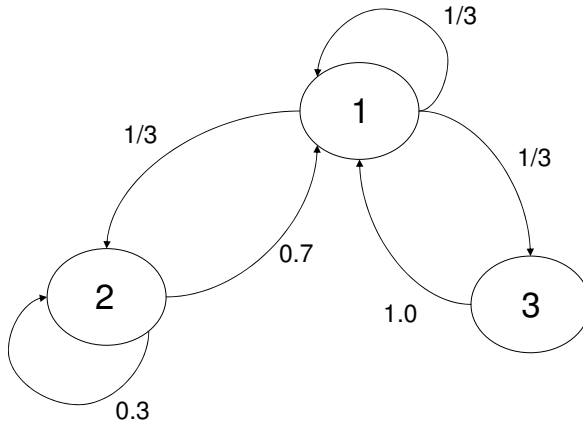
Figure 8.2: State transition diagram for Example 8.2.

The state transition probabilities p_{ij} are given by

$$p_{11} = \frac{1}{3}$$
$$p_{12} = \frac{1}{3}$$
$$p_{13} = \frac{1}{3}$$
$$p_{21} = 0.7$$
$$p_{22} = 0.3$$
$$p_{31} = 1,$$

and the rest are zero. It is customary to write the p_{ij} in matrix form:

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & \cdots & p_{1m} \\ \vdots & & \vdots \\ p_{m1} & \cdots & p_{mm} \end{pmatrix}. \tag{8.2}$$

Here

$$P = \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0.7 & 0.3 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Another convenient method is called the *state transition diagram* (see Figure 8.2). This makes it easy to visualize the Markov chain as a sequence of random jumps. Suppose $X_0 = 1$. Then $X_1 = 1$, 2, or 3 with probability 1/3 each. The probability that $X_2 = 1$ is obtained by calculating the probability associated with each

individual sequence of jumps that transitions from state 1 to state 1 in two steps. Thus,

$$\Pr\{X_2 = 1\} = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)(0.7) + \left(\frac{1}{3}\right)(1) = 0.67\overline{7}.$$

Similarly,

$$\Pr\{X_2 = 2\} = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) + \left(\frac{1}{3}\right)(0.3) = 0.21\overline{1},$$

and

$$\Pr\{X_2 = 3\} = \left(\frac{1}{3}\right)\left(\frac{1}{3}\right) = \frac{1}{9}.$$

To calculate $\Pr\{X_n = j\}$ for larger n, it is useful to observe that

$$\Pr\{X_{n+1} = j\} = \sum_i p_{ij} \Pr\{X_n = i\}. \tag{8.3}$$

The only way to get to state j at time $n + 1$ is to be in some state i at time n, and then jump from i to j. Hence, we could have calculated

$$\Pr\{X_2 = 1\} = p_{11} \Pr\{X_1 = 1\} + p_{21} \Pr\{X_1 = 2\} + p_{31} \Pr\{X_1 = 3\},$$

and so forth. This is where the matrix notation comes in handy.

If we let

$$\pi_n(i) = \Pr\{X_n = i\},$$

then Eq. (8.3) can be written in the form

$$\pi_{n+1}(j) = \sum_i p_{ij}\, \pi_n(i). \tag{8.4}$$

If we let π_n denote the vector with entries $\pi_n(1), \pi_n(2), \ldots$ and let P denote the matrix in Eq. (8.2), then the set of equations relating π_{n+1} to π_n can be written most compactly in the form

$$\pi_{n+1} = \pi_n P. \tag{8.5}$$

For example, we have that $\pi_2 = \pi_1 P$, or

$$(0.67\overline{7}, 0.21\overline{1}, \frac{1}{9}) = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)\begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0.7 & 0.3 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Now we can calculate $\pi_3 = \pi_2 P$ to get

$$\pi_3 = (0.485, 0.289, 0.226)$$

to three decimal places. Continuing on, we obtain

$$\pi_4 = (0.590,\ 0.248,\ 0.162)$$
$$\pi_5 = (0.532,\ 0.271,\ 0.197)$$
$$\pi_6 = (0.564,\ 0.259,\ 0.177)$$
$$\pi_7 = (0.546,\ 0.266,\ 0.188)$$
$$\pi_8 = (0.556,\ 0.262,\ 0.182)$$
$$\pi_9 = (0.551,\ 0.264,\ 0.185)$$
$$\pi_{10} = (0.553,\ 0.263,\ 0.184)$$
$$\pi_{11} = (0.553,\ 0.263,\ 0.184)$$
$$\pi_{12} = (0.553,\ 0.263,\ 0.184).$$

Notice that the probabilities $\pi_n(i) = \Pr\{X_n = i\}$ tend to a specific limiting value as n increases. We say that the stochastic process approaches steady state. This concept of steady state or equilibrium differs from that for a deterministic dynamic model. Because of random fluctuations, we cannot expect that the state variable will stay at one value when the system is in equilibrium. The best we can hope for is that the probability distribution of the state variable will tend to a limiting distribution. We call this the *steady–state distribution*. In Example 8.2 we have

$$\pi_n \to \pi,$$

where the steady–state probability vector

$$\pi = (0.553,\ 0.263,\ 0.184) \tag{8.6}$$

to three decimal places.

A faster way to calculate the steady–state vector π is as follows. Suppose that

$$\pi_n \to \pi.$$

Certainly

$$\pi_{n+1} \to \pi$$

too, so if we let $n \to \infty$ on both sides of Eq. (8.5), we obtain the equation

$$\pi = \pi P. \tag{8.7}$$

We can calculate π simply by solving this linear system of equations. For Example 8.2 we have

$$(\pi_1,\ \pi_2,\ \pi_3) = (\pi_1,\ \pi_2,\ \pi_3) \begin{pmatrix} 1/3 & 1/3 & 1/3 \\ 0.7 & 0.3 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

and it is not hard to calculate that Eq. (8.6) is the only solution to this system of equations for which

$$\sum \pi_i = 1.$$

Not every Markov chain tends to steady state. For example, consider the two–state Markov chain for which

$$\Pr\{X_{n+1} = 2 | X_n = 1\} = 1$$

and

$$\Pr\{X_{n+1} = 1 | X_n = 2\} = 1.$$

The state variable alternates between states 1 and 2. Certainly π_n does not tend to a single limiting vector. We say that this Markov chain is *periodic* with period two. Generally we say state i is periodic with period δ if, starting at $X_n = i$, the chain can return to state i only at times $n + k\delta$. If $\{X_n\}$ is *aperiodic* (every state i has period $\delta = 1$), and if, for each i and j, it is possible to transition from i to j in a finite number of steps, we say that X_n is *ergodic*. There is a theorem that guarantees that an ergodic Markov chain tends to steady state. Furthermore, the distribution of X_n tends to the same steady–state distribution regardless of the initial state of the system (see, e.g., Çinlar (1975) p. 152). Thus, in Example 8.2, if we had started with $X_0 = 2$ or $X_0 = 3$, we would still see π_n converge to the same steady–state distribution π given by Eq. (8.6). The problem of calculating the steady–state probability vector π is mathematically equivalent to the problem of locating the equilibrium of a discrete–time dynamical system with state space $\pi \in \mathbb{R}^m$, $0 \le \pi_j \le 1$,

$$\sum \pi_i = 1$$

and iteration function

$$\pi_{n+1} = \pi_n P.$$

The previously cited theorem states that there is a unique asymptotically stable equilibrium π for this system whenever P represents an ergodic Markov chain.

We return now to the inventory problem of Example 8.1. We will model this problem using a Markov chain. Step 3 is to formulate the model. We begin with a consideration of the state space. The concept of state here is much the same as for deterministic dynamical systems. The state contains all of the information necessary in order to predict the (probabilistic) future of the process. We will take $X_n = S_n$, the number of 20–gallon aquariums in stock at the beginning of our sales week, as the state variable. The demand D_n relates to the dynamics of the model and will be used to construct the state transition matrix P. The state space is

$$X_n \in \{1, 2, 3\}.$$

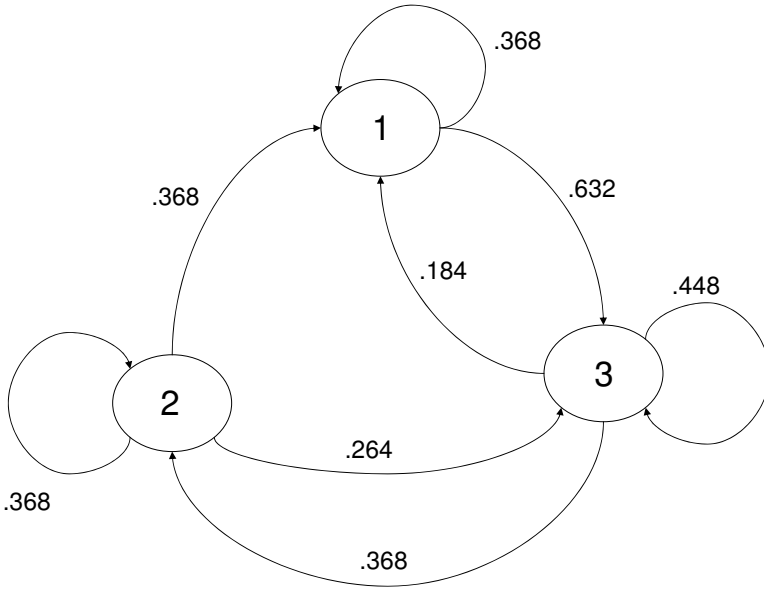Figure 8.3: State transition diagram for the inventory problem.

We do not know the initial state, but it seems reasonable to assume that $X_0 = 3$. In order to determine P, we will begin by drawing the state transition diagram. See Figure 8.3. The distribution of the demand D_n yields

$$\begin{aligned}
\Pr\{D_n = 0\} &= 0.368 \\
\Pr\{D_n = 1\} &= 0.368 \\
\Pr\{D_n = 2\} &= 0.184 \\
\Pr\{D_n = 3\} &= 0.061 \\
\Pr\{D_n > 3\} &= 0.019,
\end{aligned} \tag{8.8}$$

so that if $X_n = 3$, then

$$\begin{aligned}
\Pr\{X_{n+1} = 1\} &= \Pr\{D_n = 2\} = 0.184 \\
\Pr\{X_{n+1} = 2\} &= \Pr\{D_n = 1\} = 0.368 \\
\Pr\{X_{n+1} = 3\} &= 1 - (0.184 + 0.368) = 0.448.
\end{aligned}$$

The remaining state transition probabilities are computed similarly. The state transition matrix is

$$P = \begin{pmatrix} .368 & 0 & .632 \\ .368 & .368 & .264 \\ .184 & .368 & .448 \end{pmatrix}. \tag{8.9}$$

Now for step 4. The analysis objective was to calculate the probability

$$\Pr\{D_n > S_n\}$$

that demand exceeds supply. In general this probability depends on n. More specifically, it depends on X_n. If $X_n = 3$, then

$$\Pr\{D_n > S_n\} = \Pr\{D_n > 3\} = 0.019,$$

and so forth. To get a better idea of how often demand will exceed supply, we need to know more about X_n.

Since $\{X_n\}$ is an ergodic Markov chain, we know that there is a unique steady–state probability vector π that can be computed by solving the steady–state equations. Substituting Eq. (8.9) back into Eq. (8.7), we obtain

$$
\begin{aligned}
\pi_1 &= .368\pi_1 + .368\pi_2 + .184\pi_3 \\
\pi_2 &= .368\pi_2 + .368\pi_3 \\
\pi_3 &= .632\pi_1 + .264\pi_2 + .448\pi_3,
\end{aligned}
\tag{8.10}
$$

which we need to solve along with the condition

$$\pi_1 + \pi_2 + \pi_3 = 1,$$

to obtain the steady–state distribution of X_n. Since we now have four equations in three variables, we can delete one of the equations in Eq. (8.10) and then solve, which yields

$$\pi = (\pi_1, \pi_2, \pi_3) = (.285, .263, .452).$$

For all large n it is approximately true that

$$
\begin{aligned}
\Pr\{X_n = 1\} &= .285 \\
\Pr\{X_n = 2\} &= .263 \\
\Pr\{X_n = 3\} &= .452.
\end{aligned}
$$

Putting this together with our information about D_n, we obtain

$$
\begin{aligned}
\Pr\{D_n > S_n\} &= \sum_{i=1}^{3} \Pr\{D_n > S_n | X_n = i\} \Pr\{X_n = i\} \\
&= (.264)(.285) + (.080)(.263) + (.019)(.452) = .105
\end{aligned}
$$

for large n. In the long run, demand will exceed supply about 10% of the time.

It is easy to compute the steady–state probabilities using a computer algebra system. Figure 8.4 illustrates the use of the computer algebra system Maple to solve the system of equations in Eq. (8.10) to find the steady–state probabilities.

Computer algebra systems are quite useful in such problems, especially when performing sensitivity analysis. If you have access to a computer algebra system, it will be useful in the exercises at the end of this chapter. Even if you prefer to solve systems of equations by hand, you will have the ability to verify your results.

Finally, step 5. The current inventory policy results in lost sales about 10% of the time, or at least five lost sales per year. Most of this is due to the fact

```
> s:={pi1=.368*pi1+.368*pi2+.184*pi3,
>       pi2=.368*pi2+.368*pi3,
>       pi1+pi2+pi3=1};
     s := {π1 = 0.368 π1 + 0.368 π2 + 0.184 π3, π2 = 0.368 π2 + 0.368 π3, π1 + π2 + π3 = 1}
> solve(s,{pi1,pi2,pi3});
                    {π2 = 0.2631807648, π1 = 0.2848348783, π3 = 0.4519843569}
>
```

Figure 8.4: Calculation of the steady–state distribution of the number of 20–gallon aquariums in stock at the beginning of the week for the inventory problem, using the computer algebra system Maple.

that we do not order more aquariums when only one is left. Although we only sell an average of one unit per week, the actual number of potential sales per week (demand) fluctuates from one week to the next. Hence, when we start the week with only one unit in stock, we run a significant risk (about a one in four chance) of losing potential sales due to insufficient inventory. In the absence of other factors, such as a discount for orders of three or more, it seems reasonable to try out a new inventory policy in which we never start out a week with only one aquarium.

We now come to the subject of sensitivity analysis and robustness. The main sensitivity issue is the effect of the arrival rate λ of potential buyers on the probability that demand exceeds supply. Currently $\lambda = 1$ customer per week. For arbitrary λ, the state transition matrix for X_n is given by

$$P = \begin{pmatrix} e^{-\lambda} & 0 & 1 - e^{-\lambda} \\ \lambda e^{-\lambda} & e^{-\lambda} & 1 - (1 + \lambda)e^{-\lambda} \\ \lambda^2 e^{-\lambda}/2 & \lambda e^{-\lambda} & 1 - (\lambda + \lambda^2/2)e^{-\lambda} \end{pmatrix}, \tag{8.11}$$

using the fact that D_n has a Poisson distribution. While it would be possible to carry through the calculation of $p = \Pr\{D_n > S_n\}$ from this point, it would be very messy. It makes more sense simply to repeat the calculations of step 4 for a few selected values of λ near 1. The results of this exercise are shown in Figure 8.5. They confirm that our basic conclusions are not particularly sensitive to the exact value of λ. The sensitivity $S(p, \lambda)$ is around 1.5. (Another sensible option for sensitivity analysis is to use a computer algebra system to perform the messy calculations. See Exercise 2 at the end of this chapter.)

Finally, we should consider the robustness of our model. We have assumed a Markov chain model based on a Poisson process model of the arrival process. The robustness of the Poisson process model as a representative of a more general arrival process was discussed briefly at the end of Section 7.2. It is reasonable to conclude that our results would not be altered significantly if the arrival process were not exactly Poisson. The basic assumption here is that the arrival process represents the merging of a large number of independent arrival processes. Many kinds of customers arrive at the shop from time to
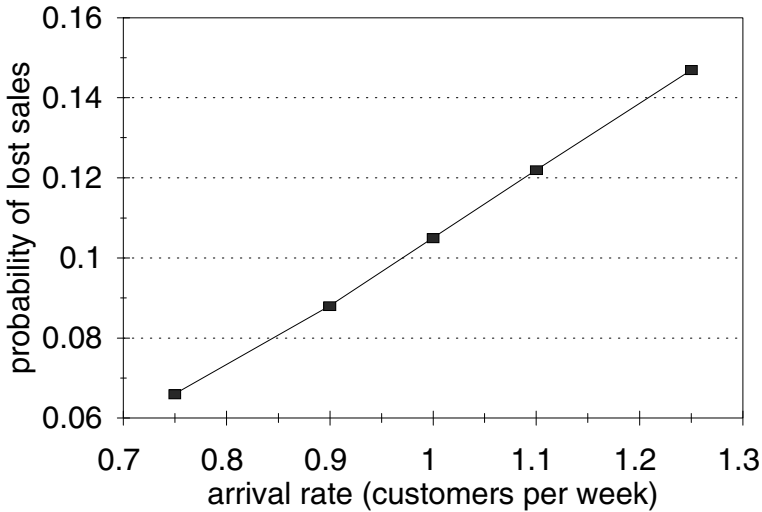
Figure 8.5: Graph showing the sensitivity of the probability of lost sales to arrival rate in the inventory problem.

time to buy a 20–gallon aquarium, and it is reasonable to assume that they do not coordinate this activity with each other. Of course, certain store activities, such as an advertised sale price on 20–gallon aquariums, would invalidate this assumption, resulting in the need to reexamine the conclusions of our modeling exercise. It may also be that there are significant seasonal variations in the demand for this item.

The other basic modeling assumption is that the inventory level S_n represents the state of the system. A more sophisticated model might take into account the store manager's response to long–term fluctuations in sales, such as seasonal variations. The mathematical analysis of such a model is more complex, but not essentially different from what we did here. We just expand the state space to include information on past sales; say, S_n, S_{n-1}, S_{n-2}, S_{n-3}. Of course, our transition matrix P is now 81×81 instead of 3×3.

Many different inventory policies are possible. Several of these are explored in the exercises at the end of the chapter. Which inventory policy is the best? One way to approach this question is to formulate an optimization model based on a generalized version of our Markov chain model. A range of inventory policies is described in terms of one or more decision variables, and the objective is defined in terms of the resulting steady–state probabilities. The study of such models is called *Markov decision theory*. Details can be found in any introductory text on operations research (e.g., Hillier et al. (1990)).

Variables: X_t = number of forklifts in repair
at time t months.

Assumptions: Vehicles arrive for repair at
the rate of 4.5 per month.
Maximum repair rate is 7.3
vehicles per month.

Objective: Calculate EX_t, $\Pr\{X_t > 0\}$

Figure 8.6: Results of step 1 of the forklift problem.

8.2 Markov Processes

A Markov process model is the continuous–time analogue of the Markov chain model introduced in the preceding section. It may also be considered as a stochastic analogue to a continuous–time dynamical system model.

Example 8.3. A mechanic working for a heavy equipment repair facility is responsible for the repair and maintenance of forklift trucks. When forklifts break down, they are taken to the repair facility and serviced in the order of their arrival. There is space for 27 forklift trucks at the facility, and last year the facility repaired 54 trucks. Average repair time for a single vehicle is about three days. In the past few months certain questions have been raised about the effectiveness and the efficiency of this operation. The two central issues are the time it takes to have a machine repaired, and the percentage of time the mechanic devotes to this part of his duties.

We will analyze the situation using a mathematical model of the repair facility. Forklift trucks arrive at the facility for repair at a rate of $54/12 = 4.5$ per month. The maximum rate at which they can be repaired is $22/3 \approx 7.3$ vehicles per month, based on an average of 22 working days per month. Let X_t denote the number of vehicles in the repair shop at time t. We are interested in the average number in service, EX_t, and the proportion of time that the mechanic is busy repairing machines, represented by $\Pr\{X_t > 0\}$. Figure 8.6 summarizes the results of step 1.

We will model the repair facility using a Markov process.

A *Markov process* is the continuous–time analogue of the Markov chain introduced in the previous section. As before, we will assume that the state space is finite; i.e., we will suppose

$$X_t \in \{1, 2, 3, \ldots, m\}.$$

The stochastic process $\{X_t\}$ is a Markov process if the current state X_t really represents the state of the system; i.e., it totally determines

the probabilistic future of the process. This condition is formally written as

$$\Pr\{X_{t+s} = j | X_u : u \leq t\} = \Pr\{X_{t+s} = j | X_t\}. \tag{8.12}$$

The Markov property, Eq. (8.12), has two important implications. First of all, the time until the next transition does not depend on how long the process has been in the current state. In other words, the distribution of time spent in a particular state has the memoryless property. Let T_i denote the time spent in state i. Then the Markov property says that

$$\Pr\{T_i > t + s | T_i > s\} = \Pr\{T_i > t\}. \tag{8.13}$$

In Section 7.2 we showed that the exponential distribution has this property, so T_i could have density function

$$F_i(t) = \lambda_i e^{-\lambda_i t}. \tag{8.14}$$

In fact, the exponential distribution is the only continuous distribution having the memoryless property. (This is a deep theorem in real analysis. See Billingsley (1979) p. 160.) Hence, for a Markov process the distribution of time in a particular state is exponential with parameter λ_i, which depends in general on the state i.

The second important implication of the Markov property has to do with state transition. The probability distribution that describes the identity of the next state can depend only on the current state. Thus, the sequence of states visited by the process forms a Markov chain. If we let p_{ij} denote the probability that the process jumps from state i to state j, then the embedded Markov chain has state transition probability matrix $P = (p_{ij})$.

Example 8.4. Consider a Markov chain with state transition probability

$$P = \begin{pmatrix} 0 & 1/3 & 2/3 \\ 1/2 & 0 & 1/2 \\ 3/4 & 1/4 & 0 \end{pmatrix}, \tag{8.15}$$

and form a Markov process by assuming that the jumps of $\{X_t\}$ follow this Markov chain with the average times in states 1, 2, 3 equal to 1, 2, and 3, respectively.

Solving the steady–state equation $\pi = \pi P$ shows that the proportion of jumps that land in states 1, 2, and 3 are 0.396, 0.227, and 0.377, respectively. However, the proportion of time spent in each state also depends on how long we wait in one state before the next jump. Correcting for this produces the relative proportions 1(.396), 2(.227), and 3(.377). If we normalize to 1 (divide each term by the

sum), we get 0.200, 0.229, and 0.571. Hence the Markov process spends about 57.1% of the time in state 3, and so on. We call this the steady–state distribution for the Markov process. Generally, if $\pi = (\pi_1, \ldots, \pi_m)$ is the steady–state distribution for the embedded Markov chain and $\lambda = (\lambda_1, \ldots, \lambda_m)$ is the vector of rates, then the proportion of time spent in state i is given by

$$P_i = \frac{(\pi_i/\lambda_i)}{(\pi_1/\lambda_1) + \cdots + (\pi_m/\lambda_m)}. \tag{8.16}$$

The reciprocal of the rate λ_i represents mean time in state i. In summary, a Markov process can be thought of as a Markov chain where the time between jumps has an exponential distribution, depending on the current state.

An equivalent model can be formulated as follows. Given that $X_t = i$, let T_{ij} be exponential with parameter $a_{ij} = \lambda_i p_{ij}$. Furthermore, suppose that $T_{i1}, \ldots, T_{im}$ are independent. Then the time T_i until the next jump is the minimum of $T_{i1}, \ldots, T_{im}$, and the next state is the state j such that T_{ij} is the minimum of $T_{i1}, \ldots, T_{im}$. The mathematical equivalence between the two forms of the Markov process model follows from the fact that

$$T_i = \min(T_{i1}, \ldots, T_{im})$$

is exponential with parameter

$$\lambda_i = \sum_j a_{ij}$$

and that

$$\Pr\{T_i = T_{ij}\} = p_{ij}.$$

(The proof is left to the reader. See Exercise 7 at the end of the chapter.) The parameter

$$a_{ij} = \lambda_i p_{ij}$$

denotes the rate at which the process tends to go from state i to state j. It is customary to depict the rates a_{ij} in a *rate diagram*. The rate diagram for Example 8.4 is given in Figure 8.7. Often the structure of a Markov process is originally specified by such a diagram.

This alternative formulation for the Markov process leads to a convenient method of computing the steady–state distribution, based on the rate diagram. As before we define

$$a_{ij} = \lambda_i p_{ij}$$

to be the rate at which the process tends to jump from state i to state j. Also, let

$$a_{ii} = -\lambda_i$$

denote the rate at which the process tends to leave state i. It can be shown that the probability functions

$$P_i(t) = \Pr\{X_t = i\} \tag{8.17}$$

must satisfy the differential equations

$$P_1'(t) = a_{11}P_1(t) + \cdots + a_{m1}P_m(t)$$
$$\vdots \tag{8.18}$$
$$P_m'(t) = a_{1m}P_1(t) + \cdots + a_{mm}P_m(t)$$

(see Çinlar (1975) p. 255). This basic condition can be most easily understood by using the *fluid–flow analogy*. Visualize the probabilities $P_i(t)$ as the amount of fluid (probability mass) at each state i. The rates a_{ij} represent the rate of fluid flow, and the fact that

$$P_1(t) + \cdots + P_m(t) = 1$$

means that the total amount of fluid stays equal to 1. In the case of Example 8.4, we have

$$P_1'(t) = -P_1(t) + \frac{1}{4} P_2(t) + \frac{1}{4} P_3(t)$$
$$P_2'(t) = \frac{1}{3} P_1(t) - \frac{1}{2} P_2(t) + \frac{1}{12} P_3(t) \tag{8.19}$$
$$P_3'(t) = \frac{2}{3} P_1(t) + \frac{1}{4} P_2(t) - \frac{1}{3} P_3(t).$$

The steady–state distribution for the Markov process corresponds to the steady–state solution to this system of differential equations. Setting $P_i' = 0$ for all i, we obtain

$$0 = -P_1 + \frac{1}{4} P_2 + \frac{1}{4} P_3$$
$$0 = \frac{1}{3} P_1 - \frac{1}{2} P_2 + \frac{1}{12} P_3 \tag{8.20}$$
$$0 = \frac{2}{3} P_1 + \frac{1}{4} P_2 - \frac{1}{3} P_3.$$

Solving the system of linear equations in Eq. (8.20) together with the condition

$$P_1 + P_2 + P_3 = 1$$

yields

$$P = \left(\frac{7}{35}, \frac{8}{35}, \frac{20}{35} \right),$$

which is the same as before.
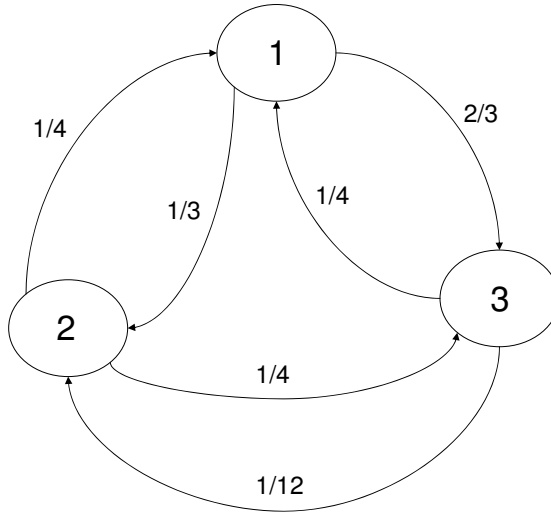
Figure 8.7: Rate diagram for Example 8.4, illustrating the rate at which the process tends to jump from one state to another.

One simple way to determine the system of equations that must be solved to get the steady–state distribution for a Markov process is to use the fluid–flow analogy. Fluid flows into and out of each state. In order for the system to remain in equilibrium, the rate at which fluid flows into each state must equal the rate at which it flows back out. For example, in Fig. 8.7 fluid is flowing out of state 1 at the rate of $1/3 + 2/3 = 1 \times P_1$. Fluid flows from state 2 back into state 1 at the rate of $1/4 \times P_2$ and from state 3 to state 1 at the rate of $1/4 \times P_3$, so we have the condition

$$P_1 = 1/4 \ P_2 + 1/4 \ P_3.$$

By applying this principle of

$$[\text{Rate out}] = [\text{Rate in}]$$

to the other two states as well, we obtain the system of equations

$$\begin{aligned}
P_1 &= \frac{1}{4} \ P_2 + \frac{1}{4} \ P_3 \\
\frac{1}{2} \ P_2 &= \frac{1}{3} \ P_1 + \frac{1}{12} \ P_3 \\
\frac{1}{3} \ P_3 &= \frac{2}{3} \ P_1 + \frac{1}{4} \ P_2.
\end{aligned} \tag{8.21}$$

which is equivalent to the system of equations in Eq. (8.20). We call the system of equations in Eq. (8.21) the *balance equations* for our Markov process model. They express the condition that the rates into and out of each state are in balance.

In Section 8.1 we remarked that an ergodic Markov chain will always tend to steady state. Now we will state the corresponding result for Markov processes. A Markov process is called *ergodic* if, for every pair of states i and j, it is possible to jump from i to j in a finite number of transitions. There is a theorem that guarantees that an ergodic Markov process always tends to steady state. Furthermore, the distribution of X_t tends to the same steady–state distribution regardless of the initial state of the system (see, for example, Çinlar (1975) p. 264). Let

$$P(t) = (P_1(t), \ldots, P_m(t))$$

denote the current probability distribution of our Markov process. Then, this theorem says that for any initial probability distribution $P(0)$ on the state space, we will always see the probability distribution $P(t)$ of the Markov process state vector X_t converge to the same steady–state distribution

$$P = (P_1, \ldots, P_m)$$

as $t \to \infty$. The system of differential equations in Eq. (8.18) that describes the dynamics of the probability distribution $P(t)$ can be written in matrix form as

$$P(t)' = P(t)A, \tag{8.22}$$

where $A = (a_{ij})$ is the matrix of rates. This is a linear system of differential equations on the space

$$S = \{x \in \mathbb{R}^m : 0 \le x_i \le 1; \sum x_i = 1\}. \tag{8.23}$$

Our theorem says that if this dynamical system in Eq. (8.22) represents an ergodic Markov process, then there exists a unique stable equilibrium solution P. Furthermore, for any initial condition $P(0)$ we will have $P(t) \to P$ as $t \to \infty$. More detailed information about the transient (time–dependent) behavior of the Markov process can be obtained by explicitly solving the linear system in Eq. (8.22) by the usual methods.

Now we return to the forklift problem of Example 8.3. We want to formulate a Markov process model for $X_t =$ the number of forklifts in repair at time t months. Since the facility can handle only 27 forklifts, we have
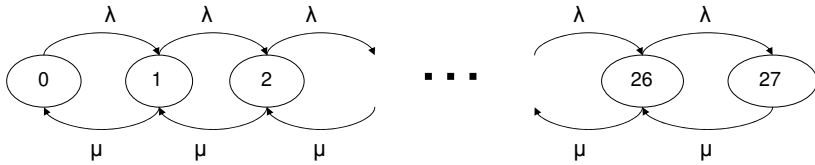
$$X_t \in \{0, 1, 2, \ldots, 27\}.$$

Figure 8.8: Rate diagram for the forklift problem, illustrating the rate at which the number of forklifts in service tends to increase or decrease.

The only allowed transitions are from $X_t = i$ to $X_t = i + 1$ or $i - 1$. The rates up and down are $\lambda = 4.5$ and $\mu = 7.3$, respectively, except that we cannot transition up from state 27 or down from state 0. The rate diagram for this problem is shown in Figure 8.8.

The steady–state equations $PA = 0$ can be obtained from the rate diagram by using the

$$[\text{Rate out}] = [\text{Rate in}]$$

principle. From Fig. 8.8 we obtain

$$\lambda P_0 = \mu P_1$$
$$(\mu + \lambda)P_1 = \lambda P_0 + \mu P_2$$
$$(\mu + \lambda)P_2 = \lambda P_1 + \mu P_3$$
$$\vdots \qquad\qquad (8.24)$$
$$(\mu + \lambda)P_{26} = \lambda P_{25} + \mu P_{27}$$
$$\mu P_{27} = \lambda P_{26}.$$

Solving along with

$$\sum P_i = 1$$

will yield the steady–state $\Pr\{X_t = i\}$. We are interested in

$$\Pr\{X_t > 0\} = 1 - \Pr\{X_t = 0\} = 1 - P_0$$

and in

$$EX_t = \sum i P_i.$$
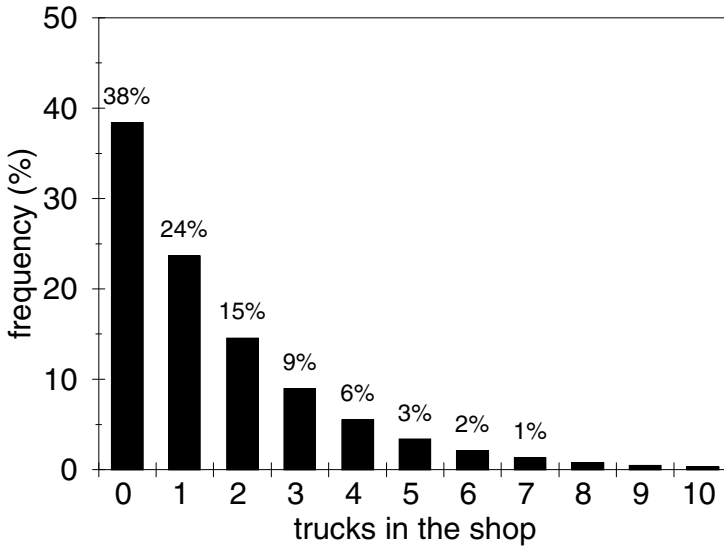
Figure 8.9: Histogram showing the distribution of the number of forklifts in service at any given time.

Moving on to step 4, we first solve for P_1 in terms of P_0, then for P_2 in terms of P_1, and so forth, to obtain

$$P_n = (\lambda/\mu)\, P_{n-1}$$

for all $n = 1$, 2, 3, ..., 27. Then

$$P_n = (\lambda/\mu)^n \, P_0$$

for all such n. Since

$$\sum_{n=0}^{27} P_n = P_0 \sum_{n=0}^{27} \left(\frac{\lambda}{\mu}\right)^n = 1,$$

we must have

$$P_0 = \frac{1 - \rho}{1 - \rho^{28}},$$

where $\rho = \lambda/\mu$. Here we have used the standard formula for the sum of a finite geometric series. For $n = 1$, 2, 3, ..., 27 we have

$$P_n = \rho^n P_0 = \frac{\rho^n (1 - \rho)}{1 - \rho^{28}}. \tag{8.25}$$

Now $\rho = \lambda/\mu = 4.5/7.3 \approx 0.616$, so $1 - \rho^{28} \approx 0.9999987$. Hence, we may assume for all practical purposes that $P_0 = 1 - \rho$ and $P_n = \rho^n(1 - \rho)$ for $n \geq 1$.

Now we calculate our two measures of performance. First, we have

$$\Pr\{X_t > 0\} = 1 - P_0 = \rho \approx 0.616.$$

Next, we have

$$
\begin{aligned}
EX_t &= \sum_{n=0}^{27} n P_n \\
&= \sum_{n=0}^{27} n \rho^n (1 - \rho),
\end{aligned}
\tag{8.26}
$$

which yields $EX_t = 1.607$.

To summarize (step 5), we consider a system where forklift trucks break down at a rate of 4.5 per month and are taken to a repair facility with the capacity to service up to 7.3 per month. Since the rate at which vehicles arrive for repair is only about 60% of the potential service rate, the mechanic is busy with this activity only about 60% of the time. However, since breakdowns occur essentially at random, there will be times when there is more than one vehicle in the shop at one time, through no fault of the mechanic. In fact, on an average day we would expect to see 1.6 vehicles in the repair facility. By this, we mean that if we kept track on a daily basis of the number of vehicles in the repair facility, then at the end of the year this number would average about 1.6. More specifically, we would expect the distribution indicated by Fig. 8.9.

Murphy's Law is certainly in evidence here. Even if the mechanic does his job flawlessly, on 8% of the working days in a year, there will be 5 or more forklift trucks in the shop. Since it takes three days to fix one truck, this represents a backlog of about three weeks. Assuming 250 working days a year, this unfortunate situation will occur about 20 days out of the year. Meanwhile, a time study will show the mechanic busy on this part of his job only about 60% of the time. This seeming discrepancy can be explained by the simple fact that breakdowns are not always nicely spaced out. Sometimes, due to sheer bad luck, several machines will break down in rapid succession, and the mechanic will be swamped. At other times there will be long periods between breakdowns, and the mechanic will have no repairs to perform.

There are really two management problems here. One is the problem of idle time. There is no way to avoid the sporadic nature of this job activity, so idle time will continue to be a significant concern. Management might consider the extent to which the mechanic can perform other duties during this idle time. Of course, these duties should be easily interruptible in case there is a breakdown.

The second problem is backlogs. Sometimes a number of forklift trucks will be waiting for repair. Additional study may be required to determine whether this problem should be addressed by devoting additional manpower to the repair activity during particularly busy periods. (This problem is the subject of Exercise 8 at the end of this chapter.)

Finally, we come to the very important questions of sensitivity and robustness. Both measures of performance depend on the ratio

$$\rho = \frac{\lambda}{\mu},$$

currently equal to 0.616. The relation

$$\Pr\{X_t > 0\} \approx \rho \tag{8.27}$$

requires no further comment. If we let $A = EX_t$ denote the average number in the system, then (using $1 - \rho^{28} \approx 1$) we have

$$A = \sum_{n=0}^{27} n\rho^n(1 - \rho), \tag{8.28}$$

which reduces to

$$A = \rho + \rho^2 + \rho^3 + \cdots + \rho^{27} - 27\rho^{28}.$$

This is approximately

$$\rho(1 + \rho + \rho^2 + \rho^3 + \cdots) = \frac{\rho}{1 - \rho},$$

and so

$$\frac{dA}{d\rho} \approx \frac{1}{(1 - \rho)^2},$$

so that

$$S(A, \rho) \approx 2.6.$$

A small error in ρ would not significantly alter our conclusions based on $A = EX_t$.

We should also consider the size of the facility storage area, currently at $K = 27$ forklift trucks. We have seen that for moderate ρ (not too close to 1) this parameter makes little difference, in the sense that we have continually used the approximation

$$1 - \rho^{K+1} \approx 1.$$

Indeed, this approximation is equivalent to the assumption that there is unlimited storage capacity for forklift trucks to repair. Essentially, we are assuming that $K = \infty$. If more storage capacity is made available, this will not change the rate λ at which vehicles arrive to be repaired. Thus, $\rho = \lambda/\mu$ remains the same, so both our measures of performance are insensitive to K. The only effect of increasing K (i.e., obtaining aditional storage space) will be to increase the number of trucks that can be waiting for service.

The model we have used to represent the repair facility is a special case of a *queuing model*. A queuing model represents a system consists of one or more service facilities at which arrivals are processed, and those arrivals that

cannot be processed have to wait for service in a queue. There is a large body of literature on queuing models, including much current research. A textbook on operations research (e.g., Hillier et al. (1990)) is a good place to begin. The most important assumption we might try to relax is that service times have an exponential distribution. There is ample reason to suspect that the arrivals are more or less random. There is a result (based on the simplifying assumption $K = \infty$, which we have made before) for a general service time distribution with variance σ^2 that states that $\rho = \lambda/\mu$ is the probability that the server is busy and the steady–state

$$EX_t = \rho + \frac{\lambda^2\sigma^2 + \rho^2}{2(1-\rho)}. \tag{8.29}$$

Of course, this reduces to $\rho/(1-\rho)$ in the case of an exponential service time. In that case, $\sigma = 1/\mu$. The general conclusion to be drawn from this formula is that the average number of vehicles in repair grows with the variance of service time. Thus, more uncertainty about the length of a repair will result in longer waiting times.

8.3   Linear Regression

The single most commonly used stochastic model assumes that the expected value of the state variable is a linear function of time. The model is attractive not only because of its wide range of applications, but also because of the availability of good software implementations.

Example 8.5. Adjustable–rate mortgages on private homes are commonly based on one of several market indices tabulated by the federal home loan bank. The author's mortgage is adjusted yearly on the basis of the U.S. Treasury one–year Constant Maturity (CM1) index for May of each year. Historical data for the three–year period beginning June 1986 are shown in Table 8.1 (source: Board of Governors of the Federal Reserve). Use this information to project the estimated value of this index in May 1990, the date of the next adjustment.

We will use the five-step method. Step 1 is to ask a question. We are attempting to estimate the future trend in a variable that exhibits a tendency to grow with time, along with some random fluctuation. Let X_t denote the U.S. Treasury one–year Constant Maturity (CM1) index at time t months after May 1986. A graph of X_t for $t = 1, \ldots, 37$ is shown in Figure 8.10. We want to estimate X_{48}. If we assume that X_t depends in part on a random element, then we cannot expect to predict X_{48} exactly. The best we can hope for is an average value EX_{48}, along with some kind of measure of the magnitude of uncertainty. For the moment we will concentrate on obtaining an estimate of EX_{48}. We will leave the other matter for the section on sensitivity analysis.

Step 2 is to select the modeling approach. We will model this problem using linear regression.

month	TB3	TB6	CM1	CM2	CM3	CM5
6/86	6.21	6.28	6.73	7.18	7.41	7.64
7/86	5.84	5.85	6.27	6.67	6.86	7.06
8/86	5.57	5.58	5.93	6.33	6.49	6.80
9/86	5.19	5.31	5.77	6.35	6.62	6.92
10/86	5.18	5.26	5.72	6.28	6.56	6.83
11/86	5.35	5.42	5.80	6.28	6.46	6.76
12/86	5.49	5.53	5.87	6.27	6.43	6.67
1/87	5.45	5.47	5.78	6.23	6.41	6.64
2/87	5.59	5.60	5.96	6.40	6.56	6.79
3/87	5.56	5.56	6.03	6.42	6.58	6.79
4/87	5.76	5.93	6.50	7.02	7.32	7.57
5/87	5.75	6.11	7.00	7.76	8.02	8.26
6/87	5.69	5.99	6.80	7.57	7.82	8.02
7/87	5.78	5.86	6.68	7.44	7.74	8.01
8/87	6.00	6.14	7.03	7.75	8.03	8.32
9/87	6.32	6.57	7.67	8.34	8.67	8.94
10/87	6.40	6.86	7.59	8.40	8.75	9.08
11/87	5.81	6.23	6.96	7.69	7.99	8.35
12/87	5.80	6.36	7.17	7.86	8.13	8.45
1/88	5.90	6.31	6.99	7.63	7.87	8.18
2/88	5.69	5.96	6.64	7.18	7.38	7.71
3/88	5.69	5.91	6.71	7.27	7.50	7.83
4/88	5.92	6.21	7.01	7.59	7.83	8.19
5/88	6.27	6.53	7.40	8.00	8.24	8.58
6/88	6.50	6.76	7.49	8.03	8.22	8.49
7/88	6.73	6.97	7.75	8.28	8.44	8.66
8/88	7.02	7.36	8.17	8.63	8.77	8.94
9/88	7.23	7.43	8.09	8.46	8.57	8.69
10/88	7.34	7.50	8.11	8.35	8.43	8.51
11/88	7.68	7.76	8.48	8.67	8.72	8.79
12/88	8.09	8.24	8.99	9.09	9.11	9.09
1/89	8.29	8.38	9.05	9.18	9.20	9.15
2/89	8.48	8.49	9.25	9.37	9.32	9.27
3/89	8.83	8.87	9.57	9.68	9.61	9.51
4/89	8.70	8.73	9.36	9.45	9.40	9.30
5/89	8.40	8.39	8.98	9.02	8.98	8.91
6/89	8.22	8.00	8.44	8.41	8.37	8.29

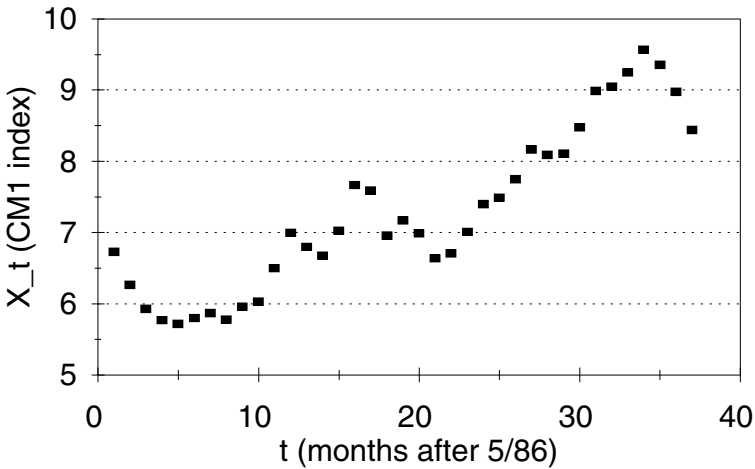Table 8.1: Possible adjustable–rate mortgage loan indices.

Figure 8.10: Graph of CM1 index versus time for the ARM problem.

The *linear regression* model assumes that

$$X_t = a + bt + \varepsilon_t, \tag{8.30}$$

where a and b are real constants and ε_t is a random variable that represents the effect of random fluctuations. It is assumed that

$$\varepsilon_1, \, \varepsilon_2, \, \varepsilon_3, \, \ldots$$

are independent and identically distributed with mean zero. It is also common to assume that ε_t is normal; i.e., that for some $\sigma > 0$ the random variable

$$\varepsilon_t / \sigma$$

has the standard normal density. In the case where the random fluctuations represented by ε_t involve the additive effects of a fairly large number of independent random factors, this normal assumption is justified by the central limit theorem. (The normal density and the central limit theorem were introduced in Section 7.3.)

Since the error term ε_t has mean zero,

$$EX_t = a + bt, \tag{8.31}$$

so the problem of estimating EX_t reduces to the estimation of the parameters a and b. If we were to graph the line

$$y = a + bt \tag{8.32}$$

on the graph of Fig. 8.10, we would expect the data points to lie near this line, with some above and some below. The best–fitting line, representing our best estimate of the parameters a and b, should minimize the extent to which the data points deviate from the line.

Given a set of data points

$$(t_1, y_1), \ldots, (t_n, y_n),$$

we measure the goodness of fit of the regression line in terms of the vertical distance

$$|y_i - (a + bt_i)|$$

between the data point (t_i, y_i) and the point on the regression line, Eq. (8.32), at $t = t_i$. To avoid absolute value signs, which are troublesome in an optimization problem, we measure the overall goodness of fit by

$$F(a, b) = \sum_{i=1}^{n} (y_i - (a + bt_i))^2. \tag{8.33}$$

The best–fitting line is characterized by a global minimum of the objective function in Eq. (8.33). Setting the partial derivatives $\partial F/\partial a$ and $\partial F/\partial b$ equal to zero yields

$$\sum_{i=1}^{n} y_i = na + b \sum_{i=1}^{n} t_i$$

$$\sum_{i=1}^{n} t_i y_i = a \sum_{i=1}^{n} t_i + b \sum_{i=1}^{n} t_i^2. \tag{8.34}$$

Solving these two linear equations in two unknowns determines a and b.

An estimate of the predictive power of the regression equation (8.32) can be obtained as follows. Let

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \tag{8.35}$$

denote the average or mean value of the y data points, and for each i let

$$\hat{y}_i = a + bt_i. \tag{8.36}$$

The total variation $y_i - \bar{y}$ between any one data value and the mean value can be expressed as the sum

$$(y_i - \bar{y}) = (y_i - \hat{y}_i) + (\hat{y}_i - \bar{y}). \tag{8.37}$$

The first term on the right-hand side of Eq. (8.37) represents the error (vertical distance of the data point from the regression line), and

the second term represents the amount of deviation in y accounted
for by the regression line. A little algebra shows that

$$\sum_{i=1}^{n} (y_i - \bar{y})^2 = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 + \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2. \qquad (8.38)$$

The statistic

$$R^2 = \frac{\displaystyle\sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2}{\displaystyle\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (8.39)$$

measures the portion of the total variation in the data accounted for
by the regression line. The remaining portion of the total variation
is attributed to random errors; i.e., the effect of ε_t. If R^2 is close to
1, then the data are very nearly linear. If R^2 is close to 0, the data
are very nearly random.

Most educational computer laboratories have statistical packages
that will automatically compute a, b, and R^2 from a data set. In-
expensive software of the same kind is available for most personal
computers, and some hand–held calculators have built–in linear re-
gression functions. For the linear regression problems in this book,
any of these methods will suffice. It is not recommended that these
problems be solved by hand.

Step 3 is to formulate the model. We will let X_t represent the value of the
CM1 index t months after May 1986, and we will assume the linear regression
model in Eq. (8.30). The data are

$$(t_1, y_1) = (1, 6.73)$$
$$(t_2, y_2) = (2, 6.27)$$
$$\vdots \qquad\qquad\qquad (8.40)$$
$$(t_{37}, y_{37}) = (37, 8.44).$$

The best–fitting regression line can be obtained by solving the linear system of
equations in Eq. (8.34) to obtain a and b. Then the goodness–of–fit statistic
R^2 can be obtained from Eq. (8.39). Using a computer implementation of this
linear regression model will allow us to avoid a lot of tedious calculation.

Step 4 is to solve the model. We used the Minitab statistical package to
obtain the regression line

$$y = 5.45 + 0.0970t \qquad (8.41)$$

and

$$R^2 = 83.0\%$$

(see Figure 8.11 for selected outputs). To do this, first we entered the CM1 data
into one column of a Minitab worksheet and entered the time index numbers

$t = 1, 2, 3, \ldots, 37$ into another column. Then we used the pull-down menus to
issue the command `Stat > Regression > Regression` and specified the CM1
data as the response and the time index data as the predictor. To get the
prediction interval for $t = 48$, we selected the `Options` button in the regression
window and entered "48" in the box labelled `Prediction intervals for new`
`observations`. The details of this procedure and the resulting outputs are quite
similar if you use a different statistical package, a spreadsheet with a regression
tool, or even a hand calculator (although some features, like prediction intervals,
may not be available).

```
The regression equation is
cm1 = 5.45 + 0.0970 t

Predictor       Coef    SE Coef       T      P
Constant      5.4475     0.1615   33.73  0.000
t            0.096989  0.007409   13.09  0.000

S = 0.481203   R-Sq = 83.0%   R-Sq(adj) = 82.6%

Unusual Observations

Obs    t       cm1      Fit   SE Fit   Residual   St Resid
  1  1.0   6.7300   5.5445   0.1551     1.1855       2.60R

R denotes an observation with a large standardized residual.

Predicted Values for New Observations

New
Obs     Fit   SE Fit          95% CI              95% PI
  1  10.1030   0.2290  (9.6381, 10.5678)  (9.0211, 11.1848)X

X denotes a point that is an outlier in the predictors.

Values of Predictors for New Observations

New
Obs     t
  1   48.0
```

Figure 8.11: Solution to the ARM problem using the statistical package Minitab.

Equation (8.41) represents the best–fitting straight line through the data
points in Eq. (8.40). See Figure 8.12 for a graphical illustration. The average
trend in the CM1 index over the period June 1986 to June 1989 has been to
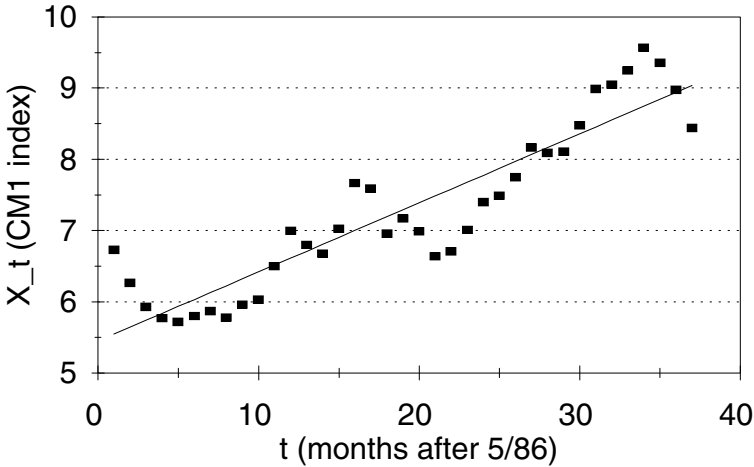
Figure 8.12: Graph of CM1 index versus time showing the regression line for the ARM problem.

increase by 0.0970 per month. Substituting $t = 48$ into Eq. (8.41), we obtain the estimate

$$EX_{48} = 5.45 + 0.0970(48) = 10.1$$

for the May 1990 CM1 index value. Since $R^2 = 83.0\%$, the regression equation accounts for 83% of the total variation in our CM1 index data. This gives us a fairly high level of confidence in our estimate of EX_{48}. Of course, the actual value of X_{48} will differ due to random fluctuations. More details about the likely magnitude of these fluctuations will be provided below in the sensitivity analysis.

Finally, step 5. We have concluded that the CM1 index shows a general trend, increasing by about 0.097 points per month. This figure is based on historical observations over the last three years. Projecting on this basis, we obtain the estimate of 10.1 for the May 1990 index figure. This is about 1.1 points higher than the May 1989 index, so in 1990 the author should expect his ARM payments to increase again.

The most important sensitivity analysis question here is the amount of random fluctuation in X_t. We are assuming the linear regression model in Eq. (8.30), where ε_t is mean zero normal. Our regression package estimated the standard deviation $\sigma \approx 0.4812$ on the basis of the data. In other words, $\varepsilon_t/0.4812$ is approximately standard normal. About 95% of the data points were no more than $\pm 2\sigma$ away from the line in Eq. (8.41). If this is representative of the magnitude of future fluctuations, then we would expect X_{48} to lie between $10.1 \pm 2\sigma$ with

95% confidence; i.e., we should see

$$9.1 \leq X_{48} \leq 11.1.$$

There is also a more sophisticated method built into the statistics package that takes into account the additional uncertainty involved in estimating EX_{48}. This method yields $9.02 \leq X_{48} \leq 11.19$ at the 95% confidence level. See Fig. 8.11.

Next we consider the sensitivity of our model to unusual data values. We are assuming the linear regression model from Eq. (8.30). Most of the time the random error ε_t will be small, but there is a small probability that ε_t will be rather large, so that one or more of our data points may lie far off the regression line. We need to consider the sensitivity of our procedure to such anomalous points, which are called *outliers*.

It is not hard to show that the regression line for a set of data points (t_1, y_1), ..., (t_n, y_n) will always pass through the point $(\bar{t}, \bar{y})$ defined by

$$\begin{aligned}
\bar{t} &= \frac{t_1 + \cdots + t_n}{n} \\
\bar{y} &= \frac{y_1 + \cdots + y_n}{n}.
\end{aligned} \tag{8.42}$$

In our model we have $\bar{t} = 19$ and $\bar{y} = 7.29$. The regression procedure selects the best–fitting line through the point $(19, 7.29)$. Since the essence of the procedure is to minimize the vertical distance between the regression line and the data points, an outlier will tend to pull the regression line toward itself, regardless of the location of the other data points. The situation gets worse as n gets smaller, because then each individual data point has more influence. It also gets worse as the distance from the base point $(\bar{t}, \bar{y})$ gets larger, since points farther out on the regression line have more leverage.

In Fig. 8.11 the statistical package Minitab has flagged the data point $(1, 6.73)$ as being unusual. If we plug $t = 1$ into the regression equation, we get

$$\hat{y}_1 = 5.45 + 0.0970(1) = 5.547.$$

The vertical distance or *residual* $y_1 - \hat{y}_1$ is 1.18, which means that this data point is about 2.6 standard deviations above the regression line. In order to ascertain the sensitivity of our model to this outlier, we repeat the regression calculation, leaving out the data point $(1, 6.73)$. Figure 8.13 shows the results of this sensitivity run.

The new regression equation is

$$EX_t = 5.30 + 0.103t$$

with $R^2 = 86.2\%$. The predicted value $EX_{48} = 5.30 + 0.103(48) = 10.24$ is our new estimate of the CM1 index on May 1990. The new residual standard deviation is 0.438450, so we expect that the CM1 index on May 1990 will be between $10.24 \pm 2(0.44)$. In other words, we are 95% sure that we will see

$$9.36 < X_{48} < 11.12.$$

```
The regression equation is
cm1 = 5.30 + 0.103 t

Predictor      Coef    SE Coef       T      P
Constant     5.3045     0.1554   34.13  0.000
t           0.102634  0.007034   14.59  0.000

S = 0.438450   R-Sq = 86.2%   R-Sq(adj) = 85.8%
```

Figure 8.13: Sensitivity analysis for the ARM problem using the statistical package Minitab.

The more sophisticated prediction interval computed by Minitab (not shown in Figure 8.13) is a bit wider: $9.24 \leq X_{48} \leq 11.22$. Either way, these conclusions are about the same as before, so we conclude that our model is not too sensitive to this outlier.

The main robustness issue here concerns our choice of a linear model in Eq. (8.30). More generally, we might assume that

$$X_t = f(t) + \varepsilon_t, \tag{8.43}$$

where $f(t)$ represents the true value of one–year U.S. Treasury bonds at time t, and ε_t represents fluctuations in the market. In this more general setting the linear regression model represents a linear approximation

$$f(t) \approx a + bt, \tag{8.44}$$

which is good near the base point $(\bar{t}, \bar{y})$. In Figure 8.11, Minitab has flagged the point $t = 48$ as being far away from the center point $t = 19$. We have data for $1 \leq t \leq 37$ and strong evidence of a linear relationship over this interval ($R^2 = 83\%$). In other words, the linear approximation in Eq. (8.44) involves at most a small percent error over this interval. As we move away from this interval, however, it must be expected that the error involved in this linear approximation gets worse. Another robustness issue arises from our assumption that the random errors ε_t are independent and identically distributed. A more complex model might take into account the dependence between these random variables. We will explore this issue in Section 8.5.

Our linear regression model is a simple example of a *time series model*. A time series model is a stochastic model of one or more variables that evolve over time. Most economic forecasting is done using time series models. More complex time series models represent the interaction of several variables and dependence in the random fluctuations of these variables. Time series analysis is a branch of statistics. In the next section, we provide an introduction to the essentials of time series analysis. For more details about time series models, a good reference is Box et al. (1976).

8.4   Time Series

A time series is a stochastic process that varies over time, usually observed at fixed intervals. Daily temperature and rainfall, monthly unemployment levels, and annual income are some typical examples of time series. The basic tool for modeling time series is linear regression, introduced in Section 8.3. For that reason, this section may be considered as a follow-up to Section 8.3 that introduces some additional applications and methods. Indeed, the example we consider next is an extension of the ARM problem from that section. This section requires a numerical implementation of multiple regression; i.e., linear regression with more than one predictor. This is available in a statistical package (e.g., Minitab, SAS, SPSS) or a spreadsheet (e.g., Excel). Since it is not reasonable to perform these problems by hand, a full set of computational formulas will not be given.

Example 8.6. Reconsider the ARM problem of Example 8.5 but now consider the relationship between the mortgage index at different times. Answer the same question as before: Estimate the value of the CM1 index in May 1990 using the data provided on the CM1 index values from June 1986 to June 1989.

We will use the five-step method. Step 1 is the same as before, except that now we also want to take into account the dependence between the CM1 index at different times. We are attempting to estimate the future trend in a variable that exhibits a tendency to grow with time, along with some random fluctuation. Let X_t denote the U.S. Treasury one–year Constant Maturity (CM1) index at time t months after May 1986. A graph of X_t for $t = 1, \ldots, 37$ was shown in Figure 8.10. We want to estimate X_{48}. We will assume that X_t depends on the time t, the previous values $X_{t-1}, X_{t-2}, \ldots$ and a random element. Then we want to predict the average EX_{48} along with a suitable estimate of uncertainty.

Step 2 is to select the modeling approach. We will model this problem as a time series and fit an autoregressive model.

> A time series is a sequence of random variables $\{X_t\}$ that varies over time $t = 0, 1, 2, \ldots$ according to some random pattern. The key to time series modeling is to recognize the pattern. A typical assumption is that the pattern involves a *trend* added to a *stationary time series*. The trend is a non-random function that varies over time, and it represents the mean value of the series. Once the trend is removed, we are left with a mean zero time series, and we want to model its dependence structure. The simplest case is where the remaining time series consists of independent random variables. However, it is typical to find dependence between these variables. Dependence is measured in terms of the *covariance*. For two random variables X_1 and X_2, the covariance
>
> $$\text{Cov}(X_1, X_2) = E[(X_1 - \mu_1)(X_2 - \mu_2)],$$
>
> where $\mu_i = E(X_i)$, the expected value or mean. The covariance measures the linear relation between the two variables. If X_1 and

X_2 are independent, then the covariance $\text{Cov}(X_1, X_2) = 0$. A positive covariance indicates that a higher–than–average value of X_1 is likely to be found in the company of a higher–than–average value of X_2. Likewise, a low value of X_1 will typically be found along with a low value of X_2. For example, if X_1 is the income of an individual and X_2 is their income tax, then $\text{Cov}(X_1, X_2)$ would be positive. You cannot infer the income tax knowing only the income, but it is a good bet that someone with a high income pays more tax, and someone with a low income pays less. Mathematically, $(X_1 - \mu_1)$ is the deviation of income from the average, and likewise for $(X_2 - \mu_2)$. The covariance averages the product of these. If one tends to be positive when the other is positive, and negative when the other is negative, then the covariance is positive, indicating a *positive relation*. For another example, the median price X_1 of a home in a town is negatively correlated to the percentage of families X_2 owning their own home. Here μ_1 is the median home price over all towns, and μ_2 is the average percentage of families owning their own home, taken over all the towns. When $X_1 - \mu_1$ is positive, then $X_2 - \mu_2$ is likely to be negative, and vice versa, so their average $\text{Cov}(X_1, X_2)$ will be negative, indicating a *negative relation* between these two variables. We also note that correlation only captures a linear relation. Consider the case where X_1 is the air pressure in a car tire and X_2 is the tread life. If X_1 is near the mean μ_1, the recommended air pressure, then X_2 will be the highest. If X_1 is either smaller or larger than its mean, then X_2 will decrease. The covariance does not capture this kind of dependence. Finally, we note that the covariance is also a generalization of the variance, in the sense that $\text{Var}(X) = \text{Cov}(X, X)$.

A close cousin of the covariance is the *correlation*

$$\rho = \text{corr}(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sigma_1 \sigma_2} = E\left[\frac{(X_1 - \mu_1)}{\sigma_1} \frac{(X_2 - \mu_2)}{\sigma_2}\right]$$

which is a dimensionless version of the covariance. Here $\sigma_i^2 = \text{Var}(X_i)$ is the variance, so that σ_i is the standard deviation of the random variable X_i. Since both μ_i and σ_i have the same units as X_i, the units cancel, leaving a dimensionless measure of dependence. Again, if X_1 and X_2 are independent, then $\text{corr}(X_1, X_2) = 0$, and we say that X_1 and X_2 are uncorrelated. It can also be shown that the correlation satisfies $-1 \leq \rho \leq 1$ in all cases, the extreme cases $\rho = \pm 1$ corresponding to the case of perfect dependence where X_2 is a linear function of X_1. If $\rho > 0$, we say that X_1 and X_2 are *positively correlated*, and if $\rho < 0$, we say that they are *negatively correlated*.

Correlation is a useful measure of dependence in a time series. In that context we call $\rho(t, h) = \text{corr}(X_t, X_{t+h})$ the *autocorrelation*

function of the time series. It measures the serial dependence between the times series at different times. The time series is called *stationary* (or sometimes *weakly stationary*) if the mean $E(X_t)$ and the autocorrelation function $\rho(h, t)$ are constant over time. In time series analysis, it is often necessary to *detrend* the series to get something that is stationary. In Example 8.5 we detrended the CM1 times series X_t by using regression to identify a linear trend $a + bt$, leaving a zero mean (centered) error term ε_t that we modeled as independent and identically distributed. In the context of time series analysis, this is called *random noise*. It is the simplest centered times series. More generally, one might hope that the centered series is at least stationary, with a correlation structure that remains the same over time. There are various tests for stationarity, but the simplest is just to plot the errors ε_t over time to see whether they appear to follow a consistent (random) pattern. One typical indication of nonstationarity would be a widening or narrowing of the distribution of ε_t over time. This is called *heteroscedasticity*, which just means that the variance changes. Once we are satisfied that a centered time series is stationary, we attempt to model its covariance structure. The simplest useful model for this is called an *autoregressive process*

$$X_t = a + bt + c_1 X_{t-1} + \cdots + c_p X_{t-p} + \varepsilon_t \qquad (8.45)$$

where for convenience, we have also included the trend. The parameter p is called the *order* of the autoregressive process, sometimes abbreviated as AR(p). In the context of linear regression, we can fit the parameters of an autoregressive process by regressing the observations X_t against multiple predictors. The first is the time t, as in Section 8.3. The remaining predictors are the previous observations $X_{t-1}, \ldots, X_{t-p}$. Now the trick is to choose a reasonable value for the parameter p, and there are two main ways to do this. One way is to look at the R^2 value, which measures how much of the variability in X_t is captured by the predictors. Since any additional data will give at least a slightly better prediction, adding a predictor will always make R^2 increase. However, a very small increase is not worth the bother of carrying another predictor, so one could just add predictors X_{t-1}, X_{t-2}, and so forth one by one, until the additional improvement in R^2 is minimal. Some packages also output an adjusted R^2 value that includes a penalty for an increased number of predictors. If this is available, one can simply add predictors until the adjusted R^2 begins to decrease (or more generally, consider several models and pick the one with the largest adjusted R^2). Such packages also typically list the *sequential sums of squares* that can be interpreted as an extension of the R^2 statistic. Recall that for a simple (one predictor) linear regression model, the formula (8.39) for R^2 is the sum of the squares of the regressions $(\hat{y}_i - \bar{y})$ divided by the sum of the squares of the total variations $(y_i - \bar{y})$. The sequential sums of

squares are just the individual components in the sum of the squares of the regressions, one predictor at a time, to measure the additional variation captured by each additional predictor. Looking at the sequential sums of squares is another way to measure the additional value of another predictor, essentially equivalent to looking at the change in the R^2 value. Additional information is contained in the *p–values* for each of the estimated regression coefficients a, b, and c_i. The *p*–value indicates the likelihood that this parameter value could have occurred by chance, even though the predictor does not belong in the model (or equivalently, it belongs in the model with a coefficient of zero). Hence, a small *p*–value (say, $p < 0.05$) indicates strong evidence that the predictor belongs in the model. However, this is less important then the R^2 indicator, since a predictor can be statistically significantly related to the dependent variable X_t we are trying to predict, and still add just a small amount of information. Hence, we might not consider it worthwhile to include. This is the principle of *parsimony*: Make the model as simple as possible, without sacrificing the ability to predict.

The second method for autoregressive modeling is to consider the *residuals* from the time series model, which are simply our estimates of the error term ε_t. Once we have determined our estimates of the parameters a, b, and c_i from the regression, we can use the formula

$$\varepsilon_t = X_t - (a + bt + c_1 X_{t-1} + \cdots + c_p X_{t-p})$$

to estimate the errors. Since our goal is to include enough predictors in the model to capture the dependence structure, we would like the resulting ε_t sequence to be an uncorrelated noise sequence. We can check this by computing the autocorrelation function of the residuals. Most statistical packages will automatically compute the residuals and their autocorrelation function, as well as *p*–values or error bars for the autocorrelation function $\rho(h)$. The error bars indicate the likely values of $\rho(h)$ in the case of a noise sequence, and values outside these bars (also indicated by a low *p*–value, typically less than 0.05) indicate a statistically significant correlation. Since the autocorrelation function assumes a stationary process, it is also advisable to check for this using graphical displays of the residuals.

Step 3 is to formulate the model. We will model the CM1 index X_t at time t months after May 1986 as an autoregressive time series with a linear trend. Hence, we are assuming that Eq. (8.45) holds for some constants a, b, $c_1, \ldots, c_p$ and some noise sequence ε_t. In order to choose a suitable value for the parameter p, we will consider a sequence of increasingly complex models $p = 0, 1, 2, \ldots$ until we achieve a satisfactory result, indicated by a model with (hopefully) a small number of predictors that appears to have an uncorrelated noise sequence of residuals. Then we can proceed to estimate (forecast) the
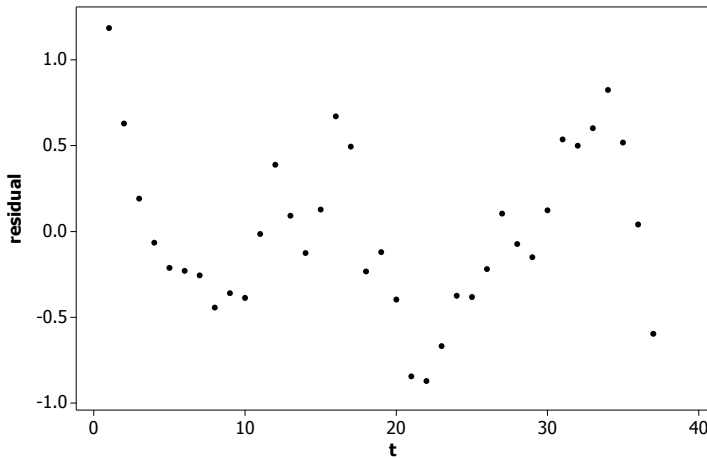
Figure 8.14: Graph of CM1 residuals ε_t versus time for the simple regression model of equation (8.30).

value X_t of the CM1 index at month $t = 48$, which is May 1990, along with appropriate error bounds.

Step 4 is to solve the problem. We will use the statistical package Minitab which has convenient facilities for multiple regression, time series analysis, and graphical display. We begin by examining in more detail the results of the modeling in Example 8.5. There, we fit a simple linear regression model of the form $X_t = a + bt + \varepsilon_t$. The results are summarized in Figure 8.11. The best–fit regression line is given by $a = 5.45$ and $b = 0.097$, indicating an upward trend in the CM1 index. The statistic $s = 0.48$ estimates the standard deviation of the errors ε_t, and the R^2 statistic indicates that the trend predicts 83% of variations in the CM1 index X_t. The p–values for both a and b are given as 0.000, indicating strong statistical evidence that these parameters are different from zero. Our subsequent analysis and predictions were based on the simple regression model, which assumes that the errors ε_t form a noise sequence, independent and identically distributed. We will now test this assumption using graphical displays and the correlation function.

The residuals or estimated errors are computed as described in Section 8.3, using the estimated values of a and b and the equation $\varepsilon_t = X_t - (a + bt)$. The predicted values $\hat{y}_t = a + bt$ are plotted in Figure 8.12 along with the original data. The residuals are simply the vertical deviations $(y_t - \hat{y}_t)$ of the data values from the regression line, where $y_t = X_t$ is the t-th observation of the CM1 index. For example, the second data value is $y_2 = 6.27$ and the fitted value is $\hat{y}_2 = 5.45 + 0.097(2) = 5.64$, so the residual is $y_2 - \hat{y}_2 = 0.63$, indicating

the the second data point in Figure 8.12 is 0.63 units above the regression line.

Figure 8.14 shows a plot of the residuals. They were computed in Minitab by clicking the `Storage` button in the regression window and checking the box labelled `Residuals`. The graph was also prepared in Minitab using the command `Graph > Scatterplot`. The data appear stationary in the sense that the spread of values does not seem to increase or decrease with t. However, it does appear that there may be some serial dependence, particularly for $t \geq 20$ where there seems to be an upward trend. This could indicate many things, including nonstationarity, a change in the dependence structure of the series, or some correlation.

To investigate further, we compute the autocorrelation function of the residuals. Figure 8.15 shows the results of this calculation using the Minitab command `Stat > Time Series > Autocorrelation` applied to the residuals stored as part of the regression calculation. The vertical bars indicate the correlation function $\rho(h) = \text{corr}(\varepsilon_t, \varepsilon_{t+h})$ as a function of the time lag $h = 1, 2, 3, \ldots$ and the dotted curve represents the 95% error bars. The autocorrelation function plotted here is, of course, a statistical estimate of the model autocorrelation, and the error bars show the range of normal variation for this statistical estimate for an uncorrelated noise sequence. Hence, a value outside the bars indicates strong statistical evidence for a nonzero autocorrelation. In Figure 8.15, the first value $\rho(1)$ is well outside the error bars. This is an indication of serial correlation in the residuals ε_t from our simple regression, evidence that a more complex model is needed to obtain a simple uncorrelated (white) noise sequence. This kind of strong positive correlation can cause a pattern that looks like a trend, as seen in Figure 8.14, since a large positive value of ε_t makes it more likely that the next value is large, and so forth.

Now we proceed to consider a more complicated autoregressive times series model (8.45) for the CM1 index time series data. Our goal is to find a number of predictors p in terms of the past history of the process that is in some sense optimal. We begin by repeating the regression procedure with $p = 1$ additional predictor X_{t-1}. First, we prepare another column of data X_{t-1}; that is, the CM1 index shifted downward by one place. Use a simple copy and paste or the Minitab command `Stat > Time Series > Lag`. It is necessary to leave off the last CM1 index value in the lag one column X_{t-1}, since the statistical package requires predictors to be the same data length as the predicted data (and, of course, the first entry is blank or missing, in Minitab this is denoted by a $*$ in that data cell). The details vary with different packages, but the steps are similar. Now we repeat the regression command and store the residuals.

Figure 8.16 displays portions of the Minitab computer output. The regression equation is $X_t = 1.60 + 0.033t + 0.698X_{t-1}$, where X_t is the CM1 index t months after May 1986. This indicates an upward trend along with some serial dependence between the CM1 index on subsequent months. The statistic $R^2 = 94.1\%$ indicates that the combination of trend and the CM1 index from last month predicts 94.1% of the variations in the CM1 index this month. This is a significant improvement over the value $R^2 = 83.0\%$ for the simple linear regression model of Example 8.5.
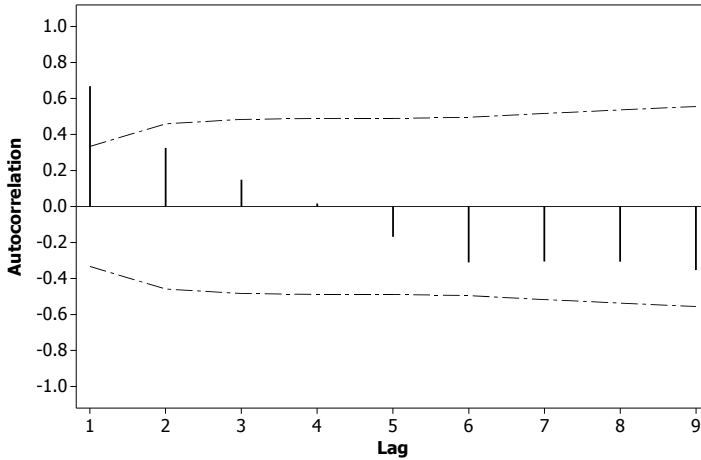
Figure 8.15: Autocorrelation function for CM1 residuals ε_t versus lag for the simple regression model of equation (8.30).

The adjusted $R^2 = 93.8\%$ here is also higher than the 82.6% figure from Example 8.5, additional evidence that the autoregressive model is superior. The analysis of variance statistics give more detail on the R^2 calculation. Recalling the formula (8.39), we see that R^2 is the ratio of two sums of squares. The sum of squared regressions $\sum_i (\hat{y}_i - \bar{y})^2 = 44.676$ and the sum of squares of total variation is $\sum_i (y_i - \bar{y})^2 = 47.460$. The ratio of these two is $R^2 = 44.676/47.460 = 0.941$. The sequential sums of squares table indicates that 40.924 of the 44.676 residual sum of squares comes from the first predictor t, and an additional 3.752 comes from the addition of a second predictor X_{t-1}. If we had listed the two predictors in the reverse order, then these two figures would change, since the two predictors t and X_{t-1} are not completely independent, but they would still add to 44.767. We also note that the p-value for the constant $a = 1.5987$ is 0.008, the p-value for the coefficient $b = 0.03299$ of the predictor t is 0.007, and the p-value for the coefficient $c_1 = 0.6977$ of X_{t-1} is 0.000. This is additional evidence that all these coefficients are statistically significantly different than zero and should be included in the model.

Figure 8.17 shows the fitted model $1.60 + 0.033t + 0.698X_{t-1}$ plotted along with the data X_t to check the quality of the fit. It appears that the fit is better than the simple regression line; compare Figure 8.12. Now a prediction of X_{48} can be accomplished by iterating the equation $X_t = 1.60 + 0.033t + 0.698X_{t-1}$

The regression equation is
CM1 = 1.60 + 0.0330 t + 0.698 X(t-1)

```
Predictor      Coef  SE Coef    T       P
Constant     1.5987   0.5652  2.83   0.008
t           0.03299  0.01144  2.88   0.007
X(t-1)       0.6977   0.1046  6.67   0.000
```

S = 0.290471   R-Sq = 94.1%   R-Sq(adj) = 93.8%

Analysis of Variance

```
Source            DF     SS      MS       F       P
Regression         2  44.676  22.338  264.75  0.000
Residual Error    33   2.784   0.084
Total             35  47.460
```

```
Source  DF  Seq SS
t        1  40.924
X(t-1)   1   3.752
```

Figure 8.16: Autoregressive model for the ARM problem using the statistical package Minitab.

starting at $t = 37$ (e.g., use a hand calculator or a spreadsheet). This yields

$$X_{38} = 1.60 + 0.0330(38) + 0.698(8.44) = 8.75$$
$$X_{39} = 1.60 + 0.0330(39) + 0.698(8.75) = 8.99$$
$$\vdots$$
$$X_{48} = 1.60 + 0.0330(48) + 0.698(10.16) = 10.28$$

so we predict that the value of the CM1 index for May 1990 will be 10.28. Using the model $X_t = 1.60 + 0.033t + 0.698X_{t-1} + \varepsilon_t$ along with the fact that the estimated standard deviation of ε_t is 0.290471, we predict (with 95% certainty) that the value of the CM1 index on May 1990 will be between $10.28 \pm 2(0.29)$, or in other words, between 9.7 and 10.9. This is a more precise estimate than Example 8.5 because the regression model fits more closely, giving a smaller error standard deviation. Graphically, the standard deviation is simply the typical vertical variation of the data points from the prediction line, so a closer fit gives a smaller standard deviation. However, the bottom line is the same as before: The author's ARM interest rate is likely to increase again in 1990. In fact, this improved model gives a slightly higher estimate for the May 1990 CM1 index.

Next we examine the model residuals to determine whether they resemble an
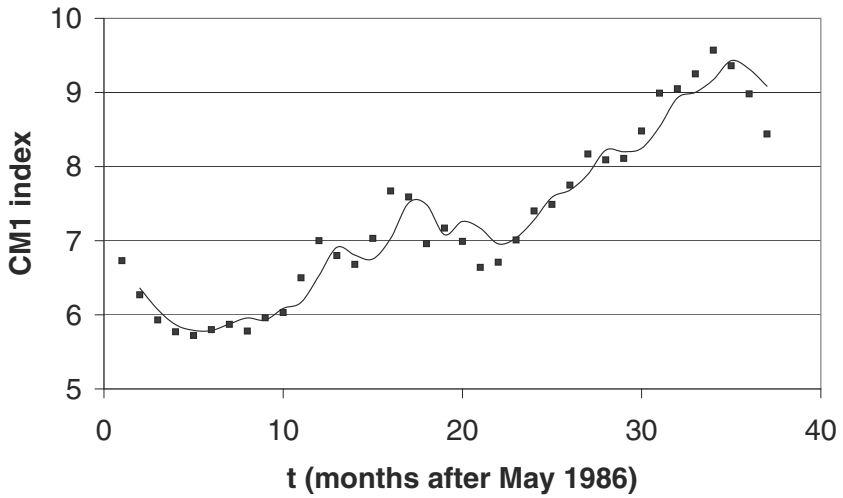
Figure 8.17: Graph of CM1 index versus time t showing the fitted autoregressive model of equation (8.45) with $p = 1$.
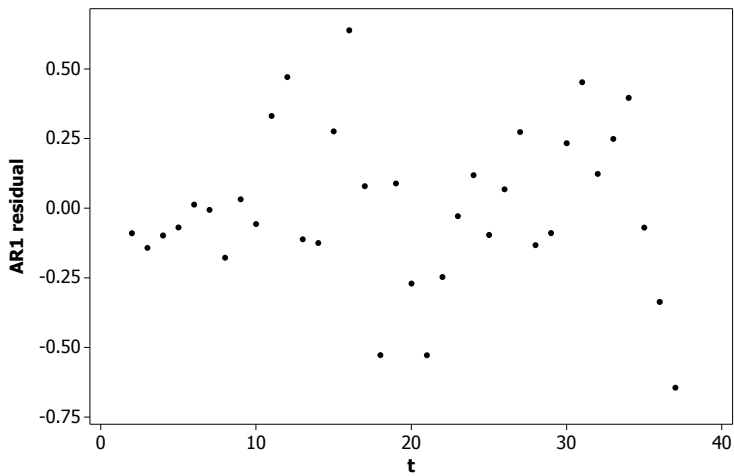
Figure 8.18: Graph of CM1 residuals ε_t versus time for the autoregressive model of equation (8.45) with $p = 1$.

uncorrelated (white) noise sequence. Figure 8.18 shows a plot of the residuals for this AR(1) model versus time. This plot is more satisfactory than Figure 8.14, with no significant evidence of serial correlation. We can also examine the autocorrelation function of these residuals as we did before, and this plot (not shown) indicates no serial correlation, since all the autocorrelations are within the error bars. Hence, it is reasonable to conclude that the AR(1) model (8.45) with $p = 1$ captures nearly all of the dependence in the CM1 time series. To validate this conclusion, we may also consider a further regression model with three predictors: t, X_{t-1}, and X_{t-2}. We leave the details for an exercise (see Exercise 8.19).

Finally, we come to Step 5. Example 8.6 may be considered as part of the robustness analysis of Example 8.5. There we predicted that the value of the CM1 index on May 1990 would be 10.1, about 1.1 points higher than the May 1989 index value of 8.98. The more sophisticated model considered here gives a refined estimate of 10.2, a little higher. We also recall that in Example 8.5 we produced a prediction interval of (9.1,11.1) with 95% confidence. The refined model yields a significantly tighter interval of (9.7,10.9). Putting this all together, we expect that the CM1 index for May 1990 is likely to be a bit more than one point higher than the May 1989 value, and we are reasonably sure it will go up by at least 3/4 of a percent.

Sensitivity analysis could include many factors. For example, two of the observations at $t = 16, 37$ were flagged by Minitab for large residuals (both were more than 2.2 standard deviations from the mean). Hence, we could repeat the analysis with these values deleted, to see if this makes any significant difference. We could consider a different trend function like $a + bt + ct^2$ (see Exercise 16) or at^b (see Exercise 18 where we apply this trend to the response time data for the facility location problem). We could also add more predictors $X_{t-2}, X_{t-3}, \ldots$ and check to see if this makes a big difference (see Exercise 8.19). The possibilities are literally endless, and hence, some judgment is called for. This is where the principle of parsimony comes in. Our goal is to get a reasonable estimate of the CM1 index for May 1990, based on the available data. Refining the simple regression model of Example 8.5 to the AR(1) model of Example 8.6 was probably worthwhile, not so much for the improved point estimate of 9.2 instead of 9.1, but rather for the significantly tighter prediction interval. If we only cared about the point estimate, then simple linear regression is probably good enough. Whether it is worth while to continue in this vein, to consider more alternative models with more predictors and/or more complicated trend functions, is less clear. In the real world, sensitivity analysis can continue as long as time and funds permit, but at some point the intelligent modeler will declare victory and move on to a new challenge.

Many interesting robustness questions are important in the analysis of real–world time series. One question is the trend. While we use a linear trend, other options such as a higher order polynomial (see Exercise 16) or a nonlinear trend (see Exercise 18) can also be considered. Adding more parameters will always improve the fit, so care is required. Examining the adjusted R^2 value is one way to avoid over-parameterization. Often the choice of trend function will

depend on the applications. For example, for income or population data, one might expect exponential rather than linear growth. Another important issue is change point analysis: Did the underlying correlation structure or trend function of the time series change at some point during the data collection period? For example, this is an important part of the debate over global warming. Time series is a growing field both in applications and theory. A good place to start learning more about the underlying theory is Brockwell and Davis (1991).

8.5   Exercises

1. Reconsider the inventory problem of Example 8.1, but now suppose that the store policy is to order more aquariums if there are less than two left in stock at the end of the week. In either case (zero or one remaining) the store orders enough to bring the total number of aquariums in stock back up to three.

   (a) Calculate the probability that the demand for aquariums in a given week exceeds the supply. Use the five-step method, and model as a Markov chain in steady state.

   (b) Perform a sensitivity analysis on the demand rate λ. Calculate the steady–state probability that demand exceeds supply, assuming $\lambda = 0.75$, 0.9, 1.0, 1.1, and 1.25, and display in graphical form, as in Fig. 8.5.

   (c) Let p denote the steady–state probability that demand exceeds supply. Use the results of part (b) to estimate $S(p, \lambda)$.

2. (Requires a computer algebra system) Reconsider the inventory problem of Example 8.1. In this problem we will explore the sensitivity of the probability p that demand exceeds supply to the demand rate λ.

   (a) Draw the state transition diagram for an arbitrary λ. Show that Eq. (8.11) is the appropriate state transition probability matrix for this problem.

   (b) Write down the system of equations analogous to Eq. (8.10) that must be solved to get the steady–state distribution for a general λ. Use a computer algebra system to solve these equations.

   (c) Let p denote the steady–state probability that demand exceeds supply. Use the results of part (b) to obtain a formula for p in terms of λ. Graph p versus λ for the range $0 \leq \lambda \leq 2$.

   (d) Use a computer algebra system to differentiate the formula for p obtained in part (c). Calculate the exact sensitivity $S(p, \lambda)$ at $\lambda = 1$.

3. Reconsider the inventory problem of Example 8.1, but now suppose that the inventory policy depends on recent sales history. Whenever inventory

drops to zero, the number of units ordered is equal to two plus the number sold over the past week, up to a maximum of four.

(a) Determine the steady–state probability distribution of the number of aquariums in stock. Use the five-step method, and model as a Markov chain.

(b) Determine the steady–state probability that demand exceeds supply.

(c) Determine the average size of a resupply order.

(d) Repeat parts (a) and (b), but now suppose that weekly demand is Poisson with a mean of 2 customers per week.

4. Reconsider the inventory problem of Example 8.1, but now suppose that three additional aquariums are ordered any time that there are less than two in stock at the end of the week.

(a) Determine the probability that demand exceeds supply on any given week. Use the five-step method, and model as a Markov chain in steady state.

(b) Use the steady–state probabilities from part (a) to calculate the expected number of aquariums sold per week under this inventory policy.

(c) Repeat part (b) for the inventory policy in Example 8.1.

(d) Suppose that the store makes a profit of $5 per 20–gallon aquarium sold. How much would the store gain by implementing the new inventory policy?

5. For the purposes of this problem, we will consider the stock market to be in one of three states:

$$
\begin{array}{ll}
1 & \text{Bear market} \\
2 & \text{Strong bull market} \\
3 & \text{Weak bull market}
\end{array}
$$

Historically, a certain mutual fund gained -3%, 28%, and 10% annually when the market was in states 1, 2, and 3 respectively. Assume that the state transition probability matrix

$$
P = \begin{pmatrix}
0.90 & 0.02 & 0.08 \\
0.05 & 0.85 & 0.10 \\
0.05 & 0.05 & 0.90
\end{pmatrix}
$$

applies to the weekly change of state in the stock market.

(a) Determine the steady–state distribution of market state.

(b) Suppose that $10,000 is invested in this fund for ten years. Determine the expected total yield. Does the order of state transitions make any difference?

(c) In the worst–case scenario, the long–run expected proportion of time in each market state is 40%, 20%, and 40%, respectively. What is the effect on the answer to part (b)?

(d) In the best–case scenario, the long–run expected proportion of time in each market state is 10%, 70%, and 20%, respectively. What is the effect on the answer to part (b)?

(e) Does this mutual fund offer a better investment opportunity than a money market fund currently yielding about 8%? Consider that the money market fund offers a lower risk.

6. A Markov chain model of floods uses the state variable $X_n = 0, 1, 2, 3, 4$ where state 0 means the average daily flow is below 1,000 cubic feet per second, state 1 is 1,000–2,000 cfs, state 2 is 2,000–5,000 cfs, 3 is 5,000–10,000 cfs, and 4 is over 10,000 cfs. The state transition probability matrix for this model is

$$P = \begin{pmatrix} 0.9 & 0.05 & 0.025 & 0.015 & 0.01 \\ 0.3 & 0.7 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0.6 & 0.4 & 0 \\ 0 & 0 & 0 & 0.8 & 0.2 \end{pmatrix}.$$

(a) Draw the state transition probability diagram for this model.

(b) Find the steady–state probability distribution for this model.

(c) How often are severe floods (over 10,000 cfs) expected to occur?

(d) A reservoir used for drought storage depends on the flow of this river. When the flow is over 5,000 cfs, the reservoir is allowed to store 1,000 acre–feet per day. When the flow is under 1,000 cfs, the reservoir is required to release 100 acre–feet per day back into the river. Find the expected annual number of acre–feet of water stored in the reservoir. Is this number positive or negative? What does this mean?

7. This exercise shows the equivalence of the two formulations of a Markov process model.

(a) Suppose that $T_{i1}, \dots, T_{im}$ are independent random variables and that T_{ij} is exponential with rate parameter $a_{ij} = p_{ij}\lambda_i$. Assume that $\sum p_{ij} = 1$ and $\lambda_i > 0$. Show that

$$T_i = \min(T_{i1}, \dots, T_{im})$$

has an exponential distribution with rate parameter λ_i. [Hint: Use the fact that

$$\Pr\{T_i > x\} = \Pr\{T_{i1} > x, \dots, T_{im} > x\}$$

for all $x > 0$.]

(b) Suppose that $m = 2$ so that $T_i = \min(T_{i1}, T_{i2})$ and show that $\Pr\{T_i = T_{i1}\} = p_{i1}$. [Hint: Use the fact that

$$\Pr\{T_i = T_{i1}\} = \Pr\{T_{i2} > T_{i1}\}$$
$$= \int_0^\infty \Pr\{T_{i2} > x\} f_{i1}(x)\, dx,$$

where $f_{i1}(x)$ is the probability density function of the random variable T_{i1}.]

(c) Use the results of (a) and (b) to show that in general $\Pr\{T_i = T_{ij}\} = p_{ij}$.

8. Reconsider the forklift problem of Example 8.3, but now suppose that a second mechanic is called in whenever two or more trucks need repair.

   (a) Determine the steady–state distribution of the number of trucks needing repair. Use the five-step method and model as a Markov process.

   (b) Use the results of part (a) to calculate the steady–state expected number of forklift trucks in the repair facility, the probability that the first mechanic is busy, and the probability that the second mechanic is called in. Compare to the results obtained for one mechanic in Example 8.3.

   (c) The cost of the second mechanic is $250 per day, and the mechanic is only paid for the days worked. Another option which can be used in case of backlog (two or more trucks in repair) is to lease replacement vehicles for customers whose trucks are in repair, at a cost of $125 per week per vehicle. Which of the two plans is most cost-effective?

   (d) There is some uncertainty as to the actual cost of bringing in a second mechanic during periods of backlog. What is the minimum cost per day for a second mechanic that makes leasing a better alternative?

9. Five locations are connected by radio. The radio link is active 20% of the time, and there is no radio activity the remaining 80% of the time. The main location sends radio messages with an average duration of 30 seconds, and the remaining four locations send messages that average 10 seconds in length. Half of all radio messages originate at the main location, with the remaining proportion equally divided among the other four locations.

   (a) For each location, determine the steady–state probability that this location is sending a message at any given time. Use the five-step method, and model as a Markov process.

   (b) A monitoring station samples from the radio emissions on this network once every five minutes. How long on average does it take until the monitor finds a message in progress from one particular location?

(c) The monitor can identify the source of a radio message if the message lasts at least three seconds after the monitor begins listening. How long does it take until the monitor locates one particular location?

(d) Perform a sensitivity analysis to see how the results of part (c) are affected by the percent utilization of the radio frequency (currently 20%).

10. A gasoline service station has two pumps, each of which can service two cars at a time. If all pumps are busy, cars will queue up in a single line to await a free pump. Since the station operates in a competitive environment, it can be expected that customers encountering long lines at this station will take their business elsewhere.

(a) Construct a model that can be used to predict both the steady–state probability of a waiting line and the expected length of the line. Use the five-step method, and model as a Markov process. You will have to make some additional assumptions about customer demand, service time, and balking (refusing to join a queue).

(b) Use the model of part (a) to estimate the fraction of potential business lost because of customer balking. Consider a range of possible levels of customer demand.

(c) What is the easiest way to infer the level of customer demand (i.e., potential sales) from data to which the station manager has access?

(d) Under what circumstances would you recommend that the station purchase additional pumps?

11. A certain form of one–celled creature reproduces by cell division, producing two offspring. The mean lifetime before cell division is one hour, and each individual cell has a 10% chance of dying before it can reproduce.

(a) Construct a model that represents the evolution of population size over time. Use a Markov process, and draw a rate diagram.

(b) Describe in general terms what you would expect to happen to the population level over time.

(c) What is the problem with applying steady–state results to this model?

12. Table 8.2 gives per capita income in 1982 dollars and population density in population per square miles for the ten poorest counties in Asia (source: Webster's New World Atlas, (1988)).

(a) Does the data support the proposition that prosperity is linked to population density? Use linear regression to obtain a formula that predicts per capita income as a linear function of population density.

(b) What percentage of the total variation in per capita income can be attributed to variations in population density?

Country	Per Capita Income (in 1982 dollars)	Population Density (pop. per sq. mi.)
Nepal	168	290
Kampuchea	117	101
Bangladesh	122	1740
Burma	171	139
Afghanistan	172	71
Bhutan	142	76
Vietnam	188	458
China	267	284
India	252	578
Laos	325	45

Table 8.2: Per capita income versus population density for the 10 poorest Asian nations.

(c) What is the effect on your answers to parts (a) and (b) if we leave out of our analysis the country (Bangladesh) with population density of more than 1,000 per square mile?

(d) On the basis of your regression model, estimate the potential benefit for the citizens of a poor Asian country that manages to reduce its population by 25%.

13. Reconsider Exercise 12(a), but now obtain the regression line by solving the underlying optimization problem by hand. Letting t_i denote the population density and y_i the per capita income of country i, the goodness–of–fit for a candidate regression line

$$y = a + bt$$

is given by Eq. (8.33) in the text. Plug in the data points (t_i, y_i) and compute the function $F(a, b)$. Then obtain the best–fitting line by minimizing $F(a, b)$ over the set of all $(a, b) \in \mathbb{R}^2$.

14. Reconsider the ARM problem of Example 8.5. Suppose that only the data for the period June 1986 to June 1988 were known, and we were attempting to predict the May 1989 value of the CM1 index.

(a) Use a computer or calculator implementation of linear regression to obtain the regression line for this data.

(b) What is the predicted value of the May 1989 CM1 index according to the regression model of part (a)?

(c) What is the value of R^2 for the model in part (a)? How would you interpret this value?

(d) Compare to the actual value of the May 1989 CM1 index. How close was the predicted value? Was it within two standard deviations?

15. Repeat Exercise 14, but use the numbers for the TB3 index.

16. Reconsider the ARM problem of Example 8.5. Use a computer program for multiple regression to predict future trends in the CM1 index by fitting a second–degree polynomial to the data. After you input the time index $t = 1, 2, 3, \ldots$ in one column, and the CM1 data in a second column, prepare a third column of data containing the second power of the time index numbers $t^2 = 1, 4, 9, \ldots$ Multiple regression will give the best–fitting second–degree polynomial

$$\text{CM1} = a + bt + ct^2,$$

and R^2 can be interpreted as before. This technique is known as *polynomial least squares.*

   (a) Use a computer implementation of multiple linear regression to obtain a formula that predicts the CM1 index as a quadratic function of t.

   (b) Use the formula obtained in part (a) to predict the expected May 1990 value for the CM1 index.

   (c) What percent of the total variation in the CM1 index can be accounted for by using this model?

   (d) Compare the R^2 value for this multiple regression model with what was done in Section 8.3. Which model gives the best fit?

17. (Response–time formula from Example 3.2) A suburban community intends to replace its old fire station with a new facility. As part of the planning process, response–time data were collected for the past quarter. It took an average of 3.2 minutes to dispatch the fire crew. The dispatch time was found to vary only slightly. The time for the crew to reach the scene of the fire (drive time) was found to vary significantly depending on the distance to the scene. The data on drive time are displayed in Table 8.3.

   (a) Use linear regression to obtain a formula that predicts drive time as a linear function of distance traveled. Then determine a formula for total response time, including dispatch time.

   (b) What percentage of the total variation in drive time is accounted for by the formula you came up with in part (a)?

   (c) Draw a graph of drive time versus distance for the data in Table 8.3. Do the data seem to indicate a linear trend?

   (d) Plot the regression line from part (a) on the graph made in part (c). Does the line seem to be a good predictor for this data?

Distance (miles)	Drive Time (minutes)
1.22	2.62
3.48	8.35
5.10	6.44
3.39	3.51
4.13	6.52
1.75	2.46
2.95	5.02
1.30	1.73
0.76	1.14
2.52	4.56
1.66	2.90
1.84	3.19
3.19	4.26
4.11	7.00
3.09	5.49
4.96	7.64
1.64	3.09
3.23	3.88
3.07	5.49
4.26	6.82
4.40	5.53
2.42	4.30
2.96	3.55

Table 8.3: Response time data for the facility location problem.

18. (Continuation of Exercise 17) Another way to get a formula that relates drive time d to distance r is to use a power law model. Suppose that the underlying relationship between d and r is of the form $d = ar^b$. Taking logarithms of both sides yields the relationship

$$\ln d = \ln a + b \ln r.$$

Then linear regression can be used to estimate the parameters $\ln a$ and b in this linear equation.

(a) Transform the data in Table 8.3 by taking logarithms of both the drive time d and the distance r. Plot $\ln d$ versus $\ln r$. Does your graph suggest a linear relationship between $\ln d$ and $\ln r$?

(b) Use linear regression to obtain a formula that predicts $\ln d$ as a linear function of $\ln r$. Then determine a formula for total response time, including dispatch time, as a function of the distance r to the fire. Compare to the formula in Example 3.2.

(c) What is the value of R^2 for your regression model in part (b)? How do you interpret this number?

(d) Plot drive time d versus distance r, and then sketch a graph of the formula for d as a function of r, which you determined in part (b). Does the power law model seem to give a good fit to this data?

(e) Compare the results of parts (c) and (d) with the results of Exercise 17. Which model seems to give a better fit to this data? Justify your answer.

19. (Model selection for CM1 data) Reconsider the ARM problem of Example 8.6, but now consider a time series model with two previous months of data as predictors.

(a) Use a computer package for multiple linear regression to fit the model

$$X_t = a + bt + c_1 X_{t-1} + c_2 X_{t-2} + \varepsilon_t$$

to the CM1 data.

(b) Interpret the R^2 value for this model, and compare with the results of Example 8.6.

(c) Make a plot for the residuals for this model, similar to Figure 8.18 in the text. Do the residuals appear to form a stationary uncorrelated noise sequence?

(d) Iterate the model equation from part (a) and estimate the CM1 index on May 1990. Use the reported standard deviation to give a 95% prediction interval. Compare with the results of Example 8.6. Is this new model significantly better?

20. (Model selection for TB3 data) Reconsider the ARM problem of Examples 8.5 and 8.6, but now consider the TB3 data from Table 8.1.

(a) Use simple linear regression to fit the model $X_t = a + bt + \varepsilon_t$ as in Example 8.5. Predict the TB3 index for September 1989 and give a 95% prediction interval.

(b) Compute the residuals for the model in part (a) and check for stationarity and serial dependence. Does the model seem adequate?

(c) Repeat parts (a) and (b) using the autoregressive model $X_t = a + bt + c_1 X_{t-1} + \varepsilon_t$ as in Example 8.6 and compare with the results of (a) and (b).

(d) Which of the two models would you recommend? Justify your answer.

Further Reading

1. Arrow, K. et al. (1958) *Studies in the Mathematical Theory of Inventory and Production.* Stanford University Press, Stanford, California.

2. Billingsley, P. (1979) *Probability and Measure.* Wiley, New York.

3. Box, G. and Jenkins, G. (1976) *Time Series Analysis, Forecasting, and Control.* Holden-Day, San Francisco.

4. Brockwell, P. and Davis, R. (1991) *Time Series: Theory and Methods.* 2nd Ed., Springer–Verlag, New York.

5. Çinlar, E. (1975) *Introduction to Stochastic Processes.* Prentice–Hall, Englewood Cliffs, New Jersey.

6. Cornell, R., Flora, J. and Roi, L. *The Statistical Evaluation of Burn Care.* UMAP module 553.

7. Freedman, D. et al. (1978) *Statistics.* W. W. Norton, New York.

8. Giordano, F., Wells, M. and Wilde, C. *Dimensional Analysis.* UMAP module 526.

9. Giordano, M., Jaye, M. and Weir, M. *The Use of Dimensional Analysis in Mathematical Modeling.* UMAP module 632.

10. Hillier, F. and Lieberman, G. (1990) *Introduction to Operations Research.* 5th Ed., Holden-Day, Oakland CA.

11. Hogg, R. and Tanis, E. (1988) *Probability and Statistical Inference.* Macmillan, New York.

12. Huff, D. (1954) *How to Lie with Statistics.* W. W. Norton, New York.

13. Kayne, H. *Testing a Hypothesis: t–Test for Independent Samples.* UMAP module 268.

14. Keller, M. *Markov Chains and Applications of Matrix Methods: Fixed–Point and Absorbing Markov Chains.* UMAP modules 107 and 111.

15. Knapp, T. *Regression Toward the Mean.* UMAP module 406.

16. Meerschaert, M. and Cherry, W. P. (1988) Modeling the behavior of a scanning radio communications sensor. *Naval Research Logistics Quarterly*, Vol. 35, 307–315.

17. Travers, K., and Heeler, P. *An Iterative Approach to Linear Regression.* UMAP module 429.

18. Yates, F. *Evaluating and Analyzing Probabilistic Forecasts.* UMAP module 572.

Chapter 9

SIMULATION OF PROBABILITY MODELS

Computational methods for optimization are important, because most optimization problems are too difficult to be solved analytically. For dynamic models, it is often possible to determine steady–state behavior analytically, but the study of transient (time–dependent) behavior requires computer simulation. Probability models are even more complex. Models with no time dynamics can sometimes be solved analytically, and steady–state results are available for the simplest stochastic models. But for the most part, probability models are solved by simulation. In this chapter we will discuss some of the most generally useful simulation methods for probability models.

9.1   Monte Carlo Simulation

Questions involving the transient or time–dependent behavior of stochastic models are difficult to resolve analytically. Monte Carlo simulation is a general modeling technique that is usually effective for such problems. The construction of Monte Carlo simulation software can be time–consuming, and the repeated simulation runs needed for accuracy and sensitivity analysis can become prohibitively expensive. Even so, Monte Carlo simulation models continue to enjoy a very wide appeal. They are easy to conceptualize, easy to explain, and they are the only viable method for the modeling of many complex stochastic systems. A Monte Carlo simulation models random behavior. It can be based on any simple randomizing device, such as a coin flip or the roll of dice, but typically a computer pseudorandom number generator is used. Because of the random element, each repetition of the model will produce different results.

Example 9.1. Arriving on your vacation, you are dismayed to learn that the local weather service forecasts a 50% chance of rain every day this week. What are the chances of 3 consecutive rainy days?

Variables: $X_t = \begin{cases} 0 & \text{if no rain on day } t \\ 1 & \text{if rain on day } t \end{cases}$

Assumptions: $X_1, X_2, \ldots, X_7$ are independent
$\Pr\{X_t = 0\} = \Pr\{X_t = 1\} = 1/2$

Objective: Determine the probability that
$X_t = X_{t+1} = X_{t+2} = 1$ for some
$t = 1, 2, 3, 4,$ or 5

Figure 9.1: Results of step 1 of the rainy day problem.

We will use the five-step method. Step 1 is to formulate a question. In the process we assign variable names to quantities of interest, and we clarify our assumptions about these variables. Then we state a question in terms of these variables. See Figure 9.1 for the results of step 1. Step 2 is to select the modeling approach. We will use Monte Carlo simulation.

> *Monte Carlo simulation* is a technique that can be applied to any probability model. A probability model includes a number of random variables and must also specify the probability distribution for each of these random variables. Monte Carlo simulation uses a randomizing device to assign a value to each random variable, in accordance with its probability distribution. Since the results of the simulation depend on random factors, subsequent repetitions of the same simulation will reproduce different results. Usually a Monte Carlo simulation will be repeated a number of times in order to determine an average or expected outcome.
>
> Monte Carlo simulation is typically used to estimate one or more measures of performance (MOPs) of the system. Repeated simulations can be considered as independent random trials. For the moment let us consider the situation where there is only one simulation parameter Y to be examined. Repeated simulation produces the results $Y_1, Y_2, \ldots, Y_n$, which we may consider as independent and identically distributed random variables whose distribution is unknown. By the strong law of large numbers we know that

$$\frac{Y_1 + \cdots + Y_n}{n} \to EY \qquad (9.1)$$

> as $n \to \infty$. Hence, we should use the average of $Y_1, \ldots, Y_n$ to estimate the true expected value of Y. We also know that, letting

$$S_n = Y_1 + \cdots + Y_n,$$

> the central limit theorem implies that

$$\frac{S_n - n\mu}{\sigma\sqrt{n}}$$

is approximately standard normal for large n, where $\mu = EY$ and $\sigma^2 = VY$. For most cases the normal approximation is fairly good when $n \geq 10$. Even though we do not know μ or σ, the central limit theorem still gives some important insight. The difference between the observed average S_n/n and the true mean $\mu = EY$ is

$$\frac{S_n}{n} - \mu = \frac{\sigma}{\sqrt{n}} \left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \right), \qquad (9.2)$$

so we can expect the variation in the observed mean to tend to zero about as fast as $1/\sqrt{n}$. In other words, to get one more decimal place of accuracy in EY would require 100 times as many repetitions of the simulation. More sophisticated statistical analysis is possible, but the basic idea is now very clear. We will have to be satisfied with fairly rough approximations of average behavior if we are to use Monte Carlo simulation.

As a practical matter, there are many sources of error and variation in a modeling problem, and the additional variation produced by Monte Carlo simulation is not typically the most serious of these. A judicious application of sensitivity analysis will suffice to ensure that the results of a simulation are used properly.

Moving on to step 3, we now need to formulate the model. Figure 9.2 gives an algorithm for Monte Carlo simulation of our vacation problem.

As in Chapter 3, the notation Random $\{S\}$ denotes a point selected at random from the set S. In our simulation, each day's weather is represented by a random number from the interval $[0, 1]$. If the number turns out to be less than p, we assume that this is a rainy day. Otherwise, it is a sunny day. Then p is the probability that any one day is rainy. The variable C simply counts the number of consecutive rainy days. Figure 9.3 shows a slightly modified algorithm. The modified version repeats the Monte Carlo simulation n times and counts the number of rainy weeks (i.e., the number of weeks in which it rains 3 days in a row). The notation

$$Y \leftarrow \text{Rainy Day Simulation } (p)$$

indicates that we evaluate the output variable Y by running the rainy day simulation from Fig. 9.2 with input variable p.

Step 4 is to solve the problem. We ran a computer implementation of the algorithm in Fig. 9.3 with $p = 0.5$ and $n = 100$. The simulation counted 43 rainy weeks out of 100. On this basis we would estimate a 43% chance of a rainy week. Several additional runs were made to confirm these results. In every case the simulation counted around 40 rainy weeks out of 100. Given the likely magnitude of error in the 50% estimated chance of rain, this is about as much accuracy as we will need for this problem. More details on the sensitivity of our simulation results to random factors will be discussed later, in the section on sensitivity analysis.

Algorithm: RAINY DAY SIMULATION

Variables: p = probability of one rainy day

$$X(t) = \begin{cases} 1 & \text{if rain on day } t \\ 0 & \text{otherwise} \end{cases}$$

$$Y \quad = \begin{cases} 1 & \text{if } \geq 3 \text{ consecutive rainy days} \\ 0 & \text{otherwise} \end{cases}$$

Input: p

Process: Begin
    $Y \leftarrow 0$
    $C \leftarrow 0$
    for $t = 1$ to 7 do
        Begin
        if Random $\{[0,\ 1]\} < p$ then
            $X(t) = 1$
        else
            $X(t) = 0$
        if $X(t) = 1$ then
            $C \leftarrow C + 1$
        else
            $C \leftarrow 0$
        if $C \geq 3$ then $Y \leftarrow 1$
        End
    End

Output: Y

Figure 9.2: Pseudocode for Monte Carlo simulation of the rainy day problem.

Finally, step 5. Arriving on your vacation, you find that the local weather service predicts a 50% chance of rain every day for a week. A simulation indicates that, if this forecast is correct, there is a 40% chance that there will be at least 3 consecutive rainy days this week. These results apply to sunshine as well as rain, and so, to end on a somewhat more optimistic note, let us point out that there is a 50% chance of sunshine every day this week, and a 40% chance of at least 3 consecutive days of sunshine. Enjoy your vacation!

We will begin our sensitivity analysis by examining the sensitivity of our simulation results to random factors. Each model run simulates $n = 100$ weeks of vacation and counts the number of rainy weeks. In the terminology of step 2, our MOP is Y, where $Y = 1$ indicates a rainy week, and $Y = 0$ indicates otherwise. Our model simulates $n = 100$ independent random variables $Y_1, \ldots, Y_n$,

Algorithm: REPEATED RAINY DAY SIMULATION

Variables: $p =$ probability of 1 rainy day
$n =$ number of weeks to simulate
$S =$ number of rainy weeks

Input: p, n

Process: Begin
$S \leftarrow 0$
for $k = 1$ to n do
        Begin
        $Y \leftarrow$ Rainy Day Simulation (p)
        $S \leftarrow S + Y$
        End

Output: S

Figure 9.3: Pseudocode for repeated Monte Carlo simulation to determine average behavior in the rainy day problem.

all of which have the same distribution as Y. Here $Y_k = 1$ indicates that week k was a rainy week. Our model outputs the random variable $S_n = Y_1 + \cdots + Y_n$, which represents the number of rainy weeks. Let

$$q = \Pr\{Y = 1\} \tag{9.3}$$

denote the probability of a rainy week. It is not hard to calculate that

$$\begin{aligned} \mu &= EY = q \\ \sigma^2 &= VY = q(1-q). \end{aligned} \tag{9.4}$$

Our first model run output $S_n = 43$. On this basis, we can use the strong law of large numbers, Eq. (9.1), to estimate that

$$q = EY \approx S_n/n = 0.43. \tag{9.5}$$

How good is this estimate? By the central limit theorem, we obtain from Eq. (9.2) that S_n/n is unlikely to differ from $\mu = q$ by more than $2\sigma/\sqrt{n}$, since a standard normal random variable is 95% certain to have absolute value less than 2. Using Eqs. (9.4) and (9.5), we would conclude that our estimate in Eq. (9.5) is within

$$2\sqrt{(0.43)(0.57)/100} \approx 0.1 \tag{9.6}$$

of the true value of q.

A more elementary way to investigate the sensitivity of our simulation results to random factors is to compare the results of a number of model runs. Figure
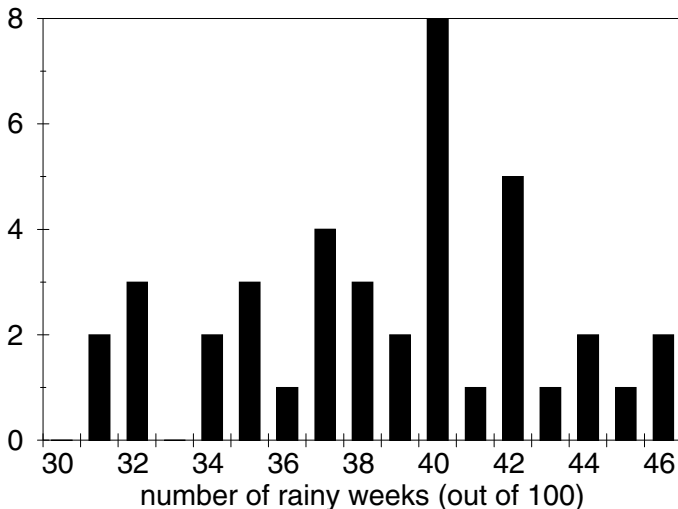
Figure 9.4: Histogram showing the distribution of the number of rainy weeks out of 100 for the rainy day problem.

9.4 shows the results of 40 model runs, each of which simulates 100 weeks of vacation. All of these model runs lead to estimates $S_n/n \approx 0.4$, and none is outside of the interval 0.4 ± 0.1. It also appears that the distribution of S_n is approximately normal.

We should also examine the sensitivity of our simulation results to the forecast 50% chance of rain. Figure 9.5 shows the results of 10 additional model runs for each of the cases $p = 0.3$, 0.4, 0.5, 0.6, and 0.7, where p is the probability of one rainy day. The boxes connected by the dotted line show the average outcome (fraction of rainy weeks), and the vertical bars show the range of outcomes for each case. For a 40% chance of rain each day, the probability of at least 3 consecutive rainy days in a week is around 20%, and so forth. While the probability of 3 straight days of rain varies quite a bit, it seems safe to say that if the chance of rain each day is moderate, then so is the chance of rain 3 straight days in one week.

What about robustness? We should examine the critical assumptions that make up the structure of the model. In step 1 we assumed that the indicator variables $X_1, \ldots, X_7$ were independent and identically distributed. In other words, the chance of rain is the same each day, and the weather on one day is independent of the weather on any other day. Suppose instead that $\Pr\{X_t = 1\}$ varies with t, still keeping the independence assumption. We can use our sensitivity analysis results to obtain upper and lower bounds on the probability of 3 consecutive days of rain, reasoning that this probability should be monotone
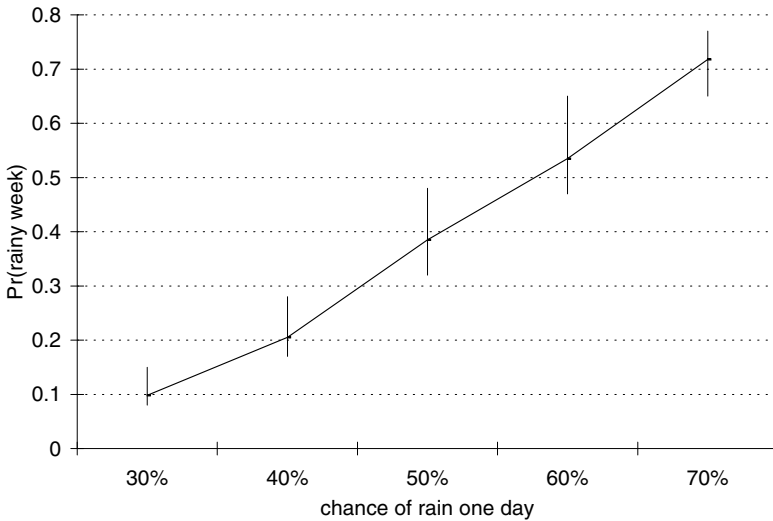
Figure 9.5: Graph of the probability of a rainy week versus the chance of rain one day for the rainy day problem.

increasing as any $\Pr\{X_t = 1\}$ increases. A higher chance of rain on day t means a higher probability of 3 straight days of rain, all else being equal. Then, if all $\Pr\{X_t = 1\}$ are between 0.4 and 0.6 (a 40% to 60% chance of rain), the probability of 3 days of rain is between 0.2 and 0.6. Our model is quite robust in this regard, partly due to the fact that we are not looking for a lot of accuracy in our answer.

Suppose now that $\{X_t\}$ are not independent. For example, we could model $\{X_t\}$ as a Markov chain on the state space $\{0, 1\}$. This implies that the chance of rain today depends on the weather yesterday. Local weather forecasts rarely contain the kind of information necessary to formulate such a model, particularly the state transition probabilities. However, we could always guess at these and then use sensitivity analysis to ensure that we have not assumed too much. The kind of question we are asking here cannot be answered using a steady–state analysis, because it concerns the time–dependent or transient behavior of the stochastic process. Such questions are usually difficult to handle analytically, even by advanced techniques. This is one reason why Monte Carlo simulation is so widely used. On the other hand, it is sometimes possible to develop analytical solutions; see Exercise 9.19.

9.2   The Markov Property

A stochastic process is said to have the *Markov property* if the information contained in the current state of the process is all that is needed to determine the probability distribution of future states. Markov chains and Markov processes were introduced in Chapter 8. Both have the Markov property. Monte Carlo simulation of a stochastic process is much simpler in the presence of the Markov property, because it reduces the quantity of information that needs to be stored in the computer.

Example 9.2. We reconsider the docking problem of Example 4.3, now taking into account the random element. Our basic assumptions are summarized in Fig. 4.7. Our goal or objective, as before, is to determine the success of our control procedure to match velocities.

We will use the five-step method. Our starting point is Fig. 4.7, but we must also make some assumptions about the variables a_n, c_n, and w_n. Ideally we would perform some experiments and collect data on such factors as astronaut response time and the time it takes to read or manipulate controls. In the absence of such data, we would attempt to make some reasonable assumptions consistent with what is known about similar situations.

The random variable that represents the most uncertainty (i.e., the greatest variance) is c_n, the time it takes to make a control adjustment. This variable represents the time it takes to observe the rate of closing, calculate the desired acceleration adjustment, and carry out the adjustment. We will assume that it takes roughly 1 second to observe the rate of closing, 2 seconds to calculate the adjustment, and 2 seconds to make the adjustment. The actual time to carry out each phase is random. Let R_n denote the time to read the closing velocity, S_n the time to calculate the desired adjustment, and T_n the time to make the adjustment. Then $ER_n = 1$ second and $ES_n = ET_n = 2$ seconds. We need to make a reasonable assumption about the distributions of these random variables. It seems reasonable to suppose that they are all nonnegative, mutually independent, and that outcomes close to the mean are most likely. There is a wide variety (infinite!) of distributions that fit this general description, and as of now, we have no particular reason to prefer one over another. This being the case, we will refrain from specifying the exact distribution at this time. One of our random variables is $c_n = R_n + S_n + T_n$. The others are w_n, the waiting time before the next control adjustment, and a_n, the acceleration after this control adjustment. We will assume that $a_n = -kv_n + \varepsilon_n$, where ε_n is a (small) random error. We assume that ε_n is normally distributed with mean zero. The variance of ε_n depends on both the skill of the human operator and on the sensitivity of the control mechanism. We will assume that we can typically achieve an accuracy of about $\pm$ 0.05 m/sec^2, and so we set the standard deviation of ε_n at $\sigma = 0.05$. The waiting time w_n will depend on c_n if we are trying to maintain a fixed time between control adjustments of 15 seconds total. We will assume that $w_n = 15 - c_n + E_n$, where E_n is a small random error. Assume E_n

Variables: t_n = time of nth velocity observation (sec)
v_n = velocity at time t_n (m/sec)
c_n = time to make nth control adjustment (sec)
a_n = acceleration after nth control adjustment (m/sec^2)
w_n = wait before $(n+1)$st observation (sec)
R_n = time to read velocity (sec)
S_n = time to calculate adjustment (sec)
T_n = time to make adjustment (sec)
ε_n = random error in control adjustment (m/sec^2)
E_n = random error in waiting time (sec)

Assumptions: $t_{n+1} = t_n + c_n + w_n$
$v_{n+1} = v_n + a_{n-1}c_n + a_n w_n$
$a_n = -kv_n + \varepsilon_n$
$c_n = R_n + S_n + T_n$
$w_n = 15 - c_n + E_n$
$v_0 = 50, \ t_0 = 0$
$ER_n = 1, \ ES_n = ET_n = 2$, and the
distribution of R_n, S_n, T_n is yet
to be specified
ε_n is normal mean 0, standard
deviation 0.05
E_n is normal mean 0, standard
deviation 0.1

Objective: Determine $T = \min\{t_n : |v_n| \le 0.1\}$

Figure 9.6: Results of step 1 of the docking problem with random factors.

has mean zero, is normal, and that astronaut response–time limitations imply a standard deviation of 0.1 seconds.

We should now consider the analysis objectives. We want to determine the success of our control procedure. Certainly, we are interested in seeing that $v_n \to 0$. Simulation can determine that. But we also have the opportunity to gather information on other aspects of system performance. We should decide at this time, as a part of step 1, what measures of performance (MOPs) we want to track. The selected MOPs should represent important quantitative information that can be used as the basis for comparison with computing control procedure options. Assume that our initial closing velocity is 50 m/sec, and that the velocity–matching process is considered successful when the closing velocity has been reduced to 0.1 m/sec. We would be most interested in the total time it takes to succeed. This will be our measure of performance. At this point, to conclude step 1, we summarize our results in Figure 9.6.

Step 2 is to select the modeling approach. We will use a Monte Carlo simulation based on the Markov property.

The general idea is as follows. At each time step n there is

Begin
Read data
Initialize X_0
While (not done) do
  Begin
  Determine distribution of X_{n+1} using X_n
  Use Monte Carlo method to determine X_{n+1}
  Update records for MOPs
  End
Calculate and output MOPs
End

Figure 9.7: Algorithm for the general Markovian simulation.

a vector X_n that describes the current state of the system. The sequence of random vectors $\{X_n\}$ is assumed to have the Markov property. In other words, the current state X_n contains all of the information needed to determine the probability distribution of the next state, X_{n+1}.

The general structure of the simulation is as follows. First, we initialize variables and read data files. At this stage we must specify the initial state X_0. Next, we enter a loop that repeats until an end condition is satisfied. In the loop we use X_n to specify the distribution of X_{n+1}, and then we use a random number generator to determine X_{n+1} according to that distribution. We must also calculate and store any information needed to generate the simulation MOPs. Once the end condition occurs, we exit the loop and output the MOPs. Then we are done. The algorithm for this simulation is illustrated by Figure 9.7.

We need to discuss the inner loop in some detail. Suppose for now that the state vector X_n is one–dimensional. Let

$$F_\Theta(t) = \Pr\{X_{n+1} \leq t | X_n = \Theta\}.$$

The value of $\Theta = X_n$ determines the probability distribution of X_{n+1}. The function F_Θ maps the state space

$$E \subseteq \mathbb{R}$$

onto the interval [0, 1]. There are widely available methods for generating a random number in [0, 1], and these can be used to generate a random variable with distribution F_Θ. Since

$$y = F_\Theta(x)$$

maps

$$E \to [0, 1],$$

the inverse function
$$x = F_\Theta^{-1}(y)$$
maps
$$[0, 1] \to E.$$

If U is a random variable uniformly distributed over $[0, 1]$ (i.e., the density function of U is the function equal to 1 on $[0, 1]$ and equal to 0 elsewhere), then $X_{n+1} = F_\Theta^{-1}(U)$ has distribution F_Θ, since given $X_n = \Theta$,

$$
\begin{aligned}
\Pr\{X_{n+1} \le t\} &= \Pr\{F_\Theta^{-1}(U) \le t\} \\
&= \Pr\{U \le F_\Theta(t)\} \qquad (9.7) \\
&= F_\Theta(t)
\end{aligned}
$$

because
$$\Pr\{U \le x\} = x \text{ for } 0 \le x \le 1.$$

Example 9.3. Let $\{X_n\}$ denote a stochastic process where X_{n+1} has an exponential distribution with rate parameter X_n. Determine the first passage time

$$T = \min\{n : X_1 + \cdots + X_n \ge 100\},$$

assuming that $X_0 = 1$.

We will present a computer simulation to solve this problem once we discuss the details of generating X_{n+1} from X_n. Letting $\Theta = X_n$, the density function of X_{n+1} is

$$f_\Theta(x) = \Theta e^{-\Theta x}$$

on $x \ge 0$. The distribution function is

$$F_\Theta(x) = 1 - e^{-\Theta x}.$$

Setting
$$y = F_\Theta(x) = 1 - e^{-\Theta x}$$
and inverting, we obtain

$$x = F_\Theta^{-1}(y) = -\ln(1 - y)/\Theta.$$

Hence we can let
$$X_{n+1} = -\ln(1 - U)/\Theta,$$

where U is a random number between 0 and 1. See Figure 9.8 for the complete simulation algorithm.

The discussion above provides a method of generating random variables with any prescribed distribution. While useful in theory,

Algorithm: FIRST PASSAGE TIME SIMULATION (Example 9.3)

Variables: X = initial state variable
                 N = first passage time

Input: X

Process: Begin
                 $S \leftarrow 0$
                 $N \leftarrow 0$
                 until $(S \geq 100)$ do
                       Begin
                       $U \leftarrow$ Random $\{[0, 1]\}$
                       $R \leftarrow X$
                       $X \leftarrow -\ln(1 - U)/R$
                       $S \leftarrow S + X$
                       $N \leftarrow N + 1$
                       End
                 End

Output: N

Figure 9.8: Pseudocode for the Markovian simulation for Example 9.3.

in practice there is sometimes a catch. For many distributions, such as the normal distribution, it is not easy to compute the inverse function F_{Θ}^{-1}. We can always circumvent this difficulty by interpolating from a table of functional values, but in the case of a normal distribution, there is an easier way.

The central limit theorem guarantees that for any sequence of independent, identically distributed random variables $\{X_n\}$ with mean μ and variance σ^2, the normalized partial sums

$$\frac{(X_1 + \cdots + X_n) - n\mu}{\sigma\sqrt{n}}$$

tend to a standard normal distribution. Suppose $\{X_n\}$ are uniform on $[0, 1]$. Then

$$\mu = \int_0^1 x \cdot dx = 1/2$$

$$\sigma^2 = \int_0^1 (x - 1/2)^2 \, dx = 1/12, \tag{9.8}$$

and so for n sufficiently large, the random variable

$$Z = \frac{(x_1 + \cdots + X_n) - n/2}{\sqrt{n/12}} \tag{9.9}$$

Algorithm: NORMAL RANDOM VARIABLE

Variables: μ = mean
σ = standard deviation
Y = normal random variable with mean μ, standard deviation σ

Input: μ, σ

Process: Begin
$S \leftarrow 0$
for $n = 1$ to 12 do
    Begin
    $S \leftarrow S + \text{Random } \{[0, 1]\}$
    End
$Z \leftarrow S - 6$
$Y \leftarrow \mu + \sigma Z$
End

Output: Y

Figure 9.9: Pseudocode for Monte Carlo simulation of a normal random variable.

is approximately standard normal. For most purposes a value of $n \geq 10$ is sufficient. We will use $n = 12$ to eliminate the denominator in Eq. (9.9). Given a standard normal random variable Z, another normal random variable Y with mean μ and standard deviation σ can be obtained by setting

$$Y = \mu + \sigma Z. \tag{9.10}$$

Figure 9.9 shows a simple algorithm for generating normal random variables with a specified mean and variance.

Returning to the docking problem, we begin step 3. Our first concern is to identify our state variables. In this case we can take

$$\begin{aligned} T &= t_n \\ V &= v_n \\ A &= a_n \\ B &= a_{n-1} \end{aligned} \tag{9.11}$$

as our state variables. Since our only MOP is already a state variable, we will not need to initialize or update any additional variables for that purpose. Figure 9.10 gives the algorithm for our docking simulation. The notation Normal (μ, σ) denotes the output of the normal random variable algorithm described in Fig. 9.9.

Algorithm: Docking Simulation

Variables: k = control parameter
                  n = number of control adjustments
                  $T(n)$ = time (sec)
                  $V(n)$ = current velocity (m/sec)
                  $A(n)$ = current acceleration (m/sec^2)
                  $B(n)$ = previous acceleration (m/sec^2)

Input: $T(0)$, $V(0)$, $A(0)$, $B(0)$, k

Process: Begin
                  $n \leftarrow 0$
                  while $|V(n)| > 0.1$ do
                      Begin
                      $c \leftarrow$ Normal $(5, 1)$
                      $B(n) \leftarrow A(n)$
                      $A(n) \leftarrow$ Normal $(-k\,V(n), 0.05)$
                      $w \leftarrow$ Normal $(15 - c, 0.1)$
                      $T(n) \leftarrow T(n) + c + w$
                      $V(n) \leftarrow V(n) + c\,B(n) + w\,A(n)$
                      $n \leftarrow n + 1$
                      End
                  End

Output: $T(n)$

Figure 9.10: Pseudocode for Monte Carlo simulation of the docking problem.

For now we will assume that

$$c_n = R_n + S_n + T_n$$

has a normal distribution with a mean of $\mu = 5$ seconds and a standard deviation of $\sigma = 1$ second.

Figure 9.11 shows the results of 20 simulation runs. In these runs the docking time ranged between 156 and 604 seconds, with an average docking time of 305 seconds.

We have constructed a Monte Carlo simulation of a velocity–matching exercise for spacecraft manual docking. Our simulation takes into account the random factors inherent in this man–machine system. Based on what we believe to be a reasonable set of assumptions on operator performance, the model indicates a wide variance in the time to complete the docking procedure. For example, the time to match velocities starting from a relative velocity of 50 m/sec, and using a 1 : 50 control factor, averaged about 5 minutes. But outcomes of less than 3 or greater than 7 minutes are not uncommon. The major source
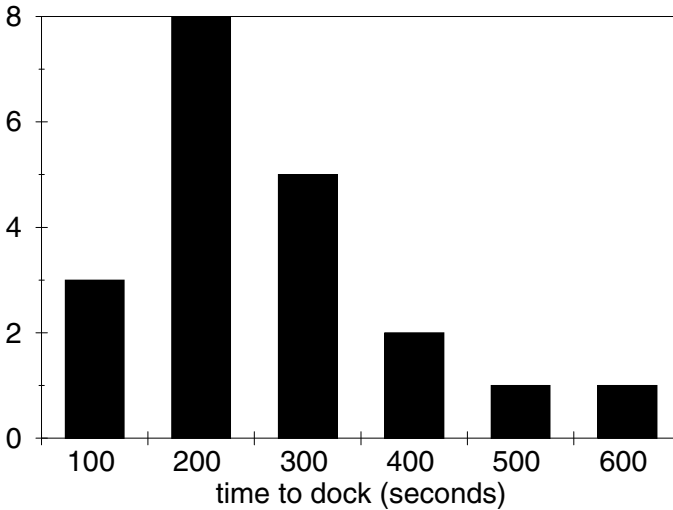
Figure 9.11: Histogram showing the distribution of docking times for the docking problem: case $k = 0.02$, $\sigma = 0.1$.

of variation is the time it takes the pilot to complete the control adjustment procedure.

One important parameter for sensitivity analysis is the standard deviation of c_n, the time to make a control adjustment. We have assumed that the standard deviation of c_n is $\sigma = 1$. Figure 9.12 shows the results of 20 additional model runs for a few values of σ near 1. As in Fig. 9.5, we let the vertical bars represent the range of outcomes, and the boxes connected by the dotted line indicate the average outcome for each value of σ. It seems that our overall conclusions are fairly insensitive to the exact value of σ. In every case the average docking time is around 300 seconds (5 minutes), and the variation in docking times is quite large.

Probably the most important parameter for sensitivity analysis is the control parameter k. We have assumed $k = 0.02$, which results in an average docking time of around 300 minutes. Figure 9.13 shows the results of some additional model runs in which we varied k. We made 20 additional model runs for each new value of k. As before, the vertical bars denote the range of outcomes, and the boxes represent the average outcome for each value of k. As we would expect, docking time is reasonably sensitive to the control parameter k. Of course, it is of considerable interest to determine the best value of k. We will leave this problem for the exercises.

The main robustness question for this model is the distribution of c_n, which we assumed was normal. Since our results did not vary significantly with small
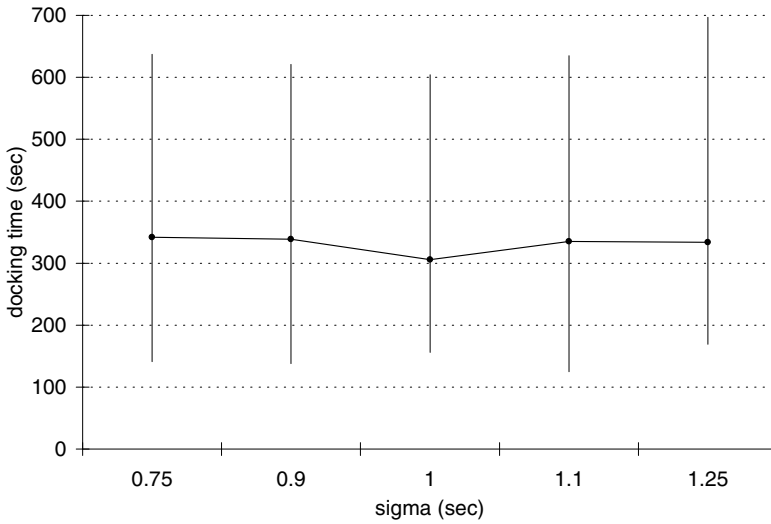
Figure 9.12: Graph of docking time versus standard deviation σ of time to make nth control adjustment for the docking problem.
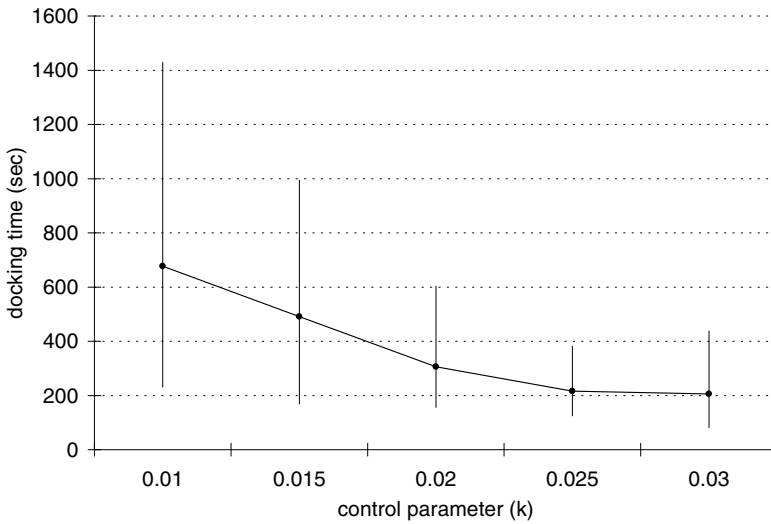
Figure 9.13: Graph of docking time versus control parameter k for the docking problem: case $\sigma = 1$.

changes in σ, and since we are not demanding much accuracy, there is ample reason to expect the model to exhibit robustness with respect to the distribution of c_n. There are a few simple experiments to verify robustness that suggest themselves. We leave these to the exercises.

9.3 Analytic Simulation

Monte Carlo simulation models are relatively easy to formulate, and they are intuitively appealing. Their major drawback is that a very large number of model runs is required to obtain reliable results, especially in the area of sensitivity analysis. Analytic simulations are more difficult to formulate, but are computationally more efficient.

Example 9.4. A military operations analyst plans an air strike against a well–defended target. High–altitude strategic bombers will be sent to attack this important target. It is important to ensure the success of this attack early in the battle, preferably on the first day. Each individual aircraft has a 0.5 probability of destroying the target, assuming that it can get through the air defense and then acquire (i.e., find) the target. The probability that a single aircraft will acquire the target is 0.9. The target is defended by 2 surface–to–air missile (SAM) batteries and a number of air defense guns. The flight profile of the aircraft will prevent the air defense guns from being effective (because the planes will be too high). Each SAM battery has its own tracking radar and computer guidance equipment, which is capable of tracking 2 aircraft and guiding 2 missiles simultaneously. Intelligence estimates a 0.6 probability that 1 missile will disable its target aircraft. Both SAM batteries share a target acquisition radar that is highly effective against high–altitude bombers at a range of up to 50 miles. The effective range of the tracking radar is 15 miles. The bombers will travel at 500 miles/hour at an altitude of 5 miles, and the attack requires that they loiter in the target area for 1 minute. Each SAM battery can launch 1 missile every 30 seconds, and the missiles travel at 1,000 miles/hour. How many bombers should be sent against this target to ensure its destruction?

   We will use the five-step method. Step 1 is to formulate a question. We want to know how many aircraft to commit to this mission. The goal is to destroy the target, but it is immediately clear that we cannot demand a 100% guarantee of success. Let us say now that we want to be 99% sure of destroying the target. We will perform a sensitivity analysis on this number later on. Suppose that N aircraft are committed to this mission. The number of planes destroyed by air defenses before they can complete their attack on the target is a random variable. We will denote this random variable by X. In order to obtain an expression for the probability S that the mission is a success, we will proceed in two stages. First we will obtain an expression for the probability P_i of mission success, given that $X = i$ planes are destroyed before they can attack the target.

Then we will obtain the probability distribution of X so that we can compute

$$S = \sum_i P_i \Pr\{X = i\}. \tag{9.12}$$

If N aircraft are sent and $X = i$ are destroyed before they reach the target, then there are $(N - i)$ attacking aircraft. If p is the probability that 1 attacking aircraft can destroy the target, then $(1 - p)$ is the probability that 1 attacking aircraft will fail to destroy the target. The probability that all $(N - i)$ attacking aircraft fail to destroy the target is

$$(1 - p)^{N-i}.$$

Thus, the probability that at least one of the attacking aircraft succeeds in destroying the target is

$$P_i = 1 - (1 - p)^{N-i}. \tag{9.13}$$

The total exposure time to air defense prior to completing the attack is

$$\frac{15 \text{ mi}}{500 \text{ mi/hr}} \cdot \frac{60 \text{ min}}{\text{hr}} = 1.8 \text{ min}$$

on the way to the target, and an additional minute in the target area, for a total of 2.8 minutes, in which time each SAM battery fires 5 shots. Thus, the attacking aircraft will be exposed to $m = 10$ shots total. We assume the number of aircraft X destroyed will have a *binomial distribution*

$$\Pr\{X = i\} = \binom{m}{i} q^i (1 - q)^{m-i}, \tag{9.14}$$

where

$$\binom{m}{i} = \frac{m!}{i!(m - i)!} \tag{9.15}$$

is the binomial coefficient. (The distribution in Eq. (9.14) is the analytic model for the number of successes in m trials, where q represents the probability of success. See Exercise 12 for more details.) Now, in order to compute the probability of mission success S, we need to substitute Eqs. (9.13) and (9.14) back into Eq. (9.12). Then our objective is to determine the smallest N for which $S > 0.99$. This concludes step 1. We summarize our results in Figure 9.14.

Step 2 is to select our modeling approach. We will use an *analytic simulation* model. In a Monte Carlo simulation we draw random numbers to simulate events and use repeated trials to estimate probabilities and expected values. In an analytic simulation we use a combination of probability theory and computer programming to calculate probabilities and expected values. Analytic simulations are more mathematically sophisticated, which accounts for their greater efficiency. The feasibility of analytic simulation depends on both the problem complexity and the skill of the modeler. Most highly skilled analysts consider

Variables: N = number of bombers sent
m = number of missiles fired
p = probability 1 bomber can destroy target
q = probability 1 missile can disable bomber
X = number of bombers disabled prior to attack
P_i = probability of mission success given $X = i$
S = overall probability of mission success

Assumptions: $p = (0.9)(0.5)$
$q = 0.6$
$m = 10$
$P_i = 1 - (1 - P)^{N-i}$
$\Pr\{X = i\} = \binom{m}{i} q^i (1-q)^{m-i}; i = 0, 1, 2, \ldots, m$
$S = \sum_{i=0}^{m} P_i \Pr\{X = i\}$

Objective: Find the smallest N for which $S > 0.99$

Figure 9.14: Results of step 1 of the bombing run problem.

Monte Carlo simulation as a last resort, to be employed only if they are unable to formulate a suitable analytic model.

Step 3 is to formulate the model. Figure 9.15 shows an algorithm for analytic simulation of the bombing run problem. The notation Binomial (m, i, q) denotes the value of the binomial probability defined by Eqs. (9.14) and (9.15).

Step 4 is to solve the model. We used a computer implementation of the algorithm in Fig. 9.15, with inputs

$$m = 10$$
$$p = (0.9)(0.5)$$
$$q = 0.6,$$

and we varied N to obtain the results shown in Figure 9.16.

A minimum of $N = 15$ planes is required to ensure a 99% chance of mission success. Step 5 is to answer the question, which is how many bombers we must commit to this mission in order to be 99% sure of success. The answer is 15. Now we need to conduct a sensitivity analysis to get at the broader question of what would be a good number of bombers to commit to this mission.

First let us consider the desired probability of success $S = 0.99$. This is a number we literally just made up. Figure 9.17 shows the effect of varying this parameter. It appears that $N = 15$ is a reasonable decision, although any number greater than 10 and less than 20 would be fine. Sending more than 20 aircraft would be overkill.

Bad weather would decrease the detection probability, which factored into the equation $p = (0.9)(0.5)$. If the detection probability decreases to 0.5, then $p = 0.25$, and for $S > 0.99$ we will need at least $N = 23$ aircraft. If the detection

Algorithm: BOMBING RUN PROBLEM

Variables: N = number of bombers sent
                  m = number of missiles fired
                  p = probability 1 bomber can destroy target
                  q = probability 1 missile can disable bomber
                  S = probability of mission success

Input: N, m, p, q

Process: Begin
                  $S \leftarrow 0$
                  for $i = 0$ to m do
                      Begin
                      $P \leftarrow 1 - (1 - p)^{N-i}$
                      $B \leftarrow$ Binomial (m, i, q)
                      $S \leftarrow S + P \cdot B$
                      End
                  End

Output: S

Figure 9.15: Pseudocode for analytic simulation of the bombing run problem.
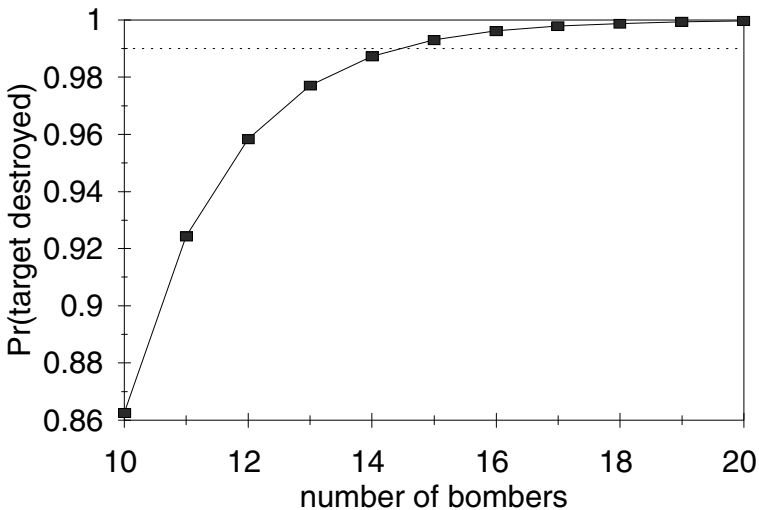
Figure 9.16: Graph of mission success probability S versus number of bombers sent N for the bombing run problem.
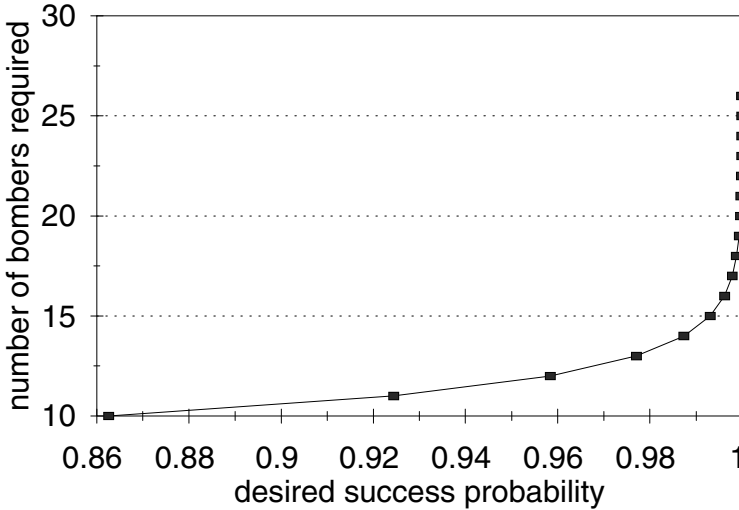
Figure 9.17: Graph of minimum number of bombers N required to obtain mission success probability S in the bombing run problem.

probability is 0.3, then $N = 35$ aircraft are required. The overall relationship between the detection probability and the number of planes required for mission success is illustrated in Figure 9.18. It is unlikely that we will want to fly this mission in bad weather.

One of the applications for models of this kind is to analyze the potential operational impact of engineering advances. Suppose we had a bomber that flew at 1,200 miles/hour and that reduced the loitering time at the target to 15 seconds. Now the aircraft will be exposed to SAM fire for only 1 minute, so that the air defense can only shoot $m = 4$ missiles. Now $N = 11$ bombers are required to get a 99% chance of success. Figure 9.19 shows the relationship between the number of bombers sent and the probability of mission success in the baseline case ($m = 10$ missiles are fired by the air defense) and in the case of the advanced concept aircraft ($m = 4$). For the sake of comparison, we also include the case $m = 0$ missiles fired. This curve represents the maximum potential benefit of the proposed technology. If bombers are exposed to no threat from air defense, it still takes at least 8 planes to ensure a 99% chance of success.

Suppose that we would produce a better targeting system that increased the probability of one aircraft destroying the target to 0.8. Now $p = (0.9)(0.8) = 0.72$ and, all else remaining the same, it would take $N = 13$ aircraft to achieve $S > 0.99$. This is not much of an improvement. If we combine the high–speed bomber with the high–accuracy bombs, we can trim the number of planes needed to 8.
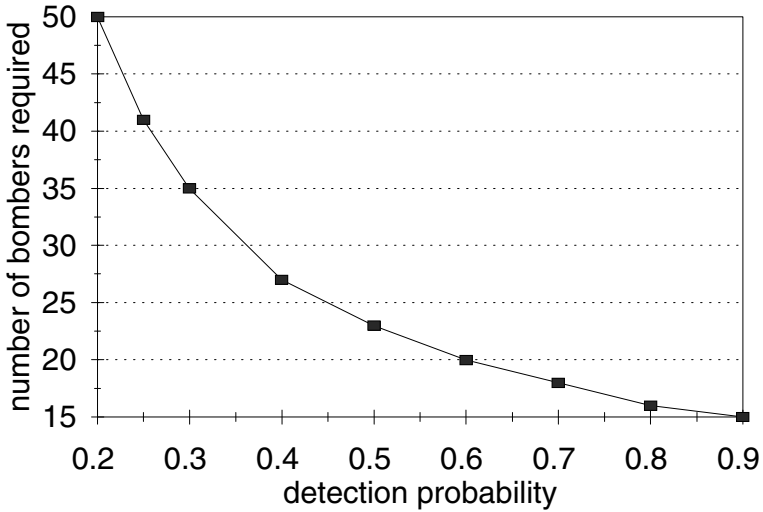
Figure 9.18: Graph of minimum number of bombers N required for mission success versus probability bombers can detect target for the bombing run problem.
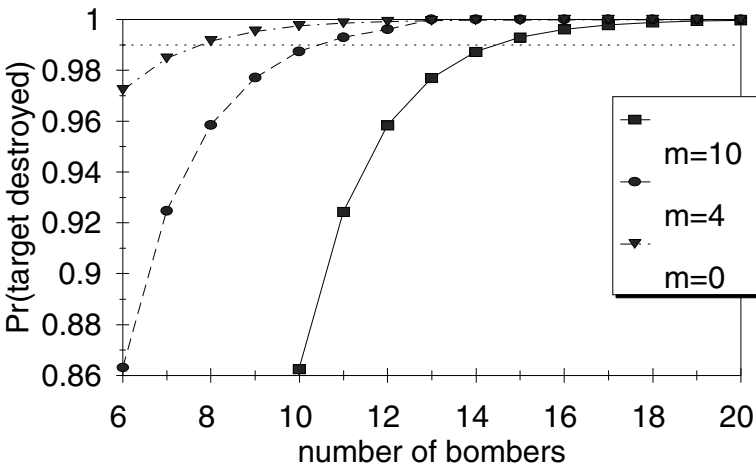
Figure 9.19: Graph of probability of mission success S versus number of bombers sent N for the bombing run problem: comparison of the three cases where the number of missiles fired is $m = 0$, 4, and 10.

What if we have underestimated the effectiveness of enemy air defenses? If $q = 0.8$, then $N = 17$ aircraft are required for $S > 0.99$. If $q = 0.6$ but $m = 15$ (assume that there are 3 SAM batteries), then 18 planes are needed. In either case our general conclusions are not seriously altered.

On the subject of robustness, we have made a number of simplifying assumptions in our model. We have assumed that each of the attacking aircraft acquires the target independently and with the same probability. In reality the first bombs may throw up smoke and dust, obscuring the target area. This may reduce the probability that the remaining bombers can acquire the target. We have already performed a sensitivity analysis on this parameter and found that N is quite sensitive to this probability. Although our model cannot represent a dependence here, it does provide bounds. If each individual aircraft has at least a 50% chance of target acquisition, then $N = 23$ planes will suffice.

Our model also assumed that the SAM batteries never shoot twice at the same aircraft. This is certainly the optimal strategy in a target–rich environment, where the number of potential targets exceeds the maximum number of shots. But suppose that a new stealth technology is able to reduce significantly the number of aircraft detected. Now the air defense may have more shots than targets. Assume that the air defense is able to tell whether they have disabled an aircraft, so that they will not waste any shots. If d aircraft are detected, then the number of aircraft kills in m shots may be represented using a Markov chain model where

$$X_n \in \{0, 1, \ldots, d\}$$

is the number of hits after n shots. The state transition probabilities are

$$\Pr\{X_{n+1} = i + 1 | X_n = i\} = q \text{ and } P_r\{X_{n+1} = i | X_n = i\} = 1 - q$$

for $0 \le i < d$, and of course

$$\Pr\{X_{n+1} = d | X_n = d\} = 1.$$

If $X_0 = 0$, then X_m is the number of aircraft lost. It would be possible to calculate the probability distribution of X_m for each $d = 1, \ldots, N$ and incorporate this into a modified version of our model. This is an example of the application of transient analysis for a Markov chain. It is also much easier to implement a Monte Carlo simulation for this problem. See Exercises 14 and 15 below. Our analytic simulation model is not very robust in this regard.

9.4 Particle Tracking

Particle tracking is a method for solving partial differential equations by simulating the associated stochastic process. This Monte Carlo simulation method is much simpler to code than other numerical solution methods. It is particularly useful for models with variable coefficients, irregular domains, or boundary values. For diffusion problems, it also provides a useful physical model for the motion of individual particles.

Variables: t = Time since release of pollutant (hrs)
                    d = Distance between pollution particle and town (km)
                    s = Plume spread at time t (km)
                    P = Pollution concentration in town (ppm)
                    v = Wind speed at distance d from town (km/hr)

Assumptions: $v = 3$ if $d > 10$
                    $v = 8 - 5d/10$ if $d \leq 10$
                    Peak concentration $P = 1000$ ppm at $t = 1$ hour
                    Plume spread is $s = 2000$ meters at $t = 1$ hour

Objective: Determine the maximum pollution level in town, and
                    time until pollution falls to safe level of 50 ppm.

Figure 9.20: Results of Step 1 of the pollution problem.

Example 9.5. Reconsider the pollution problem of Example 7.5, but now consider that the wind speed increases as we get closer to the city. Suppose that the wind speed increases from 3 kilometers per hour at the spill site to 8 kilometers per hour in the center of town. This is due to a *heat island* effect, where the buildings and pavement in the city retain more heat, leading to warmer conditions and higher wind speeds. Taking this into account, what is the maximum concentration expected in town, when will it occur, and how long until the concentration of pollutant falls back below a safe level? For this Category 1 pollutant (the most dangerous kind), the US Environmental Protection Agency safe level is 50 parts per million (ppm) by volume.

We will use the five-step method. The first step is to ask a question. We want to know the concentration of pollutant in town and how it varies over time. We will assume that the wind speed varies depending on the distance from the center of town. Since we do not know how the actual wind speed varies, we will make the simple assumption that wind speed varies linearly between a value of 3 km/hr at distances 10 or more km from town, up to 8 km/hr at the center of town. Then later we will test this assumption through sensitivity and robustness analysis (see also Exercise 21 at the end of this chapter). The results of Step 1 are summarized in Figure 9.20.

Step 2 is to select the modeling approach. We will use a diffusion model, and solve the model by the method of particle tracking.

The diffusion equation

$$\frac{\partial C}{\partial t} = \frac{D}{2}\frac{\partial^2 C}{\partial x^2}, \qquad (9.16)$$

was introduced in Section 7.4. Now we modify this equation to explicitly represent plume velocity. Recall that $C(x, t)$ represents the relative concentration of contaminants at location x at time t.

The diffusion equation (9.16) models spreading away from the plume center of mass, and for a constant velocity, the center of mass is located at $\mu = vt$ where v is the average velocity. A probability model for the diffusion equation was also explained in Section 7.4. In that model, we track the particle plume in a moving coordinate system, with the origin at the plume center of mass. Each particle makes a small random movement X_i over a small time interval Δt, and these movements are assumed to be independent with mean 0 and variance $D\Delta t$. Then the central limit theorem implies that the particle has deviated by an amount $X_1 + \cdots + X_n \approx \sqrt{Dt}Z$ away from the center of mass by time $t = n\Delta t$ for n sufficiently large, where Z is standard normal. Now consider the actual particle location after n steps, assuming a constant plume velocity v. A randomly selected particle moves by an amount $v\Delta t + X_i$ over the time interval Δt, so after n jumps the particle is located at

$$X_1 + \cdots + X_n + vt \approx vt + \sqrt{Dt}Z.$$

The probability density function of $\sqrt{Dt}Z$ was given in equation (7.32) in Section 7.4. Since the term vt is not random, a simple change of variable shows that $vt + \sqrt{Dt}Z$ has probability density function

$$C(x,t) = \frac{1}{\sqrt{2\pi Dt}}\, e^{-(x-vt)^2/(2Dt)} \tag{9.17}$$

for any $t > 0$. This change of variable shifts the center of the graph from $x = 0$ to $x = vt$.

Next we will use Fourier transforms to see how (9.17) solves a diffusion equation with drift. Using (7.30) along with a change of variable $y = (x - vt)/\sqrt{Dt}$ yields

$$\hat{C}(k,t) = \int_{-\infty}^{\infty} e^{-ikx} C(x,t)\, dx = e^{-ikvt - Dtk^2/2}$$

when $C(x,t)$ is given by (9.17). Then

$$\frac{d\hat{C}}{dt} = -v(ik)\hat{C} + \frac{D}{2}(ik)^2\hat{C}, \tag{9.18}$$

and inverting the Fourier transform leads to

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial x} + \frac{D}{2}\frac{\partial^2 C}{\partial x^2}, \tag{9.19}$$

the diffusion equation with drift.

To solve this equation by particle tracking, simulate a large number N of random particles, using a Monte Carlo simulation of $(X_1 + v\Delta t) + \cdots + (X_n + v\Delta t)$ with $t = n\Delta t$. Then a relative frequency

Algorithm: PARTICLE TRACKING CODE

Variables: N = number of particles
T = time of final particle jump
M = number of particle jumps
v = particle velocity
D = particle dispersivity
$t(j)$ = time of jth jump
$S(i, j)$ = location of ith particle at time $t(j)$

Input: N, T, M, v, D

Process: Begin
$\Delta t \leftarrow T/M$
for $j = 0$ to M do
  Begin
  $t(j) \leftarrow j\Delta t$
  End
for $i = 1$ to N do
  Begin
  $S(i, 0) \leftarrow 0$
  for $j = 0$ to $M - 1$ do
    Begin
    $S(i, j + 1) \leftarrow S(i, j) + \text{Normal}\,(v\Delta t, \sqrt{D\Delta t})$
    End
End

Output: $t(1), \ldots, t(M)$
$S(1, 1), \ldots, S(N, 1)$
$$\vdots$$
$S(1, M), \ldots, S(N, M)$

Figure 9.21: Pseudocode for particle tracking simulation of Section 7.5.
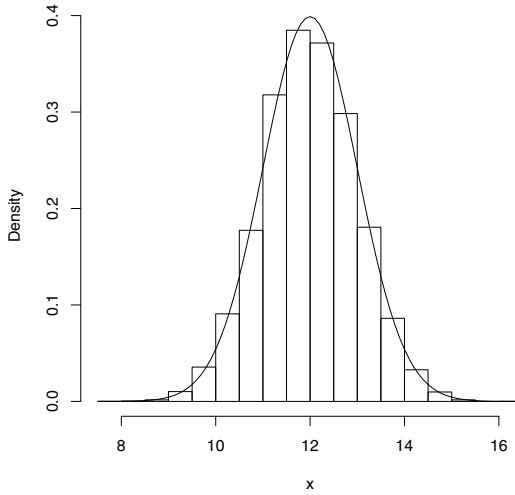
Figure 9.22: Relative frequency histogram showing results of the particle track-ing code in Figure 9.21 with $N = 10,000$, $T = 4$, $M = 50$, $v = 3.0$, and $D = 0.25$. The corresponding analytical solution (solid line) to the diffusion equation with drift (9.19) at time $t = 4$ is also shown for comparison.

histogram of these particle locations will approximate the probabil-ity density curve $C(x,t)$ that gives the theoretical relative concen-tration for an infinite number of particles. Figure 9.21 shows the pseudocode for this algorithm. Figure 9.22 shows the results of a computer implementation, to solve the pollution problem of Exam-ple 7.5 with $v = 3.0$ km/hr and $D = 0.25$ km^2/hr. The simulation used $N = 10,000$ particles, a final time of $T = 4$ hours, and $M = 50$ time steps. Figure 9.22 also shows the analytical solution (9.17) for comparison. It is apparent that the histogram gives a reasonable ap-proximation to the normal density curve. The particle model can be considered as the basic physical model for this problem. The density curve is an approximation, justified by the central limit theorem.

The particle tracking algorithm can easily be extended to allow the coefficients v, D in (9.19) to vary in space. When the velocity v varies in space, the only difference is that the mean jump size $v\Delta t$ for any given particle depends on the current location. In other words, the particle jumps become a Markov chain on the continuous state space $-\infty < x < \infty$ and the jump distribution depends on the current state in a very simple way. As $\Delta t \to 0$, the graph with these Markov chain jumps converges to a *diffusion process*, whose density function $C(x,t)$ solves the diffusion equation with drift (9.19), see Friedman (1975) for more details.

Now we continue with Step 3 of the pollution problem. Using the method of particle tracking, we will model the pollution plume as a cloud of random particles, and we will use Monte Carlo simulation to track the plume movements. We assume a coordinate system where the pollutant release occurred at location 0 and the town is at location 10 kilometers. Then the distance between a particle at location x and the center of town is given by $d = |x - 10|$. All N particles are located at $x = 0$ at time $t = 0$. At each time step Δt, each particle makes a random movement $v\Delta t + X_i$ where

$$v = v(x) = \begin{cases} 3 & \text{if } |x - 10| > 10, \\ 8 - 0.5|x - 10| & \text{if } |x - 10| \le 10. \end{cases} \tag{9.20}$$

Each X_i has mean 0 and variance $D\Delta t$ where $D = 0.25$. For our simulation, we will take each X_i to be normally distributed. We will also assign a mass to each particle. Since the maximum pollution concentration at time $t = 1$ hour is 20×50 ppm $= 1000$ ppm, the pollution concentration $P(x, t)$ is related to the relative concentration $C(x, t)$ by $P = P_0 C$ where $P_0 = 1000\sqrt{0.5\pi}$, using the calculations of Section 7.4. Hence we will assign each particle a concentration $\Delta P = P_0/N$. In our problem setup, $x = 0$ represents the pollution source, and $x = 10$ the center of town. We will estimate concentration in the center of town based on the number of particles at location $9.5 < x \le 10.5$. If K out of N particles lie in this interval, the relative concentration $C \approx K/N$, and the pollution level is $P \approx P_0(K/N) = K\Delta P$. If the time interval $[0, T]$ is divided into M jumps, then the time increment is $\Delta t = T/M$. The pseudocode for this particle tracking simulation is summarized in Figure 9.23. The variable $t(j)$ is the time after j times steps. The state variable $S(i, j)$ represents the location of particle i at time $t(j)$. The simulation only requires us to store the current state, i.e., we can replace $S(i, j)$ by S in the code. The variable $P(j)$ represents the estimated pollution level in town at time $t(j)$.

Step 4 is to implement the particle tracking code developed in Step 3. After implementing the algorithm in Figure 9.23, one run of the resulting code with final time $T = 4.0$ hours, $M = 100$ time steps, and $N = 1,000$ particles produced estimates of $Pmax = 397$, $Tmax = 1.96$, and $Tsafe = 2.32$ for our three MOP. Figure 9.24 shows the simulated concentration levels $P(j)$ at each time $t(j)$. The roughness in the graph is due to the particle tracking simulation. Using more particles (i.e., increasing N) would result in a smoother graph.

Step 5 is to answer the question. We estimate that the maximum pollution level in town will occur about 2 hours after the accident. The maximum pollution level in town is predicted to reach about 400 parts per million (ppm) by volume, which is about 8 times greater than the maximum safe level of this Category 1 contaminant. It is estimated that the pollution plume will continue through town, carried by the wind, and after 2 hours and 40 minutes, the toxic level in town will fall back down below the safety threshold. These estimates are based on a particle tracking simulation that approximates the path of pollution particles headed toward the town. The simulation takes variations in wind speed into account, since wind speeds are generally greater in the city. Figure 9.24

Algorithm: PARTICLE TRACKING CODE (Example 9.5)

Variables: N = number of particles
T = time of final particle jump (hrs)
M = number of particle jumps
$t(j)$ = time of jth jump (hrs)
$P(j)$ = pollution concentration in town after j jumps (ppm)
$P\,max$ = maximum pollution concentration in town (ppm)
$T\,max$ = time of maximum pollution in town (hrs)
$Tsafe$ = time pollution in town falls to safe level (hrs)

Input: N, T, M

Process: Begin
$\Delta t \leftarrow T/M$
$\Delta P \leftarrow 1000\sqrt{0.5\pi}/N$
for $j = 1$ to M do
   $t(j) \leftarrow j\Delta t$
   $P(j) \leftarrow 0$
for $i = 1$ to N do
   $S(i,0) \leftarrow 0$
   for $j = 1$ to M do
     $v \leftarrow 3$
     if $|S(i,j-1) - 10| \leq 10$ then $v \leftarrow 8 - 0.5|S(i,j-1) - 10|$
     $S(i,j) \leftarrow S(i,j-1) + \text{Normal}\,(v\Delta t, 0.5\sqrt{\Delta t})$
     if $9.5 < S(i,j) \leq 10.5$ then $P(j) = P(j) + \Delta P$
$T\,max \leftarrow 0$
$P\,max \leftarrow 0$
$Tsafe \leftarrow 0$
for $j = 1$ to M do
  if $P(j) > P\,max$ then
    $P\,max \leftarrow P(j)$
    $T\,max \leftarrow t(j)$
  if $P(j) > 50$ then $Tsafe = t(j) + \Delta t$
End

Output: $T\,max$, $P\,max$, $Tsafe$

Figure 9.23: Pseudocode for particle tracking simulation of the pollution problem with variable wind speed.
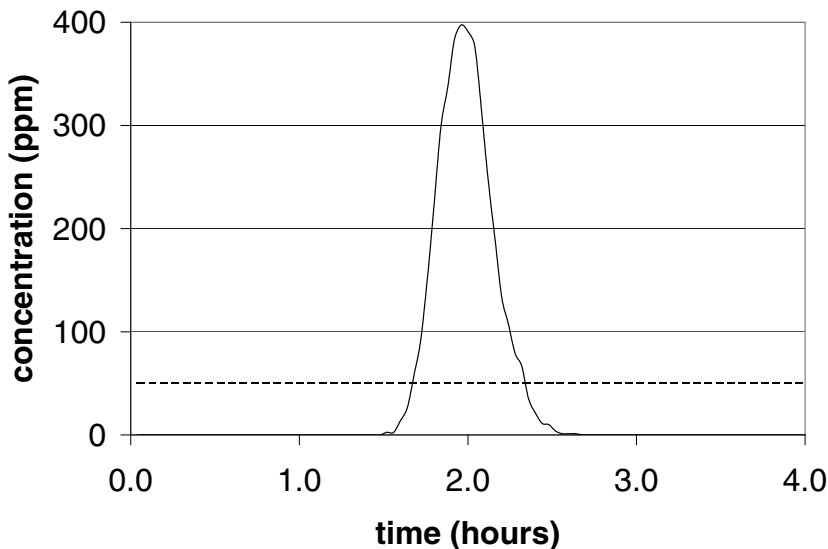
Figure 9.24: Estimated pollution concentration in center of town for the pollution problem with variable wind speed. The dotted line indicates the maximum safe level.

illustrates the predicted contamination levels at the center of town over time, starting at the time of the accident. The dotted line in that figure indicates the maximum safe level of 50 ppm.

For sensitivity analysis, we begin by investigating the sensitivity of our simulation results to random factors. Figure 9.25 illustrates the result of $R = 100$ simulation runs of the algorithm in Figure 9.23. Each simulation assumes a final time of $T = 4.0$ hours, using $M = 100$ time steps, and $N = 1,000$ particles. The average value of $P\,max$ for these simulations is around 400 ppm, with a typical range of ± 20 ppm. More specifically, the $R = 100$ simulation runs produced 100 values of $P\,max$, with sample mean 404.12 and sample standard deviation 16.18. The sample mean and sample standard deviation were computed using built-in functions in our programming platform. For R data points $X_1, \ldots, X_R$, the *sample mean* $\bar{x}$ is just the average,

$$\bar{x} = \frac{1}{R} \sum_{j=1}^{R} X_j.$$

The *sample variance* is given by the formula

$$s^2 = \frac{1}{R-1} \sum_{j=1}^{R} (X_j - \bar{x})^2,$$

which averages the squared deviations from the sample mean. Then the *sample*
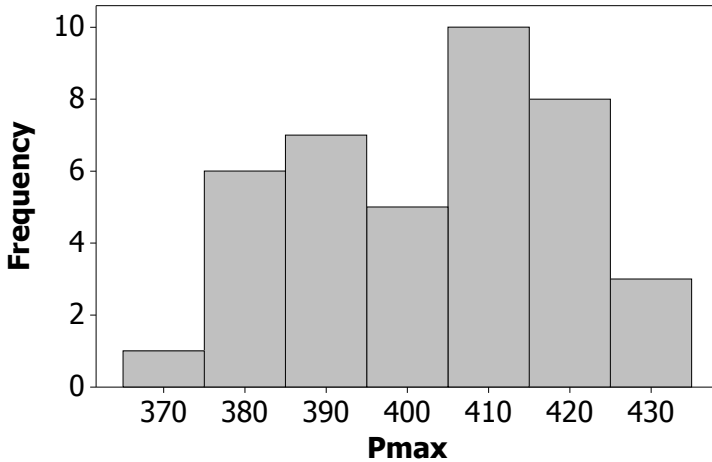
Figure 9.25: Estimated pollution concentration in center of town for the pollution problem with variable wind speed. Results of 100 simulations.

standard deviation s is the square root of the sample variance. Dividing by $R - 1$ instead of R in the formula for the sample variance ensures that the estimates of σ^2 are unbiased, i.e., repeated application of this formula in the same setting gives estimates that average to the correct value. The sample mean can be considered a typical value, and the sample standard deviation measures the typical spread of values. See Ross (1985) for more details.

The most common value of the time $Tmax$ of the maximum concentration was 1.92 hours, and every trial gave a value between 1.88 and 1.96. The most common value of the time $Tsafe$ until a safe level was 2.32 hours, and all values were between 2.28 and and 2.36. Since the time step was $T/M = 0.04$, the difference between the most common values and the largest or smallest value was a single time step. We conclude that the $Tmax$ and $Tsafe$ are relatively unaffected by random factors. The sensitivity of $Pmax$ to random factors is somewhat greater. This standard deviation depends on the number of particles N used in each simulation.

To test this, we conducted $R = 10$ additional simulations with $N = 10,000$ particles, using the same values of the remaining parameters. The resulting values for the maximum concentration $Pmax$ were all between 395 and 408, with a sample mean of 400 and a sample standard deviation of 5.6. A simple probability model explains how the variation in outcomes depends on the sample size. The estimated value of $Pmax$ is $K\Delta P$ where $\Delta P = P_0/N$ is not random, and K is the random number of particles that lie in the interval $9.5 < x \leq 10.5$.

The quantity $Pmax = P_0 q$ where

$$q = \int_{9.5}^{10.5} C(x,t)\,dx$$

is the theoretical probability that a given particle will lie in this interval, at time $t = Tmax$. As in Section 9.1, set $Y_i = 1$ if the ith particle lies in this interval, and $Y_i = 0$ otherwise, so that each Y_i has mean q and variance $\sigma^2 = q(1-q)$. Then $K = Y_1 + \cdots + Y_N$ is the number of particles that lie in this interval. The strong law of large numbers guarantees that

$$K\Delta P = P_0(K/N) = P_0\frac{Y_1 + \cdots + Y_N}{N} \to P_0 q = Pmax$$

as the number of particles $N \to \infty$. The central limit theorem implies that K/N is unlikely to vary from its mean q by more than $2\sigma/\sqrt{N}$ for N large. Hence increasing the number of particles by a factor of 10 should decrease the variability of simulation results by a factor of $\sqrt{10} \approx 3$, which is consistent with our results. To go a bit further, since the total mass of all particles is

$$P_0 = \int_{-\infty}^{\infty} P(x,t)\,dx = 1000\sqrt{0.5\pi} \approx 1253 \text{ ppm},$$

the maximum concentration of 400 ppm represents about 30% of the particles. Then we could predict that around 68% of our simulated estimates for $Pmax$ should lie within $P_0\sqrt{(0.3)(0.7)/N}$ ppm of the true value, which evaluates to around 18 ppm for $N = 1000$ and 5.7 ppm for $N = 10,000$. This is reasonably consistent with the observed standard deviations (16 and 5.6, respectively) for those simulations. A similar analysis shows that variations in the estimated concentration curve Figure 9.24 decrease like the square root of the number of particles in the simulation. Hence a larger number of particles will result in a smoother curve. However, since the variations fall off like the square root of the sample size, it may not always be practical to simulate enough particles to get a completely smooth curve.

Next we consider the robustness of our results to the assumed form of the velocity function $v = v(x)$. A comparison with the results of Section 7.4 shows that, compared to the case of a constant velocity, the particle tracking model predicts a faster arrival of the peak concentration, and a lower level of that concentration. Since the safe level is 50 ppm, the highest concentration in town is $50 \times 11 = 550$ ppm at a constant wind speed of 3.0 km/hr, arriving after 3.3 hours. This can also be checked using a simple modification of the particle tracking code in Figure 9.23, by deleting the line:

$$\text{if } |S(i,j) - 10| \le 10 \text{ then } v \leftarrow 8 - 0.5|S(i,j) - 10|.$$

Then the velocity is fixed at 3.0 km/hr throughout the simulation. One simulation run of the modified code with final time $T = 6.0$ hours, $M = 150$ time steps, and $N = 10,000$ particles produced estimates of $Pmax = 528$ ppm at
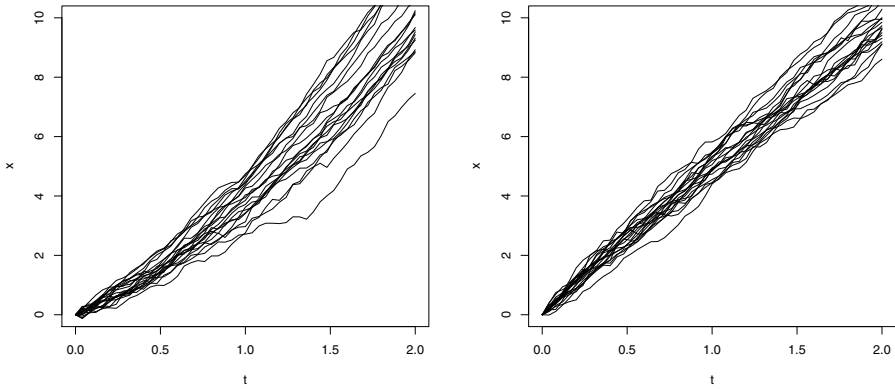
Figure 9.26: Left panel: Particle traces for particle tracking model with variable wind speed (9.20). Right panel: Particle traces for particle tracking model with constant wind speed $v = 5$ km/hr.

time $Tmax = 3.28$ hours, and $Tsafe = 4.04$ hours, in good agreement with the results from Section 7.4. In summary, it seems that the variation in wind speed causes the peak to arrive sooner, with a lower concentration. Is this due to the fact that the "average" wind speed is increasing?

The sensitivity analysis in Table 7.1 shows that the peak concentration in town ranges between $50 \times 11 = 315$ ppm at a wind speed of 1.0 km/hr to $50 \times 14.2 = 710$ ppm at a wind speed of 5.0 km/hr. That is, a higher wind speed leads to higher peak concentration in town, arriving earlier. This makes sense, because at a higher wind speed, the plume arrives sooner, with less time to spread. Indeed, the particle tracking simulation with the variable wind speed $v(x) \geq 3$ km/hour given by (9.20) predicts that the plume peak will arrive sooner. However, the model also predicts a *lower* peak concentration. To understand this paradox, we will take a closer look at the individual particle traces.

The left panel in Figure 9.26 graphs the paths of $N = 20$ particles from the particle tracking model with variable wind speed (9.20). These graphs were produced using the algorithm in Figure 9.23 by plotting $S(i, j)$ versus $t(j)$ for $j = 1, 2, \ldots, M$ for particles $i = 1, 2, \ldots, N$. Compare the right panel in Figure 9.26, which graphs the paths of $N = 20$ particles with constant wind speed $v = 5$, chosen so that both plumes reach the city, located at coordinate $x = 10$, at around the same time. Compared to the constant velocity model, the variable velocity paths are significantly more dispersed. A closer inspection of the left panel shows the reason for this dispersion. Because of random factors, some particle begin more slowly, and therefore experience lower velocities in the subsequent jumps. Other particles begin faster, and remain on a faster path. The term *dispersion* is used to describe the spreading of particles due to
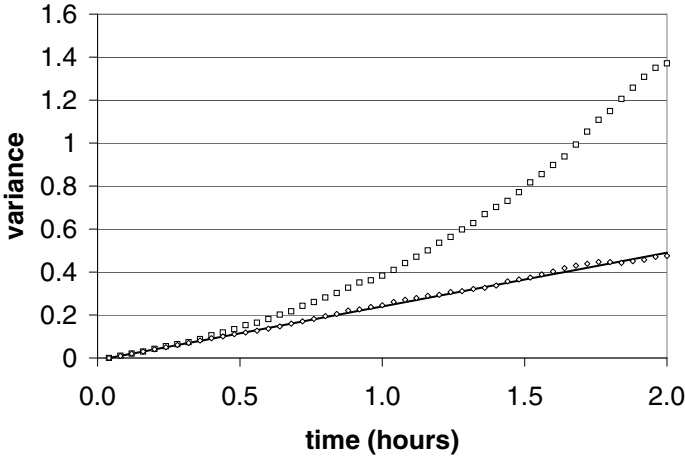
Figure 9.27: Observed variance for particle tracking model (boxes) with variable wind speed (9.20) compared to observed variance for the same model with constant wind speed $v = 5$ km/hr (diamonds), and theoretical variance for the model with constant wind speed (solid line).

differences in velocity. For pollution problems in air and water, dispersion is often the main driving factor behind plume spreading.

In engineering models of contaminant transport, it is common to use the diffusion equation with drift (9.19) as the basic model for plume evolution. It has been noted in many studies that the parameter D, determined by fitting a Gaussian probability density curve to data, tends to grow with time (or with distance traveled by the plume center of mass). This super-diffusion is often attributed to variations in the particle velocities. An interesting fractal model for porous media, used in ground water contamination studies, has been proposed to explain the observation that D tends to grow like a power law, $D = D_0 t^p$ for some $p > 0$, in many studies. See Wheatcraft and Tyler (1988) for more details.

Figure 9.27 documents super-dispersion in the particle tracking model with variable velocity. The graph shows the results of a particle tracking simulation with $N = 1000$ particles with final time $T = 2.0$ hours, and $M = 50$ time steps, so that the time increment $t\Delta t = 0.04$ is the same as before. The boxes shows the variance of the particle locations $\{S(i, j) : 1 \le i \le N\}$ at each time $t(j)$ for $j = 1, \ldots, M$, using the algorithm in Figure 9.23. For a constant velocity, the variance grows linearly, since the particle location at time t is $vt + Z_t$, where Z_t is normally distributed with mean zero and variance $= Dt$. This is the theoretical variance line on the graph. The results of another particle tracking simulation, using the algorithm in Figure 9.23 with constant velocity $v = 5$ km/hr, and all other parameters the same, was run to check this theoretical result. The diamonds on the graph show the resulting variance in particle

locations at each time step, to demonstrate how closely the observed variance matches the theoretical model. Note that a constant velocity simulation with $v = 3$ or $v = 8$ or any other value would produce the same result, since the constant velocity only changes the mean particle location, and has no effect on the variance.

The super-dispersion in contaminant plumes creates a significant problem in practice, since it means that the constant coefficient equation (9.19) is inadequate to predict plume behavior. A variety of methods have been applied to address this problem, including more complicated systems of differential equations in one or more dimensions with variable coefficients, that can be solved by particle tracking or finite difference codes similar to those developed in Chapter 6 of this book. Recently, an interesting new idea has been proposed to model super-diffusive plumes using fractional calculus. The new model replaces the second derivative in (9.19) by a fractional derivative of order $1 < \alpha < 2$. In the next section, we will investigate an application of this fractional diffusion model to ground water pollution.

9.5 Fractional Diffusion

Fractional derivatives were invented by Leibnitz in 1695, soon after their integer order cousins, but have only recently found practical applications. Now they are used in models of anomalous diffusion in water and air pollution, heat transfer in complex materials, invasive species, cell membranes, and electronics. In this section, we will use particle tracking to explore a fractional calculus model for water pollution.

Example 9.6. An experiment to study dispersion in ground water was conducted on a United States Air Force base in Columbus, Mississippi in 1993 (see Boggs et al. (1993) for further details). Water containing a radioactive tracer (tritium) was injected underground, and the movement of the tritium plume through the ground water was tracked over time. Figure 9.28 shows the relative concentration $C(x, t)$ of tritium tracer versus distance x meters downstream from the injection site, based on measurements taken $t = 224$ days after injection. The thin curve near the bottom of the graph is the best fitting Gaussian density, the solution to the diffusion equation with drift (9.19). The movement of the center of mass over the course of the experiment was roughly proportional to elapsed time. The plume spread proportional to $t^{0.9}$, evidence of anomalous super-dispersion. It is also clear that the plume shape is strongly skewed in the positive (downstream) direction, so that the symmetric Gaussian solution to the traditional diffusion equation (9.19) cannot provide an adequate model for this data. Tritium levels were be measured in Curies, a unit of radioactivity. One Curie (Ci) is defined as 3.7×10^{10} decays per second. In this experiment, $P_0 = 540$ Ci of tritium were injected. Estimate the time and amount of the maximum tritium level at a point $x = 20$ meters downstream of the injection site, as well as the time required for the tritium concentration at that location to fall below 2 Ci.

c) Snapshot 3 (day #224)
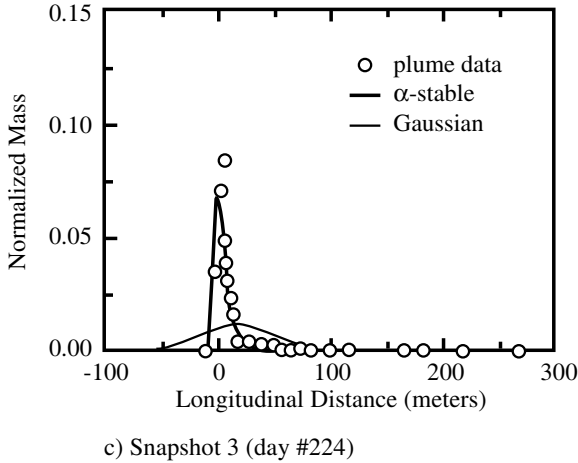
Figure 9.28: Measured tritium concentration (hollow circles) from the fractional diffusion problem of Example 9.6 with fitted stable density (thick solid line) and fitted Gaussian density (thin black line), from Benson et al. (2001).

We will use the five-step method. The first step is to ask a question. We adopt a coordinate system where $x = 0$ is the injection site, and water flows in the positive x direction, consistent with Figure 9.28. We want to predict the concentration of tritium tracer at location $x = 20$ meters downstream from the injection site, and how it varies over time. The results of step 1 are summarized in Figure 9.29.

Step 2 is to select the modeling approach. We will use a fractional diffusion model, and solve the model by the method of particle tracking.

The *fractional derivative* $\partial^\alpha C / \partial x^\alpha$ can be defined as the function with Fourier transform $(ik)^\alpha \hat{C}$. The fractional diffusion equation with drift is

$$\frac{\partial C}{\partial t} = -v\frac{\partial C}{\partial x} + D\frac{\partial^\alpha C}{\partial x^\alpha} \qquad (9.21)$$

where $1 < \alpha < 2$. Take the Fourier transform on both sides to get

$$\frac{d\hat{C}}{dt} = -v(ik)\hat{C} + D(ik)^\alpha \hat{C}$$

and solve, using the point source initial condition $\hat{C}(k, 0) \equiv 1$, to arrive at

$$\hat{C}(k, t) = \int_{-\infty}^{\infty} e^{-ikx} C(x, t)\, dx = e^{-ikvt + Dt(ik)^\alpha} \qquad (9.22)$$

for any $t > 0$. This Fourier transform cannot be inverted in closed form, but it is known from probability theory that $C(x, t)$ is an α-*stable* density function.

Variables: t = Time since release of tracer (days)
                     x = Distance downstream from injection site (m)
                     P = Pollution concentration at location x at time t (Ci)

Assumptions: Average plume velocity is constant
                   Plume spreads proportional to $t^{0.9}$
                   Plume is positively skewed

Objective: Determine the maximum tracer concentration 20 m
                   downstream, time the maximum occurs, and time until
                   concentration falls below 2 Ci.

<div style="text-align:center">

Figure 9.29: Results of Step 1 of the water pollution problem.

</div>

The stable density appears in an *extended central limit theorem.* Suppose that $X, X_1, X_2, \ldots$ are independent and identically distributed random variables. If $\Pr\{X > x\} = Ax^{-\alpha}$ for some $A > 0$ and some $1 < \alpha < 2$, then the extended central limit theorem implies that

$$\lim_{n\to\infty} \Pr\left\{\frac{X_1 + \cdots + X_n - n\mu}{n^{1/\alpha}} \le x\right\} \to \Pr\{Z_\alpha \le x\} \qquad (9.23)$$

where $\mu = E(X)$, and Z_α is a stable random variable with index α, whose density $g_\alpha(x)$ has Fourier transform

$$\hat{g}(k) = \int_{-\infty}^{\infty} e^{-ikx} g(x)\, dx = e^{D(ik)^\alpha} \qquad (9.24)$$

for some $D > 0$ depending on A and α. Because $\sigma^2 = V(X) = \infty$ in this case, the usual central limit theorem, discussed in Section 7.3, does not apply here. See Meerschaert and Sikorskii (2012) for more details about stable laws, the extended central limit theorem, and fractional derivatives.

The spreading rate of the solution $C(x,t)$ to the fractional diffusion equation can be determined from the Fourier transform. A simple change of variables shows that the mean-centered concentration $C_0(x,t) = C(x + vt, t)$ has Fourier transform

$$\int_{-\infty}^{\infty} e^{-ikx} C_0(x,t)\, dx = e^{Dt(ik)^\alpha}, \qquad (9.25)$$

and another change of variables shows that $t^{-1/\alpha} C_0(t^{-1/\alpha}x, 1)$ has the same Fourier transform as $C_0(x,t)$. This shows that the peak concentration falls like $t^{1/\alpha}$, and the plume spreads like $t^{1/\alpha}$ away from its center of mass at $x = vt$. Since $\alpha < 2$, this means that the

plume spreads faster than a traditional diffusion, so that the fractional diffusion equation with drift (9.21) is a model for anomalous super-diffusion. Finally, note that in the case $\alpha = 2$, the fractional diffusion model reduces to traditional diffusion.

Step 3 is to formulate the model. We will model the tritium plume using the fractional diffusion equation (9.21). We will solve this equation by the method of particle tracking introduced in Section 9.4, using the extended central limit theorem (9.23). The random variables $X_1, \ldots, X_n$ in (9.23) are identically distributed with a random variable X with cumulative distribution function

$$F(x) = \begin{cases} 0 & \text{for } x < A^{1/\alpha} \\ 1 - Ax^{-\alpha} & \text{for } x \geq A^{1/\alpha}. \end{cases} \tag{9.26}$$

Its probability density function is

$$f(x) = \begin{cases} 0 & \text{for } x < A^{1/\alpha} \\ A\alpha x^{-\alpha-1} & \text{for } x \geq A^{1/\alpha}, \end{cases}$$

and its mean is

$$\mu = \int_{A^{1/\alpha}}^{\infty} x\, A\alpha x^{-\alpha-1} dx = A^{1/\alpha} \frac{\alpha}{\alpha - 1}.$$

To simulate the random variable X, we will use the inverse cumulative distribution method developed in Example 9.3. Set $y = F(x) = 1 - Ax^{-\alpha}$ and invert to get

$$x = F^{-1}(y) = \left(\frac{A}{1 - y} \right)^{1/\alpha}$$

and then take $X = F^{-1}(U)$, where U is uniform on $(0, 1)$. To confirm that X has the desired distribution, write

$$\Pr\{X \leq x\} = \Pr\left\{ \left(\frac{A}{1 - U} \right)^{1/\alpha} \leq x \right\}$$
$$= \Pr\left\{ U \leq 1 - Ax^{-\alpha} \right\} = 1 - Ax^{-\alpha}$$

when $x \geq A^{1/\alpha}$, so that $0 < 1 - Ax^{-\alpha} < 1$.

For the particle tracking model, assume that each particle makes a small random movement $v\Delta t + (\Delta t)^{1/\alpha} X_i$ over a small time interval $\Delta t = t/n$, where

$$X_i = \left(\frac{A}{1 - U_i} \right)^{1/\alpha} - A^{1/\alpha} \frac{\alpha}{\alpha - 1}$$

and $U_1, \ldots, U_n$ are independent uniform $(0, 1)$ random variables. Use the extended central limit theorem (9.23) with $\mu = 0$ to see that $X_1 + \cdots + X_n \approx n^{1/\alpha} Z_\alpha$ for large n. Since $t = n\Delta t$, we have

$$\left(v\Delta t + (\Delta t)^{1/\alpha} X_1 \right) + \cdots + \left(v\Delta t + (\Delta t)^{1/\alpha} X_n \right) \approx vt + (n\Delta t)^{1/\alpha} Z_\alpha = vt + t^{1/\alpha} Z_\alpha.$$

The distribution function of Z_α is

$$G(x) = \int_{-\infty}^{x} g(u)\, du$$

and so the distribution function of the limit $vt + t^{1/\alpha} Z_\alpha$ is

$$\Pr\{vt + t^{1/\alpha} Z_\alpha \leq x\} = \Pr\left\{ Z_\alpha \leq \frac{x - vt}{t^{1/\alpha}} \right\}$$

$$= \int_{-\infty}^{t^{-1/\alpha}(x-vt)} g_\alpha(u)\, du$$

$$= \int_{-\infty}^{x} g_\alpha(t^{-1/\alpha}(y - vt)) t^{-1/\alpha}\, dy$$

using the substitution $u = t^{-1/\alpha}(y - vt)$ in the last line. Then the density of $vt + t^{1/\alpha} Z_\alpha$ has Fourier transform

$$\hat{C}(k, t) = \int_{-\infty}^{\infty} e^{-ikx} g(t^{-1/\alpha}(y - vt)) t^{-1/\alpha}\, dy$$

$$= \int_{-\infty}^{\infty} e^{-ik(vt + t^{1/\alpha} u)} g_\alpha(u)\, du$$

$$= e^{-ikvt} \int_{-\infty}^{\infty} e^{-i(kt^{1/\alpha})u} g_\alpha(u)\, du$$

$$= e^{-ikvt + Dt(ik)^\alpha}$$

using the same substitution $u = t^{-1/\alpha}(y - vt)$ again, along with (9.24). Since this is identical to (9.22), it follows that $C(x, t)$ solves the fractional diffusion equation with drift (9.21). Hence a relative frequency histogram of particle locations will approximate the probability density curve $C(x, t)$ that gives the theoretical relative concentration for an infinite number of particles.

The α-stable curve in Figure 9.28 was obtained by numerically inverting the Fourier transform (9.22) for the stable density, using a fitted velocity of $v = 0.12$ m/day. The robustness of this velocity assumption will be discussed later in this section. The stable index $\alpha = 1.1$ was chosen to give a spreading rate of $t^{1/1.1} \approx t^{0.9}$. The fitted value of α was also verified by checking the rate of peak concentration decline, and the leading tail of the concentration curve. Figure 9.30 shows the same tracer data from Figure 9.28 on a log-log scale, along with the fitted Gaussian and stable models. The α-stable density $C = C(x, t)$ has the property that $C \approx t\alpha A x^{-\alpha-1}$ for x sufficiently large, so that $\log C \approx \log(t\alpha A) - (\alpha + 1)\log x$. Hence a log-log plot of the fitted stable density will appear linear for large x. (The slope of this line can be used to estimate the parameter α.) The fact that the concentration data also follows this line provides additional evidence in favor of the fractional diffusion model. Since the normal estimate of tracer concentration is too low by a factor of 10^6 at the plume leading edge, the traditional diffusion model seriously under-estimates the risk
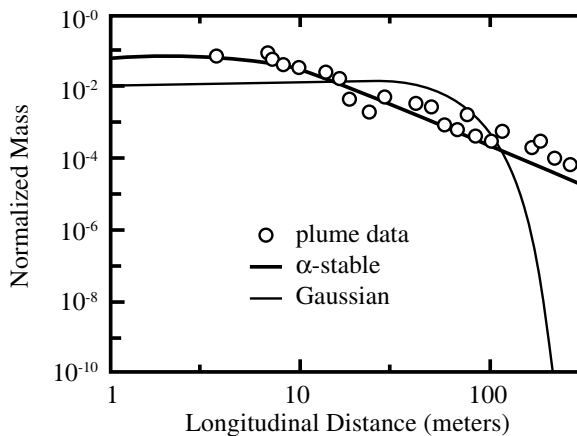
a) Snapshot 3 (day #224)

Figure 9.30: Tracer data and fitted models from Figure 9.28 on a log-log scale, to show the power law leading tail.

of downstream contamination. The fractional dispersivity $D = 0.14$ m$^\alpha$/day in (9.21) was chosen to give the best fitting stable density curve for all time. (Data was collected at day 27, 132, 224, and 328). This lead to an estimate $A = 0.0131$ for the particle jumps, using the formula $A = D(\alpha - 1)/\Gamma(2 - \alpha)$ from Theorem 3.41 in Meerschaert and Sikorskii (2012). The gamma function is discussed in Exercise 23 at the end of this chapter.

The code for our particle tracking simulation is listed in Figure 9.31. The code is similar to Figure 9.23, with a few modifications. The total mass of the particles is $P_0 = 540$ Ci, and the target concentration is 2 Ci. The tritium concentration at location $x = 20$ is estimated by counting the number of particles in the interval $15 < x \leq 25$, and then dividing by the length of this interval. The two distribution parameters are also listed. The parameter α governs the tail, and then A determines the scale, e.g., the mean value is proportional to $A^{1/\alpha}$. Since $P_0 = 540$ Ci of tritium were injected, we model the tritium concentration using $P(x,t) = P_0 C(x,t)$.

Before proceeding to Step 4, we tested the code in Figure 9.31, to see if we could reproduce the concentration curve in Figure 9.28. Figure 9.32 shows a relative frequency histogram of the locations $S(i, M)$ of particles $i = 1, \ldots, N$ at time $T = t(M) = 224$ days from a particle tracking simulation with $N = 10,000$ particles and $M = 100$ time steps, based on a computer implementation of the code in Figure 9.31. We are satisfied that the histogram of particle locations in Figure 9.32 matches the α-stable density curve in Figure 9.28 reasonably well.

Step 4 is to solve the model. Figure 9.33 summarizes the results of running the code in Figure 9.31 with $N = 10,000$ particles, with a final time of $T = 4000$ days and $M = 100$ time steps. The graph could be made smoother by increasing

Algorithm: PARTICLE TRACKING CODE (Example 9.6)

Variables: N = number of particles
T = time of final particle jump (days)
M = number of particle jumps
$t(j)$ = time of jth jump (days)
$P(j)$ = tritium concentration after j jumps (Ci)
$Pmax$ = maximum tritium concentration (Ci)
$Tmax$ = time of maximum concentration (hrs)
$Tsafe$ = time concentration falls to 2.0 Ci (hrs)
$A = 0.0131$ scale parameter in Eq. (9.26)
$\alpha = 1.1$ tail parameter in Eq. (9.26)

Input: N, T, M

Process: Begin
$\Delta t \leftarrow T/M$
$\Delta P \leftarrow 540/N$
for $j = 1$ to M do
  $t(j) \leftarrow j\Delta t$
  $P(j) \leftarrow 0$
for $i = 1$ to N do
  $S(i, 0) \leftarrow 0$
  for $j = 1$ to M do
    $U \leftarrow \text{Random}(0, 1)$
    $X \leftarrow (A/(1 - U))^{1/\alpha} - A^{1/\alpha}\alpha/(\alpha - 1)$
    $S(i, j) \leftarrow S(i, j - 1) + 0.12\Delta t + \Delta t^{1/\alpha}X$
    if $15 < S(i, j) \leq 25$ then $P(j) = P(j) + \Delta P/10$
$Tmax \leftarrow 0$
$Pmax \leftarrow 0$
$Tsafe \leftarrow 0$
for $j = 1$ to M do
  if $P(j) > Pmax$ then
    $Pmax \leftarrow P(j)$
    $Tmax \leftarrow t(j)$
  if $P(j) > 2.0$ then $Tsafe = t(j) + \Delta t$
End

Output: $Tmax, Pmax, Tsafe$

Figure 9.31: Pseudocode for particle tracking simulation of the fractional diffusion problem in Example 9.6.
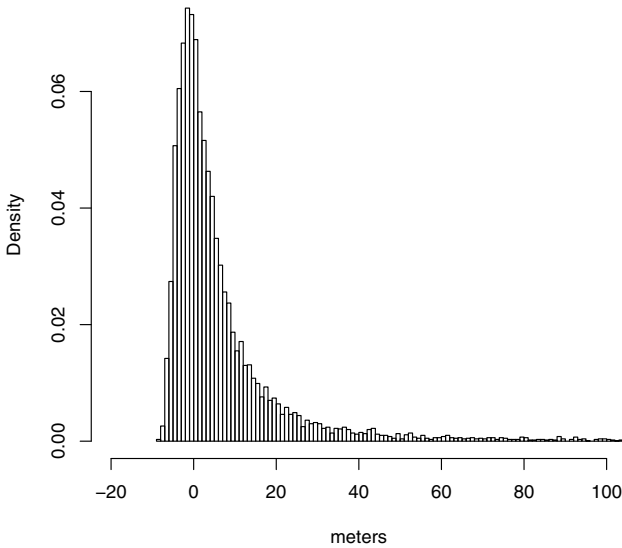
Figure 9.32: Results of particle tracking simulation for the fractional diffusion problem of Example 9.6 at time $t = 224$ days. The histogram approximates the solution $C(x, t)$ to the fractional diffusion equation with drift (9.21).

the number of particles. The maximum concentration of $Pmax = 9.14$ Ci occurred at time $Tmax = 1080$ days. It took $Tsafe = 2840$ days for the tritium level to fall back below 2 Ci.

Step 5 is to answer the question. We estimate that the maximum concentration of tritium at a point $x = 20$ meters downstream of the injection site will occur in approximately 3 years. At that time, the estimated concentration will be around 9 Curies. It will take almost 8 years for the tritium concentration at this point to fall back below 2 Curies, as the tritium plumes is carried downstream by the underground water. Figure 9.33 illustrates the predicted tritium concentration levels over time at the point $x = 20$ meters downstream, starting at the time of injection. The dotted line in that figure indicates the nominal level of 2 Curies. It is also relevant to note that the predicted tritium plume is stretched out in the direction of the ground water flow. Figure 9.32 illustrates the predicted plume shape 224 days after injection.

Since the results reported in Step 5 come from a Monte Carlo simulation, it is important to explore the sensitivity to random factors. We repeated the simulation 30 times with $N = 1,000$ particles, with a final time of $T = 4000$ days and $M = 100$ time steps. These are the same values as in Step 4, except that we decreased the number of particles per run, in order to speed up the simulation. The resulting values of the three MOP are $Pmax = 9.6 \pm 0.6$, $Tmax = 944 \pm 125$, and $Tsafe = 2830 \pm 110$, where we report the sample mean
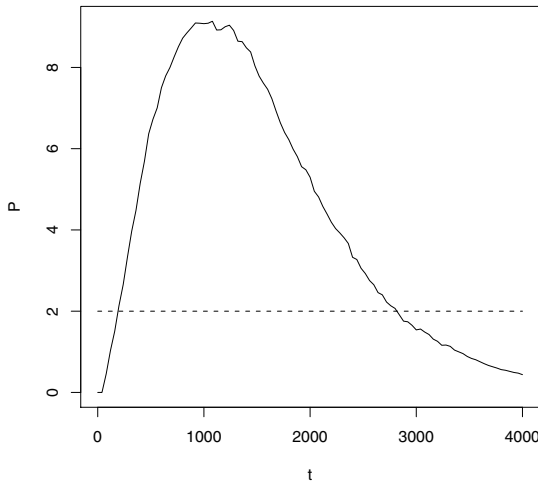
Figure 9.33: Estimated tritium levels at location $x = 20$ meters downstream of the injection point, for the water pollution problem of Example 9.6.

$\pm$ the sample standard deviation. Since the mean values from our 30 simulations are consistent with the values reported in Step 5, and the standard deviations are less than one unit (Curies or years), we conclude that the values reported in Step 5 are reliable, to the accuracy stated there.

Next we want to consider the robustness of the estimation procedure for the mean particle velocity. Recall that the velocity $v = 0.12$ m/day was estimated based on the mean (weighted average) of concentration measurements. Since the concentration curve is highly skewed, the mean (plume center of mass) is far from the mode (point of highest concentration). For example, at time $t = 224$ days, the mean is $x = vt = (0.12)(224) = 26.88$ meters. However, the peak of the α-stable density curve in Figure 9.28 lies far to the left of this point. The mean deviates significantly from the mode, because the concentration curve has a *heavy tail*. Consider the histogram in Figure 9.32. For the $N = 10,000$ particles simulated, the sample mean location was 18.5 meters, and the sample standard deviation was 355.2 meters in that Monte Carlo simulation. This absurdly large sample standard deviation is due to the fact that the theoretical standard deviation is infinite (since the second moment does not exist). Hence the sample standard deviation provides no useful information, other than to highlight the fact that there is a considerable spread in the particle location data. In fact, particle locations ranged from -8.45 to 28,352.3 meters. A very small number of particles travel a very long distance. Recall that the probability of jumping more than r meters downstream falls off proportional to $r^{-\alpha}$ in our particle jump model. Since $\alpha \approx 1$, this means that about 1 out of 10,000 particles will jump 10,000 times farther than usual. In statistics, these extreme data points

are called *outliers.*

In the presence of outliers, the mean can be an unreliable estimator of typical behavior. Consider a company in which the owner nets 1,000,000 per year, and each of 20 employees makes 50,000 per year. The average salary, including the owner, is $(2,000,000)/21 \approx 95,000$ per year, but this is hardly a good indicator of a typical salary. Likewise, the average particle location in the tritium plume is not a reliable indicator of plume center, because a small number of outliers inflate the average. The median of particle location is around 2.0 meters (50% of particles travel further downstream). The median is a more reliable measure of center than either the mean or the peak (mode), since it is not affected by outliers. However, there is no simple way to relate the median to the parameters in the model (9.21).

Note that the concentration readings in Figure 9.28 were all taken within approximately 300 meters of the injection site. This measurement procedure truncates the concentration plume. What affect does this have on the estimated velocity? Recall that the $N = 10,000$ particle locations from Figure 9.28 had a sample mean of 18.5 meters, which is not too far from the theoretical mean of 26.88 meters (considering the huge sample standard deviation). We sorted this location data, and found that 58 out of 10,000 data points exceeded 300 meters. Then we recomputed the mean particle location, omitting these values from the calculation (i.e., we added up the remaining data values, and divided by 9942). The resulting mean particle location was 8.33 meters, less than half the overall average of 18.5 meters. Although we used 99.4% of all the data values, the remaining 0.6% have a profound effect on the mean particle location. Since this calculation was based on a Monte Carlo simulation, we repeated this process several more times. The resulting mean location for all particles varied significantly. The standard deviation, and the location of the largest particle, varied widely. However, the mean location of those particles that end up less than 300 m downstream was always around 8 m. When the velocity parameter v was estimated from the concentration data in Figure 9.28 (and three other snapshots), this truncation effect was taken into account. In fact, the mean value estimated from the data in Figure 9.28 is 8.0 meters. The Gaussian curve in Figure 9.28 has a mean of approximately 8 m, since the truncation effect is negligible for this normal probability density function.

In modeling, the estimated parameter values in a given model must never be confused with the actual "truth" about these parameters. While both the traditional and fractional diffusion models contain velocity parameters, the meaning of those parameters can vary in the context of an application. For this reason, it would be a mistake to estimate the velocity parameter in the context of one model (traditional diffusion) and assume the same value is valid for a different model (fractional diffusion). The same applies to all model parameters. For example, it is not appropriate to estimate the tail parameter α in the power law model (9.26), and then assume that the same parameter value pertains to the stable model. In the finance literature, this simple truth has caused some serious confusion, see McCulloch (1997) for further details.

We conclude this example with a few comments about the connection be-

tween fractional derivatives, power law jumps, and fractals. We have already
seen that the fractional diffusion equation, with a fractional derivative of order
α, governs the sum of random power law jumps, where the number α is also
the tail index of the power law jump probability. The probability model with
power law jumps gives a concrete understanding of the fractional derivative. A
good way to think about what the fractional derivative means is to consider
the particle tracking model for a diffusion. The second derivative codes particle
jumps with mean zero and finite variance. The fractional derivative codes power
law jumps.

In Section 7.4, we derived the traditional diffusion equation (7.28) by ap-
plying the conservation of mass equation (7.26) along with Fick's Law (7.27).
Recall that Fick's Law is empirical, i.e., it is based on observations of real world
data in certain controlled experimental settings. The fractional diffusion equa-
tion can be derived in a similar manner, using a fractional Fick's law. In the
fractional Fick's Law, the particle flux

$$q = -D\frac{\partial^{\alpha-1}C}{\partial x^{\alpha-1}} \tag{9.27}$$

is based on empirical observations from chaotic dynamical systems. There the
proportion of particles that jump across j boxes of size Δx during one time step
Δt falls off like a power law. This deterministic model for fractional diffusion is
closely connected to the particle tracking model, based on random jumps with
a power law distribution. See Meerschaert and Sikorskii (2012) for more details.

In Section 6.4 we discussed the interesting subject of fractals. Particles
traces in both traditional and fractional diffusion are also random fractals. The
right panel in Figure 9.26 shows some representative particle traces, obtained
by plotting $S(n)$ versus $t(n)$ where

$$S(n) = \sum_{j=1}^{n}\left(v\Delta t + \sqrt{D\Delta t}\ Z_j\right), \tag{9.28}$$

and Z_j are independent standard normal random variables with mean zero and
variance 1. The sum $S(n)$ has a normal distribution with mean vt and variance
Dt, where $t = n\Delta t$. As $n \to \infty$, the discrete time stochastic process with value
$S(n)$ at time $t(n)$ converges to a *Brownian motion* with drift $B(t) + vt$. The
Brownian motion $B(t)$ is a Markov process whose state space is the entire real
line. Its density functions $C(x,t)$ solve the diffusion equation (7.28). Particle
traces of a Brownian motion are random fractals with dimension $d = 3/2$. If we
replace the normal random variables in (9.28) with any other random variables
with the same mean and variance, the central limit theorem still applies, and we
get the same Brownian motion process in the limit. In two or more dimensions,
the particle traces of a Brownian motion are random fractals with dimension
$d = 2$. Although the dimension is an integer, these particle trajectories are still
fractals, since their dimension is not equal to one.

Figure 9.34 shows a typical particle trace of Brownian motion. Since the
particle trace is a fractal, zooming in on one portion of the graph will reveal
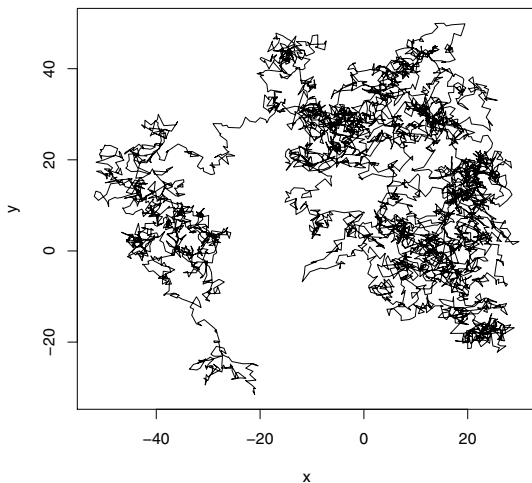
Figure 9.34: Particle tracking simulation of Brownian motion, a random fractal of dimension $d = 2$.

additional structure, similar to the larger picture. This is related to the self-similarity property of Brownian motion: The process $B(ct)$ at time scale $c > 0$ is probabilistically identical to the process $c^{1/2}B(t)$. Decreasing the time scale c decreases the spatial scale $c^{1/2}$, which is equivalent to zooming in on the graph.

If we replace the normal random variables in (9.28) with power law jumps that follow distribution (9.26) with $1 < \alpha < 2$, the extended central limit theorem shows that the limit is stable with index α. As $n \to \infty$, the discrete time stochastic process with value $S(n)$ at time $t(n)$ converges to a *stable Lévy motion with drift* $L(t) + vt$. The stable Lévy motion $L(t)$ is a close cousin of the Brownian motion. Its density functions $C(x, t)$ solve the fractional diffusion equation (9.21) with $v = 0$. In two or more dimensions, the particle traces of a α-stable Lévy motion are random fractals with dimension $d = \alpha$. Figure 9.35 shows a typical particle trace of stable Lévy motion, with $\alpha = 1.8$ in this case. The picture is similar to Brownian motion, except for occasional large jumps. The behavior is somewhat similar to the weather problem of Example 6.6. The trajectory remains localized for a period of time, and then jumps quickly to a different neighborhood. The process $L(ct)$ at time scale $c > 0$ is probabilistically identical to the process $c^{1/\alpha}L(t)$. As the fractal dimension α decreases, the graph becomes smoother. In summary, the parameter α codes the power law jumps, the order of the fractional derivative, and the fractal dimension of the particle traces.
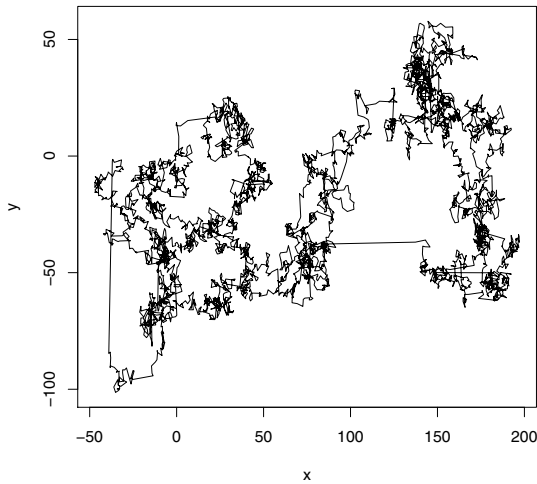
Figure 9.35: Particle tracking simulation of stable Lévy motion, a random fractal of dimension $d = \alpha = 1.8$.

9.6 Exercises

1. A simple game of chance is played by flipping a coin. The house flips a coin and the player calls it in the air. If the coin lands the way the player called it, the house pays the player \$1; otherwise, the player pays the house \$1. The player begins with \$10.

   (a) What are the odds that the player will go broke before he doubles his money? Use the five-step method, and model using a Monte Carlo simulation.

   (b) How long on average does the game described in part (a) last?

   (c) How much on average does the player have after 25 coin flips?

2. On a roll of two dice, a total of seven occurs with probability 1/6.

   (a) In 100 rolls of the dice, what is the probability that five consecutive rolls of seven will occur? Use the five-step method, and model using Monte Carlo simulation.

   (b) What is the average number of rolls until a roll of seven occurs? Use any method.

3. Reconsider the inventory problem of Example 8.1. In the text we stated that if the time between customer arrivals is exponential with mean one, the distribution of the number of arrivals in one week is Poisson with mean one.

(a) Use Monte Carlo simulation to model one week of arrivals. Assume that the time between arrivals is exponential with mean one. Simulate to determine the mean number of arrivals in one week.

(b) Modify the simulation program to keep track of the fraction of the simulated weeks in which there were 0, 1, 2, 3, or more than 3 arrivals. Compare with the Poisson probabilities given in Section 8.1.

4. (a) Repeat Exercise 3, but now assume that the time between customer arrivals is uniformly distributed between 0 and 2 weeks.

(b) Use the probabilities from your computer program output to modify the state transition probabilities for the Markov chain in Example 8.1.

(c) Solve $\pi = \pi P$ for the steady–state probabilities, assuming a uniform customer interarrival time.

(d) Determine the steady–state probability that demand exceeds supply for this modified example.

(e) Compare the results of (d) to the calculations of Section 8.1, and comment on the robustness of our original model with respect to the assumption of random arrivals.

5. Reconsider the docking problem of Example 9.3, but now assume that the time to make a control adjustment is uniformly distributed between 4 and 6 seconds.

(a) Modify the algorithm in Fig. 9.10 to reflect this change in the distribution of c_n.

(b) Implement the algorithm in part (a) on a computer.

(c) Make 20 simulation runs with $k = 0.02$, and tabulate your results. Estimate the mean time to dock.

(d) Compare the results of (c) to those obtained in Section 9.2. Would you say that the model is robust with respect to the assumption that c_n is normally distributed?

6. Reconsider the docking problem of Example 9.3.

(a) Implement the algorithm in Fig. 9.10 on a computer. Make a few model runs, and compare with the results in the text.

(b) Vary the parameter k to determine the optimal value of this control parameter. You will need to make several model runs for each value of k to determine average behavior.

(c) Explore the sensitivity of your answer in part (b) to the initial velocity of the spacecraft, which we assumed to be 50 m/sec.

(d) Explore the sensitivity of your answer in (b) to the docking threshold, which we assumed was 0.1 m/sec.

7. Reconsider the rainy day problem of Example 9.1, but now assume that if today is rainy, there is a 75% chance that tomorrow will be rainy, and likewise, if today is sunny, there is a 75% chance that tomorrow will be sunny. The day we arrived on vacation was sunny ($X_0 = 0$).

   (a) Determine the steady–state probability that any future day will be rainy. Model $\{X_t\}$ as a Markov chain.

   (b) Estimate the probability of three consecutive days of rain using a Monte Carlo simulation.

8. Reconsider the cell division problem of Exercise 11 in Chapter 8.

   (a) Use a Monte Carlo simulation to model the cell division process. How long will it take on average for one cell to grow to 100 cells?

   (b) What is the probability that the family line of one cell will die out before reaching 100 cells?

9. (Continuation of Chapter 7, Exercise 8) Simulate a random arrival process with an arrival rate of five minutes using the Monte Carlo method. Determine the average time between the last arrival prior to $t = 1$ hour and the next arrival after $t = 1$ hour.

10. (Continuation of Chapter 7, Exercise 9) Simulate the supermarket checkout stand problem. Draw a random number to represent your service time, and then see how many random numbers it takes to exceed that. Repeat the simulation a large number of times and determine the average number it takes to find a random number that exceeds yours. Justify the difference between your answer and the one obtained in part (c) of Chapter 7, Exercise 9.

11. Reconsider the inventory problem of Example 8.1, and determine the average number of lost sales per week.

   (a) Assume that on each day a customer arrives with probability 0.2. Construct a Monte Carlo simulation based on a time step of one day to simulate one week of sales activity. For a beginning inventory of 1, 2, or 3 aquariums determine the average number of lost sales by repeated simulation.

   (b) Combine the results of part (a) with the steady–state probabilities calculated in Section 8.1 to determine the overall average number of lost sales per week.

12. This exercise explains the binomial model. Suppose m independent random trials are conducted, each of which has a probability q of success. Let $X_i = 1$ if the ith trial is successful, and $X_i = 0$ otherwise. Then $X = X_1 + \cdots + X_m$ is the number of successes.

   (a) Show that $EX = mq$ and $VX = mq(1 - q)$. [Hint: First show that $EX_i = q$ and $VX_i = q(1 - q)$.]

   (b) Explain why there are $\binom{m}{i}$ possible ways for $X = i$ to occur and why each one has probability $q^i(1 - q)^{m-i}$.

   (c) Explain why Eq. (9.14) in the text represents the probability of i successes in m trials.

13. Reconsider the bombing run problem of Example 9.4. Modify the model to print out the expected number of aircraft lost during this mission. Take into account the possibility that additional planes are lost after the attack is concluded, as the bombers are leaving the target area.

   (a) How many planes on average are lost during the mission if $N = 15$ are sent?

   (b) Perform a sensitivity analysis with respect to N.

   (c) What happens if we can use advanced bombers that fly at 1,200 miles/hour and only need to loiter in the target area for 15 seconds?

   (d) Perform a sensitivity analysis on the probability q that one missile kills one plane. Consider $q = 0.4, 0.5, 0.6, 0.7,$ and 0.8. State your general conclusions. Under what circumstances would a responsible commander order his pilots to fly this mission?

14. Reconsider the bombing run problem of Example 9.4. Suppose that superior technology allows most bombers to get through the air defense undetected.

   (a) Suppose that four aircraft are detected. Let Y denote the number of these aircraft that survive eight shots by the air defense. Determine the probability distribution of Y and calculate the mean number of aircraft lost prior to completion of the attack. Use a Monte Carlo simulation based on a Markov chain model, as discussed at the end of Section 9.3.

   (b) Repeat part (a), but now use an analytic simulation.

   (c) Suppose that you were required to incorporate the possibility of multiple shots at a single aircraft into the model of Example 9.4. You have two options available to you. You may write a purely analytic simulation incorporating the results of part (b), or you may use a generalized version of the Monte Carlo simulation model of part (a) to obtain the probability distribution of Y for $d = 1, \ldots, 7$ aircraft detected, and incorporate these results into the model as data. Which option would you choose? Explain.

15. (Hard problem) Carry out the model enhancements described in problem 14(c).

16. Reconsider the bombing run problem of Example 9.4.

    (a) Use a Monte Carlo simulation to find the probability of mission success if $N = 15$ aircraft are sent.

    (b) Perform a sensitivity analysis on N. Determine the approximate probability of mission success for $N = 12, 15, 18$, and 21.

    (c) Compare the relative advantages of Monte Carlo and analytic simulations in terms of the difficulty of both model formulation and sensitivity analysis.

17. Reconsider the bombing run problem of Example 9.4. This problem shows how the binomial formula

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}$$

can be used to simplify the analytic simulation model presented in the text.

    (a) Use the binomial formula to derive the equation

$$S = 1 - (1 - p)^{N-m}(q + (1 - p)(1 - q))^m.$$

    (b) Show that the number N of planes required to ensure a success probability S is the smallest integer greater than or equal to

$$N = \log \left[\frac{(1 - S)(1 - p)^m}{(q + (1 - p)(1 - q))^m} \right] / \log(1 - p).$$

    (c) Use this formula to verify the sensitivity analysis results reported in Fig. 9.17.

18. A radio communications channel is active 20% of the time and idle 80% of the time. The average message lasts 20 seconds. A scanning sensor monitors the channel periodically in an attempt to detect the location of emitters using this channel. An analytic simulation model of scanner performance is to be constructed. It would greatly simplify the model if it were at least approximately true that the state of the channel (busy or idle) when one scan is made is independent of the state found during the previous scan. Model the channel using a two–state Markov process and use an analytic simulation to determine how long it takes until the process settles down into steady state. After this point the process essentially forgets its original state.

    (a) Determine the steady–state distribution for this Markov process.

    (b) Derive the set of differential equations satisfied by the state probabilities $P_t(i) = \Pr\{X_t = i\}$. See Section 8.2.

(c) If $X_t = 0$ (channel idle), how long does it take for the state probabilities to get within 5% of their steady–state values?

(d) Repeat part (c), assuming that $X_t = 1$ (channel busy).

(e) How far apart should successive scans be in order for the Markov property to apply (at least approximately) for this model?

19. This problem suggests a way to solve the rainy day problem of Example 9.1 using an analytic model.

   (a) Let C_t denote the number of consecutive rainy days by day t. Show that $\{C_t\}$ is a Markov chain. Write down the transition diagram and the transition matrix for this Markov chain.

   (b) We are interested in the probability that $\max\{C_1, \ldots, C_7\} \geq 3$. Alter the Markov chain model from part (a) by restricting the state space to just $\{0, 1, 2, 3\}$. Change the state transition probabilities so that 3 is an absorbing state; i.e., set

   $$\Pr\{C_{t+1} = 3 | C_t = 3\} = 1.$$

   Explain why the probability of at least three consecutive rainy days this week is the same as $\Pr\{C_7 = 3 | C_0 = 0\}$.

   (c) Use the methods of Chapter 8 to calculate the probability of at least three consecutive rainy days in one week, assuming a 50% chance of rain each day.

   (d) Perform a sensitivity analysis on the 50% assumption. Compare with the results shown in Fig. 9.5.

   (e) Compare this analytic model to the Monte Carlo model used in Section 9.1. Which do you prefer, and why? If you had just now come across this problem, which modeling approach would you have selected?

20. Use the method of particle tracking to solve the pollution problem of Example 7.5.

   (a) Implement the particle tracking code in Figure 9.23. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.5 of the text.

   (b) Modify the code so that the velocity $v = 3.0$ km/hr is a constant. Repeat part (a), and compare to the results of Example 7.5. Are the results consistent with those reported in the text?

   (c) Perform a sensitivity analysis on the Monte Carlo simulation results in part (b), to determine how the three measures of performance $Tmax$, $Pmax$ and $Tsafe$ depend on random factors. How confident are you of the results reported in part (b)?

(d) Perform a sensitivity analysis on the plume velocity. Repeat the simulation of part (b) several times for each value of v in Table 7.1, and use the average to reproduce the results in Table 7.1. How confident are you of the tabulated results?

(e) Compare this Monte Carlo simulation model to the analytical model used in Section 7.4. Which do you prefer, and why? If you had just now come across this problem, which modeling approach would you have selected?

21. Reconsider the pollution problem of Example 7.5, but now assume that the wind speed v (km/hr) is given by the formula

$$v = 3 + \frac{M - 3}{1 + 0.1d^2} \tag{9.29}$$

where d (km) is the distance from the center of town, and $M = 8$ km/hr is the wind speed in the center of town.

(a) Plot the new wind speed function together with the wind speed function from Example 7.5. Do they seem comparable?

(b) Implement the particle tracking code in Figure 9.23. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.5 of the text.

(c) Modify the code in part (b) using the new wind speed. Repeat part (b) and compare the results. Is the model of Example 7.5 robust with respect to the assumed wind speed function?

(d) Perform a sensitivity analysis on the maximum wind speed in town. Repeat part (c) for $M = 4, 6, 8, 10, 12$ km/hr. How sensitive are the three measures of performance, $Tmax$, $Pmax$ and $Tsafe$, to the maximum wind speed?

22. This problem investigates a power law model for dispersivity in the pollution problem of Example 7.5.

(a) Implement the particle tracking code in Figure 9.23. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.5 of the text.

(b) Modify the code in part (a) to compute the sample variance $s^2(j)$ of the particle location $\{S(i, j) : 1 \le i \le N\}$ at each time $t(j)$, for $j = 1, 2, \ldots, M$.

(c) Plot the variance $s^2(j)$ versus time $t(j)$ from part (b), similar to Figure 9.27. Comment on main features of the graph.

(d) Make a log-log plot of the variance $s^2(j)$ versus time $t(j)$ from part (b). That is, plot $\log s^2(j)$ versus $\log t(j)$. Do the points on this log-log plot seem to follow a straight line?

(e) Fit a power law model $\sigma = Ct^p$ to the results of part (b). One way to do this is to apply linear regression to the log-transformed data, compare Exercise 18 in Chapter 8. Plot the power law model from part (e) together with the variance data from part (c). Is it reasonable to adopt a power law model for the variance of the particle plume? What is the practical utility of this model?

23. This problem introduces the gamma function, and the Laplace transform, in the context of fractional derivatives. The gamma function is defined by

$$\Gamma(b) = \int_0^\infty x^{b-1} e^{-x} dx$$

for $b > 0$. The Laplace transform of a function $f(x)$ is defined by

$$F(s) = \int_0^\infty e^{-sx} f(x)\, dx.$$

(a) Use integration by parts to show that $\Gamma(b + 1) = b\Gamma(b)$. Conclude that $\Gamma(n + 1) = n!$

(b) Use a substitution $y = sx$ to show that the function $f(x) = x^p$ has Laplace transform $s^{-p-1}\Gamma(p+1)$.

(c) Use integration by parts to show that $sF(s) - f(0)$ is the Laplace transform of the first derivative $f'(x)$.

(d) The Caputo fractional derivative of order $0 < \alpha < 1$ has Laplace transform $s^\alpha F(s) - s^{\alpha-1} f(0)$. Use this formula to show that $f(x) = x^p$ has fractional derivative

$$\frac{\Gamma(p + 1)}{\Gamma(p + 1 - \alpha)} x^{p-\alpha}.$$

(e) Explain why the same formula also holds for positive integers α.

24. This problem introduces two integral formulas for fractional derivatives. The Riemann-Liouville fractional derivative of order $n - 1 < \alpha < n$ is defined by

$$\frac{1}{\Gamma(n - \alpha)} \frac{d^n}{dx^n} \int_0^\infty f(x - y) y^{n-\alpha-1} dy, \qquad (9.30)$$

using the gamma function introduced in Exercise 23. The Caputo fractional derivative of order $n - 1 < \alpha < n$ is defined by

$$\frac{1}{\Gamma(n - \alpha)} \int_0^\infty \frac{d^n}{dx^n} f(x - y) y^{n-\alpha-1} dy, \qquad (9.31)$$

by moving the derivative inside the integral.

(a) Apply the formula (9.30) to compute the Riemann-Liouville fractional derivative of the function $f(x) = e^{ax}$ for $a > 0$.

(b) Explain why the resulting formula is also true for positive integers α.

(c) Apply the formula (9.31) to compute the Caputo fractional derivative of the function $f(x) = e^{ax}$ for $a > 0$. Compare with the answer in part (a).

(d) Apply the formula (9.30) to compute the Riemann-Liouville fractional derivative of the constant function $f(x) = 1$ for all $x \geq 0$.

(e) Apply the formula (9.31) to compute the Caputo fractional derivative of the constant function $f(x) = 1$ for all $x \geq 0$. Compare with the answer in part (d). Which formula agrees with the case where α is a positive integer?

25. Reconsider the water pollution problem of Example 9.6, and now consider a simple analytical model for downstream contamination at early time.

(a) Implement the particle tracking code in Figure 9.31. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.6 of the text.

(b) Modify the code to estimate the *earliest* time $T\,risk$ that downstream concentration at the point $x = 20$ m downstream exceeds 2 Ci.

(c) Repeat part (b) several times to get an average value for $T\,risk$. How accurate is this value?

(d) Repeat part (c) for $x = 25, 30, 35, 40$ m downstream. Plot $T\,risk$ versus x.

(e) It was stated in the text that the α-stable density $C(x, t) \approx t\alpha A x^{-\alpha-1}$ for x sufficiently large. Use this asymptotic approximation to estimate $T\,risk$. Compare to the results of the Monte Carlo simulation. Does the analytical formula provide a reasonable estimate?

26. Reconsider the water pollution problem of Example 9.6, and perform a sensitivity analysis on the tail parameter α.

(a) Implement the particle tracking code in Figure 9.31. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.6 of the text.

(b) Repeat part (a) with $\alpha = 1.2, 1.3, 1.5, 1.8$. How sensitive are the results to the tail parameter α?

(c) Repeat part (b) several times for each value of $\alpha = 1.1, 1.2, 1.3, 1.5, 1.8$. Tabulate the average value of each measure of performance, $T\,max$, $P\,max$ and $Tsafe$.

(d) Give an accuracy estimate for the numbers you tabulated in part (c).

(e) Estimate the sensitivities $S(T\,max, \alpha)$, $S(P\,max, \alpha)$ and $S(Tsafe, \alpha)$, and interpret in the context of this problem.

27. Reconsider the water pollution problem of Example 9.6. Some scientists believe that there is a natural upper bound to the magnitude of particle jumps.

    (a) Implement the particle tracking code in Figure 9.31. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.6 of the text.

    (b) Modify the code to enforce a maximum particle jump which is $J = 100$ times the average jump size. Repeat part (a) and compare the results.

    (c) Repeat part (b) several times, and tabulate the average value of each measure of performance, $Tmax$, $Pmax$ and $Tsafe$.

    (d) Repeat part (c) with $J = 25$. How does the maximum jump size affect the results?

28. Reconsider the water pollution problem of Example 9.6, and consider how the tail parameter α affects the distribution of particle location.

    (a) Implement the particle tracking code in Figure 9.31. Run the code, and verify that the outputs are reasonably consistent with those reported in Example 9.6 of the text.

    (b) Plot a relative frequency histogram of particle locations at time $t = 224$ days, similar to Figure 9.32. Repeat several times, and comment on how the histogram shape varies depending on random factors.

    (c) Repeat part (b) for $\alpha = 1.2, 1.3, 1.5, 1.8$. How does the distribution of particle locations vary with the tail parameter α?

    (d) Repeat part (b) with $\alpha = 4$. How does the resulting histogram shape compare?

    (e) What does the central limit theorem of Section 7.3 imply about the distribution of particle location in part (d)?

29. This problem explores the fractal particle traces discussed in Section 9.5.

    (a) Implement the particle tracking code in Figure 9.31 with $N = 20$ to obtain simulated particle locations $S(i, j)$ at times $t(j)$ for particles $i = 1, 2, \ldots, N$.

    (b) Use the results of part (a) to plot the particle traces for $N = 20$ particles, similar to Figure 9.26. Comment on the features of these plots. How do they compare to the particle traces in Figure 9.26?

    (c) Repeat part (b) for $\alpha = 1.2, 1.3, 1.5, 1.8$, and recall from Section 9.5 that each graph is a fractal of dimension $2 - 1/\alpha$. How does the graph vary with the parameter α?

    (d) Modify the code to produce two sets of traces for each particle, representing the x and y coordinates of particle location. Plot the y location versus the x location, to get a graph similar to Figure 9.35.

(e) Repeat part (d) for $\alpha = 1.2, 1.3, 1.5, 1.8$, and recall from Section 9.5 that each graph is a fractal of dimension α. How does the appearance of the graph change as α increases?

Further Reading

1. Benson, D. A., Schumer, R., Meerschaert, M. M. and Wheatcraft, S. W. (2001) Fractional dispersion, Lévy motion, and the MADE tracer tests, *Transport in Porous Media* Vol. 42, 211–240.

2. Boggs, J. M., Beard, L. M., Long, S. E. and McGee, M. P. (1993) Database for the second macrodispersion experiment (MADE-2), EPRI report TR-102072, Electric Power Res. Inst., Palo Alto, CA.

3. Bratley, P. et al. (1983) *A Guide to Simulation.* Springer-Verlag, New York.

4. Friedman, A. (1975) *Stochastic differential equations and applications. Vol. 1*, Academic Press, New York.

5. Hoffman, D. *Monte Carlo: The Use of Random Digits to Simulate Experiments.* UMAP module 269.

6. McCulloch, J. H. (1997) Measuring tail thickness to estimate the stable index α: a critique. *J.* Business Econom. Statist. Vol. 15, 74-81.

7. Meerschaert, M. and Cherry, W. P. (1988) Modeling the behavior of a scanning radio communications sensor. *Naval Research Logistics Quarterly*, Vol. 35, 307–315.

8. Meerschaert, M. M. and Sikorskii, A. (2012) *Stochastic Models for Fractional Calculus.* De Gruyter Studies in Mathematics **43**, De Gruyter, Berlin.

9. Molloy, M. (1989) *Fundamentals of Performance Modeling.* Macmillan, New York.

10. Press, W. et al. (1987) *Numerical Recipies.* Cambridge University Press, New York.

11. Ross, S. (1985) *Introduction to Probability Models.* 3rd ed., Academic Press, New York.

12. Rubenstein, R. (1981) *Simulation and the Monte Carlo Method.* Wiley, New York.

13. Shephard, R., Hartley, D., Haysman, P., Thorpe, L. and M. Bathe. (1988) *Applied Operations research: Examples from Defense Assessment.* Plenum Press, London.

14. Wheatcraft, S. W. and Tyler, S. (1988) An explanation of scale-dependent dispersivity in heterogeneous aquifers using concepts of fractal geometry, *Water Resources Research*, Vol. 24, 566–578.

Afterword

Mathematics is the language of problem solving. It is at the heart of all science and technology. The beauty of an education in mathematics is that it gives you the freedom to pursue just about any technical career you can imagine. In the next few pages we will provide a brief discussion of some of the more common career opportunities for students with advanced mathematical training. This advice applies to, but is not limited to, those students who earn a degree in mathematics. It is also relevant to students who choose to gain advanced training in mathematics while they earn degrees in other areas. We will also provide some suggestions about how to achieve success in a job that uses mathematics to solve real–world problems. The number one question in the mind of most students is whether to go directly to work, or whether to pursue another, more advanced degree. We will begin by describing the most plentiful job opportunities for bachelor's and master's degree mathematics graduates.

Computer jobs are the number–one job opportunity for students with advanced mathematical training. Students who combine advanced mathematics with courses in computing, including advanced programming, operating systems, and data structures, will find a variety of opportunities for jobs in the industry. In fact, as the job market in computing becomes more competitive, the savvy computer student will seek out complementary areas of expertise to enhance their resume, and mathematics is one of the best ways to do this. It is also important to make sure that your computer studies include commonly useful programming languages like Java and C. Computing is a marketable skill that can open many doors for you. After you get a job and prove yourself, you will find that many other opportunities present themselves.

Another good job opportunity for mathematics students is actuarial work. Actuarial firms will often hire a good mathematics major just on the basis of his or her academic record, but if you are really interested, it is a good idea to take the first actuarial exam (covering probability) before you graduate. To become a full–fledged actuary requires passing a series of exams. There are graduate courses in actuarial science that will help you to prepare for the exams, or you can study on your own. If you have the ability and self-discipline, you can rise rapidly to the top of a very interesting and lucrative profession. Most Fortune 500 companies have at least one vice president who is an actuary, not to mention the opportunities at insurance companies and independent actuarial firms. Actuaries do the mathematical modeling for these firms. This is an

especially good track for people with a strong interest in business, but you do not need to have any course work in business, economics, or accounting to qualify.

Many students choose to combine their degree in mathematics with a second degree in engineering, computer science, or accounting. Of course, a good student in one of these areas will have no problem landing a respectable job. One question he or she may have is to what extent that job will use the full measure of his or her knowledge in mathematics. To be honest, there is good news and bad news here. The bad news is that during the first year or two of employment, there may be little opportunity to use much college–level mathematics. Everybody has to start somewhere, and most people have to start at the bottom. The important thing to remember is to perform with energy, enthusiasm, and accuracy. Look at this part of your working career as a test to see whether you are ready for a more sophisticated job. The good news is that once you get past the opening stanza, there will be a wealth of genuine opportunities to put your math to work on significant and interesting problems. Of course, you will have to prove that you can handle the challenge.

Next, we will discuss some of the opportunities for mathematics students who intend to pursue an advanced degree. Students with advanced mathematical training are welcome in virtually every graduate program, especially in science and engineering. When you look at the kind of work being done at the advanced levels in most of these areas, you will find that it involves a lot of mathematics. There is not enough room here to describe the incredible variety of opportunities available. We will concentrate on those fields primarily concerned with the mathematical aspects of solving real–world problems.

Advanced graduate education in mathematics is an obvious choice for someone with a degree in mathematics. If you are interested in solving real–world problems, you should look for a graduate school that offers a program in applied mathematics, statistics, or operations research. Computer science is another attractive choice, for reasons stated earlier. Bright mathematics majors with a good background in computing will find ample opportunity to use their mathematical skills during computer science graduate school. After graduation, these folks will possess a powerful combination of mathematical and computer skills that can be brought to bear on a wide variety of fascinating real–world problems. Statistics is another good option. Most statisticians started out by earning a degree in mathematics. There are plenty of job openings, and the work is much more varied than most people believe. Some of the most interesting work in mathematical modeling is done by statisticians. In fact, the author of this textbook is now employed in a statistics department!

Operations research is a very broad field of study that encompasses much of what we usually think of as mathematical modeling. It includes the study of problems in optimization, queuing theory, and inventory theory. It is possible to enter the field with a degree in mathematics, but it is better to start with an advanced degree in the field of operations research. The biggest problem here is to find the right program. Mathematics, statistics, computer science, engineering, and even MBA programs often offer a major or a concentration in this

area. Choose any one you like; it does not make too much difference which department grants the degree. Different schools have different philosophies about where such a program belongs. Another confusing problem is the name of the program. Operations research, operations management, management science, and systems science are all different names for essentially the same thing. Once again, it does not really matter which one of these appears on your degree.

If you are interested in research and teaching, consider a doctoral program. A doctorate in mathematics is one of the more marketable degrees in academics, and a doctorate in some branch of applied mathematics (e.g., numerical analysis or partial differential equations) can lead to a very good job in an industry research laboratory. These jobs are also very marketable in academics, since applied mathematicians are valuable both to teach applied courses and to participate in interdisciplinary research projects. If you are interested in an academic job, you should also consider earning your doctorate in statistics, computer science, or operations research. You will find that the work is mostly math, the job market is better, and the salaries are higher. In fact, a doctoral program in almost any field of science or engineering can provide a wealth of opportunities to use mathematics to solve real–world problems. You should not be afraid to pursue anything that captures your imagination. Finally, do not overlook the possibility of a combined doctorate. Programs in mathematical physics, mathematical biology, mathematical psychology, and mathematical economics offer unique challenges. You may also want to begin thinking about the choice between an academic job and an industry job. While academic jobs offer much in the way of lifestyle benefits, industry jobs typically pay about twice as much. You will want to be in a position to choose.

Further Reading

1. *101 Careers in Mathematics*, edited by Andrew Sterrett, Mathematical Association of America, 1529 18th Street NW, Washington DC 20036–1385, `www.maa.org/careers`

2. *Careers in Applied Mathematics*, Society for Industrial and Applied Mathematics, 3600 Market Street, 6th Floor, Philadelphia, PA 19104-2688, `www.siam.org/careers/`

3. *Careers in Operations Research*, Institute for Operations Research and the Management Sciences, 7240 Parkway Drive, Suite 300, Hanover MD 21076 USA, `www.informs.org/Build-Your-Career/INFORMS-Student-Union/`

4. *Careers in Statistics*, American Statistical Association, 732 North Washington Street, Alexandria VA 22314-1943, `www.amstat.org/careers`

5. *Occupational Outlook Handbook*, Computer and Mathematical Occupations, U.S. Bureau of Labor Statistics, Office of Occupational Statistics and Employment Projections, PSB Suite 2135, 2 Massachusetts Avenue NE, Washington DC 20212-0001, `www.bls.gov/oco/`

6. *Mathematical Sciences Career Information*, American Mathematical Society, 201 Charles Street, Providence RI 02904-2294, `www.ams.org/careers`

7. *The Actuarial Profession*, Society of Actuaries, 475 North Martingale Rd., Schaumburg IL 60173–2226, `www.soa.org/careers`

Index